Annual Review of Biochemistry

How Does the Ribosome Fold the Proteome?

Anaïs M.E. Cassaignau, Lisa D. Cabrita, and John Christodoulou

Institute of Structural and Molecular Biology, University College London and Birkbeck College, London WC1E 7HX, United Kingdom; email: anais.cassaignau.09@ucl.ac.uk, l.cabrita@ucl.ac.uk, j.christodoulou@ucl.ac.uk

Annu. Rev. Biochem. 2020. 89:389–415

The *Annual Review of Biochemistry* is online at biochem.annualreviews.org

https://doi.org/10.1146/annurev-biochem-062917-012226

Keywords

co-translational folding, protein synthesis, ribosome-bound nascent chain, protein folding, structural biology, NMR spectroscopy, protein misfolding

Abstract

Folding of polypeptides begins during their synthesis on ribosomes. This process has evolved as a means for the cell to maintain proteostasis, by mitigating the risk of protein misfolding and aggregation. The capacity to now depict this cellular feat at increasingly higher resolution is providing insight into the mechanistic determinants that promote successful folding. Emerging from these studies is the intimate interplay between protein translation and folding, and within this the ribosome particle is the key player. Its unique structural properties provide a specialized scaffold against which nascent polypeptides can begin to form structure in a highly coordinated, co-translational manner. Here, we examine how, as a macromolecular machine, the ribosome modulates the intrinsic dynamic properties of emerging nascent polypeptide chains and guides them toward their biologically active structures.

Contents

1. THE RIBOSOME COORDINATES A MYRIAD OF CO-TRANSLATIONAL EVENTS

A central process within the cell is the folding and assembly of proteins into their biologically active structures. The cell continuously balances the concentration of ~20,000 structurally and functionally diverse proteins, which collectively underpin nearly every cellular process. Within this, the cell's macromolecular machine, the ribosome, is solely dedicated to successfully producing polypeptide chains and facilitating their folding—this is key to all biological activity within the cell.

As established from a range of seminal biochemical studies (reviewed in 1), the relatively fast timescales of protein-folding events (μs–s timescale) and comparatively slower translation rates (4–20 amino acids/s in bacteria) allow both of these processes to occur concurrently for many nascent polypeptides. In silico predictions estimate that approximately one-third of cytosolic *Escherichia coli* proteins fold co-translationally (2). Within eukaryotic systems, co-translational folding is an even more important feature, as translation rates associated with eukaryotes are comparatively slower (1 amino acid/s), and the eukaryotic proteome is generally composed of larger proteins (average size ~500 amino acids) with complex topologies. It is likely that co-translational folding has facilitated the evolution of multidomain proteins (70% of the proteome) (3), enabling each domain to fold sequentially as it emerges from the ribosomal exit tunnel.

At a molecular level, protein folding involves the collapse of solvent-exposed hydrophobic residues to form structure, and as such, co-translational folding is one cellular mechanism by which aberrant interactions between amino acids can be minimized at the earliest stages of nascent chain biosynthesis. This is ultimately a strategy to maintain proteostasis and mitigate the risks of human misfolding diseases (4). The cell has evolved a myriad of complementary processes designed to optimize the temporal and spatial regulation of protein biosynthesis. Not least of these is the pseudohelical arrangement of ribosomes in polyribosome complexes on a single mRNA, which minimizes internascent chain contacts that may lead to unfavorable interactions (5). The cell's extensive chaperone networks [reviewed elsewhere (6, 7)] are also tightly coupled to protein biosynthesis, as a means of assisting substrates during co- and posttranslational stages of folding,

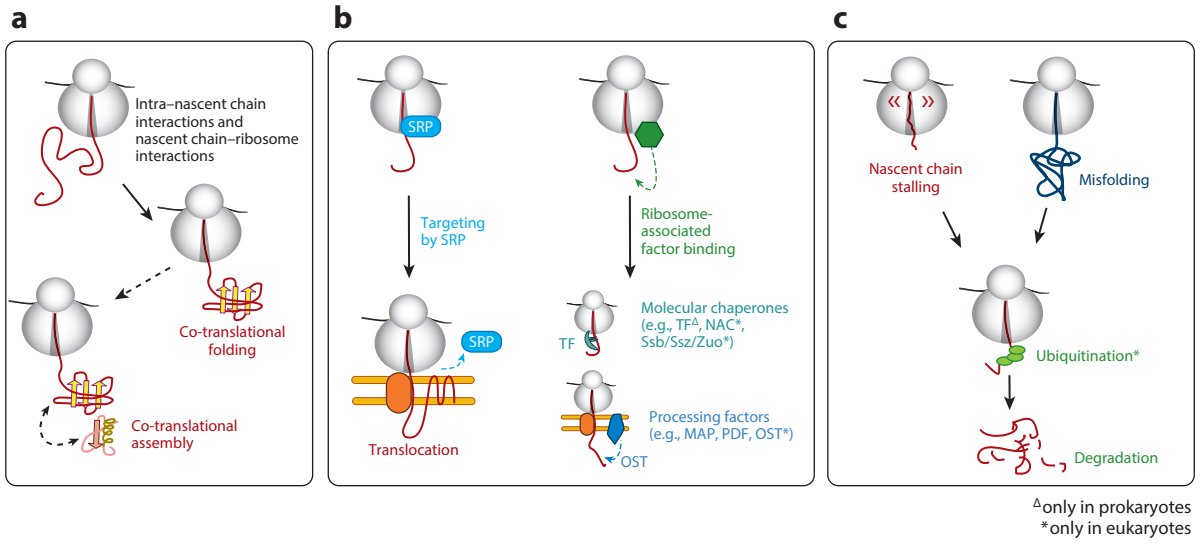

Figure 1

The co-translational processes associated with nascent chains during protein biosynthesis on the ribosome: (*a*) folding and complex assembly, (*b*) translocation, chaperoning, and processing, and (*c*) stalling, misfolding, and degradation. The ribosome is the platform for choreographing the interplay with ribosome-associated proteins that assist in these concurrent and competing processes. Abbreviations: MAP, methionine aminopeptidase; NAC, nascent polypeptide-associated complex; OST, oligosaccharyl-transferase; PDF, peptidyl deformylase; SRP, signal recognition protein; TF, trigger factor.

in an apparently seamless process. Collectively, these processes are advantageous for the cell, as shown in yeast where the failure at the initial stages of folding results in ∼1–6% of all nascent chains (8) [and in 12–15% of nascent proteins in humans (9)] being co-translationally ubiquitinated on the ribosome itself and subsequently degraded. Therefore, the absolute necessity for the cell to meet the continuous demand to regulate protein biosynthesis, and make a return for the highly costly process of translation (10) [∼75% of the total cellular energy budget (11)], rests on a crucial relationship among the ribosome particle, the process of polypeptide synthesis, and the efficiency in forming a biologically active structure.

The formation of structure co-translationally occurs within a crowded cellular environment, against a backdrop of a range of other co-translational processes occurring on the ribosome (**Figure 1**), each of which competes for the attention of the nascent chain to achieve obligatory processing (e.g., methionine removal) and modifications (e.g., N-terminal acetylation, glycosylation) and to direct its intracellular localization (e.g., to membranes) or extracellular export. The timing of these processes is a key strategy of the cell to optimize the efficient regulation of protein production. In *E. coli*, the interactions of the emerging nascent chain with the early engaging ribosome-associated auxiliary proteins (nascent chain lengths <40 residues), peptidyl deformylase (PDF), methionine aminopeptidase (MAP), and signal recognition protein (SRP), are choreographed around those with the later-engaging molecular chaperone trigger factor (TF) (nascent chain lengths >50 residues) to facilitate folding. Such a carefully orchestrated series of sequential events is dictated by the extent of nascent chain emergence at any given time and the concomitant increase in relative affinities that these auxiliary proteins have for the ribosome (12, 13). A similar cohort of auxiliary proteins and processes exist in eukaryotes: MAP and SRP act alongside N-terminal acetylases (NATs) and oligosaccharyl-transferases (within the rough endoplasmic reticulum), and the mechanisms for these are beginning to emerge (6). Remarkably, these

TF: trigger factor

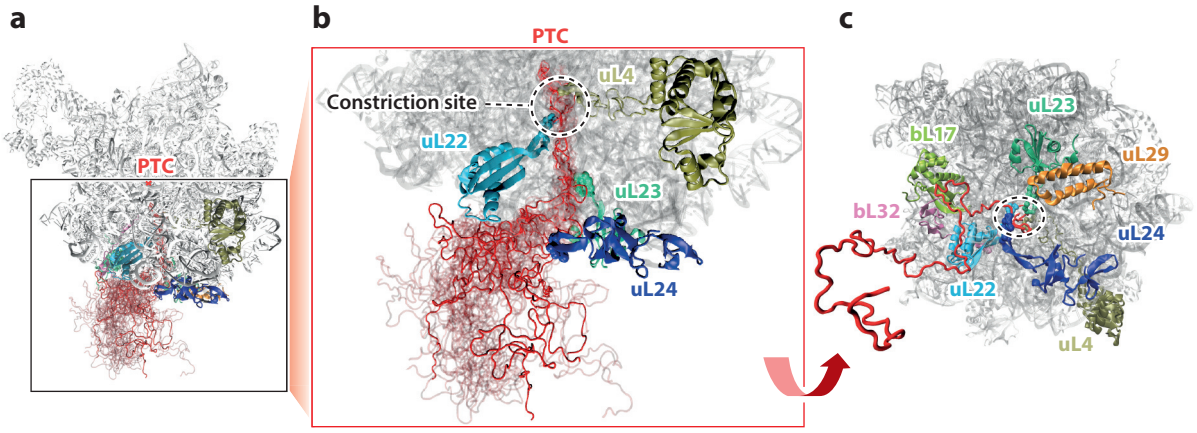

Figure 2

The emergence of a nascent chain from the ribosome tunnel. (*a*) A structural ensemble of a disordered FLN5-variant nascent chain on the *Escherichia coli* ribosome. The MD trajectory was kindly provided by T. Wlodarski (unpublished data). (*b*) This enhanced region, selected in panel *a*, shows how a nascent chain traverses the tunnel from the PTC, through the constriction site formed by the loops of uL4 and uL22 (*circled*) and past the loops of uL23 and uL24 within the vestibule. The nascent chain mobility increases sharply on reaching the exit vestibule. (*c*) View from the top of the ribosome with the exit tunnel circled. The r-proteins in the proximity of the nascent chain are shown. Abbreviations: MD, molecular dynamics; PTC, peptidyltransferase center; r-proteins, ribosomal proteins.

ribosome-associated partners transiently interact with the ribosome at all times and seamlessly engage at a handful of key ribosomal proteins (r-proteins) (e.g., uL23, uL29) that crown the ribosome's exit port when they encounter a region of interest within a substrate nascent chain (**Figure 2**).

Within these processes, the nascent chain is not a passive entity. Its emerging sequence and energetic desire to fold result in a continuous flux that governs the types of contacts it can make with ribosome-associated factors. Exposed hydrophobic and aromatic nascent chain residues (which typically become buried upon nascent chain folding) (14) as well as positively charged residues cause recruitment of the bacterial chaperone, TF (15). Through its holdase action, TF shields the nascent chain from forming unproductive interactions. Other chaperones such as Hsp70 (DnaK in prokaryotes) and the eukaryotic chaperonin TriC are similarly invoked in a responsive manner to target the folding status of the nascent polypeptide, which elicits a hierarchical approach for their recruitment (16). Similarly, SRP predominantly recognizes exposed hydrophobic α-helical transmembrane domains (17), although intriguingly its engagement may be solicited by nascent chains as short as 20 amino acids, still confined within the tunnel (18). Moreover, the folding status of the nascent chain itself appears to be recognized by co-translational quality control mechanisms [as opposed to ribosome-associated quality control mechanisms that sense the state of the ribosome (19)]; these specifically ubiquitinate aberrantly folded nascent chains, resulting in their degradation by the proteasome (8, 20, 21).

The incipient structure of the nascent chain as well as the nature of its exposed sequence (especially in the case of unstructured nascent chain segments) are the determinants for the successful navigation of the array of downstream events ranging from co-translational complex assembly (22–24) to interactions with molecular chaperones and processing factors (6). Immediately upon its emergence from the ribosome, the nascent chain is able to sample conformational space not only to fold but also to interact and bind with potential partners. The ribosome is thus the critical molecular machine that exists at the interface between an organism's genotype and its phenotype,

with a fundamental role in setting up each protein toward its biologically active state according to the cell's needs, by promoting efficient processing and folding of the nascent chain.

In recent years, significant progress has been made toward providing a detailed molecular understanding of the processes by which nascent polypeptides can fold co-translationally. This field of study, deemed intractable just over a decade ago [see Stu Borman's commentary (25)], is instead now offering a detailed structural, dynamic, and mechanistic understanding of this fundamental process. These studies provide a much-needed missing piece in understanding the protein folding problem as highlighted by Linus Pauling and colleagues (26) and in which the mechanistic details of the folding of isolated proteins have been teased out of experimental and computational studies carried out over the last 50 years. Capturing the multiple states that a nascent chain samples during its translation, however, presents challenges—the nascent polypeptide accounts for <0.1% of the mass of a ribosome-bound nascent chain (RNC) complex, the result of which typically limits experimental sensitivity and impedes the design of experiments that exclusively observe the nascent chain. Homogenous preparations of uniformly, translationally arrested nascent chains on ribosomes (27) using naturally occurring stalling sequences (28) or nonstop mRNA-mediated stalling (29, 30) now permit long-lived biosynthetic snapshots to be produced for such investigations. Coinciding with the very significant leaps in structural biology, including the resolution improvements in cryo–electron microscopy (cryo-EM) (31), the boosts in the sensitivity of NMR spectroscopy (32), as well as the expansion in computational power and methods, our capacity to describe RNCs has undergone substantial strides. These combined strategies have shifted the once phenomenological descriptions of co-translational folding of nascent chains toward increasingly structure-led molecular models of this process at high resolution (33–35). Crucially, these observations are also complemented by exquisite quantitative biochemical and biophysical measurements (36–39), which offer unique mechanistic details, and some are able to report events on a biological timescale (40).

RNC: ribosome-bound nascent chain

PTC: peptidyltransferase center

2. WHEN DOES FOLDING BEGIN?

2.1. Structure Formation on the Ribosome Is Strongly Modulated by Ribosome Sterics

A contemporary question that continues to challenge the field is: When and where does protein folding begin on the ribosome? The vectorially emerging nascent polypeptide chain on its parent ribosome particle faces unique physical constraints on how and when it folds. These constraints are imposed by the features of the ribosome itself.

2.1.1. The geometry of the ribosome tunnel and its effect on the nascent chain. During biosynthesis, amino acids sequentially combine at the peptidyltransferase center (PTC), and the nascent chain emerges vectorially from the 80–100-Å ribosomal exit tunnel (**Figure 2**), which occludes between ~25 amino acids for a fully extended polypeptide, and up to ~40 amino acids for strongly compacted conformations. The width of the tunnel is ~15 Å in bacteria, with a narrow constriction site (<10-Å width) formed by the protruding loops from r-proteins uL4 and uL22 (41) (**Figure 2*b***). The dynamics of the nascent polypeptide chain thus change dramatically as it transitions from the confines of the prohibitively narrow ribosome tunnel, via a more accommodating wider vestibule region (~20-Å width), and ultimately toward the cytosol. Upon exposure to the cytosol, the emerged N-terminal segments of the nascent chain become highly dynamic and can begin to sample conformational space more freely while still tethered to the ribosome (**Figure 2*a*,*b***). By contrast, in the case of membrane-associated polypeptides, nascent chains are comparatively limited in their conformational freedom upon exit from the ribosome,

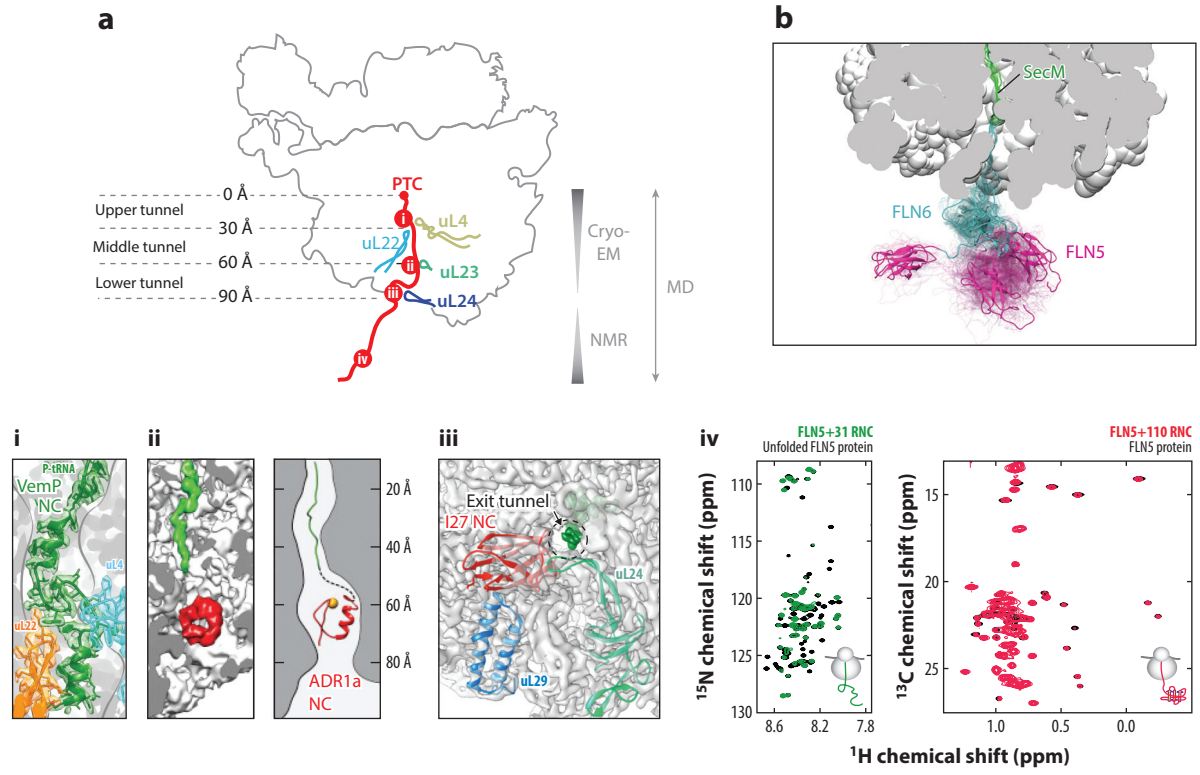

Figure 3

High-resolution structures of co-translationally folding NCs using cryo-EM, MD, and NMR spectroscopy. (*a*) The ribosomal exit tunnel is shown, delineated into upper (PTC to uL4-uL22 constriction site), middle (uL4-uL22 constriction site to uL23 loop), and lower (uL23 loop to uL24 loop) regions. The locations of experimentally determined NC structures, as further described in subpanels *i–iv*, are shown. (*i*) The VemP stalling motif forms a compact, helical structure above the uL4-uL22 constriction site as shown by cryo-EM (local map resolution ≤3 Å). Panel *a*, subpanel *i* adapted with permission from Reference 34. (*ii*) The ADR1a zinc finger NC forms native tertiary structure within the vestibule (local cryo-EM map resolution ≤5 Å). Panel *a*, subpanel *ii* adapted with permission from Reference 35. (*iii*) The native structure formed by the titin I27 NC beyond the ribosomal exit (local cryo-EM map resolution ≤8 Å). Panel *a*, subpanel *iii* adapted with permission from Reference 66. (*iv*) NMR spectroscopy of immunoglobulin-like FLN5 RNCs: $^{1}H,^{15}N$ correlation NMR spectrum of a uniformly ^{15}N-labeled disordered FLN5+31 RNC (*green*). $^{1}H,^{13}C$ correlation spectrum of natively folded FLN5 within an FLN5+110 RNC, with the Ile, Leu, and Val methyl groups of the NC selectively $^{13}CH_3$-labeled against a perdeuterated background (*red*) (A.M.E. Cassaignau, unpublished data). (*b*) NMR-restrained structural ensemble of a pair of Ig-like domains, FLN5–6 from the gelation factor filamin during biosynthesis, showing natively folded FLN5 (*pink*) and the compact ensemble of disordered FLN6 (*cyan*). Panel *b* adapted with permission from Reference 33. Abbreviations: cryo-EM, cryo–electron microscopy; MD, molecular dynamics; NC, nascent chain; PTC, peptidyltransferase center; RNC, ribosome-bound nascent chain; VemP, Vibrio export monitoring polypeptide.

as they rapidly engage with the translocation machinery (e.g., SRP/SecA/translocon) and are then immediately integrated into a membrane. The variation in physical environments experienced by emerging nascent chains, therefore, places limits on any single technique's ability to describe the complete structure of the RNC with sufficient resolution (see Section 2.2). This is most apparent within cryo-EM analyses, in which the nascent chain residues in the wider vestibule region give rise to weaker electron density compared with residues closer to the PTC area (**Figure 3*a***), because the chain is able to gain more motional freedom (**Figure 2*b***). Integrating experimental and computational approaches (see Section 2.2) is now viewed as providing an essential basis for

molecular descriptions of nascent chain structure and folding mechanisms from the PTC to the cytosol.

2.1.2. Structure formation inside the ribosomal exit tunnel. The narrow geometry of the ribosome exit tunnel (**Figure 2*b***) severely restricts extensive intranascent chain contacts from being formed (42). However, compaction of the nascent chain and persistent structure formation have been described: Helix formation, predicted to be stabilized entropically by the ribosome (43), has been shown experimentally in defined zones of the tunnel using cysteine-mapping PEGylation studies (44), fluorescence resonance energy transfer (FRET) observations of integral membrane proteins (45), and more recently, cryo-EM reconstructions of RNCs (46). In particular, the ability to form persistent helices may be specific to SRP-associated ribosomes translating transmembrane helices, which can form rather early during their biosynthesis 20–30 Å from the PTC; these may also be further stabilized by nonpolar interactions with elements of the ribosomal tunnel and, later, the membrane (47). This ability to form rudimentary structures inside the tunnel may in select cases offer a functional advantage—for example, SRP recognition is facilitated by the presence of a transmembrane helix (47). Conversely, hydrophilic, soluble helices of cytosolic proteins are more likely formed beyond the tunnel (47).

The formation of tertiary structure within the tunnel is limited but appears to be supported closer to the wide vestibule region, located ~80 Å from the PTC at the end of the tunnel (see Section 2.1.3). Initially predicted computationally for a number of small proteins (48), this has since been shown experimentally in the formation of a β-hairpin motif via PEGylation (49) and for a 3-kDa zinc finger of 29 residues that was observed by cryo-EM to be in a folded conformation (35) (**Figure 3*a***).

2.1.3. The onset of co-translational folding. For the vast majority of proteins, persistent tertiary structure forms only once a sufficient number of amino acids has emerged beyond the exit vestibule, after which the space available to the polypeptide abruptly expands. Seminal early biochemical (3, 37) and biophysical (36) studies of multidomain proteins showed that the folding of nascent chains can occur domain-by-domain; these findings also provided compelling evidence for the ribosome directly influencing nascent chain folding, which begins at a variable minimum linker length (typically 30–40 residues) tethering each nascent chain domain to the ribosome.

Experimentally, the readout of the force generated by a polypeptide as a function of nascent chain length has been recently exploited as an effective means of measuring the onset of folding. Optical tweezer studies of the 90-residue protein Top7 showed that the folding of this small domain generates an ~12-pN pulling force that appears sufficient to cause release of translationally arrested nascent chains (50). The tether (~38 residues) at which significant force starts being generated closely correlates with the length at which the 128-residue ileal binding protein (ILBP) nascent chain becomes resistant to limited proteolysis (51). Similarly, the folding onset of the 105-residue FLN5 nascent chain when linked by 42 residues to the PTC, as directly observed by NMR spectroscopy (33), is consistent with that measured using force assays (35–47 residues) (52). Broadly, the extent of emergence required for nascent chain structure to form is found to be correlated with protein size; the larger the protein, the greater the required extrusion before native structure can form, with small proteins (~75 residues in length) folding close to the exit vestibule (51).

2.1.4. Emerging co-translational folding mechanisms. The increasing body of experimental data is allowing an examination of the nature of folding mechanisms on the ribosome. The force assays described above appear to be able to monitor complex folding phenomena, well exemplified

in the study of the 217-residue pentarepeat β-helix protein PENT for which a multipeak force profile was attributed to the detection of putative folding intermediates (53). These force assays are also highly sensitive to additional processes—partial structure formation (53a), membrane insertion (54, 55), the binding of interaction partners (24), as well as interactions within the tunnel and the external ribosome surface. Such a sensitivity to the environment may constitute a challenge for these rich assays to draw generalized mechanistic interpretations, but when combined with orthogonal data [as described in Section 2.1.3], the continued evolution of this strategy shows great promise, including to further deconvolute complex nascent chain behavior and understand length-dependent structure formation. In the case of P22 tailspike nascent chains, conformational antibodies (56) were used to show the sampling of significant extents of native-like structure at early stages of emergence from the ribosome. A series of collapsed, intermediate states were identified by both FRET and photoinduced electron transfer (40) in a study of an all-α-helical HemK nascent chain, findings that were further supported by a multipeak force profile (52). Other examples feature proteins that only sample detectable structure once well emerged from the ribosome, such as T4 lysozyme (58). Similarly, the all-α-helical EF-hand protein calerythrin (59) and the β-sandwich immunoglobulin domain FLN5 (33) both show a delay between the point at which they emerge and the point at which they begin to sample native structure (see Section 2.2.2). The effect of the ribosome in shaping these conformational preferences appears in many cases to be critical to the success of protein folding (30, 58). To help delineate these mechanisms, the emerging body of structural data (see Section 2.2) is shaping our understanding of the free-energy landscape of nascent chains, which is in continuous flux during their elongation on the ribosome. By building on the established kinetic and thermodynamic principles of protein folding and tailoring these to the study of the nascent polypeptide bound to its parent ribosome (discussed in Section 3.1), the progression from a view of protein folding based on isolated polypeptides to that of a more complex co-translational folding landscape is on the horizon.

2.2. Toward High-Resolution Structural Ensembles of Nascent Chains to Describe the Influence of the Ribosome: Integrating Structural Methods

The X-ray crystallographic (60, 61) and cryo-EM (62) structures of the ribosome particle have provided a comprehensive backdrop to current efforts to describe the coupled processes of nascent chain emergence, ribosome interactions, folding, and interactions with cellular factors. As the nascent polypeptide emerges from the exit tunnel, NMR spectroscopy is able to uniquely probe incipient nascent chain structure and dynamics. As observed for residues of intrinsically disordered α-synuclein, the first hint of sufficiently increased local mobility [typically a local rotational correlation time of <50 ns is required for solution-state NMR observability (63)] coincides with the first residue of the nascent chain (Asp135, 28 residues from the peptidyl transferase center) that is sufficiently flexible to be visible by NMR (64). The emerging application of solid-state NMR to nascent chains will allow less-mobile nascent chain regions to be probed, in particular using dynamic nuclear polarization (DNP) magic-angle spinning (MAS) NMR, as has been demonstrated for a rigid nascent signal peptide within the exit tunnel (65).

2.2.1. Cryo–electron microscopy in co-translational folding. In recent years, significant attention has been invested into the use of cryo-EM (62) to study the ribosome during translation, including nascent chain folding (35, 57, 66), chaperone binding (67, 68), and nascent chain translocation (69–73). These high-resolution snapshots of the nascent chain during biosynthesis capture its earliest co-translational conformations as they form on the ribosome.

Structural investigations of RNCs have been enabled through the use of stalling sequences (28), which form interactions within the tunnel that lead to translational arrest. Frequently utilized stalling sequences are those from TnaC (74) and SecM (75) in bacteria, and Xbp1 (76) and hCMV (77) from eukaryotes. A cryo-EM map of SecM-stalled RNC shows the continuous backbone density arising from the 22-residue arrest sequence from the PTC to the vestibule, beyond the protruding loop of uL23 (75), and identifies arrest as a relay of core contacts made with the ribosome. These interactions ultimately lead to a shift of the C-terminal glycine residue's ester linkage to the P-site transfer RNA (tRNA), a key determinant to the halting of translation, demonstrating that the nascent chain can constrain the geometry of the PTC site in the ribosome. Further displays of the interplay between the ribosome and the nascent chain are evident in the RNC of the secretory protein VemP (*Vibrio* export monitoring polypeptide). VemP adopts a highly compact structure within the tunnel that folds back onto itself within the zone just above the constriction site (34) (**Figure 2*b***). Here, it is the α-helical conformation of the nascent polypeptide chain formed in the upper region in the tunnel close to the PTC (see **Figure 3*a***) that both stabilizes the nascent chain and prevents accommodation of the incoming tRNA, and as a result halts translation. These and other structures (74, 76, 77) showing translation arrest are extreme examples of the nascent chain modulating translation in a two-way communication between the nascent peptide and the ribosome tunnel with potentially critical downstream consequences for nascent chain folding, targeting, and interactions.

Toward the exit vestibule, a lower local resolution is typically observed in cryo-EM maps for the nascent chain (**Figure 3*a***) as a result of a less constrained and more dynamic nascent chain segment capable of sampling a range of trajectories within the exit tunnel (78). Despite these factors, however, the formation of both secondary structure (described in Section 2.1.2) and tertiary structure, such as the aforementioned 29-residue ARD1a zinc finger domain, has been observed in the exit vestibule (35). Beyond the vestibule, whose boundary can be defined by the position of the protruding loop of uL24, where the nascent chain becomes even more dynamic, the likely positions of the globular domains of an 89-residue nascent immunoglobulin-like titin domain (66) and of a 109-residue α-spectrin protein (57) can be mapped into electron densities to ~8 Å-resolution. These latter observations correlate reasonably well with the corresponding force-based folding profiles of these nascent chains that reveal the point at which structure is said to form.

2.2.2. NMR spectroscopy in co-translational folding. The dynamic, flexible segments of the emerging polypeptide outside the ribosome (i.e., beyond the uL24 loop) have remained elusive to electron density techniques but have been amenable to NMR spectroscopic investigations (32), despite the 2.4-MDa molecular weight of the ribosome. NMR investigation of uniformly ^{15}N-labeled unprogrammed ribosomes revealed well-resolved resonances attributable to its most dynamic structural elements that originate predominantly from the flexible parts of the bL12 stalk and disordered regions of the bS1 protein (79–81). These observations help to explain the capacity to observe the similarly dynamic nascent chains when selectively isotopically labeled against a backdrop of NMR-silent ribosomes (27, 32). This strategy has produced fingerprint spectra of high quality, rich in residue-specific structural and dynamic information, for nascent chains of a multidomain immunoglobulin-like protein (33) (**Figure 3*a***), SH3 (82), barnase (83), as well as the intrinsically disordered α-synuclein (64).

As the majority of protein domains undergo a significant collapse only after emerging from the exit tunnel, these NMR observations provide a powerful means to provide high-resolution descriptions of the process of co-translational folding. Using a series of RNC biosynthetic snapshots as a function of nascent chain length of the immunoglobulin domain FLN5 attached to the ribosome

via its disordered subsequent domain, FLN6, a detailed description of the progressive folding was delineated (33). During the emergence of the FLN5 nascent chain, its disordered state was probed via uniform ^{15}N-labeling selectively of the nascent chain, the ^{1}H-^{15}N spectra providing close to 100 residue-specific probes of the nascent polypeptide, informing quantitatively on its structure and dynamics (**Figure 3*a***). Such probing of structural dynamics and interactions across different timescales of nascent chains by NMR are providing an illuminating portrait of co-translational folding processes. Site-specific resonance broadenings in ^{1}H-^{15}N RNC fingerprint spectra have been exploited to provide an understanding of the local interactions with the slowly tumbling ribosome particle (as further described in this section). As it compacts and folds toward its native structure, the nascent chain is best probed via ^{1}H-^{13}C NMR spectra of RNCs selectively labeled at the methyl group side chains, such as those of Ile, Leu, and Val residues, against a perdeuterated background (**Figure 3*a***). Using ^{1}H-^{13}C and ^{1}H-^{15}N NMR, the observation of a natively folded FLN5 domain appears to be delayed relative to the nascent chain length at which the entire domain becomes solvent accessible (33). These observations provide an indicator that the ribosome particle itself is able to influence structure formation in a dramatic manner.

With improvements in both RNC sample preparations (27) and tailored NMR capabilities (32, 84), quantitative NMR dynamics measurements are on the horizon. Moreover, for advanced 3D molecular models of dynamic RNCs, which are beyond the limits of detection by cryo-EM, the application of paramagnetic relaxation enhancement (PRE)-NMR and residual dipolar coupling (RDC)-NMR to probe the dynamic nascent chain is offering promise for increasingly accurate structure determination of the nascent chain. The weak anisotropic alignment of 70S ribosomes in filamentous bacteriophage has enabled RDC-NMR measurements for the bL12 stalk protein attached to the ribosome (81) and is likely to be applicable to RNCs (see Section 2.2.3).

2.2.3. Restraining ribosome-bound nascent chain simulations with experimental data. Molecular dynamics (MD) simulations and modeling, in particular those that incorporate experimental structural data, have an essential role in providing atomistic descriptions of co-translational folding (**Figure 3*b***). The RNC system, however, presents significant challenges to MD modeling through both the sheer size of the ribosome ($\sim 3 \times 10^6$ atoms) and the timescales on which folding events occur (typically μs–ms). Inevitably here, a trade-off exists between the size of the system that is tractable for simulation and the timescale that can be achieved. Advances continue to be made in describing the ribosome and RNCs using MD via various strategies, ranging from all-atom systems to simulations of a subset of the atoms, or via variations of the use of coarse-graining methods (86). The latter have permitted an analysis of the geometry of the ribosomal exit tunnel and its imposed effects on the nascent chain during folding, revealing that significant tertiary structure can form in the close vicinity of the vestibule (48) and in a manner that is strongly modulated by translation rates, as also supported by theoretical and computational investigations (87, 88). In a major effort undertaken toward simulating the entire ribosome, an all-atom simulation of the 70S particle in complex with SecYEß permitted the detailed dissection of their dynamic interactions (89).

MD simulations also enhance the capacity to describe the conformational preferences of nascent chains by integrating experimental data, in particular those obtained from cryo-EM and NMR studies. This combination currently allows the most advanced descriptions of less-well-structured nascent chain regions that are typically of a highly dynamic nature and sample a large distribution of conformations at the exit vestibule and beyond. Particularly useful are experimentally derived restraints of nascent chains on the ribosome used in MD simulations, which include NMR parameters, such as chemical shift–based restraints (33, 90, 91), residual dipolar couplings (81), and PRE-derived distance restraints (92), as well as cryo-EM maps. This

powerful combination will serve as an important basis for detailed structure determinations that are capable of bridging the rigid segments of the nascent chain buried within the ribosome tunnel, with the dynamic ensembles describing the conformations and motions of the emerged parts of the nascent chain. Presently, of the available experimental parameters, the NMR chemical shift is the simplest NMR parameter to obtain and is itself a direct probe of structure (90, 91, 93). Chemical shift–restrained, all-atom simulations of RNCs have been insightful (33) in explorations of the ribosomal surface that an emerging polypeptide is likely to contact upon exit from the particle (**Figure 2*b*,*c***). However, simulating the entire multimillion atom ribosome system is computationally currently expensive (94) and simplification of the simulations of the large ribosomal complex through coarse-graining provides a powerful and expedient strategy (95). Coarse-grained Gō-model simulations have recapitulated the onset of folding of immunoglobulin-like I27 nascent chains (from the muscle protein titin) as measured by force-based folding assays strikingly well, when limiting interactions between protein and ribosome to entirely repulsive ones, hinting at a predominantly steric influence of the ribosome in this particular case (66). Moreover, a coarse-grained model of the intrinsically disordered α-synuclein RNC was also used as a basis to study its interactions with the ribosome and also with the molecular chaperone, TF, during biosynthesis (64). A good agreement is seen with NMR measurements that described a weak bimolecular association and defined a minimum nascent chain length of ∼50 residues from the PTC for it to reach the substrate-binding cradle at the center of the TF chaperone. Understanding and predicting even transient contacts within disordered nascent chains has important ramifications for defining the initial steps of co-translational folding processes. The continued development of algorithm design and hardware is likely to further unlock the predictive capacity of MD to be applied to increasingly longer timescales including the use of deep learning methods (95a), provided sufficient training sets are available, and an integrative structural biology approach combining cryo-EM, NMR, and MD. These strategies will certainly be yielding in terms of defining the entire molecular mechanism of protein folding on the ribosome (**Figure 3*b***).

3. THE STRUCTURE OF THE RIBOSOME MODULATES NASCENT CHAIN FOLDING

An emerging body of data is providing insight into the manner in which the ribosome uniquely remodels the energy landscape of the nascent chain as compared with that of the corresponding isolated polypeptide.

3.1. Native Structure Is Destabilized on the Ribosome

Determining the thermodynamic stability of RNCs to describe the folding potential of the nascent chain has been studied via a range of approaches including NMR spectroscopy, pulse proteolysis, and optical tweezers. NMR spectroscopy has been insightful (33) through its capacity to reveal residue-specific details: Using NMR fingerprint spectra of a pair of FLN5-FLN6 immunoglobulin domain RNCs, the folding transition of the FLN5 domain was described using the observation of discrete resonances corresponding to the disordered domain that diminished as the population of this state decreased during the folding transition. The folding equilibrium of these FLN5 nascent chains was shown to be strongly influenced by the ribosome in a length-dependent manner (**Figure 4*a***): The disordered state of short nascent chains (<43-residue linker) showed site-specific interactions with the ribosome particle, whereas the folded state is destabilized by the ribosome, as the latter sterically restricts the conformational freedom of the polypeptide. With greater

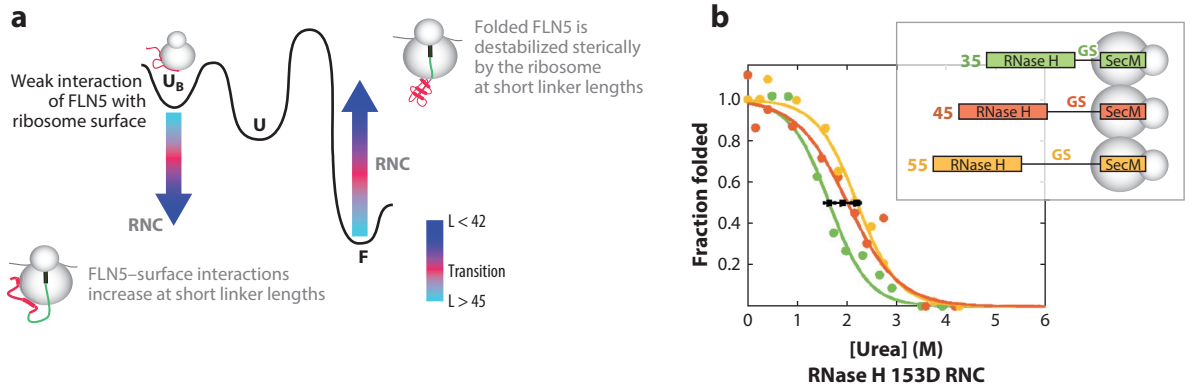

Figure 4

The energetics of native structure formation differ *on* versus *off* the ribosome. (*a*) The energy landscape of RNCs of an Ig-like FLN5 domain changes in a length-dependent manner. Panel *a* adapted with permission from Reference 33. (*b*) Pulse proteolysis experiments describe the stability of an RNase H nascent chain with increasing distance from the ribosome (35, 45, and 55 amino-acid linker lengths). Panel *b* adapted with permission from Reference 38. Abbreviations: GS, Gly-Ser linker; RNC, ribosome-bound nascent chain.

emergence of the polypeptide chain, these effects subside, at which point NMR spectra reveal fingerprints consistent with the nascent chain sampling native-like structure (see also Section 2.2.2).

This destabilizing impact of the ribosome has also been observed quantitatively in several other systems by measuring the differences in thermodynamic stability of RNCs relative to isolated proteins, using an elegant pulse proteolysis approach. By taking RNCs of different lengths incubated in urea and subjecting them to protease for a short time, ΔG values of folding were obtained (38) (**Figure 4*b***). For destabilized variants of dihydrofolate reductase (DHFR), RNase H, and barnase, which were selected as they unfold at urea concentrations at which the ribosome remains intact, these studies revealed a perturbation of native structure during biosynthesis on the ribosome with $\Delta\Delta G$ values up to >2 kcal/mol. The extent of this effect diminishes with increasing nascent chain linker length, as the chain increasingly behaves more like an isolated protein. Insightful kinetics-based experiments using optical tweezers ascribe the destabilizing effect to the ribosome both decelerating the nascent chain's folding rate and accelerating its unfolding rate (58, 59, 96).

3.2. The Stabilization of Disordered Nascent Chains via Interactions with the Ribosome Surface

Among the factors that perturb the energy landscape of nascent polypeptide chains on the ribosome are interactions formed between disordered nascent chains and the ribosome surface (33, 58), the impact of which can modulate the onset of folding. Although it is important to understand the behavior of incomplete sequences on the ribosome during translation, such studies present challenges, as current descriptions of disordered polypeptides even in isolation are complicated by the ability to define meaningful parameters that accurately describe the distributions of unstructured, dynamically interconverting states (97). Experimentally, high-resolution NMR data are the most potent for describing the topological preferences within disordered ensembles and are typically used to restrain structure determination methods. The averaging within the NMR parameters does, however, complicate their definitive use and render such processes reliant on

the quality of the force field used (98, 99). Intrinsically disordered proteins (IDPs) and truncations of proteins (100–102) provide useful disordered models without requiring denaturants. On the ribosome, nascent chains maintain nonglobular conformations through the occlusion of their C-terminal portion by the ribosome tunnel (**Figure 2**). The N-terminal emerged, solvent-accessible sequence can in principle be mimicked off the ribosome by truncation of the C terminus to generate isolated N-terminal fragments (100), but these are unable to capture the effects of tethering to the ribosome. Comparative studies of disordered proteins on and off the ribosome thus have the potential to yield valuable insight into the early stages of co-translational folding but may also be limited, particularly as the disordered states of isolated proteins likely do not model the dynamic characteristics of the nascent chain within the confines and vicinity of the exit vestibule. Additional important features unique to RNCs include the mechanical tension exerted on the nascent chain through its tethering to the ribosome via the P-site tRNA, which may contribute to its unfolded nature (50). Such effects will require elucidation to understand their impact on folding.

3.2.1. Nascent chain interactions in the tunnel (upper and middle regions). The steric confinement and nascent chain interactions with the ribosome within the exit tunnel contribute to maintaining the majority of nascent polypeptides in a disordered state. The interactions mediated by naturally occurring stalling sequences permit the dissection of the relevance of nascent chain–ribosome contacts (28). Although the molecular details differ between nascent chains, some binding sites within the tunnel appear common to some substrates. In the upper region of the tunnel (**Figure 3*a***), adjacent to the PTC, the 23S ribosomal RNA (rRNA) nucleotides A2062 and U2585 (*E. coli* numbering) make contact with the bacterial stalling systems (see Section 2.2.1) of ErmC and TnaC (103, 104). Both A2062 and U2585 are required for facilitating rapid polypeptide release during termination (105). In cryo-EM structures of human ribosomes, the equivalent nucleotides (*Homo sapiens* A3879 and U4493) are seen to make extensive contacts with the hCMV stalling motif (77). At the narrow constriction site where the middle region of the tunnel begins, strong contacts are consistently seen between the nascent chain and the protruding loops of uL4 and uL22, as well as the 23S rRNA nucleotide A751 (*E. coli* numbering), which also contributes to the tunnel narrowing (46, 74, 75, 77).

3.2.2. Nascent chain interactions in the lower tunnel and outside. Upon emergence from the tunnel, the nascent chain still experiences the ribosomal surface at very high local concentrations—the effective concentration of the ribosome surface approaches 20 mM for a residue located 10 residues from the protruding uL24 loop, beyond the exit tunnel, owing to its covalent linkage to the ribosome (33). Even a weak affinity between the nascent chain sequence and the ribosome surface, as observed within the immunoglobulin domain FLN5, can therefore result in extensive nascent chain–ribosome interactions through the additional effect of tethering (33). Recent studies have also begun to elucidate some of the molecular details of nascent chain interactions made with the outer ribosomal surface. RNCs of IDP systems provide the ideal scaffold in delineating such interactions, as these are devoid of competing folding processes. For RNCs of the intrinsically disordered PIR protein, fluorescence depolarization anisotropy measurements of an N-terminal fluorophore demonstrated interactions between positively charged regions of the nascent chain and the negatively charged ribosome surface (106) (**Figure 5**). Such interactions can have important effects: Two oppositely charged IDPs (ACTR and NCBD) have been observed to assemble co-translationally, where the negatively charged ACTR nascent chains interact with the isolated, positively charged NCBD. The assembly between isolated ACTR and

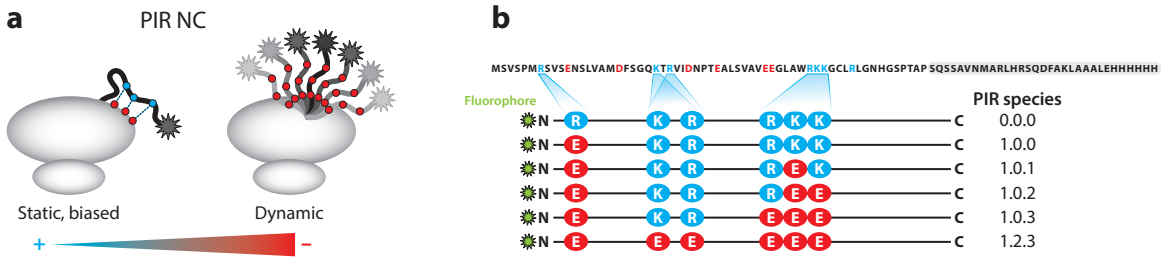

Figure 5

Acidic side chains in the PIR nascent chain (NC) mediate interactions with the negatively charged ribosome surface and result in significant populations of spatially biased conformations. Figure adapted with permission from Reference 106; copyright 2013 American Chemical Society.

ribosome-bound NCBD, however, is not observed, because during its translation, NCBD undergoes interactions with the ribosome surface that are thought to out-compete co-translational assembly (24). NMR spectroscopy is also able to describe dynamic chemical exchange processes and has been used to dissect residue-specific-level information on RNCs. A study on the intrinsically disordered α-synuclein RNC revealed a propensity for sequence segments rich in basic residues to form electrostatics-based interactions with the ribosome surface (64). The quantitation of such contacts using NMR is an ongoing area of activity, with likely relevance to the folding onset and contributing to the delay in folding observed as shown in the case of the FLN5 domain (Section 2.2.2). NMR studies show that most of these ribosome–nascent chain interactions are highly transient in nature (33, 64). Ribosome–nascent chain interactions are also observed in optical tweezer measurements of a T4 lysozyme RNC, where an ∼450-fold attenuation in folding rate, relative to the isolated protein, can be partially mitigated by weakening these interactions using high-salt conditions (58).

In accord with these observations, the likely ribosomal feature mediating interactions with nascent chains is its high net-negative charge conferred largely by the sugar–phosphate RNA backbone [∼5,000 phosphate charges (107)]. R-proteins also contribute to this negative surface charge via their solvent-exposed globular domains, whereas their basic extensions—resulting in near-to-neutral overall charge of these proteins (108)—presumably allow their insertion into the particle and thereby facilitate ribosome assembly. The resulting globally negatively charged ribosome surface attracts positively charged counterions, mainly Mg^{2+} but also K^+, both of which have significant importance not only for the ribosome structure and function (109) but also for cellular function, to the extent that the ribosome has been described as an important reservoir of these cations and contributes to regulating their concentrations within the cell (110). Mg^{2+} homeostasis has been directly linked to co-translational events in *E. coli:* A recently described process termed intrinsic ribosome destabilization (IRD) allows, under low Mg^{2+} conditions, the synthesis of a regulatory MgtL nascent chain to cause ribosome dissociation, thereby aborting its translation (111). Such sensing would appear to be part of a feedback mechanism in which downregulation of MgtL translation causes upregulation of the expression of the Mg^{2+} transporter, MgtA. A molecular understanding of the role of Mg^{2+} ions in regulating this intimate two-way communication between the nascent chain and the ribosome will begin to shed light on the modes by which electrostatic interactions can be modulated by metals on the ribosome.

It is likely that Mg^{2+} and K^+ cations also contribute to forming the counterion shell surrounding the ribosome, together with polyanions such as putrescine and spermidine (112). How this shell influences nascent chain folding and specific interactions between the nascent chain and the

ribosome surface remains close to entirely unexplored. The high negative charge of the ribosome surface may also have a significant role in supporting quinary interactions with cellular components, in particular those that are highly basic (113). Ribosomes have been shown to bind a range of substrates: in some cases, including for DHFR and adenylate kinase, micromolar affinity interactions are suggested to reduce enzymatic activity in the cell, whereas thymidylate synthase activity is activated through ribosome binding (114).

These features of the ribosome surface suggest that interaction sites for the nascent chain likely also exist at the vestibule and on the outer ribosomal surface. It is less clearly understood whether nascent chain segments preferentially interact with specific sites or whether the sterically accessible space imposed by the ribosome on the nascent chain is the overarching determinant of such interactions. Several possible r-protein interaction sites have been identified via cross-linking (115). The protruding β-hairpin loop of uL23 and, at approximately the same cross-section of the tunnel, nucleotide A1321 of helix 50 are seen to form contacts with the N-terminal segments of the SecM nascent chain. Helix 50 of the 28S rRNA is also an area contacted by eukaryotic, helical nascent sequences (46).

At the vestibule, the β-hairpin loop of uL24 protrudes into the space delimiting the end of the tunnel. Recent force-based folding assays have identified a uL23 loop deletion to allow the ADR1a zinc finger to sample structure at a slightly earlier point in its translation than in wild-type ribosomes (116), in which a cryo-EM map shows the folded structure in the vestibule close to the uL23 loop (35) (**Figure 3*a***). Titin I27, which according to force-based assays folds when it is at a greater distance from the PTC relative to the ADR1a zinc finger (66), is able to fold at marginally shorter distances when the protruding loop of uL24 is removed (116). These assays suggest an important steric role for the tunnel-exposed loops of uL23 and uL24 in modulating the onset of folding of nascent polypeptides.

In bacteria, the uL24 loop is extended relative to that in eukaryotes, potentially partitioning the exit vestibule into distinct, possible paths for the nascent chain (T. Wlodarski, personal communication). Cryo-EM maps of titin I27 (66) and an α-spectrin domain (57) suggest in both cases that the emerging structured domain is nudged by the uL24 loop toward helix H59 of the 23S rRNA and uL23; therefore, it is possible that uL24 dictates a specific path for the emergent polypeptide. MD simulations show the immunoglobulin domain FLN5 nascent chain frequently contacting uL24 and, to a lesser extent, uL23, uL29, and uL22 on the outer ribosome surface (33).

Further understanding of the incipient structure on the ribosome will likely require a detailed molecular description of the nascent chain at the exit vestibule, where the cumulative effects of ion concentrations and steric confinement combined with the tethering of the nascent chain are likely to remodel the conformational preferences of the dynamic nascent polypeptide in unique ways.

3.3. The Ribosome Attenuates Misfolding by Acting as a Solubility Tag

Although the folding of small proteins (<100 residues) is frequently driven by thermodynamics alone, and in many cases occurs posttranslationally, larger proteins with more complex topologies, such as multidomain proteins (70% of the eukaryotic proteome), largely initiate their folding during their synthesis on the ribosome. In many cases, their free energy landscapes are likely to harbor partially folded intermediate conformations, including kinetically trapped states that arise through the interplay between the rate of synthesis and that of folding (117) (see Section 4). This becomes more likely when the sequence similarity between neighboring domains is significant, as occurs in tandem repeat domains (118), where interactions between adjacent domains during folding (119) can result in misfolding (100, 117, 120). Misfolding of multidomain proteins on

the ribosome has been observed in the two-domain, four-EF-hand protein calerythrin (59) via optical tweezer experiments. The mispairing of the first two EF hands comprising the emerged N-terminal domain was suggested to be the basis of the misfolding, which then resolved into the native state through a domain rearrangement after the very C-terminal EF hand had emerged from the ribosomal exit tunnel (**Figure 6*a***).

Another optical tweezer study showing the ribosome modulating multidomain protein folding is a description of the apparent refolding rates of elongation factor G nascent chains (121) (**Figure 6*b***). At short nascent chain lengths, prior to an optimal length at which the folding rate is maximal, the disordered nascent polypeptide undergoes interactions with the ribosome that effectively slow its folding rate. At lengths beyond this optimal length, additional emerged portions of the neighboring II-domain become available to interact with the G-domain, which also disfavors folding. These intra-chain interactions at long lengths are also shown to lead to misfolded conformations, although the rate of their formation is decreased by the ribosome. The net result of this ribosome-mediated destabilization is thus to favor productive folding and destabilize misfolding pathways. Among proteins observed to misfold, T4 lysozyme (58), apoMb (122), tailspike protein P22 (56), and HaloTag (39) have been studied co-translationally and, on the ribosome, all appear to be relatively protected from such effects. A possible rationale for these observations may relate to the described general tendency of nascent proteins to interact with the ribosome surface, which may destabilize some elements of intramolecular structure formation to mitigate misfolding events.

The amino acid groups on the nascent chain that are seen to support interactions with the ribosome particle [i.e., basic and aromatic residues (64, 106)] are also those typically known to mediate interactions with well-described chaperones such as TF (14) and DnaK (123), both of which sequester polypeptides to assist nascent chains in their pathway to folding. Analogously, the ribosome may be regarded as the first chaperone the nascent chain encounters before handover to downstream chaperones, which act in *trans* if the substrate requires further assistance to fold correctly. This holdase effect is likely to contribute to a solubilizing capability of the ribosome and to occur via relatively weak and frequently nonspecific interactions with its disordered substrates (albeit with a preference for positively charged and aromatic amino acid types) to enable efficient folding and release of the nascent chain, such as is seen for parts of α-synuclein (64), disordered FLN5 (33), and PIR (106) nascent chains.

Is the ribosome an efficient chaperone? Some useful insights into this question may be obtained by considering aspects of the heterologous soluble expression of proteins. The extent of aggregation of the protein bovine enterokinase was reported to be correlated with the acidity of the solubilizing partner to which it was fused (124) and, indeed, some of the most commonly employed solubility tags are themselves highly negatively charged—for example, *NusA*, which has a net charge of −40. The high net negative charge of the ribosome [possibly in the negative thousands (107)] may therefore make it a fusion partner that protects against insoluble expression. The simple effect of steric restriction imposed by the large ribosome particle is also likely to play a role, such as through preventing undesired inter- and intramolecular association of an emerging polypeptide chain. These effects that describe the ribosome as a giant solubility tag may also in part explain observations that some proteins such as eukaryotic receptors that have not been purified successfully from *E. coli* are successfully co-translationally folded and protected from aggregation processes as long as they remain bound to the bacterial ribosome during ribosome display experiments (125). This apparent resilience of the ribosome to maintain diverse polypeptide chains in a largely aggregation-free state would also appear to be beneficial in mitigating the potentially detrimental effects of nonproductive mutations during gene evolution (126).

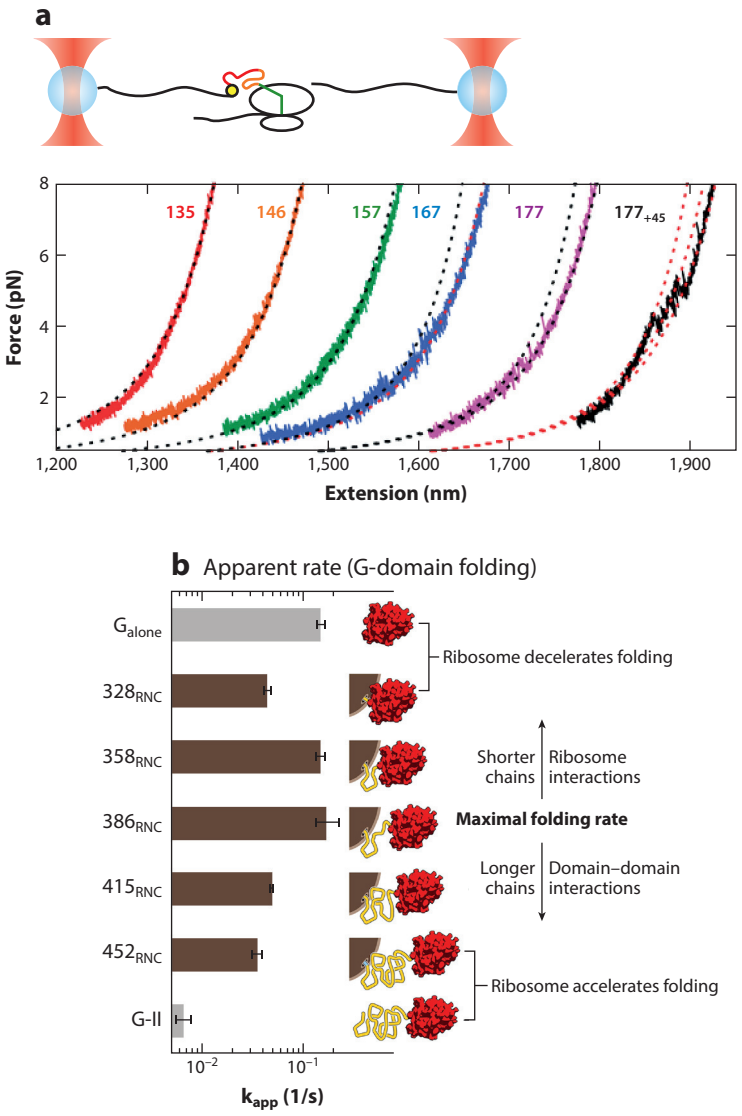

Figure 6

Observations of misfolded nascent chains on the ribosome. (*a*) Refolding force–extension curves using optical tweezers at varying lengths of calerythrin RNCs. At polypeptide chain lengths of 135, 146, and 157, the nascent chain does not fold despite an entire pair of EF hands (EF1 and EF2) having emerged from the exit tunnel. At chain lengths of 167 and 177, a misfolded state is sampled. Upon further translation of an additional 45-residue linker, the native state is sampled via a C-terminal intermediate. Panel *a* adapted with permission from Reference 59. (*b*) Apparent folding rates of the G-domain within RNCs of varying lengths (*brown*) versus the isolated G_{alone} and G-II polypeptides (*gray*) measured using repeated force-ramp cycles via optical tweezers, showing the length at which the RNC folding rate is maximal. Panel *b* adapted with permission from Reference 121. Abbreviation: RNC, ribosome-bound nascent chain.

The ribosome particle's protein/RNA composition further influences how efficiently certain substrates are translated and/or folded (127). Although the structural core and catalytic sites are highly conserved in the ribosome particle, organisms and even organelle-specific ribosomes are characterized by distinct compositions and overall architecture (128). A particularly striking example is that of the ribosomal exit tunnel, which is among the most conserved regions of the ribosome—except in mitochondrial ribosomes (mitoribosomes), where it has undergone a remarkable evolution. In fungal mitoribosomes, an evolutionary loss of a single RNA helical stretch (eight nucleotides of mitochondrial rRNA helix H24) has led to the creation of an alternative exit tunnel that is wider than that of the bacterial ribosome and is suggested to have conferred a functional advantage in enabling mitochondrial proteins to fold early within this tunnel (129). Both the tunnel's size and the proteins that line it are thought to have evolved to ideally accommodate the mitochondrion's proteome, which is particularly enriched in membrane proteins with high helical content. An understanding of how such organelle-specific ribosomes cope with distinct proteomes will be an exciting area of future study, as will be the study of the impact on productive folding of evolved orthogonal ribosomes (130) such as those that exploit quadruplet codons for the translation of non-natural amino acids (131).

4. THE INFLUENCE OF PROTEIN TRANSLATION KINETICS

The rate at which the ribosome translates proteins has the potential to profoundly affect folding outcomes. Within the stochastic search of a protein for its native fold, the rate of translation governs the pathways visited by a nascent protein across its energy landscape (117). An identical amino acid sequence produced by synonymous codons for which the decoding rate differs can, in principle, lead to distinct folding outcomes. This was exemplified in a synthetic system, in which two N- and C-terminal fluorescent proteins (FPs) half-domains kinetically compete for a central half-domain to complete their fluorophore fold (132). The substitution of synonymous codons within the central half-domain resulted in modulation of the decoding rate and altered the relative yields of the possible N- and C-terminal products YFP (yellow) and BFP (blue).

Infrequently used codons are also often linked to rare tRNA isoacceptors, which statistically take more time to diffuse toward the ribosome and therefore slow translation (133). Superficially, using codons with high-abundance tRNA isoacceptors might seem a useful strategy to maximize protein production; indeed, rapidly growing bacteria (e.g., in rich media) have been suggested to do this for a large proportion of their genes coding for highly expressed proteins to maximize growth rates (134). For a number of proteins studied [SufI (135), EgFABP1 (136), NBD1 from CFTR (30), chloramphenicol acetyltransferase (CAT) (137), the clock protein FRQ (138), γ-B crystallin (139)], substitution of stretches of rare codons leads to adverse outcomes ranging from aberrant folding to alternative conformations. Synonymous versions of γ-B crystallin lead to different structures as described using NMR spectroscopy (139). The CAT nascent chain requires specific stretches of rare codons to promote co-translational folding, and the failure to do so results in posttranslational degradation, as demonstrated through antibiotic resistance experiments (137). Aberrant folding due to non-optimal codon usage is also a feature of the NBD1 domain (from CFTR) nascent chains, in which FRET studies demonstrated that the kinetics of translation of the α-subdomain are critical to allowing for successful β-sheet core folding. Indeed, the NBD1 example presents a case for how subdomain folding may be kinetically coupled: NBD1 avoids misfolding through optimized codons that ensure a slow translation rate window to enable the α-subdomain to form (30). Indeed, rare codon stretches are typically identified at precisely defined positions, strategically clustering ∼30–40 amino acids downstream from the C terminus of independently folding entities (or foldons) (135); this length corresponds to the

approximate sequence length needed to span the ribosomal exit tunnel. The use of rare codons, particularly within large or multidomain proteins, is therefore suggested to allow each foldon to acquire structure before the emergence of additional sequence.

Collectively, these studies demonstrate a link between protein stability and codon choice; hence, the degeneracy of the genetic code appears to have been exploited by cells to impart an additional level of temporal control by fine-tuning between rapid synthesis and high-fidelity folding under conditions of cellular stress. This level of control has likely been achieved via evolutionary pressures through the optimization of the timing of biosynthesis and concurrent co-translational folding (140, 141).

Overall, how important are these effects? Given that synonymous codon changes do not alter the energy landscape of the nascent chain but only affect the choice of folding pathway down the same energy funnel—or in other words, bias the accessible part of the energy landscape—in most cases, the nascent chain is still able to fold into its native state. In *E. coli*, it takes ~50–250 ms to translate a single codon (142), and ribosome profiling data suggest that the decoding rate maximally varies only approximately two-fold in yeast (143) [although estimates vary between studies (144)]. In the large majority of cases, a single codon substitution would therefore not cause a significant effect on the structure of the protein product. Indeed, in the majority of the experiments described above, in which codon substitutions caused aberrant protein folding or alternative conformations, stretches of tens of residues or more have been substituted (30, 137–139). A theoretical prediction of a possible switch from folding posttranslationally to folding co-translationally was made on the basis of changing every single codon to its slowest translating counterpart (144). Generally, therefore, folding is likely to be somewhat resilient to a certain amount of variation in translation rates. In some select cases, however, even the substitution of single codons can perturb folding outcomes dramatically. The protein SufI, whose N-terminal domain can fold co-translationally, becomes progressively more labile to proteolysis with 1–2 codon substitutions within a slow-translating cluster (135). Another notable exception is the single nucleotide polymorphism (SNP) in the *MDR1* gene, which alters the conformation of its encoded P-glycoprotein, leading to a different substrate specificity and an altered sensitivity to inhibitor compounds (145). This striking effect from a single codon substitution may be rationalized by the fact that MDR proteins are able to readily alter substrate specificity through small changes in the conformation of the drug-binding pocket (146) to confer promiscuous resistance, implying a shallow energy landscape for MDR allowing different conformations to be accessed even with very subtle changes in translation kinetics.

5. OUTLOOK

The dynamic nature of the nascent chain, the large size of the ribosome, the complex interplay of the various integrated systems of processing and quality control, and the kinetic aspect of biosynthesis all contribute to the challenge of creating a high-resolution, mechanistic understanding of co-translational folding. Considerable strides have been made, underpinned by comprehensive studies of the refolding of isolated proteins in vitro. Nevertheless, the field of protein folding on the ribosome is still in its infancy relative to the plethora of descriptions of protein translation, in which ribosome motion has been extensively probed by a range of studies examining translation initiation through to termination. An understanding, therefore, of how proteins begin their journey toward forming the functional structures that underpin every biological process in the cell is clearly of importance and likely to revolutionize our present understanding of proteostasis and its close ties to the onset of human disease.

One of the frontiers in understanding co-translational folding is advancement of our high-resolution structural description of nascent chains. We expect that quasi-atomic-resolution structural ensembles of stalled nascent chains will be increasingly provided by combining cryo-EM maps with NMR experimental data (e.g., chemical shift, RDC, or PRE measurements; see Section 2.2) as restraints for MD simulations. A faithful representation of the large-scale motions sampled by dynamic nascent chains will likely only be captured using structural ensembles, especially in the case of the disordered nascent polypeptides that are the mainstay during synthesis on the ribosome. Understanding their subtle conformational preferences will likely be key, as these may constitute transient contacts that initiate structure formation in a fledgling nascent chain and ultimately guide downstream folding. The challenges in such MD simulations are manifold (as described in Section 2.2.3) and include the convergence of simulations being complicated by the very long timescale (ms–s) of translation itself. Moreover, an accurate representation of electrostatics, experimentally shown to be important in regulating co-translational events, presents significant modelling challenges, especially for very highly charged systems such as the ribosome (147).

Another frontier in understanding folding on the ribosome is to describe co-translational events directly and in a time-resolved manner as they occur during polypeptide elongation. Although high-resolution structural studies are only realistically accessible on nascent chains captured at thermodynamic equilibrium, a holistic view of co-translational folding must consider how the kinetics of translation, relative to the kinetics of folding, modulate the structures actually sampled during biosynthesis (59). A study of the fast-folding protein HaloTag is an example that bridges kinetically resolved and equilibrium observations: The folding of HaloTag was measured via the binding of a ligand with a very rapid on-rate (39) to determine its rate of co-translational folding during synthesis and the extent of folding at discrete nascent chain lengths as measured at equilibrium.

With bulk biophysical studies typically unable to resolve heterogeneous nascent chain behavior within a single sample, another particularly exciting prospect is that of dissecting events at the single-molecule level. Fluorescence-based strategies can track translation with single-codon resolution in real time (148, 149) and will increasingly be combined with descriptions of nascent chain motions during elongation (35, 150). Watching rounds of translation and folding occur on single molecules will allow the dissection of the distribution of populations sampling different conformations that define distinct co-translational folding pathways.

In vivo data of co-translational protein folding are sparse and exceedingly challenging to collect but ultimately will be needed to understand the influence of the cellular environment. Indeed, seminal studies using in vivo pulse-chase assays first demonstrated the existence of co-translational folding via the autocatalytic cleavage of the cytosolic Semliki Forest virus protein (SFVP) in both prokaryotic *E. coli* and in eukaryotic Chinese hamster ovary (CHO) cells (151). The more recent biochemical approaches using force-based measurements of folding have emerged as a tractable method to compare in vitro versus in vivo folding (50). Moreover, the continuous development of high-sensitivity NMR (32) and in-cell NMR (152) offers the promise of observing RNCs folding within live cells.

It is likely that by integrating a range of experimental and computational approaches, which permit measurements of thermodynamics and kinetics to be merged with molecular-level structural models, a detailed mechanistic description of fundamental co-translational protein folding processes, both in vitro and in vivo, will emerge over the next decade. Such knowledge of the process underpinning activity of all living systems will undoubtedly reveal novel and remarkable precedents in cellular biology, and importantly, will serve as a means for uncovering the crucial links that exist between protein synthesis and folding on the ribosome and many human diseases.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We dedicate this manuscript to Chris Dobson (1949–2019), thanking him for many treasured years of collaboration, countless inspirational discussions on this topic, and his unerring support, generosity, and friendship. We thank the Wellcome Trust for an Investigator Award (to J.C.).

LITERATURE CITED

1. Cabrita LD, Dobson CM, Christodoulou J. 2010. Protein folding on the ribosome. *Curr. Opin. Struct. Biol.* 20:33–45
2. Ciryam P, Morimoto RI, Vendruscolo M, Dobson CM, O'Brien EP. 2013. In vivo translation rates can substantially delay the cotranslational folding of the *Escherichia coli* cytosolic proteome. *PNAS* 110:E132–40
3. Netzer WJ, Hartl FU. 1997. Recombination of protein domains facilitated by co-translational folding in eukaryotes. *Nature* 388:343–49
4. Chiti F, Dobson CM. 2017. Protein misfolding, amyloid formation, and human disease: a summary of progress over the last decade. *Annu. Rev. Biochem.* 86:27–68
5. Brandt F, Etchells SA, Ortiz JO, Elcock AH, Hartl FU, Baumeister W. 2009. The native 3D organization of bacterial polysomes. *Cell* 136:261–71
6. Kramer G, Shiber A, Bukau B. 2019. Mechanisms of cotranslational maturation of newly synthesized proteins. *Annu. Rev. Biochem.* 88:337–64
7. Hartl FU, Bracher A, Hayer-Hartl M. 2011. Molecular chaperones in protein folding and proteostasis. *Nature* 475:324–32
8. Duttler S, Pechmann S, Frydman J. 2013. Principles of cotranslational ubiquitination and quality control at the ribosome. *Mol. Cell* 50:379–93
9. Wang F, Durfee LA, Huibregtse JM. 2013. A cotranslational ubiquitination pathway for quality control of misfolded proteins. *Mol. Cell* 50:368–78
10. Lynch M, Marinov GK. 2015. The bioenergetic costs of a gene. *PNAS* 112:15690–695
11. Lane N, Martin W. 2010. The energetics of genome complexity. *Nature* 467:929–34
12. Kramer G, Boehringer D, Ban N, Bukau B. 2009. The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins. *Nat. Struct. Mol. Biol.* 16:589–97
13. Yang C-I, Hsieh H-H, Shan S-O. 2019. Timing and specificity of cotranslational nascent protein modification in bacteria. *PNAS* 116:23050–60
14. Patzelt H, Rüdiger S, Brehmer D, Kramer G, Vorderwülbecke S, et al. 2001. Binding specificity of *Escherichia coli* trigger factor. *PNAS* 98:14244–49
15. Kaiser CM, Chang H-C, Agashe VR, Lakshmipathy SK, Etchells SA, et al. 2006. Real-time observation of trigger factor function on translating ribosomes. *Nature* 444:455–60
16. Stein KC, Kriel A, Frydman J. 2019. Nascent polypeptide domain topology and elongation rate direct the cotranslational hierarchy of Hsp70 and TRiC/CCT. *Mol. Cell* 75:1117–30
17. Schibich D, Gloge F, Pöhner I, Björkholm P, Wade RC, et al. 2016. Global profiling of SRP interaction with nascent polypeptides. *Nature* 536:219–23
18. Berndt U, Oellerer S, Zhang Y, Johnson AE, Rospert S. 2009. A signal-anchor sequence stimulates signal recognition particle binding to ribosomes from inside the exit tunnel. *PNAS* 106:1398–403
19. Brandman O, Hegde RS. 2016. Ribosome-associated protein quality control. *Nat. Struct. Mol. Biol.* 23:7–15
20. Wang F, Canadeo LA, Huibregtse JM. 2015. Ubiquitination of newly synthesized proteins at the ribosome. *Biochimie* 114:127–33

21. Turner GC, Varshavsky A. 2000. Detecting and measuring cotranslational protein degradation in vivo. *Science* 289:2117–20
22. Shiber A, Döring K, Friedrich U, Klann K, Merker D, et al. 2018. Cotranslational assembly of protein complexes in eukaryotes revealed by ribosome profiling. *Nature* 561:268–72
23. Shieh Y-W, Minguez P, Bork P, Auburger JJ, Guilbride DL, et al. 2015. Operon structure and cotranslational subunit association direct protein assembly in bacteria. *Science* 350:678–80
24. Marino J, Buholzer KJ, Zosel F, Nettels D, Schuler B. 2018. Charge interactions can dominate coupled folding and binding on the ribosome. *Biophys. J.* 115:996–1006
25. Borman S. 2007. Protein baby pictures. *Chem. Eng. News* 85:56–57
26. Pauling L, Corey RB, Branson HR. 1951. The structure of proteins: two hydrogen-bonded helical configurations of the polypeptide chain. *PNAS* 37:205–11
27. Cassaignau AME, Launay HMM, Karyadi M-E, Wang X, Waudby CA, et al. 2016. A strategy for cotranslational folding studies of ribosome-bound nascent chain complexes using NMR spectroscopy. *Nat. Protoc.* 11:1492–507
28. Ito K, Chiba S. 2013. Arrest peptides: *cis*-acting modulators of translation. *Annu. Rev. Biochem.* 82:171–202
29. Kelkar DA, Khushoo A, Yang Z, Skach WR. 2012. Kinetic analysis of ribosome-bound fluorescent proteins reveals an early, stable, cotranslational folding intermediate. *J. Biol. Chem.* 287:2568–78
30. Kim SJ, Yoon JS, Shishido H, Yang Z, Rooney LA, et al. 2015. Translational tuning optimizes nascent protein folding in cells. *Science* 348:444–48
31. Nogales E, Scheres SHW. 2015. Cryo-EM: a unique tool for the visualization of macromolecular complexity. *Mol. Cell* 58:677–89
32. Waudby CA, Launay H, Cabrita LD, Christodoulou J. 2013. Protein folding on the ribosome studied using NMR spectroscopy. *Prog. Nucl. Magn. Reson. Spectrosc.* 74:57–75
33. Cabrita LD, Cassaignau AME, Launay HMM, Waudby CA, Wlodarski T, et al. 2016. A structural ensemble of a ribosome–nascent chain complex during cotranslational protein folding. *Nat. Struct. Mol. Biol.* 23:278–85
34. Su T, Cheng J, Sohmen D, Hedman R, Berninghausen O, et al. 2017. The force-sensing peptide VemP employs extreme compaction and secondary structure formation to induce ribosomal stalling. *eLife* 6:e25642
35. Nilsson OB, Hedman R, Marino J, Wickles S, Bischoff L, et al. 2015. Cotranslational protein folding inside the ribosome exit tunnel. *Cell Rep.* 12:1533–40
36. Tsalkova T, Odom OW, Kramer G, Hardesty B. 1998. Different conformations of nascent peptides on ribosomes. *J. Mol. Biol.* 278:713–23
37. Kleizen B, van Vlijmen T, de Jonge HR, Braakman I. 2005. Folding of CFTR is predominantly cotranslational. *Mol. Cell* 20:277–87
38. Samelson AJ, Jensen MK, Soto RA, Cate JHD, Marqusee S. 2016. Quantitative determination of ribosome nascent chain stability. *PNAS* 113:13402–7
39. Samelson AJ, Bolin E, Costello SM, Sharma AK, O'Brien EP, et al. 2018. Kinetic and structural comparison of a protein's cotranslational folding and refolding pathways. *Sci. Adv.* 4:eaas9098
40. Holtkamp W, Kokic G, Jäger M, Mittelstaet J, Komar AA, et al. 2015. Cotranslational protein folding on the ribosome monitored in real time. *Science* 350:1104–7
41. Dao Duc K, Batra SS, Bhattacharya N, Cate JHD, Song YS. 2019. Differences in the path to exit the ribosome across the three domains of life. *Nucleic Acids Res.* 47:4198–210
42. Voss NR, Gerstein M, Steitz TA, Moore PB. 2006. The geometry of the ribosomal polypeptide exit tunnel. *J. Mol. Biol.* 360:893–906
43. Ziv G, Haran G, Thirumalai D. 2005. Ribosome exit tunnel can entropically stabilize α-helices. *PNAS* 102:18956–61
44. Lu J, Deutsch C. 2005. Folding zones inside the ribosomal exit tunnel. *Nat. Struct. Mol. Biol.* 12:1123–29
45. Woolhead CA, McCormick PJ, Johnson AE. 2004. Nascent membrane and secretory proteins differ in FRET-detected folding far inside the ribosome and in their exposure to ribosomal proteins. *Cell* 116:725–36

46. Bhushan S, Gartmann M, Halic M, Armache J-P, Jarasch A, et al. 2010. α-Helical nascent polypeptide chains visualized within distinct regions of the ribosomal exit tunnel. *Nat. Struct. Mol. Biol.* 17:313–17
47. Bañó-Polo M, Baeza-Delgado C, Tamborero S, Hazel A, Grau B, et al. 2018. Transmembrane but not soluble helices fold inside the ribosome tunnel. *Nat. Commun.* 9:5246
48. O'Brien EP, Hsu S-TD, Christodoulou J, Vendruscolo M, Dobson CM. 2010. Transient tertiary structure formation within the ribosome exit port. *J. Am. Chem. Soc.* 132:16928–37
49. Kosolapov A, Deutsch C. 2009. Tertiary interactions within the ribosomal exit tunnel. *Nat. Struct. Mol. Biol.* 16:405–11
50. Goldman DH, Kaiser CM, Milin A, Righini M, Tinoco I Jr., et al. 2015. Mechanical force releases nascent chain–mediated ribosome arrest in vitro and in vivo. *Science* 348:457–60
51. Farías-Rico JA, Ruud Selin F, Myronidi I, Frühauf M, von Heijne G. 2018. Effects of protein size, thermodynamic stability, and net charge on cotranslational folding on the ribosome. *PNAS* 115:E9280–87
52. Kemp G, Kudva R, de la Rosa A, von Heijne G. 2019. Force-profile analysis of the cotranslational folding of HemK and filamin domains: comparison of biochemical and biophysical folding assays. *J. Mol. Biol.* 431:1308–14
53. Notari L, Martínez-Carranza M, Farías-Rico JA, Stenmark P, von Heijne G. 2018. Cotranslational folding of a pentarepeat β-helix protein. *J. Mol. Biol.* 430:5196–206
53a. Jensen MK, Samelson AJ, Steward A, Clarke J, Marqusee S. 2020. The folding and unfolding behavior of ribonuclease H on the ribosome. bioRxiv 2020.04.16.044867. **https://doi.org/10.1101/2020.04.16.044867**
54. Ismail N, Hedman R, Lindén M, von Heijne G. 2015. Charge-driven dynamics of nascent-chain movement through the SecYEG translocon. *Nat. Struct. Mol. Biol.* 22:145–49
55. Yap M-N, Bernstein HD. 2011. The translational regulatory function of SecM requires the precise timing of membrane targeting. *Mol. Microbiol.* 81:540–53
56. Evans MS, Sander IM, Clark PL. 2008. Cotranslational folding promotes β-helix formation and avoids aggregation in vivo. *J. Mol. Biol.* 383:683–92
57. Nilsson OB, Nickson AA, Hollins JJ, Wickles S, Steward A, et al. 2017. Cotranslational folding of spectrin domains via partially structured states. *Nat. Struct. Mol. Biol.* 24:221–25
58. Kaiser CM, Goldman DH, Chodera JD, Tinoco I, Bustamante C. 2011. The ribosome modulates nascent protein folding. *Science* 334:1723–27
59. Alexander LM, Goldman DH, Wee LM, Bustamante C. 2019. Non-equilibrium dynamics of a nascent polypeptide during translation suppress its misfolding. *Nat. Commun.* 10:2709
60. Ban N, Nissen P, Hansen J, Moore PB, Steitz TA. 2000. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* 289:905–20
61. Schluenzen F, Tocilj A, Zarivach R, Harms J, Gluehmann M, et al. 2000. Structure of functionally activated small ribosomal subunit at 3.3 Å resolution. *Cell* 102:615–23
62. Brown A, Shao S. 2018. Ribosomes and cryo-EM: a duet. *Curr. Opin. Struct. Biol.* 52:1–7
63. Serber Z, Corsini L, Durst F, Dötsch V. 2005. In-cell NMR spectroscopy. *Meth. Enzymol.* 394:17–41
64. Deckert A, Waudby CA, Wlodarski T, Wentink AS, Wang X, et al. 2016. Structural characterization of the interaction of α-synuclein nascent chains with the ribosomal surface and trigger factor. *PNAS* 113:5012–17
65. Lange S, Franks WT, Rajagopalan N, Döring K, Geiger MA, et al. 2016. Structural analysis of a signal peptide inside the ribosome tunnel by DNP MAS NMR. *Sci. Adv.* 2:e1600379
66. Tian P, Steward A, Kudva R, Su T, Shilling PJ, et al. 2018. Folding pathway of an Ig domain is conserved on and off the ribosome. *PNAS* 115:E11284–93
67. Deeng J, Chan KY, van der Sluis EO. 2016. Dynamic behavior of trigger factor on the ribosome. *J. Mol. Biol.* 428:3588–602
68. Zhang Y, Ma C, Yuan Y, Zhu J, Li N, et al. 2014. Structural basis for interaction of a cotranslational chaperone with the eukaryotic ribosome. *Nat. Struct. Mol. Biol.* 21:1042–46
69. Schaffitzel C, Oswald M, Berger I, Ishikawa T. 2006. Structure of the *E. coli* signal recognition particle bound to a translating ribosome. *Nature* 444:503–6
70. Becker T, Bhushan S, Jarasch A, Armache J-P, Funes S, et al. 2009. Structure of monomeric yeast and mammalian Sec61 complexes interacting with the translating ribosome. *Science* 326:1369–73

71. Frauenfeld J, Gumbart J, van der Sluis EO, Funes S, Gartmann M, et al. 2011. Cryo-EM structure of the ribosome-SecYE complex in the membrane environment. *Nat. Struct. Mol. Biol.* 18:614–21
72. Voorhees RM, Fernandez IS, Scheres SHW, Hegde RS. 2014. Structure of the mammalian ribosome-Sec61 complex to 3.4 Å resolution. *Cell* 157:1632–43
73. Gogala M, Becker T, Beatrix B, Armache J-P, Barrio-Garcia C, et al. 2014. Structures of the Sec61 complex engaged in nascent peptide translocation or membrane insertion. *Nature* 506:107–10
74. Bischoff L, Berninghausen O, Beckmann R. 2014. Molecular basis for the ribosome functioning as an L-tryptophan sensor. *Cell Rep.* 9:469–75
75. Bhushan S, Hoffmann T, Seidelt B, Frauenfeld J, Mielke T, et al. 2011. SecM-stalled ribosomes adopt an altered geometry at the peptidyl transferase center. *PLOS Biol.* 9:e1000581
76. Shanmuganathan V, Schiller N, Magoulopoulou A, Cheng J, Braunger K, et al. 2019. Structural and mutational analysis of the ribosome-arresting human XBP1u. *eLife* 8:e46267
77. Matheisl S, Berninghausen O, Becker T, Beckmann R. 2015. Structure of a human translation termination complex. *Nucleic Acids Res.* 43:8615–26
78. Javed A, Cabrita LD, Cassaignau AME, Wlodarski T, Christodoulou J, Orlova EV. 2019. Visualising nascent chain dynamics at the ribosome exit tunnel by cryo-electron microscopy. bioRxiv 722611. **https://doi.org/10.1101/722611**
79. Mulder FAA, Bouakaz L, Lundell A, Venkataramana M, Liljas A, et al. 2004. Conformation and dynamics of ribosomal stalk protein L12 in solution and on the ribosome. *Biochemistry* 43:5930–36
80. Christodoulou J, Larsson G, Fucini P, Connell SR, Pertinhez TA, et al. 2004. Heteronuclear NMR investigations of dynamic regions of intact *Escherichia coli* ribosomes. *PNAS* 101:10949–54
81. Wang X, Kirkpatrick JP, Launay HMM, de Simone A, Häussinger D, et al. 2019. Probing the dynamic stalk region of the ribosome using solution NMR. *Sci. Rep.* 9:567–69
82. Eichmann C, Preissler S, Riek R, Deuerling E. 2010. Cotranslational structure acquisition of nascent polypeptides monitored by NMR spectroscopy. *PNAS* 107:9111–16
83. Rutkowska A, Beerbaum M, Rajagopalan N, Fiaux J, Schmieder P, et al. 2009. Large-scale purification of ribosome-nascent chain complexes for biochemical and structural studies. *FEBS Lett.* 583:2407–13
84. Chan SHS, Waudby CA, Cassaignau AME, Cabrita LD, Christodoulou J. 2015. Increasing the sensitivity of NMR diffusion measurements by paramagnetic longitudinal relaxation enhancement, with application to ribosome–nascent chain complexes. *J. Biomol. NMR* 63:151–63
85. Deleted in proof
86. Bock LV, Kolář MH, Grubmüller H. 2018. Molecular simulations of the ribosome and associated translation factors. *Curr. Opin. Struct. Biol.* 49:27–35
87. O'Brien EP, Vendruscolo M, Dobson CM. 2012. Prediction of variable translation rate effects on cotranslational protein folding. *Nat. Commun.* 3:868
88. O'Brien EP, Ciryam P, Vendruscolo M, Dobson CM. 2014. Understanding the influence of codon translation rates on cotranslational protein folding. *Acc. Chem. Res.* 47:1536–44
89. Gumbart J, Trabuco LG, Schreiner E, Villa E, Schulten K. 2009. Regulation of the protein-conducting channel by a bound ribosome. *Structure* 17:1453–64
90. Camilloni C, Cavalli A, Vendruscolo M. 2013. Replica-averaged metadynamics. *J. Chem. Theory Comput.* 9:5610–17
91. Shen Y, Lange O, Delaglio F, Rossi P, Aramini JM, et al. 2008. Consistent blind protein structure generation from NMR chemical shift data. *PNAS* 105:4685–90
92. Clore GM, Iwahara J. 2009. Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes. *Chem. Rev.* 109:4108–39
93. Robustelli P, Kohlhoff K, Cavalli A, Vendruscolo M. 2010. Using NMR chemical shifts as structural restraints in molecular dynamics simulations of proteins. *Structure* 18:923–33
94. Sanbonmatsu KY, Tung CS. 2007. High performance computing in biology: multimillion atom simulations of nanoscale systems. *J. Struct. Biol.* 157:470–80
95. Trovato F, O'Brien EP. 2016. Insights into cotranslational nascent protein behavior from computer simulations. *Annu. Rev. Biophys.* 45:345–69

95a. Senior AW, Evans R, Jumper J, Kirkpatrick J, Sifre L, et al. 2020. Improved protein structure prediction using potentials from deep learning. *Nature* 577:706–10
96. Liu K, Rehfus JE, Mattson E, Kaiser CM. 2017. The ribosome destabilizes native and non-native structures in a nascent multidomain protein. *Protein Sci.* 26:1439–51
97. Toal S, Schweitzer-Stenner R. 2014. Local order in the unfolded state: conformational biases and nearest neighbor interactions. *Biomolecules* 4:725–73
98. Robustelli P, Piana S, Shaw DE. 2018. Developing a molecular dynamics force field for both folded and disordered protein states. *PNAS* 115:E4758–66
99. Bowler BE. 2012. Residual structure in unfolded proteins. *Curr. Opin. Struct. Biol.* 22:4–13
100. Waudby CA, Wlodarski T, Karyadi M-E, Cassaignau AME, Chan SHS, et al. 2018. Systematic mapping of free energy landscapes of a growing filamin domain during biosynthesis. *PNAS* 115:9744–49
101. Neira JL, Fersht AR. 1999. Exploring the folding funnel of a polypeptide chain by biophysical studies on protein fragments. *J. Mol. Biol.* 285:1309–33
102. de Prat Gay G, Ruiz-Sanz J, Neira JL, Corrales FJ, Otzen DE, et al. 1995. Conformational pathway of the polypeptide chain of chymotrypsin inhibitor-2 growing from its N terminus in vitro. Parallels with the protein folding pathway. *J. Mol. Biol.* 254:968–79
103. Cruz-Vera LR, Rajagopal S, Squires C, Yanofsky C. 2005. Features of ribosome-peptidyl-tRNA interactions essential for tryptophan induction of *tna* operon expression. *Mol. Cell* 19:333–43
104. Vázquez-Laslop N, Thum C, Mankin AS. 2008. Molecular mechanism of drug-dependent ribosome stalling. *Mol. Cell* 30:190–202
105. Youngman EM, Brunelle JL, Kochaniak AB, Green R. 2004. The active site of the ribosome is composed of two layers of conserved nucleotides with distinct roles in peptide bond formation and peptide release. *Cell* 117:589–99
106. Knight AM, Culviner PH, Kurt-Yilmaz N, Zou T, Ozkan SB, Cavagnero S. 2013. Electrostatic effect of the ribosomal surface on nascent polypeptide dynamics. *ACS Chem. Biol.* 8:1195–204
107. Record MT Jr., Courtenay ES, Cayley DS, Guttman HJ. 1998. Responses of *E. coli* to osmotic stress: large changes in amounts of cytoplasmic solutes and water. *Trends Biochem. Sci.* 23:143–48
108. Fedyukina DV, Jennaro TS, Cavagnero S. 2014. Charge segregation and low hydrophobicity are key features of ribosomal proteins from different organisms. *J. Biol. Chem.* 289:6740–50
109. Selmer M, Dunham CM, Murphy FV IV, Weixlbaumer A, Petry S, et al. 2006. Structure of the 70*S* ribosome complexed with mRNA and tRNA. *Science* 313:1935–42
110. Nierhaus KH. 2014. Mg^{2+}, K^{+}, and the ribosome. *J. Bacteriol.* 196:3817–19
111. Chadani Y, Niwa T, Izumi T, Sugata N, Nagao A, et al. 2017. Intrinsic ribosome destabilization underlies translation and provides an organism with a strategy of environmental sensing. *Mol. Cell* 68:528–39
112. Turnock G, Birch B. 1973. Binding of putrescine and spermidine to ribosomes from *Escherichia coli*. *FEBS J.* 33:467–74
113. Schavemaker PE, Śmigiel WM, Poolman B. 2017. Ribosome surface properties may impose limits on the nature of the cytoplasmic proteome. *eLife* 6:21
114. DeMott CM, Majumder S, Burz DS, Reverdatto S, Shekhtman A. 2017. Ribosome mediated quinary interactions modulate in-cell protein activities. *Biochemistry* 56:4117–26
115. Peterson JH, Woolhead CA, Bernstein HD. 2010. The conformation of a nascent polypeptide inside the ribosome tunnel affects protein targeting and protein folding. *Mol. Microbiol.* 78:203–17
116. Kudva R, Tian P, Pardo-Avila F, Carroni M, Best RB, et al. 2018. The shape of the bacterial ribosome exit tunnel affects cotranslational protein folding. *eLife* 7:19
117. Waudby CA, Dobson CM, Christodoulou J. 2019. Nature and regulation of protein folding on the ribosome. *Trends Biochem. Sci.* 44:P914–26
118. Wright CF, Teichmann SA, Clarke J, Dobson CM. 2005. The importance of sequence diversity in the aggregation and evolution of proteins. *Nature* 438:878–81
119. Han J-H, Batey S, Nickson AA, Teichmann SA, Clarke J. 2007. The folding and evolution of multidomain proteins. *Nat. Rev. Mol. Cell Biol.* 8:319–30
120. Borgia MB, Borgia A, Best RB, Steward A, Nettels D, et al. 2011. Single-molecule fluorescence reveals sequence-specific misfolding in multidomain proteins. *Nature* 474:662–65

121. Liu K, Maciuba K, Kaiser CM. 2019. The ribosome cooperates with a chaperone to guide multi-domain protein folding. *Mol. Cell* 74:310–319.e7
122. Chow CC, Chow C, Raghunathan V, Huppert TJ, Kimball EB, Cavagnero S. 2003. Chain length dependence of apomyoglobin folding: structural evolution from misfolded sheets to native helices. *Biochemistry* 42:7090–99
123. Rüdiger S, Germeroth L, Schneider-Mergener J, Bukau B. 1997. Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. *EMBO J.* 16:1501–7
124. Su Y, Zou Z, Feng S, Zhou P, Cao L. 2007. The acidity of protein fusion partners predominantly determines the efficacy to improve the solubility of the target proteins expressed in *Escherichia coli*. *J. Biotechnol.* 129:373–82
125. Plückthun A. 2011. Ribosome display: a perspective. In *Ribosome Display and Related Technologies*, ed. JA Douthwaite, RH Jackson, pp. 3–28. Methods Mol. Biol. Ser. 805. New York: Springer
126. Camps M, Herman A, Loh E, Loeb LA. 2007. Genetic constraints on protein evolution. *Crit. Rev. Biochem. Mol. Biol.* 42:313–26
127. Segev N, Gerst JE. 2018. Specialized ribosomes and specific ribosomal protein paralogs control translation of mitochondrial proteins. *J. Cell Biol.* 217:117–26
128. Melnikov S, Manakongtreecheep K, Söll D. 2018. Revising the structural diversity of ribosomal proteins across the three domains of life. *Mol. Biol. Evol.* 35:1588–98
129. Petrov AS, Wood EC, Bernier CR, Norris AM, Brown A, Amunts A. 2019. Structural patching fosters divergence of mitochondrial ribosomes. *Mol. Biol. Evol.* 36:207–19
130. Chin JW. 2017. Expanding and reprogramming the genetic code. *Nature* 550:53–60
131. Neumann H, Wang K, Davis L, Garcia-Alai M, Chin JW. 2010. Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature* 464:441–44
132. Sander IM, Chaney JL, Clark PL. 2014. Expanding Anfinsen's principle: contributions of synonymous codon selection to rational protein design. *J. Am. Chem. Soc.* 136:858–61
133. Zhang G, Ignatova Z. 2011. Folding at the birth of the nascent chain: coordinating translation with co-translational folding. *Curr. Opin. Struct. Biol.* 21:25–31
134. Kurland CG. 1993. Major codon preference: theme and variations. *Biochem. Soc. Trans.* 21:841–46
135. Zhang G, Hubalewska M, Ignatova Z. 2009. Transient ribosomal attenuation coordinates protein synthesis and co-translational folding. *Nat. Struct. Mol. Biol.* 16:274–80
136. Cortazzo P, Cerveñansky C, Marín M, Reiss C, Ehrlich R, Deana A. 2002. Silent mutations affect in vivo protein folding in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* 293:537–41
137. Walsh IM, Bowman MA, Soto Santarriaga IF, Rodriguez A, Clark PL. 2020. Synonymous codon substitutions perturb cotranslational protein folding in vivo and impair cell fitness. *PNAS* 117:3528–34
138. Zhou M, Guo J, Cha J, Chae M, Chen S, et al. 2013. Non-optimal codon usage affects expression, structure and function of clock protein FRQ. *Nature* 495:111–15
139. Buhr F, Jha S, Thommen M, Mittelstaet J, Kutz F, et al. 2016. Synonymous codons direct cotranslational folding toward different protein conformations. *Mol. Cell* 61:341–51
140. Jacobs WM, Shakhnovich EI. 2017. Evidence of evolutionary selection for cotranslational folding. *PNAS* 114:11434–39
141. Pechmann S, Frydman J. 2013. Evolutionary conservation of codon optimality reveals hidden signatures of cotranslational folding. *Nat. Struct. Mol. Biol.* 20:237–43
142. Sørensen MA, Pedersen S. 1991. Absolute in vivo translation rates of individual codons in *Escherichia coli*. *J. Mol. Biol.* 222:265–80
143. Gardin J, Yeasmin R, Yurovsky A, Cai Y, Skiena S, et al. 2014. Measurement of average decoding rates of the 61 sense codons in vivo. *eLife* 3:e03735
144. Nissley DA, Sharma AK, Ahmed N, Friedrich UA, Kramer G, et al. 2016. Accurate prediction of cellular co-translational folding indicates proteins can switch from post- to co-translational folding. *Nat. Commun.* 7:10341
145. Kimchi-Sarfaty C, Oh JM, Kim I-W, Sauna ZE, Calcagno AM, et al. 2007. A "silent" polymorphism in the *MDR1* gene changes substrate specificity. *Science* 315:525–28

146. Tsai C-J, Sauna ZE, Kimchi-Sarfaty C, Ambudkar SV, Gottesman MM, Nussinov R. 2008. Synonymous mutations and ribosome stalling can lead to altered folding pathways and distinct minima. *J. Mol. Biol.* 383:281–91
147. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. *PNAS* 98:10037–41
148. Aitken CE, Puglisi JD. 2010. Following the intersubunit conformation of the ribosome during translation in real time. *Nat. Struct. Mol. Biol.* 17:793–800
149. Chen J, Tsai A, Petrov A, Puglisi JD. 2012. Nonfluorescent quenchers to correlate single-molecule conformational and compositional dynamics. *J. Am. Chem. Soc.* 134:5734–37
150. Prabhakar A, Puglisi EV, Puglisi JD. 2019. Single-molecule fluorescence applied to translation. *Cold Spring Harb. Perspect. Biol.* 11:a032714
151. Nicola AV, Chen W, Helenius A. 1999. Co-translational folding of an alphavirus capsid protein in the cytosol of living cells. *Nat. Cell Biol.* 1:341–45
152. Selenko P, Wagner G. 2007. Looking into live cells with in-cell NMR spectroscopy. *J. Struct. Biol.* 158:244–53

Annual Review of Biochemistry

Volume 89, 2020

Contents

Errata

An online log of corrections to *Annual Review of Biochemistry* articles may be found at http://www.annualreviews.org/errata/biochem