

# Cotranslational Folding Promotes $\beta$ -Helix Formation and Avoids Aggregation *In Vivo*

Michael S. Evans, Ian M. Sander and Patricia L. Clark\*

Department of Chemistry and Biochemistry, University of Notre Dame, 251 Nieuwland Science Hall, Notre Dame, IN 46556, USA

Received 20 March 2008;  
received in revised form  
10 July 2008;  
accepted 12 July 2008  
Available online  
22 July 2008

Newly synthesized proteins must form their native structures in the crowded environment of the cell, while avoiding non-native conformations that can lead to aggregation. Yet, remarkably little is known about the progressive folding of polypeptide chains during chain synthesis by the ribosome or of the influence of this folding environment on productive folding *in vivo*. P22 tailspike is a homotrimeric protein that is prone to aggregation via misfolding of its central  $\beta$ -helix domain *in vitro*. We have produced stalled ribosome:tailspike nascent chain complexes of four fixed lengths *in vivo*, in order to assess cotranslational folding of newly synthesized tailspike chains as a function of chain length. Partially synthesized, ribosome-bound nascent tailspike chains populate stable conformations with some native-state structural features even prior to the appearance of the entire  $\beta$ -helix domain, regardless of the presence of the chaperone trigger factor, yet these conformations are distinct from the conformations of released, refolded tailspike truncations. These results suggest that organization of the aggregation-prone  $\beta$ -helix domain occurs cotranslationally, prior to chain release, to a conformation that is distinct from the accessible energy minimum conformation for the truncated free chain in solution.

© 2008 Elsevier Ltd. All rights reserved.

**Keywords:** cotranslational protein folding ribosome;  $\beta$ -helix;  $\beta$ -sheet; nascent chain

Edited by S. Radford

## Introduction

As a protein is synthesized by the ribosome, it begins to fold into a three-dimensional shape and must simultaneously avoid aggregation with other proteins in the crowded cell. In this seemingly hostile environment, where total protein concentrations can exceed 200–300 mg/mL, *de novo* protein folding during and after translation is, for many proteins, more efficient than *in vitro* refolding.<sup>1</sup> In other words, many proteins that can fold productively in the cell will aggregate severely under *in vitro* refolding conditions, presumably due to differences between the dominant folding pathway used. For example, experiments with both bacterial and firefly luciferase have demonstrated that these nascent chains adopt conformations cotranslationally that are not populated during refolding from denaturant,<sup>2,3</sup> and these

cotranslational conformations fold to the native state much more efficiently than the conformations populated during refolding from denaturant. However, there continues to be debate as to what extent large, multidomain proteins can fold cotranslationally in prokaryotes.<sup>4–7</sup>

While molecular chaperones certainly play a role in efficient protein folding in the cell, less than 20% of *Escherichia coli* cytoplasmic proteins require an interaction with one of the three major chaperone systems [trigger factor (TF), DnaK/DnaJ, or GroEL/ES] in order to fold correctly under normal growth conditions.<sup>8</sup> Remarkably, both chaperone systems responsible for the recognition of newly synthesized polypeptide chains in the *E. coli* cytoplasm (TF and DnaK/J) can be deleted simultaneously without abolishing cell viability.<sup>9–11</sup> Hence, for many proteins, cotranslational folding alone may be sufficient to partition newly synthesized polypeptide chains away from aggregation pathways and towards a productive folding pathway.<sup>12</sup>

How might protein folding be affected by the vectorial appearance of the polypeptide chain? Studies of purified, C-terminal truncations of single-

\*Corresponding author. E-mail address: pclark1@nd.edu.

Abbreviations used: mAb, monoclonal antibody; RNC, ribosome: nascent chain complex; TF, trigger factor.

domain proteins suggest that the formation of stable, folded, native-like structure requires the presence of an entire domain,<sup>13,14</sup> for many proteins, deleting only a few C-terminal residues is sufficient to globally destabilize domain structure. Yet, some native structural topologies will presumably be more affected by cotranslational folding than others. For example, the formation of  $\beta$ -sheet structure can require contacts between residues that are distant in the primary sequence and, hence, is inherently more complex than the formation of  $\alpha$ -helical structure. While recent reports have described the potential for  $\alpha$ -helices to form cotranslationally, even within the ribosomal exit tunnel,<sup>15,16</sup> it remains to be determined how the residues comprising a single  $\beta$ -strand or other incomplete portions of a  $\beta$ -sheet avoid nonnative intra- or intermolecular contacts while waiting for their native partner residues to appear outside the ribosome.<sup>17</sup> Similar issues arise when vectorial folding of multimeric proteins is considered.

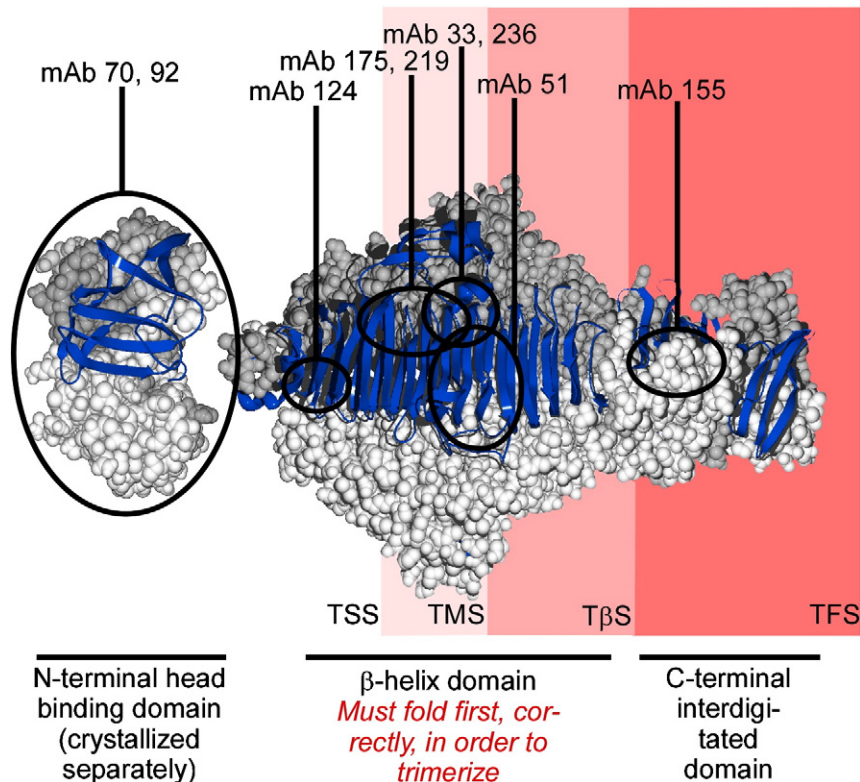
The homotrimeric tailspike protein from *Salmonella* phage P22 is composed almost entirely of  $\beta$ -sheet structure; its largest structural motif is a 13-rung parallel  $\beta$ -helix (Fig. 1). *In vitro*, tailspike folds first via the organization of the  $\beta$ -helix domain, followed by chain trimerization, and, finally, via additional folding steps that include the interdigitation of the C-terminal domain, which imparts high thermostability.<sup>20–22</sup> The proper formation of the  $\beta$ -helix domain must occur prior to trimerization, and

tailspike temperature sensitive for folding (*tsf*) and suppressor (*su*) mutations are found exclusively in the  $\beta$ -helix domain.<sup>23–25</sup> *In vivo*, tailspike is expressed primarily as a soluble protein at 30–37 °C, although some aggregation occurs even under physiological conditions.<sup>20</sup> *In vitro*, tailspike refolding is far less efficient, with significant aggregation occurring at 20 °C.<sup>26</sup> Neither GroEL nor DnaK has been shown to participate in productive tailspike folding *in vivo*;<sup>27,28</sup> the role of TF is undetermined. Previous qualitative studies from our laboratory<sup>29</sup> and others<sup>30,31</sup> have demonstrated that full-length tailspike nascent chains have significant native-like organization of the  $\beta$ -helix domain prior to release of the nascent chain from the ribosome. Trimerization, however, does not occur cotranslationally.<sup>32</sup> Here, we show that tailspike  $\beta$ -sheet formation occurs in a processive manner, during translation of successive rungs, even prior to the presence of the entire  $\beta$ -helix domain. In addition, we show that cotranslational folding traps the nascent chain in a conformation that is distinct from the conformations accessible to the free polypeptide chain.

## Results

### Selection of tailspike nascent chain lengths

We used the 17-residue “stall sequence” from the *E. coli* SecM protein to efficiently stall tailspike



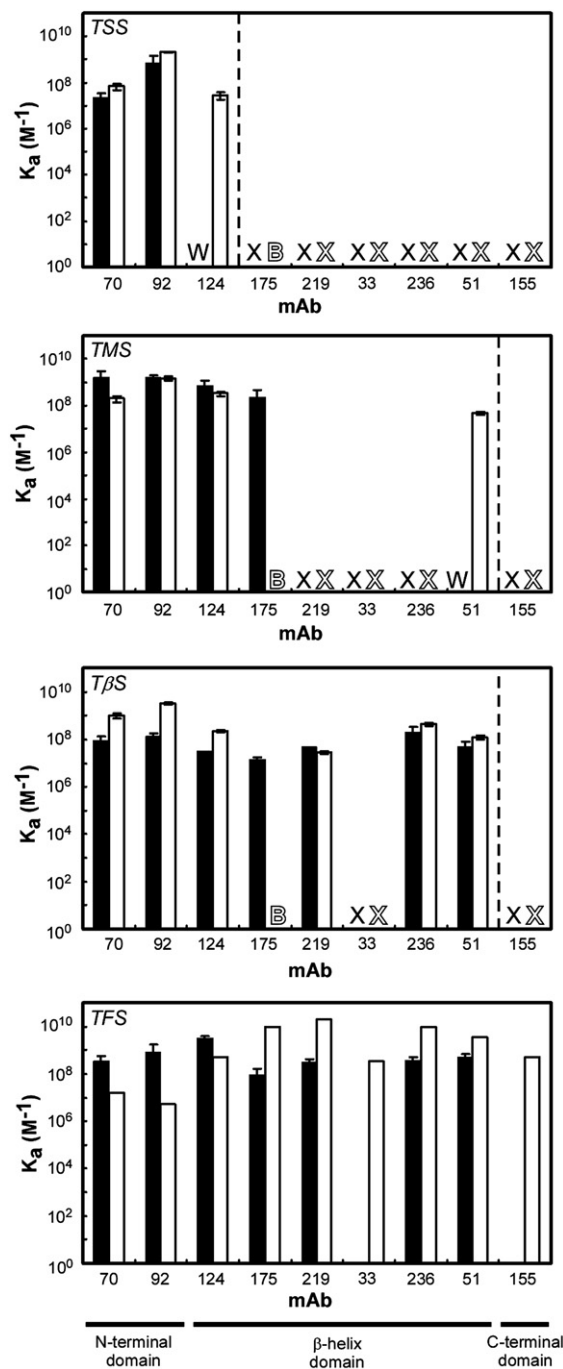
**Fig. 1.** Tailspike structure and truncation constructs. Tailspike trimer crystal structure,<sup>18</sup> one monomer chain is depicted as a blue ribbon, and the other two chains are depicted as light and dark gray space-filling models. mAb epitope boundaries<sup>19</sup> are indicated by black circles. The portion of the tailspike polypeptide chain exposed outside the ribosome for each tailspike nascent chain construct is indicated by the red shading behind the crystal structure.

translation *in vivo*<sup>29</sup> at selected locations. Yet, in any study of nascent chain conformations on stalled ribosomes, there is a risk that the equilibrium conformation of a stalled nascent chain might not accurately reflect the conformation of that nascent chain populated during unstalled translation. In order to minimize this possible discrepancy, we selected nascent chain lengths that correspond to the locations of statistically significant clusters of rare codons in the endogenous tailspike mRNA sequence, including detected translation pause sites (Clarke and Clark, submitted). Due to translational pausing, nascent chains of these lengths have more time to reach an equilibrium conformation, even during unstalled translation. The tailspike nascent chains described here, therefore, may more closely resemble biologically relevant translation intermediates than nascent chains of other lengths. Moreover, nascent chains of these lengths expose structurally interesting portions of the tailspike polypeptide chain outside the ribosome exit tunnel (Fig. 1): TSS (residues 1–229) exposes the N-terminal domain and the first three rungs of the  $\beta$ -helix domain; TMS (residues 1–406) exposes the N-terminal domain and the first half (eight rungs) of the  $\beta$ -helix; T $\beta$ S (residues 1–556) exposes the N-terminal domain and the entire  $\beta$ -helix; and TFS (residues 1–666) exposes the entire tailspike chain, with the exception of the extreme C-terminal residues that are masked by the ribosome exit tunnel.

### Many conformation-sensitive mAbs bind tailspike nascent chains as they emerge from the ribosome

A panel of nine anti-tailspike monoclonal antibodies (mAbs) developed in the Goldberg laboratory binds to a range of tailspike structural features, with the majority recognizing epitopes in the  $\beta$ -helix domain<sup>19,33</sup> (Fig. 1). Previous studies have shown that, in most cases, mAb binding is sensitive to the folding status of tailspike,<sup>31,33,34</sup> including a broad range of recognition profiles for tailspike *in vitro* refolding intermediates.<sup>19</sup> The anti-tailspike mAbs, therefore, provide a selective probe with which to assay tailspike conformational states in complicated mixtures such as cell lysates. Overall, antibody binding as a function of refolding time suggests a model for *in vitro* tailspike refolding (and particularly the organization of the  $\beta$ -helix domain) that begins with the folding of the extreme N- and C-termini of the  $\beta$ -helix.<sup>19</sup> The following work demonstrates how the cotranslational appearance of the  $\beta$ -helix affects this organizational pattern.

Previous qualitative results have shown that full-length tailspike nascent chains form some native-like structure while associated with the ribosome, based on the binding of three anti-tailspike mAbs.<sup>29,30</sup> We have extended this observation, quantitatively characterizing the binding of nine mAbs to four tailspike nascent chain lengths, using the competition ELISA method<sup>35</sup> (Fig. 2; see also Supplementary Fig. S1). Of course, for each nascent chain length, only the subset



**Fig. 2.** Binding of anti-tailspike mAbs to ribosome-bound tailspike nascent chains and corresponding released chains. mAb binding affinity to ribosome-bound tailspike nascent chains (solid bars) and mAb binding affinity to either the corresponding released tailspike chain (for TSS, TMS, and T $\beta$ S) or native tailspike (for TFS) (open bars). 'X' indicates that no measurable binding affinity was detected; 'W' indicates a weak affinity, too low to be quantified ( $K_a < 10^6 \text{ M}^{-1}$ ); 'B' signifies high background binding of mAb 175 to other components in the cell lysate supernatant. Error bars represent the standard deviation between experiments performed in triplicate. mAbs to the right of the vertical broken lines have epitopes not yet synthesized at these nascent chain lengths. Results for mAb 236 binding to TFS, T $\beta$ S, and TSS nascent chains were previously reported.<sup>29</sup>

of the antibodies whose epitopes have been synthesized can be used to provide information on the tailspike nascent chain conformation.<sup>33</sup>

For the shortest tailspike nascent chain construct (TSS), two of the three mAbs with epitopes encoded within this nascent chain bound tightly. Indeed, mAbs 70 and 92 bind with a higher affinity to TSS than to native trimeric tailspike. In addition, mAb 124 binding can be detected qualitatively, but the affinity constant is too low to be measured accurately ( $<10^6 \text{ M}^{-1}$ ). For TMS, which truncates tailspike in the middle of the  $\beta$ -helix domain, binding was detected for four of the expected seven mAbs. The most N-terminal of the anti-native mAbs, 175, binds TMS nascent chains tightly ( $K_a = 2 \times 10^8 \text{ M}^{-1}$ ), though not as tightly as that to native trimer ( $K_a = 10^{10} \text{ M}^{-1}$ ). mAb 51 binding to TMS nascent chains can be detected qualitatively, but the affinity constant is too low to be measured. For T $\beta$ S, seven of the expected eight mAbs (i.e., all except mAb 33) bound to the nascent chain. Finally, for TFS, all mAbs except mAb 155, which recognizes the most C-terminal interdigitated portion of the tailspike polypeptide chain, and mAb 33, in the center of the  $\beta$ -helix, showed binding affinities comparable to those measured for native tailspike.<sup>19</sup> In all cases, mAbs that bound to TFS nascent chains bound very tightly ( $\geq 10^7 \text{ M}^{-1}$ ), and in some cases, they bound more tightly than to native trimeric tailspike. Interestingly, mAbs 175, 219, and 236, whose epitopes are all located in the  $\beta$ -helix and show low to no affinity for early *in vitro* tailspike refolding intermediates,<sup>19</sup> all bound tightly to ribosome-bound tailspike nascent chains of several lengths, including (for mAb 175) TMS, which contains an incomplete  $\beta$ -helix domain.

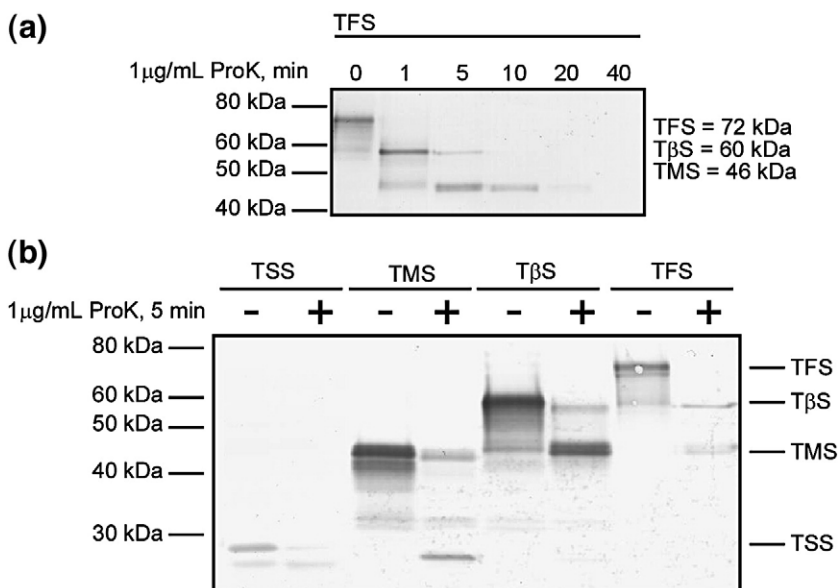
### Limited proteolysis reveals a stable fragment present in longer nascent chains

As an alternative means of measuring structure formation, we subjected stalled ribosome: nascent

chain complexes (RNCs) to limited proteolytic digestion with proteinase K, a nonspecific protease. We treated RNCs with proteinase K for varying times at 4 °C (Fig. 3). We performed control experiments to insure that protease digestion did not affect the structure of the ribosome itself (data not shown). TSS was rapidly digested with no detectable intermediates or protected fragments. TMS showed more resistance to protease digestion: after 5 min, detectable amounts of full-length TMS remained, although a significant amount was digested to a protease-resistant fragment of approximately 25 kDa. Both T $\beta$ S and TFS were rapidly digested to a relatively protease-resistant 47-kDa fragment, indistinguishable in size from the intact TMS nascent chain. Interestingly, at intermediate times, TFS was digested to a metastable 60-kDa fragment as well, corresponding to the size of T $\beta$ S (Fig. 3a). Identical results were obtained for Western blotting performed using the two anti-N-terminal antibodies or a mixture of all anti-tailspike mAbs (not shown).

### Released, refolded truncated chains populate conformations distinct from the conformations of ribosome-bound nascent chains

Tailspike C-terminal truncations bearing the SecM stall sequence are eventually released from the ribosome. The majority aggregate (not shown), but a small fraction remain soluble and can be separated from RNCs by sucrose gradient ultracentrifugation.<sup>29</sup> We measured the binding of the anti-tailspike mAbs to these released, truncated tailspike chains (Fig. 2). For released TSS, the two N-terminal, less conformationally sensitive mAbs, 70 and 92, still bind with high affinity. For TMS, the four mAbs that bind the nascent chain attached to the ribosome still bind to the released chain, with one exception: we observed high background binding in the empty vector control samples for mAb 175 binding to released chains, making accurate quantification impossible. In addi-



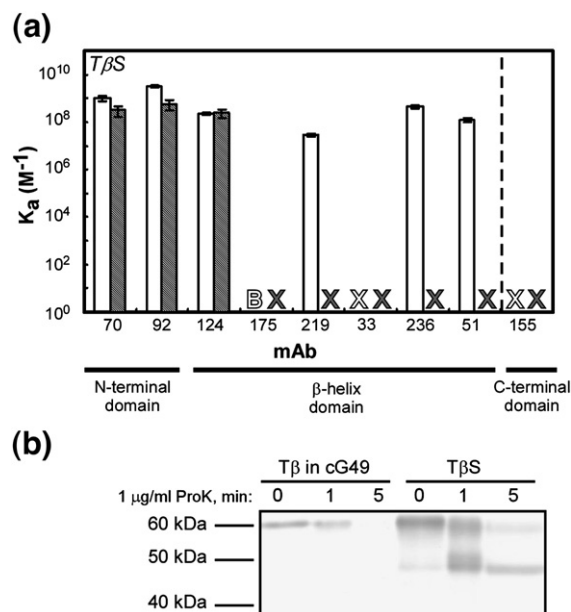
**Fig. 3.** Protease digestion of ribosome-bound tailspike nascent chains. (a) Ribosomes bearing TFS nascent chains were treated with 1  $\mu\text{g}/\text{mL}$  proteinase K for the indicated times at 4 °C. The resulting fragments were separated on 5–14% denaturing polyacrylamide gels, transferred to a polyvinylidene fluoride membrane, and immunoblotted using mAbs 70 and 92. (b) Ribosomes bearing TFS, T $\beta$ S, TMS, or TSS nascent chains were incubated for 5 min at 4 °C with or without 1  $\mu\text{g}/\text{mL}$  proteinase K. Digestion products were detected as in (a). Similar results were observed for room-temperature reactions, although the reactions proceeded more rapidly (data not shown).

tion, both released TSS and released TMS are recognized by an additional mAb: mAb 124 to released TSS and mAb 51 to released TMS. In both cases, the epitope for this additional mAb is located very close to the C-terminus of the construct, and ribosome release may increase antibody access to the nascent chain. For T $\beta$ S, all seven mAbs that bind to T $\beta$ S nascent chains bind with similar affinities to T $\beta$ S released chains, again with the caveat that background binding for mAb 175 to released chains was unusually high.

To test whether these tailspike truncations can also adopt ordered, native-like structure independent of the ribosome, we expressed and purified the tailspike truncation constructs as free chains (without the SecM stall sequence; designated TS, TM, and T $\beta$ ). As with released chains bearing the SecM stall sequence, all three chain lengths aggregated severely when expressed intracellularly, even at 30 °C (data not shown). After purification under denaturing conditions, we attempted to refold each truncated chain length. However, only the T $\beta$  construct produced detectable levels of soluble protein. We measured the binding of the anti-tailspike mAbs to this refolded T $\beta$ . The three mAbs with the most N-terminal epitopes (70, 92, and 124) bound refolded T $\beta$  with high affinity. In contrast, the other mAbs that bind to the T $\beta$ S nascent chain, all of which have epitopes located further into the  $\beta$ -helix domain, showed essentially no binding affinity to T $\beta$  (Fig. 4a).

To further investigate the relationship between the conformations of refolded truncated chains and ribosome-bound nascent chains, we subjected truncated, refolded T $\beta$  to limited proteolysis with proteinase K under similar conditions to those used for digesting ribosome-bound tailspike nascent chains. To account for the possibility that the protease resistance of ribosome-bound T $\beta$ S was due to ribosomal proteins competing with T $\beta$ S as substrate for proteinase K, we digested T $\beta$  in the presence of ribosomes expressing an unrelated ribosome-bound stalled nascent chain. Under these conditions, refolded T $\beta$  remains resistant to proteolytic digestion for approximately the same length of time as ribosome-bound T $\beta$ S, indicating that the refolded, truncated construct has adopted a well-ordered conformation. However, proteinase K treatment of refolded T $\beta$  did not produce a protease-resistant 47-kDa fragment, unlike digestion of ribosome-bound T $\beta$ S (Fig. 4b).

The lack of mAb binding to the  $\beta$ -helix and the change in protease digestion pattern of refolded T $\beta$  were surprising, as previous studies on the isolated  $\beta$ -helix domain of tailspike (Bhx) demonstrated that this domain refolds under low salt conditions to an active, native-like conformation.<sup>36,37</sup> When we refolded T $\beta$  using the low salt buffer used for Bhx refolding, we obtained far-UV CD and tryptophan fluorescence emission spectra largely similar to those obtained for Bhx,<sup>36,37</sup> despite the additional N-terminal domain present in the T $\beta$  construct. Yet, T $\beta$  refolded under low salt conditions also did not produce the characteristic 47-kDa protease-resistant



**Fig. 4.** Conformational differences between refolded T $\beta$  and the corresponding nascent, or released, polypeptide chain. (a) mAb binding affinity to T $\beta$ S released chains (open bars) and T $\beta$  refolded, truncated protein (shaded bars). An 'X' indicates that no measurable binding affinity could be detected, and a 'B' signifies high background binding (see Fig. 2). Error is reported as in Fig. 2. (b) Proteolysis of refolded T $\beta$ . Refolded T $\beta$  was diluted to a concentration similar to that of ribosome-bound T $\beta$ S in a solution containing ribosomes expressing an unrelated stalled nascent chain (cG49). Digestion of ribosome-bound T $\beta$ S is shown for comparison. The proteolysis reactions were performed at 4 °C and were visualized by Western blotting as described in Fig. 3.

fragment (Supplementary Fig. S2). To control for the influence of other cellular components (e.g., molecular chaperones) on the conformation of soluble T $\beta$ , we also refolded T $\beta$  in the presence of a cleared *E. coli* cell lysate. However, the presence of cellular factors, including molecular chaperones and ribosomes, had no effect on the yield of refolded T $\beta$ , the appearance of the 47-kDa protease-resistant fragment (Supplementary Fig. S3), or the mAb binding profile (not shown). Taken together with the mAb binding results above, these results indicate that the  $\beta$ -helix domain, upon dilution from denaturant, does not populate the same conformations as the corresponding nascent chain when associated with the ribosome or even after ribosome release.

### Molecular chaperones are not required for cotranslational folding of tailspike nascent chains

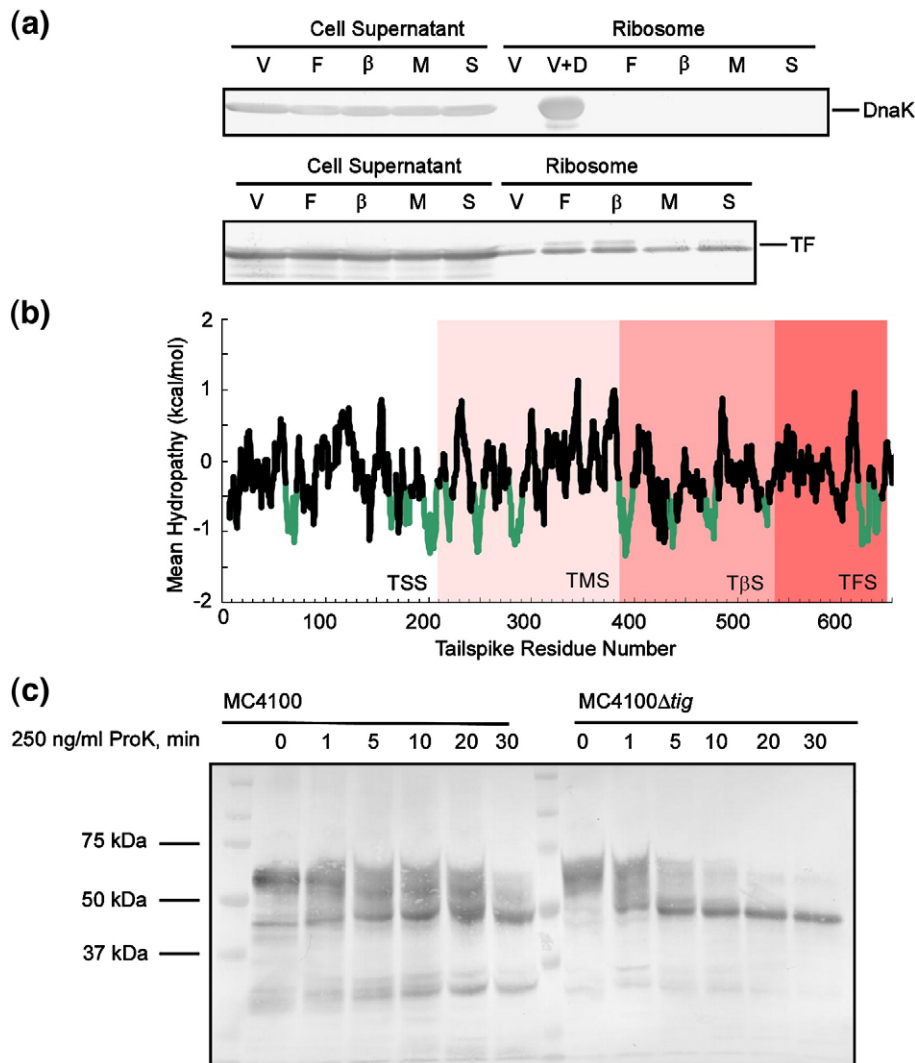
Molecular chaperones are known to play an important role in the cotranslational folding of several proteins. In order to account for the effects of chaperones on tailspike folding, we first measured recruitment of the ribosome-associated chaperones DnaK and TF. While we were unable to demonstrate recruitment of DnaK to ribosomes bearing tailspike

nascent chains, we were able to detect some recruitment of TF to tailspike RNCs (Fig. 5a). TF has been shown to bind hydrophobic stretches of nascent polypeptides.<sup>38</sup> When we analyzed the hydropathy of the tailspike amino acid sequence, we found several stretches of hydrophobic amino acids, which may provide binding sites for TF (Fig. 5b). To test whether TF plays a role in tailspike cotranslational folding, we expressed T $\beta$ S nascent chains in an *E. coli* strain lacking TF (MC4100 $\Delta$ tf) and a parent strain with TF (MC4100). Time-resolved proteinase K digestion of these T $\beta$ S RNCs revealed that the 47-kDa stable fragment appears regardless of the presence or absence of TF and is the major digestion product of T $\beta$ S from either cell line (Fig. 5c). MC4100 is not a protease-deficient

strain, and the presence of many other bands besides the major digestion products might be due to endogenous *E. coli* proteases. Moreover, the digestion of T $\beta$ S is slightly faster in the TF deletion strain, which could be indicative of TF restricting access of the protease to the ribosome-bound nascent chain. Regardless, these experiments show that tailspike nascent chains are able to form native-like structures in a cotranslational manner, without a requirement for TF chaperone activity.

## Discussion

The mAb binding and protease protection studies reported here demonstrate that the tailspike  $\beta$ -helix



**Fig. 5.** Chaperone association with ribosome-bound tailspike nascent chains. (a) Cell lysates were centrifuged through sucrose cushions. The unpelleted lysate supernatant was removed, and the ribosomes were resuspended. Aliquots of both the unpelleted lysate and the ribosomes were examined by Western blotting using anti-DnaK (top) or anti-TF (bottom) antibodies. V, empty vector control; F, TFS;  $\beta$ , T $\beta$ S; M, TMS; S, TSS. For DnaK, a vector control sample spiked with DnaK (1:1 DnaK:ribosome) is included as a positive control (V+D). (b) Hydropathy plot of tailspike. Mean hydropathy of the tailspike sequence was calculated according to published procedures.<sup>37</sup> Putative TF binding sites, as indicated by a mean hydropathy of less than  $-0.5$  kcal/mol, are shown in green. (c) Proteinase K digestion of T $\beta$ S RNCs from MC4100 or MC4100 $\Delta$ tf cells. Proteolysis reactions were performed at 4 °C and visualized by Western blotting as described in Fig. 3.

domain can fold to a native-like, compact conformation prior to release from the ribosome. Moreover, formation of this compact, native-like conformation begins prior to the complete synthesis of the  $\beta$ -helix domain, as demonstrated by the proteolytic protection and high-affinity mAb binding to the TMS nascent chain. Some threshold length of the  $\beta$ -helix appears to be required for the formation of stable structure, however, as the TSS nascent chain, with only three rungs of the  $\beta$ -helix exposed from the ribosome, is quickly digested.

Two mAbs, 33 and 155, did not recognize any of the four tailspike nascent chain lengths studied here. Under *in vitro* refolding conditions, both recognize only native trimer.<sup>19,31</sup> Moreover, the epitope for mAb 155 lies in the C-terminal interdigitated domain. It is likely that native-like structure in the interdigitated C-terminus forms only after post-translational trimerization, as even the TFS nascent chains are unlikely to adopt a native-like structure that requires interactions between the three monomer chains. The conformational epitope for mAb 33, like mAb 155, might require trimerization in order to form, as suggested by the *in vitro* refolding experiments. Surprisingly, however, the mAb 33 epitope is located in the center of the  $\beta$ -helix, near the mAb 236 epitope, and mAb 236 binds tightly to T $\beta$ S and TFS nascent chains. Nevertheless, subtle conformational differences between these epitopes might exist; alternatively, the mAb 33 epitope may be masked by an interaction between the nascent chain and the surface of the ribosome.

It is interesting to note that the size of the fragment produced upon proteinase K digestion of T $\beta$ S and TFS is indistinguishable from the size of the TMS species (47 kDa); that is, this 47-kDa fragment corresponds to the chain length produced during an endogenous rare-codon-derived translation pause point (Clarke and Clark, submitted). While the exact identity of the TFS and T $\beta$ S 47-kDa fragment is unknown, it is recognized by mAbs 70 and 92, both of which recognize epitopes located in the tailspike N-terminal domain. Because the size of the fragment is much larger than the N-terminal domain itself (16 kDa), a fragment of 47 kDa could include the entire N-terminal domain and still accommodate a sizeable portion (seven rungs) of the  $\beta$ -helix domain. These results indicate that proteinase K digestion occurs primarily at the C-terminus of the nascent chain, suggesting that this more recently synthesized portion is less stably folded than the N-terminus of the nascent chain (Fig. 6).

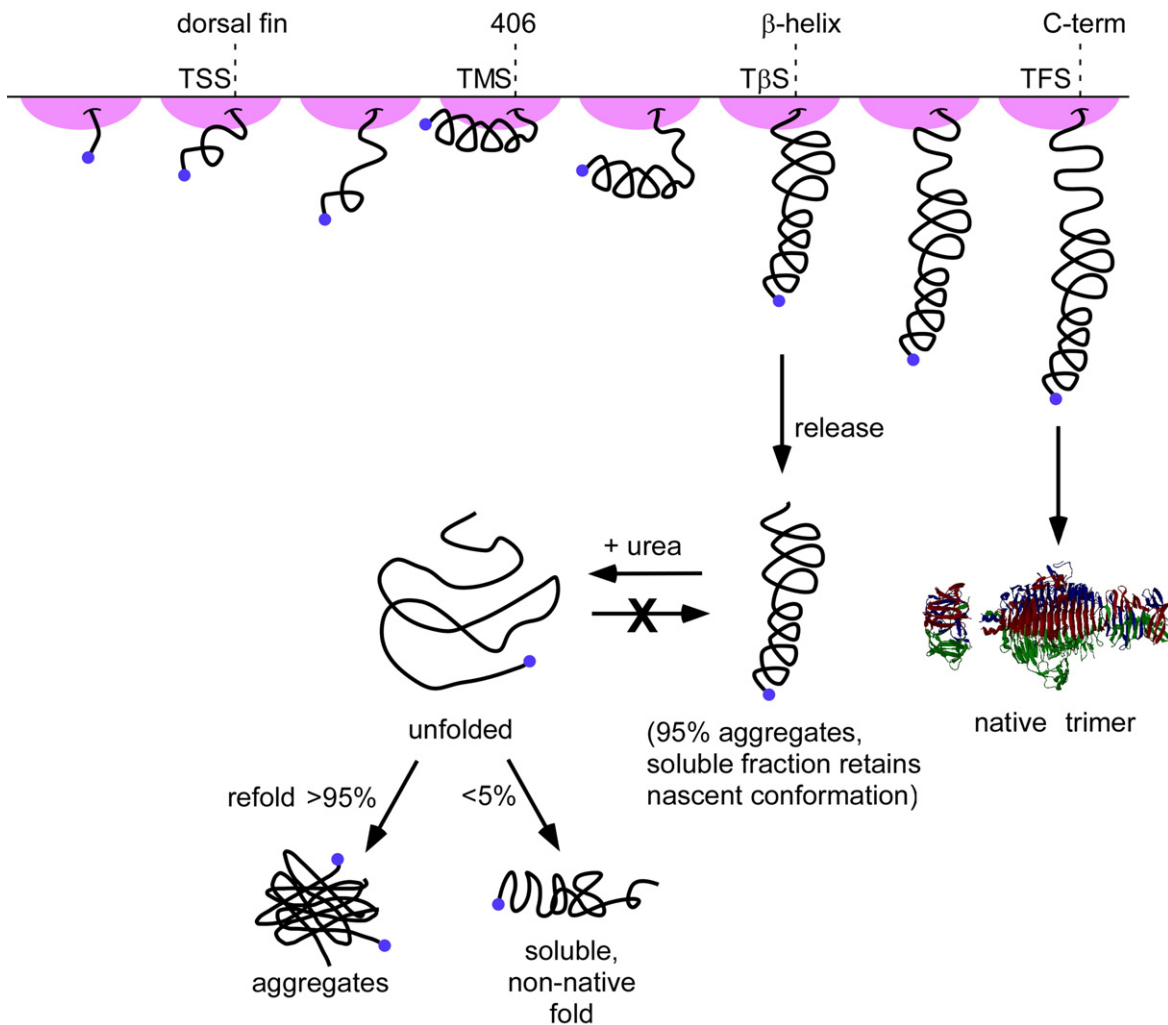
When tailspike nascent chains are released prematurely from the ribosome, the truncated polypeptide chains primarily aggregate, although a small fraction remain soluble and maintain conformations similar to those observed for the corresponding ribosome-bound nascent chain. A small fraction of the aggregated released T $\beta$ S chains, after solubilization in chemical denaturant and dilution into buffer, are able to refold into a soluble, “well-

behaved” conformation. Yet, surprisingly, this conformation is distinct from the conformation of nascent T $\beta$ S chains on the ribosome or released chains (Fig. 6). This suggests that the refolding of denatured, truncated tailspike chains results in a fundamentally different final conformation (or ensemble of conformations) than the folding of newly synthesized, ribosome-bound tailspike chains. More generally, it appears that cotranslational folding ushers the ribosome-bound tailspike polypeptide chain into an energy minimum that is inaccessible to free truncated polypeptide chains with untethered termini.

The tendency of released (and refolded) C-terminal tailspike truncations to aggregate is not surprising. For example, *in vitro*, the TM and TS truncations are expected to be aggregation-prone via intermolecular interactions involving their uncapped C-terminal ends of the  $\beta$ -helix domain, as has been predicted for uncapped  $\beta$ -structure in general.<sup>39</sup> Tethering the C-terminus of the nascent chain to the ribosome would shield the partially folded  $\beta$ -helix from such associations. Alternatively, the ribosome may play a more active role, functioning as a scaffold or template for incompletely synthesized tailspike chains. While it is unclear what portions of the ribosome surface might serve as a scaffold, interactions between full-length tailspike chains and 30S ribosome subunits have been reported previously.<sup>30</sup> The exit tunnel and the immediate surroundings likely play a role in protecting the nascent chains, as mAb 51, whose epitope is located near the C-terminus of TMS, can bind released, but not ribosome-associated, TMS. The same is likely true for mAb 124 binding to released (but not nascent) TSS.

Some have speculated that interactions with TF delay cotranslational folding in prokaryotes.<sup>38</sup> TF might therefore explain inefficiencies reported for the folding of multidomain proteins in prokaryotic translation systems.<sup>4,7</sup> Yet, while protease digestion of nascent T $\beta$ S on ribosomes lacking TF is slightly accelerated, we did not observe an effect of TF on tailspike nascent chain conformations. Hence, while TF is bound to ribosomes translating tailspike nascent chains, it is not required for the formation of protease-resistant structure. Tailspike nascent chains fold to a native-like structure cotranslationally, on prokaryotic ribosomes, and this process is neither aided nor retarded by the presence of TF.

Cotranslational folding may represent a general strategy for proteins—and, in particular, large, multidomain proteins—to fold productively and avoid aggregation during biogenesis. For tailspike folding *in vivo*, we propose that cotranslational folding of the  $\beta$ -helix domain allows tailspike to preferentially partition into the productive folding pathway and away from the aggregation-prone conformations that dominate refolding *in vitro* (Fig. 6). Previous studies have demonstrated that early, proper folding of the  $\beta$ -helix domain is crucial to the assembly of the native trimer.<sup>40,41</sup> More generally, cotranslational folding may represent a strategy for



**Fig. 6.** Model for cotranslational folding of tailspike nascent chains. When the first rungs of the  $\beta$ -helix emerge from the ribosome exit tunnel (TSS), little native-like or compact structure can form. As more of the  $\beta$ -helix is translated, a relatively stable compact structure containing native-like interactions forms throughout the first seven rungs of the  $\beta$ -helix domain (TMS). Once the entire  $\beta$ -helix domain is exposed from the ribosome (T $\beta$ S), native-like compact structure forms throughout a significant portion of the chain, including the mAb 51 epitope (around residue 406). Based on the results of the protease protection experiments, the tailspike C-terminal interdigitated domain may be in a less compact conformation than the C-terminus of the  $\beta$ -helix domain. Free from the ribosome, truncated chains tend to aggregate, but the fraction remaining in solution maintains the conformation adopted while ribosome-bound. Chemical denaturation and refolding yield primarily insoluble aggregates, but the small amount of T $\beta$  that refolds to a soluble, compact structure adopts a conformation that is distinct from the conformation of T $\beta$ S while it is ribosome-bound.

efficient folding of other  $\beta$ -sheet topologies. The ribosome, perhaps in concert with TF, can act as a temporary cap on the C-terminal strands as they emerge from the ribosome. As seen for free, truncated tailspike polypeptides, exposure of uncapped  $\beta$ -strands provides an excellent nucleus for aggregation that can be shielded during cotranslational folding on the ribosome.

## Methods

### Plasmids

Construction of plasmids pET21b/TFS, pET21b/T $\beta$ S, and pET21b/TSS was described previously.<sup>29</sup> pET21b/

TMS was constructed similarly, using the primers 5'-GGGAATCCATATGACAGACATCACTGCAAACG-3' and 5'-CCCGAGCTCCTCCGGATTCATGTCAGTGTG-3'. Plasmids pET21b/T $\beta$ , pET21b/TM, and pET21b/TS, lacking the C-terminal SecM stall sequence, were constructed by amplifying the tailspike coding regions from the corresponding stall plasmids using the same 5' primer and a 3' primer incorporating an XhoI restriction site. PCR products were cloned into the pET21b vector (Novagen) between the NdeI and XhoI sites, placing the tailspike coding sequence in frame with the C-terminal His tag.

### Stalled RNC preparation

Ribosomes bearing nascent chains stalled at the SecM stall sequence were produced and purified as described previously.<sup>29</sup>

### Preparation of truncated, released tailspike polypeptide chains

Cleared lysates from cells expressing the tailspike truncations with the SecM stall sequence were centrifuged through a sucrose cushion as described previously for purifying stalled RNCs.<sup>29</sup> Released, truncated tailspike chains were collected from the top of the sucrose cushion following ultracentrifugation.

### Purification of truncated, refolded tailspike polypeptide chains

*E. coli* strain BL21(DE3)pLysS transformed with pET21b/T $\beta$ , pET21b/TM, or pET21b/TS was grown in LB supplemented with 100  $\mu$ g/mL ampicillin at 37 °C to an OD<sub>600</sub> (optical density at 600 nm) of ~0.5. Expression was induced by adding IPTG to 500  $\mu$ M; growth continued for 4 h. Cells were collected by centrifugation, and the pellet was frozen overnight at -20 °C. The pellet was thawed on ice and resuspended in CR buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub> and 300 mM NaCl, pH 8). Lysozyme was added to 1 mg/mL, and the mixture was incubated on ice for 30 min, followed by sonication (twelve 30-s bursts separated by 30-s rests) on ice. The resulting partially lysed cells were frozen for 90 min at -80 °C, thawed, and sonicated as before. The lysate was then centrifuged at 10,000g for 25 min. The pellet was resuspended in PR buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, and 6 M guanidinium hydrochloride, pH 8) and filter sterilized.

His-tagged proteins were purified by passage through Ni-NTA superflow resin (Qiagen). The column was washed with 50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 6 M guanidinium hydrochloride, and 20 mM imidazole, pH 8 (wash buffer), followed by an imidazole gradient (20–250 mM) to elute the His-tagged protein.

Fractions containing truncated tailspike proteins were concentrated using an Amicon Ultra 10,000 MWCO (molecular weight cutoff) filter. Imidazole was removed by washing with PR buffer. Concentrated protein was refolded by rapid 50-fold dilution into refolding buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 2 mM ethylenediaminetetraacetic acid, and 3 mM  $\beta$ -mercaptoethanol, pH 8), followed by incubation at 0 °C for at least 48 h. Low salt refolding was performed as described previously.<sup>37</sup> Purified protein was concentrated using Amicon Ultra 10,000 MWCO filters.

### ELISA/*K<sub>d</sub>* measurements

Competition ELISAs were used to measure the binding of anti-tailspike mAbs to stalled ribosome nascent chain complexes and free truncated tailspike proteins, performed as previously described.<sup>19,29</sup> Dissociation constants were converted to association constants to facilitate graphical comparison of binding results (Fig. 2). Due to experimental limitations, it was not possible to measure affinity constants <10<sup>6</sup> M<sup>-1</sup>. The concentration of free truncated chains was determined by absorbance at 280 nm using the molar extinction coefficient determined from the following equation:<sup>42</sup>

$$\epsilon = 5500(\text{Trp}) + 1490(\text{Tyr}) + 125(\text{disulfide}).$$

### Proteolysis

Limited proteolysis of RNCs bearing TFS, T $\beta$ S, TMS, or TSS nascent chains was initiated by mixing an aliquot of

each ribosome sample with a solution of proteinase K. The final protease concentration was 1  $\mu$ g/mL. Longer time points were initiated first, and reactions were terminated at the end of the experiment by addition of a one-fifth volume of 100% trichloroacetic acid (w/v). Samples were incubated for 30 min at -20 °C to completely precipitate proteins. The precipitates were centrifuged and washed with acetone:hydrochloric acid (19:1). Washed pellets were then prepared for SDS-PAGE and Western blotting as described below. Refolded T $\beta$  was first diluted to 200 nM in R buffer (50 mM Tris, 10 mM MgCl<sub>2</sub>, and 150 mM KCl, pH 7.5) containing ribosomes bearing an unrelated stalled nascent chain (cG49) prior to addition of proteinase K.

### Acknowledgements

We are grateful to Mary C. Finn and Daniel Lano for help with the preparation of truncated polypeptide chains, Krastyu Ugrinov for ribosome-bound cG49 nascent chains, Bernd Bukau for providing the anti-TF antibody, and Jose Barral and Ulrich Hartl for providing the MC4100-based *E. coli* strains. This work was supported by an award from the National Institutes of Health (GM 74807).

### Supplementary Data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.jmb.2008.07.035](https://doi.org/10.1016/j.jmb.2008.07.035)

### References

1. Ellis, R. J. (2001). Macromolecular crowding: obvious but underappreciated. *Trends Biochem. Sci.* **26**, 597–604.
2. Frydman, J., Erdjument-Bromage, H., Tempst, P. & Hartl, F. U. (1999). Co-translational domain folding as the structural basis for the rapid de novo folding of firefly luciferase. *Nat. Struct. Biol.* **6**, 697–705.
3. Fedorov, A. N. & Baldwin, T. O. (1999). Process of biosynthetic protein folding determines the rapid formation of native structure. *J. Mol. Biol.* **294**, 579–586.
4. Netzer, W. J. & Hartl, F. U. (1997). Recombination of protein domains facilitated by co-translational folding in eukaryotes. *Nature*, **388**, 343–349.
5. Nicola, A. V., Chen, W. & Helenius, A. (1999). Co-translational folding of an alphavirus capsid protein in the cytosol of living cells. *Nat. Cell Biol.* **1**, 341–345.
6. Kolb, V. A., Makeyev, E. V. & Spirin, A. S. (2000). Co-translational folding of an eukaryotic multidomain protein in a prokaryotic translation system. *J. Biol. Chem.* **275**, 16597–16601.
7. Hirano, N., Sawasaki, T., Tozawa, Y., Endo, Y. & Takai, K. (2006). Tolerance for random recombination of domains in prokaryotic and eukaryotic translation systems: limited interdomain misfolding in a eukaryotic translation system. *Proteins*, **64**, 343–354.
8. Hartl, F. U. & Hayer-Hartl, M. (2002). Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science*, **295**, 1852–1858.

9. Deuerling, E., Schulze-Specking, A., Tomoyasu, T., Mogk, A. & Bukau, B. (1999). Trigger factor and DnaK cooperate in folding of newly synthesized proteins. *Nature*, **400**, 693–696.
10. Deuerling, E., Patzelt, H., Vorderwulbecke, S., Rauch, T., Kramer, G., Schaffitzel, E. *et al.* (2003). Trigger factor and DnaK possess overlapping substrate pools and binding specificities. *Mol. Microbiol.* **47**, 1317–1328.
11. Genevaux, P., Keppel, F., Schwager, F., Langendijk-Genevaux, P. S., Hartl, F. U. & Georgopoulos, C. (2004). In vivo analysis of the overlapping functions of DnaK and trigger factor. *EMBO Rep.* **5**, 195–200.
12. Clark, P. L. (2004). Protein folding in the cell: reshaping the folding funnel. *Trends Biochem. Sci.* **29**, 527–534.
13. Neira, J. L. & Fersht, A. R. (1999). Exploring the folding funnel of a polypeptide chain by biophysical studies on protein fragments. *J. Mol. Biol.* **285**, 1309–1333.
14. Chow, C. C., Chow, C., Raghunathan, V., Huppert, T. J., Kimball, E. B. & Cavagnero, S. (2003). Chain length dependence of apomyoglobin folding: structural evolution from misfolded sheets to native helices. *Biochemistry*, **42**, 7090–7099.
15. Woolhead, C. A., McCormick, P. J. & Johnson, A. E. (2004). Nascent membrane and secretory proteins differ in FRET-detected folding far inside the ribosome and in their exposure to ribosomal proteins. *Cell*, **116**, 725–736.
16. Lu, J. & Deutsch, C. (2005). Folding zones inside the ribosomal exit tunnel. *Nat. Struct. Mol. Biol.* **12**, 1123–1129.
17. Evans, M. S., Clarke, T. F. I. & Clark, P. L. (2005). Conformations of co-translational folding intermediates. *Protein Pept. Lett.* **12**, 189–195.
18. Steinbacher, S., Miller, S., Baxa, U., Budisa, N., Weintraub, A., Seckler, R. & Huber, R. (1997). Phage P22 tailspike protein: crystal structure of the head-binding domain at 2.3 Å, fully refined structure of the endorhamnosidase at 1.56 Å resolution, and the molecular basis of O-antigen recognition and cleavage. *J. Mol. Biol.* **267**, 865–880.
19. Jain, M., Evans, M. S., King, J. & Clark, P. L. (2005). Monoclonal antibody epitope mapping describes tailspike  $\beta$ -helix folding and aggregation intermediates. *J. Biol. Chem.* **280**, 23032–23040.
20. Goldenberg, D. & King, J. (1982). Trimeric intermediate in the in vivo folding and subunit assembly pathways of the tailspike endorhamnosidase of bacteriophage P22. *Proc. Natl. Acad. Sci. USA*, **79**, 3403–3407.
21. King, J., Haase-Pettingell, C., Robinson, A. S., Speed, M. & Mittraki, A. (1996). Thermolabile folding intermediates: inclusion body precursors and chaperonin substrates. *FASEB J.* **10**, 57–66.
22. Benton, C. B., King, J. & Clark, P. L. (2002). Characterization of the protrimer intermediate in the folding pathway of the interdigitated  $\beta$ -helix tailspike protein. *Biochemistry*, **41**, 5093–5103.
23. Mittraki, A., Fane, B., Haase-Pettingell, C., Sturtevant, J. & King, J. (1991). Global suppression of protein folding defects and inclusion body formation. *Science*, **253**, 54–58.
24. Steinbacher, S., Seckler, R., Miller, S., Steipe, B., Huber, R. & Reinemer, P. (1994). Crystal structure of P22 tailspike protein: interdigitated subunits in a thermally stable trimer. *Science*, **265**, 383–386.
25. Haase-Pettingell, C. & King, J. (1997). Prevalence of temperature sensitive folding mutations in the parallel  $\beta$ -coil domain of phage P22 tailspike endorhamnosidase. *J. Mol. Biol.* **267**, 88–102.
26. Betts, S. D. & King, J. (1998). Cold rescue of the single-chain tailspike intermediate at the junction between productive folding and off-pathway aggregation. *Protein Sci.* **7**, 1516–1523.
27. Sather, S. K. & King, J. (1994). Intracellular trapping of a cytoplasmic folding intermediate of the phage P22 tailspike using iodoacetamide. *J. Biol. Chem.* **269**, 25268–25276.
28. Gordon, C. L., Sather, S. K., Casjens, S. & King, J. (1994). Selective in vivo rescue by GroEL/ES of thermolabile folding intermediates to phage P22 structural proteins. *J. Biol. Chem.* **269**, 27941–27951.
29. Evans, M. S., Ugrinov, K. G., Frese, M. A. & Clark, P. L. (2005). Homogeneous stalled ribosome nascent chain complexes produced in vivo or in vitro. *Nature Meth.* **2**, 757–762.
30. Clark, P. L. & 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.
31. Friguet, B., Djavadi-Ohanian, L., King, J. & Goldberg, M. E. (1994). In vitro and ribosome-bound folding intermediates of P22 tailspike protein detected with monoclonal antibodies. *J. Biol. Chem.* **269**, 15945–15949.
32. Goldenberg, D., Berget, P. & King, J. (1982). Maturation of the tail spike endorhamnosidase of *Salmonella* phage P22. *J. Biol. Chem.* **257**, 7864–7871.
33. Friguet, B., Djavadi-Ohanian, L., Haase-Pettingell, C. A., King, J. & Goldberg, M. E. (1990). Properties of monoclonal antibodies selected for probing the conformation of wild type and mutant forms of the P22 tailspike endorhamnosidase. *J. Biol. Chem.* **265**, 10347–10351.
34. Speed, M. A., Morshead, T., Wang, D. I. & King, J. (1997). Conformation of P22 tailspike folding and aggregation intermediates probed by monoclonal antibodies. *Protein Sci.* **6**, 99–108.
35. Friguet, B., Chaffotte, A. F., Djavadi-Ohanian, L. & Goldberg, M. E. (1985). Measurements of the true affinity constant in solution of antigen–antibody complexes by enzyme-linked immunosorbent assay. *J. Immunol. Methods*, **77**, 305–319.
36. Danner, M., Fuchs, A., Miller, S. & Seckler, R. (1993). Folding and assembly of phage P22 tailspike endorhamnosidase lacking the N-terminal, head binding domain. *Eur. J. Biochem.* **215**, 653–661.
37. Miller, S., Schuler, B. & Seckler, R. (1998). A reversibly unfolding fragment of P22 tailspike protein with native structure: the isolated  $\beta$ -helix domain. *Biochemistry*, **37**, 9160–9168.
38. Kaiser, C. M., Chang, H. C., Agashe, V. R., Lakshminpathy, S. K., Etchells, S. A., Hayer-Hartl, M. *et al.* (2006). Real-time observation of trigger factor function on translating ribosomes. *Nature*, **444**, 455–460.
39. Richardson, J. S. & Richardson, D. C. (2002). Natural  $\beta$ -sheet proteins use negative design to avoid edge-to-edge aggregation. *Proc. Natl. Acad. Sci. USA*, **99**, 2754–2759.
40. Fuchs, A., Seiderer, C. & Seckler, R. (1991). In vitro folding pathway of phage P22 tailspike protein. *Biochemistry*, **30**, 6598–6604.
41. Simkovsky, R. & King, J. (2006). 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. USA*, **103**, 3575–3580.
42. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G. & Gray, T. (1995). How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* **4**, 2411–2423.