Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Navigating the complexities of multi-domain protein folding

Nandakumar Rajasekaran[1] and Christian M. Kaiser[2,3]

Abstract

Proteome complexity has expanded tremendously over evolutionary time, enabling biological diversification. Much of this complexity is achieved by combining a limited set of structural units into long polypeptides. This widely used evolutionary strategy poses challenges for folding of the resulting multi-domain proteins. As a consequence, their folding differs from that of small single-domain proteins, which generally fold quickly and reversibly. Co-translational processes and chaperone interactions are important aspects of multi-domain protein folding. In this review, we discuss some of the recent experimental progress toward understanding these processes.

Addresses

[1] CMDB Graduate Program, Johns Hopkins University, Baltimore, MD, United States

[2] Department of Biology, Johns Hopkins University, Baltimore, MD, United States

[3] Bijvoet Center for Biomolecular Research, Utrecht University, Utrecht, Netherlands

Corresponding author: Kaiser, Christian M. (c.m.kaiser@uu.nl)

Current Opinion in Structural Biology 2024, **86**:102790

This review comes from a themed issue on **Folding and Binding (2024)**

Edited by **H. Jane Dyson** and **Peter E. Wright**

For a complete overview see the Issue and the Editorial

Available online xxx

https://doi.org/10.1016/j.sbi.2024.102790

Introduction

A drastic expansion in proteome complexity over evolutionary time has fueled biological diversification. Along the tree of life, the number of proteins per proteome has increased by more than two orders of magnitude [1]. The number of unique structural units, or domains, expanded less than fivefold from the simplest prokaryotes (archaea and bacteria) to humans, whereas the number of multi-domain proteins increased approximately 50-fold, yielding many novel domain combinations [1]. The enormous diversity across extant proteomes is thus achieved with a relatively small number of unique domains [2,3], which can be viewed as the "words" that make up multi-domain protein "sentences" [4]. Stringing together several domains into one long polypeptide appears evolutionarily favorable. For instance, synthesis as one long polypeptide ensures exactly defined stoichiometry, enables physical linkage of distinct activities in novel combinations without the need for specific binding interfaces, and — by creating very high local concentrations of interacting domains — facilitates tuning of interaction dynamics and allosteric regulation.

While clearly a successful and prevalent evolutionary strategy, combining domains into a single polypeptide is constrained by the ability of the resulting multi-domain proteins to adopt their native, functional structures. The number of possible non-native interactions within a polypeptide chain increases rapidly with its length, resulting in highly frustrated folding energy landscapes [5]. Even in the relatively simple case of tandem repeat proteins with largely independent domains in a "beads-on-a-string" configuration, non-native domain—domain interactions frustrate productive folding [6]. Organizing individual polypeptides into homo- and heterooligomeric complexes is prevalent [7] and circumvents some of the foldability problems. Both folding of multi-domain proteins and assembly of protein complexes cannot be separated from cellular protein biogenesis processes, including translation and protein quality control. Impressive advances in understanding oligomeric complex assembly have recently been reviewed elsewhere [8,9]. This short review focuses on the folding of multi-domain proteins. (For the sake of conciseness, we mostly use the term 'folded' here to refer to states with native tertiary structures and add appropriate qualifiers when discussing other states, e.g. off-pathway structures or secondary structure within the ribosome exit tunnel.) Globular multi-domain proteins, which feature extensive inter-domain interfaces that are buried in the native structure, are highly prone to misfolding [10,11] and aggregation [12] *in vitro*, illustrating the importance of cellular context for understanding their folding.

It is technically challenging to measure multi-domain protein folding, and comparatively few examples have been studied so far. Interactions among non-native domains within one polypeptide commonly give rise to off-

pathway states that compete with folding [13–16]. How, then, do multi-domain proteins fold into their native structures efficiently and robustly? Here, we discuss some of the recent progress that brought us closer to understanding multi-domain protein folding in a biological context. While folding studies of membrane proteins, which constitute one third of the proteome, have led to exciting advances [17], including for large multi-domain proteins [18], we limit our discussion to soluble proteins. A comprehensive review of co-translational folding is beyond the scope of this article, and recent reviews cover this topic from various angles [8,19–26]. We emphasize recent progress and open questions, focusing on co-translational folding and chaperone function, which we view as key aspects that have enabled the evolution of functional multi-domain proteins. We first discuss interactions of the nascent protein with itself and how they form during protein synthesis, then touch upon interactions with the ribosome and molecular chaperones that facilitate native structure formation.

Coupling of protein synthesis and folding

Co-translational folding is a key aspect of multi-domain protein folding. As the ribosome sequentially adds amino acids, the growing nascent polypeptide emerges vectorially from N- to C-terminus and begins to form tertiary structures, domain by domain. Beyond this segmentation of folding, translation elongation can be coupled to folding kinetically. During active elongation, the calcium binding protein calerythrin does not access folded and misfolded conformations that readily form under equilibrium conditions [27]. This observation suggests non-equilibrium effects resulting from the coupling of nascent chain elongation and folding. Non-equilibrium effects might explain why the folding pathways of at least some proteins (including HemK [28], gamma-B crystallin [29], and HaloTag [30]) are profoundly affected by translation elongation. Due to the difficulty of studying non-equilibrium phenomena in co-translational folding, their prevalence and mechanistic basis remains very poorly understood.

Multiple studies have established the impact of synonymous changes in the coding sequence on protein synthesis and folding [31] as well as their potentially deleterious consequences [32,33]. For instance, silent polymorphisms in the coding sequence for the cystic fibrosis transmembrane conductance regulator interfere with the protein's functionality [34]. Elongation rate changes caused by synonymous changes [35,36] have been demonstrated to not only change the folding efficiency, but also to potentially alter the final conformation [29,37–39]. Advances in biophysical approaches have made it possible to detect the resulting structural differences in a few cases [29,39], demonstrating that changes in co-translational folding can elicit phenotypic changes. To establish general rules, novel methods are needed that preserve native translation kinetics and resolve the coupling of elongation and folding.

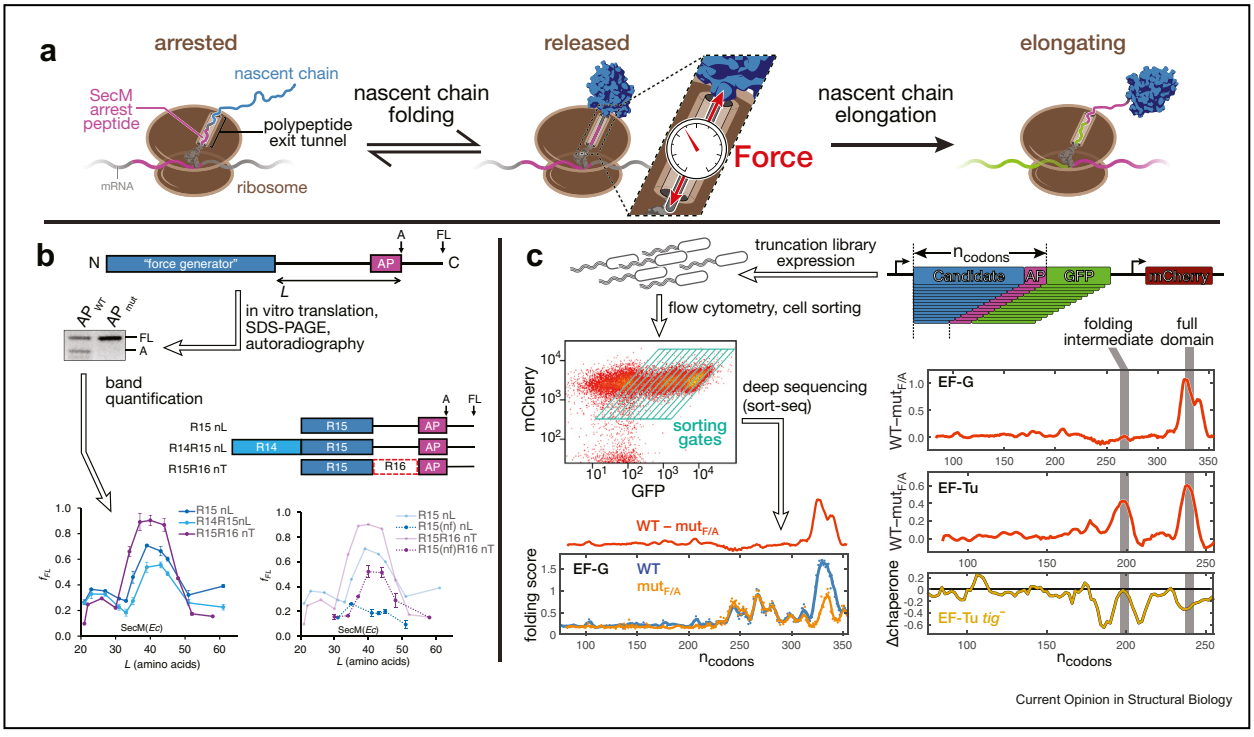
In addition to codon choice, interactions of the nascent protein with the tunnel modulate elongation rate kinetics. Arrest peptides (APs) [40] are an extreme example of this phenomenon. APs interact with the ribosome tunnel to reversibly inhibit elongation. Amino acid substitutions can modulate arrest strength [41,42], illustrating that arrest strength is not a binary parameter, but continuously tunable. While strongly arresting sequence motifs are eliminated from most proteins by negative selection to ensure efficient protein synthesis [35], sequences that slow down elongation to varying degrees (at either the nucleic acid or polypeptide level) are likely prevalent in proteomes [43–45]. It is intriguing to speculate that weak, strategically placed arrest motifs afford a form of elongation rate control that can be feedback-regulated by nascent chain folding (see following paragraph). However, such a mechanism awaits experimental validation.

Nascent protein folding in the cell

Recent studies have harnessed the SecM AP [46], which serves a regulatory function in bacteria, for monitoring nascent chain folding *in vivo* [47–51]. Folding of a protein domain near the exit tunnel generates mechanical force (Figure 1(a)) that disrupts the interactions of the arrest peptide with the ribosome and allows elongation to resume [47]. Fusing the 17 amino acid SecM AP to the C-terminus of a candidate protein thus creates a sensor for nascent protein folding that can be read out by monitoring release rates at distinct nascent chain lengths, yielding a codon-resolved map of co-translational folding events. The approach has so far mainly been used for *in vitro* folding studies (see, e.g., Refs. [52–55]), for instance to dissect the co-translational folding of spectrin repeats [53] (Figure 1(b)). However, bacterial *in vitro* translation systems are inefficient at translating long, multi-domain proteins, limiting the AP approach to relatively short proteins. Moreover, the complex network of nascent chain interactions with cellular factors (see below) cannot be reconstituted *in vitro*.

AP-based folding detection has recently been adapted to probe folding waypoints in live *Escherichia coli* cells, using fluorescent or luminescent reporters [48,51] (Figure 1(c)). One such study indicated that the 294 amino acid long GTPase domain (G-domain) of elongation factor G (EF-G) remains unfolded until the entire domain is fully extruded from the ribosome [48]. The absence of detectable stable tertiary structure in an almost 300 amino acid long polypeptide suggests that the nascent chain close to the ribosome behaves differently from the isolated polypeptide, perhaps mirroring the

Figure 1



AP-based mapping co-translational folding *in vitro* and *in vivo*. (a)**.** The SecM arrest peptide stalls elongation. Arrest is released by mechanical force, which can be generated by nascent chain folding, and elongation restarts. (**b).** AP measurement *in vitro*. Folding-mediated arrest release during *in vitro* translation is measured as the fraction of full-length protein (f_{FL}) after quantification of arrested and full-length translation products (A and FL) by autoradiography. Folding of spectrin repeat 15 (R15) is analyzed at several lengths L. L describes the separation between R15 and the translation arrest position. f_{FL} for R15 is maximal near L = 40 amino acids, in line with folding outside the exit tunnel. The R15 f_{FL} does not change much when the upstream repeat R14 is present ('R15 nL' vs. 'R14R15 nL'). However, the identity of the downstream sequence (spanning the exit tunnel) has a profound effect on the measured signal ('nL' vs 'nT'), even for a non-folding version of R15 ('R15(nf)'), indicating that the helical segment of R16 in the 'nT' constructs stabilizes R15 and/or accelerates arrest release by secondary structure folding. Modified from Ref. [53]. (**c).** AP profiling *in vivo*. Detection of arrest release by fluorescent reporter expression enables multiplexed co-translational folding measurements in *E. coli* cells. After expressing a truncation library of the candidate protein, cells are sorted by ratiometric flow cytometry, and the distribution of each truncation variant across sorting gates is determined by deep sequencing. A 'folding score' for each construct is calculated from the fluorescence ratio of AP-regulated GFP and constitutively expressed mCherry. Several folding score peaks are observed along the G-domain region of the EF-G ORF (note that the x-axis is different from that in panel b). Some of the peaks remain unchanged in a mutant with destabilized tertiary structure ($mut_{F/A}$) and thus likely result from secondary structure formation. The difference between wild-type and mutant ('WT−$mut_{F/A}$') measures stable tertiary structure formation. For EF-G, a single tertiary structure peak is apparent upon full extrusion of the G-domain from the exit tunnel. In contrast, the G-domain from EF-Tu exhibits peaks for both the full domain and a folding intermediate. Genetic ablation of trigger factor results in changes that are mainly localized to the region around the folding intermediate (EF-Tu *tig*$^{-}$). Modified from ref. [51].

observations for calerythrin [27]. Folding commences immediately upon full extrusion from the ribosome, suggesting that folding initiates at the C-terminus. The C-terminal alpha helix of the EF-G G-domain is mostly buried in the native structure, in part due to the presence of a subdomain, termed G′. The resulting structure might necessitate a folding pathway in which the C-terminal helix forms early and the enclosing regions fold around it. The G-domain of EF-G might thus provide an example of how a folding pathway is dictated by a particular native topology or structure [48].

It is often difficult to generalize conclusions from folding studies of one or a few specific protein(s). At the same time, most experimental approaches currently available for studying co-translational folding are not scalable, precluding high-throughput formats. By combining cellular AP-based detection of co-translational folding with cell sorting and deep sequencing, a large number of constructs can be screened. This approach, termed AP profiling, enables mapping of co-translational folding waypoints at a larger scale with codon resolution [51] (Figure 1(c)). Application to translation factor GTPases showed domain-wise folding of EF-G and related proteins. The G-domain of elongation factor Tu (EF-Tu), which is homologous to EF-G, populates a co-translational folding intermediate [51]. In this case, the C-terminal alpha-helix is exposed to the solvent, consistent with the conjecture that the folding pathway is matched to the native structure. More studies are needed to corroborate this concept. AP profiling is bringing into reach studies at the proteome scale. Obtaining information on the folding pathways of multiple related domains or proteins

will likely pave the way for extracting general principles of co-translational folding and might, combined with ancestral sequence reconstruction [56], shed light on the evolution of (co-translational) folding pathways.

One caveat regarding AP measurements *in vitro* and *in vivo* is that mechanisms of arrest release remain incompletely understood. Stably folded globular domains in the nascent chain near the ribosome elicit arrest release, perhaps by generating 'entropic force' [47]. However, secondary structures and possibly other features in the nascent protein also affect arrest strength [54]. The identity of the linker sequence upstream of the AP profoundly affects arrest release in some contexts (compare the 'nL' and 'nT' constructs in the bottom right panel in Figure 1(b) [53]), highlighting the importance of experimental design for interpreting AP measurements. Perturbing tertiary structure formation by replacing phenylalanine with alanine in a globular domain leaves some detected folding events unaffected (Figure 1(c)) [51], suggesting that secondary structures can overcome arrest. Secondary structure formation inside the polypeptide exit tunnel has indeed been mapped using AP measurements [54]. Whether secondary structures accelerate arrest release via force generation and to what degree arrest release is modulated by the identity of AP-adjacent residues remains to be determined. Regardless of the underlying mechanisms, appropriately controlled AP experiments have tremendous potential for probing co-translational folding processes.

Native domain–domain interactions

As outlined above, domain-wise co-translational folding avoids frustration resulting from accumulation of unfolded polypeptide. In addition, contacts between native domains can be required for energetic stabilization. How energetic coupling affects the overall folding pathway of multi-domain proteins has been investigated for EF-G. Its second domain (domain II) requires native contacts with the natively folded N-terminal G-domain to be stably structured [16], imposing a folding pathway whose order matches the order of synthesis. In the C-terminal half of EF-G, however, folding runs counter to synthesis, because the third domain (domain III) requires native contacts with its C-terminal neighbors (domains IV and V) for stability, perhaps reflecting a functional requirement [57]. The arrangement of domains along the primary structure and their interactions may generally evolve to meet the requirements of both folding and function. Energetic dependencies among domains are thus an important yet understudied aspect of multi-domain protein folding.

Nascent chain-chaperone interactions

Specialized molecular chaperones in bacteria and eukaryotes directly bind to the ribosome near the polypeptide exit tunnel and engage with nascent proteins [8,58]. The vicinity of the exit tunnel also serves as a docking site for several other proteins involved in nascent chain processing and targeting [8,59], as well as surveillance and quality control [60,61]. The bacterial chaperone trigger factor directly binds to ribosomal protein L23 near the exit tunnel, positioning it for engagement with emerging nascent proteins [8,62]. NMR experiments with ribosome-nascent chain complexes showed that trigger factor binds weakly to a relatively short, disordered nascent chain, mostly through positively charged residues that also bind to the ribosome [63,64]. The chaperone has several distinct binding sites that interact with multiple short segments in an isolated (non-ribosome-bound) client protein [65]. However, trigger factor may not interact with all client proteins in the same manner. In measurements with an unfolded nascent chain, it appears to stabilize the unfolded state, functioning as a holdase [63,64]. For other client proteins, the chaperone binds fully or partially folded structures [66,67] or even promotes folding [68,69]. Another trigger factor function was observed in a recent study of EF-G. Interactions with a nascent, still unfolded domain can denature already folded structures [16], complicating co-translational multi-domain protein folding in an unanticipated way. Trigger factor prevented this denaturation, presumably by sequestering unfolded segments. Unfolded polypeptide is inevitably present during translation. One largely unexplored function of ribosome-binding chaperones might be to shield folded domains from denaturing interactions with unfolded polypeptide that is being extruded from the exit tunnel.

Eukaryotes lack a known trigger factor homolog. The heterodimeric ribosome-associated complex (RAC) may be the eukaryotic equivalent of trigger factor as a general chaperone for nascent proteins. RAC is a heterodimer of specialized Hsp70- and Hsp40-type proteins that cooperate with the Hsp70 chaperone Ssb [59]. Recent structural and functional studies illuminated how this chaperone triad engages with nascent proteins [70–72]. Another heterodimer, the nascent chain-associated complex (NAC), coordinates nascent chain processing and targeting [59]. NAC is positioned to scan nascent proteins and helps in the recruitment of the appropriate processing [73] and targeting [74] factors. A parallel between trigger factor and NAC is that they both increase the fidelity of processing and targeting by competing with other factors [73–76], including the signal recognition particle for nascent chain binding. Native folding and proper targeting are thus promoted by nascent chain-binding chaperones (trigger factor in bacteria, NAC and RAC in eukaryotes) whose binding disfavors spurious off-pathway contacts. In addition to these general chaperone systems, eukaryotes may contain specialized chaperones for specific nascent multi-domain proteins, as exemplified by the role of

Hgh1 in eEF2 folding [77]. Few mechanistic studies have been reported to date describing how general or specialized chaperones support nascent multi-domain protein folding in eukaryotes. The biophysical approaches developed for bacterial systems should prove useful for investigating their function and fill this knowledge gap.

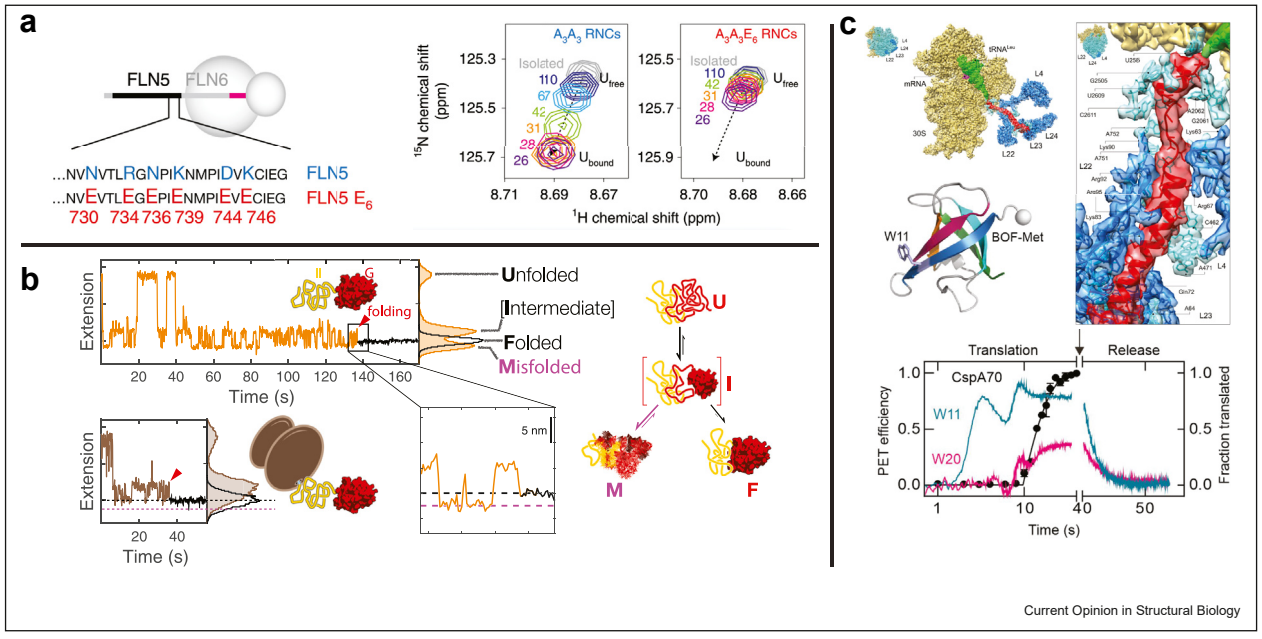
Ribosome interactions

In addition to its function in decoding and catalysis, the ribosome contributes to nascent chain folding in distinct ways. Through a bevy of specific binding sites on its surface, the ribosome recruits factors that interact with the nascent protein (see above). Moreover, the nascent chain itself interacts with the ribosome. Compared to interactions within the exit tunnel that elicit stalling, interactions with the surface are probably less specific. The nascent protein remains tethered to the ribosome via tRNAs throughout elongation. This tethering results in millimolar local concentrations of ribosomal binding sites around the nascent chain [78], so that even weak interactions are significantly populated and impact folding [16,79–83].

The negatively charged ribosome surface interacts differentially with positively and negatively charged nascent chain segments. Basic segments exhibit decreased flexibility, compared with identical free polypeptides in isolation [84]. The conformational constraints on the nascent protein resulting from these interactions slow some, but not all folding steps [79]. In addition to basic residues that interact electrostatically (Figure 2(a)), aromatic side chains also contribute [64,82]. Due to their dependence on local concentration, ribosome interactions are likely strongest for the segments of the nascent chain that are closest to the ribosome surface. Sequestering the proximal, still unfolded nascent chain segment can increase overall folding efficiency by reducing non-productive interactions within the nascent chain [16] (Figure 2(b)). The strength of these interactions changes as the nascent chain becomes longer and longer. How the continuously changing folding landscape relates to overall folding and whether it contributes to non-equilibrium effects remains poorly understood.

The long tunnel running through the large ribosomal subunit also presents numerous sites for nascent chain interactions [85], as illustrated by AP sequences (see above). Engineered changes in the tunnel environment can change the stability of structures formed by the nascent chain [86]. Folding inside the tunnel is possible to a limited degree in an environment that is distinct from that outside the ribosome [26]. Large regions of the tunnel can not only accommodate but stabilize alpha-helical structure in the nascent chain [87,88]. Even segments that form beta-sheets in the native structure were recently found to be helical inside the

Figure 2



Spectroscopic characterization of nascent chain folding. **(a).** NMR spectroscopy resolves nascent chain structure and ribosome interactions, revealing that positively charged nascent chain residues interact with the ribosome surface. Modified from Cassaignau et al. [82] **(b).** Single-molecule force spectroscopy resolves folding pathways kinetically. Ribosome interactions sequester the proximal nascent chain, reducing inter-domain misfolding. Modified from Liu et al. [16]. **(c).** PET spectroscopy resolves nascent chain compaction during early stages of nascent chain elongation, and cryogenic electron microscopy reveals that the all-beta CspA protein forms a helix inside the ribosome exit tunnel. Modified from Agirrezabala et al. [89].

exit tunnel [89] (Figure 2(c)). Whether and how secondary structure formed inside the tunnel can propagate or persist long enough to guide downstream folding steps remains to be investigated. Another (albeit speculative) function of helix formation inside the exit tunnel could be to provide a "mechanical buffer" that insulates the peptidyl transferase center from mechanical forces that generally accompany nascent chain folding [47]. High forces could potentially distort the position of the amino acid on the P-site tRNA and thus reduce peptide bond formation rates. An alpha-helix spanning the exit tunnel would be well-suited to act as a damper. More generally, how the mechanical forces that act across the exit tunnel influence elongation and folding remains to be explored.

Conclusion

Multi-domain proteins are ubiquitous and essential building blocks of life. Understanding how they are synthesized and how they fold may not only result in a better understanding of cellular function, but also help to optimize recombinant expression of biologicals and aid in the design of artificial proteins with novel functions. Several key questions remain to be addressed. How are nascent chain elongation and folding coupled to elicit non-equilibrium effects? What are the relevant regulatory circuits that ensure robust folding, including yet unknown feedback mechanisms, e.g. through mechanical cues? Do nascent chain-binding chaperones generally contribute beyond their classical role of shielding non-native segments from inappropriate interactions? The most promising approach will be to study authentic multi-domain proteins in the context of their native interaction partners. Extant sequences not only encode the native structure, but also the folding pathway that has evolved in the presence of all the interactors that the protein is exposed to in the cell. Combining cellular methods (which capture biological context) with complementary *in vitro* spectroscopic approaches (which inform molecular mechanisms) is a particularly promising route for understanding multi-domain protein folding.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

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

Acknowledgements

This work was supported by a grant from the National Institutes of Health (5R01GM121567) to C.M.K.

References

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

* of special interest
** of outstanding interest

1. Rebeaud ME, Mallik S, Goloubinoff P, Tawfik DS: **On the evolution of chaperones and cochaperones and the expansion of proteomes across the Tree of Life**. *Proc Natl Acad Sci U S A* 2021, **118**.
2. Andreeva A, Kulesha E, Gough J, Murzin AG: **The SCOP database in 2020: expanded classification of representative family and superfamily domains of known protein structures**. *Nucleic Acids Res* 2020, **48**:D376–D382.
3. Sillitoe I, *et al.*: **CATH: increased structural coverage of functional space**. *Nucleic Acids Res* 2021, **49**:D266–D273.
4. Buchan DWA, Jones DT: **Learning a functional grammar of protein domains using natural language word embedding techniques**. *Proteins* 2020, **88**:616–624.
5. Zheng W, Schafer NP, Wolynes PG: **Frustration in the energy landscapes of multidomain protein misfolding**. *Proc Natl Acad Sci U S A* 2013, **110**:1680–1685.
6. Borgia MB, *et al.*: **Single-molecule fluorescence reveals sequence-specific misfolding in multidomain proteins**. *Nature* 2011, **474**:662–665.
7. Hu P, *et al.*: **Global functional atlas of *Escherichia coli* encompassing previously uncharacterized proteins**. *PLoS Biol* 2009, **7**:e96.
8. Kramer G, Shiber A, Bukau B: **Mechanisms of cotranslational maturation of newly synthesized proteins**. *Annu Rev Biochem* 2019, **88**:337–364.
9. Badonyi M, Marsh JA: **Hallmarks and evolutionary drivers of cotranslational protein complex assembly**. *FEBS J* 2023 (Early View. Online Version of Record before inclusion in an issue.).
10. Han JH, Batey S, Nickson AA, Teichmann SA, Clarke J: **The folding and evolution of multidomain proteins**. *Nat Rev Mol Cell Biol* 2007, **8**:319–330.
11. Rajasekaran N, Kaiser CM: **Co-translational folding of multidomain proteins**. *Front Mol Biosci* 2022, **9**:869027.
12. Louros N, Schymkowitz J, Rousseau F: **Mechanisms and pathology of protein misfolding and aggregation**. *Nat Rev Mol Cell Biol* 2023, **24**:912–933.
13. Scholl ZN, Yang W, Marszalek PE: **Chaperones rescue luciferase folding by separating its domains**. *J Biol Chem* 2014, **289**:28607–28618.
14. * Jahn M, Buchner J, Hugel T, Rief M: **Folding and assembly of the large molecular machine Hsp90 studied in single-molecule experiments**. *Proc Natl Acad Sci U S A* 2016, **113**:1232–1237.
Single-molecule experiments on a large multidomain protein show how frustration between domains leads to misfolded states.
15. Garcia-Manyes S, *et al.*: **Single-molecule force spectroscopy predicts a misfolded, domain-swapped conformation in human gammaD-crystallin protein**. *J Biol Chem* 2016, **291**:4226–4235.
16. ** Liu K, Maciuba K, Kaiser CM: **The ribosome cooperates with a chaperone to guide multi-domain protein folding**. *Mol Cell* 2019, **74**:310–319 e7.
This study demonstrates how the ribosome and the chaperone trigger factor together mitigate inter-domain misfolding and domain denaturation in nascent proteins.
17. Hong H, Choi HK, Yoon TY: **Untangling the complexity of membrane protein folding**. *Curr Opin Struct Biol* 2022, **72**:237–247.
18. Im J, *et al.*: **ABC-transporter CFTR folds with high fidelity through a modular, stepwise pathway**. *Cell Mol Life Sci* 2023, **80**:33.

19. Komar AA: **A pause for thought along the co-translational folding pathway**. *Trends Biochem Sci* 2009, **34**:16–24.

20. Waudby CA, Dobson CM, Christodoulou J: **Nature and regulation of protein folding on the ribosome**. *Trends Biochem Sci* 2019, **44**:914–926.

21. Bustamante C, Alexander L, Maciuba K, Kaiser CM: **Single-molecule studies of protein folding with optical tweezers**. *Annu Rev Biochem* 2020, **89**:443–470.

22. Liu Y: **A code within the genetic code: codon usage regulates co-translational protein folding**. *Cell Commun Signal* 2020, **18**:145.

23. Maciuba K, Rajasekaran N, Chen X, Kaiser CM: **Co-translational folding of nascent polypeptides: multi-layered mechanisms for the efficient biogenesis of functional proteins**. *Bioessays* 2021, **43**:2100042.

24. Morales-Polanco F, Lee JH, Barbosa NM, Frydman J: **Cotranslational mechanisms of protein biogenesis and complex assembly in eukaryotes**. *Annu Rev Biomed Data Sci* 2022, **5**:67–94.

25. Moss MJ, Chamness LM, Clark PL: **The effects of codon usage on protein structure and folding**. *Annu Rev Biophys* 2023.

26. Samatova E, Komar AA, Rodnina MV: **How the ribosome shapes cotranslational protein folding**. *Curr Opin Struct Biol* 2023, **84**:102740.

27. ** Alexander LM, Goldman DH, Wee LM, Bustamante C: **Non-equilibrium dynamics of a nascent polypeptide during translation suppress its misfolding**. *Nat Commun* 2019, **10**:2709.
One of the only mechanistic studies of the coupling of elongation and folding, this paper shows how non-equilibrium effects prevent folding and misfolding.

28. Holtkamp W, *et al.*: **Cotranslational protein folding on the ribosome monitored in real time**. *Science* 2015, **350**:1104–1107.

29. Buhr F, *et al.*: **Synonymous codons direct cotranslational folding toward different protein conformations**. *Mol Cell* 2016, **61**:341–351.

30. Samelson AJ, *et al.*: **Kinetic and structural comparison of a protein's cotranslational folding and refolding pathways**. *Sci Adv* 2018, **4**, eaas9098.

31. Rodnina MV, Wintermeyer W: **Protein elongation, co-translational folding and targeting**. *J Mol Biol* 2016, **428**:2165–2185.

32. Sauna ZE, Kimchi-Sarfaty C: **Understanding the contribution of synonymous mutations to human disease**. *Nat Rev Genet* 2011, **12**:683–691.

33. Davyt M, Bharti N, Ignatova Z: **Effect of mRNA/tRNA mutations on translation speed: implications for human diseases**. *J Biol Chem* 2023, **299**:105089.

34. Rauscher R, *et al.*: **Positive epistasis between disease-causing missense mutations and silent polymorphism with effect on mRNA translation velocity**. *Proc Natl Acad Sci U S A* 2021, **118**.

35. Choi J, *et al.*: **How messenger RNA and nascent chain sequences regulate translation elongation**. *Annu Rev Biochem* 2018, **87**:421–449.

36. Liu Y, Yang Q, Zhao F: **Synonymous but not silent: the codon usage code for gene expression and protein folding**. *Annu Rev Biochem* 2021, **90**.

37. Kim SJ, *et al.*: **Protein folding. Translational tuning optimizes nascent protein folding in cells**. *Science* 2015, **348**:444–448.

38. Yu CH, *et al.*: **Codon usage influences the local rate of translation elongation to regulate co-translational protein folding**. *Mol Cell* 2015, **59**:744–754.

39. * Walsh IM, Bowman MA, Soto Santarriaga IF, Rodriguez A, Clark PL: **Synonymous codon substitutions perturb cotranslational protein folding in vivo and impair cell fitness**. *Proc Natl Acad Sci U S A* 2020, **117**:3528–3534.
Synonymous mutations in chloramphenicol acetyl-transferase yield active enzyme but with reduced stability *in vivo*, presumably through an altered folding mechanism.

40. Ito K, Chiba S: **Arrest peptides: cis-acting modulators of translation**. *Annu Rev Biochem* 2013, **82**:171–202.

41. Yap MN, Bernstein HD: **The plasticity of a translation arrest motif yields insights into nascent polypeptide recognition inside the ribosome tunnel**. *Mol Cell* 2009, **34**:201–211.

42. Cymer F, Hedman R, Ismail N, von Heijne G: **Exploration of the arrest peptide sequence space reveals arrest-enhanced variants**. *J Biol Chem* 2015, **290**:10208–10215.

43. Buskirk AR, Green R: **Ribosome pausing, arrest and rescue in bacteria and eukaryotes**. *Philos Trans R Soc Lond B Biol Sci* 2017, **372**.

44. Collart MA, Weiss B: **Ribosome pausing, a dangerous necessity for co-translational events**. *Nucleic Acids Res* 2020, **48**:1043–1055.

45. Samatova E, Daberger J, Liutkute M, Rodnina MV: **Translational control by ribosome pausing in bacteria: how a non-uniform pace of translation affects protein production and folding**. *Front Microbiol* 2021, **11**.

46. Nakatogawa H, Ito K: **The ribosomal exit tunnel functions as a discriminating gate**. *Cell* 2002, **108**:629–636.

47. Goldman DH, *et al.*: **Ribosome. Mechanical force releases nascent chain-mediated ribosome arrest in vitro and in vivo**. *Science* 2015, **348**:457–460.

48. * Chen X, Rajasekaran N, Liu K, Kaiser CM: **Synthesis runs counter to directional folding of a nascent protein domain**. *Nat Commun* 2020, **11**:5096.
In vivo assays combined with single-molecule experiments suggest how co-translational multi-domain protein folding pathways are dictated by the native structure.

49. Elfageih R, *et al.*: **Cotranslational folding of alkaline phosphatase in the periplasm of *Escherichia coli***. *Protein Sci* 2020, **29**:2028–2037.

50. Mermans D, Nicolaus F, Baygin A, von Heijne G: **Cotranslational folding of human growth hormone in vitro and in *Escherichia coli***. *FEBS Lett* 2023, **597**:1355–1362.

51. Chen X, Kaiser CM: **AP profiling resolves co-translational folding pathway and chaperone interactions in vivo**. *bioRxiv* 2023, https://doi.org/10.1101/2023.09.01.555749.

52. Notari L, Martinez-Carranza M, Farias-Rico JA, Stenmark P, von Heijne G: **Cotranslational folding of a pentarepeat beta-helix protein**. *J Mol Biol* 2018, **430**:5196–5206.

53. Kemp G, Nilsson OB, Tian P, Best RB, von Heijne G: **Cotranslational folding cooperativity of contiguous domains of alpha-spectrin**. *Proc Natl Acad Sci U S A* 2020, **117**:14119–14126.

54. * Liutkute M, Maiti M, Samatova E, Enderlein J, Rodnina MV: **Gradual compaction of the nascent peptide during cotranslational folding on the ribosome**. *Elife* 2020, **9**:e60895.
Detailed *in vitro* analysis, combining arrest peptides and spectroscopy assays, show how the exit tunnel environment promotes formation of α-helices and accommodates tertiary structures in the vestibule region.

55. Jensen MK, Samelson AJ, Steward A, Clarke J, Marqusee S: **The folding and unfolding behavior of ribonuclease H on the ribosome**. *J Biol Chem* 2020, **295**:11410–11417.

56. Chisholm LO, Orlandi KN, Phillips SR, Shavlik MJ, Harms MJ: **Ancestral reconstruction and the evolution of protein energy landscapes**. *Annu Rev Biophys* 2023.

57. Liu K, Chen X, Kaiser CM: **Energetic dependencies dictate folding mechanism in a complex protein**. *Proc Natl Acad Sci U S A* 2019, **116**:25641–25648.

58. Deuerling E, Gamerdinger M, Kreft SG: **Chaperone interactions at the ribosome**. *Cold Spring Harb Perspect Biol* 2019.

59. Gamerdinger M, Deuerling E: **Cotranslational sorting and processing of newly synthesized proteins in eukaryotes**. *Trends Biochem Sci* 2023.

60. Sitron CS, Brandman O: **Detection and degradation of stalled nascent chains via ribosome-associated quality control**. *Annu Rev Biochem* 2020, **89**:417–442.

61. Filbeck S, Cerullo F, Pfeffer S, Joazeiro CAP: **Ribosome-associated quality-control mechanisms from bacteria to humans**. *Mol Cell* 2022, **82**:1451–1466.

62. Kramer G, *et al.*: **L23 protein functions as a chaperone docking site on the ribosome**. *Nature* 2002, **419**:171–174.

63. Deckert A, *et al.*: **Structural characterization of the interaction of alpha-synuclein nascent chains with the ribosomal surface and trigger factor**. *Proc Natl Acad Sci U S A* 2016, **113**: 5012–5017.

64. * Deckert A, *et al.*: **Common sequence motifs of nascent chains engage the ribosome surface and trigger factor**. *Proc Natl Acad Sci U S A* 2021, **118**.
NMR spectroscopy experiments identify sequence motifs in nascent α-synuclein that interact with the ribosome and trigger factor.

65. Saio T, Guan X, Rossi P, Economou A, Kalodimos CG: **Structural basis for protein antiaggregation activity of the trigger factor chaperone**. *Science* 2014, **344**, 1250494.

66. Martinez-Hackert E, Hendrickson WA: **Promiscuous substrate recognition in folding and assembly activities of the trigger factor chaperone**. *Cell* 2009, **138**:923–934.

67. Mashaghi A, *et al.*: **Reshaping of the conformational search of a protein by the chaperone trigger factor**. *Nature* 2013, **500**: 98–101.

68. Haldar S, Tapia-Rojo R, Eckels EC, Valle-Orero J, Fernandez JM: **Trigger factor chaperone acts as a mechanical foldase**. *Nat Commun* 2017, **8**:668.

69. Wu K, Minshull TC, Radford SE, Calabrese AN, Bardwell JCA: **Trigger factor both holds and folds its client proteins**. *Nat Commun* 2022, **13**:4126.

70. Zhang Y, *et al.*: **The ribosome-associated complex RAC serves in a relay that directs nascent chains to Ssb**. *Nat Commun* 2020, **11**:1504.

71. Kisonaite M, *et al.*: **Structural inventory of cotranslational protein folding by the eukaryotic RAC complex**. *Nat Struct Mol Biol* 2023, **30**:670–677.

72. Lee K, *et al.*: **Pathway of Hsp70 interactions at the ribosome**. *Nat Commun* 2021, **12**:5666.

73. Gamerdinger M, *et al.*: **NAC controls cotranslational N-terminal methionine excision in eukaryotes**. *Science* 2023, **380**: 1238–1243.

74. Jomaa A, *et al.*: **Mechanism of signal sequence handover from NAC to SRP on ribosomes during ER-protein targeting**. *Science* 2022, **375**:839–844.

75. Bornemann T, Holtkamp W, Wintermeyer W: **Interplay between trigger factor and other protein biogenesis factors on the ribosome**. *Nat Commun* 2014, **5**:4180.

76. Ariosa A, Lee JH, Wang S, Saraogi I, Shan SO: **Regulation by a chaperone improves substrate selectivity during cotranslational protein targeting**. *Proc Natl Acad Sci U S A* 2015, **112**: E3169–E3178.

77. Monkemeyer L, *et al.*: **Chaperone function of Hgh1 in the biogenesis of eukaryotic elongation factor 2**. *Mol Cell* 2019.

78. Waudby CA, Burridge C, Cabrita LD, Christodoulou J: **Thermodynamics of co-translational folding and ribosome-nascent chain interactions**. *Curr Opin Struct Biol* 2022, **74**, 102357.

79. Kaiser CM, Goldman DH, Chodera JD, Tinoco Jr I, Bustamante C: **The ribosome modulates nascent protein folding**. *Science* 2011, **334**:1723–1727.

80. Cabrita LD, *et al.*: **A structural ensemble of a ribosome-nascent chain complex during cotranslational protein folding**. *Nat Struct Mol Biol* 2016, **23**:278–285.

81. Liu K, Rehfus JE, Mattson E, Kaiser CM: **The ribosome destabilizes native and non-native structures in a nascent multidomain protein**. *Protein Sci* 2017, **26**:1439–1451.

82. Cassaignau AME, *et al.*: **Interactions between nascent proteins and the ribosome surface inhibit co-translational folding**. *Nat Chem* 2021, **13**:1214–1220.

83. Chan SHS, *et al.*: **The ribosome stabilizes partially folded intermediates of a nascent multi-domain protein**. *Nat Chem* 2022, **14**:1165–1173.

84. Knight AM, *et al.*: **Electrostatic effect of the ribosomal surface on nascent polypeptide dynamics**. *ACS Chem Biol* 2013, **8**: 1195–1204.

85. Kosolapov A, Deutsch C: **Tertiary interactions within the ribosomal exit tunnel**. *Nat Struct Mol Biol* 2009, **16**:405–411.

86. ** Ahn M, *et al.*: **Modulating co-translational protein folding by rational design and ribosome engineering**. *Nat Commun* 2022, **13**:4243.
Engineered changes in proteins that line the ribosome exit tunnel change the thermodynamics of nascent chain folding, establishing directed changes to the ribosome as a powerful tool for investigating co-translational folding.

87. Woolhead CA, McCormick PJ, Johnson AE: **Nascent membrane and secretory proteins differ in FRET-detected folding far inside the ribosome and in their exposure to ribosomal proteins**. *Cell* 2004, **116**:725–736.

88. Lu J, Deutsch C: **Folding zones inside the ribosomal exit tunnel**. *Nat Struct Mol Biol* 2005, **12**:1123–1129.

89. * Agirrezabala X, *et al.*: **A switch from alpha-helical to beta-strand conformation during co-translational protein folding**. *EMBO J* 2022, **41**, e109175.
Cryogenic electron microscopy reveals that a protein with an all-beta native structure forms helices inside the ribosome exit tunnel, suggesting that the tunnel environment may generally favor alpha-helical structure.