Buried hydrophobic side-chains essential for the folding of the parallel β -helix domains of the P22 tailspike

SCOTT BETTS,¹ CAMERON HAASE-PETTINGELL, KRISTEN COOK, AND JONATHAN KING

Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

(RECEIVED February 18, 2004; FINAL REVISION May 27, 2004/SCAP; ACCEPTED June 3, 2004)

Abstract

The processive β -strands and turns of a polypeptide parallel β -helix represent one of the topologically simplest β -sheet folds. The three subunits of the tailspike adhesin of phage P22 each contain 13 rungs of a parallel β -helix followed by an interdigitated section of triple-stranded β -helix. Long stacks of hydrophobic residues dominate the elongated buried core of these two β -helix domains and extend into the core of the contiguous triple β -prism domain. To test whether these side-chain stacks represent essential residues for driving the chain into the correct fold, each of three stacked phenylalanine residues within the buried core were substituted with less bulky amino acids. The mutant chains with alanine in place of phenylalanine were defective in intracellular folding. The chains accumulated exclusively in the aggregated inclusion body state regardless of temperature of folding. These severe folding defects indicate that the stacked phenylalanine residues are essential for correct parallel β -helix folding. Replacement of the same phenylalanine residues with valine or leucine also impaired folding *in vivo*, but with less severity. Mutants were also constructed in a second buried stack that extends into the intertwined triple-stranded β -helix and contiguous β -prism regions of the protein. These mutants exhibited severe defects in later stages of chain folding or assembly, accumulating as misfolded but soluble multimeric species. The results indicate that the formation of the buried hydrophobic stacks is critical for the correct folding of the parallel β -helix, triple-stranded β -helix, and β -prism domains in the tailspike protein.

Keywords: protein folding; β -sheet; β -helix; β -prism; hydrophobic stacks; buried residues

Deciphering how amino acid sequences direct the folding of polypeptide chains into the β -sheet conformation has proven surprisingly difficult. The topologically simplest class of β -sheet proteins are the parallel β -helix proteins, in which the β -strands pack processively with respect to the amino acid sequence (Yoder et al. 1993; Jurnak et al. 1994; Pickersgill et al. 1998). Single chain β -helical motifs are

found among polysaccharide-binding proteins, some of which are virulence factors associated with bacterial pathogens (Bradley et al. 2001). Triple chain β -helical folding motifs have been reported for viral adhesins, including the penton fiber of adenovirus (Mitraki et al. 2002; Weigele et al. 2003). The parallel β -helical fold is also a candidate for the polypeptide chain fold in amyloid fibrils from the Alzheimer's peptide (Wetzel 2002; Williams et al. 2004).

A large number of amino acid residues important for controlling β -helix folding at high temperatures have been identified in the P22 tailspike (King et al. 1996). Such temperature-sensitive folding (tsf) mutants identify residues in the parallel β -helix domain that are required for chain folding at the higher temperature range of phage growth (Goldenberg et al. 1983; Danner and Seckler 1993; King et al. 1996; Haase-Pettingell and King 1997). These sites are pre-

Reprint requests to: Jonathan King, Department of Biology 68-330, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139, USA; e-mail: jaking@mit.edu; fax: (617) 252-1843.

¹Present address: Syngenta Biotechnology, Inc., 3054 Cornwallis Road, Research Triangle Park, North Carolina 27709, USA.

Abbreviations: tsf, temperature-sensitive folding; CD, circular dichroism spectroscopy; SDS, sodium dodecylsulfate; IPTG, isopropyl- β -D-thiogalactopyranoside; IB, inclusion body.

Article and publication are at <http://www.proteinscience.org/cgi/doi/10.1110/ps.04676704>.

dominantly surface residues at turns, loops, or strand transitions (Haase-Pettingell and King 1997). Because these mutant chains fold into biologically active thermostable tailspike proteins at low temperature, the wild-type residues at the tsf positions do not carry the essential information directing β -helix folding (Goldenberg and King 1981; Sturtevant et al. 1989; Danner and Seckler 1993).

Jurnak and colleagues identified stacks of hydrophobic residues as a prominent feature of β -helical proteins (Jurnak et al. 1994). Disruption of the packing of hydrophobic side-chains in the buried core of β -sheet and β -barrel proteins disrupts their folding or destabilizes the native state (Sauer 1996; Gunasekaran et al. 2001). Thus, the buried and stacked hydrophobic side-chains are candidates for residues essential for correct β -helix folding.

The 210-kD tailspike trimer of *Salmonella* phage P22 is the adhesin for binding to the host cell surface (Steinbacher et al. 1996). Ribbon diagrams of amino-terminal (residues 1–113) and carboxy-terminal portions (residues 109–666) of the P22 tailspike trimer are shown in Figure 1A. Residues 143–540 form the single right-handed parallel β -helix domain. The three single parallel β -helices line up side-by-side in the native trimer (Steinbacher et al. 1994; Seckler 1998). After residue 540, the three polypeptide chains wrap around each other to form an interdigitated triple-stranded β -helix (Fig. 1C, top). This region extends approximately from residues 541–557. Following this intertwined region, residues 558–619 of the chains fold back on themselves, forming the sides of a triple β -prism, but still maintaining a single buried core (Kreisberg et al. 2000). Before diverging, the chains wrap around each other once again. This region contains Cys 613 and Cys 635. In the native trimer, these two cysteine residues from each subunit are in close proximity and together generate a ring-like motif of six cysteine residues that we have termed the cysteine annulus. After this the chains form the carboxy-terminal caudal fin of antiparallel β -sheets.

The parallel β -helix domain consists of 13 consecutive coils or rungs, with individual rungs ranging in length from 20–60 amino acids. The rungs are flattened and stack up to form an elongated ribbon-like structure; each rung/coil contributes β -strands to each of three parallel β -sheets.

Figure 1B shows two of three hydrophobic side-chain stacks, each containing 13 consecutive hydrophobic residues contributing to the buried core of the parallel tailspike β -helix. For ease of nomenclature we have named the stacks after the sheet from which they are derived: buried stack A (cyan) is in β -sheet A and buried stack C (green) is in β -sheet C. The three stacked phenylalanines at positions 284, 308, and 336 represent a local packing motif only rarely observed in globular proteins, where edge-to-face orientation is more common (Burley and Petsko 1988).

Note that two types of β -helix domains (single- and triple-stranded) and a β -prism domain are present in the

tailspike structure (Steinbacher et al. 1994; Seckler 1998). In the *single parallel β -helix* each chain forms its own 13-runged coil; this is the dominant structural motif of the tailspike. The short *triple-stranded β -helix* starts after residue 540, where the individual polypeptide chains emerge from the single β -helix coils, and then wrap around each other. The chains separate again starting at about residue 559 to form three faces of the *triple β -prism* in which each chain contributes hydrophobic side-chains to the common core (Fig. 1C).

Partially folded intermediates identified in the *in vivo* and *in vitro* folding pathway are shown in Figure 2. Spectroscopic studies by Seckler and co-workers (Fuchs et al. 1991; Miller et al. 1998) suggest that an early rapid change in fluorescence and CD signals is caused by a rapid collapse of unfolded chains [U] to a partially folded monomeric conformation [I_o]. This rapidly transforms to the earliest structured species [I] presumed to be in the β -helix fold. In the higher range of physiological temperatures, this early intermediate [I] generates a misfolded conformation [I^*] that self-associates into oligomeric aggregation intermediates and ultimately ends up in the aggregated electron-dense inclusion body state (Speed et al. 1995, 1996; Betts et al. 1999).

The tsf mutations appear to act by further destabilizing this already thermolabile single-chain folding intermediate [I] (Haase-Pettingell and King 1988; Danner and Seckler 1993; Schuler and Seckler 1998). Global suppressors of the tsf mutations stabilize [I] or otherwise inhibit association of these intermediates along the aggregation pathway (Mitraki et al. 1991b; Beissinger et al. 1995). This partially folded single-chain intermediate [I] has been identified directly using native gel electrophoresis, both *in vivo* (Goldenberg et al. 1983; Haase-Pettingell and King 1988) and *in vitro* (Betts and King 1998).

As folding proceeds, the subunits develop enough structure to associate into a partially folded protrimer intermediate, in which the strands are associated but not intertwined into the mature triple-stranded β -helix (Goldenberg et al. 1983; Fuchs et al. 1991). The intertwining appears to function as a molecular clamp, raising the melting temperature more than 40°C to 88°C (Kreisberg et al. 2002). The protrimer intermediate contains transient disulfide bonds, which may keep the chains in register before the intertwining step (Robinson and King 1997). Chains carrying amino acid substitutions disrupting these later steps in tailspike folding do not aggregate into the inclusion body state but assemble into protrimer-like species (Schwarz and Berget 1989; Haase-Pettingell et al. 2001; Kreisberg et al. 2002).

Temperature-sensitive folding mutations have rarely been recovered at sites forming the buried hydrophobic core of the tailspike (Haase-Pettingell and King 1997). The failure to recover tsf mutations affecting the hydrophobic buried stack residues suggested that these residues are absolutely

Figure 1. Features of the P22 tailspike protein. (A) Ribbon diagram of the P22 tailspike protein (Steinbacher et al. 1994, 1996). The trimer is oriented with the amino termini up and the carboxyl termini down. The subunits are in different colors. The single parallel β -helix domain (residues 143–540) is *below* the amino-terminal capsid-binding domain (Steinbacher et al. 1996). The narrower *bottom* portion of the structure includes the triple-stranded β -helix (residues 541–558), an intertwined region in which the three independent strands wrap around each other. Following this region, the chains fold back on themselves for 5 rungs—but still with a common buried core—called the triple β -prism (residues 559–619) and then twist into the cysteine annulus. They then separate to form the β -sheet caudal fins. Residues targeted by mutagenesis are labeled; these and other residues that form the triple β -prism subunit interface are shown in space-filling mode. (B) Depiction of the single parallel β -helix domain showing two of the three interior side-chain stacks, stack A (cyan) and stack C (green); the backbone atoms are shown using CPK color scheme residues mutated in this paper are indicated. (C) Axial views of three stacked coils that form the triple-stranded β -helix and the triple β -prism domains. Isoleucine and leucine side-chains from each subunit can be seen at the *center* of each coil forming the hydrophobic core of the interdigitated domain. Sequences in connecting loops and turns are omitted. Purple stars identify Ile 548 (*top*) and Ile 560 (*middle*). Yellow stars mark Glu 543.

essential at low and high temperatures for correct β -helix folding or tailspike stability. Here we report the results of experiments investigating the role of buried side-chain stacks in tailspike folding *in vivo*.

Results

Effects of decreased side-chain volume at aromatic positions within the parallel β -helix core

The intracellular folding and assembly of mutant tailspike chains was examined by monitoring the partitioning of the

expressed polypeptide chains between the competing productive folding and nonproductive inclusion body pathways. Five phenylalanine and two leucine residues, located within two side-chain stacks and two different β -sheets, were targeted for investigation.

Stacked phenylalanine residues

Residues Phe 181, Phe 284, Phe 308, and Phe 336 participate in one stack within β -sheet C (Figs. 1B and 3, green).

Figure 2. Tailspike folding and aggregation pathways, in vivo and in vitro. Newly synthesized chains released from ribosomes form a single chain partially folded intermediate [I] (Clark and King 2001). For chains refolded out of denaturant an earlier intermediate [I_o] is suggested from spectroscopy (Fuchs et al. 1991). The [I] intermediate folds further to a species that can self-associate to form the protrimer intermediate [pT], in which the chains are associated but not fully folded (Goldenberg and King 1982; Fuchs et al. 1991; Betts and King 1999). The protrimer, a disulfide bonded species, then transforms to the native tailspike [Nt]. Interchain disulfide bonds stabilize the protrimer, but are reduced in the transition to the native trimer (Robinson and King 1997). The early single chain species [I] is thermolabile, and as temperature increases, forms a species [I^*] that associates into multimers that polymerize into the kinetically trapped inclusion body state.

Phe 181 is in rung 2; its side-chain is stacked between the side-chains of Val 160 and Ile 264. The other three phenylalanine residues form a pancake-like phenylalanine stack of aromatic rings that spans rungs 4–6. Phe 336 is followed by

Thr 364 in rung 7. The phenylalanine side-chains at these four sites were replaced with hydrophobic side-chains of different volumes: leucine, valine, alanine, as well as a polar residue, threonine. Phe 284 was also replaced with isoleu-

rung	residue	Stack A				Stack B				turn/loop	Stack C				turn/loop
		i	o	i	o	i	o	i	o		i	o	i	o	
1	143	G	L	L	-	I	DRD	Y	N	-	F	Y	G	E	T
2	165	L	T	I	E	-	C	K	A	K	F	I	G	-	D
3	188	S	R	I	A	G	V	F	M	E	-	-	-	-	dorsal fin
4	270	V	E	V	H	R	A	S	G	L	-	-	-	-	MAG
5	290	C	K	M	V	D	A	N	N	P	S	G	G	-	KDGI
6	319	N	Y	V	I	G	R	T	S	Y	G	-	-	-	SVSS
7	348	-	G	V	I	G	F	T	S	Y	R	A	-	-	GESG
8	379	L	Q	F	R	D	S	V	V	I	Y	P	-	-	VWDG
9	426	H	L	I	D	N	L	L	V	R	G	A	-	-	LGVG
10	448	M	Y	V	S	N	I	T	V	E	D	C	-	-	AGSG
11	470	S	V	F	T	N	I	A	I	I	D	T	-	-	NTKDFQANQ
12	496	C	R	V	N	G	L	R	L	-	-	-	-	-	IGIRSTDGQGLT
13	521	S	T	V	S	G	I	T	G	-	-	-	-	-	MVDPSR
14	538	A	N	L	-	-	-	-	-	-	-	-	-	-	

Figure 3. Structure-based sequence alignment of the parallel β -helix domain. The figure is a simplified alignment to guide the reader. The first column indicates the rung number (1–14), and the second column indicates the position of the first residue in each rung. The column labels indicate if the side-chains point to the inside (i) or the outside (o) of the β -helix. Colors identify residues in the three β -sheets (cyan, sheet A; yellow, sheet B; green, sheet C). The boxes identify the seven residues targeted by site-directed mutagenesis as well as the two stacks shown in corresponding colors in Figure 1B. Residues that are the sites of tsf mutants are shown in red, i.e., tsfT307A and tsfE309V, and the locations of the two global suppressors of tsf mutations (V331A and A334V) are in blue.

Figure 4. Substitutions of stacked phenylalanine residues in parallel β -sheet C; SDS gel electrophoresis of the soluble and aggregated chains. Tailspike expression in *E. coli* was induced by addition of IPTG, and cells were harvested after 2 h expression at 30°C. Supernatant and pellet fractions of cell lysates expressing single amino acid replacements at tailspike residues Phe 284 (A), Phe 308 (B), and Phe 336 (C) were electrophoresed through SDS-polyacrylamide gels. The gels were stained with Coomassie blue. The small amount of native tailspike in the pellet is due to incomplete lysis of the cells.

cine. The effects of the substitutions on intracellular tailspike folding were examined.

The mutant and wild-type tailspike polypeptide chains were expressed from a pET plasmid in *Escherichia coli* at 30°C. The resulting lysate was fractionated into supernatant and pellet fractions. Samples were mixed with SDS sample buffer—but not heated—and electrophoresed through polyacrylamide gels. Under these conditions native trimeric tailspikes remain native, whereas folding intermediates and inclusion bodies are dissociated into SDS-polypeptide chain complexes. Figure 4 shows representative SDS-polyacrylamide gels. These gels demonstrate the effects of substituting residues with decreasing side-chain volume and polarity for phenylalanine stack residues 284, 308, and 336, on the folding of tailspike chains between soluble trimer and pelleted inclusion body fractions. A quantitative analysis of the partitioning is shown in Figure 5.

In control cultures, wild-type tailspike chains accumulated primarily as soluble native (SDS-resistant) trimers.

Substitution of Phe 284 with alanine, isoleucine, or threonine completely blocked formation of native tailspikes at 30°C (Figs. 4 and 5). Substitution with leucine also increased partitioning into the inclusion body pathway, but to a lesser degree, resulting in the maturation of a small fraction of total tailspike chains into SDS-resistant trimers.

Figure 5. In vivo partitioning of mutant tailspike chains between productive folding and aggregated inclusion bodies. The levels of tailspike chains were determined by laser densitometry as described in Materials and Methods. The data are plotted as averages of duplicate cultures expressing tailspike chains that contained single amino acid replacements at residues Phe 181 (A), Phe 284 (B), Phe 308 (C), and Phe 336 (D). Cultures expressing wild-type tailspike were included in each experiment as controls. The SDS-polypeptide chain complexes found in the pellet represent chains solubilized from the inclusion body state.

Substitution of Phe 308 with alanine or threonine completely blocked tailspike folding (Figs. 4 and 5). F308L was slightly less severe, allowing accumulation of trace amounts of native trimers to levels detectable by visual inspection of the gel in Figure 4. However, these mutant trimers were not detected above background by laser densitometry (Fig. 5). F308V was the least severe replacement at this site, but chains carrying this mutation still strongly favored the inclusion body pathway.

Substitution of Phe 336 with alanine or threonine completely blocked productive folding. F336V also increased partitioning onto the inclusion body pathway, but allowed a fraction of chains to reach the native trimeric conformation. Replacement of Phe 336 with leucine did not significantly increase partitioning onto the inclusion body pathway.

Replacement of Phe 181 with alanine resulted in reduced yields of native tailspikes, with about half the total chains ending up as tailspikes and half as inclusion bodies. Replacement of Phe 181 with leucine or valine had little or no effect on the partitioning between folding and inclusion body pathways (Fig. 5).

Triple mutants were also constructed with alanines in place of the three phenylalanines and with threonines substituted for the phenylalanines. As expected, these triple mutant chains failed to form native trimers and accumulated in the inclusion body state (data not shown).

As an additional test, we isolated the inclusion bodies, solubilized them with urea, and tried refolding mutant chains recovered from inclusion bodies under conditions that yield SDS-resistant trimeric tailspikes with wild-type chains. Inclusion bodies containing mutant chains were dissolved by incubation in 6 M urea (at pH 8.0) and 1 mM DTT, and refolding was initiated by dilution into buffer. After 24 h at 20°C samples of the refolding reactions were electrophoresed through native gels and probed by antinative and anti-intermediate antibodies on a Western blot (Friguet et al. 1984). In the control refolding reaction, wild-type tailspike chains recovered from inclusion bodies refolded to native tailspikes. No native tailspikes were formed in refolding reactions, even after 24 h, using chains containing the single mutation F308A (data not shown).

Examination of in vivo folding as a function of temperature

The tailspike chains with substitutions at the buried phenylalanine sites and expressed at 30°C, failed to fold efficiently in vivo. These mutant chains were tested to determine whether they could be rescued by expression at lower or higher temperatures. Cells containing the tailspike expression vector pET-tsp encoding the F308V or F308A mutant chains were grown at 37°C. The culture was induced, and samples were shifted immediately to 18°C, 30°C, 37°C, and 39°C. After further incubation, cell samples were harvested

and resuspended in lysis buffer. The lysates were centrifuged to generate pellet and supernatant fractions, and the resulting samples—without heating—were electrophoresed through an SDS-polyacrylamide gel. The gel was analyzed by laser densitometry, and the data were corrected for differences in cell density as determined by OD₆₀₀.

The resulting bar graph, shown in Figure 6, shows the efficiencies of chain folding into SDS-resistant tailspikes as a function of temperature. The percentage of wild-type native tailspike with respect to total wild-type tailspike chains at each temperature is defined as 100%. The amounts of native (SDS-resistant) trimer produced by cultures expressing the F308V and F308A mutant chains grown at 30°C, 37°C, and 39°C were close to background. However, at 18°C, the culture expressing the F308V mutant chains had a native tailspike yield that was similar to the yield observed in the wild-type strain. However, the F308A substitution was not rescued by growth at lower temperature, indicating that this substitution is absolutely lethal. The F308V substitution produced native-like tailspike chains, but the stability of these chains has not yet been examined.

Replacement of buried hydrophobic residues in β -sheet A

Fourteen hydrophobic residues form a well-defined stack across from the phenylalanine stack. Three hydrophobic residues located in rungs 1, 8, and 14 were selected for further study (cyan shading in Figs. 1B and 3, boxed residues). Leu 144 is located in the first rung and is followed in

Figure 6. Effect of temperature on in vivo folding of tailspike chains. Cultures expressing wild-type, F308A, and F308V tailspike chains were induced with IPTG and portions were immediately transferred to different temperatures for continued incubation (18°C for 5 h, 30°C for 3.5 h; and 37° and 39°C for 3 h). Cultures were harvested, lysed, and separated into supernatant and pellet fractions. The wild-type value, indicated by the solid gray bar, is defined as 100%. F308A strain, black bar with white spots; F308V strain, white bar with black spots.

the stack by an isoleucine residue; Phe 381 in rung 8 is stacked between a valine and an isoleucine; and Leu 540 follows a valine and ends the stack. The contribution of these residues to tailspike folding was investigated by constructing site-specific mutants. As in the experiments described in the previous section, mutant tailspike chains were expressed in *E. coli* for 2 h at 30°C and the partitioning of tailspike chains between soluble and insoluble fractions was measured using SDS-PAGE and laser densitometry.

Replacement of Leu 144 with alanine resulted in a severe defect in tailspike folding, although a small fraction of chains accumulated as SDS-resistant tailspikes (Fig. 7).

Figure 7. In vivo partitioning of tailspike chains containing substitutions in a single stack in parallel β -sheet A. Cultures of Leu 144 (*top*), Phe 381 (*middle*), and Phe 540 (*bottom*) were fractionated into low-speed pellet and supernatant fractions, and the polypeptide chains electrophoresed through SDS-polyacrylamide gels. The majority of the chains recovered from the supernatant migrated similarly to SDS-resistant native-like trimers. The chains in the pellet migrated as SDS-soluble inclusion body chains. The positions of the targeted substitutions are indicated in Figure 1B and the alignment is Figure 3 (boxed residues). The levels of tailspike chains were determined by laser densitometry as described in Materials and Methods. The data are plotted as averages of duplicate cultures expressing tailspike chains. Cultures expressing wild-type tailspike were included in each experiment as controls.

Substitution of isoleucine at this position also caused a serious folding defect and resulted in a majority of chains ending up in inclusion bodies. In contrast, a majority of chains carrying the L144V substitution folded correctly into native trimers, indicating that this substitution was tolerated.

Substitution of Phe 381, located in the middle of the stack, with alanine prevented the mutant polypeptide chains from reaching the native state. Substitution of Phe 381 by valine increased partitioning onto the inclusion body pathway, but still allowed about half of the chains to fold correctly (Fig. 7).

The Leu 540 is the terminal residue in this stack. Replacement of Leu 540 with isoleucine, valine, and alanine had little or no effect on the partitioning between folding and inclusion body pathways (Fig. 7).

In summary, targeted residues located closer to the center of the parallel β -helix (rungs 4, 5, 6, 8) were more sensitive to substitution with alanine than residues closer to the ends of the parallel β -helix (rungs 1, 2, 14). Of the five phenylalanine residues targeted, only Phe 181 in rung 2 tolerated substitution with alanine.

Substitutions within the triple-stranded β -helix stack and the triple β -prism

The tailspike folding pathway contains a protrimer folding intermediate ([pT] in Fig. 2), in which the three chains are associated but not fully folded. Recent results suggest that the protrimer represents the intermediate in which the interdigitated region has not yet wrapped (Benton et al. 2002; Kreisberg et al. 2002). In the native trimer, side-chains from the three subunits form a single buried core in the triple-stranded β -helix domain (Fig. 1C, top; Steinbacher et al. 1994; Seckler 1998; Kreisberg et al. 2000). The side-chains of Ile 548 (in the triple-stranded β -helix) and Ile 560 (in the triple β -prism) from the three subunits occupy the first and second positions in the six residue side-chain stacks that represent the subunit interfaces in this region of the protein (Fig. 1A,C, purple stars).

Neither tsf nor missense mutants had previously been recovered in this region of the polypeptide chain (Fane and King 1987; Villafane and King 1988; C. Haase-Pettingell and J. King, unpubl.), suggesting that the intertwined chain sections might be hypersensitive to amino acid replacement.

We were interested whether mutations in this threefold hydrophobic core and subunit interface might disrupt later steps in subunit folding or assembly. The buried isoleucine residues are oriented similarly with their hydrophobic side-chains forming the core of the trimer at this location. Ile 548 and Ile 560 were each replaced with valine and alanine.

Residue Glu 543 is in the transition region between parallel and triple-stranded β -helix structures, and points outward toward solvent. As a comparison site for the buried core residues in this region, we substituted Glu 543 (Fig. 1A,C, yellow stars) with isoleucine.

The effects of amino acid substitutions at residues Glu 543, Ile 548, and Ile 560 on the folding and accumulation of tailspike chains *in vivo* are shown in Figure 8. Pellet and supernatant fractions of lysates from cells harvested 2 h after induction at 30°C were analyzed by SDS-PAGE without heating (Fig. 8A). The first two lanes show soluble and insoluble proteins from a culture transformed with vector only and no insert. In cultures expressing the wild-type tailspike gene, tailspike chains accumulated primarily as the soluble SDS-resistant native trimer (supernatant fraction), with only a small fraction of chains accumulated as inclusion bodies. The quantitative analysis of tailspike chain con-

centration in Figure 8B demonstrated that >80% of wild-type chains reached the SDS-resistant conformation. The mutation E543I at the transition between parallel and triple-stranded β -helix domains showed a similar profile on the SDS gel as wild-type tailspike. This result indicates that replacement of the charged surface residue did not affect tailspike folding or assembly.

Replacement of Ile 548 with alanine blocked the formation of native tailspike trimers. Polypeptide chains with this substitution accumulated as SDS-sensitive non-native species, which were soluble and not in the pelleted inclusion body state. The I548V chains behaved similarly to wild

Figure 8. *In vivo* partitioning of tailspike chains containing replacements of the buried side-chain residues of the triple-stranded β -helix and triple β -prism domains. (A) Cultures expressing vector, wild-type, E543I, I548V, I548A, I560V, I560A, and F284I chains were fractionated into low-speed pellet and supernatants. Samples were electrophoresed through SDS-polyacrylamide gels and stained with Coomassie blue. (B) Quantification of the tailspike bands in the *above* gel by laser densitometry. (C) Examination of the conformational properties of tailspike species recovered from supernatant fractions, as analyzed by native gel electrophoresis in the cold. The slightly retarded mobility of E543I is expected for a substitution that changes the net surface charge on the tailspike (Yu and King 1988). The slower mobilities of the I548A chains are associated with the protrimer intermediate (Benton et al. 2002) or multimeric aggregation intermediates (Speed et al. 1996).

type. Chains containing F284I in the parallel β -helix domain are included on the same gel in Figure 8A (far right lane) as an example of a replacement that increases partitioning onto the inclusion body pathway. Polypeptide chains containing this mutation were detected only in the insoluble pellet fraction as SDS-polypeptide chain complexes.

The I560A chains were also unable to form native tailspike, and accumulated as soluble species similar to I548A. The I560V substitution was similar to wild type, showing no apparent difference in tailspike accumulation. Both the I548A and I560A supernatant species may have been released from tailspike trimers as a result of the mutations destabilizing the native structure. Or they may represent chains that failed to complete folding to the SDS-resistant trimeric conformation. The accumulation of tailspike mutant chains in a soluble but SDS-sensitive conformation is consistent with a defect in protrimer assembly or maturation.

Both mutations I548A and I560A thus appeared to completely block the formation of SDS-resistant native tailspikes. Because the level of insoluble aggregated chains did not increase, the mutations do not appear to affect partitioning between folding and inclusion body pathways.

Adjacent isoleucine residues in the triple-stranded β -helix and triple β -prism internal stacks control different folding steps

The conformations of E543I, I560A, and I548A tailspike chains in the soluble fractions of cell lysates were analyzed by native polyacrylamide gel electrophoresis. The supernatant fractions shown in Figure 8A were analyzed under non-denaturing conditions in the gel shown in Figure 8C. Wild-type native tailspikes migrated as a single sharp band in the non-denaturing gels. Tailspikes containing the E543I mutation showed a retarded migration as compared to the native wild-type tailspikes. This retarded mobility presumably represents the effect of surface charge change on the electrophoretic behavior of the native-like trimer (Yu and King 1984). The last lane contains the supernatant fraction from a culture expressing the parallel β -helix stack mutant F284I, in which tailspike chains are aggregated in the pellet.

Tailspikes containing the I548A mutation showed a retarded electrophoretic mobility. This striking retardation in electrophoretic mobility relative to SDS-resistant native tailspikes is characteristic of the protrimer folding intermediate (Goldenberg et al. 1982; Benton et al. 2002; Kreisberg et al. 2002). This result, combined with the demonstration in the SDS gel that I548A polypeptide chains are soluble but do not form SDS-resistant tailspikes, suggests that this mutation blocks the conversion of the protrimer to the native tailspike conformation. This phenotype is similar to that

reported for the nearby G546D and G546N substitutions (Kreisberg et al. 2002).

The I560A mutation had the most drastic conformational effects on soluble tailspike chains. The predominant species in the supernatant fraction migrated near the top of the native gel (Fig. 8C). This non-native tailspike species is most likely an oligomer. Additional staining intensity was observed in the region of the gel where single chain tailspike folding intermediates migrate. The accumulation of tailspike chains in two predominant conformations (unassembled and oligomeric) suggests that the I560A substitution disrupts normal subunit association and results in an incorrect but nevertheless specific oligomerization reaction. This phenotype is similar but more pronounced than that previously reported for β -prism mutations R563Q and A575T (Kreisberg et al. 2002).

The native-like trimers formed by I548V and I560V chains have similar mobilities as wild-type trimers (Fig. 8A,C). These results demonstrate that residues Ile 548 and Ile 560 are not critical for nucleation of subunit folding but are essential for subunit assembly.

Discussion

Role of buried hydrophobic residues in the elongated core of parallel β -helix domains

β strands are characterized by side-chains of successive residues alternately pointing inward to a hydrophobic buried core, or outward to solvent or other secondary structures. The buried core residues of parallel β -helix, triple-stranded β -helix, and triple β -prism domains form long, uninterrupted stacks of bulky hydrophobic side-chains (Jurnak et al. 1994; Steinbacher et al. 1994). In the tailspike, substitutions at many of the solvent-exposed strand and turn residues result in tsf defects (Haase-Pettingell and King 1997). Because these mutants fold into thermostable biologically active proteins at low temperatures, these surface residues cannot be absolutely essential for directing the chains into the parallel β -helix fold.

In genetic screens, tsf mutations have been recovered at very few buried side-chain sites in the parallel β -helix or triple-stranded β -helix, or triple β -prism domains of the tailspike adhesin. In globular proteins some buried hydrophobic residues are essential for reaching the native state (Lim and Sauer 1991; Gronenborn and Clore 1994; Sauer 1996). These observations suggested that the β -helix buried hydrophobic residues might be essential for correct folding and assembly of the native SDS-resistant tailspike trimer. However, because the buried cores of these domains are elongated and essentially cylindrical, it seemed possible that the folding of this domain would be tolerant to single amino acid substitutions in the buried core.

Replacement of aromatic branched side-chains with alanine was not tolerated at four of the five targeted phenylalanine sites. These four sites included the three-residue phenylalanine stack consisting of Phe 284, Phe 308, and Phe 336 (stack C, rungs 4, 5, and 6), as well as Phe 381, located in the central region of the parallel β -helix domain (stack A, rung 8). In contrast, Phe 181 was the only phenylalanine targeted here that tolerated the alanine substitution. Perhaps its peripheral location near the end of stack C is less critical for folding or trimer stability compared to the other four, more centrally located, phenylalanine residues.

The aromatic side-chains of residues Phe 284, Phe 308, and Phe 336 are stacked almost in parallel, like the bases in nucleic acids. This orientation is in fact not common in the buried cores of globular proteins (Burley and Petsko 1985, 1988; Hunter et al. 1991; Serrano et al. 1991). The substitution of these three phenylalanines by the smaller residues alanine and threonine resulted in the complete failure of the substituted chains to reach the native trimeric state. This result strongly suggests that this phenylalanine stack is essential for the correct folding of the β -helical domain of the native tailspike.

Bulky hydrophobic side-chains appear to be essential for productive tailspike folding *in vivo* at positions of Phe 284, Phe 308, Phe 336, and Phe 381 in the parallel β -helix domain. In contrast, the three targeted residues occupying rungs closer to the ends of the parallel β -helix (Leu 144, Phe 181, Leu 540) were tolerant to substitutions with alanine.

A collapsed molten globule state is thought to be an early intermediate in the *in vitro* folding of the buried core of globular β -sheet proteins. Thus, three phenylalanine side-chains form an "aromatic cluster" in the hydrophobic core of the villin headpiece subdomain. Frank and colleagues (2002) showed that single replacements of the phenylalanine residues with leucine, although destabilizing, were still tolerated by the native fold. This toleration for substitution within the hydrophobic core has been reported for many globular proteins (Sauer 1996). In such proteins the role of individual hydrophobic residues may be dispensable, as other residues can drive the collapse.

The extended buried core of a parallel β -helix seems unlikely to be formed by a concerted collapse mechanism. We suspect that the regularity in the fold may reflect a kinetic processive folding mechanism, with formed rungs nucleating folding of the next rung in the polypeptide chain. The BetaWrap algorithm, which efficiently predicts the parallel β -helix fold from primary sequence, follows this model (Bradley et al. 2001). In such a mechanism key substitutions in any rung would prevent further productive folding of the overall helix.

If β -helix folding is processive or zipper-like, it could initiate from either end, or from the middle. One can imagine a hydrophobic collapse of the central portion of the tailspike parallel β -helix (including rungs 3–8) happening

first, followed rapidly by β -sheet formation propagating in both amino-terminal and carboxy-terminal directions. This folding mechanism would also explain the result reported here that hydrophobic residues further from the central region, including Phe 181, appear more tolerant to amino acid substitutions. The central region of the parallel β -helix domain may be critical for nucleation or folding of the tailspike β -helix.

Comparison of in vivo and in vitro early folding stages

In addition to the phenylalanine stack, a number of sites of tsf mutants (Fig. 3, red residues) and the two global suppressor sites Val 331 and Ala 334 (Fig. 3, blue residues) are located in rungs 3–8 of sheet C. The tsf substitutions affect an early single-chain folding intermediate, resulting in accumulation of the chains in the misfolded and aggregated inclusion body state. A number of tsf mutants are rescued by the second-site global suppressor mutations, suV331A and suA334V. The suppressors stabilize the thermolabile intermediates or inhibit the steps leading off-pathway (Mitraki et al. 1991a,b,c; Baxa et al. 1999).

Although the global suppressors effectively correct tsf folding defects, they did not efficiently suppress the folding defect in F308A. Chains carrying suV331A/F308A or suA334V/F308A accumulated in the non-native inclusion body state when expressed *in vivo* (data not shown). One interpretation of the failure of global suppressors to efficiently rescue stack mutants is that the defect in folding of the stack mutants occurs earlier in the folding pathway than the tsf defects. This implicates the partially folded single chain species [I_o] (Fig. 2) as the conformation in which side-chain stacking, and thus parallel β -helix formation, initiates.

Ala 334 is in the buried core and the suppressing substitution is valine, a bulky hydrophobic side-chain. Baxa and colleagues proposed that the increased buried volume improves hydrophobic stacking in this region of the core stabilizing a critical parallel β -helical folding intermediates (Baxa et al. 1996; Schuler and Seckler 1998).

The existence of such an early intermediate species during *in vitro* refolding is indicated by the rapid increase in CD signal and fluorescence (Fuchs et al. 1991; Danner and Seckler 1993). *In vitro* [I_o] transforms to the partially folded thermolabile intermediate [I], characterized by a predominantly β -sheet spectroscopic signature. This single chain species can also be detected by native gel electrophoresis (Betts and King 1998).

Within cells, the early folding events occur in the nascent polypeptide chain still attached to the ribosome (Clark and King 2001). This species must of course differ from the free solution species [I_o] formed after dilution from urea. However, characterization of the conformation of ribosome-bound nascent tailspike chains using monoclonal antibodies reveals that they differ from the solution species [I]. Thus,

there is likely to be an earlier precursor to [I] in vivo as well as in vitro.

Formation of the three-stranded regions

Once the parallel β -helices have formed in the unassembled partially folded subunits, the intermediate [I] presumably has enough structure to form partially folded dimers and then the protrimer (Goldenberg et al. 1983; Benton et al. 2002). Gage and Robinson (2003) have described mutants at the carboxy-terminal positions L606 and L663, which prevent protrimer formation, suggesting that this region is responsible for initiating trimerization. The protrimer has an increased radius of gyration and presumably represents the trimer intermediate before the intertwining of the three strands. The resulting triple-stranded β -helix appears to function as a molecular clamp, sharply increasing the thermal stability and protease resistance of the folded trimer (Kreisberg et al. 2002).

Residues Ile 548 and Ile 560 from each subunit are stacked at the subunit interfaces and contribute to the tightly packed hydrophobic core of the triple-stranded β -helix and triple β -prism domains. Replacement of Ile 548 with alanine blocked folding to the native state. I548A chains did not accumulate as inclusion bodies but remained soluble. Analysis of lysate samples by native gel electrophoresis revealed that mutant chains accumulated as protrimer-like species. The mutation G546D/N located in the same region also forms a protrimer-like species (Kreisberg et al. 2002). These results suggest that a bulky hydrophobic side-chain at residue 548 is absolutely required for formation of the triple-stranded β -helix, and that polar side-chains replacing Gly 546 are not allowed.

Ile 560 is in the triple β -prism core stacked below Ile 548 (Fig. 1A,C). Substitution of Ile 560 with alanine resulted in chains that could not form the native trimer but did accumulate as a soluble species. This soluble species appears to be a higher order oligomer. Additional staining in the monomeric region most likely represents unassembled chains. It is not clear whether the oligomeric species forms directly from a partially folded monomer, or from a trapped protrimer. Although the Ile 560 residue is oriented toward the interior of the triple β -prism, it has a similar phenotype to two external mutants R563Q and A575T, described by Kreisberg et al (2002). Mutations R563Q and A575T cause a similar defect resulting in the accumulation of soluble but misfolded tailspike chains.

Some aspect of the processes of alignment, oligomerization, and the transition from protrimer to stable trimer, involves transient disulfide bonds (Robinson and King 1997; Haase-Pettingell et al. 2001). At present we do not know whether the interchain disulfide bonds found in the protrimer are reduced before the wrapping step, or coupled to the wrapping step. Because disulfide bonds have never been

observed in the native trimer, it is unlikely that their reduction would occur as the last stage of maturation. The mutations I548A and G546D/N appear to trap a protrimer-like intermediate. It is possible that these mutations do not block subunit intertwining, as argued here, but instead prevent the reduction of interchain disulfides that may be a prerequisite to tailspike maturation.

The tailspike contains three regions with elongated hydrophobic cores: the single parallel β -helix domain, the triple-stranded β -helix, and the triple β -prism. The mutants studied in this paper identify residues critical to the correct folding, association, and assembly of each of these three domains.

Materials and methods

Materials and DNA sequencing

Gene 9 of the *Salmonella* phage P22 was subcloned into the vector pET11a (Novagen) as described previously (Kreisberg et al. 2002; A. S. Robinson and J. A. King, unpubl.). The resulting vector was designated pET (*gene9*) and encoded the full-length tailspike polypeptide. Plasmid DNA was purified for mutagenesis and sequencing using Qiagen kits. DNA sequences were determined by automated sequencing at the MIT Biopolymers Laboratory.

Gene 9 cloning and mutagenesis

Degenerate primers were designed to introduce one or two missense mutations at the targeted codon. Where possible silent restriction sites were included to distinguish specific missense mutations by restriction mapping, before confirmation by DNA sequence analysis. All mutations were introduced using the pET(*gene9*) vector (Kreisberg et al. 2002) as mutagenesis template and the QuikChange mutagenesis kit (Stratagene). Mutant clones were identified by restriction mapping and confirmed by DNA sequencing.

The hydrophobic stacks were used as reference points to generate the structure-based sequence alignment in Figure 3. Rows 1–13 in the alignment correspond to a complete turn or rung of the parallel β -helix. The three β -sheets that form the sides of the parallel β -helix are shown in cyan (sheet A), yellow (sheet B), and green (sheet C).

Tailspike expression in E. coli

Recombinant tailspike polypeptides were overexpressed in *E. coli* BL21(DE3) using the pET vector system (Novagen). Cultures in LB broth supplemented with 50 μ g/mL ampicillin (LB-Amp) were inoculated either with colonies or frozen cells and grown overnight at 37°C with shaking. A portion of the overnight culture was diluted 50-fold into fresh LB-Amp medium and incubated at 28° or 30°C with shaking (250 rpm). Recombinant protein expression was induced by addition of IPTG to 400 μ M when cultures reached an optical density at 600 nm of about 0.6.

After 2 h of further shaking at 30°C the cells were harvested by centrifugation (7 min, 8000g, 4°C) and placed on ice. Cells were resuspended in ice-cold lysis buffer (50 mM Tris-HCl, at pH 8.0,

100 mM NaCl, 2 mM EDTA) and stored at -20°C . Frozen suspensions were thawed at room temperature and immediately supplemented with PMSF (1 mM final concentration from 100 mM stock in ethanol) and DTT (2 mM final concentration from 1 M aqueous solution). Lysis was initiated by addition of lysozyme (100 $\mu\text{g}/\text{mL}$ final concentration from 2 mg/mL aqueous solution) and incubated at room temperature for 30 min with occasional mixing. Lysed cells were incubated for an additional 20–30 min at room temperature after addition of DNase I (Sigma) (25 $\mu\text{g}/\text{mL}$ final concentration from 0.5 mg/mL stock in 0.5 M MgCl_2).

Soluble and insoluble fractions of the cell lysates were separated by centrifugation at 4°C for 5 min at 17,000g for volumes less than 1.5 mL and 12,000g for higher volumes. Lysate pellets were resuspended in lysis buffer by gentle agitation (small volumes) or with a Dounce homogenizer (large volumes). Lysate fractions were kept on ice and either analyzed immediately or stored at -20°C .

Inclusion body refolding and analysis

The concentration of tailspike polypeptide chains in IB preparations was determined by SDS-PAGE and laser densitometry. Known quantities of native tailspikes were boiled in SDS sample buffer to yield an identical quantity of SDS-polypeptide chain complexes. These samples were used as standards to estimate the concentration of tailspike polypeptide chains present in IB preparations. Suspensions of detergent-washed IB preparations containing 20 μg tailspike chains were pelleted and the buffer was aspirated. Pellets were resuspended in 50 μL of a solution containing 6 M urea, 50 mM Tris-HCl (pH 7.6), 25 mM NaCl, 2 mM EDTA, and 10 mM DTT. After incubation for 60 min at 20°C , protein refolding was initiated by dilution with 950 μL of ice-cold dilution buffer (50 mM Tris, 25 mM NaCl, 2 mM EDTA, at pH 7.6, supplemented with 10 mM DTT). The concentration of tailspike polypeptide chains was 20 $\mu\text{g}/\text{mL}$ during refolding. After 60 min of incubation on ice, a portion of the refolding reaction (200 μL) was transferred to a chilled tube on ice and the reaction tube containing the remaining 800 μL was transferred to 20°C . After 10 min of incubation at 20°C , 200 μL of each refolding reaction were mixed with 100 μL ice-cold native sample buffer and electrophoresed through a polyacrylamide gel in the cold.

Gel electrophoresis and laser densitometry

The protocol for native gel electrophoresis was as described in Betts et al. (1999). Lysate fractions were prepared for electrophoresis by dilution in one-half volume ice-cold native gel sample buffer (14 mM Tris base, 109 mM glycine, 30% glycerol, 50 mM DTT, and 0.1% bromphenol blue). Samples were kept on ice and electrophoresis was performed in a 4°C cold room using a Tris-glycine running buffer. Samples for SDS-PAGE were mixed with one-half volume SDS sample buffer (188 mM Tris at pH 6.8, 2% SDS, 15% β -mercaptoethanol, and bromphenol blue) and kept at room temperature until electrophoresis. The discontinuous buffer system of King and Laemmli (1971) was used. Coomassie-stained gels were analyzed using a Personal Densitometer SI (Molecular Dynamics, Amersham Bioscience) and ImageQuant software. Concentrations of tailspike chains were determined from Coomassie-stained SDS gels using tailspike standards as described previously (Betts and King 1998).

Acknowledgments

This research was supported by the NIH (GM 17,980 to J.K.) and by the NSF's Engineering Research Center Initiative (8803014).

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

- Baxa, U., Steinbacher, S., Miller, S., Weintraub, A., Huber, R., and Seckler, R. 1996. Interactions of phage P22 tails with their cellular receptor, Salmonella O-antigen polysaccharide. *Biophys. J.* **71**: 2040–2048.
- Baxa, U., Steinbacher, S., Weintraub, A., Huber, R., and Seckler, R. 1999. Mutations improving the folding of phage P22 tailspike protein affect its receptor binding activity. *J. Mol. Biol.* **293**: 693–701.
- Beissinger, M., Lee, S.C., Steinbacher, S., Reinemer, P., Huber, R., Yu, M.H., and Seckler, R. 1995. Mutations that stabilize folding intermediates of phage P22 tailspike protein: Folding in vivo and in vitro, stability, and structural context. *J. Mol. Biol.* **249**: 185–194.
- Benton, C.B., King, J., and Clark, P.L. 2002. Characterization of the protrimmer intermediate in the folding pathway of the interdigitated β -helix tailspike protein. *Biochemistry* **41**: 5093–5103.
- Betts, S.D. and King, J. 1998. Cold rescue of the thermolabile tailspike intermediate at the junction between productive folding and off-pathway aggregation. *Protein Sci.* **7**: 1516–1523.
- . 1999. There's a right way and a wrong way: In vivo and in vitro folding, misfolding and subunit assembly of the P22 tailspike. *Structure Fold Des.* **7**: R131–R139.
- Betts, S., Speed, M., and King, J. 1999. Detection of early aggregation intermediates by native gel electrophoresis and native western blotting. *Methods Enzymol.* **309**: 333–350.
- Bradley, P., Cowen, L., Menke, M., King, J., and Berger, B. 2001. BETAWRAP: Successful prediction of parallel β -helices from primary sequence reveals an association with many microbial pathogens. *Proc. Natl. Acad. Sci.* **98**: 14819–14824.
- Burley, S.K. and Petsko, G.A. 1985. Aromatic-aromatic interaction: A mechanism of protein structure stabilization. *Science* **229**: 23–28.
- . 1988. Weakly polar interactions in proteins. *Adv. Prot. Chem.* **39**: 125–189.
- Clark, P.L. and King, J. 2001. A newly synthesized, ribosome-bound polypeptide chain adopts conformations dissimilar from early in vitro refolding intermediates. *J. Biol. Chem.* **276**: 25411–25420.
- Danner, M. and Seckler, R. 1993. Mechanism of phage P22 tailspike protein folding mutations. *Protein Sci.* **2**: 1869–1881.
- Fane, B. and King, J. 1987. Identification of sites influencing the folding and subunit of the p22 tailspike polypeptide chain using nonsense. *Genetics* **117**: 157–172.
- Frank, B.S., Vardar, D., Buckley, D.A., and McKnight, C.J. 2002. The role of aromatic residues in the hydrophobic core of the villin headpiece subdomain. *Protein Sci.* **11**: 680–687.
- Friguet, B., Djavadi-Ohanian, L., and Goldberg, M.E. 1984. Some monoclonal antibodies raised with a native protein bind preferentially to the denatured antigen. *Mol. Immunol.* **21**: 673–677.
- Fuchs, A., Seiderer, C., and Seckler, R. 1991. In vitro folding pathway of phage P22 tailspike protein. *Biochemistry* **30**: 6598–6604.
- Gage, M.J. and Robinson, A.S. 2003. C-terminal hydrophobic interactions play a critical role in oligomeric assembly of the P22 tailspike trimer. *Protein Sci.* **12**: 2732–2747.
- Goldenberg, D.P. and King, J. 1981. Temperature-sensitive mutants blocked in the folding or subunit assembly of the bacteriophage P22 tail spike protein. II. Active mutant proteins matured at 30 degrees C. *J. Mol. Biol.* **145**: 633–651.
- . 1982. Trimeric intermediate in the in vivo folding and subunit assembly of the tail spike endorhamnosidase of bacteriophage P22. *Proc. Natl. Acad. Sci.* **79**: 3403–3407.
- Goldenberg, D.P., Berget, P.B., and King, J. 1982. Maturation of the tail spike endorhamnosidase of *Salmonella* phage P22. *J. Biol. Chem.* **257**: 7864–7871.
- Goldenberg, D.P., Smith, D.H., and King, J. 1983. Genetic analysis of the folding pathway for the tail spike protein of phage P22. *Proc. Natl. Acad. Sci.* **80**: 7060–7064.
- Gronenborn, A.M. and Clore, G.M. 1994. Experimental support for the "hydrophobic zipper" hypothesis [letter]. *Science* **263**: 536.
- Gunasekaran, K., Eyles, S.J., Hagler, A.T., and Gierasch, L.M. 2001. Keeping it in the family: Folding studies of related proteins. *Curr. Opin. Struct. Biol.* **11**: 83–93.

- Haase-Pettingell, C. and King, J. 1988. Formation of aggregates from a thermolabile in vivo folding intermediate in P22 tailspike maturation. A model for inclusion body formation. *J. Biol. Chem.* **263**: 4977–4983.
- . 1997. Prevalence of temperature sensitive folding mutations in the parallel β coil domain of the phage P22 tailspike endorhamnosidase. *J. Mol. Biol.* **267**: 88–102.
- Haase-Pettingell, C., Betts, S., Raso, S.W., Stuart, L., Robinson, A., and King, J. 2001. Role for cysteine residues in the in vivo folding and assembly of the phage P22 tailspike. *Protein Sci.* **10**: 397–410.
- Hunter, C.A., Singh, J., and Thornton, J.M. 1991. π - π interactions: The geometry and energetics of phenylalanine-phenylalanine interactions in proteins. *J. Mol. Biol.* **218**: 837–846.
- Jurnak, F., Yoder, M.D., Pickersgill, R., and Jenkins, J. 1994. Parallel β -domains: A new fold in protein structures. *Curr. Opin. Struct. Biol.* **4**: 802–806.
- King, J. and Laemmli, U.K. 1971. Polypeptides of the tail fibers of bacteriophage T4. *J. Mol. Biol.* **62**: 465–477.
- King, J., Haase-Pettingell, C., Robinson, A.S., Speed, M., and Mitraki, A. 1996. Thermolabile folding intermediates: Inclusion body precursors and chaperonin substrates. *FASEB J.* **10**: 57–66.
- Kreisberg, J.F., Betts, S.D., and King, J. 2000. β -helix core packing within the triple-stranded oligomerization domain of the P22 tailspike. *Protein Sci.* **9**: 2338–2343.
- Kreisberg, J.F., Betts, S.D., Haase-Pettingell, C., and King, J. 2002. The interdigitated β -helix domain of the P22 tailspike protein acts as a molecular clamp in trimer stabilization. *Protein Sci.* **11**: 820–830.
- Lim, W.A. and Sauer, R.T. 1991. The role of internal packing interactions in determining the structure and stability of a protein. *J. Mol. Biol.* **219**: 359–376.
- Miller, S., Schuler, B., and Seckler, R. 1998. A reversibly unfolding fragment of P22 tailspike protein with native structure: The isolated β -helix domain. *Biochemistry* **37**: 9160–9168.
- Mitraki, A., Fane, B., Haase-Pettingell, C., and King, J. 1991a. Mutations affecting protein folding and misfolding in vivo. In *Application in enzyme biotechnology* (eds. T. Baldwin and J. Kelly), pp. 129–136. Plenum Press, New York.
- Mitraki, A., Fane, B., Haase-Pettingell, C., Sturtevant, J., and King, J. 1991b. Global suppression of protein folding defects and inclusion body formation. *Science* **253**: 54–58.
- Mitraki, A., Haase-Pettingell, C., and King, J. 1991c. Mechanisms of inclusion body formation. In *Protein refolding* (eds. G. Georgiou, and E. De Bernardez-Clark), pp. 35–49. American Chemical Society, Washington, DC.
- Mitraki, A., Miller, S., and van Raaij, M.J. 2002. Review: Conformation and folding of novel β -structural elements in viral fiber proteins: The triple β -spiral and triple β -helix. *J. Struct. Biol.* **137**: 236–247.
- Pickersgill, R., Harris, G., Lo Leggio, L., Mayans, O., and Jenkins, J. 1998. Superfamilies: The 4/7 superfamily of β α -barrel glycosidases and the right-handed parallel β -helix superfamily. *Biochem. Soc. Trans.* **26**: 190–198.
- Robinson, A.S. and King, J.A. 1997. Disulfide-bonded intermediate on the folding and assembly pathway of a non-disulfide bonded protein. *Nat. Struct. Biol.* **4**: 450–455.
- Sauer, R.T. 1996. Protein folding from a combinatorial perspective. *Fold. Des.* **1**: R27–30.
- Schuler, B. and Seckler, R. 1998. P22 tailspike folding mutants revisited: Effects on the thermodynamic stability of the isolated β -helix domain. *J. Mol. Biol.* **281**: 227–234.
- Schwarz, J.J. and Berget, P.B. 1989. Characterization of bacteriophage P22 tailspike mutant proteins with altered endorhamnosidase and capsid assembly activities. *J. Biol. Chem.* **264**: 20112–20119.
- Seckler, R. 1998. Folding and function of repetitive structure in the homotrimeric phage P22 tailspike protein. *J. Struct. Biol.* **122**: 216–222.
- Serrano, L., Bycroft, M., and Fersht, A.R. 1991. Aromatic-aromatic interactions and protein stability. Investigation by double-mutant cycles. *J. Mol. Biol.* **218**: 465–475.
- Speed, M.A., Wang, D.I., and King, J. 1995. Multimeric intermediates in the pathway to the aggregated inclusion body state for P22 tailspike polypeptide chains. *Protein Sci.* **4**: 900–908.
- . 1996. Specific aggregation of partially folded polypeptide chains: The molecular basis of inclusion body composition. *Nat. Biotech.* **14**: 1283–1287.
- Steinbacher, S., Seckler, R., Miller, S., Steipe, B., Huber, R., and Reinemer, P. 1994. Crystal structure of P22 tailspike protein: Interdigitated subunits in a thermostable trimer. *Science* **265**: 383–386.
- Steinbacher, S., Baxa, U., Miller, S., Weintraub, A., Seckler, R., and Huber, R. 1996. Crystal structure of phage P22 tailspike protein complexed with *Salmonella* sp. O-antigen receptors. *Proc. Natl. Acad. Sci.* **93**: 10584–10588.
- Sturtevant, J.M., Yu, M.-H., Haase-Pettingell, C., and King, J. 1989. Thermostability of temperature-sensitive folding mutants of the P22 tailspike protein. *J. Biol. Chem.* **264**: 10693–10698.
- Villafane, R. and King, J. 1988. Nature and distribution of sites of temperature-sensitive folding mutations in the gene for the P22 tailspike polypeptide chain. *J. Mol. Biol.* **204**: 607–619.
- Weigele, P.R., Scanlon, E., and King, J. 2003. Homotrimeric, β -stranded viral adhesins and tail proteins. *J. Bacteriol.* **185**: 4022–4030.
- Wetzel, R. 2002. Ideas of order for amyloid fibril structure. *Structure (Camb)* **10**: 1031–1036.
- Williams, A.D., Portelius, E., Kheterpal, I., Guo, J.T., Cook, K.D., Xu, Y., and Wetzel, R. 2004. Mapping a β amyloid fibril secondary structure using scanning proline mutagenesis. *J. Mol. Biol.* **335**: 833–842.
- Yoder, M.D., Keen, N.T., and Jurnak, R. 1993. New domain motif: The structure of pectate lyase C, a secreted plant virulence factor. *Science* **260**: 1503–1507.
- Yu, M.-H. and King, J. 1984. Single amino-acid substitutions influencing the folding. *Proc. Natl. Acad. Sci.* **81**: 6584–6588.
- . 1988. Surface amino acids as sites of temperature-sensitive folding mutations in the P22 tailspike protein. *J. Biol. Chem.* **263**: 1424–1431.