

Complementary Role of Co- and Post-Translational Events in *De Novo* Protein Biogenesis

Rayna M. Addabbo, Matthew D. Dalphin, Miranda F. Mecha, Yue Liu, Alexios Staikos, Valeria Guzman-Luna, and Silvia Cavagnero*



Cite This: *J. Phys. Chem. B* 2020, 124, 6488–6507



Read Online

ACCESS |



Metrics & More

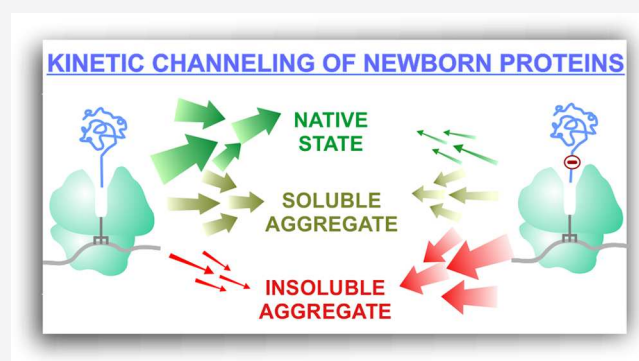


Article Recommendations



Supporting Information

ABSTRACT: The relation between co- and post-translational protein folding and aggregation in the cell is poorly understood. Here, we employ a combination of fluorescence anisotropy decays in the frequency domain, fluorescence-detected solubility assays, and NMR spectroscopy to explore the role of the ribosome in protein folding within a biologically relevant context. First, we find that a primary function of the ribosome is to promote cotranslational nascent-protein solubility, thus supporting cotranslational folding even in the absence of molecular chaperones. Under these conditions, however, only a fraction of the soluble expressed protein is folded and freely tumbling in solution. Hence, the ribosome alone is insufficient to guarantee quantitative formation of the native state of the apomyoglobin (apoMb) model protein. Right after biosynthesis, nascent chains encoding apoMb emerge from the ribosomal exit tunnel and undergo a crucial irreversible post-translational kinetic partitioning between further folding and aggregation. Mutational analysis in combination with protein-expression kinetics and NMR show that nascent proteins can attain their native state only when the relative rates of soluble and insoluble product formation immediately upon release from the ribosome are tilted in favor of soluble species. Finally, the outcome of the above immediately post-translational kinetic partitioning is much more sensitive to amino acid sequence perturbations than the native fold, which is rather mutation-insensitive. Hence, kinetic channeling of nascent-protein conformation upon release from the ribosome may be a major determinant of evolutionary pressure.



INTRODUCTION

The mechanism by which proteins attain their native structure in the cell is currently not well understood. Many proteins of variable size, secondary structure and topology start their folding cotranslationally and complete it posttranslationally, as illustrated in Figure 1a. This conclusion was reached by analyzing ribosome-bound nascent chains (RNCs) via time-resolved fluorescence depolarization,^{1–3} Förster resonance energy transfer (FRET),^{4,5} optical tweezers,⁶ antibody-based assays,^{7–9} protease digestion,^{7,10,11} mass spectrometry,¹⁰ multi-dimensional NMR,^{12–14} computation,¹⁵ and arrest assays coupled with single-particle cryo-EM.^{16,17}

Both native and non-native conformations can be generated cotranslationally.^{4,7} The ribosomal surface and liquid layers in its proximity provide a unique highly polar environment for RNC conformational sampling.^{3,14} Interactions between nascent chains and the ribosome were reported for peptides carrying an N-terminal signal sequence,^{18–21} and they may even occur in the absence of signal sequences.^{14,22} The presence of these interactions is sometimes invoked to rationalize the experimentally observed nascent chain dynamics.^{3,6,14,22}

Despite the presence of cotranslational folding,^{1–14,16,17} typical single-domain nascent proteins populate dynamic incomplete structures when ribosome-bound, and they are not able to attain their native state until they have been fully released from the ribosome (Figure 1a).^{1,4,23–28} It has been proposed that the extent of post-translational protein folding that occurs upon release from the ribosome depends on protein size and other physical properties, including the complexity of the fold and the fraction of residues buried in the ribosome exit tunnel.^{1–3}

Most protein amino acid sequences evolved to fold efficiently²⁹ in response to evolutionary pressure for improved function and higher chances of cell survival. As a result of this driving force, newly synthesized bacterial proteins typically fold without aggregating, *in vivo*. While the cell is equipped with

Received: April 6, 2020

Revised: May 26, 2020

Published: May 26, 2020



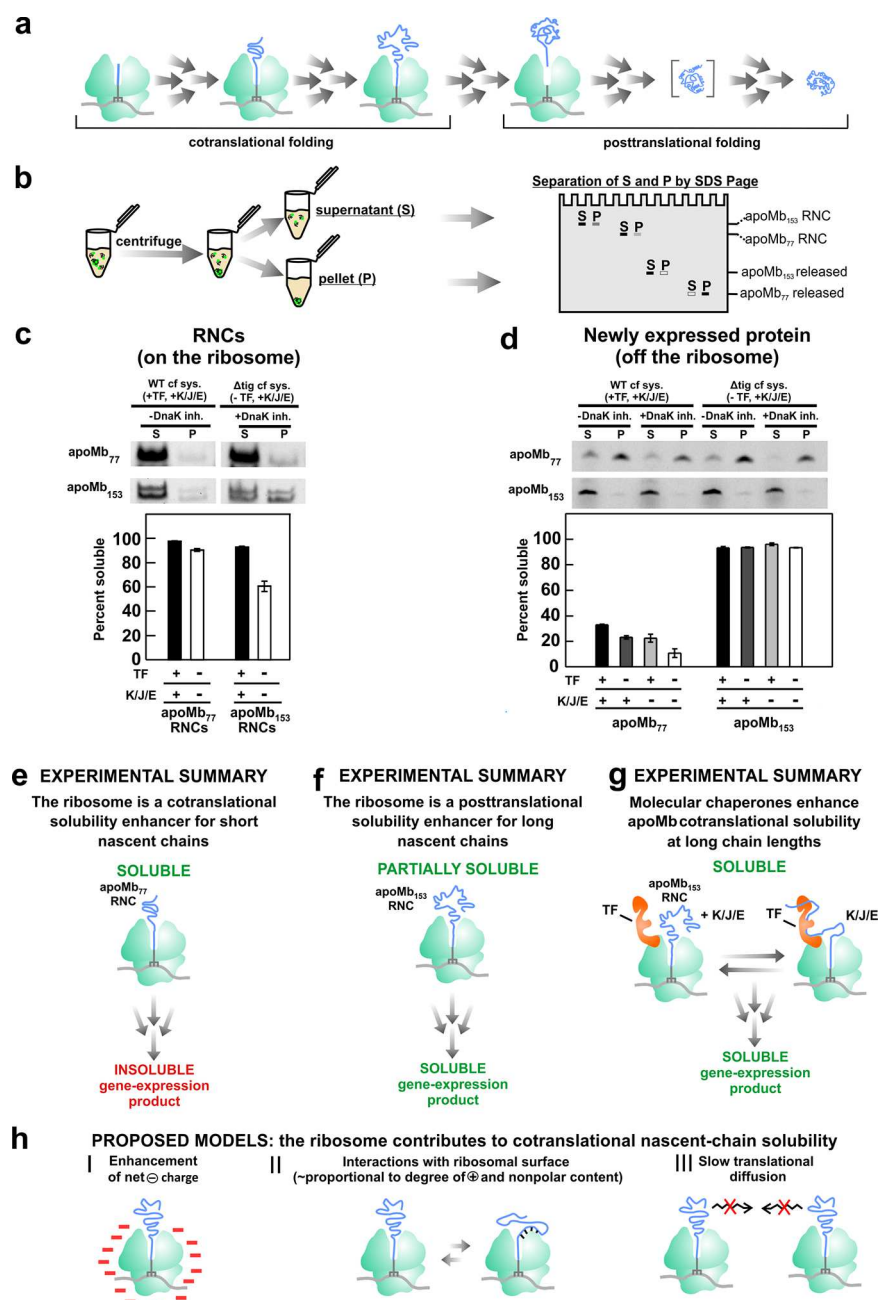


Figure 1. The ribosome is a facilitator of protein solubility. (a) Cartoon representation of typical co- and post-translational folding events for medium-size single-domain proteins. Proteins fold partially cotranslationally as the chain elongates, and complete their folding posttranslationally. The species in square brackets denotes any putative post-translational folding intermediate(s). The nascent chain is not drawn to scale. (b) Cartoon representation of solubility assay employed in this work. N-terminally fluorophore-labeled protein-containing solutions are centrifuged followed by separation of soluble and insoluble fractions. Fluorescent-protein-containing supernatant (S) and pellet (P) are isolated and visualized by low-pH SDS-PAGE and fluoroimaging.³⁰ (c) Solubility assay of cell-free-generated apoMb₇₇ and apoMb₁₅₃ RNCs (i.e., ribosome-bound nascent chains, unpurified, total concentration ca. 30 nM) in the absence and presence of DnaK and TF chaperones (top). The double bands observed for apoMb₁₅₃ are due to different in-gel tRNA conformations of the same RNC, as they release to a single band upon treatment with puromycin (see also Figure S1). Bar graph (bottom) illustrating the quantitative analysis of RNCs upon multiple independent experiments. Symbols are as follows: WT, wild-type; cf sys, cell-free system; Δ tig, strain devoid of the gene encoding the TF chaperone; inh., inhibitory peptide. (d) Representative low-pH SDS-PAGE (top) and bar graph (bottom) illustrating the solubility of cell-free generated apoMb₇₇ and apoMb₁₅₃ (total concentration ca. 0.3 μ M) in the absence and presence of molecular chaperones. Note that the nascent protein was directly ribosome-released by endogenous release factors as part of cell-free transcription–translation. (e, f) Cartoons highlighting the role of the ribosome as a cotranslational solubility tag and as a post-translational protein solubilizing agent, respectively. (g) Cartoon illustrating the influence of the K/J/E and TF molecular chaperones on the solubility of long nascent chains.¹ Data in panels c and d are avg $\pm$ SE ($n = 2$ –4). (h) Cartoon illustration of the proposed modes of action enabling the ribosome to promote nascent-protein solubility.

quality control mechanisms involving protein disaggregases,^{31–34} this machinery requires a considerable free-energy

expenditure, and has an overall limited scope during the early stages of protein life. In summary, the cell needs to produce

and maintain a structurally accurate proteome without overburdening the quality-control apparatus.

At present, there is still a critical lack of knowledge on the key parameters that govern successful protein folding devoid of aggregation on- and off- the ribosome. For instance, while it is known that both co- and post-translational folding is necessary, it is not understood how these events are coordinated to enable protein channeling to the native state. Further, it is not known which specific processes fail and why, upon *de novo* protein misfolding and aggregation.

Here, we address the above questions by investigating how co- and post-translational folding events grant solubility and structural accuracy to single-domain medium-size proteins. We define a protein as structurally accurate when it encompasses a 100% native-state population.

We find that an important cotranslational role of the ribosome is to solubilize short nascent proteins, even in the absence of molecular chaperones. In addition, we identify a key irreversible kinetic partitioning process that takes place upon nascent-protein release from the ribosome, during which nascent chains are irreversibly channeled toward either native or aggregated conformations. Further, we find that the balance between post-translational kinetic channeling between native state and aggregated species upon release from the ribosome is extremely sensitive to variations in amino acid sequence. This result suggests the presence of specific evolutionary pressure for the generation of protein sequences able to fold without aggregating immediately after release from the ribosome.

METHODS

Cell-Free Transcription-Translation and Production of Ribosome-Released Proteins. Ribosome-released protein were produced as described.¹ In summary, pET-Blue1 plasmids containing *Escherichia coli*-codon-optimized genes³⁵ encoding wild-type (WT) sperm whale apoMb, apoMb₇₇, or M131D apoMb³⁶ were added to either a wild-type or Δ tig (i.e., lacking the trigger factor gene) A19 *E. coli* transcription-translation cell-free system prepared *in-house*, as described.³⁷ Site-specific N-terminal labeling of the translated protein was achieved cotranslationally upon addition of an *in-house* prepared BODIPY FL-Met-tRNA^{Met} according to known procedures.¹ Unless otherwise stated, cell-free transcription-translations were carried out at 37 °C for 20 min. When relevant, a peptide inhibitor of DnaK³⁸ was added to a final concentration sufficient to inhibit the large majority of this chaperone in our *E. coli* cell-free system (140–400 μ M). The concentration of DnaK, DnaJ and GrpE in the cell-free transcription-translation mixtures are ca. 0.5 μ M, 0.04 μ M and 0.05 μ M, respectively, as assessed by Western blotting. In the experiments of Table 1, Figure Sd, and Figure S7, full-length ribosome-released WT apoMb was produced in the presence of additional purified DnaK, DnaJ, and GrpE. The final concentration of DnaK in these experiments was on average 21 or 75 μ M as listed, and the final concentration of DnaJ and GrpE were 0.04–12 μ M and 19 μ M, respectively. Cell-free expression samples produced for fluorescence lifetime and anisotropy-decay experiments were centrifuged for 20 min at 15,800 rcf at 4 °C prior to data collection. Data were acquired on the supernatant. Cell-free samples denoted as “diluted” in Table 1 were prepared by diluting the soluble-cell-free expression product by 10–13-fold into resuspension buffer (10 mM Tris HCl, 11.2 mM Mg(OAc)₂, 60 mM NH₄Cl, 0.5 mM EDTA, 1 mM dithiothreitol, pH 7.0). The samples that

Table 1. Frequency-Domain Fluorescence Anisotropy Decay Parameters of Cell-Free-Expressed WT apoMb in the Absence and Presence of Molecular Chaperones^{a,b}

	time scale of motion ^{a,b}								
	slow	intermediate			fast		reduced χ^2		
	$\tau_{c,s}$ (ns)	$\tau_{c,i}$ (ns)	S_i	$\theta_{c,i}$ (deg)	$\tau_{c,f}$ (ns)	S_f	$\theta_{c,f}$ (deg)	2-comp ^c	3-comp ^c
Δ tig cf sys. + DnaK inh. (item 1)	260 $\pm$ 180				0.104 $\pm$ 0.005	0.38 $\pm$ 0.01	59.6 $\pm$ 0.7	0.57 $\pm$ 0.06	0.37 $\pm$ 0.02
WT cf sys. + 75 μ M DnaK (item 2)	62 $\pm$ 9	1.28 $\pm$ 0.09	0.76 $\pm$ 0.02	33 $\pm$ 2	0.080 $\pm$ 0.005	0.629 $\pm$ 0.006	43.2 $\pm$ 0.4	2.8 $\pm$ 0.4	0.37 $\pm$ 0.03
Δ tig cf sys. + DnaK inh., diluted (item 3)	700 $\pm$ 200				0.122 $\pm$ 0.004	0.25 $\pm$ 0.02	68.5 $\pm$ 0.8	0.47 $\pm$ 0.03	0.44 $\pm$ 0.04
WT cf sys. + 75 μ M DnaK, diluted (item 4)	34 $\pm$ 2	1.3 $\pm$ 0.1	0.80 $\pm$ 0.01	30.6 $\pm$ 0.1	0.108 $\pm$ 0.008	0.58 $\pm$ 0.01	46.8 $\pm$ 0.7	1.65 $\pm$ 0.07	0.4 $\pm$ 0.1
WT cf sys. + 75 μ M DnaK + DnaK inh. added post-translationally (item 5)	44 $\pm$ 5	1.25 $\pm$ 0.07	0.72 $\pm$ 0.02	37 $\pm$ 1	0.07 $\pm$ 0.01	0.62 $\pm$ 0.01	43.9 $\pm$ 0.9	3.2 $\pm$ 0.2	0.39 $\pm$ 0.07

^aAccuracies for $\tau_{c,s}$ and $\theta_{c,s}$ are reported as either propagated error or standard error (SE), whichever of the two was larger, out of 2–7 independent experiments. ^bEach independent experiment included 2–4 replicates. The reduced χ^2 of each independent experiment is the average reduced χ^2 for the replicates. The reduced χ^2 values listed in the table are the average and SE for all independent experiments. Three-component models were preferred to two-component models if the corresponding fits resulted in a ≥ 2.5 -fold reduction in reduced χ^2 , relative to the two-component model. Accordingly, a three-component model was selected for WT apoMb produced in the presence of chaperones, and a two-component model was selected for WT apoMb produced in the absence of chaperones. ^cThe 2-comp and 3-comp captions denote 2- and 3-component curve fits, respectively. The chosen number of components for any given experimental condition (based on the criteria listed in footnote ^c) is shown in bold. ^dThe subscripts S, I, and F denote slow global motions, intermediate local motions and fast local motions occurring on the μ s, ns and sub-ns time scales, respectively. The symbols $\tau_{c,s}$, $\tau_{c,i}$ and $\tau_{c,f}$ denote the rotational correlation times for the global (S) and local (I, F) motions, respectively. S_i and S_f are the order parameters, which, in combination with the $\theta_{c,i}$ and $\theta_{c,f}$ cone semi-angles, serve as reporters of the amplitude of local-motions.^{2,88}

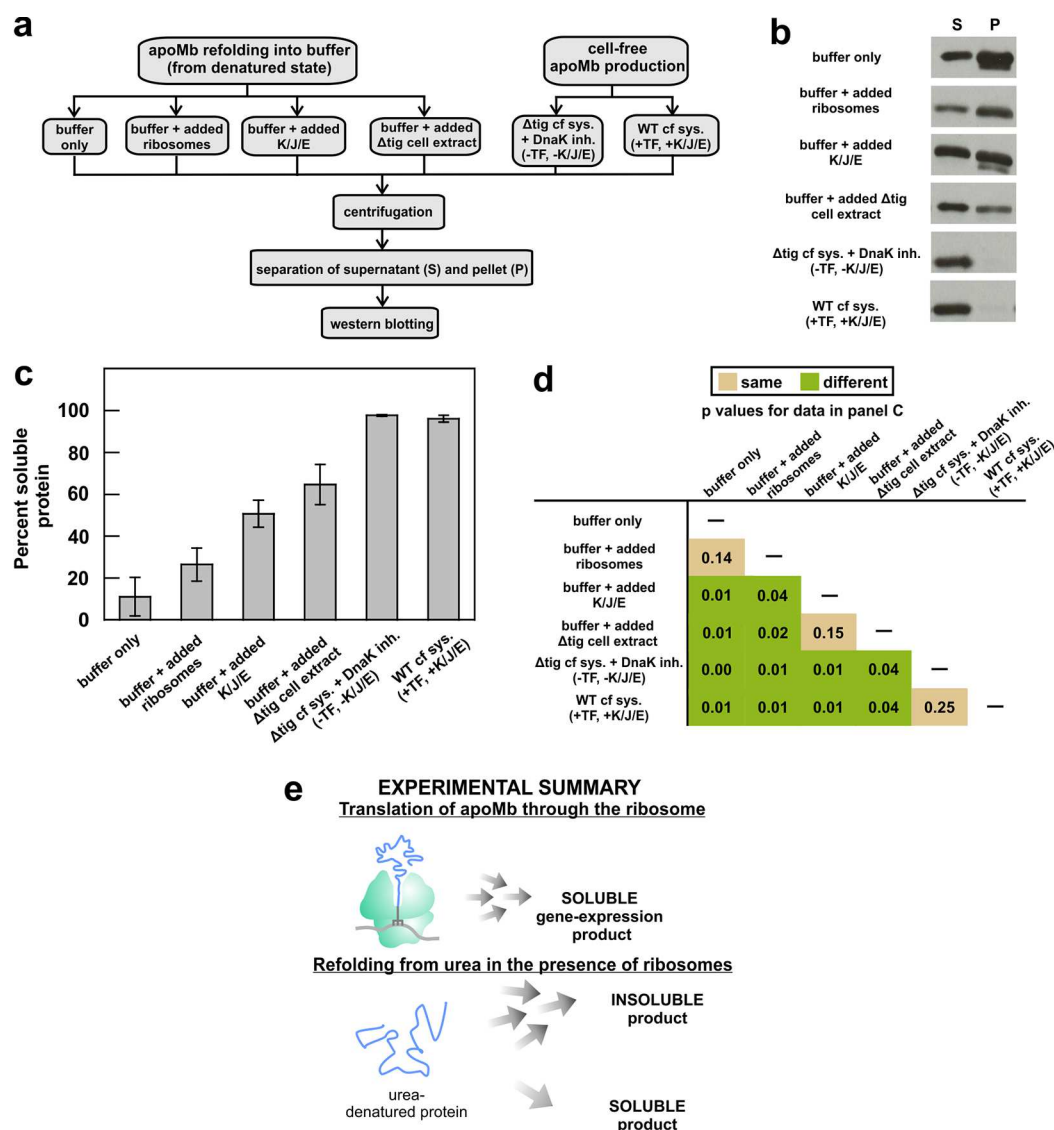


Figure 2. Biosynthesis through the ribosome is necessary and sufficient to generate fully soluble newly synthesized apoMb (a.k.a. WT apoMb, apoMb₁₅₃). (a) Summary of experimental design to isolate the determinants of ribosome-released apoMb solubility. (b) Representative Western blot images comparing soluble (S, supernatant) and insoluble (P, pellet) full-length apoMb refolded from denaturant or expressed in an *E. coli* cell-free system under the conditions listed in panel a. The final total-protein concentration was ca. 0.3 μ M in all experiments documented in the blots. (c) Bar-graph quantification of experimental results according to the panel-a scheme. Data represent avg $\pm$ SE ($n = 2-3$). (d) Statistical p values of data in panel c. All p values were determined via the one-tailed Student's test assuming unequal variances. P values ≤ 0.05 or > 0.05 denote data that are statistically different (green) or not (orange), respectively. (e) Cartoon summarizing the importance of translation through the ribosome to promote the solubility of newly synthesized proteins.

included the post-translationally added DnaK inhibitor were prepared by generating apoMb in the presence of 76 μ M DnaK, 0.04–12 μ M DnaJ and 19 μ M GrpE, followed by isolation of the soluble fraction by centrifugation (20 min, 15,800 rcf, 4 $^{\circ}$ C) and addition of the DnaK inhibitor and ATP to a final concentration of 400 μ M and 1 mM, respectively, assuming the latter components to be anhydrous.

Generation of Purified and Unpurified RNCs. Ribosome-bound nascent protein chains (RNCs) of desired length were prepared in an *E. coli* cell-free transcription–translation system (see previous section) via oligodeoxynucleotide-directed mRNA cleavage¹ at 37 $^{\circ}$ C for 20–30 min. In Figure 4d–f, in which RNCs were employed in spectroscopic analysis, RNCs were purified through a sucrose cushion as described.¹ The RNC pellet was then dissolved in resuspension buffer (10 mM Tris HCl, 11.2 mM Mg(OAc)₂, 60 mM NH₄Cl, 0.5 mM

EDTA, 1 mM dithiothreitol, pH 7.0). The purified nascent chains were released from the ribosome upon treatment with hydroxylamine (see Tables S3 and S4 for details). Nascent chains generated for solubility analysis (Figure 1c and Figure 4g) were neither purified nor further released from the ribosome. The RNC status of all nascent chains in this study was confirmed by their unique position following low-pH SDS-PAGE.³⁰ In Figures 1c and 4g, spontaneous ester bond cleavage of the peptidyl tRNA to the free protein state occurred at 37 $^{\circ}$ C on time scales of ca. $\sim$ 30 min. Chaperone (or lack thereof) conditions were as specified in the figures and (or) figure legends.

Protein Biosynthesis with a Reconstituted Transcription–Translation System (PURExpress). Transcription–translation in the absence of any chaperones (see Figure S4) was carried out with the PURExpress *In Vitro* Protein

Synthesis kit (New England BioLabs) according to the manufacturer's specifications. Briefly, 10 μL of a solution containing amino acids, tRNAs, nucleotides, creatine phosphate, K^+ and Mg^{2+} salts, spermidine, dithiothreitol (DTT), and a formyl-exchange agent was mixed with 7.5 mL of another solution containing aminoacyl synthetases, initiation/elongation/release factors, ribosome recycling factor, creatine kinase, myokinase, nucleotide diphosphate kinase, pyrophosphatase, methionyl-tRNA formyltransferase, T7 RNA polymerase, and ribosomes. Then WT-apoMb-encoding plasmid (tot. concentration 40 $\mu\text{g}/\text{mL}$) and 0.1 units of murine RNase inhibitor were added, together with a sufficient volume of RNase-free water to bring the total mixture volume to 25 mL. To ensure a final protein concentration equal to that achieved via regular cell-free expression, PURExpress transcription–translation was carried out at 37 $^\circ\text{C}$ for 60 min. The solubility of the translated protein was assessed via the centrifugation assay described below. In some cases, PURExpress protein production was carried out in the presence of added chaperones. The desired chaperones were added at the following final concentrations: GroEL and GroES (Sigma-Aldrich) at 0.25 μM and 0.43 μM , respectively, and DnaK, DnaJ, and GrpE at 3.9 μM , 0.3 μM , and 0.5 μM , respectively.

In Vivo apoMb Expression and Purification. Large-scale (4 L growth) overexpression and purification of WT sperm whale and M131D apoMb were carried out as described.^{36,39} Either a pET-Blue1 or a pET-17b plasmid harboring the WT apoMb gene was used. Alternatively, as needed, a pET-Blue1 plasmid harboring the M131D apoMb variant was employed. Either pET-Blue-1-carrying BL21 Turner (DE3) pLac1 cells or pET-17b-carrying BL21 (DE3) cells were used. Inoculation with 1 mM IPTG was carried out at $\text{OD}_{600} \sim 0.5$ and $\text{OD}_{600} \sim 0.9$ with BL21 Turner (DE3) and BL21 (DE3) cells, respectively. Proteins were purified by reverse phase HPLC on a VYDAC Protein C4 preparative column (Grace Materials Technologies), followed by lyophilization. Identity of the purified proteins was assessed by MALDI mass spectrometry. The WT apoMb avg. expected and experimental molecular weights were 17,331 amu. The M131D apoMb expected average molecular weight was 17 315 amu, and the experimental molecular weight was 17 312 amu.

Generation of M131D apoMb. The QuikChange II Site-Directed Mutagenesis kit (Agilent) was used to generate the M131D apoMb DNA point mutation, starting from the pET-Blue1 vector harboring the WT apoMb gene. M131D apoMb was overexpressed and purified by the same procedures employed for the WT protein.

Refolding of Purified Native apoMb from a Chemically Denatured State. Lyophilized WT apoMb (160 μM ; purified as above) was incubated at room temperature for 1 h in a denaturing solution (6 M urea in DNase/RNase-free H_2O adjusted to pH 2.3 with HCl). Next, the denatured apoMb was filtered through a 0.22 μm membrane and then bath-sonicated for five min. The denatured apoMb solution was then diluted 1:1000 in refolding buffer (1 mM TrisCl pH 7.0, 1 mM $\text{Mg}(\text{OAc})_2$, 6 mM ammonium chloride, and 50 μM EDTA). When needed, the refolding buffer also contained additional purified ribosome, extra DnaK (and its cochaperones DnaJ and GrpE), or Δtig A19 *E. coli* cell extract. Ribosomes were purified as described below. Additional DnaK, DnaJ, and GrpE were added to a final concentration of 3–5 μM , 0.6–1 μM , and 1.2–2 μM , respectively.

Purification of Empty Ribosomes. In experiments requiring purified ribosomes, we started from a Δtig A19 *E. coli* cell extract, which contains predominantly empty ribosomes, and purified its ribosomes via sucrose cushion centrifugation at 160 000 rcf at 4 $^\circ\text{C}$ for 1 h, similarly to the RNC purification described above.

Solubility Assays. Cell-free expressed apoMb bearing an N-terminal BODIPY FL fluorophore was centrifuged at 15 800 rcf and 4 $^\circ\text{C}$ for 20 min. The resulting pellet was separated from the supernatant and resuspended in resuspension buffer, 60 mM ammonium chloride and 0.5 mM EDTA. The resulting supernatant and pellet were then analyzed by low-pH SDS-PAGE.³⁰ The fluorophore-labeled apoMb in each fraction was visualized via a FLA 9500 Typhoon Gel Imager (473 nm excitation laser, BPB1 emission filter). ApoMb band intensities were then analyzed with the ImageJ software package. The percent of soluble protein was computed according to the expression below, where I_{soluble} and $I_{\text{insoluble}}$ denote fluorescent band intensities of the soluble and insoluble fractions

$$\% \text{ solubility} = \frac{I_{\text{soluble}}}{I_{\text{soluble}} + I_{\text{insoluble}}} \times 100 \quad (1)$$

Western-Blot Solubility Assays. Western blotting was employed to analyze the soluble and insoluble fractions of WT apoMb that was either *in vitro* refolded or cell-free expressed under different conditions (Figure 2). Samples were subject to conventional SDS-PAGE prior to transfer to a PVDF membrane. The membrane was then blocked for 1 h at room temperature. Subsequent 1st (1:500 rabbit-antimyoglobin from Santa Cruz Biotechnology) and 2nd (1:10000 goat-antirabbit-HRP, GE Healthcare Life Sciences) antibody incubations were carried out in fresh blocking solution at room temperature for 3 h or 40 min, respectively. Western blot development was carried out via an Amersham ECL Western blot detection kit (GE Healthcare Life Sciences) according to the manufacturer's instructions. ApoMb band intensities corresponding to soluble and insoluble fractions were assessed with the ImageJ software package.

NMR Sample Preparation. To improve apoMb cell-free expression yields and enable NMR analysis, selectively labeled apoMb containing ^{15}N A, E, L, K, V, and ^{15}N – ^{13}C W was generated in a dialysis device similar to the one employed by Bakke et al.,³⁷ except that an *E. coli* cell-free extract lacking the TF gene (Δtig) and an added DnaK inhibitor³⁸ were employed. This procedure enabled generating apoMb in an environment devoid of TF and active DnaK. In short, four identical 100 μL cell-free mixtures were each dialyzed against 8 mL of dialysis buffer in Pierce 96-well microdialysis strips (10 kDa molecular weight cutoff, Thermo Scientific). The dialysis buffer consisted of the same concentrations of amino acids, nucleotides, buffers, salts, energy sources, DnaK inhibitor, and folic acid as the cell-free mixtures. The concentration of each amino acid was 1 mM. The concentrations of plasmid and DnaK inhibitor were 50 $\mu\text{g}/\text{mL}$ and 145 μM , respectively. The concentration of T7 RNA polymerase in the cell-free mixture was 5.2 U/ μL . Cell-free transcription–translation was carried out overnight (ca. 12 h) at 37 $^\circ\text{C}$ under gentle shaking. After dialysis, each cell-free mixture was recovered and centrifuged at 15 800 rcf for 20 min at 4 $^\circ\text{C}$. The supernatants of each of the four cell-free mixtures were combined to generate 400 μL of cell-free mixture containing soluble, selectively labeled apoMb. Prior to NMR analysis, 40 μL of D_2O were added to 360 μL of

the cell-free mixture and loaded onto a 5-mm-diameter Shigemitsu tube (Wilma, BMS-005B).

A negative control sample was prepared as above except that a cell-free extract derived from wild-type A19 *E. coli* strain was used, and no DnaK inhibitor nor T7 RNA polymerase (to prevent gene expression) were added.

The total apoMb concentration of NMR samples was assessed via SDS-PAGE upon comparisons with standards consisting of cell-free reactions spiked with apoMb of known concentration. Gels were stained with Coomassie blue and imaged on a Typhoon FLA 9500 (GE Healthcare) upon excitation with a 532 nm laser in the presence of a 575 nm long-pass filter. ApoMb gel band intensities were assessed with the ImageJ software.^{40,41} A standard curve of apoMb of known concentrations was generated to deduce the total apoMb concentration of the NMR sample (59.8 μ M).

NMR Data Collection, Processing, and Volume Analysis. NMR spectra were recorded at 36 °C on an Avance III HD 600 MHz NMR spectrometer (Bruker Biospin Corp) equipped with the triple resonance inverse (TCI) cryogenic probe $^1\text{H}/(^{19}\text{F}/^{13}\text{C}/^{15}\text{N})$ fitted with a z-axis gradient. The pulse sequence used was sensitivity-enhanced HSQC employing pulse-field gradients for coherence selection and a square water flipback pulse for solvent suppression.^{42–45} Data in parts a and b of Figure 3 were collected with 740 and 700 scans per row, respectively. Data were acquired with 2048 (t_2) $\times$ 128 (t_1) complex points. The direct dimension was centered on the residual solvent resonance. Sweepwidths of 10 000 and 2200 Hz were used for the direct and indirect dimensions, respectively. The relaxation delay was set to 1.5 s. All data were zero-filling twice in both dimensions. A 18°-shifted cosine-bell square window function was applied to the indirect dimension. ^1H chemical shift referencing was done with external 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS). Data were processed and visualized with the NMRPipe (v: 9.0.0-b108)⁴⁶ and NMRView J (v: 2009.015.15.35)⁴⁷ software packages.

Volume analysis of NMR resonances led to estimating the fraction of spectroscopically undetectable native state due to extremely slow tumbling. The NMR resonance volume, as opposed to intensity, does not depend on molecular tumbling rates.⁴⁸ Detectable resonances have measurable volumes, which are directly proportional to the populations of fast-tumbling species regardless of the NMR exchange regime. Note that undetectable populations arise from losses of transverse-coherence either during pulses, delays or acquisition time. Briefly, the 2D NMR volume of the ^1H – ^{15}N resonance corresponding to the free-Trp indole (Vol_{Trp}) was measured with the NMR View J software. The total free-Trp concentration in the sample (Conc_{Trp} , 900 μ M) was regarded as effectively identical with the total free-Trp concentration in the dialysis-assisted transcription–translation, given the large excess of free Trp. Next, Conc_{Trp} was divided by Vol_{Trp} . The resulting ratio was multiplied by the volume of one of the clearly resolved apoMb resonances (Vol_{res}): V17, A19, A22, A57, K62, K78, or A130, according to

$$\frac{\text{Conc}_{\text{Trp}}}{\text{Vol}_{\text{Trp}}} \times \text{Vol}_{\text{res}} = \text{Conc}_{\text{res}} \quad (2)$$

This procedure led to eight separate native-apoMb concentrations, which were then averaged. The resulting concentration was multiplied by 100 and divided by the total

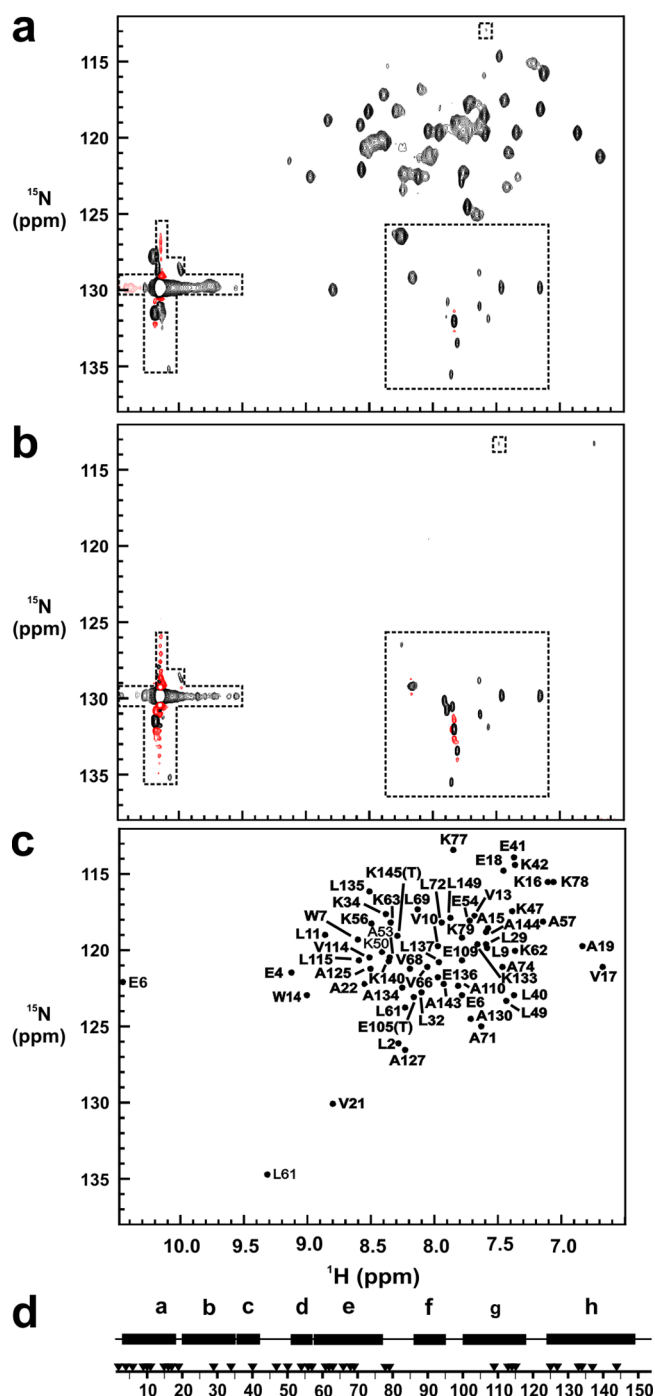


Figure 3. The ribosome alone is not sufficient to guarantee formation of fully structurally accurate apoMb. (a) ^1H – ^{15}N HSQC 2D NMR spectrum of selectively labeled apoMb (^{15}N -labeled A, E, L, K, V, and ^{13}C - and ^{15}N -labeled W) newly synthesized in a cell-like medium lacking the trigger factor chaperone and containing a DnaK inhibitor. Cell-free expression was carried out in a dialysis-enhanced mode,³⁷ and NMR data were collected directly within the cell-free mixture. Unboxed resonances are consistent with known native-apoMb assignments.⁵⁹ (b) Negative-control ^1H – ^{15}N HSQC NMR spectrum of a sample similar to panel a, except that no transcription–translation was allowed to occur (see details in Methods). (c) Simulated native-apoMb ^1H – ^{15}N HSQC NMR spectrum comprising the selectively labeled backbone resonances according to published assignments.⁵⁹ (d) Mapping of the labeled ApoMb resonances onto the protein's primary structure. ApoMb resonances are denoted by the (▼) and the eight individual helices of native apoMb are shown as black horizontal

Figure 3. continued

bars (a–h). Boxed resonances in panels a and b denote products of amino acid metabolism⁸⁷ and the ¹H–¹⁵N indole pair of free tryptophan. NMR-resonance volume analysis revealed that only 23% of the total protein is spectroscopically detectable. Therefore, only a moderate fraction of the newly synthesized apoMb generated in the absence of trigger factor and in the presence of a DnaK inhibitor has native-like conformation.

protein concentration of the NMR sample, determined via Coomassie-blue-detected SDS-PAGE as described in the previous section, to yield the percent of NMR-detectable folded, soluble protein (23.5%). The remaining NMR-undetectable population (76.5%) was ascribed to a slow-tumbling fraction of the apoMb NMR sample, due to soluble self- or heteroassociated species.

Circular Dichroism. Purified WT and M131D apoMb were resuspended in 20 mM NaOAc at pH 6.0, bath-sonicated for 5 min, and eluted through a 0.2 μm pore-size filter. The resulting samples were analyzed on a MOS-450 spectropolarimeter (Bio-Logic Science Instruments) in circular dichroism mode over the 200–260 nm spectral range with 1 nm steps and data-collection time of 20 s per step, at room temperature. The resulting spectra were baseline-corrected and converted to mean residue ellipticity (MRE) according to

$$\text{MRE} = \frac{\theta}{10C(N_r - 1)b} \quad (3)$$

where θ denotes the experimentally determined ellipticity (in mdeg), b is the cuvette path length in cm, C is the apoMb molar concentration, and N_r is the number of apoMb residues, which was set to 154. The molar concentration of apoMb in each sample was determined on an HP 8452A Diode Array UV/vis spectrophotometer (Hewlett-Packard) by

$$[\text{apoMb}] = \frac{A_{280}}{\epsilon b} \quad (4)$$

where A_{280} is the baseline-corrected sample absorbance at 280 nm, b is the cuvette path length in cm, and ϵ is the known WT apoMb extinction coefficient,⁴⁹ set to 14 134 M^{−1}cm^{−1}. The parameter ϵ was assumed to be the same for WT and M131D apoMb.

Assessment of WT and M131D apoMb Thermodynamic Stability. The thermodynamic stability of purified WT and M131D apoMb was assessed by Trp fluorescence. Lyophilized powders of purified protein (4–7 μM) were resuspended in 20 mM NaOAc at pH 6.0. The resulting samples were bath-sonicated for 5 min and eluted through a 0.2 μm pore-size filter, followed by incubation in the presence of 0–7 M urea (separate samples at each denaturant concentration, assessed by refractometry⁵⁰) for 1 h at room temperature. The Trp fluorescence emission of each sample (from 300 to 500 nm) was monitored at room temperature on a PC1 photon-counting steady-state fluorimeter (ISS Inc.) upon excitation at 285 nm. Excitation and emission monochromators had slit widths of 4 nm. Frequency shifts in spectral center of mass ($\langle \nu_i \rangle$) upon unfolding were assessed according to Jameson⁵¹ and Weber⁵² via relation

$$\langle \nu_i \rangle = \frac{\sum \nu_i F_i}{\sum F_i} \quad (5)$$

where F_i denotes the fluorescence intensity at wavenumber ν_i and the summation is performed across the range of fluorescence intensities of each emission spectrum. The fraction of unfolded protein at each urea concentration (f_{unf}) was determined according to^{52,53}

$$f_{\text{unf}} = \left(1 + Q \left(\frac{\langle \nu_{\text{XM}} \rangle - \langle \nu_{7\text{M}} \rangle}{\langle \nu_{0\text{M}} \rangle - \langle \nu_{\text{XM}} \rangle} \right) \right)^{-1} \quad (6)$$

where Q denotes the ratio between the quantum yields of the unfolded and native states, $\langle \nu_{\text{XM}} \rangle$ is the spectral center of mass at XM urea, and $\langle \nu_{0\text{M}} \rangle$ and $\langle \nu_{7\text{M}} \rangle$ denote the spectral centers of mass of the native and unfolded states, respectively. Quantum-yield changes were estimated from changes in total fluorescence intensity at the maximum. Q values of WT and M131D apoMb were assumed to be the same, given the structural similarity between the two proteins assessed by far-UV circular dichroism. Data were fit according to a two-state model according to Santoro and Bolen^{54,55} with Kaleidagraph (version 4, Synergy Software). Apparent $\Delta G^\circ_{\text{NU}}$ and m_{NU} were deduced from the fits and reported as average $\pm$ SE for $n = 3$. P values were assessed via the two-tailed Student test assuming unequal variances. Additional rationale justifying the choice of an apparent two-state model is provided in the [Supporting Information](#).

Kinetic Experiments Assessing the Timecourse of Soluble and Insoluble Protein Production in an *E. coli* Cell-Free System. Experiments shown in panels a and b of [Figure 5](#) were carried out as follows. All cell-free transcription–translation components were mixed on ice (total volume 82 ± 2 μL). Twelve identical aliquots (5.7 μL each) of this mixture were then generated in separate RNase-free tubes. The tubes were then immediately placed in a water bath at 37 °C to initiate transcription–translation. At each of the time intervals of interest (4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 50, and 60 min), transcription–translation of one of the aliquots was quenched upon treatment with 2 mM puromycin (pH = 7.0) and the sample was immediately placed on ice. Puromycin causes instantaneous interruption of translation,⁵⁶ and acts as an excellent quencher. Each sample was centrifuged for 20 min at 15 800 rcf at 4 °C, and supernatants were separated from pellets. Pellets were dissolved in a resuspension-buffer volume equal to the total sample volume before centrifugation. Equal volumes of all samples were subject to low-pH SDS-PAGE³⁰ analysis. Precision Plus molecular-weight markers (Biorad) were run in parallel. Gels were imaged with an FLA 9500 Typhoon Gel Imager with a 473 nm excitation laser and a BPB1 emission filter, while a 635 nm excitation laser and an LPR filter were used to image the molecular-weight markers. Quantification of gel-band intensities corresponding to soluble and insoluble protein was carried out with ImageJ.^{40,41}

Processing of Kinetic Data. To guide the eye and to assist with the generation of kinetic-flux information, all the timecourse data for soluble and insoluble proteins production ([Figure 5a,b](#)) were fit to generic analytical expressions for multistep unimolecular reactions. Kinetic fluxes for each of the timecourse data ([Figure S6c,d](#)) were estimated as the first derivative of the curves corresponding to each of the above analytical expressions.

Post-translational Kinetic Trapping of Newly Synthesized Protein. Experiments shown in [Figure 5e](#) were performed as follows. Ribosome-released M131D apoMb was generated as described above, except that transcription–

translation was carried out for 1 h instead of 20 min. The process was quenched by placement on a water/ice bath. Soluble and insoluble protein fractions were separated by centrifugation at 15 800 rcf at 4 °C for 20 min. The pellet was resuspended in a cell-free transcription–translation solution containing all components except for T7 RNA polymerase, BODIPY FL-Met-tRNA^{Met} and plasmid DNA, to prevent further protein production. Additionally, puromycin (1 mM, in 10 mM Tris, 11 mM Mg(OAc)₂, 60 mM NH₄Cl, 0.5 mM EDTA at pH 7) was added to both the supernatant and resuspended pellet to further ensure that no additional protein synthesis would occur. Supernatant and resuspended pellet were incubated at 37 °C for ca. 14 h, and then placed on ice. The solutions were then centrifuged at 15 800 rcf at 4 °C for 20 min. Pellets were resuspended in buffer (10 mM Tris, 11.2 mM Mg(OAc)₂, 60 mM NH₄Cl, 0.5 mM EDTA, 1 mM dithiothreitol). All samples were subject to low-pH SDS-PAGE,³⁰ and gel-band intensities were quantified with ImageJ.^{40,41}

Fluorescence-Lifetime and Anisotropy-Decay: Data Collection and Analysis. Fluorescence lifetime and anisotropy-decay data were collected as described³ on a Chronos frequency-domain fluorometer (ISS Inc.) equipped with calcite prism polarizers. BODIPY-FL-labeled samples were excited with a 477 nm laser diode. A 480 ± 5 nm band-pass and a 495 nm long-pass (Chroma Technology) filter were applied to the excitation and emission channels, respectively. In both lifetime and anisotropy-decay experiments, the excitation polarizer was set to vertical (i.e., ||). In lifetime experiments, the emission polarizer was set to the magic angle (54.7°). Temperature control was achieved with a circulating water bath set to 25 °C. Samples were inserted in the cuvette housing of the fluorometer and thermally equilibrated for ≥30 min prior to data collection. Fluorescence depolarization data were corrected with polynomials derived from frequency-dependent G-factor experiments, performed on each experimental day as described.³

Lifetime and depolarization data were fit with the Globals Software Suite (Laboratory for Fluorescence Dynamics, LFD). Reduced χ^2 values were determined assuming frequency-independent instrumental errors of 0.2° and 0.004 in the phase and modulation, respectively. All lifetime data required fitting to three-discrete exponentially decaying components, two of which reported on the actual fluorophore lifetimes. A third fictitious component, which was set to a fixed value of 1 ps, served to account for an extremely small degree of elastic light scattering. Anisotropy data were fit to multiexponential anisotropy decays with two or three discrete components. The fundamental anisotropy $r(0)$ was fixed at a value of 0.37, while fractional amplitudes and rotational correlation times were allowed to float. The 3-component anisotropy fits were only used if they resulted in a 2.5-fold or greater reduction in reduced χ^2 relative to 2-component fits. Order parameters and cone semiangles were determined as described.²

Lifetime phasors⁵⁷ were computed at 23.08 MHz for full-length ribosome-released WT apoMb in an *E. coli* cell-free system (Figure Sd and Figure S7), according to the expressions

$$G = M \cos \phi \quad (7)$$

$$S = M \sin \phi \quad (8)$$

where M and ϕ denote modulation and phase shift, respectively.

RESULTS

Experimental Design. We focused on the biogenesis of medium size (i.e., ca. 17 kDa) single-domain proteins, which encompass a highly represented class in bacteria.⁵⁸ Our model protein, apomyoglobin (apoMb), is a member of the ubiquitous globin fold. Pure apoMb has a well-defined structure in solution under physiologically relevant conditions,⁵⁹ and its *in vitro* refolding pathways have been extensively characterized.^{60–64} ApoMb populates a compact partially folded structure cotranslationally but completes its folding posttranslationally.¹ This protein is prone to insoluble inclusion-body formation upon *in vivo* overexpression³⁷ but is fully soluble when generated at more moderate expression levels in an *E. coli* cell-free system.³⁷ ApoMb is known to interact cotranslationally with the trigger factor (TF) and DnaK/DnaJ/GrpE (K/J/E) chaperone systems.¹ The concentration of apoMb produced in this study (ca. 0.3 μM) is within the range of natural *in vivo* expression levels in the *E. coli* proteome.⁵⁸ The plasmid encoding for apoMb has codon usage optimized for expression in *E. coli*.³⁵ We also analyzed an aggregation-prone variant bearing a single-point mutation in a region buried inside the ribosomal exit tunnel upon completion of biosynthesis.

Unlike studies performed *in vivo*, cell-free strategies enable discriminating effects due to the ribosome alone from contributions arising from the *E. coli* chaperones trigger factor (TF) and DnaK. We employed an efficient wild-type *E. coli* cell-free system,³⁷ as well as a cell-free system generated from an *E. coli* strain lacking TF (Δtig) or containing an *in-house*-designed DnaK inhibitor (see Methods). The latter setup enabled ruling out contributions from the TF and DnaK chaperones, respectively.

Cotranslational properties were probed by stalling the nascent protein on the ribosome, and post-translational characteristics were assessed by allowing translation to proceed naturally, including its termination via the bacterial release factors in the cell-free system. The use of ribosome-stalled nascent protein chains instead of the corresponding transiently populated chains is reasonable, given that (a) spectroscopic analysis of transient chains in real time is technically challenging, (b) conformational sampling is typically faster than translation^{36,39,65} including on the ribosome,⁶⁶ and (c) codon-dependent translation rates may govern cotranslational folding rates when rare codons or codon clusters are present,⁶⁷ yet the gene encoding our model protein has no rare codons.⁵⁸ Therefore, no such effects are expected. Specifically, translation of nascent apoMb takes a few seconds at 37 °C,^{35,36} providing plenty of time for conformational sampling of the nascent protein before the onset of ribosomal stalling.

Newly synthesized protein quality was evaluated in terms of solubility and structural accuracy. ApoMb solubility was characterized by centrifugation (Figure 1b and Supporting Information), and the structural accuracy of the ensemble was probed by multidimensional NMR. Time-resolved fluorescence depolarization was employed to examine co- and post-translational protein compaction. In all experiments involving fluorescence detection, apoMb was cotranslationally site-specifically labeled at its N terminus with the BODIPY FL fluorophore.

The Ribosome Enhances Nascent-Protein Solubility. Parts c and d of Figure 1 illustrate the co- and post-translational solubility of a 77-residue N-terminal fragment of

apoMb, denoted as apoMb₇₇, and the corresponding solubility of the full-length protein (153 residues), referred to as apoMb₁₅₃ or WT apoMb. ApoMb₇₇ ribosome-bound nascent complexes (RNCs) are highly soluble (~30 nM, Figure 1c and Figure S1) regardless of the presence of the trigger factor (TF) and DnaK/DnaJ/GrpE (K/J/E) chaperones. The RNC ribosomal complexes carrying nascent apoMb₇₇ are known to associate with molecular chaperones.¹ This association, however, does not play any significant role in apoMb₇₇ RNC solubility (Figure 1c). The solubility of apoMb₇₇ RNCs may be contributed by the protective environment of the ribosomal-tunnel core, shielding 30–35 C-terminal residues,⁶⁸ and by the tunnel vestibule, potentially shielding the remaining nascent-chain residues from bulk solution.

In contrast, newly expressed ribosome-released ApoMb₇₇ is highly insoluble (Figure 1d and Figure S1) both in the absence and presence of the DnaK and TF chaperones (~0.5 μ M and 0.2 μ M, respectively). We conclude that the ribosome increases the solubility of partially synthesized aggregation-prone nascent apoMb, even in the absence of molecular chaperones, as illustrated in the cartoon of Figure 1e. Control experiments verified that the nascent-protein solubilizing activity of the ribosome is not an artifact due to differences in concentration between RNCs and ribosome-released proteins (Figure S2).

RNCs encoding full-length apoMb₁₅₃, on the other hand, are only partially soluble in the absence of chaperones (Figure 1c and Figure S1). This incomplete solubility may be due to higher exposure of the longer apoMb₁₅₃ RNCs to a more solvent-accessible region outside of the ribosomal vestibule, relative to apoMb₇₇ RNCs. Alternatively, apoMb₁₅₃ nascent chains may simply have an increased amount of nonpolar content relative to apoMb₇₇ chains (Supporting Information and Figure S3a,b). The nascent-chain N-terminal protein fragment comprising residues emerging out of the ribosomal tunnel of apoMb₁₅₃ RNCs (ca. 118–123 amino acids)⁶⁸ is significantly aggregation-prone.^{36,49} Hence we conclude that the ribosome has an RNC-solubilizing effect whose extent depends on the length and physical properties of the nascent chain.

Upon release from the ribosome, apoMb₁₅₃ regains fully soluble status even in the absence of chaperones (parts d and f of Figure 1 and Figure S1). The latter effect is not surprising, as the added availability of the last C-terminal residues for conformational sampling, when the chain emerges from the ribosomal exit tunnel, promotes completion of intramolecular folding.^{1,68} See later sections for additional kinetic considerations.

In summary, our data show that the ribosome promotes RNC solubility.

Not surprisingly, apoMb₁₅₃ RNCs achieve complete solubility in the presence of the DnaK and TF chaperones (Figure 1c,g and Figure S1), consistent with the fact that these chaperones contribute to the solubility of *de novo* synthesized proteins.⁶⁹

The role of the ribosome on nascent-protein solubility can be rationalized according to the three proposed modes of action illustrated in Figure 1h. First, the large net charge of the ribosome⁷⁰ serves as a solubilizing tag for nascent chains, much like soluble covalently bound proteins (e.g., maltose-binding protein) that are known to serve as solubility-enhancing fusion partners.⁷¹ Second, by interacting with the ribosomal surface,^{18–21} nascent proteins may be protected from deleterious

nonspecific interactions with other nascent proteins. Finally, the large, multimegadalton size of the ribosome is expected to disfavor molecular collisions among RNCs due to slow translational diffusion. While our data prove that the ribosome plays a role in RNC solubilization, more detailed future experiments are necessary to establish which of the above models is most pertinent.

Emergence through the Ribosome Is Necessary to Prevent Insoluble-Aggregate Formation. Next, we explored the extent to which the ribosome is required for the solubility of newly synthesized ribosome-released proteins. The basic experimental scheme is shown in Figure 2a. A centrifugation assay (Figure 1b; see Methods and further details in the Supporting Information) was employed to separate newly synthesized soluble from insoluble protein. Visualization of apoMb was performed by Western blotting (Figure 2b). The final apoMb concentration is the same (~0.3 μ M), in all data of Figure 2. ApoMb is a particularly suitable protein for this experiment because its solution conformation,⁵⁹ refolding from denaturant⁶⁰ and folding behavior as a function of chain elongation^{36,72–75} are well-known. This wealth of knowledge provides crucial reference information.

First, as a control, purified apoMb₁₅₃ was refolded from a chemically denatured state (in urea) into an excess of buffered solution lacking denaturant. A second class of experiments involved chemically denatured apoMb₁₅₃ refolded into a buffer containing either purified empty ribosomes or the K/J/E chaperone system, which comprises the DnaK (K), DnaJ (J), and GrpE(E) chaperones. Alternatively, refolding into buffer was carried out in the presence of added Δ tig cell extract, which lacks TF but contains other chaperones in addition to K/J/E, to test the overall importance of cellular components other than TF.

Protein behavior under the above conditions was compared to apoMb₁₅₃ generated *in situ* via ribosome-assisted transcription–translation in a bacterial cell-free system. Newly synthesized proteins were produced directly in the cell-free environment; i.e., no generation of stalled RNCs was involved. This class of experiments was carried out either in the absence or presence of TF and K/J/E, to discriminate the role of the ribosome from that of chaperones.

Panels b–d of Figure 2 show that the only experiment that produces nearly 100% soluble folded apoMb₁₅₃ is the one where the nascent protein has been generated *in situ* and has emerged through the ribosome. Interestingly, no statistically relevant differences were observed whether or not the TF and DnaK chaperones were present, highlighting that the ribosome alone is sufficient to grant full solubility to the newly synthesized protein, and that chaperones are not needed for solubility purposes. Control experiments performed with a fully reconstituted PURE cell-free system lacking any co- and post-translationally active chaperones confirmed this conclusion (Figure S4).

Purified apoMb₁₅₃ refolded from denaturant into buffer is highly insoluble upon refolding. This behavior is a shared characteristic of many *in vitro*-refolded proteins.^{76–78} Note that previous reports of apoMb refolding from denaturant under comparable conditions (similar protein concentration, pH, buffer systems) reported either no aggregation^{60–64} or soluble⁷⁹ aggregates. This is either due to the fact that solution conditions were not identical with those reported here, to the fact that the protein had been filtered (via 0.2 μ m devices or sephadex G25 spin columns), or to the fact that

insoluble aggregates tend to remain suspended in solution in the absence of centrifugation. Interestingly, even our apoMb samples of Figure 2c, refolded into buffer, appeared only slightly opalescent to eye inspection and displayed no visible precipitate. The presence of the insoluble aggregate could only be detected explicitly upon centrifugation.

Added ribosomes do not lead to any statistically significant increases in solubility ($p = 0.14$, Figure 2d). Hence the mere presence of the ribosome in solution is not sufficient to grant solubility to newly synthesized protein, and the known chaperone activity of the ribosomal peptidyl transferase center^{80–86} does not appear to play a significant role in our experiments, out of the context of the translating ribosome. Refolding in the presence of either purified K/J/E or entire cellular components except for TF (Δ tig extract) shows improvement in solubility relative to refolding from buffer ($p = 0.00–0.01$, Figure 2d). Hence, as expected, chaperones like K/J/E, and possibly others, promote solubility.

In summary, translation *through* the ribosome alone, is necessary and sufficient to grant newly synthesized apoMb a 100% soluble status, even in the absence of molecular chaperones. This important result is summarized in Figure 2e.

Emergence *through* the Ribosome Is Not Sufficient to Guarantee the Full Structural Accuracy of Newly Synthesized apoMb: NMR Analysis. To gain additional insights into the conformational characteristics of the soluble fraction of the newly synthesized-protein, we acquired a 2D ¹H, ¹⁵N HSQC NMR spectrum of cell-free-generated apoMb *in situ*, in the absence of the TF and DnaK molecular chaperones. To simplify spectral features and limit costs, only selected residues were ¹⁵N-labeled (Figure 3a). Control data on a cell-free system devoid of apoMb transcription–translation (Figure 3b) were also collected, and a computed reference 2D NMR spectrum based on published assignments⁵⁹ was generated (Figure 3c).

We assigned the majority of the resonances in Figure 3a (56 out of 61) to native apoMb (Table S1).⁵⁹ In addition, we observed 7 new resonances not attributable to metabolites or free amino acids. The latter may be due to small pH differences between the cell-free environment and the conditions used in the published assignments, or to alternative apoMb conformational states in slow exchange on the NMR chemical-shift time scale. Overall, there is excellent agreement between panels a and c of Figure 3, showing that the major detectable resonances are due to native apoMb.

Next, the volume of the detectable protein resonances was compared to the corresponding volume of the free-tryptophan indole ¹⁵N NMR resonance, employed as an internal standard of known concentration. This procedure exploits the independence of NMR-resonance volumes, as opposed to intensities, on molecular tumbling rates⁴⁸ (see Methods for additional details and theoretical considerations). NMR volume analysis combined with evaluation of gel-band intensities enabled estimating that ca. 77% of the soluble cell-free-generated apoMb is not spectroscopically detectable. This lack of detectability is due to line-broadening beyond detection.

Importantly, this broadening arises from species that must be in slow conformational exchange with the native state on the NMR chemical-shift time scale. If this were not the case, the native-state resonances would be significantly broadened. This feature is not observed in our case, as testified by the sharp native-state resonances (Figure 3a). The NMR-

undetectable species in slow exchange with the native state are either slow-tumbling large adducts (e.g., homo- or heteroaggregates), small-size apoMb conformers interconverting on the microsecond–millisecond time scale (intermediate exchange regime), or apoMb species that may weakly bind to the cell-free milieu on any time scale yet interconvert slowly ($\sim$ faster than a few milliseconds) with the native state. A combination of the above limiting cases is also possible. Yet, again, the undetectable populations must be in slow exchange relative to the native state on the NMR chemical-shift time scale.

Emergence *through* the Ribosome Is Not Sufficient to Guarantee the Full Structural Accuracy of Newly Synthesized apoMb: Fluorescence Analysis.

Fluorescence anisotropy decay data, collected under conditions comparable to the NMR experiments, shed additional light on this topic. Newly ribosome-released nascent proteins encoding WT apoMb were monitored within the cell-free environment via frequency-domain fluorometry. This technique is a powerful tool to characterize local environment, compaction and dynamics of nascent proteins in complex environments.^{1–3,88} The proteins were specifically labeled at their N termini with the small BODIPY FL fluorophore. As shown in item 1 of Table 1 (see also Table S2), the rotational correlation time for the global tumbling of apoMb ($\tau_{c,s}$) in the absence of trigger factor and in the presence of DnaK inhibitor, i.e., under chaperone-free conditions, is 260 ± 180 ns. This is a large number, which is consistent with the presence of a very slow-tumbling apoMb population. This value supports the presence of a large species in solution, i.e., either a self- or heteroassociated apoMb soluble aggregate. A similar conclusion is supported by repeating the same experiment after a post-translation 10–13-fold dilution in buffer, leading to a $\tau_{c,s}$ of 700 ± 200 ns under chaperone-free conditions (Table 1, item 3). Despite the fact that the above rotational correlation times (260 and 700 ns) are considerably larger than the average fluorophore lifetime, the frequency-domain anisotropy decay approach was able to reproducibly detect the presence of these large slow-tumbling species. Interestingly, fluorescence anisotropy decays were, on the other hand, unable to explicitly detect the monomeric native state, whose presence is abundantly evident by NMR.

As shown in Figure S7a, the average fluorescence lifetimes of ribosome-released WT apoMb are reproducibly different upon addition of increasing amounts of DnaK and its cochaperones. On the other hand, these differences are not large (i.e., less than 2-fold). This similarity in average fluorescence lifetimes is likely responsible for the dominance of the large τ_c species upon fluorescence data fitting. Hence, an associative data fitting,^{3,89–91} capable of explicitly discriminating small and large species having sufficiently different fluorescence lifetimes, could not be applied.

In all, fluorescence and NMR complement each other in identifying both small and large populated species in solution under chaperone-free conditions.

Importantly, the observed rotational correlation time decreases to lower values (62 ± 9 and 34 ± 2 ns, when physiologically relevant concentrations (75 μ M) of the DnaK/DnaJ/GrpE (K/J/E) chaperones are added to the medium (Table 1 items 2 and 4). Given that the remaining solution conditions were identical with the chaperone-free samples (items 1 and 2 in Table 1, respectively) except for the presence of molecular chaperones, this result strongly suggests that

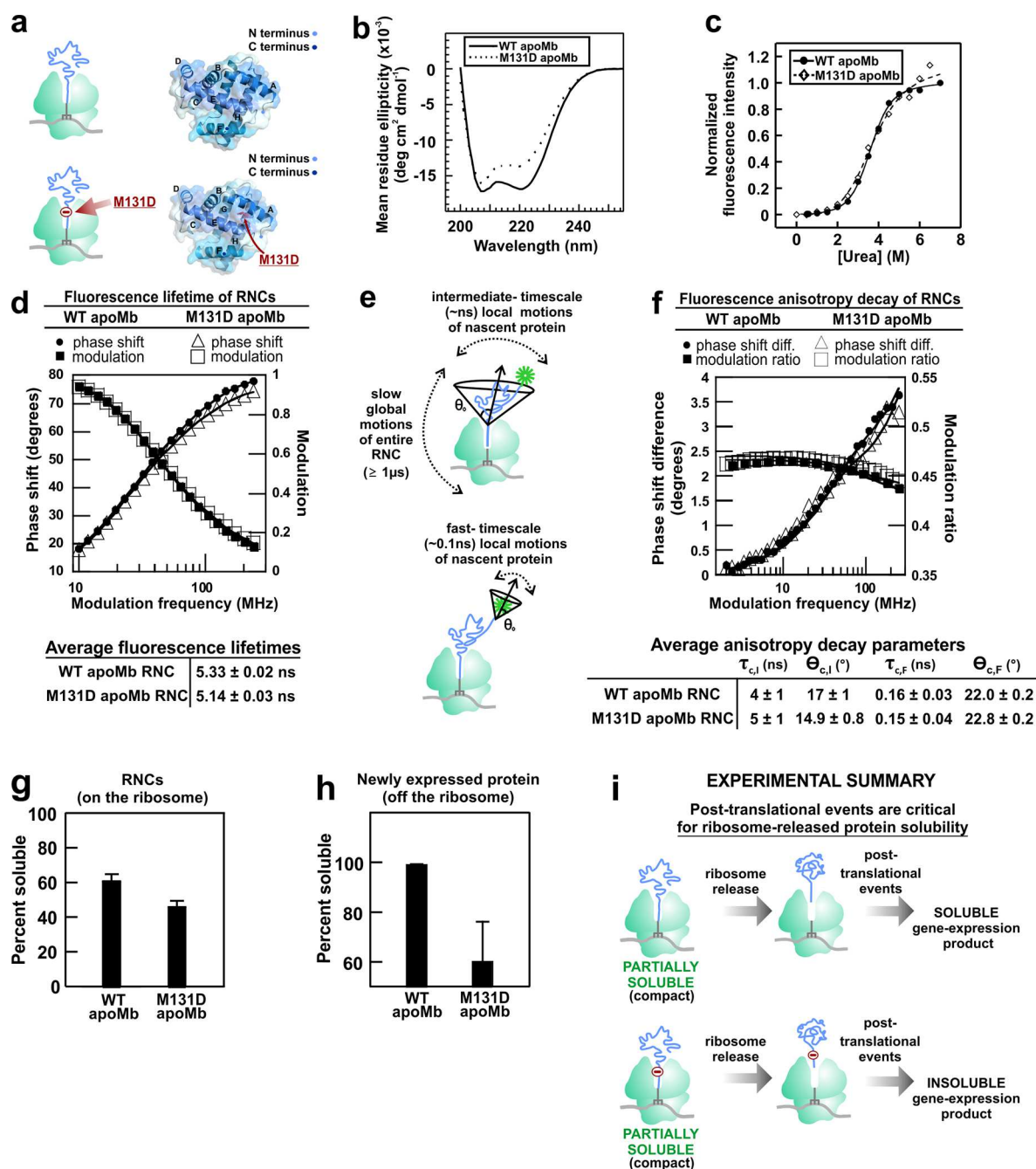


Figure 4. Biophysical Characterization of apoMb M131D variant. (a) Cartoon illustrating the approximate location of the M131D mutation within the ribosomal exit tunnel for full-length apoMb RNCs (left, bottom). Crystal structure of folded myoglobin highlighting the M131D mutation site (right, bottom, PDB code 1MBC⁹⁹). For comparison, the corresponding images for WT apoMb are also shown (top left, top right). (b) Far-UV circular dichroism spectra of purified WT apoMb (continuous line) and the M131D apoMb (dotted line) variant, each averaged over three independent experiments. (c) Representative tryptophan fluorescence urea titration curves of WT apoMb (black circles, continuous line) and M131D apoMb (open diamonds, dotted line). Data were collected in triplicates. (d) Representative frequency-domain fluorescence lifetime decay profiles of WT apoMb (black circle and squares) and M131D apoMb RNCs (open triangles and squares) in resuspension buffer (WT cell-free system, see [Methods](#)), and a table listing average fluorescence lifetimes. Data represent average ± SE for $n = 3$. (e) Cartoon illustrating the detected global and local motions experienced by ribosome bound nascent chains. θ_0 denotes the cone semiangle, an experimentally measurable parameter that characterizes the amplitude of the nascent-chain local motions. (f) Representative frequency-domain fluorescence anisotropy decay profiles of WT apoMb (black circles and squares) and M131D apoMb (open triangles and squares) RNCs in resuspension buffer (WT cell-free system, see [Methods](#)). To facilitate visualization, only every other modulation-ratio data point of WT apoMb is shown. Average anisotropy-decay parameters are reported in the Table below the plot. Data represent average ± SE for $n = 2-3$. (g) Solubility of WT and M131D apoMb RNCs in the absence of TF and DnaK. Data represent average ± SE for $n = 3-4$. (h) Solubility of full-length ribosome-released WT and M131D apoMb synthesized under -TF and -DnaK conditions. Data are reported as average ± SE for $n = 2$. (i) Cartoon highlighting the importance of post-translational events for ribosome-released protein solubility.

addition of chaperones to the medium prior to transcription–translation increases the structural accuracy of soluble newly synthesized apoMb by decreasing the overall size of the translation product. This conclusion is particularly statistically significant upon comparing items 3 and 4 in Table 1 and Table S2. The high dilution conditions of both the chaperone-free (item 3) and chaperone-containing (item 4) samples likely reduces effects due to viscosity and nonspecific binding (see also discussion below) relative to the corresponding concentrated samples (items 1 and 2 in Table 1). Hence, we propose that the diluted solution conditions enable a more explicit assessment of the presence of large and small species in solution.

The lifetime phasor plot of Figure S7a^{57,92} sheds further light on the anisotropy decay data. In general, lifetime-phasor values shift in position along a straight line if the relative amounts of two coexisting populations vary⁹² (see also Methods and brief tutorial on phasors in Supporting Information). Indeed, the lifetime-phasors of Figure S7a undergo a linear shift upon addition of increasing amounts of K/J/E. This result supports the concept that addition of molecular chaperones decreases the fraction of one species in solution. Given the matching decrease in rotational correlation time $\tau_{c,S}$, we propose that addition of K/J/E leads to a decrease in fraction of soluble aggregate. A brief tutorial on lifetime phasors is provided in the Supporting Information.

The positively charged newly expressed apoMb (in WT cell-free systems) in the presence of *E. coli* relevant concentrations of K/J/E chaperones has a somewhat large rotational correlation time (Table 1, item 2), for a monomeric folded protein. This value may either arise from the residual presence of low-order small aggregates (e.g., a folded dimer, as already reported for this protein⁹³) or from the viscous and highly crowded cell-free environment, potentially including nonspecific binding to negatively charged cell-free components.^{94–96} Interestingly, post-translational dilution (ca. 13-fold) of the translation mixture in the presence, but not in the absence, of added K/J/E chaperones (Table 1, item 4) leads to a significant decrease in observed rotational correlation time $\tau_{c,S}$. This result, combined with the fact that the soluble aggregates are kinetically trapped relative to the native state (see discussion below, especially Figure 5d), shows that dilution speeds up apoMb global tumbling due to a decrease in either solution viscosity, molecular crowding or nonspecific binding to cell-free components. These effects may either be present alone or in combination.^{96,97}

In summary, while NMR shows that ribosome-assisted biosynthesis in the absence of molecular chaperones grants only a small percent (ca. 23%) of structurally accurate protein, some or the entirety of the remaining apoMb population is expressed as a soluble aggregate. Therefore, apoMb biosynthesis via the ribosome in the absence of molecular chaperones is of paramount importance, yet it guarantees the structural accuracy of only a fraction of the total newly synthesized protein.

Mutations Close to the C Terminus Have No Effect on Some of the Major Physical Properties of the Ribosome-Released Native State and Ribosome-Bound Nascent Chains. To evaluate the relative importance of co- and post-translational protein-folding in apoMb biogenesis, we selectively disrupted post-translational folding/aggregation events. Toward this end, we designed a variant (M131D apoMb) bearing an additional negative charge in the

N-terminal region. Given that the ribosomal exit tunnel hosts 30–40 nascent-chain residues,^{68,98} the mutated apoMb site is designed to be housed within the ribosomal exit tunnel at the end of translation for apoMb₁₅₃ RNCs (see apoMb structure and mutation site in Figure 4a). In this way, the M131D mutation site is only available for conformational sampling post-translationally, after the full-length nascent protein has been released from the ribosomal exit tunnel. If cotranslational protein folding is sufficient to guarantee the solubility of ribosome-released proteins, the M131D apoMb variant is expected to be fully soluble, just like the WT protein.

In order to assess whether the M131D mutation introduces any structural variations in the folded state, we performed a series of structure and thermodynamic stability studies. We found that the far-UV circular dichroism (CD) spectra of purified native WT and M131D apoMb are similar and are consistent with predominantly α -helical secondary structures. The slightly decreased CD signal of the M131D variant indicates loss of a small fraction of α -helical content (Figure 4b), consistent with the fact that proteins undergo subtle structural rearrangements in response to the introduction of charged residues at interior positions.^{100,101} Urea denaturation profiles (Figure 4c) show that the unfolding free energies of variant and WT protein are the same within error ($\Delta G^{\circ}_{\text{NU,WT}} = 4.7 \pm 1.1$ kcal/mol, $\Delta G^{\circ}_{\text{NU,M131D}} = 4.2 \pm 1.1$ kcal/mol, $p = 0.76$). Hence the M131D mutation does not impact overall protein stability. The WT and M131D apoMb RNCs also have identical m_{NU} values ($m_{\text{NU,WT}} = -1.2 \pm 0.3$, $m_{\text{NU,M131D}} = -1.2 \pm 0.3$, $p = 0.89$). Therefore, the two species display similar differences in solvent-accessible surface area in their native and unfolded states. We conclude that the M131D mutation does not introduce any significant variations in the structure and stability of apoMb.

Ribosome-bound nascent proteins encoding both WT and variant apoMb were monitored via frequency-domain fluorometry. Fluorescence lifetimes were assessed first, given that this parameter is a sensitive probe of local electronic environment (in proximity of the fluorophore). Individual and average fluorescence lifetimes of WT and M131D apoMb RNCs were found to be similar to each other (Figure 4d and Table S3). Therefore, WT and M131D ribosome-bound nascent chains experience a comparable local electronic environment. In addition, the environment of WT and M131D apoMb RNCs is also of equal polarity, given that the wavelengths of fluorescence emission maxima of both species are indistinguishable (Figure S8). It is worth noting that the polarity of the local environment of WT and M131D apoMb does not significantly vary upon ribosome release, as shown by the wavelengths of fluorescence emission maxima, which are the same within error (Figure S8). This result suggests that these compact RNCs are not strongly influenced by the highly polar counterion layers adjacent to the ribosomal surface.⁷⁰ In contrast, RNCs of intrinsically disordered proteins show no compact conformation³ and directly interact with the ribosomal surface (Guzman-Luna et al., submitted). Conversely, the average fluorescence lifetimes of both WT and M131D significantly decrease upon ribosome release (Table S3). This finding is not surprising because fluorescence lifetimes are affected by a much broader range of environmental effects than mere polarity.

In addition, fluorescence anisotropy decays in the frequency domain were measured. Fluorescence anisotropy decays provide valuable information about local and global tumbling

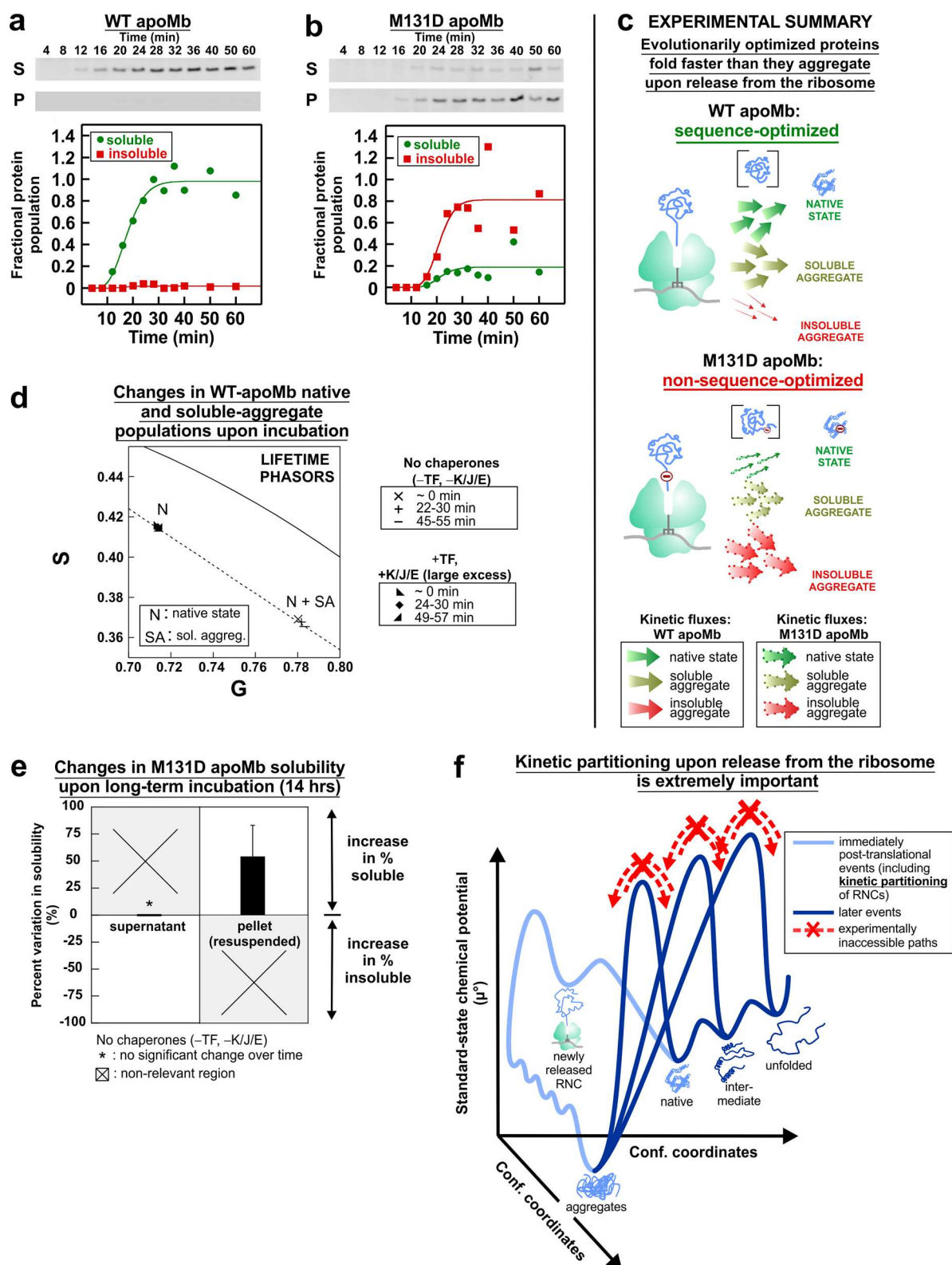


Figure 5. Kinetic partitioning and kinetic trapping upon release from the ribosome dictates nascent-protein solubility. (a) Representative SDS-PAGE analysis of the kinetics of soluble and insoluble cell-free protein production (i.e., transcription–translation) under $-TF$ and $-DnaK$ conditions (i.e., in a Δtig cell-free system treated with DnaK inhibitor) for a, WT, and b, M131D apoMb. Upper and lower lanes show data for soluble (supernatant, S) and insoluble (pellet, P) protein, respectively. Plots illustrating normalized gel-band intensities are also shown. To guide the eye, data were fit to generic kinetic expressions for multistep unimolecular reactions. Two independent experiments were carried out for each of the conditions in panels a and b. The average percentages of final soluble protein, resulting from all experiments, were $98.7 \pm 0.4\%$ and $21 \pm 2\%$ for WT and M131D apoMb, respectively. See Figure S6 for full gels of kinetic time-courses and for plots of kinetic fluxes. (c) Cartoon illustrating the fact that the relative kinetic flux toward formation of soluble native, soluble aggregated and insoluble aggregated protein is exquisitely dependent on amino acid sequence. (d) *In situ* incubation-time dependence of fluorescence-lifetime phasors of N-terminally BODIPY-FL-labeled (i) WT apoMb

Figure 5. continued

produced in an *E. coli* cell-free system in the presence of molecular chaperones (75 μ M DnaK, 0.4–12 μ M DnaJ, 19 μ M GrpE, $\sim$ 0.2 μ M Trigger factor, and (ii) WT apoMb produced in the absence of DnaK and Trigger factor. See Figure S7 for additional details. (e) Bar graph showing the change in solubility of M131D apoMb supernatant and pellet upon 14-h incubation. (f) Standard-state chemical-potential energy landscape of apoMb, highlighting that once the protein is kinetically channeled toward either the native or aggregated state immediately after release from the ribosome, it remains kinetically trapped in the respective state, barring later disaggregation or degradation events.

rates and about the spatial amplitude of local motions. The fluorescence anisotropy decays of WT and M131D RNCs showed that nascent chains of both species have a similar overall compaction. This assessment was based on a comparison between the rotational correlation times for the intermediate-time scale ($\sim$ nanosecond) motions of WT and M131D apoMb RNCs, which were deduced from fluorescence depolarization decays, and were found to be 4–5 ns for both species (Figure 4e,f Figure S1, and Table S4). Rotational correlation times are diagnostic of nascent-chain tumbling time scales, and values in the low-nanosecond regime are characteristic of compact conformations, for small-to-midsized globular proteins.⁸⁸

Cone-semiangle analysis provided additional insights. Fluorescence-anisotropy cone semiangles serve as reporters of the amplitude of local-motions^{88,2} and their numerical values are proportional to the amplitude of the motions, as schematically shown in Figure 4e. We found that the amplitude of the compact-domain motions of WT and M131D apoMb are alike, given the very similar cone semiangles (Figure 4f and Table S4). Small semiangles (ca. 20–25 deg), are diagnostic of a highly spatially constrained environment, as expected for nascent folded-protein domains close to the ribosomal surface. We deduce that RNCs of WT and M131D apoMb are comparably compact on the ribosome, and they show no detectable differences in the polarity of their environment and overall chain dynamics.

Note that detection of compact ribosome-bound nascent-protein conformations does not preclude the presence of a nascent-protein fraction that interacts with the ribosomal surface.^{3,88} This population can only be detected by fluorescence depolarization if its lifetime is significantly different from that of the compact species. This was found not to be the case for apoMb. In contrast, two populations with largely different fluorescence lifetimes and rotational dynamics could be explicitly detected in the case of the intrinsically disordered protein PIR.³

Immediately Post-translational Events Are Important for Folding. Figure 4g shows that the M131D mutation leads to only a slight decrease in the solubility of apoMb RNCs. This result is somewhat expected, given that the mutation site is buried within the ribosomal exit tunnel in M131D apoMb RNCs. However, upon release from the ribosome, the M131D variant has a higher fraction of insoluble component than WT apoMb, by ca. 40% (Figure 4h and Figure S1).

We conclude that the presence of cotranslational compaction is not sufficient to guarantee the solubility of ribosome-released medium-size single-domain proteins. Given that the M131D mutation site becomes available for conformational sampling only after the nascent chain has been cleaved from the tRNA and has partially or fully emerged out of the ribosomal exit tunnel, the onset of aggregation observed for the M131D variant must be post-translational. Hence, we deduce that post-translational events must play a critical role in the

genesis of insoluble protein aggregates, as illustrated in Figure 4i.

It is worth noticing that the above result does not undermine the importance of cotranslational protein compaction and folding (Figure 4f, Figure S1), which likely serves to decrease the extent of additional folding necessary to reach the native state.

Kinetic Channeling upon Release from the Ribosome Dictates the Solubility and Structural Accuracy of apoMb. Figure 5 shows the timecourse of soluble and insoluble full-length protein production, based on the assessment of ribosome-released gel bands, in the absence of molecular chaperones. Full representative gels are provided in Figure S6. N-terminally fluorophore-labeled WT and M131D apoMb were produced independently. Cell-free transcription–translation was quenched at various time points followed by a centrifugation assay (Figure 5, parts a and b). The bottom graphs of parts a and b of Figure 5 provide a quantitative assessment of gel-band intensities. Data were normalized to show the fraction of soluble and insoluble species, out of total protein produced at each time point. The slow buildup of protein production is consistent with the known kinetics of transcription–translation in *E. coli* cell-free systems.^{36,102} Importantly, panels c and d of Figure S6 show the corresponding kinetic fluxes, i.e., the rates of production of soluble and insoluble species.

Fully soluble WT apoMb is generated with a high kinetic flux, while the flux leading to insoluble protein is negligible. Conversely, the M131D variant displays a net inversion of fluxes, with the generation of insoluble ribosome-released protein surpassing by far the corresponding flux for the production of soluble species. These strikingly different trends are illustrated in Figure 5c, and they are particularly interesting because they cannot be ascribed to large variations in structure or changes in thermodynamic stability of either the native state or the fully synthesized ribosome-bound nascent proteins. A sequential folding mechanism, where the soluble aggregate is a precursor to the insoluble aggregate, can be ruled out (see Supporting Information).

Due to the M131D mutation-site location, i.e., buried inside the exit tunnel before translation termination, the observed inversion of fluxes is ascribed to exclusively post-translational events. Importantly, the kinetic channeling of the newly ribosome-released nascent protein (toward either a soluble or insoluble product) is effectively irreversible, given that revisiting the partially folded protein conformations populated immediately after translation termination is highly improbable. Therefore, the initial rates¹⁰³ of soluble and insoluble protein production dictate the overall amounts of soluble and insoluble products.

We conclude that, to avoid generation of insoluble aggregates, the initial kinetic flux (i.e., the initial rate) for the generation of soluble species must exceed the initial flux to generate insoluble protein (Figure 5c). This condition is experimentally fulfilled for WT apoMb (Figure 5c).

Considering the data in Figure 3, there are three major conformational pathways that proteins can undergo upon ribosome release. Namely, fold to the native state, form a soluble aggregate, or generate an insoluble aggregate. Hence, release from the ribosome is a critical stage in the life of a protein, as it dictates whether a protein will be soluble or insoluble and, if soluble, whether it will be 100% folded.

We propose that, upon release from the ribosome, the M131D variant is not able to efficiently bury its hydrophobic residues, resulting in slower folding and faster aggregation than WT apoMb. Consistent with this idea, it was previously reported that placing a charged residue within the hydrophobic core of a protein destabilizes folding intermediates, slowing down folding.¹⁰⁴

Native State and Soluble Aggregates Are Kinetically Trapped Relative to Each Other, after Their Initial Biogenesis. In order to test for changes in native-state and soluble-aggregate populations over time, we performed fluorescence phasor analysis on newly synthesized ribosome-release protein.⁵⁷ The results are shown in Figure 5d. Lifetime-phasor values are expected to shift in position along a straight line if the relative amounts of two coexisting populations (in this case native state and soluble aggregate) vary with time.⁹² An overview of fluorescence lifetime-phasors is provided in the Supporting Information text. More details are available in the literature.⁵⁷

When ribosome-released apoMb is generated in the presence of chaperones (TF and a large excess of K/J/E), the native state is predominantly produced (Figure 5d). Conversely, in the absence of molecular chaperones a different species is generated. This species is assigned to a soluble aggregate, given its larger observed rotational correlation time, as assessed by fluorescence anisotropy decay in the frequency domain. As illustrated in Figure 5d, the fluorescence lifetime phasors of WT apoMb do not shift in position (along the straight line) upon incubation. Hence the relative populations of the native state and soluble aggregate do not change after ribosome release, strongly suggesting that the native state and soluble aggregate are kinetically trapped with respect to one another at room temperature on time scales longer than the doubling time of *E. coli* (e.g., ≥ 20 –30 min). The above is consistent with recent data on purified apoMb and the majority of *E. coli* proteins.^{105,106} In addition, the *in vitro* literature on purified monomeric refolded apoMb reports no aggregate formation upon incubation.^{60,76} Note that apoMb is kinetically trapped relative to aggregates, and not relative to its unfolded state. In the experiments shown here, there is neither post-translational disaggregation nor covalent degradation of soluble aggregates by disaggregation factors or degradases, on the ~ 1 h experimental time scale. See Figure S1 for a representative SDS-PAGE gel demonstrating sample integrity and purity.

As previously discussed, the M131D apoMb variant has a strong tendency to form both soluble and insoluble species immediately upon release from the ribosome (Figure 5b). Importantly, after long-term incubation, the soluble species does not further evolve into insoluble protein (Figure 5e). Interestingly, long-term incubation of the M131D apoMb insoluble fraction led to partial resolubilization in the cell-free environment (Figure 5e). This effect may be due to protein disaggregation by ClpB,^{33,107} and the presence of a DnaK inhibitor in our cell-free system may be responsible for the slow observed disaggregation rates.^{108,109}

Together, the above experiments underscore the importance of the irreversible kinetic steps that immediately follow translation termination, as illustrated in Figure 5f. Importantly, once the native state is formed for the first time, it is protected from aggregation by the kinetic-trapping effect.^{105,106} Thus, post-translational kinetic trapping of the native state (relative to soluble and insoluble aggregates) is beneficial to proteins and likely plays a protective role in Nature.

DISCUSSION

In this work, we identified critical co- and immediately post-translational events taking place in the bacterial translation machinery upon expression of apoMb, a model medium-size single-domain protein. The main take-home messages are illustrated in Figure 6 and summarized below.

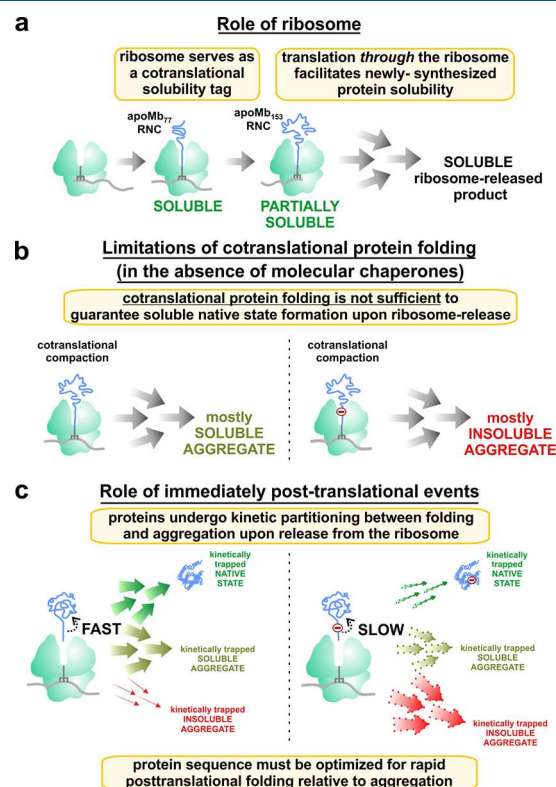


Figure 6. Cartoon summarizing the complementary role of co- and post-translational events in protein biogenesis. This cartoon emphasizes the role of the ribosome in the absence of cotranslationally active molecular chaperones. (a) The ribosome facilitates protein solubility during translation. This process facilitates protein solubility upon release from the ribosome. (b) Cotranslational formation of a compact structure is not sufficient to guarantee solubility and native state formation upon release from the ribosome. (c) As they depart from the ribosome, proteins kinetically partition between folding and aggregation routes. In order to successfully attain the native state, proteins must bear evolutionarily advanced amino acid sequences so that they fold faster than aggregating, upon release from the ribosome.

First, the ribosome enhances cotranslational nascent-protein solubility, thus supporting intramolecular compaction and conformational sampling during protein biosynthesis (Figure 6a). Second, the ribosomal machinery is not sufficient to produce 100% native protein upon release from the ribosome in the absence of molecular chaperones (Figure 6b). Third, immediately after translation termination, an irreversible

kinetic competition between completion of protein folding and aggregation takes place. During this critical step, proteins achieve their native state only if they fold faster than aggregating (Figure 6c).

Our data show that the kinetic balance between protein-folding completion and aggregation upon release from the ribosome is much more sensitive to amino acid sequence than the final folded state. The above results identify (a) explicit constraints for the successful folding of apoMb, and possibly other single-domain proteins, and (b) a novel source of evolutionary pressure for primary structure optimization.

Incompletely synthesized nascent chains derived from the single-domain globin apoMb are aggregation-prone and marginally stable, in isolation.³⁶ Furthermore, significant thermodynamic stability is only expected when the last C-terminal residues of a globular protein domain become available for conformational sampling.¹¹⁰ Thus, there is a compelling need to keep nascent proteins soluble during translation, and the ribosome is a major contributor to achieve this goal (Figure 1c,e).

WT apoMb is most efficiently generated in fully soluble form (at $\sim 0.3 \mu\text{M}$ total protein concentration) when it emerges through the ribosome (Figure 2), yet a fraction of this protein is present in solution as a soluble aggregate (Figure 3). Thus, even at physiologically relevant expression levels, biosynthesis by the ribosome is insufficient to guarantee successful protein folding. Hence, it is important to realize that protein solubility does not guarantee structural accuracy.

The frequent cytotoxic nature of soluble aggregates¹¹¹ motivates the need to prevent their formation and highlights the importance of molecular chaperones for the biogenesis of folded proteins. In addition, the heme cofactor, which is capable of cotranslational association with nascent globins^{112,113} but is not required within the final apoMb folded structure,⁵⁹ may play a role in enhancing the soluble native-state population upon release from the ribosome.

Only 30% of *E. coli* proteins are known to express in soluble form when biosynthesized by the ribosome in the absence of molecular chaperones (at ca. $33 \mu\text{g/mL}$ average protein concentration).¹¹⁴ On the other hand, our work suggests that a portion of the soluble proteome generated without molecular chaperones is expressed as soluble aggregate.

The M131D apoMb variant has a similar structure and thermodynamic stability to WT apoMb (Figure 4b,c). However, the relative rates of folding-completion and aggregation upon release from the ribosome are not optimized for this variant (Figure 5a,b). Hence, the relative rates of folding and aggregation upon ribosome release are much more sensitive to amino acid sequence than protein structure and thermodynamic stability.

Protein aggregation is generally assumed to occur later than biosynthesis.¹¹⁵ However, after ribosome-release, kinetic trapping of the native state relative to soluble and insoluble aggregates is known to protect apoMb and many other proteins from aggregation.¹⁰⁵

We show here that the most aggregation-sensitive steps of apoMb's life are not later in life, but immediately after nascent-protein release from the ribosome. At this stage, the parallel fluxes leading to folding or aggregation can be large, comparable in magnitude, and easily tunable via variations in amino acid sequence (Figure 5 and Figure S6). Later steps lead to no further aggregation. In short, apoMb is more likely to aggregate immediately upon release from the ribosome than

later, when expressed at a concentration relevant to regular expression levels in *E. coli*.

Protein expression levels are known to be inversely proportional to aggregation propensities.¹¹⁶ We propose that this concept can be further refined for single-domain proteins, upon defining relevant aggregation propensities as the ratio between protein folding and aggregation fluxes immediately upon release from the ribosome. In essence, we propose that protein expression levels have been evolutionarily tuned to prevent aggregation immediately after protein biosynthesis, in addition to later in life.

Protein evolution *in vivo* is believed to proceed via errors in biogenesis (e.g., mistakes in transcription, translation)¹¹⁷ and may be driven by multiple factors including folding kinetics and the achievement of structures compatible with function.^{29,118,119}

Here, we propose that single-domain proteins like apoMb undergo an additional specific element of evolutionary pressure, i.e., the drive to fold faster than aggregating upon release from the ribosome. In this way, population of the correct region of the standard-state chemical-potential landscape is ensured (Figure 5f), enabling function. Interactions with molecular chaperones may tune the above kinetic balance by transiently capturing the monomeric *de novo* synthesized protein. The resulting transient decrease in free-protein population is expected to slow down the aggregation flux more than the folding flux, given the steeper dependence of aggregation kinetics on free-protein concentration.

We speculate that protein folding (instead of aggregation) upon release from the ribosome should ideally proceed with 100% efficiency in live cells. This is because, despite the protective role of kinetic trapping,¹⁰⁵ even small traces of some aggregates may serve as seeds and enable further post-translational aggregation,¹⁰⁵ requiring the intervention of energetically expensive disaggregation³³ and/or degradation¹²⁰ machineries.

CONCLUSIONS

Here, we employed apoMb and its M131D variant as model systems to address the importance of the ribosome as well as the relative role of co- and post-translational events in protein folding. This work highlights the role of the bacterial ribosome as a nascent-protein (i.e., RNC) solubilizing agent. The ribosome was also shown to be a required conduit for the generation of soluble protein after completion of translation. On the other hand, molecular chaperones are needed to enhance the population of newly expressed proteins, thus contributing to their structural accuracy within the solution ensemble. Importantly, we showed that the kinetic steps immediately following translation termination include parallel paths leading to post-translational folding and aggregation. The kinetic channeling toward either a fully soluble or fully insoluble protein is highly dependent upon amino acid sequence and can be readily reversed upon point mutation. This inversion in kinetic channeling takes place under conditions where thermodynamic stability and structure of the native state stay essentially unchanged. The latter finding underscores the importance of kinetic control upon ribosome release across evolution of protein sequence and folding.

In the case of holo-proteins, the heme cofactor may further influence folding and aggregation outcomes, and its role will be addressed in future studies. Extension of our findings beyond the globin fold will also be addressed elsewhere.

■ ASSOCIATED CONTENT

■ Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpcb.0c03039>.

Effect of centrifugation parameters on protein solubility, relations between physical properties of nascent proteins and experimentally observed solubility, apparent thermodynamic stability of WT and M131D apoMb, justification for the presence of parallel post-translational pathways leading to soluble and insoluble protein upon release from the ribosome, the use of fluorescence lifetime phasors for identifying the presence of different protein conformations in solution, representative SDS Page gels, concentration control for apoMb77, chain-length-dependent properties of apoMb RNCs, control on the role of the ribosome in protein solubility, spectral center-of-mass fluorescence-emission analysis, representative SDS PAGE transcription–translation timecourses, fluorescence lifetime phasor evidence for kinetic trapping of the native state and soluble-aggregate states of apoMb₁₅₃, fluorescence emission spectra of labeled RNCs and corresponding ribosome-released proteins, summary table of ¹H, ¹⁵N HSQC NMR analysis, fluorescence lifetime decay parameters of ribosome-bound and ribosome-released nascent chains, fluorescence anisotropy decay parameters of ribosome-bound and ribosome-released nascent chains, and a table of *p*-values of ribosome-released nascent-protein global rotational correlation times (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Silvia Cavagnero – Biophysics Graduate Degree Program and Department of Chemistry, University of Wisconsin–Madison, Madison, Wisconsin 53706, United States; orcid.org/0000-0002-4290-2331; Phone: 608- 262-5430; Email: cavagnero@chem.wisc.edu

Authors

Rayna M. Addabbo – Biophysics Graduate Degree Program, University of Wisconsin–Madison, Madison, Wisconsin 53706, United States

Matthew D. Dalphin – Biophysics Graduate Degree Program, University of Wisconsin–Madison, Madison, Wisconsin 53706, United States

Miranda F. Mecha – Biophysics Graduate Degree Program, University of Wisconsin–Madison, Madison, Wisconsin 53706, United States

Yue Liu – Department of Chemistry, University of Wisconsin–Madison, Madison, Wisconsin 53706, United States

Alexios Staikos – Department of Chemistry, University of Wisconsin–Madison, Madison, Wisconsin 53706, United States

Valeria Guzman-Luna – Department of Chemistry, University of Wisconsin–Madison, Madison, Wisconsin 53706, United States

Complete contact information is available at:

<https://pubs.acs.org/doi/10.1021/acs.jpcb.0c03039>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors are grateful to the National Science Foundation (NSF) for funding (Grants MCB-1616459 and CBET 1912259 to S.C.). In addition, R.M.A. thanks the National Science Foundation for a Graduate Research Fellowship (Grant DGE-1256259). The authors are also thankful for a grant from the Graduate School and the Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin-Madison (to S.C.), and for an award from the Dorothy Powelson Scholarship Fund (to R.M.A.). Finally, the authors thank all the members of the Cavagnero research group for helpful discussions and Anna Allen, Teddy Jennaro, and Andrew Fuchs for their invaluable assistance with the generation of computer graphics.

■ REFERENCES

- (1) Ellis, J. P.; Bakke, C. K.; Kirchdoerfer, R. N.; Jungbauer, L. M.; Cavagnero, S. Chain Dynamics of Nascent Polypeptides Emerging from the Ribosome. *ACS Chem. Biol.* **2008**, *3*, 555–566.
- (2) Ellis, J. P.; Culviner, P. H.; Cavagnero, S. Confined Dynamics of a Ribosome-Bound Nascent Globin: Cone Angle Analysis of Fluorescence Depolarization Decays in the Presence of Two Local Motions. *Protein Sci.* **2009**, *18*, 2003–2015.
- (3) Knight, A. M.; Culviner, P. H.; Kurt-Yilmaz, N. e.; Zou, T.; Ozkan, S. B.; Cavagnero, S. Electrostatic Effect of the Ribosomal Surface on Nascent Polypeptide Dynamics. *ACS Chem. Biol.* **2013**, *8*, 1195–1204.
- (4) Holtkamp, W.; Kokic, G.; Jäger, M.; Mittelstaet, J.; Komar, A. A.; Rodnina, M. V. Cotranslational Protein Folding on the Ribosome Monitored in Real Time. *Science* **2015**, *350*, 1104–1107.
- (5) Kim, S. J.; Yoon, J. S.; Shishido, H.; Yang, Z.; Rooney, L. A.; Barral, J. M.; Skach, W. R. Translational Tuning Optimizes Nascent Protein Folding in Cells. *Science* **2015**, *348*, 444–448.
- (6) Kaiser, C. M.; Goldman, D. H.; Chodera, J. D.; Tinoco, I.; Bustamante, C. The Ribosome Modulates Nascent Protein Folding. *Science* **2011**, *334*, 1723–1727.
- (7) Evans, M. S.; Sander, I. M.; Clark, P. L. Cotranslational Folding Promotes B-Helix Formation and Avoids Aggregation in Vivo. *J. Mol. Biol.* **2008**, *383*, 683–692.
- (8) Clark, P. L.; King, J. A Newly Synthesized, Ribosome-Bound Polypeptide Chain Adopts Conformations Dissimilar from Early in Vitro Refolding Intermediates. *J. Biol. Chem.* **2001**, *276*, 25411–25420.
- (9) Friguet, B.; Djavadi-Ohanian, L.; King, J.; Goldberg, M. E. In Vitro and Ribosome-Bound Folding Intermediates of P22 Tailspike Protein Detected with Monoclonal Antibodies. *J. Biol. Chem.* **1994**, *269*, 15945–15949.
- (10) Rajabi, K.; Reuther, J.; Deuerling, E.; Radford, S. E.; Ashcroft, A. E. A Comparison of the Folding Characteristics of Free and Ribosome-Tethered Polypeptide Chains Using Limited Proteolysis and Mass Spectrometry. *Protein Sci.* **2015**, *24*, 1282–1291.
- (11) Samelson, A. J.; Jensen, M. K.; Soto, R. A.; Cate, J. H.; Marqusee, S. Quantitative Determination of Ribosome Nascent Chain Stability. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 13402–13407.
- (12) Eichmann, C.; Preissler, S.; Riek, R.; Deuerling, E. Cotranslational Structure Acquisition of Nascent Polypeptides Monitored by Nmr Spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 9111–9116.
- (13) Cabrita, L. D.; Cassaignau, A. M.; Launay, H. M.; Waudby, C. A.; Wlodarski, T.; Camilloni, C.; Karyadi, M.-E.; Robertson, A. L.; Wang, X.; Wentink, A. S.; et al. A Structural Ensemble of a Ribosome–Nascent Chain Complex During Cotranslational Protein Folding. *Nat. Struct. Mol. Biol.* **2016**, *23*, 278–285.
- (14) Deckert, A.; Waudby, C. A.; Wlodarski, T.; Wentink, A. S.; Wang, X.; Kirkpatrick, J. P.; Paton, J. F.; Camilloni, C.; Kukic, P.; Dobson, C. M.; et al. Structural Characterization of the Interaction of

A-Synuclein Nascent Chains with the Ribosomal Surface and Trigger Factor. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 5012–5017.

(15) Trovato, F.; O'Brien, E. P. Insights into Cotranslational Nascent Protein Behavior from Computer Simulations. *Annu. Rev. Biophys.* **2016**, *45*, 345–369.

(16) Nilsson, O. B.; Nickson, A. A.; Hollins, J. J.; Wickles, S.; Steward, A.; Beckmann, R.; von Heijne, G.; Clarke, J. Cotranslational Folding of Spectrin Domains Via Partially Structured States. *Nat. Struct. Mol. Biol.* **2017**, *24*, 221–225.

(17) Tian, P. F.; Steward, A.; Kudva, R.; Su, T.; Shilling, P. J.; Nickson, A. A.; Hollins, J. J.; Beckmann, R.; von Heijne, G.; Clarke, J.; Best, R. B. Folding Pathway of an Ig Domain Is Conserved on and Off the Ribosome. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, E11284–E11293.

(18) Eisner, G.; Koch, H.-G.; Beck, K.; Brunner, J.; Müller, M. Ligand Crowding at a Nascent Signal Sequence. *J. Cell Biol.* **2003**, *163*, 35–44.

(19) Ullers, R. S.; Houben, E. N.; Raine, A.; ten Hagen-Jongman, C. M.; Ehrenberg, M.; Brunner, J.; Oudega, B.; Harms, N.; Luijck, J. Interplay of Signal Recognition Particle and Trigger Factor at L23 near the Nascent Chain Exit Site on the Escherichia Coli Ribosome. *J. Cell Biol.* **2003**, *161*, 679–684.

(20) Eisner, G.; Moser, M.; Schäfer, U.; Beck, K.; Müller, M. Alternate Recruitment of Signal Recognition Particle and Trigger Factor to the Signal Sequence of a Growing Nascent Polypeptide. *J. Biol. Chem.* **2006**, *281*, 7172–7179.

(21) Peterson, J. H.; Woolhead, C. A.; Bernstein, H. D. The Conformation of a Nascent Polypeptide inside the Ribosome Tunnel Affects Protein Targeting and Protein Folding. *Mol. Microbiol.* **2010**, *78*, 203–217.

(22) Hsu, S.-T. D.; Fucini, P.; Cabrita, L. D.; Launay, H.; Dobson, C. M.; Christodoulou, J. Structure and Dynamics of a Ribosome-Bound Nascent Chain by Nmr Spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 16516–16521.

(23) Fedorov, A. N.; Baldwin, T. O. Contribution of Cotranslational Folding to the Rate of Formation of Native Protein Structure. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 1227–1231.

(24) Ugrinov, K. G.; Clark, P. L. Cotranslational Folding Increases Gfp Folding Yield. *Biophys. J.* **2010**, *98*, 1312–1320.

(25) Kelkar, D. A.; Khushoo, A.; Yang, Z.; Skach, W. R. Kinetic Analysis of Ribosome-Bound Fluorescent Proteins Reveals an Early, Stable, Cotranslational Folding Intermediate. *J. Biol. Chem.* **2012**, *287*, 2568–2578.

(26) Frydman, J.; Nimmegern, E.; Ohtsuka, K.; Hartl, F. U. Folding of Nascent Polypeptide Chains in a High Molecular Mass Assembly with Molecular Chaperones. *Nature* **1994**, *370*, 111–117.

(27) Kudlicki, W.; Odom, O. W.; Kramer, G.; Hardesty, B. Activation and Release of Enzymatically Inactive, Full-Length Rhodanese That Is Bound to Ribosomes as Peptidyl-Trna. *J. Biol. Chem.* **1994**, *269*, 16549–16553.

(28) Frydman, J.; Erdjument-Bromage, H.; Tempst, P.; Hartl, F. U. Co-Translational Domain Folding as the Structural Basis for the Rapid De Novo Folding of Firefly Luciferase. *Nat. Struct. Biol.* **1999**, *6*, 697–705.

(29) Debès, C.; Wang, M.; Caetano-Anollés, G.; Gräter, F. Evolutionary Optimization of Protein Folding. *PLoS Comput. Biol.* **2013**, *9*, e1002861.

(30) Kirchdoerfer, R. N.; Huang, J. J. T.; Isola, M. K.; Cavagnero, S. Fluorescence-Based Analysis of Aminoacyl- and Peptidyl-Trna by Low-Ph Sodium Dodecyl Sulfate–Polyacrylamide Gel Electrophoresis. *Anal. Biochem.* **2007**, *364*, 92–94.

(31) Kim, Y. E.; Hipp, M. S.; Bracher, A.; Hayer-Hartl, M.; Ulrich Hartl, F. Molecular Chaperone Functions in Protein Folding and Proteostasis. *Annu. Rev. Biochem.* **2013**, *82*, 323–355.

(32) Balchin, D.; Hayer-Hartl, M.; Hartl, F. U. In Vivo Aspects of Protein Folding and Quality Control. *Science* **2016**, *353*, aac4354.

(33) Mogk, A.; Bukau, B.; Kampinga, H. H. Cellular Handling of Protein Aggregates by Disaggregation Machines. *Mol. Cell* **2018**, *69*, 214–226.

(34) Doyle, S. M.; Genest, O.; Wickner, S. Protein Rescue from Aggregates by Powerful Molecular Chaperone Machines. *Nat. Rev. Mol. Cell Biol.* **2013**, *14*, 617–629.

(35) Springer, B. A.; Sligar, S. G. High-Level Expression of Spermin Whale Myoglobin in Escherichia Coli. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84*, 8961–8965.

(36) Chow, C. C.; Chow, C.; Raghunathan, V.; Huppert, T. J.; Kimball, E. B.; Cavagnero, S. Chain Length Dependence of Apomyoglobin Folding: Structural Evolution from Misfolded Sheets to Native Helices. *Biochemistry* **2003**, *42*, 7090–7099.

(37) Bakke, C. K.; Jungbauer, L. M.; Cavagnero, S. In Vitro Expression and Characterization of Native Apomyoglobin under Low Molecular Crowding Conditions. *Protein Expression Purif.* **2006**, *45*, 381–392.

(38) Dalphin, M. D.; Stangl, A. J.; Liu, Y.; Cavagnero, S. Klr-70: A Novel Cationic Inhibitor of the Bacterial Hsp70 Chaperone. *Biochemistry* **2020**, *59*, 1946.

(39) Jennings, P. A.; Stone, M. J.; Wright, P. E. Overexpression of Myoglobin and Assignment of Its Amide, C α and C β Resonances. *J. Biomol. NMR* **1995**, *6*, 271–276.

(40) Schneider, C. A.; Rasband, W. S.; Eliceiri, K. W. Nih Image to Imagej: 25 Years of Image Analysis. *Nat. Methods* **2012**, *9*, 671–675.

(41) Abramoff, M. D.; Magalhães, P. J.; Ram, S. J. Image Processing with Imagej. *Biophotonics international* **2004**, *11*, 36–42.

(42) Grzesiek, S.; Bax, A. The Importance of Not Saturating Water in Protein Nmr. Application to Sensitivity Enhancement and Noe Measurements. *J. Am. Chem. Soc.* **1993**, *115*, 12593–12594.

(43) Kay, L.; Keifer, P.; Saarinen, T. Pure Absorption Gradient Enhanced Heteronuclear Single Quantum Correlation Spectroscopy with Improved Sensitivity. *J. Am. Chem. Soc.* **1992**, *114*, 10663–10665.

(44) Palmer, A. G., III; Cavanagh, J.; Wright, P. E.; Rance, M. Sensitivity Improvement in Proton-Detected Two-Dimensional Heteronuclear Correlation Nmr Spectroscopy. *J. Magn. Reson.* **1991**, *93*, 151–170.

(45) Schleucher, J.; Schwendinger, M.; Sattler, M.; Schmidt, P.; Schedletsky, O.; Glaser, S.; Sørensen, O.; Griesinger, C. A General Enhancement Scheme in Heteronuclear Multidimensional Nmr Employing Pulsed Field Gradients. *J. Biomol. NMR* **1994**, *4*, 301–306.

(46) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. Nmrpipe: A Multidimensional Spectral Processing System Based on Unix Pipes. *J. Biomol. NMR* **1995**, *6*, 277–293.

(47) Johnson, B. A.; Blevins, R. A. Nmr View: A Computer Program for the Visualization and Analysis of Nmr Data. *J. Biomol. NMR* **1994**, *4*, 603–614.

(48) Lee, J. H.; Zhang, D. Y.; Hughes, C.; Okuno, Y.; Sekhar, A.; Cavagnero, S. Heterogeneous Binding of the Sh3 Client Protein to the Dnak Molecular Chaperone. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, E4206–E4215.

(49) Kurt, N.; Rajagopalan, S.; Cavagnero, S. Effect of Hsp70 Chaperone on the Folding and Misfolding of Polypeptides Modeling an Elongating Protein Chain. *J. Mol. Biol.* **2006**, *355*, 809–820.

(50) Pace, C. N. [14]Determination and Analysis of Urea and Guanidine Hydrochloride Denaturation Curves. *Methods Enzymol.* **1986**, *131*, 266–280.

(51) Jameson, D. M. *Introduction to Fluorescence*; CRC Press: 2014.

(52) Silva, J. L.; Miles, E. W.; Weber, G. Pressure Dissociation and Conformational Drift of The.Beta. Dimer of Tryptophan Synthase. *Biochemistry* **1986**, *25*, 5780–5786.

(53) Paladini, A. A.; Weber, G. Pressure-Induced Reversible Dissociation of Enolase. *Biochemistry* **1981**, *20*, 2587–2593.

(54) Santoro, M. M.; Bolen, D. W. Unfolding Free Energy Changes Determined by the Linear Extrapolation Method. 1. Unfolding of Phenylmethanesulfonyl Alpha-Chymotrypsin Using Different Denaturants. *Biochemistry* **1988**, *27*, 8063–8068.

(55) Pace, C. N. Measuring and Increasing Protein Stability. *Trends Biotechnol.* **1990**, *8*, 93–8.

- (56) Wohlgemuth, I.; Beringer, M.; Rodnina, M. V. Rapid Peptide Bond Formation on Isolated 50s Ribosomal Subunits. *EMBO Rep.* **2006**, *7*, 699–703.
- (57) Stefl, M.; James, N. G.; Ross, J. A.; Jameson, D. M. Applications of Phasors to in Vitro Time-Resolved Fluorescence Measurements. *Anal. Biochem.* **2011**, *410*, 62–69.
- (58) Ishihama, Y.; Schmidt, T.; Rappsilber, J.; Mann, M.; Hartl, F. U.; Kerner, M. J.; Frishman, D. Protein Abundance Profiling of the Escherichia Coli Cytosol. *BMC Genomics* **2008**, *9*, 102.
- (59) Eliezer, D.; Wright, P. E. Is Apomyoglobin a Molten Globule? Structural Characterization by Nmr. *J. Mol. Biol.* **1996**, *263*, 531–538.
- (60) Dyson, H. J.; Wright, P. E. How Does Your Protein Fold? Elucidating the Apomyoglobin Folding Pathway. *Acc. Chem. Res.* **2017**, *50*, 105–111.
- (61) Cavagnero, S.; Nishimura, C.; Schwarzing, S.; Dyson, H. J.; Wright, P. E. Conformation and Dynamic Characterization of the Molten Globule State of an Apomyoglobin Mutant with an Altered Folding Pathway. *Biochemistry* **2001**, *40*, 14459–14467.
- (62) Garcia, C.; Nishimura, C.; Cavagnero, S.; Dyson, H. J.; Wright, P. E. Changes in the Apomyoglobin Folding Pathway Caused by Mutation of the Distal Histidine Residue. *Biochemistry* **2000**, *39*, 11227–11237.
- (63) Cavagnero, S.; Dyson, H. J.; Wright, P. E. Effect of H Helix Destabilizing Mutations on the Kinetic and Equilibrium Folding of Apomyoglobin. *J. Mol. Biol.* **1999**, *285*, 269–282.
- (64) Tsui, V.; Garcia, C.; Cavagnero, S.; Siuzdak, G.; Dyson, H. J.; Wright, P. E. Quench-Flow Experiments Combined with Mass Spectrometry Show Apomyoglobin Folds through an Obligatory Intermediate. *Protein Sci.* **1999**, *8*, 45–49.
- (65) Eun, Y. J.; Kurt, N.; Sekhar, A.; Cavagnero, S. Thermodynamic and Kinetic Characterization of Apohmp, a Fast-Folding Bacterial Globin. *J. Mol. Biol.* **2008**, *376*, 879–897.
- (66) Samelson, A. J.; Bolin, E.; Costello, S. M.; Sharma, A. K.; O'Brien, E. P.; Marqusee, S. Kinetic and Structural Comparison of a Protein's Cotranslational Folding and Refolding Pathways. *Science Advances* **2018**, *4*, eaas9098.
- (67) Thommen, M.; Holtkamp, W.; Rodnina, M. V. Co-Translational Protein Folding: Progress and Methods. *Curr. Opin. Struct. Biol.* **2017**, *42*, 83–89.
- (68) Malkin, L. I.; Rich, A. Partial Resistance of Nascent Polypeptide Chains to Proteolytic Digestion Due to Ribosomal Shielding. *J. Mol. Biol.* **1967**, *26*, 329–346.
- (69) Niwa, T.; Kanamori, T.; Ueda, T.; Taguchi, H. Global Analysis of Chaperone Effects Using a Reconstituted Cell-Free Translation System. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 8937–8942.
- (70) Fedykina, D. V.; Jennaro, T. S.; Cavagnero, S. Charge Segregation and Low Hydrophobicity Are Key Features of Ribosomal Proteins from Different Organisms. *J. Biol. Chem.* **2014**, *289*, 6740–6750.
- (71) Kapust, R. B.; Waugh, D. S. Escherichia Coli Maltose-Binding Protein Is Uncommonly Effective at Promoting the Solubility of Polypeptides to Which It Is Fused. *Protein Sci.* **1999**, *8*, 1668–1674.
- (72) Kurt, S.; Cavagnero, S. The Burial of Solvent-Accessible Surface Area Is a Predictor of Polypeptide Folding and Misfolding as a Function of Chain Elongation. *J. Am. Chem. Soc.* **2005**, *127*, 15690–15691.
- (73) Cavagnero, S.; Kurt, N. Folding and Misfolding as a Function of Polypeptide Chain Elongation: Conformational Trends and Implications for Intracellular Events. *Misbehaving Proteins: Protein (Mis)Folding, Aggregation and Stability* **2006**, 217–246.
- (74) Mounce, B.; Kurt, N.; Ellison, P. A.; Cavagnero, S. Nonrandom Distribution of Intramolecular Contacts in Native Single-Domain Proteins. *Proteins: Struct., Funct., Genet.* **2009**, *75*, 404–412.
- (75) Jennaro, T. S.; Beaty, M. R.; Kurt-Yilmaz, N.; Luskin, B. L.; Cavagnero, S. Burial of Nonpolar Surface Area and Thermodynamic Stabilization of Globins as a Function of Chain Elongation. *Proteins: Struct., Funct., Genet.* **2014**, *82*, 2318–2331.
- (76) Chow, C.; Kurt, N.; Murphy, R. M.; Cavagnero, S. Structural Characterization of Apomyoglobin Self-Associated Species in Aqueous Buffer and Urea Solution. *Biophys. J.* **2006**, *90*, 298–309.
- (77) Willis, M. S.; Hogan, J. K.; Prabhakar, P.; Liu, X.; Tsai, K.; Wei, Y. Y.; Fox, T. Investigation of Protein Refolding Using a Fractional Factorial Screen: A Study of Reagent Effects and Interactions. *Protein Sci.* **2005**, *14*, 1818–1826.
- (78) Cowieson, N. P.; Wensley, B.; Listwan, P.; Hume, D. A.; Kobe, B.; Martin, J. L. An Automatable Screen for the Rapid Identification of Proteins Amenable to Refolding. *Proteomics* **2006**, *6*, 1750–1757.
- (79) Chow, C. K.; Kurt, N.; Murphy, R. M.; Cavagnero, S. Structural Characterization of Apomyoglobin Self-Associated Species in Aqueous Buffer and Urea Solution. *Biophys. J.* **2006**, *90*, 298–309.
- (80) Das, B.; Chattopadhyay, S.; Gupta, C. D. Reactivation of Denatured Fungal Glucose 6-Phosphate Dehydrogenase and E. Coli Alkaline Phosphatase with E. Coli Ribosome. *Biochem. Biophys. Res. Commun.* **1992**, *183*, 774–780.
- (81) Chattopadhyay, S.; Das, B.; Dasgupta, C. Reactivation of Denatured Proteins by 23s Ribosomal Rna: Role of Domain V. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 8284–8287.
- (82) Kudlicki, W.; Coffman, A.; Kramer, G.; Hardesty, B. Ribosomes and Ribosomal Rna as Chaperones for Folding of Proteins. *Folding Des.* **1997**, *2*, 101–108.
- (83) Das, D.; Das, A.; Samanta, D.; Ghosh, J.; Dasgupta, S.; Bhattacharya, A.; Basu, A.; Sanyal, S.; Das Gupta, C. Role of the Ribosome in Protein Folding. *Biotechnol. J.* **2008**, *3*, 999–1009.
- (84) Samanta, D.; Mukhopadhyay, D.; Chowdhury, S.; Ghosh, J.; Pal, S.; Basu, A.; Bhattacharya, A.; Das, A.; Das, D.; DasGupta, C. Protein Folding by Domain V of Escherichia Coli 23s Rna: Specificity of Rna-Protein Interactions. *J. Bacteriol.* **2008**, *190*, 3344–3352.
- (85) Voisset, C.; Saupe, S. J.; Blondel, M. The Various Facets of the Protein-Folding Activity of the Ribosome. *Biotechnol. J.* **2011**, *6*, 668–673.
- (86) Banerjee, D.; Sanyal, S. Protein Folding Activity of the Ribosome (Pfar)—a Target for Antiprion Compounds. *Viruses* **2014**, *6*, 3907–3924.
- (87) Ozawa, K.; Headlam, M. J.; Schaeffer, P. M.; Henderson, B. R.; Dixon, N. E.; Otting, G. Optimization of an Escherichia Coli System for Cell-Free Synthesis of Selectively ¹⁵N-Labelled Proteins for Rapid Analysis by Nmr Spectroscopy. *Eur. J. Biochem.* **2004**, *271*, 4084–4093.
- (88) Weinreis, S. A.; Ellis, J. P.; Cavagnero, S. Dynamic Fluorescence Depolarization: A Powerful Tool to Explore Protein Folding on the Ribosome. *Methods* **2010**, *52*, 57–73.
- (89) Szmanski, H.; Jayaweera, R.; Cherek, H.; Lakowicz, J. R. Demonstration of an Associated Anisotropy Decay by Frequency-Domain Fluorometry. *Biophys. Chem.* **1987**, *27*, 233–241.
- (90) Hazlett, T. L.; Jameson, D. L. Time-Resolved Fluorescence Studies on Components of the Prokaryotic Protein Elongation System. *Proc. SPIE* **1988**, *909*, 412–419.
- (91) Bialik, C. N.; Wolf, B.; Rachofsky, E. L.; Ross, J. B. A.; Laws, W. R. Dynamics of Biomolecules: Assignment of Local Motions by Fluorescence Anisotropy Decay. *Biophys. J.* **1998**, *75*, 2564–2573.
- (92) James, N. G.; Ross, J. A.; Stefl, M.; Jameson, D. M. Applications of Phasor Plots to in Vitro Protein Studies. *Anal. Biochem.* **2011**, *410*, 70–76.
- (93) Ziehr, D. R.; Ellis, J. P.; Culviner, P. H.; Cavagnero, S. Production of Ribosome-Released Nascent Proteins with Optimal Physical Properties. *Anal. Chem.* **2010**, *82*, 4637–4643.
- (94) Sarkar, M.; Lu, J.; Pielak, G. Protein Crowder Charge and Protein Stability. *Biochemistry* **2014**, *53*, 1601–1606.
- (95) Smith, A.; Zhou, L.; Gorensek, A.; Senske, M.; Pielak, G. In-Cell Thermodynamics and a New Role for Protein Surfaces. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 1725–1730.
- (96) Smith, A.; Zhang, Z.; Pielak, G.; Li, C. Nmr Studies of Protein Folding and Binding in Cells and Cell-Like Environments. *Curr. Opin. Struct. Biol.* **2015**, *30*, 7–16.

- (97) Klumpp, S.; Bode, W.; Puri, P. Life in Crowded Conditions: Molecular Crowding and Beyond. *Eur. Phys. J.: Spec. Top.* **2019**, *227*, 2315–2328.
- (98) Blobel, G.; Sabatini, D. Controlled Proteolysis of Nascent Polypeptides in Rat Liver Cell Fractions: I. Location of the Polypeptides within Ribosomes. *J. Cell Biol.* **1970**, *45*, 130–145.
- (99) Kuriyan, J.; Wilz, S.; Karplus, M.; Petsko, G. A. X-Ray Structure and Refinement of Carbon-Monoxide (Fe II)-Myoglobin at 1.5 Å Resolution. *J. Mol. Biol.* **1986**, *192*, 133–154.
- (100) Robinson, A. C.; Majumdar, A.; Schlessman, J. L.; Garcia-Moreno, E. B. Charges in Hydrophobic Environments: A Strategy for Identifying Alternative States in Proteins. *Biochemistry* **2017**, *56*, 212–218.
- (101) Harms, M. J.; Castañeda, C. A.; Schlessman, J. L.; Sue, G. R.; Isom, D. G.; Cannon, B. R.; et al. The pK_a Values of Acidic and Basic Residues Buried at the Same Internal Location in a Protein Are Governed by Different Factors. *J. Mol. Biol.* **2009**, *389*, 34–47.
- (102) Jungbauer, L. M.; Bakke, C. K.; Cavagnero, S. Experimental and Computational Analysis of Translation Products in Apomyoglobin Expression. *J. Mol. Biol.* **2006**, *357*, 1121–1143.
- (103) Espenson, J. H. *Chemical Kinetics and Reaction Mechanisms*; 2nd ed.; McGraw-Hill: New York, 1995.
- (104) Spudich, G. M.; Miller, E. J.; Marqusee, S. Destabilization of the Escherichia Coli RNase H Kinetic Intermediate: Switching between a Two-State and Three-State Folding Mechanism. *J. Mol. Biol.* **2004**, *335*, 609–618.
- (105) Varela, A. E.; Lang, J. F.; Wu, Y.; Dalphin, M. D.; Stangl, A. J.; Okuno, Y.; Cavagnero, S. Kinetic Trapping of Folded Proteins Relative to Aggregates under Physiologically Relevant Conditions. *J. Phys. Chem. B* **2018**, *122*, 7682–7698.
- (106) Varela, A. E.; England, K. A.; Cavagnero, S. Kinetic Trapping in Protein Folding. *Protein Eng., Des. Sel.* **2019**, *32*, 103–108.
- (107) Foshag, D.; Henrich, E.; Hiller, E.; Schäfer, M.; Kerger, C.; Burger-Kentischer, A.; Diaz-Moreno, I.; García-Mauriño, S. M.; Dötsch, V.; Rupp, S.; et al. The E. Coli S30 Lysate Proteome: A Prototype for Cell-Free Protein Production. *New Biotechnol.* **2018**, *40*, 245–260.
- (108) Zolkiewski, M. ClpB Cooperates with DnaK, DnaJ, and GrpE in Suppressing Protein Aggregation: A Novel Multi-Chaperone System from Escherichia Coli. *J. Biol. Chem.* **1999**, *274*, 28083–28086.
- (109) Mogk, A.; Tomoyasu, T.; Goloubinoff, P.; Rüdiger, S.; Röder, D.; Langen, H.; Bukau, B. Identification of Thermolabile Escherichia Coli Proteins: Prevention and Reversion of Aggregation by DnaK and ClpB. *The EMBO journal* **1999**, *18*, 6934–6949.
- (110) Jennaro, T. S.; Beaty, M. R.; Kurt-Yilmaz, N.; Luskin, B. L.; Cavagnero, S. Burial of Nonpolar Surface Area and Thermodynamic Stabilization of Globins as a Function of Chain Elongation. *Proteins: Struct., Funct., Genet.* **2014**, *82*, 2318–2331.
- (111) Haass, C.; Selkoe, D. J. Soluble Protein Oligomers in Neurodegeneration: Lessons from the Alzheimer's Amyloid β -Peptide. *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 101–112.
- (112) Komar, A. A.; Kommer, A.; Krashennikov, I. A.; Spirin, A. S. Cotranslational Heme Binding to Nascent Globin Chains. *FEBS Lett.* **1993**, *326*, 261–263.
- (113) Komar, A. A.; Kommer, A.; Krashennikov, I. A.; Spirin, A. S. Cotranslational Folding of Globin. *J. Biol. Chem.* **1997**, *272*, 10646–10651.
- (114) Niwa, T.; Ying, B.-W.; Saito, K.; Jin, W.; Takada, S.; Ueda, T.; Taguchi, H. Bimodal Protein Solubility Distribution Revealed by an Aggregation Analysis of the Entire Ensemble of Escherichia Coli Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 4201–4206.
- (115) Fink, A. L. Protein Aggregation: Folding Aggregates, Inclusion Bodies and Amyloid. *Folding Des.* **1998**, *3*, R9–R23.
- (116) Tartaglia, G. G.; Pechmann, S.; Dobson, C. M.; Vendruscolo, M. Life on the Edge: A Link between Gene Expression Levels and Aggregation Rates of Human Proteins. *Trends Biochem. Sci.* **2007**, *32*, 204–206.
- (117) Drummond, D. A.; Wilke, C. O. The Evolutionary Consequences of Erroneous Protein Synthesis. *Nat. Rev. Genet.* **2009**, *10*, 715–724.
- (118) Lim, S. A.; Hart, K. M.; Harms, M. J.; Marqusee, S. Evolutionary Trend toward Kinetic Stability in the Folding Trajectory of RNases H. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 13045–13050.
- (119) Plaxco, K. W.; Larson, S.; Ruczinski, I.; Riddle, D. S.; Thayer, E. C.; Buchwitz, B.; Davidson, A. R.; Baker, D. Evolutionary Conservation in Protein Folding Kinetics. *J. Mol. Biol.* **2000**, *298*, 303–312.
- (120) Gottesman, S.; Wickner, S.; Maurizi, M. R. Protein Quality Control: Triage by Chaperones and Proteases. *Genes Dev.* **1997**, *11*, 815–823.