Bacteriophage 2:1, 36–49; January/February/March 2012; © 2012 Landes Bioscience

The C-terminal cysteine annulus participates in auto-chaperone function for *Salmonella* phage P22 tailspike folding and assembly

Takumi Takata, Cameron Haase-Pettingell and Jonathan King*

Department of Biology; Massachusetts Institute of Technology; Cambridge, MA USA

Keywords: tailspike, cysteines, transient disulfide bond, auto-chaperone, folding intermediates

Abbreviations: SDS, sodium dodecyl sulfate; IAA, iodoacetic acid; IAM, iodoacetamide; PAGE, polyacrylamide gel electrophoresis

Elongated trimeric adhesins are a distinct class of proteins employed by phages and viruses to recognize and bind to their host cells, and by bacteria to bind to their target cells and tissues. The tailspikes of *E. coli* phage K1F and *Bacillus* phage Ø29 exhibit auto-chaperone activity in their trimeric C-terminal domains. The P22 tailspike is structurally homologous to those adhesins. Though there are no disulfide bonds or reactive cysteines in the native P22 tailspikes, a set of C-terminal cysteines are very reactive in partially folded intermediates, implying an unusual local conformation in the domain. This is likely to be involved in the auto-chaperone function. We examined the unusual reactivity of C-terminal tailspike cysteines during folding and assembly as a potential reporter of auto-chaperone function. Reaction with IAA blocked productive refolding in vitro, but not off-pathway aggregation. Two-dimensional PAGE revealed that the predominant intermediate exhibiting reactive cysteine side chains was a partially folded monomer. Treatment with reducing reagent promoted native trimer formation from these species, consistent with transient disulfide bonds in the auto-chaperone domain. Limited enzymatic digestion and mass spectrometry of folding and assembly intermediates indicated that the C-terminal domain was compact in the protrimer species. These results indicate that the C-terminal domain of the P22 tailspike folds itself and associates prior to formation of the protrimer intermediate, and not after, as previously proposed. The C-terminal cysteines and triple β-helix domains apparently provide the staging for the correct auto-chaperone domain formation, needed for alignment of P22 tailspike native trimer.

Introduction

The elongated adhesin proteins that govern the attachment of many viruses and phages to cells use their extended lateral surfaces to recognize cell surface polysaccharides and lipopolysaccharides. Examples include the Adenovirus penton fiber, T4 short tail fiber, *Bordetella pertussis* toxin, the bacteriophage associated Hyaluronate lyase (Hylp2), and a variety of phage tailspikes.[1-8] The majority of these proteins are multimeric, folding and assembling into very stable structures that can survive the diverse environments they are exposed to.[9] The structures of three bacteriophage tailspikes have been determined to high resolution (**Fig. 1**): *Salmonella* P22 tailspike, *Bacillus* Ø29 appendage (tailspike), and *E. coli* K1F tailspike.[10-13] These three proteins have major β-helix domains, with triple β-helix regions, and complex trimeric C-terminal domains. Among these three proteins, the folding, assembly and off-pathway aggregation pathways of the *Salmonella* P22 tailspike have been characterized both in vivo and in vitro. They include monomeric, dimeric and protrimer partially folded intermediates.[14-18]

Many phage structural proteins require chaperones to assist in folding and assembly. The first identified chaperone, GroEL/ES was found associated with phage λ capsid morphogenesis.[19] Other early reports of chaperone that aid folding and assembly are associated with phage T4 head assembly,[20] T5 tail assembly,[21] and P22 coat protein folding.[22-24]

The initial efforts to identify a chaperone for P22 tailspike folding and assembly were unsuccessful,[25] and the ability of the purified fully denatured chains to refold and assemble the native biologically active state in vitro argued against such a function.[16-18,26]

Recent reports reveal that K1F and Ø29 tailspikes utilize an intra-molecular chaperone domain (IMC). An emerging class of chaperones are the IMC/auto-chaperone segments of proteins that are necessary for the efficient folding and assembly of their own chains. They were initially identified as N-terminal pro-peptides or pro-sequences needed for maturation of exported proteases such as Subtilisin.[27] Chen and Inouye refer to this class as type I.[28] Another class, type II, is represented by the C- and N-terminal pro-peptides of collagen in which the pro-peptides align and

*Correspondence to: Jonathan King; Email: jaking@mit.edu
Submitted: 12/04/11; Revised: 02/13/12; Accepted: 02/20/12
http://dx.doi.org/10.4161/bact.19775

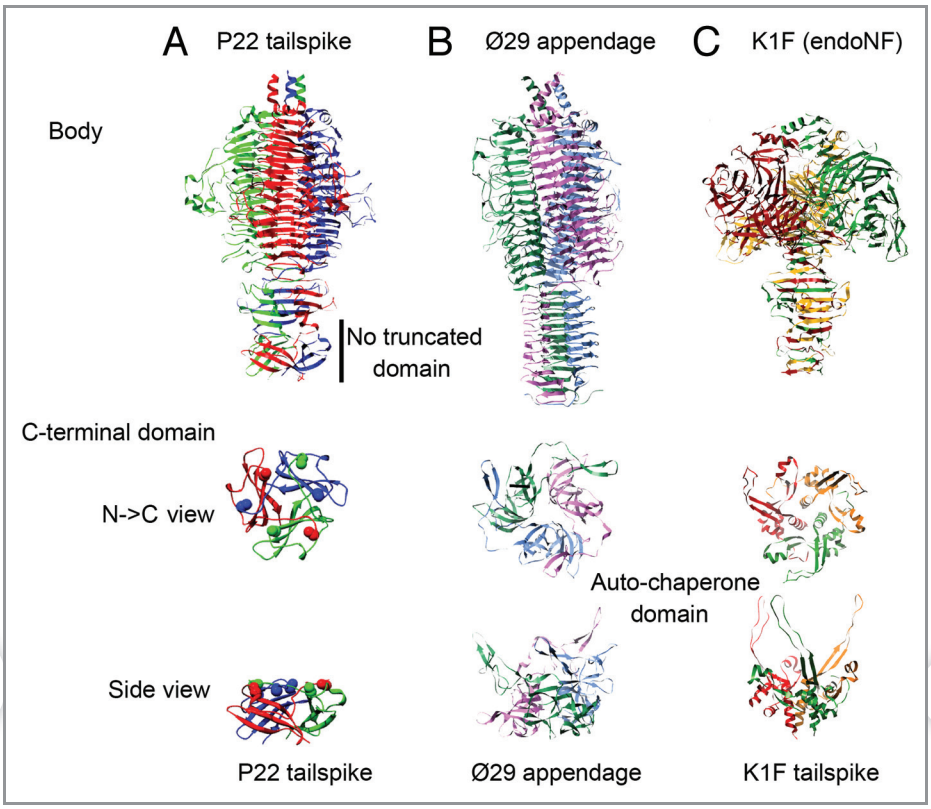

Figure 1. Structures of the C-terminal and auto-chaperone domains in *Salmonella* phage P22 tailspike and homologous proteins. (A) The structure of *Salmonela* phage P22 tailspike (PDB 1TYU). C-terminal domain remains in mature protein. (B) The structure of *Bacillus* phage Ø29 gp12 appendage (PDB 3GQ7). C-terminal domain is cleaved off after folding completion. (C) The structure of *E.coli* phage K1F tailspike (endoNF) (PDB 1V0E). C-terminal trimeric domain with long tentacles assists folding, and is cleaved off in the mature protein (3GUD, 3GW6).

tether the strands before the collagen triple helix is formed. These registration sequences are subsequently cleaved off.[29] Recently, non-cleaved type II IMC-like domains have been reported in β-helical trimeric autotransporter proteins. The autotransporter Pertactin protein of *Bordetella*, folds the C-terminus domain in the outer membrane, creating a channel for the rest of the chains to move through, and serving as the template for the folding of the N-terminal "passenger" domain.[30,31]

Recent studies of the assembly of tailspikes of the *Bacillus* subtilis phage Ø29 and *E. coli* phage K1F have shown the processes to be governed by their C-terminal domains.[13,32] The crystal structures of these adhesins are close analogs to the tailspike of the *Salmonella* bacteriophage P22 (**Fig. 1**).

The phage P22 tailspike is predominantly β-sheet (**Fig. 2**).[10,11] The N-terminal domain (residues 1–124) is comprised of anti-parallel β-sheets forming a trimeric mushroom or dome-like structure that is required for binding to the phage capsid. The major structural domain (residues 143–539) is comprised of 13 rungs of parallel β-sheets in a β-helical conformation (**Fig. 2A and C**). The length of the domain is presumably for the binding and cleavage of the lipopolysaccharide on the host cell surface.[11,33] The three β-helical domains are bound through predominantly hydrophilic interfaces.[10]

After the 13th rung the parallel β-helix motif terminates and the three chains twist around each other and intertwine to form an interdigitated triple β-helix domain (residues 540–546) (**Fig. 2C**). The chains (residues 547–612) then separate to form 3 sides of a triangular β-prism, essentially an oligomeric left handed β-helix. The subunit interfaces constitute a single buried hydrophobic core (**Fig. 2C**). The chains then twist by a loop-short α-helix-loop that put three pairs of cysteines—C613 and C635—in a distinctive even plane, forming a ring-like conformation, termed the cysteine annulus (**Fig. 2B and C**). The tailspike ends in a triple bladed motif termed the caudal fin.

The structures of the tailspikes of *Bacillus* phage Ø29, and *E. coli* phage K1F share similarities with the P22 tailspike (**Fig. 1**). The large structural domain region contains the elongated binding and catalytic domain, β-helical for P22 and Ø29 (**Fig. 1A and B**), and for K1F, three β-propellers and a β-barrel (**Fig. 1C**). Further along is the shaft region with different β-sheet structures: a triple β-helix in the P22 tailspike (**Fig. 1A**); a triple β-prism for Ø29 (**Fig. 1B**), and have similar elements to the a triangular β-prism/triple β-helix of K1F (**Fig. 1C**).[12,13]

These three tailspikes have a trimeric bladed configuration in their C-terminal domains (**Fig. 1**, bottom panels). The tailspikes of Ø29 and K1F have "tentacles" that extended up to the triple β-helix shaft. P22 tailspike lacks these tentacles.

When the Ø29 appendage and K1F tailspike are expressed without the C-termini, the resulting tailspike chains aggregated.[13,34] Ø29 and K1F mutants that retain the C-termini

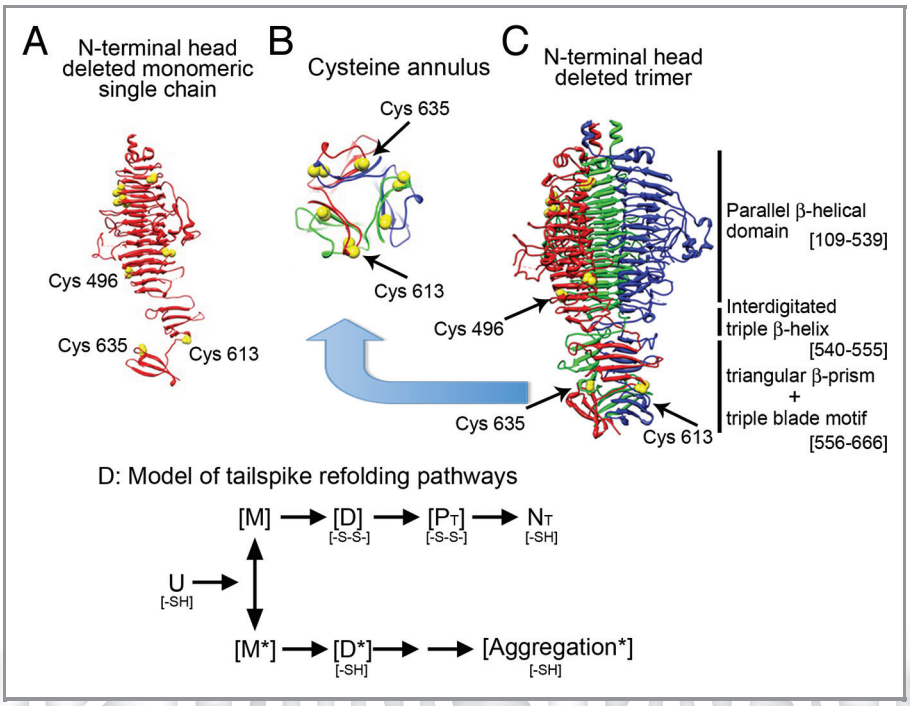

Figure 2. Structure of P22 tailspike with cysteine locations, and schematic of the folding and assembly pathway. (A) P22 tailspike single chain crystal structure with the N-terminal head domain deleted. Yellow spheres indicate cysteine residues. (B) A cross-section of the native trimer cysteine annulus bottom to top view. The three individual subunits are colored red, blue and green. Two cysteines from each subunit (C613 and C635) are located in this region. (C) The N-terminal head binding domain deleted native trimer showing the eight cysteines (PDB 1TSP). The P22 tailspike major structural domains are parallel β-helical domain (109–539), interdigitated triple β-helix domain (540–555), triangular β-prism and triple blade motif structure (556–666). All eight cysteines are reduced in the naive structure.[36] (D) Schematic diagram of intermediates in the in vitro refolding, assembly, and aggregation of tailspike chain. U: unfolded polypeptides. N_T: native trimer. The folding pathway depicts intermediates and disulfide bonds distributions: The state of the reactive cysteines in the [M] and [M*] species are described in the Results section. Dimeric intermediate: [D] and protrimer intermediate: [P_T] have been reported to contain inter-disulfide bonds: [-S-S-]. All disulfide bonds are reduced: [-SH] in native trimer. The aggregation prone intermediates are marked as [*]. No disulfide bonds have been reported in aggregation species.

domain trimerize successfully. These results show that the C-termini domains are important for folding and assembly.

The C-termini of the Ø29 and K1F chains acting as IMC/autochaperones raise the question of the role of the C-terminal trimeric domain in the folding and assembly of the P22 tailspike. Unlike Ø29 and K1F, the P22 C-terminal domain is not cleaved after native trimer folding. However, the isolated C-terminal domain (residues 537–666) of P22 tailspike can function as an independent oligomerization domain.[35] When three maltose binding protein were tethered to a trimeric P22 tailspike C-terminal domain, the chimera chains assembled into a P22 tailspike-like trimer. This suggests a type II IMC function is exhibited by the C-terminal domain of P22 tailspike.[35]

In earlier searches for cellular chaperones for P22 tailspike folding and assembly, infected cells were treated with iodoacetamide (IAM) with the intent to block ATP synthesis, in the hope of trapping chaperone complexes.[25] IAM did indeed result in blocking intracellular tailspike folding and assembly. However, it turned out to be due to direct reaction of the IAM with reactive C-terminal cysteines in tailspike folding intermediates.[25] These cysteine thiols are not reactive in the native state, and in fact are strongly hydrogen bonded.[36]

The in vivo and in vitro refolding pathway of the P22 tailspike proceeds through a number of folding intermediates (**Fig. 2D**). The polypeptide chains emerge from the ribosome or out of the chaotropic agent and proceed to form a single chain partially folded intermediate, in which the β-helical domain is somewhat structured.[17] The intermediate further folds to form a species competent for chain/chain association yielding the metastable protrimer. The protrimer then proceeds to the mature thermostable native tailspike trimer.[16-18,26,37-40]

The partially folded monomeric intermediates are very sensitive to temperature, and shift from productive folding to aggregation a few degrees above physiological temperature.[15,41,42] As a result of this thermolability of the early folding intermediate, the chain is the locus of many temperature sensitive folding (*tsf*) mutants.[43,44]

The metastable protrimer is trapped in the cold, and can be distinguished from the native state by its retarded migration in native-PAGE. The protrimer has not acquired the detergent, thermal or protease resistance, found in the native state.[37,38] The protrimer can be converted into the native state in the absence of exogenous proteins or factors.

Robinson and King reported that the protrimer intermediate formed in vitro contains transient inter-chain disulfide bonds,

which must be reduced during maturation to the native state.[45] The addition of reducing reagents to P22 tailspike folding intermediates retarded mobility of monomeric intermediate through native-PAGE, also suggesting the presence of intra-chain disulfide bonds, presumably formed from the reactive cysteines.[46]

The unusual reactivity of the C-terminal cysteine thiols may be due to interactions with side chains or backbone atoms within some distinctive local conformation, as seen in the reactive site of cysteine proteases. The thiol group of cysteine is deprotonated by nearby basic residues such as histidine before starting cleavage.[47,48] Similarly the reduction of disulfides, formed during folding or assembly, may be performed by local side chain, or backbone interaction within folding intermediates. Disulfide exchange between intra- and inter-disulfide bonds among P22 tailspike folding intermediates support this hypothesis.[49]

The presence of disulfide bonds in intermediates, yet missing from the native state has been carefully studied for bovine pancreatic trypsin inhibitor.[50-54] Similarly, the registration peptide of collagen utilizes transient intra-disulfide bonds for nucleation of C-terminal domain to stabilize the chain association stage.[55-57] The collagen zipper-like trimer folding proceeds toward the N-terminal domain with using inter-disulfide bonds formation. At the end of trimer folding, both N-terminal and C-terminal domains are processed by procollagen peptidase.

The recent identification of the auto-chaperone function in phage Ø29 and K1F adhesins suggests that the P22 tailspike cysteine annulus may be acting to organize and get in register the tailspike chains for protrimer assembly. In that case one might expect mutants of these cysteine residues to retard folding, without absolutely blocking the overall reaction. Haase-Pettingell et al. showed that the cysteine to serine mutants in vivo retarded the rate of formation of native tailspike.[58] The single serine mutations for each C-terminal cysteine decreased folding yields and kinetics, but did not completely inhibit trimer folding. However, double mutations of both cysteines for serine residues C613 and C635 sharply lowered the yield of native trimer.[46,58]

The homologous tailspike from Det7 of *Salmonella enterica* lacks cysteines in this region, Seckler and coworkers have suggested that the transient disulfide bonds are not physiological but an artifact of the relatively slow in vitro refolding process.[59] If they do not form disulfide bonds, the in vivo reactivity to IAA (iodoacetic acid) nonetheless indicates an unusual state of cysteine residues during folding and subunit assembly. This conformation dependent reactivity is likely to be a reporter or surrogate for the conformation of the C-terminal portion of the chain needed to carry out its auto-chaperone function.

As noted above, the unusual reactivity of at least two of the eight tailspike cysteines during folding, presumably represents activation by local interactions in some of the partially folded species. If the cysteine annulus motif represents an auto-chaperone function of this domain, it should form early in folding rather than later, as we previously proposed.[38,39] The reactivity of those cysteines is likely to serve as a reporter for folding of this region of the chain. Here we examine more carefully the reactivity of the cysteine residues during refolding in vitro, following conformational intermediates separated by native-PAGE, and using mass spectrometry to identify folding and assembly intermediates with reactive cysteines thiols. The results suggest that the conformation maintaining the cysteine thiols in a reactive state forms very early in folding, prior to chain association, and that our earlier model needs to be revised.

Results

Which intermediates in the folding and assembly pathway exhibit unusual cysteine reactivity? Previous work has shown that in vitro refolding of P22 tailspike proceeds through a series of folding and aggregation intermediates. Spectroscopic techniques have not resolved the monomeric and oligomeric productive intermediates or the multimeric off-pathway species. Though native-PAGE is not a traditional method for separating conformational intermediates, the 666 residues of the tailspike chain, and the complexity of the folding and aggregation pathway have made it preferable to other biophysical techniques (**Fig. 3**).[60-63] Utilizing native-PAGE, Betts and King resolved intermediates species on the productive pathway—two partially folded monomeric species M_{SLOW} and M_{FAST}, partially folded dimeric intermediates, protrimer intermediate, and of course the native trimer (**Fig. 3A**).[62] At higher temperatures, the refolding chains partition to multimeric aggregation intermediates, which have slightly different electrophoretic mobility from the related productive intermediates (**Fig. 3B**). At intermediate temperatures both productive and off-pathway intermediates are populated (**Fig. 3C**). At 0–4°C (**Fig. 3A**), two partially folded monomeric species M_{SLOW} and M_{FAST} are precursors to both productive pathway intermediates and non-productive off-pathway aggregation intermediates.[63] The complexity of the gel patterns is not an artifact of the method, but captures the population of multiple intermediates required for the chains to reach both the native trimer state, or aggregated inclusion body state.[44,62,64]

All eight cysteines are reduced and completely unreactive in the native tailspike.[36,45] The inhibition of tailspike folding by the in vivo labeling of C-terminal cysteines with radioactive iodoacetamide (IAM), as well as the characterization of intra-disulfide bond in vitro refolding intermediates, did not unambiguously identify which intermediates exhibited redox-active cysteine thiols.[25,45] To identify the partially folded in vitro intermediates with the reactive cysteines thiols, we utilized alkylation with IAA. P22 tailspike at 2.0 mg/mL was unfolded in acid urea pH 3, and refolding was initiated by dilution to 0.1 mg/mL into ice cold refolding buffer, 100 mM Tris, 2 mM EDTA, pH 7.6 for 15 min (residual concentration of urea 0.8 M).[63] Then the temperature was shifted to 24°C for further refolding and samples taken at various times. For the initial experiments, tailspike was refolded with or without IAA.

As noted above, the complexity of the band pattern represents the presence of both productive and off-pathway intermediates within the samples (**Fig. 4A**). The kinetic species resolved in this gel were monomeric intermediates, dimeric intermediates, native trimer, protrimer intermediate, as well as trimeric, and multimeric aggregates. In the left lanes at early times, the partially folded monomeric and dimeric intermediates predominate. Note that

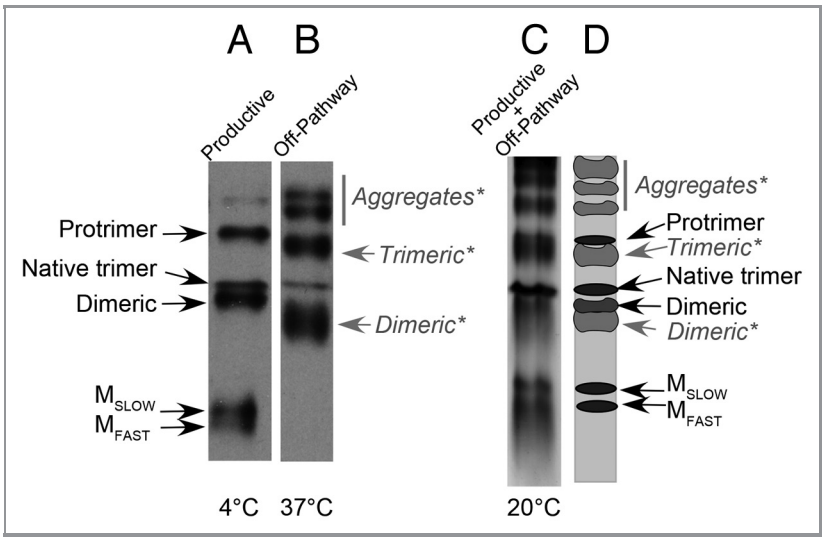

Figure 3. During the refolding of P22 tailspike, productive folding intermediates and non-productive multimeric aggregation, can be resolved with Native-PAGE. (A) At 4°C productive intermediates predominate as visualized by western blot analysis of a Native-PAGE, probed by a mixture of monoclonal antibodies recognizing non-native epitopes. (B) At 37°C non-productive folding intermediates and multimeric aggregation are detected by the same western blot analysis in (A). (C) At 20°C both productive and non-productive species are observed as seen in this silver stain of a Native-PAGE. (D) The graphic depicts the bands, the productive folding intermediates and the non-productive species from the gel in (C). Black labeled bands are productive species, gray and italic labeled bands are non-productive species. (A and B) were adapted from Betts and King 1998.[63]

due to the time taken to load the native gels, and initiate electrophoresis, the "0" time sample has incubated for additional minutes, accounting for the presence of multiple intermediate species. After 60 min of refolding at 24°C (right-hand lane), significant amounts of native trimer has accumulated. In native gels the productive protrimer forms the trailing edge of the off pathway trimeric aggregation intermediate band and is not fully resolved as seen in **Figure 3C**.

In **Figure 4B**, IAA has been added to the folding reaction at various times of incubation—0, 5, 30 and 60 min—corresponding to the samples in **Figure 4A**. These samples were further incubated with IAA as indicated in the top of **Figure 4B**

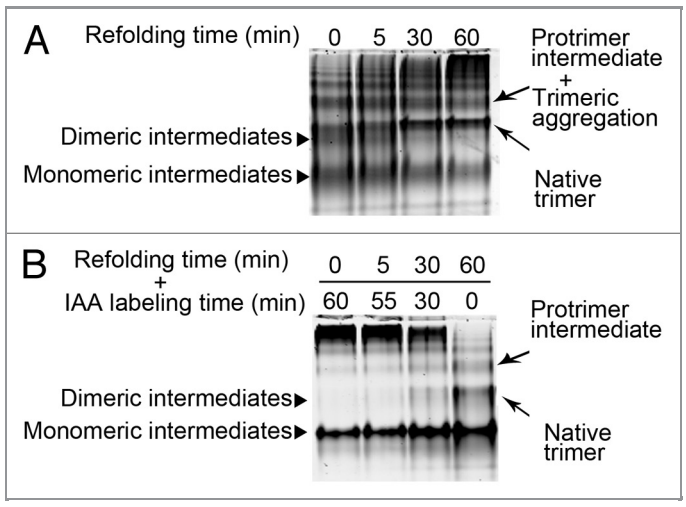

Figure 4. In vitro refolding of P22 tailspike with or without iodoacetic acid (IAA). The monomeric species were labeled and stalled by alkylation treatment. P22 tailspike was denatured and refolded as described in Materials and Methods. At indicated times, IAA was added to stall folding reaction and analyzed by non-reducing 1-D native-PAGE. (A) In vitro refolding at 24°C shows the native trimer formation as a function of the refolding time. (B) Samples were made up to 30 mM IAA at indicated times and incubated for total 60 min at 24°C. IAA reacted monomeric species that aggregated, remained at the top of gel.

[Refolding time (min) + IAA labeling time (min) to equal a total of 60 min]. Reaction of IAA with refolding tailspike and further incubation at 24°C sharply reduced the formation of native tailspike (**Fig. 4B**, 0 time). Though failing to yield native tailspikes, the samples exposed to IAA at early times accumulated higher molecular weight oligomeric aggregation species, visible at the top of the gel. In addition, the band corresponding to partially folded monomeric intermediate was not depleted upon incubation and appeared to accumulate. The simplest interpretation is (1) alkylated chains are unable to proceed through the productive pathway to the native state as found in vivo;[25] (2) polymerization to high molecular weight off-pathway aggregates is not inhibited; and may be promoted; in the presence of IAA.

P22 tailspike folding intermediates which have reacted with IAA should have an additional net charge due to the negatively charged carboxymethylated group. These modified chains can be expected to be fractionated from the unreacted chains. Two-dimensional (2-D) native-PAGE was employed in this fractionation.

During 4°C incubation, partially folded monomers, dimers, and protrimer intermediates accumulated due to suppression of aggregation at low temperature.[63] To selectively observe IAA labeled intermediates, labeling and subsequently non-reducing 2-D native-PAGE were performed under the 4°C condition (**Fig. 5**). The gel slice on the top of the gel is the lane of the first-dimensional (1-D) native-PAGE separation placed horizontally to show the species present in the refolding reaction after 3 h at 24°C in the absence of IAA. For the samples of **Figure 4B**, 30 mM IAA was added after 2 h refolding, and incubated for an additional 1 h at 4°C. Comparison of the 1-D pattern reveals that the formation of the native trimer was clearly inhibited, and monomeric species have accumulated (**Fig. 5B**, top gels). In **Figure 5** after separation through in the 1-D native-PAGE, the gel lanes were excised and put on a non-reducing 2-D native gel (in this case without further treatment), and electrophoresed in a second dimension producing the observed diagonal pattern.

In **Figure 5A**, the most intense spot in the middle of the diagonal line corresponds to native trimer (N_T). The species with

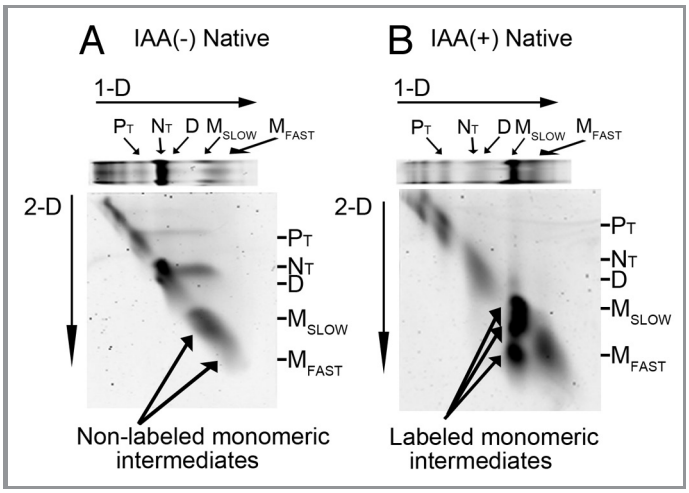

Figure 5. IAA treatment altered monomeric tailspike intermediates. (A) Non-reducing 2-D native-PAGE of refolded P22 tailspike without IAA labeling. P22 tailspike was denatured, refolded for 2 h at 24°C, and then additionally incubated for 1 h at 24°C without IAA. (B) Non-reducing 2-D native-PAGE of refolded P22 tailspike with IAA labeling. P22 tailspike was denatured, refolded for 2 h at 4°C, and then additionally incubated for 1 h at 4°C with 30 mM IAA. N_T, P_T, and D indicate native trimer, protrimer, and dimeric intermediates. M_{SLOW} and M_{FAST} refer to electrophoretically slower and faster migrating monomeric intermediates.

retarded migration in the upper left of the gel represent trimeric, and higher order aggregates. On the lower right on the diagonal line are two monomeric intermediates, slower migrating monomeric M_{SLOW} species and faster migrating monomeric M_{FAST} species.

In **Figure 5B**, with the applied sample lane on top, refolding samples treated with IAA were electrophoresed in the first dimension. The inhibition of native tailspike formation and accumulation of slower migrating monomeric M_{SLOW} intermediate was clear, in comparison with untreated control (**Fig. 5A**, top). The fractionated bands were then electrophoresed through the 2-D native-PAGE. While most bands were on the diagonal, additional spots migrated below the diagonal line indicating increased charge, or compactness (**Fig. 5B**). One of the intense spot migrated with the same mobility as faster migrating monomeric M_{FAST} species in 1-D native-PAGE. The faster migrating species has the negatively charged carboxymethylated group introduced by the IAA, which may account for the change of migration, or the reaction with the thiol may have altered the conformation and thus the migration.

Regardless of the precise origin of the altered mobility, these results confirmed that cysteine side chains with unusual chemical reactivities occur in a monomeric folding intermediate in the pathway. These results are similar to previously reported by Danek and Robinson for tailspike monomeric complex treated with oxidized glutathione.[65] Given that the cysteine reactivity must depend on some distinct chain conformation, this sensitive precursor may represent species with the C-terminal chaperone domain folded into a competent conformation.

Distinguishing monomeric precursors in the productive and aggregation pathways. Systematic study of temperature sensitive folding (*tsf*) mutants of the P22 tailspike established that the folding defect destabilizes a thermolabile partially folded intermediate leading to off-pathway aggregation.[42,43,66] Since chaperones typically function to suppress off-pathway aggregation, we wanted to examine the relationship of the two classes of partially folded monomeric intermediates to the productive and aggregation pathways.

In order to investigate productive and non-productive monomeric intermediates, non-reducing 1-D native-PAGE and non-reducing 2-D native-PAGE were employed. P22 tailspike was denatured, refolded at lower temperature without additional IAA to observe free cysteine side chain interactions, and applied to non-reducing 1-D native-PAGE. To follow the effects that temperature had on the monomeric intermediates, the 1-D native gel slice was incubated in vitro at 24°C restrictive temperature for 0 h, 1 h and 8 h. All gels were stored on wet ice to halt folding until applied to the non-reducing 2-D native-PAGE (**Fig. 6**). For clarity a graphic depiction of the 1-D gel is positioned on the 2-D gel.

Both slower migrating monomeric M_{SLOW} and faster migrating monomeric M_{FAST} intermediates were observed on a diagonal in the non-reducing 2-D native gel (**Fig. 6A**: "8 hours on ice"). The arrow points to an additional faster migrating monomeric M_{FAST} intermediate below the M_{SLOW} (**Fig. 6A**, arrow). This shifted monomeric species below the diagonal line may be produced by intra-disulfide bond formation, or folding compaction of slower migrating monomeric M_{SLOW} intermediates. After 1 h incubation at 24°C sample, the slower migrating monomeric M_{SLOW} band decreased while the intensity of faster migrating monomeric M_{FAST} increased (**Fig. 6B**, "1 hour at 24°C," arrow). After 8 h incubation at 24°C, no M_{SLOW} or M_{FAST} monomeric intermediates were visible (**Fig. 6C**, "8 hours at 24°C"). We concluded that all monomeric species entered the aggregation pathway and were unable to enter the 2-D native gel. It appears that the slower migrating M_{SLOW} intermediate is either intra-disulfide bonded, or collapsed to form faster migrating M_{FAST} species, then proceeded down the aggregation pathway. There was little evidence for thiol reactivity among the higher molecular weight multimeric intermediates, confirming previous observations.[45,60]

Which reactive cysteine-containing intermediates partition to the folding pathway? The prior experiments indicated that all partially folded monomeric intermediates upon extended incubation eventually shifted onto the aggregation pathway. However, it is evident that under the refolding conditions one of the monomeric species is the kinetic productive intermediate and the other is the off-pathway aggregation precursor. In the autochaperone model, the chain that had reached the required C-terminal conformation would be productive. Because the state of the reactive cysteines can be a report on this conformation, we would expect different reactivities between the productive and off-pathway intermediates. To this end, we employed non-reducing and reducing 2-D native-PAGE to investigate the state of the reactive thiol groups exposed in folding intermediates.

P22 tailspike was denatured, refolded, then applied to non-reducing 1-D native-PAGE. The excised lanes were incubated with or without 0.1 M DTT for 1 h at 24°C before loading onto

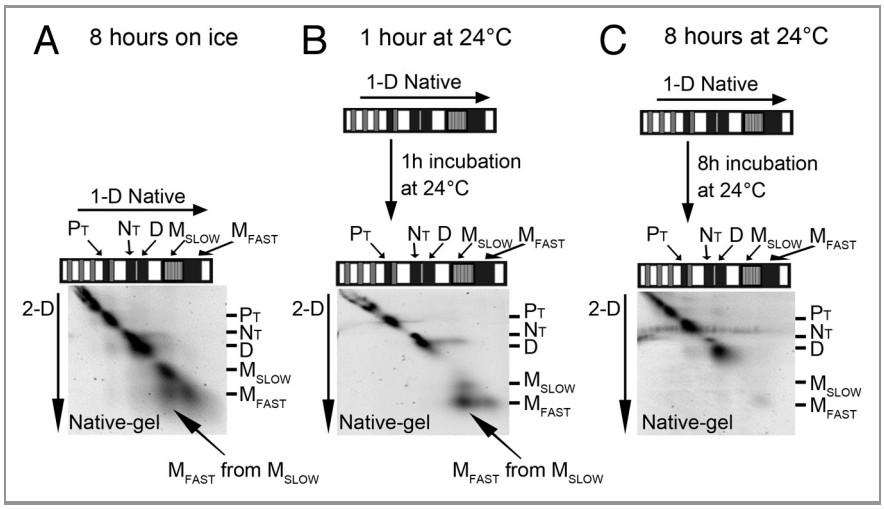

Figure 6. P22 tailspike monomeric intermediates were heat labile and aggregated in native gel. (A) Refolding tailspike was electrophoresed in the first dimension, then excised gel lane was incubated in 1 X native-PAGE running buffer on wet ice for 8 h before non-reducing 2-D native-PAGE. (B) The excised 1-D gel lane was incubated at 24°C for 1 h, then stored on wet ice until loaded on non-reducing 2-D native-PAGE. M_{SLOW} was decreased with increased M_{FAST}, indicated intra-disulfide bond formation, or folding collapse. (C) The 1-D excised gel lane was incubated at 24°C for 8 h before non-reducing 2-D native-PAGE. All monomeric species aggregated to complexes too large to enter gel.

2-D native-PAGE (**Fig. 7**). As seen in **Figure 7A**, during the 2-D native-PAGE in the absence of DTT, most of the intermediates remained on the diagonal line with the small amount of smeared M_{SLOW} unlike the M_{SLOW} shift in **Figure 6B**. This difference may be caused by the 7 h incubation of 1-D native gel on wet ice before 2-D native electrophoresis (**Fig. 6B**). Even at very low temperature in the native gel, monomeric intermediates may alter their conformations.

In the right **Figure 7B**, in which the second dimension is run on native-PAGE after DTT treatment, a number of new

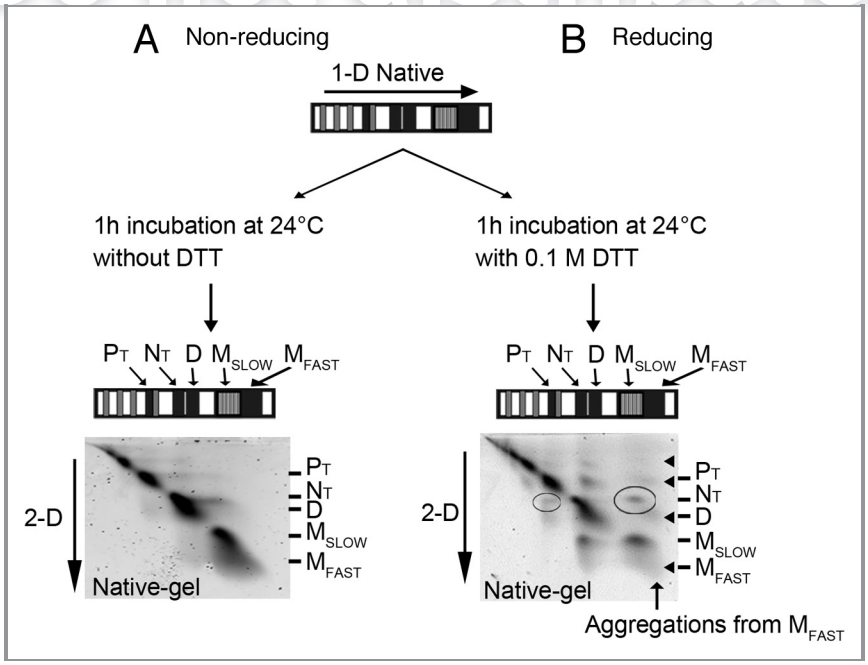

Figure 7. The effects of DTT reduction on the various tailspike folding and assembly intermediates as monitorized by 2-D native-gel electrophoresis. (A) The excised gel lane from 1-D native-PAGE was incubated in native running buffer in the absence of DTT at 24°C for 1 h before 2-D native-PAGE. The productive folding and assembly intermediates migrated on the diagonal in the second dimension. (B) The excised gel lane from 1-D native-PAGE was incubated in native running buffer in the presence of 0.1 M DTT at 24°C for 1 h before 2-D native-PAGE. Native trimer formed from slower migrating monomeric M_{SLOW} intermediates during incubation with DTT can be seen (larger circle). Reducing induced native trimer formation from protrimer P_T intermediate (smaller circle). Aggregation intermediates were formed from faster migrating monomeric M_{FAST} intermediate (triangles at the right of gel).

off-diagonal spots were observed (**Fig. 7B**). One spot appeared at the native trimer position (**Fig. 7B**, large open circle) migrating above M_{SLOW} species. This confirmed that the M_{SLOW} species was productive and proceeds to the native state in the presence of DTT. It also should be noted that the formation of a small amount of native trimer is seen below the protrimer position, as previously reported (**Fig. 7B**, small open circle).[45] Further analysis of the samples by 2-D SDS-PAGE showed that native like bands were not dissociated into monomers in the presence of SDS, confirming the existence of native tailspike trimer (data not shown). This supported the interpretation that the M_{SLOW} intermediate represents the chain conformation needed for productive folding and assembly. The conversion of M_{SLOW} to native trimer was limited, with significant levels of M_{SLOW} species still remaining after 1 h DTT treatment at 24°C (**Fig. 7B**).

The presence of DTT sharply reduced the level of the M_{FAST} species but it was not converted to native trimer. Instead, aggregated multimeric species appear above the M_{FAST} species (**Fig. 7B**, triangles on side of gel), which suggests that the M_{FAST} is a disulfide bonds containing species kinetically trapped in a non-productive conformation. Upon reduction it proceeds down the aggregation pathway.

In the position below the dimeric intermediates, M_{SLOW} monomeric species appeared, presumably due to dissociation of some dimeric intermediates rather than the productive protrimer. Since the productive and off-pathway dimeric intermediates were not resolved in the first dimension, it was difficult to interpret in the second dimension.

From these results, we concluded that tailspike folding under reducing condition is due to the presence of a partially folded monomeric intermediate with critical thiols either correctly disulfide bonded or in some critical environment in the productive conformation. This thiol environment is presumably not maintained in the species that aggregates.

When does the C-terminal domain become compact in the pathway? In the native state the cysteine annulus and triple β-helix are intimately packed and symmetrically aligned (**Fig. 2**). It seemed unlikely that the early partially folded precursors would have this packing prior to their association and assembly. If the C-terminal domain is acting as an auto-chaperone, one may expect the protrimer be arranged in a conformation closer to the native tailspike. To pursue this question we compared the resistance of the partially folded tailspike monomeric species and protrimer to trypsin digestion, using mass spectroscopy to analyze the proteolytic fragments.

Burying of cysteine residues in the folding of the C-terminal domain would be coupled with the formation of tertiary structure. To compare cysteine reactivity in monomeric and protrimer folding intermediates, a two stage thiol labeling was employed, with initial exposure to IAA followed by reduction and subsequent labeling with IAM. Free thiol groups before in-gel trypsin digestion would be carboxymethylated with initial IAA treatment. Thiol groups involved in disulfide bonds and/or are buried in tertiary structure would not react with IAA. The species would be partially denatured by the organic solvent, such as acetonitrile for dehydration of gel, these disulfides would be reduced with DTT, and the resulting thiols and buried thiol would react with IAM during a part of in-gel digestion.

The mass spectrometry of such labeled tryptic fragments identifies reactive thiols due to difference of molecular mass between carbamidomethylated and carboxymethylated cysteines in the same peptide. We will first describe the protection of the tailspike chain from tryptic digestion during folding and assembly, and then the thiol labeling results.

Trypsin fragments of the treated samples were generated and identified by mass spectrometry. The native tailspike trimer is stable to enzymatic digestion, while its folding and assembly intermediates are at least partially susceptible. The summary of the recovery of the peptides is depicted (**Fig. 8**).

As shown in **Figure 4** partially folded monomeric tailspike wild type (WT) chains accumulated after reaction with IAA. For the tryptic analysis, IAA was added at 30 min of refolding tailspike and incubated for 2 h at 4°C. After non-reducing 1-D native-PAGE, the accumulated predominant monomeric band was excised and used for in-gel trypsin digestion. 80% of total peptide of WT monomeric intermediate was recovered.

To compare with protrimer intermediate, the band corresponding to WT protrimer from the refolding reaction was excised from non-reducing 1-D native-PAGE. This gel band was incubated with IAA for 1 h to label reactive thiols in the protrimer intermediate. The gel band was applied for in-gel trypsin digestion in order to generate peptides from the protrimer intermediate. In this case only 48% of total peptide of WT tailspike protrimer intermediate was recovered.

The difference in peptide recovery between monomeric and WT protrimer intermediates is shown using the native trimer structure to depict the recovered peptides (**Fig. 8A**, "P22 tailspike monomeric intermediate vs native protrimer"). Yellow peptides were identified in both monomeric and protrimer intermediate. Red peptides were recovered only from the monomeric intermediate, and thus considered protected in the protrimer. White peptides were not identified in either species. In comparing peptide recovery from protrimer with partially folded monomeric species, the C-terminal domain of protrimer intermediate appears to be protected from trypsin digestion (**Fig. 8A**, red color region).

The protrimer in the in vitro refolding of wild type tailspike is a transient species and does not accumulate. To increase the concentration of this species, we used a mutant in the triple β-helix domain 540–548, which accumulated protrimer intermediate.[39] In this truncated tailspike variant, ΔN N547Y, the N-terminal 108 residues are deleted. The intense protrimer band of ΔN N547Y was excised after non-reducing 1-D native-PAGE, then analyzed similarly to WT protrimer.

Forty four percent of total peptide of ΔN N547Y protrimer intermediate was recovered (**Fig. 8B**, "P22 tailspike monomeric intermediate vs ΔN N547Y Protrimer"). There were some differences in peptide recovery between WT protrimer intermediate and ΔN N547Y protrimer. The N-terminal β-helix near the head binding domain was protected in ΔN N547Y protrimer (**Fig. 8B**, red color region). The protection of the C-terminus begins at the cysteine annulus in the WT protrimer, and in the ΔN N547Y extends into the triangular β-prism region.

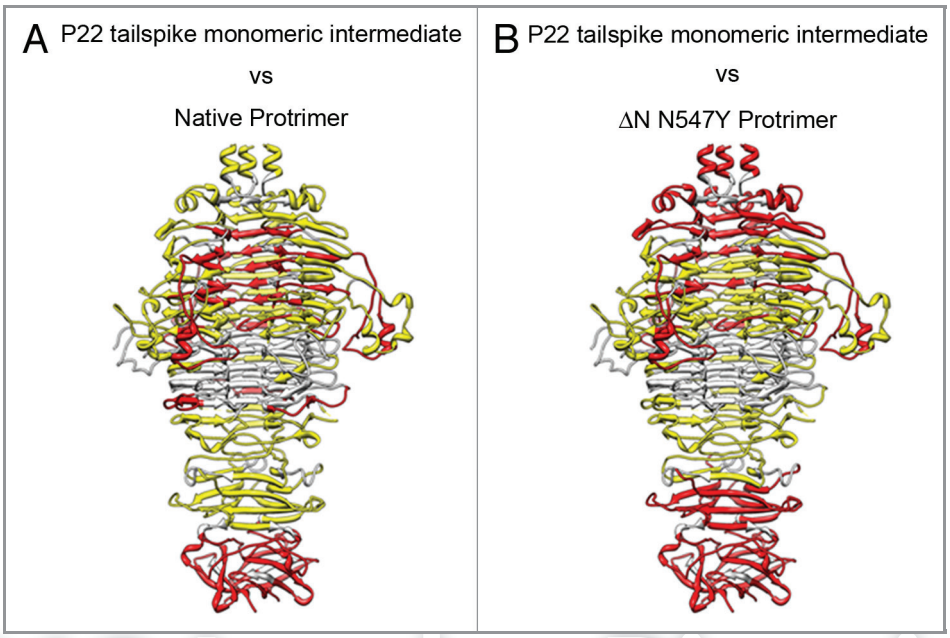

Figure 8. Comparison of tryptic peptide recovery between P22 tailspike monomeric intermediates and protrimer intermediates. Yellow region represents peptide identified in both species. Red region represents only identified in monomer species. White region peptide could not be identified in either. (A) The difference in trypsin digested peptides between WT monomeric intermediate and WT protrimer. (B) The difference in trypsin digested peptide between WT monomeric intermediate and ΔN N547Y protrimer. C-terminal domain was protected in both protrimer intermediates.

The WT protrimer intermediate and protrimer mutants may have some conformational differences, but both protrimers have stable C-terminal regions. Our interpretation is that the C-terminal domains of partially folded monomeric intermediates are relatively free and flexible and thus sensitive to the protease. Then this region begins to assemble, and collapse to a protease insensitive compact structure in the protrimer and native trimer.

Which cysteines alter their reactivity in protrimer assembly step? Kim and Robinson proposed that the disulfide exchange of transient tailspike disulfide bonds was a key step in the formation of the protrimer.[49] In the auto-chaperone model these steps would be associated with that function, similar to the disulfide exchange in procollagen folding and assembly. The earlier studies identifying transient disulfide bonds in the C-terminal cysteines were unable to resolve whether they were inter-subunit or intra-polypeptide.[46]

To identify the cysteine residues participating in putative transient disulfides, double alkylation labeled peptides from monomeric species and ΔN N547Y protrimer in prior experiment were compared (**Table 1**). From mass spectrometry of the IAA treated monomeric intermediate, three carboxymethylated cysteines C496, C613 and C635 were identified, these were in peptides in 484–497, 571–615 and 632–648 (columns 1 and 4 of **Table 1**). We used negatively charged modification to label exposed cysteines before the tryptic digestion. The negatively charged (IAA labeled) cysteine side chains may partially block trypsin access to the positively charged cleavage sites (K607 and K611). As a result, IAA labeled cysteine containing peptide were probably not completely digested into small peptides, and we could identify as a long peptide. We could not find (K_{607})VNHK$_{611}$DCR peptide containing one uncleaved site. Thus, IAA labeled cysteine inducing local structural alteration would involve both K607 and K611. The two peptides containing cysteine 613 and cysteine 635 are located in the C-terminal cysteine annulus, and peptide 484–497 containing cysteine 496 is at the end at β-helix (column 2 of **Table 1**). Sather and King confirmed reactivity of monomeric intermediate cysteines 496, 613 and 635 to IAM in vivo.[25]

Two IAM labeled cysteines C169 in 164–170 and C267 in 266–274 at β-helical domain were recovered (column 1 and 4 of **Table 1**). C169 and C267 are among the stacked side chains that appear to fold early and rapidly.[17,40] The labeling of these two cysteines by IAM indicates they were disulfide bonded in the monomeric intermediate, or were already buried and unreactive in early folding intermediates.

Two predicted peptides, 286–291 from β-helix domain containing two cysteines C287 and C290, and 447–483 from near the end of N-terminal β-helix domain containing C458 were not detected (column 4 of **Table 1**). Subsequently, endoproteinase Asp-N or Glu-C was used for the same experiment to generate appropriate length cysteine containing peptides. However, neither peptide was recovered (data not shown). Since the β-helix forms rapidly even in the partially folded monomeric intermediate,[17] the cleavage sites flanking these peptides may be inaccessible to proteolysis.

These results indicated that three cysteines in partially folded monomeric intermediates; C496 at the end of β-helix region, C613 and C635 in the C-terminal domain are free and reactive, while other cysteines are not reactive.

Table 1. Modification of cysteines in WT monomeric intermediate and ΔN N547Y protrimer mutants measured by double alkylation with mass spectrometry

Cysteine containing peptide	Structure	Tryptic peptide	Modification in WT monomeric	Modification in ΔN N547Y protrimer
164–170	β-helix	VLTIE**C$_{169}$**K	IAM	X
266–274	β-helix	E**C$_{267}$**IGVEVHR	IAM	IAM
286–291	β-helix	G**C$_{287}$**HF**C$_{290}$**K	X	X
447–483	β-helix	GMYVSNITVED**C$_{458}$**AGSGAYLLTHESVFTNIAIIDTNTK	X	X
484–497	β-helix	DFQANQIYISGA**C$_{496}$**R	IAA	IAA+IAM
571–615	C-terminal annulus	TLDSGALYSHINGGAGSGSAYTQLTAISGSTPDAVSLKVNHKD**C$_{613}$**R	IAA[†]	X
632–648	C-terminal annulus	DSS**C$_{635}$**FLPYWENNSTSLK	IAA	X

[†]The modification of C613 was identified only in monomeric intermediates 571–615 peptide containing uncleaved sites.

We proceeded to assess the state of cysteines in the ΔN N547Y protrimer mutant. For this species only two of seven cysteine containing peptides were recovered, one containing cysteines C267 and C496, (column 5 of **Table 1**). The three cysteines (C169, C613 and C635) containing peptides recovered from monomeric intermediate were not recovered. This presumably reflects the protective folded and partial assembled state of the chains in the protrimer state. The peptide containing C496, resulted in a mixture with some chains labeled by IAA, and some by IAM. The C496 thiol side chains labeled by IAA were presumably exposed in initial conformation, while the C496 thiols labeled by IAM were partially buried in the protrimer intermediate. This suggests that the lower region of the β-helical domain in the protrimer was not locked in the native like conformation, but maybe more flexible. As noted above, the peptides containing cysteines C613 and C635 were not recovered, reflecting the compact conformation in both the WT and ΔN N547Y (**Fig. 8**, and column 5 of **Table 1**). This result suggested that the sequences surrounding these two cysteines were buried or otherwise protected within the protrimer intermediate.

The reactive C-terminal annulus thiols in the monomeric intermediate, C613 and C635 were labeled with IAA, but were protected in the protrimer. These results indicated the C-terminal domain of tailspike polypeptide chain undergoes significant conformational change prior to the assembly into the protrimer, consistent with this region having an auto-chaperone function as established for Ø29 appendage and K1F tailspike.

Discussion

The C-terminal region of tailspike of P22 is a very complex topological fold, with its interdigitated regions, triple β-helix, cysteine annulus, and triangular bladed motif (**Figs. 1 and 2**). The P22 tailspike in vivo folding and assembly pathway has a $t_{1/2}$ of 5 min at 30°C, indicating complex kinetic processes.[37] Inspection of the C-terminal triple β-helix, triangular β-prism and cysteine annulus regions suggests that formation of these intimately intertwined domains may pose kinetic and/or thermodynamic barriers. Isolated β-helical (residue 109–544) domains do not form stable trimers but are in reversible trimer-monomer equilibrium.[18] It seems likely that this wrapped feature of the C-terminal of the tailspike is responsible for the more than 40°C increase in thermostability in the transition from the protrimer to the native trimer.[14,38]

The discovery that related adhesins have auto-chaperone domains (**Fig. 1**) provides a unifying model; these regions fold early and chaperone the chain assembly process to produce the intertwined domain with correct chain register. The presence of unusually reactive cysteine residues and perhaps transient disulfide bonds in the C-terminus presumably reflects some distinct local conformation that reflects the auto-chaperone function required for its in-register assembly and chain interdigitation.

P22 tailspike C-terminal domain as auto-chaperone in protrimer folding intermediate. The native tailspike trimer is completely resistant to tyrpsin digestion. However, the C-terminal region of the partially folded monomeric intermediate was susceptible to trypsin digestion. In contrast the same regions were protected in the two protrimer intermediates tested. This suggests that the protrimer intermediates have trimer-like folded C-terminal domains, while the N-terminal domains are not fully assembled. This is consistent with an auto-chaperone function within the C-terminal sequences directing the chains into the protrimer configuration.

In this auto-chaperone model the C-terminal interdigitated region forms early in protrimer formation. Our previous description of this region serving as a molecular clamp, implied C-terminal association as a late step in protrimer formation.[39] From results reported here we conclude trimer assembly is initiated from the C-terminal domains, directing trimer assembly similar to the role to the C-terminal registration peptides in procollagen triple helix precursor folding (**Fig. 9**).[56,57]

Diversity of auto-chaperone functions. The tri-bladed motif at C-terminal domain of P22 tailspike is similar in conformation to the other tailspike superfamily C-terminal auto-chaperone domains, though without sequence homology.[13,32] The P22 auto-chaperone domain remains part of mature protein after trimer folding, while those of Ø29 and K1F are not. Why these tailspike are different in terms of proteolysis is unclear.

The IMC/auto-chaperone functions have been reported in both C-terminal and N-terminal domains of many proteins, and for bacterial monomeric, or trimeric autotransporter proteins.[67,68] As noted earlier, for *Bordetella* pertussis monomeric autotransporter, the type I IMC/auto-chaperone domain at the end of N-terminal β-helix helps N-terminal "passenger domain" folding

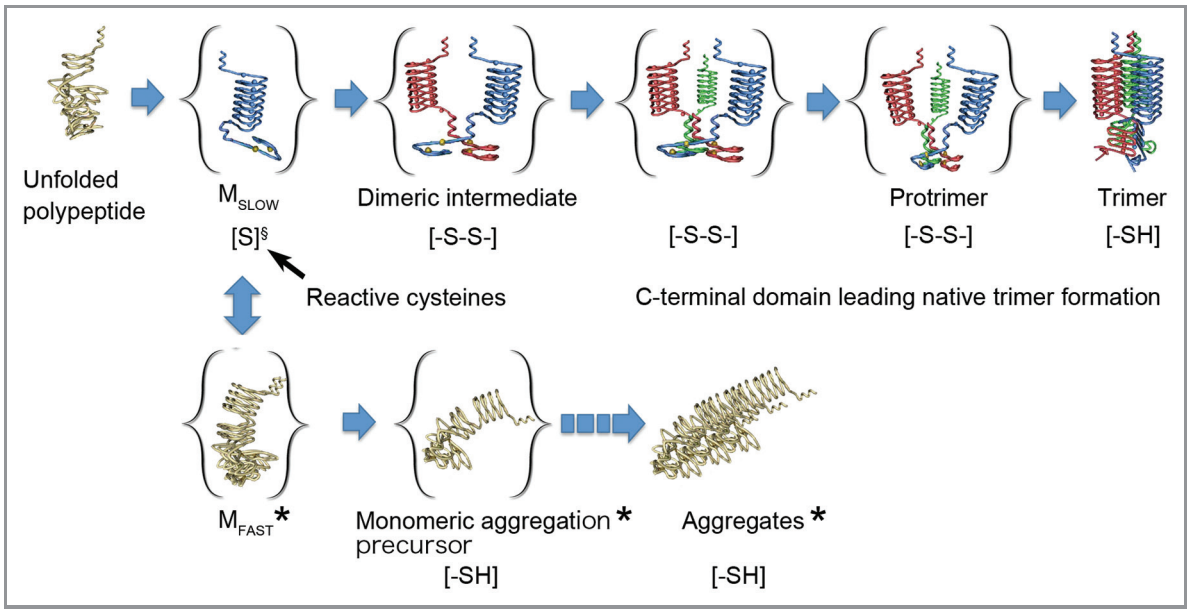

Figure 9. The folding and aggregation pathway of P22 tailspike adhesin. This schematic diagram depicts the C-terminal auto-chaperone directing productive folding and assembly. In the pathway each intermediates identified on 2-D gel are represented. The productive slower migrating monomeric (M_{SLOW}) intermediate has reactive cysteine [S]§; the non-productive faster migrating monomeric (M_{FAST}) intermediate may have intra-disulfide bond, or be more collapse; the productive dimeric and protrimer intermediates contain inter-disulfide bond [-S-S-] formed from reactive cysteines in M_{SLOW} intermediates. The aggregation prone off-pathway species (*) do not contain productive intramolecular disulfide bonds.

upon export to the cell surface.[68,69] On the other hand, both Hemophilus influenzae Hia and Yersinia enterocolitia YadA adhesin trimeric auto-chaperone domains direct trimer assembly from the end of triple N-terminal β-helix into head binding domain, as type II IMCs. These sequences remain associated and are not cleaved off.[70,71]

Many auto-chaperones and pro-peptides utilize transient disulfide bonds. The pro-peptide of Nerve growth factor (NGF) stabilizes folding intermediates to induce correct disulfide bond formation.[72] The C-terminal pro-peptide of collagen, which initiates quaternary assembly of the collagen triple helix are stabilized by intermolecular disulfide bonds.[55,57] The pro-peptide of von Willebrand factor (VWF) mediates disulfide bridge formation to complete VWF dimer folding.[73] Most reported IMCs have been classified by mode of folding as type I or type II, but structural variations suggests the presence of other types of IMC classifications in protein folding intermediates.[34,72-75]

Are transient disulfide bonds physiological or artifactual? A variety of experiments indicate a role for transient disulfide bonds in protrimer folding and assembly.[45,65] However, it has been difficult to ascertain if these are essential for efficient folding and assembly, or a byproduct of the slow folding reaction, and the proximity of the reactive 613 and 635 cysteines. The observation that double C613S/C635S are defective in folding can be interpreted as either supporting a role for transient disulfide bonds, or a role for the reactive thiol in maintaining some local critical non-disulfide bonded conformation. There are many examples of reactive cysteines in cytosols that transiently stabilize particular conformations. Li et al. has reported transient disulfide bonds-related subunits assemble in capsid protein of simian virus 40.[76] Those transient disulfide bonds are well organized in their pentamer formation under the reducing environment of the cytosol. In those proteins folding intermediates, productive intra-disulfide bonds are transferred into inter-disulfide bonds. P22 tailspike folding intermediates in the cytosol exhibit a similar mechanism between non-productive intra-disulfide bond containing monomer and productive inter-disulfide bond containing protrimer.

Regardless of the exact mechanism by which the C-terminal cysteine thiols direct the chain into the productive conformation, the interpretation that these steps are part of an auto-chaperone function as demonstrated in the K1F and Ø29 tailspikes, make it much easier to understand how these chains reach their intertwined highly stable native state.

Materials and Methods

Expression and purification of recombinant proteins. Wild type (WT), and N-terminal truncated (ΔN) N547Y of P22 tailspike were recombinantly expressed and purified from *E. coli* as previously described.[39,58] The purity of proteins was checked by SDS-PAGE. All protein concentrations were calculated by using A280 nm = 0.98.[77]

Urea unfolding and refolding. P22 tailspike at 2.0 mg/mL was unfolded in acidic urea buffer (5 M urea, 50 mM sodium citrate and 2 mM EDTA, pH 3.0) for 1 h at 24°C. Unfolded tailspike was diluted to 0.1 mg/mL into refolding buffer (0.58 M urea, 100 mM Tris and 2 mM EDTA, pH 7.6) and incubated for 15 min on wet ice to accumulate productive folding intermediates.[63] Refolding was performed as specified.

Alkylation labeling for refolding P22 tailspike. Addition of 0.1 M IAA stock solution to refolding P22 tailspike at each indicated time and further incubation was performed in the dark for a total of 1 h at 24°C. To prepare samples for non-reducing 2-D native-PAGE, IAA stock solution was added after 2 h refolding tailspike at 4°C, then incubated further for 1 h in dark. The final concentration of IAA was 30 mM for all experiments. At the end of labeling, all samples were mixed with ice-cold 3 × native-PAGE sample buffer (14 mM Tris-base, 109 mM glycine, 30% (v/v) glycerol, 0.1% bromophenol blue), and kept in wet ice until loaded into native-PAGE.

Double alkylation for mass spectrometry. To obtain alkylated peptides of monomer intermediates, IAA stock solution was added to 30 min refolding tailspike reaction and incubated for additional 2 h at 4°C. The final concentration was 30 mM IAA. After non-reducing 1-D native-PAGE, coomassie blue stained monomeric band was excised, reduced and alkylated with IAM followed by in-gel digestion methods with trypsin, Asn-N and Glu-C.

For protrimer alkylation, 2 h refolding WT, and ΔN N547Y were loaded into non-reducing 1-D native-PAGE. After electrophoresis, excised protrimer bands were immediately incubated in native-PAGE running buffer with 30 mM IAA for 1 h at 4°C in the dark, then applied for in-gel digestion and mass spectrometry analysis together with monomeric bands. Tryptic peptides were identified with QSTAR Elite mass spectrometer (Applied Biosystems) installed at the Koch Institute Proteomics lab (http://web.mit.edu/ki/facilities/core.html#biopolymers).

Reducing/non-reducing 2-D PAGE. To investigate heat stability and thiol group reactivity, refolded, and alkylated P22 tailspike was electrophoresed through non-reducing 1-D native-PAGE. Refolding and alkylation were performed as specified. After 1-D native-PAGE, the excised lane was incubated in native-PAGE running buffer in the presence, or absence of 0.1 M DTT for indicated time and temperature. After incubation, gel was loaded onto the 2-D native-gel.

Gel electrophoresis. All electrophoresis were performed as previously described.[78] The concentrations of each acrylamide/bisacrylamide in the resolving gels were 9% for native-PAGE. The concentrations of stacking gels were 3.6% for native-PAGE. All reducing and non-reducing 2-D PAGE were performed with the Bio-Rad Criterion cassette system (Bio-Rad). Native-PAGE was performed at 150 V with a constant current of 15 mA/gel for 4 h in a 4°C cold room. All proteins separated by PAGE were visualized using Krypton™ (PIERCE), and imaged on a Typhoon™ Variable Mode Imager (GE healthcare).

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

The authors would like to acknowledge Dr. Ioannis Papayannopoulos and Richard Schiavoni at the MIT Biopolymers and Proteomics Core Facility for mass spectrometry, data analysis and discussions. This work was supported by the NIH R01-GM17980 and NIH R01-AI75208.

References

1. Philipson L, Lonberg-Holm K, Pettersson U. Virus-receptor interaction in an adenovirus system. J Virol 1968; 2:1064-75; PMID:5723709
2. Kells SS, Haselkorn R. Bacteriophage T4 short tail fibers are the product of gene 12. J Mol Biol 1974; 83:473-85; PMID:4598359; http://dx.doi.org/10.1016/0022-2836(74)90508-7
3. Riede I. Receptor specificity of the short tail fibres (gp12) of T-even type Escherichia coli phages. Mol Gen Genet 1987; 206:110-5; PMID:3553859; http://dx.doi.org/10.1007/BF00326544
4. Brennan MJ, David JL, Kenimer JG, Manclark CR. Lectin-like binding of pertussis toxin to a 165-kilodalton Chinese hamster ovary cell glycoprotein. J Biol Chem 1988; 263:4895-9; PMID:3350815
5. Thomassen E, Gielen G, Schütz M, Schoehn G, Abrahams JP, Miller S, et al. The structure of the receptor-binding domain of the bacteriophage T4 short tail fibre reveals a knitted trimeric metal-binding fold. J Mol Biol 2003; 331:361-73; PMID:12888344; http://dx.doi.org/10.1016/S0022-2836(03)00755-1
6. Mishra P, Prem Kumar R, Ethayathulla AS, Singh N, Sharma S, Perbandt M, et al. Polysaccharide binding sites in hyaluronate lyase–crystal structures of native phage-encoded hyaluronate lyase and its complexes with ascorbic acid and lactose. FEBS J 2009; 276:3392-402; PMID:19438710; http://dx.doi.org/10.1111/j.1742-4658.2009.07065.x
7. Andres D, Baxa U, Hanke C, Seckler R, Barbirz S. Carbohydrate binding of Salmonella phage P22 tailspike protein and its role during host cell infection. Biochem Soc Trans 2010; 38:1386-9; PMID:20863318; http://dx.doi.org/10.1042/BST0381386
8. Andres D, Hanke C, Baxa U, Seul A, Barbirz S, Seckler R. Tailspike interactions with lipopolysaccharide effect DNA ejection from phage P22 particles in vitro. J Biol Chem 2010; 285:36768-75; PMID:20817910; http://dx.doi.org/10.1074/jbc.M110.169003
9. Weigele PR, Scanlon E, King J. Homotrimeric, beta-stranded viral adhesins and tail proteins. J Bacteriol 2003; 185:4022-30; PMID:12837775; http://dx.doi.org/10.1128/JB.185.14.4022-4030.2003
10. Steinbacher S, Seckler R, Miller S, Steipe B, Huber R, Reinemer P. Crystal structure of P22 tailspike protein: interdigitated subunits in a thermostable trimer. Science 1994; 265:383-6; PMID:8023158; http://dx.doi.org/10.1126/science.8023158
11. Steinbacher S, Miller S, Baxa U, Budisa N, Weintraub A, Seckler R, et al. 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 1997; 267:865-80; PMID:9135118; http://dx.doi.org/10.1006/jmbi.1997.0922
12. Stummeyer K, Dickmanns A, Mühlenhoff M, Gerardy-Schahn R, Ficner R. Crystal structure of the polysialic acid-degrading endosialidase of bacteriophage K1F. Nat Struct Mol Biol 2005; 12:90-6; PMID:15608653; http://dx.doi.org/10.1038/nsmb874
13. Xiang Y, Leiman PG, Li L, Grimes S, Anderson DL, Rossmann MG. Crystallographic insights into the autocatalytic assembly mechanism of a bacteriophage tail spike. Mol Cell 2009; 34:375-86; PMID:19450535; http://dx.doi.org/10.1016/j.molcel.2009.04.009
14. Goldenberg DP, Smith DH, King J. Genetic analysis of the folding pathway for the tail spike protein of phage P22. Proc Natl Acad Sci U S A 1983; 80:7060-4; PMID:6227917; http://dx.doi.org/10.1073/pnas.80.23.7060
15. Haase-Pettingell CA, King J. Formation of aggregates from a thermolabile in vivo folding intermediate in P22 tailspike maturation. A model for inclusion body formation. J Biol Chem 1988; 263:4977-83; PMID:2965152
16. Seckler R, Fuchs A, King J, Jaenicke R. Reconstitution of the thermostable trimeric phage P22 tailspike protein from denatured chains in vitro. J Biol Chem 1989; 264:11750-3; PMID:2526122
17. Fuchs A, Seiderer C, Seckler R. In vitro folding pathway of phage P22 tailspike protein. Biochemistry 1991; 30:6598-604; PMID:1828991; http://dx.doi.org/10.1021/bi00240a032
18. Miller S, Schuler B, Seckler R. A reversibly unfolding fragment of P22 tailspike protein with native structure: the isolated beta-helix domain. Biochemistry 1998; 37:9160-8; PMID:9636063; http://dx.doi.org/10.1021/bi980190e
19. Georgopoulos CP, Hendrix RW, Casjens SR, Kaiser AD. Host participation in bacteriophage lambda head assembly. J Mol Biol 1973; 76:45-60; PMID:4578100; http://dx.doi.org/10.1016/0022-2836(73)90080-6
20. van der Vies SM. Purification of the Gp31 Co-chaperonin of bacteriophage T4. Methods Mol Biol 2000; 140:51-61; PMID:11484492

21. Zweig M, Cummings DJ. Cleavage of head and tail proteins during bacteriophage T5 assembly: selective host involvement in the cleavage of a tail protein. J Mol Biol 1973; 80:505-18; PMID:4586985; http://dx.doi.org/10.1016/0022-2836(73)90418-X
22. de Beus MD, Doyle SM, Teschke CM. GroEL binds a late folding intermediate of phage P22 coat protein. Cell Stress Chaperones 2000; 5:163-72; PMID:11005374; http://dx.doi.org/10.1379/1466-1268(2000)005<0163:GBALFI>2.0.CO;2
23. Lorimer GH. A personal account of chaperonin history. Plant Physiol 2001; 125:38-41; PMID:11154291; http://dx.doi.org/10.1104/pp.125.1.38
24. Doyle SM, Anderson E, Parent KN, Teschke CM. A concerted mechanism for the suppression of a folding defect through interactions with chaperones. J Biol Chem 2004; 279:17473-82; PMID:14764588; http://dx.doi.org/10.1074/jbc.M400467200
25. Sather SK, King J. Intracellular trapping of a cytoplasmic folding intermediate of the phage P22 tailspike using iodoacetamide. J Biol Chem 1994; 269:25268-76; PMID:7929218
26. Danner M, Fuchs A, Miller S, Seckler R. Folding and assembly of phage P22 tailspike endorhamnosidase lacking the N-terminal, head-binding domain. Eur J Biochem 1993; 215:653-61; PMID:8354271; http://dx.doi.org/10.1111/j.1432-1033.1993.tb18076.x
27. Ikemura H, Takagi H, Inouye M. Requirement of prosequence for the production of active subtilisin E in Escherichia coli. J Biol Chem 1987; 262:7859-64; PMID:3108260
28. Chen YJ, Inouye M. The intramolecular chaperone-mediated protein folding. Curr Opin Struct Biol 2008; 18:765-70; PMID:18973809; http://dx.doi.org/10.1016/j.sbi.2008.10.005
29. Fessler JH, Fessler LI. Biosynthesis of procollagen. Annu Rev Biochem 1978; 47:129-62; PMID:354493; http://dx.doi.org/10.1146/annurev.bi.47.070178.001021
30. Junker M, Schuster CC, McDonnell AV, Sorg KA, Finn MC, Berger B, et al. Pertactin beta-helix folding mechanism suggests common themes for the secretion and folding of autotransporter proteins. Proc Natl Acad Sci U S A 2006; 103:4918-23; PMID:16549796; http://dx.doi.org/10.1073/pnas.0507923103
31. Rutherford N, Charbonneau ME, Berthiaume F, Betton JM, Mourez M. The periplasmic folding of a cysteineless autotransporter passenger domain interferes with its outer membrane translocation. J Bacteriol 2006; 188:4111-6; PMID:16707702; http://dx.doi.org/10.1128/JB.01949-05
32. Schulz EC, Dickmanns A, Urlaub H, Schmitt A, Mühlenhoff M, Stummeyer K, et al. Crystal structure of an intramolecular chaperone mediating triple-beta-helix folding. Nat Struct Mol Biol 2010; 17:210-5; PMID:20118935; http://dx.doi.org/10.1038/nsmb.1746
33. Baxa U, Steinbacher S, Weintraub A, Huber R, Seckler R. Mutations improving the folding of phage P22 tailspike protein affect its receptor binding activity. J Mol Biol 1999; 293:693-701; PMID:10543960; http://dx.doi.org/10.1006/jmbi.1999.3165
34. Schwarzer D, Stummeyer K, Gerardy-Schahn R, Mühlenhoff M. Characterization of a novel intramolecular chaperone domain conserved in endosialidases and other bacteriophage tail spike and fiber proteins. J Biol Chem 2007; 282:2821-31; PMID:17158460; http://dx.doi.org/10.1074/jbc.M609543200
35. Webber T, Gurung S, Saul J, Baker T, Spatara M, Freyer M, et al. The C-terminus of the P22 tailspike protein acts as an independent oligomerization domain for monomeric proteins. Biochem J 2009; 419:595-602; PMID:19196242; http://dx.doi.org/10.1042/BJ20081449
36. Raso SW, Clark PL, Haase-Pettingell C, King J, Thomas GJ, Jr.. Distinct cysteine sulfhydryl environments detected by analysis of Raman S-hh markers of Cys–>Ser mutant proteins. J Mol Biol 2001; 307:899-911; PMID:11273709; http://dx.doi.org/10.1006/jmbi.2001.4476
37. Goldenberg D, King J. Trimeric intermediate in the in vivo folding and subunit assembly of the tail spike endorhamnosidase of bacteriophage P22. Proc Natl Acad Sci U S A 1982; 79:3403-7; PMID:6954486; http://dx.doi.org/10.1073/pnas.79.11.3403
38. Kreisberg JF, Betts SD, Haase-Pettingell C, King J. The interdigitated beta-helix domain of the P22 tailspike protein acts as a molecular clamp in trimer stabilization. Protein Sci 2002; 11:820-30; PMID:11910025; http://dx.doi.org/10.1110/ps.3440102
39. Weigele PR, Haase-Pettingell C, Campbell PG, Gossard DC, King J. Stalled folding mutants in the triple beta-helix domain of the phage P22 tailspike adhesin. J Mol Biol 2005; 354:1103-17; PMID:16289113; http://dx.doi.org/10.1016/j.jmb.2005.10.007
40. Simkovsky R, King J. 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 2006; 103:3575-80; PMID:16505375; http://dx.doi.org/10.1073/pnas.0509087103
41. Smith DH, King J. Temperature-sensitive mutants blocked in the folding or subunit assembly of the bacteriophage P22 tail spike protein. III. Intensive polypeptide chains synthesized at 39 degrees C. J Mol Biol 1981; 145:653-76; PMID:7265218; http://dx.doi.org/10.1016/0022-2836(81)90308-9
42. Sturtevant JM, Yu MH, Haase-Pettingell C, King J. Thermostability of temperature-sensitive folding mutants of the P22 tailspike protein. J Biol Chem 1989; 264:10693-8; PMID:2525128
43. Haase-Pettingell C, King J. Prevalence of temperature sensitive folding mutations in the parallel beta coil domain of the phage P22 tailspike endorhamnosidase. J Mol Biol 1997; 267:88-102; PMID:9096209; http://dx.doi.org/10.1006/jmbi.1996.0841
44. King J, Haase-Pettingell C, Robinson AS, Speed M, Mitraki A. Thermolabile folding intermediates: inclusion body precursors and chaperonin substrates. FASEB J 1996; 10:57-66; PMID:8566549
45. Robinson AS, King J. Disulphide-bonded intermediate on the folding and assembly pathway of a non-disulphide bonded protein. Nat Struct Biol 1997; 4:450-5; PMID:9187652; http://dx.doi.org/10.1038/nsb0697-450
46. Danek BL, Robinson AS. Nonnative interactions between cysteines direct productive assembly of P22 tailspike protein. Biophys J 2003; 85:3237-47; PMID:14581223; http://dx.doi.org/10.1016/S0006-3495(03)74741-9
47. Polgár L. On the mode of activation of the catalytically essential sulfhydryl group of papain. Eur J Biochem 1973; 33:104-9; PMID:4691346; http://dx.doi.org/10.1111/j.1432-1033.1973.tb02660.x
48. Wallenfels K, Eisele B. Stereospecific alkylation with asymmetric reagents. Eur J Biochem 1968; 3:267-75; PMID:5650849; http://dx.doi.org/10.1111/j.1432-1033.1968.tb19526.x
49. Kim J, Robinson AS. Dissociation of intermolecular disulfide bonds in P22 tailspike protein intermediates in the presence of SDS. Protein Sci 2006; 15:1791-3; PMID:16751612; http://dx.doi.org/10.1110/ps.062197206
50. Weissman JS, Kim PS. Reexamination of the folding of BPTI: predominance of native intermediates. Science 1991; 253:1386-93; PMID:1716783; http://dx.doi.org/10.1126/science.1716783
51. Creighton. Folding pathways determined using disulfide bonds. New York: W.H. Freeman, 1992.
52. Goldenberg DP. Native and non-native intermediates in the BPTI folding pathway. Trends Biochem Sci 1992; 17:257-61; PMID:1380192; http://dx.doi.org/10.1016/0968-0004(92)90405-X
53. Darby NJ, Creighton TE. Dissecting the disulphide-coupled folding pathway of bovine pancreatic trypsin inhibitor. Forming the first disulphide bonds in analogues of the reduced protein. J Mol Biol 1993; 232:873-96; PMID:7689114; http://dx.doi.org/10.1006/jmbi.1993.1437
54. Creighton TE, Zapun A, Darby NJ. Mechanisms and catalysts of disulfide bond formation in proteins. Trends Biotechnol 1995; 13:18-23; PMID:7765801; http://dx.doi.org/10.1016/S0167-7799(00)88896-4
55. Koivu J. Disulfide bonding as a determinant of the molecular composition of types I, II and III procollagen. FEBS Lett 1987; 217:216-20; PMID:3595852; http://dx.doi.org/10.1016/0014-5793(87)80666-X
56. Engel J, Prockop DJ. The zipper-like folding of collagen triple helices and the effects of mutations that disrupt the zipper. Annu Rev Biophys Biophys Chem 1991; 20:137-52; PMID:1867713; http://dx.doi.org/10.1146/annurev.bb.20.060191.001033
57. Khoshnoodi J, Cartailler JP, Alvares K, Veis A, Hudson BG. Molecular recognition in the assembly of collagens: terminal noncollagenous domains are key recognition modules in the formation of triple helical protomers. J Biol Chem 2006; 281:38117-21; PMID:17082192; http://dx.doi.org/10.1074/jbc.R600025200
58. Haase-Pettingell C, Betts S, Raso SW, Stuart L, Robinson A, King J. Role for cysteine residues in the in vivo folding and assembly of the phage P22 tailspike. Protein Sci 2001; 10:397-410; PMID:11266625; http://dx.doi.org/10.1110/ps.34701
59. Walter M, Fiedler C, Grassl R, Biebl M, Rachel R, Hermo-Parrado XL, et al. Structure of the receptor-binding protein of bacteriophage det7: a podoviral tail spike in a myovirus. J Virol 2008; 82:2265-73; PMID:18077713; http://dx.doi.org/10.1128/JVI.01641-07
60. Speed MA, Wang DI, King J. Multimeric intermediates in the pathway to the aggregated inclusion body state for P22 tailspike polypeptide chains. Protein Sci 1995; 4:900-8; PMID:7663345; http://dx.doi.org/10.1002/pro.5560040509
61. Speed MA, Wang DI, King J. Specific aggregation of partially folded polypeptide chains: the molecular basis of inclusion body composition. Nat Biotechnol 1996; 14:1283-7; PMID:9631094; http://dx.doi.org/10.1038/nbt1096-1283
62. Speed MA, Morshead T, Wang DI, King J. Conformation of P22 tailspike folding and aggregation intermediates probed by monoclonal antibodies. Protein Sci 1997; 6:99-108; PMID:9007981; http://dx.doi.org/10.1002/pro.5560060111
63. Betts SD, King J. Cold rescue of the thermolabile tailspike intermediate at the junction between productive folding and off-pathway aggregation. Protein Sci 1998; 7:1516-23; PMID:9684883; http://dx.doi.org/10.1002/pro.5560070704

64. Betts S, King J. There's a right way and a wrong way: in vivo and in vitro folding, misfolding and subunit assembly of the P22 tailspike. Structure 1999; 7:R131-9; PMID:10404587; http://dx.doi.org/10.1016/S0969-2126(99)80078-1
65. Danek BL, Robinson AS. P22 tailspike trimer assembly is governed by interchain redox associations. Biochim Biophys Acta 2004; 1700:105-16; PMID:15210130
66. Mitraki A, King J. Amino acid substitutions influencing intracellular protein folding pathways. FEBS Lett 1992; 307:20-5; PMID:1639189; http://dx.doi.org/10.1016/0014-5793(92)80894-M
67. Cotter SE, Surana NK, St Geme JW, 3rd. Trimeric autotransporters: a distinct subfamily of autotransporter proteins. Trends Microbiol 2005; 13:199-205; PMID:15866036; http://dx.doi.org/10.1016/j.tim.2005.03.004
68. Nishimura K, Tajima N, Yoon YH, Park SY, Tame JR. Autotransporter passenger proteins: virulence factors with common structural themes. J Mol Med (Berl) 2010; 88:451-8; PMID:20217035; http://dx.doi.org/10.1007/s00109-010-0600-y
69. Oliver DC, Huang G, Nodel E, Pleasance S, Fernandez RC. A conserved region within the Bordetella pertussis autotransporter BrkA is necessary for folding of its passenger domain. Mol Microbiol 2003; 47:1367-83; PMID:12603741; http://dx.doi.org/10.1046/j.1365-2958.2003.03377.x
70. Yeo HJ, Cotter SE, Laarmann S, Juehne T, St Geme JW, 3rd, Waksman G. Structural basis for host recognition by the Haemophilus influenzae Hia autotransporter. EMBO J 2004; 23:1245-56; PMID:15029242; http://dx.doi.org/10.1038/sj.emboj.7600142
71. Cotter SE, Surana NK, Grass S, St Geme JW, 3rd. Trimeric autotransporters require trimerization of the passenger domain for stability and adhesive activity. J Bacteriol 2006; 188:5400-7; PMID:16855229; http://dx.doi.org/10.1128/JB.00164-06
72. Rattenholl A, Ruoppolo M, Flagiello A, Monti M, Vinci F, Marino G, et al. Pro-sequence assisted folding and disulfide bond formation of human nerve growth factor. J Mol Biol 2001; 305:523-33; PMID:11152610; http://dx.doi.org/10.1006/jmbi.2000.4295
73. Rosenberg JB, Haberichter SL, Jozwiak MA, Vokac EA, Kroner PA, Fahs SA, et al. The role of the D1 domain of the von Willebrand factor propeptide in multimerization of VWF. Blood 2002; 100:1699-706; PMID:12176890; http://dx.doi.org/10.1182/blood-2002-03-0789
74. Min CY, Qiao ZS, Feng YM. Unfolding of human proinsulin. Intermediates and possible role of its C-peptide in folding/unfolding. Eur J Biochem 2004; 271:1737-47; PMID:15096212; http://dx.doi.org/10.1111/j.1432-1033.2004.04079.x
75. Rozenfeld R, Muller L, El Messari S, Llorens-Cortes C. The C-terminal domain of aminopeptidase A is an intramolecular chaperone required for the correct folding, cell surface expression, and activity of this monozinc aminopeptidase. J Biol Chem 2004; 279:43285-95; PMID:15263000; http://dx.doi.org/10.1074/jbc.M404369200
76. Li PP, Nakanishi A, Clark SW, Kasamatsu H. Formation of transitory intrachain and interchain disulfide bonds accompanies the folding and oligomerization of simian virus 40 Vp1 in the cytoplasm. Proc Natl Acad Sci U S A 2002; 99:1353-8; PMID:11805304; http://dx.doi.org/10.1073/pnas.032668699
77. Sauer RT, Krovatin W, Poteete AR, Berget PB. Phage P22 tail protein: gene and amino acid sequence. Biochemistry 1982; 21:5811-5; PMID:6295443; http://dx.doi.org/10.1021/bi00266a014
78. Betts S, Speed M, King J. Detection of early aggregation intermediates by native gel electrophoresis and native western blotting. Methods Enzymol 1999; 309:333-50; PMID:10507034; http://dx.doi.org/10.1016/S0076-6879(99)09024-2