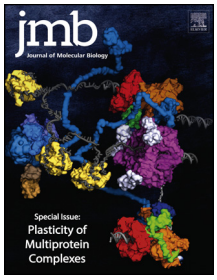

Folding up and Moving on—Nascent Protein Folding on the Ribosome

Christian M. Kaiser[1,2] **and Kaixian Liu**[1,3]

*1 - **Department of Biology,** Johns Hopkins University, 3400 N Charles St., Baltimore, MD 21218, USA*
*2 - **Department of Biophysics,** Johns Hopkins University, 3400 N Charles St., Baltimore, MD 21218, USA*
*3 - **CMDB Graduate Program,** Johns Hopkins University, 3400 N Charles St., Baltimore, MD 21218, USA*

***Correspondence to Christian M. Kaiser:** Department of Biology, 3400 N Charles St., 205 Mudd Hall, Baltimore, MD 21218, USA. kaiser@jhu.edu*
https://doi.org/10.1016/j.jmb.2018.06.050
Edited by Karl Duderstadt

Abstract

All cellular proteins are synthesized by the ribosome, an intricate molecular machine that translates the information of protein coding genes into the amino acid alphabet. The linear polypeptides synthesized by the ribosome must generally fold into specific three-dimensional structures to become biologically active. Folding has long been recognized to begin before synthesis is complete. Recently, biochemical and biophysical studies have shed light onto how the ribosome shapes the folding pathways of nascent proteins. Here, we discuss recent progress that is beginning to define the role of the ribosome in the folding of newly synthesized polypeptides.

Introduction

Proteins are linear polymers synthesized by the ribosome that generally have to fold into specific three-dimensional structures to become biologically active. Pioneering experiments by Anfinsen [1] showed that proteins fold spontaneously *in vitro*, demonstrating that the sequence of the polypeptide contains all the information necessary to specify the final structure, and that no additional components are required to guide the conformational search to the biologically active state. While Anfinsen's principle generally holds true for many small proteins, the folding of larger proteins is complicated by the population of off-pathway (or misfolded) species that compete with productive folding and often result in irreversible aggregation [2]. These competing reactions are particularly pronounced in the crowded cellular environment [3]. To promote productive folding, all cells contain an arsenal of molecular chaperones that guide initial folding, help to maintain proteins in their functional states, and are involved in protein quality control [4]. Specialized chaperones are targeted to proteins even before the ribosome has completed their synthesis [5, 6]. Recently, it has been realized that the ribosome itself modulates early folding events [7], helping to set the nascent polypeptide on the correct path for successful folding.

The ribosome is the central player in cellular protein synthesis, polymerizing amino acids in the order dictated by the genetic information in the messenger RNA (mRNA). For every codon in the mRNA, the ribosome helps to select the appropriate amino-acylated transfer RNAs (tRNAs) and subsequently catalyzes the addition of the selected amino acid to the nascent protein. By translating the information in the mRNA one codon at a time, the ribosome synthesizes polypeptides in a directed manner, adding amino acids to the carboxyl-terminus (C-terminus) of the nascent polypeptide. As a consequence, proteins emerge in a vectorial fashion, starting at the amino-terminus (N-terminus), as the ribosome moves along the mRNA (Fig. 1).

Protein synthesis by the ribosome is highly processive. Full-length translation products are released only when a stop codon in the mRNA is reached. While stalling of ribosomes during elongation can lead to ribosome rescue [14] or dissociation of the incomplete peptide [15], the growing nascent polypeptide remains stably anchored to either the A-site or the P-site tRNA

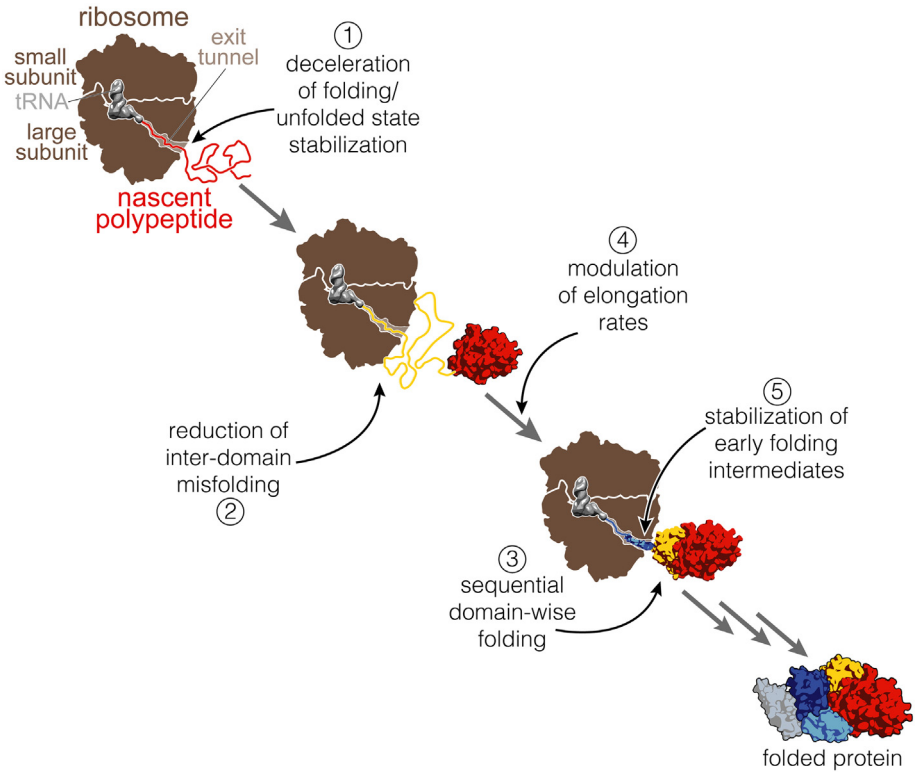

Fig. 1. Modulation of folding by the ribosome. The cartoon depicts ways in which the ribosome and the process of translation can affect the folding of nascent polypeptides. (1) Interactions with the ribosome surface reduce folding rates and stabilize the unfolded state [8–10]. (2) Presumably similar interactions might help to reduce misfolding among emerging domains in multi-domain proteins [11]. (3) Sequential domain-wise folding reduces the complexity of the conformational search problem [11]. (4) Local variations in elongation rate tune synthesis to folding, promoting the formation of native structures [12]. (5) Structures that are unstable by themselves can be stabilized in specific regions of the ribosome exit tunnel that spans the large subunit [13].

during normal elongation. At the same time, polypeptide elongation is relatively slow, proceeding at a rate of up to 20 amino acids per second in bacteria [16]. Elongation in eukaryotes is even slower, with a rate of ~6 amino acids per second as estimated from pulse chase [17] and ribosome-profiling [18, 19] experiments. Synthesis of even a small domain of 100 amino acids thus requires several seconds. Folding, on the other hand, can be much faster. While large proteins may take several minutes to complete folding, the formation of intermediates or the folding of domains is generally much faster, with time constants often in the millisecond range. Therefore, folding can in principle begin before synthesis is complete, as has been noted more than half a century ago [20]. Co-translational folding of ribosome-bound nascent proteins could indeed be observed experimentally by assaying the catalytic activity [21], the acquisition of protease-resistant structures [22], or the binding of conformation-specific antibodies [23]. These experiments also suggested that the particular environment near the ribosome and/or the vectorial process of polypeptide synthesis affect what structures the nascent polypeptide populates [22, 23]. Moreover, due to a variety of factors, the elongation rates along a given mRNA are not uniform [24]: some codons are translated faster than others, resulting local minima and maxima at specific positions during polypeptide synthesis. The modulation of elongation rates has been shown to affect the biological function [25, 26] or biochemical properties [27, 28] of the translation product. More recently, it has become possible to directly relate changes in elongation rate to the folding and structure of the newly synthesized polypeptide [12, 29] (discussed below).

Before being exposed to the cellular environment, nascent polypeptides pass through a long, narrow tunnel that spans the large subunit of the ribosome and can accommodate between 30 and 40 amino acids of a polypeptide in a largely unstructured conformation. While limited folding can occur within the interior of the ribosome, [30–32], the narrow exit tunnel does not allow extensive tertiary structure formation. Regulatory nascent peptides can adopt specific conformations within the exit tunnel and, through interactions with rRNA or protein components of the tunnel, regulate ribosome function

[33–37]. Recent advances in cryo-electron microscopy have yielded structural details of such nascent chain–ribosome interactions at very high resolution [38]. The ribosome–nascent chain interactions are required to stabilize these conformations [39], which are hence not stable outside the tunnel. Once the nascent polypeptide emerges from the tunnel, folding is less restricted, although close proximity of the ribosome limits the space available to the polypeptide and thus, in principle, the accessible conformations [40].

Experimentally measuring ribosome–nascent chain interactions and their effects on folding is challenging. Mechanistic studies of protein folding traditionally utilize thermal or chemical denaturation of the protein of interest and simultaneously monitoring folding and unfolding using optical spectroscopy. The biochemical complexity of the ribosome, which in bacteria is composed of three ribosomal RNA (rRNA) molecules and more than 50 proteins, precludes the utilization of intrinsic spectroscopic probes that are commonly employed to follow folding and unfolding (e.g., tryptophan fluorescence). Moreover, the traditional approach of adding and removing denaturants to study folding cannot be used, because globally acting denaturants destabilize not only nascent polypeptide structure, but also the ribosome itself. Over the last years, novel approaches have been developed to overcome these limitations. In the following sections, we will discuss key experiments that have begun to shed light on nascent chain folding in the unique environment of the ribosome. Figure 1 gives an overview of how interactions with the ribosome and the process of nascent chain elongation affect folding, based on these measurements. Most experiments to date have utilized ribosome–nascent chain complexes (RNCs) in which translation is stopped at precisely defined nascent chain lengths, either through the use of a genetically engineered arrest peptide from the SecM protein [41] or by using stop codon-less (nonstop) mRNA templates. Recently, a carefully designed *in vitro* translation system [42] has opened exciting avenues for simultaneously measuring of polypeptide elongation and folding.

This short review aims to highlight recent progress in understanding how ribosome–nascent chain interactions modulate protein folding. How local changes in protein elongation rates affect folding has recently been reviewed elsewhere [24, 43–45]. Similarly, we are not covering recent advances in computational modeling of co-translational folding [46]. New experimental approaches have recently enabled studies of nascent protein folding and chaperone interactions in live cells [47–49]. Here, we are focusing mainly on biochemical and biophysical experiments that have contributed mechanistic insights into nascent protein folding that can presently not be obtained from *in vivo* experiments.

Biochemical approaches

Several biochemical approaches have been used to map the onset of folding as a function of nascent chain length and the effect of the ribosome on folding and stability (Fig. 1). Hoffmann *et al.* [50] employed an experimental strategy relying on structure-dependent disulfide bond formation between strategically placed cysteine residues. These experiments revealed distance-dependent destabilization of nascent chain structure by the ribosome that was similar for a set of model proteins (including the Src homology 3 domain of α-spectrin, barnase, and β-lactamase). Interestingly, the effect is amplified by the ribosome-binding chaperone trigger factor [50], suggesting one mechanism by which the chaperone may cooperate with the ribosome to modulate nascent protein folding.

To obtain quantitative information on the destabilization caused by the ribosome, Samelson *et al.* [8] employed "pulse proteolysis" [51], a refined version of the limited proteolysis approach that had been exploited in early experiments studying the folding of nascent firefly luciferase [22]. Instead of incubating RNCs with low concentrations of protease, samples are "pulsed" with a high protease concentration for a brief period of time. The pulse duration is short relative to the mean lifetime of the folded state of the protein of interest. Consequently, only the fraction that populates the unfolded state during the pulse is digested, whereas the folded state population is protected. The equilibrium fraction of folded protein determined in this way directly yields its global stability. For the three globular proteins analyzed in this study (versions of dihydrofolate reductase, ribonuclease H and barnase), the ribosome caused a significant destabilization when the folded units were in close proximity to the ribosome. Lengthening the C-terminal extension from 35 to 55 amino acids restored stability, confirming earlier observations that the effect of the ribosome is restricted to its immediate vicinity [9, 10, 13, 32] (see below). Destabilization of folded structures in the immediate vicinity of the ribosome, presumably by interactions stabilizing the unfolded state, therefore emerges as a general phenomenon.

The SecM arrest peptide is frequently used as a tool to prepare stably stalled RNCs *in vitro* [52] or *in vivo* [53]. The mechanism underlying SecM-mediated arrest has also made this sequence a useful tool for studies of nascent chain folding dynamics. Biochemical and genetic studies have suggested that mechanical force, generated by the Sec translocon [41, 54] or insertion of transmembrane helices into lipid bilayers [55], accelerates release of SecM-induced elongation arrest. Single-molecule force spectroscopy experiments established that force is indeed sufficient to release arrest [48]. Folding of the nascent chain into a stable structure near the ribosome is thought to generate a pulling force that accelerates arrest release [48, 56]. When arrest release is coupled to an

experimental readout, this phenomenon can be exploited to define the nascent chains lengths at which folding occurs (Fig. 2). Systematically varying the distance between a folding domain and the ribosome confirmed that folding into stable structures occurs at a distance of approximately 30 to 40 amino acids from the peptidyl transferase center. Folding-mediated release of SecM arrest has been demonstrated in live *Escherichia coli* cells [48] (Fig. 2). It should be noted that the exact mechanism of force generation by nascent chain folding is not completely understood, and that arrest release is slow relative to normal elongation rates. SecM arrest release experiments are therefore difficult to interpret with certainty in terms of folding kinetics, but it seems clear that this experimental approach robustly detects the co-translational formation of stable structures in the nascent polypeptide.

Several studies have employed folding-mediated release of SecM arrest to detect folding of nascent membrane [55, 58] and soluble proteins [32, 48, 56, 59–61]. Nilsson *et al.* [61] utilized the approach, in combination with cryo-electron microscopy, to map the onset of co-translational folding. Different versions of the three-helix bundle spectrin domain begin to fold right around the position where they fully emerge from the ribosome exit tunnel. Analyzing several spectrin variants revealed that, surprisingly, the onset of nascent chain folding does not correlate with the intrinsic stabilities, folding rates or folding pathways of

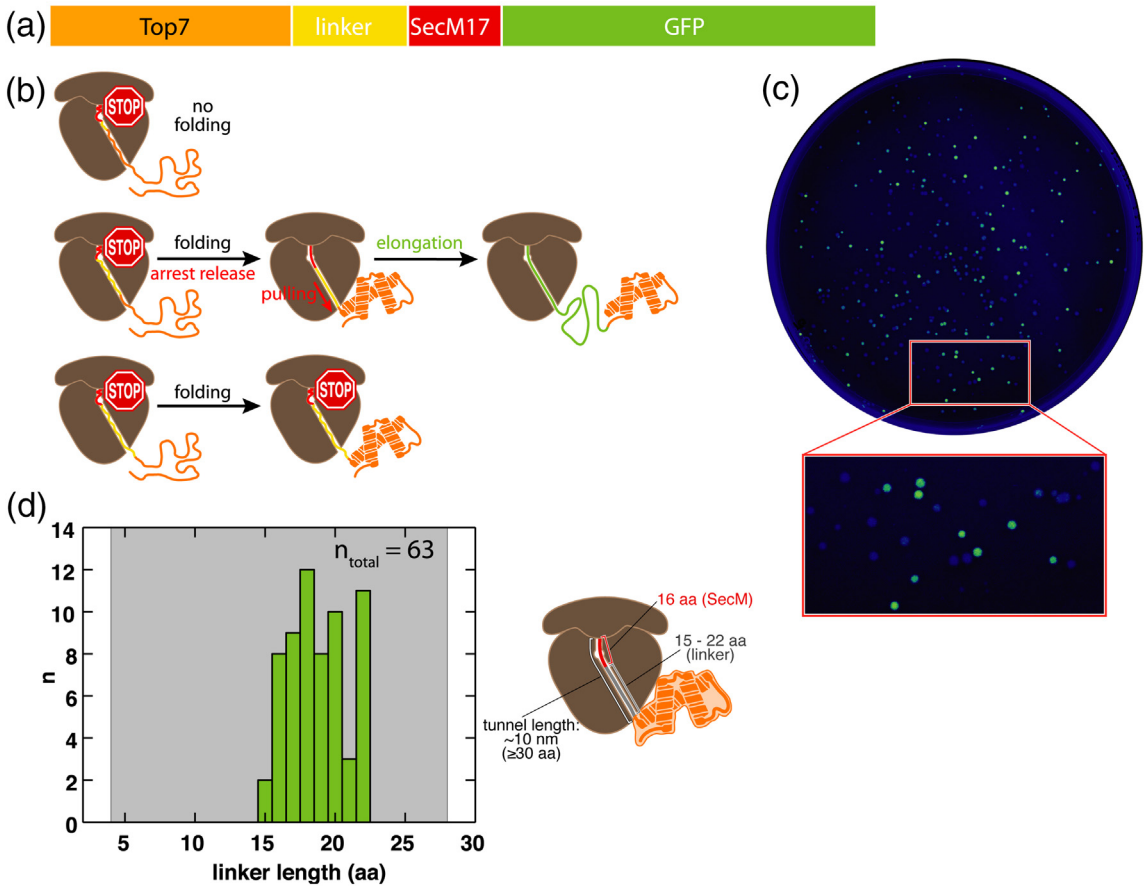

Fig. 2. The SecM arrest peptide as a tool to detect nascent chain folding. (a) Bar diagram of the construct used to assess nascent chain folding. The SecM arrest peptide (red) is fused to an independently folding domain (Top7 [57], orange) by linkers of variable length (yellow). The downstream green fluorescent protein (GFP, green) is translated only after SecM arrest has been released. Varying the linker length from 4 aa to 28 aa creates a linker "library." (b) Schematic illustrating the expected translation outcome of very short linkers (<<15 aa, top), intermediate length linkers (~15 aa, middle) and long linkers (>>15 aa, bottom). Force is generated by nascent chain folding at intermediate linker lengths, resulting in elongation restart. (c) Agar plate with colonies transformed with the linker library under UV illumination. A fraction of the colonies exhibits green fluorescence, the remaining colonies exhibit weak blue autofluorescence. (d) Distribution of linker lengths obtained by sequencing plasmid DNA from fluorescent colonies (green bars). The recovered length range represents a well-defined subset of all linker lengths present in the library (gray shaded area). The cartoon on the right illustrates that the lengths supporting elongation restart (31–38 aa, combining SecM and the flexible linker) are within the range expected based on the tunnel length. (modified from Goldman *et al.* [48]).

the isolated polypeptides. Instead, the results suggest that interactions with the ribosome can alter the folding pathway of a nascent chain compared to the same polypeptide in isolation. As a consequence, the folding pathways determined with isolated proteins may not always reflect the sequence of events during co-translational folding, and interactions with the ribosome can change the folding properties in ways that are at present not predictable. While the arrest peptide assays cannot resolve folding intermediates at the resolution provided by biophysical measurements, they nevertheless represent a very useful approach for defining "waypoints" for the co-translational folding of nascent proteins *in vitro* and in living cells.

Fluorescence spectroscopy

Incorporation of exogenous fluorophores into nascent polypeptides [62] has proven to be a useful tool for direct observations of nascent chain dynamics and folding through a variety of approaches. Measurements of the time-resolved anisotropy decay of a bodipy fluorophore at the N-terminus of ribosome-tethered apo-myoglobin showed reduced conformational flexibility, as compared to the free polypeptide after release from the ribosome [63, 64]. Experiments with a disordered protein further suggested that electrostatic interactions with the ribosome surface increase nascent chain dynamics [65], which may explain the destabilization of folded structures in ribosome-bound nascent chains that has been observed with other approaches [8–10].

By introducing two fluorophores into the nascent polypeptide, folding can be observed through fluorescence resonance energy transfer (FRET). Khushoo *et al.* [66] combined a genetically engineered N-terminal cyan fluorescent protein (CFP), serving as the FRET donor, with a small organic fluorophore incorporated at various positions in the nascent chain, serving as the FRET acceptor, to map the folding of the first nucleotide binding domain (NBD1) in nascent cystic fibrosis transmembrane conductance regulator (CFTR). FRET measurements indicated that NBD1 folded into a native-like structure even before the entire domain had emerged, suggesting the population of a folding intermediate comprising the N-terminal subdomain. Notably, binding of ATP stabilizes the intermediate, facilitating subsequent folding as more of the polypeptide becomes available during synthesis. Ligand binding therefore promotes the formation of a co-translational folding intermediate, driving vectorial folding during synthesis.

Since many proteins have binding sites for small molecules, co-translational stabilization of partially folded structures by ligand binding may be a general mechanism for facilitating co-translational folding. For a system with suitable properties, specific ligand binding itself can report on nascent chain folding and its modulation by the ribosome. Experiments exploiting the quenching of flavin mononucleotide upon binding to flavodoxin have recapitulated the previously observed destabilization of folded structures and indicated that stabilizing interactions due to ligand binding can shift the conformational equilibrium [67]. Specific binding interactions that drive nascent protein folding are not limited to small molecule ligands. Hetero-oligomeric complex formation can similarly begin while one subunit is still being synthesized, resulting in complex assembly with increased efficiency *in vitro* [68] and *in vivo* [69].

The careful design of an *in vitro* translation system that combined robust elongation rates with efficient incorporation of exogenous spectroscopic probes has recently enabled real-time measurements of co-translational folding [13]. FRET measurement, using donor and acceptor dyes incorporated into the nascent polypeptide (Fig. 3), indicated the formation of a compact intermediate early during translation that is not significantly populated in the isolated polypeptide. Based on nascent chain length, the intermediate forms within the distal part of the ribosome exit tunnel. Limited proteolysis and photo-induced electron transfer measurements confirmed the co-translational population of the intermediate that converts into the native structure upon chain elongation or release from the ribosome. More recent work [70] employing FRET and photo-induced electron transfer techniques shows that the N-terminal α-helical domain of nascent HemK, a universally conserved N^5-glutamine methyltransferase, forms at least four intermediates by a "helix docking" mechanism, which is rate-limited only by translation. These elegant experiments demonstrate that the ribosomal environment has a profound effect on nascent chain structure and pave the way to investigating the effects of elongation rate modulation on co-translational folding pathways in mechanistic detail.

NMR spectroscopy

FRET measurements provide a sensitive readout of distance changes with good time resolution, but yield only one-dimensional information about the structure formed by the nascent polypeptide. NMR spectroscopy not only provides a sensitive measurement of protein structure, but also has the potential to resolve heterogeneous ensembles of states that are likely populated during folding. While not probing co-translationally formed structures directly, NMR experiments with human gamma-B crystallin showed that synonymous changes in the coding sequence result in the formation of distinct structures of polypeptides with identical sequences, demonstrating the importance of local translation rates for correct folding [12].

Selective isotope labeling of newly translated polypeptides [53, 71] and a recently developed methyl transverse relaxation optimized spectroscopy-based pulse sequence [72] have enabled high-resolution measurements of nascent chain structure in the context

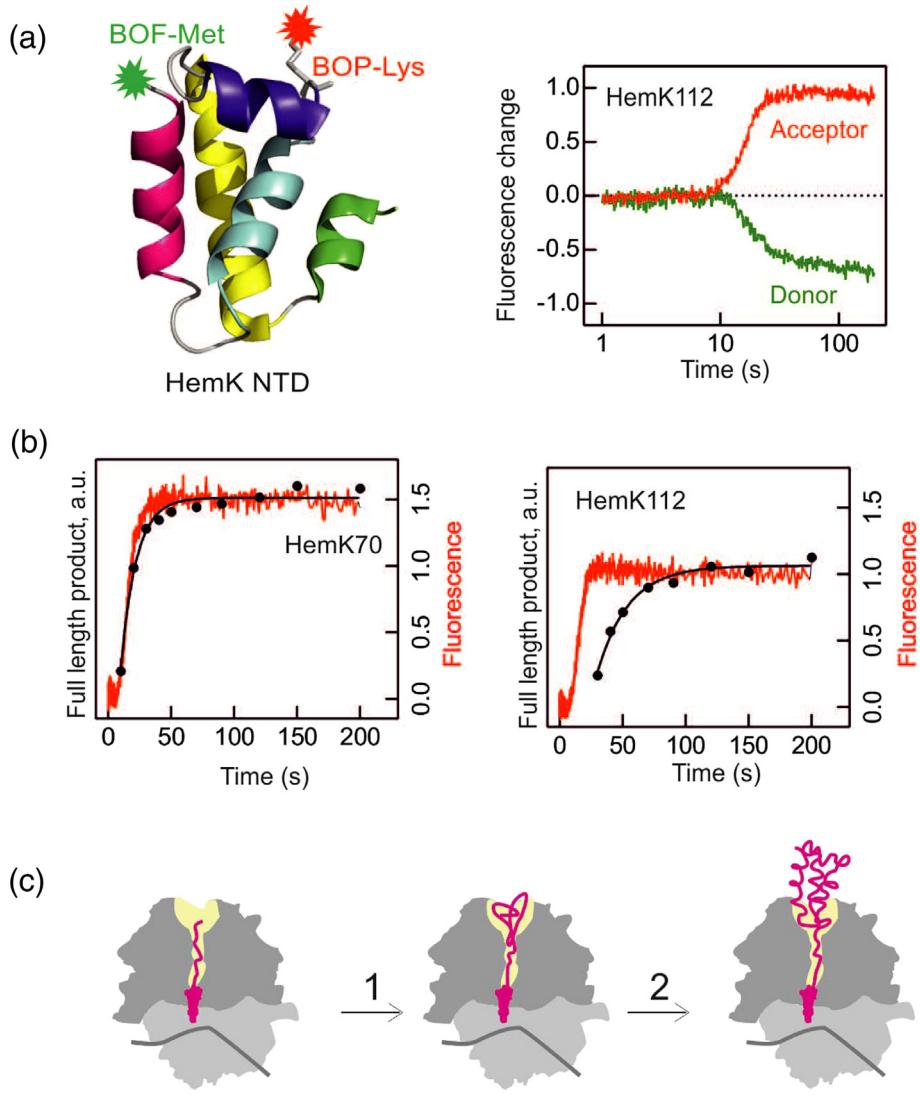

Fig. 3. Co-translational folding monitored in real-time. (a) Left: Crystal structure of the HemK N-terminal domain (residues 1–73 of the full-length protein). Green and red stars indicate the positions (amino acids 1 and 34) of the donor (BOF-Met) and acceptor (BOP-Lys) dyes incorporated into the polypeptide. Right: Changes in BOF-Met (green) and BOP-Lys (red) fluorescence during HemK synthesis. The drop in donor and simultaneous increase in acceptor fluorescence at 10 s indicate energy transfer due to folding of the nascent polypeptide. (b) Time course of acceptor fluorescence and translation product accumulation. For translation templates of varying length, the acceptor fluorescence rapidly increases after 10 s. For a construct containing the N-terminal 70 aa of HemK (HemK70), fluorescence increase and product accumulation occur simultaneously. When the entire N-terminal domain emerges from the ribosome (HemK112), the fluorescence increase precedes the completion of synthesis, demonstrating co-translational folding. The fluorescence signal in HemK70 is higher than that of shorter or longer nascent chains, suggesting that the ribosome stabilizes a compact structure at this nascent chain length. (c) Cartoon of co-translational HemK folding. The peptidyl-tRNA is shown in magenta. A compact structure is stabilized in the distal portion of the ribosome exit tunnel (yellow) that spans the large ribosomal subunit (dark gray) (step 1). After further elongation, the nascent polypeptide rearranges into a native-like structure (from Holtkamp *et al.* [13]).

of the ribosome [10]. Combining constraints from these measurements with molecular dynamics simulations yielded structural ensembles of the filamin domain 5 (FLN5) of the *Dictyostelium discoideum* gelation factor [10]. Characterizing the structure of nascent chains arrested at several defined positions revealed that FLN5 attains its native structure only after at least 47 residues of the following domain (FLN6) have been translated, well after it has emerged from the exit tunnel. In the absence of the ribosome, isolated FLN5 folds even when lacking four C-terminal residues. Ribosome–nascent chain interactions therefore destabilize the native FLN5 structure, effectively inhibiting its folding until a substantial fraction of the following

FLN6 domain has been extruded. When ribosomes are added in trans to unfolded FLN5Δ12 and FLN5Y719E, signal reduction similar to that in corresponding RNCs is observed. Full-length, folding competent FLN5 was not affected under these conditions, supporting the conclusion that the ribosome selectively interacts with non-native nascent chains, which is consistent with results from complementary experiments [9, 65]. The NMR experiments also revealed contacts between the partially synthesized, disordered FLN6 domain and ribosomal RNA and proteins, in particular L23 and L24. These measurements are thus beginning to reveal, in molecular detail, the co-translational interactions that modulate the folding landscape of nascent proteins.

NMR spectroscopy of ribosome-bound nascent chains is technically very demanding, and the approach is still being developed to its full potential. It promises to be extremely powerful because it has the capability to resolve the (equilibrium) structures populated during early co-translational folding at atomic resolution [10], which is not currently possible with other approaches. Simultaneously, such measurements yield information on their thermodynamic stability. Due to the long acquisition times, actively elongating RNCs cannot be analyzed, limiting the approach to equilibrium measurements with stalled nascent chains. Complementing high-resolution structural information from NMR experiments with the temporal resolution attainable by fluorescence measurements or force spectroscopy (see below) has the potential to yield transformative insights into the dynamics of nascent protein folding.

Force spectroscopy

Co-translational folding is intrinsically a non-equilibrium process. Defining its kinetics is therefore crucial. Single-molecule approaches are very powerful tools for dissecting protein folding landscapes, yielding not only information about the structures that are populated along folding pathways, but also about kinetic rates [73, 74]. They resolve the inherent stochasticity of folding and the transient, potentially heterogeneous states populated en route to the native state. Single-molecule force spectroscopy with optical tweezers has unique capabilities for characterizing co-translationally formed states both kinetically and thermodynamically [75]. Mechanical force acts as a denaturant, "tilting" the free energy landscape of a protein and favoring states that are more extended and less structured than the native state [76]. In contrast to chemical denaturants that globally act on all molecules in solution, force is applied locally and can therefore be used to selective perturb the stability of nascent proteins while leaving the ribosome intact. Manipulation of single molecules also circumvents the problem of inter-molecular aggregation, which complicates experiments with unfolded or partially structured states that are inevitably populated during polypeptide elongation.

Optical tweezers have been utilized to study how the ribosome modulates the folding of T4 lysozyme [9] (Fig. 4), a small protein whose folding in isolation has been extensively studied [77]. When the T4 lysozyme polypeptide is kept in close proximity to the ribosome by means of a 41-amino-acid-long C-terminal linker, folding is slowed down more than 100-fold compared to the protein in the absence of the ribosome [9]. Extending the linker to 60 amino acids partially restored folding rates, suggesting that interactions with the ribosome modulate folding. The formation of an obligatory folding intermediate, most likely corresponding to the C-terminal subdomain of the protein, is not affected, suggesting that the ribosome can modulate specific steps along the folding pathway, rather than globally destabilizing any folded structure. Whereas C-terminally truncated T4 lysozyme adopts *bona fide* misfolded states and aggregates in isolation, both misfolding and aggregation are suppressed in ribosome-bound polypeptides [9]. A potential function of the ribosome may thus be to prevent premature folding of an emerging domain until sufficient polypeptide has been synthesized to allow productive folding.

Tuning stabilities and folding rates of nascent polypeptides may be particularly important for larger, multi-domain proteins. Technical challenges have largely precluded a mechanistic understanding of how these complex proteins navigate their intricate folding landscapes. The formation of misfolded, off-pathway species that has been directly observed in single-molecule experiments [11, 78] complicates folding to the native structure. Domain-wise folding during synthesis is a straightforward mechanism of reducing the complexity of the conformational search problem for multi-domain proteins. Co-translational folding is therefore likely to be a key aspect in their biogenesis. Liu *et al.* [11] have recently reported a first step toward dissecting the co-translational folding pathway of the five-domain protein elongation factor G (EF-G). The N-terminal G-domain of EF-G folds into a stable structure both in isolation and on the ribosome. As was observed with T4 lysozyme [9], G-domain folding proceeds more slowly on the ribosome [11]. The kinetic modulation of nascent chain folding by the ribosome, presumably by sequestering part of the polypeptide into interactions with the ribosome surface, might thus be a general aspect of co-translational folding. In the context of a multi-domain protein, such interactions could also serve to limit the entanglement of several unfolded domains, which would otherwise make them prone to inter-domain misfolding.

Summary and outlook

In recent years, biochemical and biophysical experiments have begun to unravel how protein folding is coupled to synthesis. One common thread that emerges from these studies is that folded structures in the nascent polypeptide are destabilized [8, 10, 50]

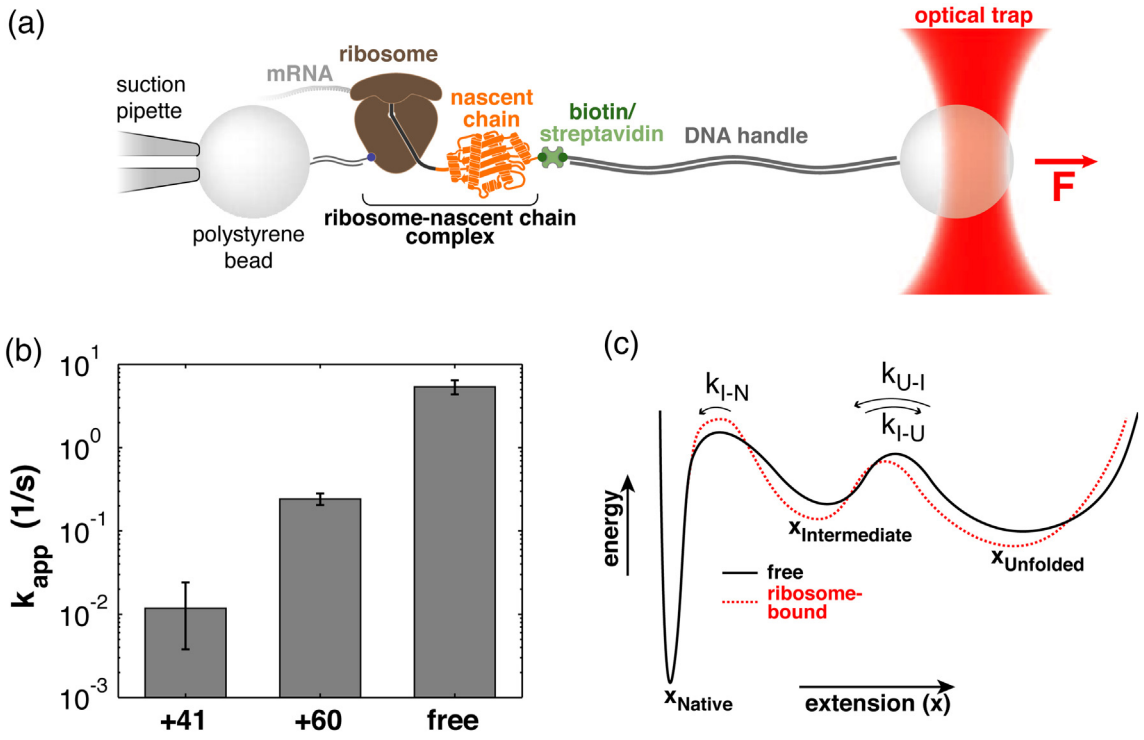

Fig. 4. Folding of T4 lysozyme on the ribosome. (a) Schematic of the molecular assembly for optical tweezers experiments. An RNC is tethered between two polystyrene beads by means of DNA handles attached to the large ribosomal subunit and the N-terminus of the nascent polypeptide. Force is applied to the tethered complex by moving the optical trap relative to the pipette. The T4 lysozyme polypeptide (orange cartoon) is linked to the ribosome by means of a C-terminal Ser/Gly linker (gray). (b) In close proximity to the ribosome (41-aa C-terminal linker, "+41"), the apparent rate of T4 lysozyme folding is reduced more than 100-fold compared to the isolated protein in the absence of the ribosome ("free"). Extending the linker to 60 aa ("+60") partially restores the folding rate, indicating that the effect of the ribosome on nascent chain folding rates strongly depends on distance. (c) Schematic of the folding energy landscape obtained from optical tweezers experiments for free (black) and ribosome-bound (red) T4 lysozyme. The ribosome stabilizes both the unfolded and an intermediate state. The equilibrium between these two states is not affected by the ribosome, but the barrier for the final transition to the native state is increased, explaining the observed deceleration of folding. The ribosome therefore modulates a specific step in the folding pathway (from Kaiser *et al.* [9]).

and that folding rates are reduced [9, 11] on the ribosome. Sequestering nascent chain residues into interactions with the ribosome that would otherwise participate in (mis-)folding can explain these observations. For small proteins, such a mechanism would ensure that structure acquisition of the nascent polypeptide is delayed until sufficient sequence has been synthesized to enter into productive folding. Because protein is highly cooperative, truncated proteins or incomplete domains are unable to attain their native structures [79]. As a consequence, they expose hydrophobic residues that are buried in the folded structure, resulting in a high propensity to aggregate [9]. Nascent chain-binding molecular chaperones are known to prevent misfolding and aggregation by shielding hydrophobic segments that are buried in the folded structure but exposed in unfolded polypeptides [4]. The ribosome itself may bolster this activity through complementary interactions with nascent polypeptides, helping to stabilize them in a folding-competent conformation. This first line of defense could serve to prevent the formation of stable misfolded states that might be difficult to resolve [9].

Most mechanistic folding studies to date have investigated small, single-domain proteins [80]. It is likely that ribosomal interactions are particularly important for the folding of more complex, multi-domain proteins, which have been shown to be prone to inter-domain misfolding [11, 78]. Continuous elongation increases the separation of a given domain in the nascent chain from the ribosome surface, and the effect of the ribosome on its folding decays [8–10]. At this point, interactions with the following domain are established, which might then prevent it from interfering with the folding of the preceding domain. The ribosome may thus help to reduce inter-domain misfolding, effectively increasing the rate of productive folding. Presumably similar interactions of the ribosome with

the nascent chain at different stages of elongation would therefore result in distinct outcomes, either slowing down folding (in proximal, partially synthesized domains) or speeding it up (in more distal, complete domains). Such a mechanism would help to ensure timely domain-wise folding and avoid the extensive accumulation of unfolded polypeptide during translation of multi-domain proteins.

The particular structure of the ribosome exit tunnel is also likely to affect early folding events. "Folding zones" [30] inside the tunnel may either promote or prevent structure formation at specific locations. The stabilization of structures near the distal end of the tunnel [13] suggests that the formation of specific folding intermediates is promoted by the ribosome. Furthermore, a growing number of nascent polypeptide sequences have been demonstrated to regulate translation elongation by adopting structures that enable specific contacts with both ribosomal RNA and proteins. Several well-characterized sequences result in robust and regulated elongation arrest [33]. It is tempting to speculate that more a subtle regulation of ribosome activity through similar interactions may play a role in tuning the rate during protein synthesis more generally. Such regulation could even provide a mechanistic underpinning for feedback regulation of elongation by folding [48].

The modulation of polypeptide elongation rates has been firmly established as an important aspect that determined folding efficiencies and outcomes for natural proteins [12, 13, 25–27, 29, 81] and engineered model systems [28]. A key challenge for mechanistically understanding co-translational folding will be to integrate the effects of elongation rate changes with the modulation of folding by ribosome–nascent chain interactions. Given the thermodynamic and kinetic effects of the ribosome on folding, when exactly during translation does folding begin, and how are folding and elongation tuned to each other? A full description of the molecular processes during co-translational folding will help to interpret bioinformatic analyses [45, 82–84] and may facilitate the rational design of coding sequences for efficient recombinant production [85, 86].

Given that both elongation and folding are inherently stochastic [87], an aspect that needs to be addressed is how robust coupling of the two processes is achieved. Single-molecule experiments are particularly well suited to dissect the mechanisms of this kinetic coupling. With the recent development of an efficient system for real-time measurements of co-translational folding [13], such experiments now appear feasible. We anticipate that single-molecule force and fluorescence spectroscopy will both contribute invaluable insights into co-translational folding in the near future, either separately or, as has recently become possible [88], in combination.

Much progress has been made in the burgeoning field of protein folding on the ribosome, in large part due to the development of methods that now allow direct measurements at high resolution. A quantitative understanding of the states populated during co-translational folding and how they are kinetically connected, that is, a description of the folding energy landscapes and how they evolve during synthesis, will be needed to mechanistically understand *de novo* protein folding. The great diversity of protein sequence properties necessitates further studies to flesh out common themes and principles. Biophysical techniques, in particular single-molecule approaches, are poised to disentangle the complex ways in which nascent chain folding is guided by cellular machinery, and promise to elucidate functions of the ribosome in protein biogenesis beyond synthesis.

Acknowledgments

We would like to thank members of the Kaiser lab for helpful discussion. C.M.K. acknowledges support from the National Institutes of Health (5R01GM121567).

Received 4 May 2018;
Received in revised form 26 June 2018;
Accepted 28 June 2018
Available online 5 July 2018

Keywords:
protein folding;
ribosome;
translation

Abbreviations used:
mRNA, messenger RNA; tRNAs, transfer RNAs; RNCs, ribosome–nascent chain complexes; FRET, fluorescence resonance energy transfer; NMR, nuclear magnetic resonance.

References

[1] C.B. Anfinsen, Principles that govern the folding of protein chains, Science 181 (1973) 223–230.
[2] C.M. Dobson, A. Sali, M. Karplus, Protein folding: a perspective from theory and experiment, Angew. Chem. Int. Ed. 37 (1998) 868–893.
[3] R.J. Ellis, A.P. Minton, Protein aggregation in crowded environments, Biol. Chem. 387 (2006) 485–497.
[4] Y.E. Kim, M.S. Hipp, A. Bracher, M. Hayer-Hartl, F.U. Hartl, Molecular chaperone functions in protein folding and proteostasis, Annu. Rev. Biochem. 82 (2013) 323–355.
[5] G. Kramer, D. Boehringer, N. Ban, B. Bukau, The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins, Nat. Struct. Mol. Biol. 16 (2009) 589–597.

[6] S. Pechmann, F. Willmund, J. Frydman, The ribosome as a hub for protein quality control, Mol. Cell 49 (2013) 411–421.
[7] M. Thommen, W. Holtkamp, M.V. Rodnina, Co-translational protein folding: progress and methods, Curr. Opin. Struct. Biol. 42 (2017) 83–89.
[8] A.J. Samelson, M.K. Jensen, R.A. Soto, J.H. Cate, S. Marqusee, Quantitative determination of ribosome nascent chain stability, Proc. Natl. Acad. Sci. U. S. A. 113 (2016) 13402–13407.
[9] C.M. Kaiser, D.H. Goldman, J.D. Chodera, I. Tinoco Jr., C. Bustamante, The ribosome modulates nascent protein folding, Science 334 (2011) 1723–1727.
[10] L.D. Cabrita, A.M. Cassaignau, H.M. Launay, C.A. Waudby, T. Wlodarski, C. Camilloni, et al., A structural ensemble of a ribosome–nascent chain complex during cotranslational protein folding, Nat. Struct. Mol. Biol. 23 (2016) 278–285.
[11] K. Liu, J.E. Rehfus, E. Mattson, C.M. Kaiser, The ribosome destabilizes native and non-native structures in a nascent multidomain protein, Protein Sci. 26 (2017) 1439–1451.
[12] F. Buhr, S. Jha, M. Thommen, J. Mittelstaet, F. Kutz, H. Schwalbe, et al., Synonymous codons direct cotranslational folding toward different protein conformations, Mol. Cell 61 (2016) 341–351.
[13] W. Holtkamp, G. Kokic, M. Jager, J. Mittelstaet, A.A. Komar, M.V. Rodnina, Cotranslational protein folding on the ribosome monitored in real time, Science 350 (2015) 1104–1107.
[14] H. Himeno, D. Kurita, A. Muto, tmRNA-mediated trans-translation as the major ribosome rescue system in a bacterial cell, Front. Genet. 5 (2014) 66.
[15] V. Heurgue-Hamard, R. Karimi, L. Mora, J. MacDougall, C. Leboeuf, G. Grentzmann, et al., Ribosome release factor RF4 and termination factor RF3 are involved in dissociation of peptidyl-tRNA from the ribosome, EMBO J. 17 (1998) 808–816.
[16] R. Young, H. Bremer, Polypeptide-chain-elongation rate in *Escherichia coli* B/r as a function of growth rate, Biochem. J. 160 (1976) 185–194.
[17] K. Bostrom, M. Wettesten, J. Boren, G. Bondjers, O. Wiklund, S.O. Olofsson, Pulse-chase studies of the synthesis and intracellular transport of apolipoprotein B-100 in Hep G2 cells, J. Biol. Chem. 261 (1986) 13800–13806.
[18] N.T. Ingolia, L.F. Lareau, J.S. Weissman, Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes, Cell 147 (2011) 789–802.
[19] K. Dao Duc, Y.S. Song, The impact of ribosomal interference, codon usage, and exit tunnel interactions on translation elongation rate variation, PLoS Genet. 14 (2018), e1007166.
[20] D.C. Phillips, Hen Egg-White Lysozyme Molecule, Proc. Natl. Acad. Sci. U. S. A. 57 (1967) 484-&.
[21] A.V. Nicola, W. Chen, A. Helenius, Co-translational folding of an alphavirus capsid protein in the cytosol of living cells, Nat. Cell Biol. 1 (1999) 341–345.
[22] J. Frydman, H. Erdjument-Bromage, P. Tempst, F.U. Hartl, Co-translational domain folding as the structural basis for the rapid de novo folding of firefly luciferase, Nat. Struct. Biol. 6 (1999) 697–705.
[23] P.L. Clark, J. King, A newly synthesized, ribosome-bound polypeptide chain adopts conformations dissimilar from early in vitro refolding intermediates, J. Biol. Chem. 276 (2001) 25411–25420.
[24] M.V. Rodnina, The ribosome in action: tuning of translational efficiency and protein folding, Protein Sci. 25 (2016) 1390–1406.
[25] C. Kimchi-Sarfaty, J.M. Oh, I.W. Kim, Z.E. Sauna, A.M. Calcagno, S.V. Ambudkar, et al., A "silent" polymorphism in the MDR1 gene changes substrate specificity, Science 315 (2007) 525–528.
[26] M. Zhou, J. Guo, J. Cha, M. Chae, S. Chen, J.M. Barral, et al., Non-optimal codon usage affects expression, structure and function of clock protein FRQ, Nature 495 (2013) 111–115.
[27] G. Zhang, M. Hubalewska, Z. Ignatova, Transient ribosomal attenuation coordinates protein synthesis and co-translational folding, Nat. Struct. Mol. Biol. 16 (2009) 274–280.
[28] I.M. Sander, J.L. Chaney, P.L. Clark, Expanding Anfinsen's principle: contributions of synonymous codon selection to rational protein design, J. Am. Chem. Soc. 136 (2014) 858–861.
[29] S.J. Kim, J.S. Yoon, H. Shishido, Z. Yang, L.A. Rooney, J.M. Barral, et al., Protein folding. Translational tuning optimizes nascent protein folding in cells, Science 348 (2015) 444–448.
[30] J. Lu, C. Deutsch, Folding zones inside the ribosomal exit tunnel, Nat. Struct. Mol. Biol. 12 (2005) 1123–1129.
[31] C.A. Woolhead, P.J. McCormick, A.E. Johnson, Nascent membrane and secretory proteins differ in FRET-detected folding far inside the ribosome and in their exposure to ribosomal proteins, Cell 116 (2004) 725–736.
[32] O.B. Nilsson, R. Hedman, J. Marino, S. Wickles, L. Bischoff, M. Johansson, et al., Cotranslational protein folding inside the ribosome exit tunnel, Cell Rep. 12 (2015) 1533–1540.
[33] K. Ito, S. Chiba, Arrest peptides: cis-acting modulators of translation, Annu. Rev. Biochem. 82 (2013) 171–202.
[34] D.N. Wilson, S. Arenz, R. Beckmann, Translation regulation via nascent polypeptide-mediated ribosome stalling, Curr. Opin. Struct. Biol. 37 (2016) 123–133.
[35] P. Gupta, B. Liu, D. Klepacki, V. Gupta, K. Schulten, A.S. Mankin, et al., Nascent peptide assists the ribosome in recognizing chemically distinct small molecules, Nat. Chem. Biol. 12 (2016) 153–158.
[36] E. Ishii, S. Chiba, N. Hashimoto, S. Kojima, M. Homma, K. Ito, et al., Nascent chain-monitored remodeling of the Sec machinery for salinity adaptation of marine bacteria, Proc. Natl. Acad. Sci. U. S. A. 112 (2015) E5513–E5522.
[37] T. Su, J. Cheng, D. Sohmen, R. Hedman, O. Berninghausen, G. von Heijne, et al., The force-sensing peptide VemP employs extreme compaction and secondary structure formation to induce ribosomal stalling, elife 6 (2017).
[38] X. Agirrezabala, E. Samatova, M. Klimova, M. Zamora, D. Gil-Carton, M.V. Rodnina, et al., Ribosome rearrangements at the onset of translational bypassing, Sci. Adv. 3 (2017), e1700147.
[39] C.A. Woolhead, A.E. Johnson, H.D. Bernstein, Translation arrest requires two-way communication between a nascent polypeptide and the ribosome, Mol. Cell 22 (2006) 587–598.
[40] J. Mittal, R.B. Best, Thermodynamics and kinetics of protein folding under confinement, Proc. Natl. Acad. Sci. U. S. A. 105 (2008) 20233–20238.
[41] H. Nakatogawa, K. Ito, The ribosomal exit tunnel functions as a discriminating gate, Cell 108 (2002) 629–636.
[42] J. Mittelstaet, A.L. Konevega, M.V. Rodnina, A kinetic safety gate controlling the delivery of unnatural amino acids to the ribosome, J. Am. Chem. Soc. 135 (2013) 17031–17038.
[43] M.V. Rodnina, W. Wintermeyer, Protein Elongation, Co-translational folding and targeting, J. Mol. Biol. 428 (2016) 2165–2185.
[44] G.N. Jacobson, P.L. Clark, Quality over quantity: optimizing co-translational protein folding with non-'optimal' synonymous codons, Curr. Opin. Struct. Biol. 38 (2016) 102–110.
[45] A.A. Komar, The Yin and Yang of codon usage, Hum. Mol. Genet. 25 (2016) R77–R85.

[46] F. Trovato, E.P. O'Brien, Insights into cotranslational nascent protein behavior from computer simulations, Annu. Rev. Biophys. 45 (2016) 345–369.
[47] E. Oh, A.H. Becker, A. Sandikci, D. Huber, R. Chaba, F. Gloge, et al., Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo, Cell 147 (2011) 1295–1308.
[48] D.H. Goldman, C.M. Kaiser, A. Milin, M. Righini, I. Tinoco Jr., C. Bustamante, Ribosome. Mechanical force releases nascent chain-mediated ribosome arrest in vitro and in vivo, Science 348 (2015) 457–460.
[49] K. Doring, N. Ahmed, T. Riemer, H.G. Suresh, Y. Vainshtein, M. Habich, et al., Profiling Ssb-nascent chain interactions reveals principles of Hsp70-assisted folding, Cell 170 (2017) 298–311 (e20).
[50] A. Hoffmann, A.H. Becker, B. Zachmann-Brand, E. Deuerling, B. Bukau, G. Kramer, Concerted action of the ribosome and the associated chaperone trigger factor confines nascent polypeptide folding, Mol. Cell 48 (2012) 63–74.
[51] C. Park, S. Marqusee, Pulse proteolysis: a simple method for quantitative determination of protein stability and ligand binding, Nat. Methods 2 (2005) 207–212.
[52] C. Schaffitzel, N. Ban, Generation of ribosome nascent chain complexes for structural and functional studies, J. Struct. Biol. 158 (2007) 463–471.
[53] L.D. Cabrita, S.T. Hsu, H. Launay, C.M. Dobson, J. Christodoulou, Probing ribosome–nascent chain complexes produced in vivo by NMR spectroscopy, Proc. Natl. Acad. Sci. U. S. A. 106 (2009) 22239–22244.
[54] M.E. Butkus, L.B. Prundeanu, D.B. Oliver, Translocon "pulling" of nascent SecM controls the duration of its translational pause and secretion-responsive secA regulation, J. Bacteriol. 185 (2003) 6719–6722.
[55] N. Ismail, R. Hedman, N. Schiller, G. von Heijne, A biphasic pulling force acts on transmembrane helices during translocon-mediated membrane integration, Nat. Struct. Mol. Biol. 19 (2012) 1018–1022.
[56] O.B. Nilsson, A. Muller-Lucks, G. Kramer, B. Bukau, G. von Heijne, Trigger factor reduces the force exerted on the nascent chain by a cotranslationally folding protein, J. Mol. Biol. 428 (2016) 1356–1364.
[57] B. Kuhlman, G. Dantas, G.C. Ireton, G. Varani, B.L. Stoddard, D. Baker, Design of a novel globular protein fold with atomic-level accuracy, Science 302 (2003) 1364–1368.
[58] F. Cymer, G. von Heijne, Cotranslational folding of membrane proteins probed by arrest-peptide-mediated force measurements, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) 14640–14645.
[59] J. Marino, G. von Heijne, R. Beckmann, Small protein domains fold inside the ribosome exit tunnel, FEBS Lett. 590 (2016) 655–660.
[60] J.A. Farias-Rico, S.K. Goetz, J. Marino, G. von Heijne, Mutational analysis of protein folding inside the ribosome exit tunnel, FEBS Lett. 591 (2017) 155–163.
[61] O.B. Nilsson, A.A. Nickson, J.J. Hollins, S. Wickles, A. Steward, R. Beckmann, et al., Cotranslational folding of spectrin domains via partially structured states, Nat. Struct. Mol. Biol. 24 (2017) 221–225.
[62] K.S. Crowley, S. Liao, V.E. Worrell, G.D. Reinhart, A.E. Johnson, Secretory proteins move through the endoplasmic reticulum membrane via an aqueous, gated pore, Cell 78 (1994) 461–471.
[63] J.P. Ellis, C.K. Bakke, R.N. Kirchdoerfer, L.M. Jungbauer, S. Cavagnero, Chain dynamics of nascent polypeptides emerging from the ribosome, ACS Chem. Biol. 3 (2008) 555–566.
[64] J.P. Ellis, P.H. Culviner, S. Cavagnero, Confined dynamics of a ribosome-bound nascent globin: cone angle analysis of fluorescence depolarization decays in the presence of two local motions, Protein Sci. 18 (2009) 2003–2015.
[65] A.M. Knight, P.H. Culviner, N. Kurt-Yilmaz, T. Zou, S.B. Ozkan, S. Cavagnero, Electrostatic effect of the ribosomal surface on nascent polypeptide dynamics, ACS Chem. Biol. 8 (2013) 1195–1204.
[66] A. Khushoo, Z. Yang, A.E. Johnson, W.R. Skach, Ligand-driven vectorial folding of ribosome-bound human CFTR NBD1, Mol. Cell 41 (2011) 682–692.
[67] J.A. Houwman, E. Andre, A.H. Westphal, W.J. van Berkel, C. P. van Mierlo, The ribosome restrains molten globule formation in stalled nascent flavodoxin, J. Biol. Chem. 291 (2016) 25911–25920.
[68] A.N. Fedorov, T.O. Baldwin, Contribution of cotranslational folding to the rate of formation of native protein structure, Proc. Natl. Acad. Sci. U. S. A. 92 (1995) 1227–1231.
[69] Y.W. Shieh, P. Minguez, P. Bork, J.J. Auburger, D.L. Guilbride, G. Kramer, et al., Operon structure and cotranslational subunit association direct protein assembly in bacteria, Science 350 (2015) 678–680.
[70] E. Mercier, M.V. Rodnina, Co-translational folding trajectory of the HemK helical domain, Biochemistry 57 (2018) 3460–3464.
[71] S.T. Hsu, P. Fucini, L.D. Cabrita, H. Launay, C.M. Dobson, J. Christodoulou, Structure and dynamics of a ribosome-bound nascent chain by NMR spectroscopy, Proc. Natl. Acad. Sci. U. S. A. 104 (2007) 16516–16521.
[72] V. Tugarinov, L.E. Kay, Ile, Leu, and Val methyl assignments of the 723-residue malate synthase G using a new labeling strategy and novel NMR methods, J. Am. Chem. Soc. 125 (2003) 13868–13878.
[73] M.T. Woodside, S.M. Block, Reconstructing folding energy landscapes by single-molecule force spectroscopy, Annu. Rev. Biophys. 43 (2014) 19–39.
[74] B. Schuler, H. Hofmann, Single-molecule spectroscopy of protein folding dynamics—expanding scope and timescales, Curr. Opin. Struct. Biol. 23 (2013) 36–47.
[75] G. Zoldak, M. Rief, Force as a single molecule probe of multidimensional protein energy landscapes, Curr. Opin. Struct. Biol. 23 (2013) 48–57.
[76] C. Bustamante, Y.R. Chemla, N.R. Forde, D. Izhaky, Mechanical processes in biochemistry, Annu. Rev. Biochem. 73 (2004) 705–748.
[77] W.A. Baase, L. Liu, D.E. Tronrud, B.W. Matthews, Lessons from the lysozyme of phage T4, Protein Sci. 19 (2010) 631–641.
[78] M. Jahn, J. Buchner, T. Hugel, M. Rief, Folding and assembly of the large molecular machine Hsp90 studied in single-molecule experiments, Proc. Natl. Acad. Sci. U. S. A. 113 (2016) 1232–1237.
[79] H. Taniuchi, C.B. Anfinsen, An experimental approach to the study of the folding of staphylococcal nuclease, J. Biol. Chem. 244 (1969) 3864–3875.
[80] E. Braselmann, J.L. Chaney, P.L. Clark, Folding the proteome, Trends Biochem. Sci. 38 (2013) 337–344.
[81] P.S. Spencer, E. Siller, J.F. Anderson, J.M. Barral, Silent substitutions predictably alter translation elongation rates and protein folding efficiencies, J. Mol. Biol. 422 (2012) 328–335.
[82] S. Pechmann, J. Frydman, Evolutionary conservation of codon optimality reveals hidden signatures of cotranslational folding, Nat. Struct. Mol. Biol. 20 (2013) 237–243.

[83] W.M. Jacobs, E.I. Shakhnovich, Evidence of evolutionary selection for cotranslational folding, Proc. Natl. Acad. Sci. U. S. A. 114 (2017) 11434–11439.
[84] J.L. Chaney, A. Steele, R. Carmichael, A. Rodriguez, A.T. Specht, K. Ngo, et al., Widespread position-specific conservation of synonymous rare codons within coding sequences, PLoS Comput. Biol. 13 (2017), e1005531.
[85] E. Angov, C.J. Hillier, R.L. Kincaid, J.A. Lyon, Heterologous protein expression is enhanced by harmonizing the codon usage frequencies of the target gene with those of the expression host, PLoS One 3 (2008), e2189.
[86] A. Rodriguez, G. Wright, S. Emrich, P.L. Clark, %MinMax: a versatile tool for calculating and comparing synonymous codon usage and its impact on protein folding, Protein Sci. 27 (2018) 356–362.
[87] C.M. Kaiser, I. Tinoco Jr., Probing the mechanisms of translation with force, Chem. Rev. 114 (2014) 3266–3280.
[88] M.J. Comstock, T. Ha, Y.R. Chemla, Ultrahigh-resolution optical trap with single-fluorophore sensitivity, Nat. Methods 8 (2011) 335–340.