How the ribosome shapes cotranslational protein folding
Ekaterina Samatova¹, Anton A. Komar²,³ and Marina V. Rodnina¹

Abstract
During protein synthesis, the growing nascent peptide chain moves inside the polypeptide exit tunnel of the ribosome from the peptidyl transferase center towards the exit port where it emerges into the cytoplasm. The ribosome defines the unique energy landscape of the pioneering round of protein folding. The spatial confinement and the interactions of the nascent peptide with the tunnel walls facilitate formation of secondary structures, such as α-helices. The vectorial nature of protein folding inside the tunnel favors local intra- and inter-molecular interactions, thereby inducing cotranslational folding intermediates that do not form upon protein refolding in solution. Tertiary structures start to fold in the lower part of the tunnel, where interactions with the ribosome destabilize native protein folds. The present review summarizes the recent progress in understanding the driving forces of nascent protein folding inside the tunnel and at the surface of the ribosome.

Addresses
¹ Department of Physical Biochemistry, Max Planck Institute for Multidisciplinary Sciences, Goettingen 37077, Germany
² Center for Gene Regulation in Health and Disease, Department of Biological, Geological and Environmental Sciences, Cleveland State University, 2121 Euclid Avenue, Cleveland, OH 44115, USA
³ Department of Biochemistry and Center for RNA Science and Therapeutics, School of Medicine, Case Western Reserve University, Cleveland, OH 44106, USA

Corresponding author: Rodnina, Marina V. (rodnina@mpinat.mpg.de)

Introduction
The central dogma of protein science is that the amino acid sequence of the protein defines its unique three-dimensional structure. Indeed, many proteins can attain their native structures spontaneously. In vitro unfolding/refolding experiments revealed fundamental principles of protein folding in solution. However, many proteins, once unfolded, fail to refold easily; tend to misfold and aggregate, and require the help of chaperones to attain their correct structure [1,2]. Those proteins may depend on the pioneering round of folding that occurs as the ribosome synthesizes the nascent chain. This raises the question as to why and how do nascent proteins begin to fold, what is the role of the ribosome, and what are the differences in co- vs. post-translational folding.

Proteins begin to form secondary structures inside the peptide exit tunnel of the ribosome. The tunnel is about ~100 Å long and 10–30 Å wide (Figure 1a). It provides a narrow passage for the protein making its way from the peptidyl transferase center of the ribosome to the exit port opened to the cytosol. The tunnel not only provides a confined folding space, but also defines the physicochemical environment for protein compaction. Another characteristic feature of cotranslational folding is its vectorial nature. Due to the directionality of protein synthesis from the N- to C-terminus, folding can start as soon as the N-terminal part of the growing nascent chain becomes available, while the C-terminus of the peptide is attached to the tRNA. In contrast, post-translational refolding (with or without chaperones) involves full-length proteins that possess all amino acid residues required to form their native structures. The vectorial nature of folding, together with the non-uniform rate of translation, shape the unique cotranslational folding pathway that can be different from the folding landscape in solution and involves unique intermediates [3,4]. In addition, the ribosome destabilizes the emerging structures, which may help N-terminal protein domains to circumvent kinetic traps of misfolding [4–9].

Segments of the nascent chain that emerge at the tunnel exit port engage in interactions with the ribosome surface, molecular chaperones, folding catalysts, ribosome-associated protein biogenesis factors, membranes, etc. At this point in translation, nascent chains can fold into their native structures, but the proximity of the charged ribosome surface modulates the dynamics of the nascent domains. With the increasing distance from
the ribosome, ribosome-bound polypeptides start to behave more like isolated proteins in solution with their characteristic refolding profiles \[10-12\]. In this review, we will summarize what we know about cotranslational folding events and driving forces of the pioneering round of protein folding on the ribosome, with the emphasis on what happens inside the exit tunnel. While local translation speed is of key importance for understanding folding, the full coverage of this topic would be beyond the focus of this article. Likewise, the action of cotranslational chaperones, ribosome biogenesis factors, protein-targeting machineries, and the cotranslational protein assembly can be found in dedicated reviews.

**Forces driving peptide compaction inside the tunnel**

The tunnel is lined with rRNA and ribosomal proteins and has alternating wider or narrower segments with a tendency to expand towards the exit port \[13\]. In bacteria, extended loops of ribosomal proteins uL4 and uL22 form a constriction that separates an upper chamber of the tunnel proximal to the peptidyl transferase center from the middle part of the tunnel and the broader lower vestibule (Figure 1). In eukaryotes, the second constriction site separates the middle part from the vestibule. uL23, uL24 and uL29 shape the geometry of the lower tunnel region. The spatial dimensions of the tunnel provide a key determinant of nascent chain compaction inside the ribosome. Computer simulations of polypeptide folding in nanotubes suggested that confinement in the roughly cylindrical tunnel can induce folding by entropic destabilization of the coiled state, thereby favoring formation of $\alpha$-helices \[14-16\]. The narrow tunnel favors short-range (local) contacts, which are established significantly faster than the long-range contacts upon protein refolding from the denatured state, explaining why cotranslational folding is a more efficient process than refolding of a polypeptide from a completely unfolded state \[17\]. On the other hand, spatial confinement inside the tunnel hinders formation of bulky tertiary structures and destabilizes native protein structures at the ribosome surface near the exit port \[8,10,18-20\]. Thus, the length of the exit tunnel may have evolved to facilitate both efficient compaction of nascent protein inside the tunnel and its rearrangements upon appearance at the exit port where it is likely to encounter macromolecular crowders and interaction partners.

Another important feature is the anisotropic electrostatic environment inside the tunnel with a higher propensity of positive charges at the L4-L22 constriction and negative charges at the exit port \[13,21\] (Figure 1). Electrostatics of the tunnel and charges in the nascent chain contribute to both protein folding inside the tunnel and the local translation speed \[22,23\]. One
interesting possibility is that the incorporation of positively charged amino acids at the N-terminus of the nascent polypeptide may facilitate its progression towards the more negatively charged part of the exit tunnel [24]. The presence of several negatively charged amino acids at the N-terminus of the nascent polypeptide can cause ribosome stalling, premature termination and ribosome destabilization [25,26]. In addition to providing a directional ‘push forward’ to the growing peptide, electrostatic interactions with the ribosome may delay tertiary folding in the lower part of the tunnel [8,18]. Folding of tertiary structures and interactions with the tunnel walls generate force [27], which can modulate folding of the downstream portions of the nascent peptide. Thus, electrostatic effects, interactions with the ribosome, and the dimensions of the tunnel have a stimulatory effect on early compaction of secondary structures, but destabilize native structures as the protein leaves the tunnel.

**Protein dynamics in the upper section of the tunnel**

A nascent peptide entering the exit tunnel interacts with the rRNA and ribosomal proteins lining the tunnel walls. The C-terminal residues of the nascent chain are stabilized by the interactions with the universally conserved nucleotides G2061, A2062, and U2506 of the 23S rRNA [28], whereas the N-terminus becomes free to compact. Biophysical experiments of nascent chain dynamics show that compaction can start as early as at 13 amino acid nascent peptide length [29]. Some proteins have a high propensity to form helices before the constriction site [30,31]. Structural studies show that VemP stalling peptide [32] and the topoisomerase subunit of the T4 bacteriophage DNA polymerase [33] form α-helices both before and after the uL4-uL22 constriction site (Figure 2). Notably, VemP and the N-terminal helix of the methyl transferase HemK have low propensity to form structures in solution, but are folded on the ribosome [3,32]. This is consistent with the computational work suggesting that the tunnel environment can induce secondary structure formation in nascent peptides [14–16]. The importance of α-helices at early stages of cotranslational folding is underscored by the finding that cold-shock protein A (CspA), which is β-stranded structure in its native conformation, forms a α-helix inside the tunnel [29] (Figure 2).

At the L4-L22 constriction site, conserved positively charged residues form an interaction network with arrest peptides such as MifM [34], SecM [35], VemP [32] and TnaC [36] or in eukaryotic XBP1u [37]. Interactions of nascent peptide with Arg67 and Arg61 in uL4 and Arg95 in uL22 play a role in stabilizing a catalytically inactive conformation of the peptidyl transferase center during a recoding event upon translation of gene 60 mRNA of the T4 phage in bacteria. They also induce the rolled conformation of the ribosome, demonstrating how interactions of the nascent chain with the tunnel are communicated to the functional centers of the ribosome [33]. The same Arg residues in uL4 and uL22 respond to changes in the conformation of nascent CspA, thereby modulating the dynamics of the rRNA residues at the peptidyl transferase center [29]. Notably, the N-terminus of CspA becomes less dynamic when the peptide grows from 14 to 19 amino acids, suggesting that the constriction not only resolves the secondary structure [31], but may also stabilize the nascent chain in a less dynamic state [29]. Together, these data suggest the role of the constriction site as a sensor to regulate the activity of the peptidyl transferase center through the networks of interactions inside the tunnel.

**Folding in the middle and lower parts of the tunnel**

After the constriction, the tunnel gradually opens up, reaching >20 Å at its widest part at the so-called vestibule (Figure 1). Here, nascent chains can compact into tertiary structures and even small domains [3,27,38–40]. In the lower part of the tunnel, extended conserved loops of uL23 and uL24 protrude towards the nascent peptide [41,42]. In contrast to uL4 and uL22, which affect the peptidyl transferase activity, but not protein folding [18], tunnel loops of uL23 and uL24 alter nascent chain compaction [18,43], multidomain
protein folding [19], and interactions with ribosome-associated protein biogenesis factors [44,45]. When uL23 and uL24 loops were truncated, the dynamics of unfolded nascent chains and their propensity to fold increased due to reduced binding to the ribosome surface and increased space volume permitting an earlier folding onset for the nascent FLN5 peptide [18,43].

As the polypeptide grows, it continues to fold by forming tertiary structure elements between adjacent secondary structures [20,29,30,46] (Figure 3). Similarly to the secondary structure elements, tertiary structure intermediates may favor the local intra- and inter-chain interactions rather than the native contacts prevalent in solution. For example, folding of a small α-helical HemK NTD occurs in a stepwise manner, as indicated by time-resolved FRET, PET, and FPA measurement [3,20,46]. Helices appear to form inside the tunnel; however, the hydrophobic core of the protein forms only after the extrusion from the tunnel. The peptide attached to the ribosome remains dynamic, with elements of secondary structure fluctuating in the µs time range, and attains its native stable fold only after the release from the ribosome [20]. Similarly, CspA remains in highly dynamic state as long as it is attached to the ribosome, but rearranges into its native, stable state upon release into solution [29].

The Zn-finger domain of alcohol dehydrogenase regulator (ADR1) provides another well-studied example of domain folding inside the tunnel. Cryo-EM visualized the domain inside the middle part of the tunnel [27] (Figure 2). Optical tweezers combined with single-molecule FRET measurements and molecular dynamics simulations suggest that interactions with the tunnel accelerate folding and stabilize the folded state by decreasing the chain entropy due to electrostatic interactions between nascent peptide and ribosomal tunnel walls, akin to the proposed function of chaperonin during protein folding [47].

Figure 3
Schematic of vectorial cotranslational folding. a. Secondary structure elements, in particular α-helices, can form in the upper and middle parts of the exit tunnel, but remain dynamic and fluctuate between alternative structures. Ribosome induces local interactions in the nascent chain. The two orientations of a nascent chain (purple and blue) represent its dynamics in the tunnel. Ribosomal proteins uL4 and uL22 interact with the nascent chains and modulate the activity of the peptidyl transferase center depending on the nascent chain conformation. b. Sequential docking of secondary structures into a dynamic tertiary fold in the lower part of the tunnel. These structures can be native-like, non-native or resemble a molten globule. Interactions with uL23, uL24 and uL29 modulate the native structure formation. c. Upon emerging from the ribosome, nascent chains adopt near-native structures that remain unstable due to interactions with the ribosome surface. d. The nascent protein domain (purple) adopts its native fold upon moving away from the ribosome surface. However, domain stability and folding can be affected by the folding of the neighboring domain (magenta) or by interactions of its binding partners (not shown). The protein adopts its stable native structure after the release from the ribosome.

Protein folding at the ribosome surface
The next level of compaction entails vectorial folding of domains. For example, folding of NBD1 of CFTR (nucleotide-binding domain 1 of human cystic fibrosis transmembrane conductance regulator) occurs vectorially starting with compaction of the ~112-aa N-terminal α/β-subdomain that forms a functional ATP binding site [48]. Also folding of the N-terminal domains of HemK, γB-crystallin, or EF-G fold into their native-like structures upon emerging from the ribosome before synthesis of the C-terminal domains of these proteins is completed [3,46,49–51]. Numerous reports indicate that the ribosome alters the dynamics of nascent polypeptides emerging from the exit tunnel. FLN5, RNase H and DHFR domain folding is destabilized at the ribosome surface, but become stable with the increasing linker length [10,52]. Destabilization is due to electrostatic interactions of nascent peptides with the ribosome surface, which competes with protein domain folding and delays it [19]. Similarly, folding experiments with ribosome-bound DHFR that used methionine oxidation as a readout for correct folding suggest that the protein attains its native fold only at a linker length of 70 aa and that the electrostatic interactions with the ribosome inhibit folding [53].

Interactions between emerging nascent polypeptide chains and the ribosome can also modulate cotranslational protein assembly. Two intrinsically disordered
proteins of opposite charge, ACTR and NCBD, form a high-affinity complex in a coupled folding-and-binding reaction and their interaction is modulated by the ribosome [54]. The negatively charged ribosomal surface binds the positively charged nascent chain of NCBD, thereby preventing ACTR binding. In contrast, the negatively charged nascent ACTR is repelled by the ribosomal surface and thus remains available for productively binding its partner [54].

In addition to the ribosome itself, also interactions between protein domains on the ribosome modulate folding (Figure 3). If the C-terminal portion of nascent peptide starts to fold inside the tunnel, it can delay folding of the N-terminal domain, because it ‘pulls’ the domain back into the narrow vestibule. Optical tweezers experiments on folding of the N-terminal GTP-binding domain of EF-G showed that the G-domain is able to fold off and on the ribosome, but ribosome destabilizes its folded structure and reduces the folding rates [55]. Neighboring domains can cause unfolding; in this case, chaperon TF can prevent misfolding of the destabilized domain by protecting it until the C-terminal part of the polypeptide chain is long enough to form a compact structure [5]. Finally, cotranslational interaction between nascent proteins emerging from nearby ribosomes can modulate cotranslational assembly of protein complexes and such interactions include both homo- and hetero-oligomers (for review see Ref. [56]).

By destabilizing incompletely synthesized proteins, the ribosome may prevent premature folding into stable non-native structures and formation of detrimental folding intermediates until the entirety of a folding domain has emerged from the ribosomal exit tunnel. While the ribosome and the chaperones can help the protein to find its native fold, not every misfolding event can be reversed by keeping the nascent chain dynamic or by spontaneous global unfolding and refolding of the protein after its release from the ribosome. This notion is particularly important given that appearance of such non-native misfolded proteins can lead to disease. For example, two point mutations in the NBD1 transiently alter cotranslational folding pathway of this domain by delaying the formation and reducing the stability of intermediate structures at a particular polypeptide length, which, in turn, causes misfolding and dysfunction of the full-length CFTR protein [57]. Similarly, a point mutation E342K in human alpha-1-antitrypsin (AAT), which stalls full-length nascent chains on the ribosome, induces formation of a distinct cotranslational folding intermediate that contributes to extensive aggregation of the protein after its release from the ribosome. This leads to the loss of the protein from the bloodstream, and causes the severe form of alpha-1-antitrypsin deficiency [58]. Notably, cotranslationally formed misfolded structures can be extremely stable, for example, the non-native conformation adopted by partially synthesized tailspike nascent chains persists after nascent chain release from the ribosome [59]. This stable misfolded intermediate is uniquely cotranslational, because it does not form once those released chains were chemically denatured and refolded by dilution from denaturant. We note that such persistent misfolding occurs not only upon ribosome stalling, but also during on-going translation due to alterations of local translation rates, which are non-uniform along the mRNA. While the natural rate of translation has evolved to ensure correct folding, altered lifetimes of metastable cotranslational folding intermediates may result in partitioning to an alternative folding pathway leading to formation of misfolded or locally entangled states, which will affect protein structure, stability and activity of the protein after its release from the ribosome. These data demonstrate that the unique pathway of cotranslational folding can determine the fate of the protein and emphasize the importance of cotranslational folding for understanding the cellular proteostasis.

Conclusions and perspectives
The recent work on cotranslational folding suggests a dual role of the ribosome. The narrow exit tunnel enhances secondary structure formation, but destabilizes the native fold as the protein emerges at the vestibule. The physical properties of the tunnel (its dimensions, charge, water distribution) and the vectorial nature of cotranslational folding define the unique pioneering folding landscape. By restricting the interchain interactions to local contacts, the tunnel may prevent misfolding of incompletely synthesized proteins into misfolded states that do not refold easily in solution. Nascent chain interactions with L4 and L22 modulate the activity of the peptidyl transferase center, whereas interactions inside the tunnel with L23 and L24 attenuate folding. Still, many questions remain open. For example, it remains unclear how ‘conventional’ (i.e., non-stalling) peptides interacting with L4 and L22 can regulate the synthetic activity of the ribosome. Docking of the secondary structure elements into a tertiary fold may proceed through unique intermediates that are not populated during refolding in solution. Identification of such intermediates is a challenging task, but would be important to understand the role of potential non-native intermediates, misfolded and entangled states and the overall cotranslational folding landscape. Finally, understanding the interplay between cotranslational protein folding and the auxiliary factors that receive the nascent peptide upon extrusion from the vestibule (SRP, TF, PDF, MAP, DnaK) and their regulatory effect on the rates of translation and the nascent protein fold is important for understanding the link between translation, quality control and protein homeostasis in the cell.

Declaration of competing interest
None.
Data availability

No data was used for the research described in the article.

Acknowledgments

We thank Dr. Niels Fischer (MPI-NAT) for providing the figures for this article. The work was funded by the European Research Council (ERC) Advanced Investigator Grant RIBOFOLD (proposal number 769726) to M.V.R., the DFG Leibniz Prize, and the Max Planck Society. AAK is supported by the National Institutes of Health (NIH) grants HL151392 and GM128981.

References

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. To P, Xia Y, Lee SQ, Devin T, Fleming KG, Fried SD; **A proteome-wide map of chaperone-assisted protein refolding in a cytosol-like milieu. Proc Natl Acad Sci USA 2022, 119, e2210536119.**


17. Usage of a combination of NMR and ribosome mutations, the authors show the role of tunnel loops in ribosomal proteins for cotranslational protein folding.


20. The study describes how α helices fold sequentially into a compact protein domain inside the ribosome and shows that these domains are highly dynamic as long as the protein remains attached to the ribosome.


The study visualizes the structural switch from the α-helical conformation inside the tunnel to the native β-stranded conformation upon extrusion of the protein out of the ribosome and identifies nascent protein conformation-dependent interactions with L4 and L22 that modulate the structure of the peptidyl transferase center.

49. Liutkute M, Mati R, Samatova E, Enderlein J, Rodnina MV: Gradual compaction of the nascent peptide during cotranslational folding on the ribosome. Elife 2020, 9,