PNAS RESEARCH ARTICLE | CELL BIOLOGY
BIOPHYSICS AND COMPUTATIONAL BIOLOGY

Molecular drivers of cotranslational protein assembly in bacteria

Jaro L. Schmitt[a,1], Josef J. Auburger[a,1,2], Florian Wruck[b,1], Jiří Koubek[a,3], Sander J. Tans[b,c,4], Bernd Bukau[a,4], and Günter Kramer[a,4]

Affiliations are included on p. 9.

Edited by Patricia L. Clark, University of Notre Dame, Notre Dame, IN; received March 24, 2026; accepted August 26, 2026 by Editorial Board Member Yale E. Goldman

Protein complexes are pivotal to all major cellular processes. Translation-coupled complex assembly is ubiquitous in eukaryotes, but its molecular basis and prevalence in bacteria remain poorly understood. Here, we combine ribosome profiling and single-molecule optical tweezers to establish principles governing cotranslational assembly of four structurally diverse complexes in *Escherichia coli*. We reveal how heterotetramers form cotranslationally via hierarchical assembly pathways, with initial heterodimer formation priming the assembly of higher-order oligomers through a topology-dependent organization of homomeric and heteromeric interfaces. Co-post assembly, defined as interaction between nascent and fully synthesized subunits, initiates when dimerization domains are completely exposed at the ribosome. The assembly order is predominantly unidirectional and matches the operon gene order, with upstream-encoded subunits being synthesized first, and then engaging downstream nascent partner subunits. The co-post assembly efficiency depends both on the native operon structure and the ribosome-associated chaperone trigger factor, which binds emerging oligomerization domains and whose dissociation marks the onset of partner subunit engagement. The handover from chaperone engagement to cotranslational assembly is mediated by nascent chain folding. Finally, we show that co-post assembly provides a fitness advantage. Together, these findings reveal how genome organization and chaperone-assisted folding coordinate cotranslational assembly to promote efficient complex biogenesis and cellular fitness in bacteria.

cotranslational protein assembly | trigger factor | molecular chaperones | nascent chain folding | operon structure

Significance

Interactions between partner subunits of protein complexes often initiate cotranslationally, i.e., while at least one subunit is still being synthesized at the ribosome. In bacteria, such genes are often grouped into operons, controlled by the same promoter and transcribed into one polycistronic mRNA. We show that the bacterial chaperone trigger factor binds the dimerization domain of a nascent polypeptide until it is fully folded and enhances the efficiency of the subsequent handover to the full-length partner subunit. Moreover, we show that the native operon structure enhances the cotranslational assembly of two partner subunits, which accelerates cell growth under conditions where this complex is essential. This reveals how genome organization, chaperones, and complex assembly work together to enhance cellular fitness.

Most proteins function as part of multiprotein complexes. However, productive complex assembly competes with diverse deleterious off-pathway interactions (1), suggesting that precise coordination is required. The formation of protein complexes during ongoing subunit synthesis has emerged as a pathway that is ubiquitous across the proteome and conserved in evolution, and may optimize the assembly process (2, 3). Cotranslational assembly may occur either between two nascent chains (co-co assembly) or between one nascent chain and a fully synthesized partner subunit (co-post assembly).

Recent experiments have revealed key features of cotranslational assembly in eukaryotes. It is governed by structural features of dimerization interfaces (4–6), can prevent partner subunit aggregation (5, 7, 8), direct alternative assembly pathways (6, 9) and is associated with the colocalization of mRNAs (8, 10–13). In contrast, bacterial cotranslational assembly is poorly understood, and remains essentially restricted to studies on the model proteins beta-galactosidase and heterologously expressed bacterial luciferase (14–17). Moreover, bacterial protein assembly occurs in a cellular context that differs fundamentally from that of eukaryotes. For instance, as they emerge from the ribosome, most nascent chains interact with the chaperone trigger factor (TF) (18, 19), which binds a dedicated ribosomal site near the tunnel exit (20), and lacks a eukaryotic analog. TF protects nascent chains from aggregation (21–26), shields subsequently synthesized protein domains (24, 27), accelerates folding by enhancing their collapse (28), and stabilizes incomplete surfaces on folding intermediates (27, 29) across the proteome (19). Yet, whether TF promotes cotranslational assembly is unknown.

Unlike eukaryotes, prokaryotes typically organize functionally related genes within operons, transcribed to polycistronic mRNAs that may facilitate the spatiotemporal coordination of complex assembly. Comparative computational analyses suggest that operon gene order reflects the predicted order in which subunits assemble (30, 31). Operon organization was found to enhance the assembly efficiency of the heteromeric LuxAB luciferase, but this process relied on heterologous expression of artificial operons and did not directly address cotranslational pathways (17). These observations suggest that operon organization has evolved to coordinate the biogenesis of protein complexes. However, it remains unclear how operon architecture, complex topology, and the chaperone TF

Author contributions: J.L.S., J.J.A., F.W., S.J.T., B.B., and G.K. designed research; J.L.S., J.J.A., F.W., and J.K. performed research; J.L.S., J.J.A., and F.W. analyzed data; and J.L.S., S.J.T., B.B., and G.K. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission. P.L.C. is a guest editor invited by the Editorial Board.

Copyright © 2026 the Author(s). Published by PNAS. This article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

[1]J.L.S., J.J.A., and F.W. contributed equally to this work.

[2]Present address: Research & Development, Crop Science, Bayer AG, Monheim am Rhein 40789, Germany.

[3]Present address: Laboratory of RNA Biology of Protists, Institute of Parasitology, Biology Centre, Czech Academy of Sciences, České Budějovice 370 05, Czech Republic.

[4]To whom correspondence may be addressed. Email: tans@amolf.nl, bukau@zmbh.uni-heidelberg.de, or g.kramer@zmbh.uni-heidelberg.de.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2608997123/-/DCSupplemental.

Published October 22, 2026.

Downloaded from https://www.pnas.org by AMOLF LIBRARY SERVICES on November 18, 2026 from IP address 194.171.111.64.

influence the cotranslational assembly of native complexes, and how this influences cellular fitness.

Here, we combined ribosome profiling with single-molecule assays to dissect cotranslational assembly mechanisms in *Escherichia coli* focusing on four operon-encoded protein complexes with distinct architectures. Three complexes (CarAB, PheST, and TrpED) assembled via co-post assembly, revealing three general principles: i) Co-post assembly initiates upon ribosome-exposure of the complete dimerization domain. ii) Co-post assembly follows a unidirectional order, with the fully synthesized subunit encoded by the upstream open reading frame (ORF) engaging their nascent partner subunit encoded by the downstream ORF. iii) Co-post assembly favors heteromer formation, while the co-co assembly pathway favors homomeric interactions, yet both pathways can converge during the cotranslational assembly of a dimer of heterodimers. Examining the connection of cotranslational folding, chaperone binding, and assembly, we show that TF binds emerging oligomerization domains and co-post assembly frequencies are reduced in cells lacking TF. Single molecule measurements detected progressively decreasing TF binding as folding of the nascent subunit advanced, while efficient co-post assembly required partial folding of the nascent chain. The folding state of the nascent polypeptide thus directs a handover from TF engagement to cotranslational dimer assembly. Finally, we show that operon-driven colocalized synthesis from a bicistronic mRNA enhances assembly efficiency and cell growth compared to expression from separate mRNAs, establishing a direct link between cotranslational protein complex assembly and increased evolutionary fitness. Together, these findings establish key mechanisms and a functional relevance of cotranslational assembly in *E. coli*, highlighting how operon architecture, folding dynamics, and chaperone interactions collectively shape protein complex biogenesis.

Results

A Dynamic Network of Cotranslational Interactions Guides Complex Assembly. To analyze molecular principles of cotranslational assembly in bacteria, we investigated four complexes that differ in size, topological complexity, interface distribution, physical entanglement, and the set of hetero- and homomeric interfaces. Hence, these complexes offer a framework for probing how architectural features shape cotranslational assembly pathways. In all the complexes, operon-encoded heterodimers form the basic structural unit. To study co-post assembly of these heterodimers by selective ribosome profiling (SeRP, Fig. 1*A*), we C-terminally tagged each subunit with monomeric YFP (mEYFP) in a scar-free manner at its native chromosomal locus. We could thus test whether and when a fully synthesized YFP-tagged subunit engaged its nascent dimerization partner, by sequencing the mRNA footprints of co-purified ribosomes.

We first investigated the co-post assembly propensity of the carbamoyl-phosphate synthetase $(CarAB)_4$, a hetero-octamer with marked structural asymmetry. CarA is a small and topologically simple glutamine amidotransferase (GATase). CarB is among the largest proteins in *E. coli*, has a complex topology and contributes with multiple domains to the large, interlocked heterodimer interface (Fig. 1*B* and *SI Appendix*, Fig. S1 *A* and *B*). We detected binding of full-length CarA, encoded upstream of *carB* within the operon, to nascent CarB but not to nascent CarA, whereas full-length CarB did not bind nascent CarA or CarB (Fig. 1*C*). The lack of CarA and CarB homodimerization is consistent with the octamer structure, as CarA subunits do not interact with each other, while the two CarB subunits have a small contact site (*SI Appendix*, Fig. S1 *A* and *B*), leading to frequent CarAB and $(CarAB)_4$ interconversion (32). CarA bound nascent CarB after full ribosome exposure of the heterodimerization interface (Asn227–Asn554), which spans domains C and D. We detected a delay between heterodimerization interface emergence and CarA-CarB assembly onset at about codon 680. This delay allows synthesis of the subsequent domain I (Asn554–Lys686), which structurally supports the two domains forming the heterodimerization interface (Fig. 1 *B* and *C*). This suggests that productive co-post assembly of the CarAB heterodimer requires both cotranslational domain maturation and correct interdomain docking within the nascent CarB polypeptide. Subsequently, the cotranslationally formed CarAB heterodimers may tetramerize posttranslationally to form the $(CarAB)_4$ octamer.

The phenylalanyl-tRNA synthetase $(PheST)_2$ has a more interdependent architecture with structurally intertwined subunits (Fig. 1*D*). We found that the upstream-encoded full-length PheS bound nascent PheT, while neither full-length PheS nor PheT bound nascent PheS (Fig. 1*E*). The structure suggested two factors contributing to this asymmetry: i) PheS-PheT dimer interactions are distributed throughout the entire upstream encoded subunit (PheS), while they are located only centrally in the downstream-encoded subunit (PheT, *SI Appendix*, Fig. S1*C*), and ii) in the intertwined PheST dimer, PheT adopts a more extended conformation, while PheS is more compact. Hence, PheT can fold onto the surface of PheS rather than the other way around. In contrast to CarAB assembly, we also detected full-length PheT associated with nascent PheT (Fig. 1*E*). This finding is notable, as the PheT-PheT interactions are limited and confined to the C terminus (*SI Appendix*, Fig. S1*C*). Moreover, both PheS-PheT and PheT-PheT binding occurred precisely at the same time, namely upon exposure of the PheT domains B5 and B6/7 (33). An important structural clue here is that one PheS protomer contacts not just one PheT protomer but both of them (*SI Appendix*, Fig. S1*C*) (33), indicating a central role in PheT folding. Overall, these data suggest that binding of PheS to nascent PheT primes the latter for the association of the second PheST dimer that is already fully formed, resulting in the cooperative cotranslational assembly of the complete $(PheST)_2$ tetramer (Fig. 1 *E*, *Top*).

Within the anthranilate synthase $(TrpED)_2$, TrpE forms extensive heteromeric contacts across the entire GATase domain at the N terminus of TrpD (Fig. 1*F* and *SI Appendix*, Fig. S1*D*). In *E. coli trpD*, the GATase domain is genetically fused to an anthranilate phosphoribosyltransferase (AnPRT), which in isolation often adopts a homodimeric architecture (35). This modular organization creates a combined set of heteromeric and homodimeric interface elements that differs from simple heteromeric GATase complexes and standalone AnPRT dimers, as well as from $(CarAB)_4$ and $(PheST)_2$. Consequently, assembly of $(TrpED)_2$ involves a complex network of cotranslational interactions. First, we found that TrpE efficiently bound nascent TrpD, yet this interaction was not strictly unidirectional (Fig. 1*G*). Indeed, we detected low but persistent binding of full-length TrpD to nascent TrpE after the translation of codon 200 of *trpE*, indicating nascent TrpE is engaged with low frequency or low affinity. Consistently, only a limited number of TrpD-contacting residues are exposed at this stage of translation (*SI Appendix*, Fig. S1*D*), as the main contacts lie in the chorismate binding domain of TrpE (Asp242–Phe520). Second, we observed concomitant binding of two subunits, full-length TrpE and TrpD, to nascent TrpD at the same translation stage (Fig. 1*G*). Binding coincided with complete emergence of the GATase domain, which harbors the heteromeric interface but no homomeric contact sites, suggesting that full-length TrpD binds nascent TrpD indirectly via TrpE. Third,

Downloaded from https://www.pnas.org by AMOLF LIBRARY SERVICES on November 18, 2026 from IP address 194.171.111.64.

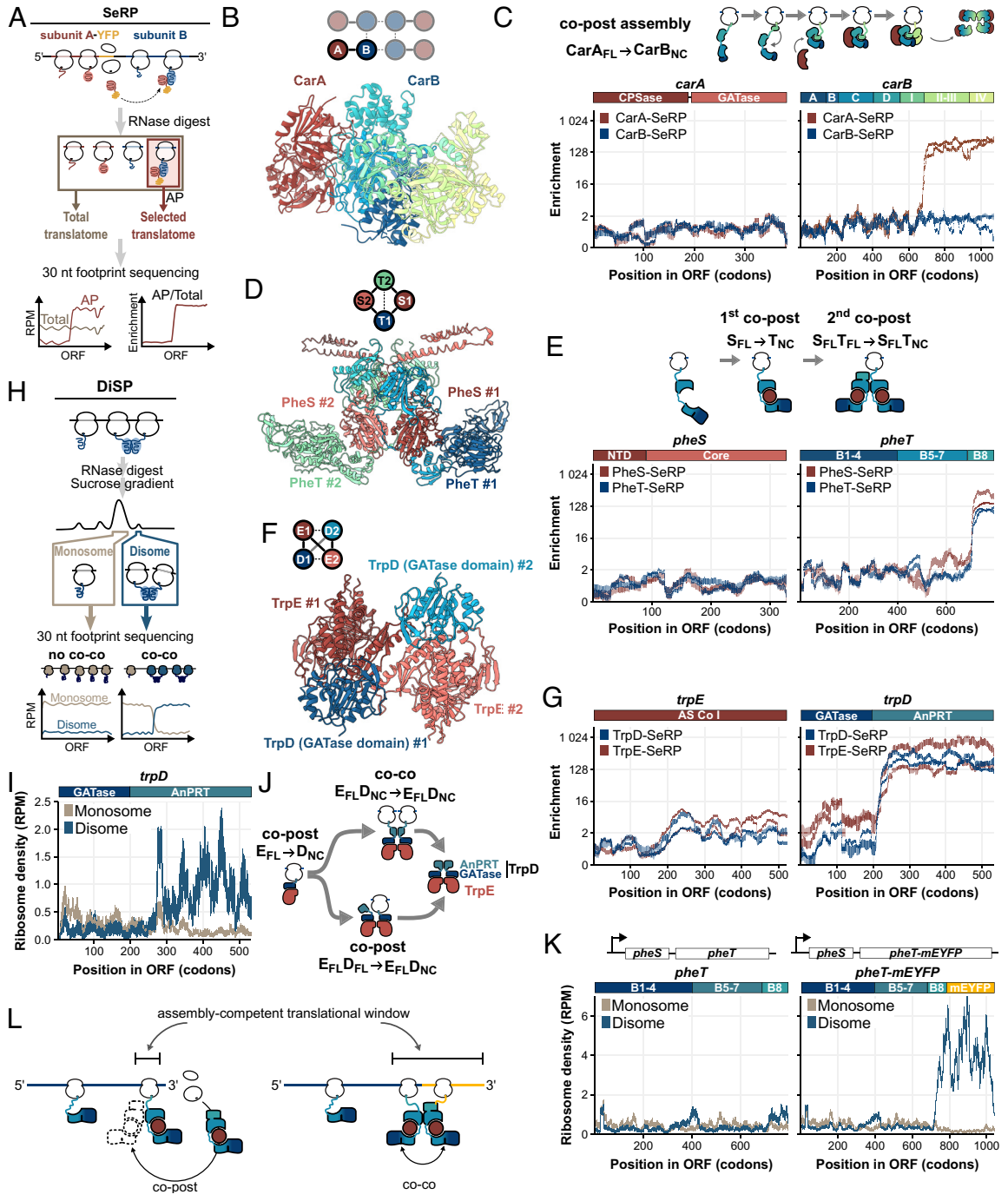


Fig. 1. Cotranslational assembly pathways revealed by ribosome profiling. (*A*) Schematic of SeRP. Ribosomes from an RNase-treated cell lysate are co-purified with the factor of interest, e.g., a C-terminally mEYFP-tagged protein complex subunit. Ribosome protected footprints of these co-purified ribosomes are sequenced (selected translatome) to reveal the position of ribosomes on a given mRNA engaged by the factor of interest. To account for the naturally occurring nonuniform distribution of ribosomes across a given ORF, the selected translatome sample is normalized at every codon position to the total translatome. (*B*) The CarAB heterodimer of the carbamoyl-phosphate synthetase octamer [PDB: 1kee (32)]. CarB domains are color coded as in *C*. (*C*) *Top*: Schematic model of co-post assembly between CarA and CarB. FL, full-length. NC, nascent chain. *Bottom*: SeRP data plotted as the codon-resolved respective enrichment of CarA or CarB-bound RNCs (selected translatome, see *A*) over the total translatome across the ORFs of *carA* and *carB*. Enrichment plot y-axis in log2 scale. Color-coded rectangles indicate the sequence position of domains. CPSase, carbamoyl-phosphate synthetase. GATase, glutamine amidotransferase. Distinct graphs show biological duplicates. Width of the graph at a specific codon indicates the 95% CI. (*D*) Phenylalanyl-tRNA synthetase $(PheST)_2$ [PDB: 3pco (33)]. PheT chain #1 domains are color coded as in *E*. (*E*) *Top*: Schematic model of two-step co-post assembly of the $(PheST)_2$ complex. FL, full-length. NC, nascent chain. *Bottom*: SeRP data for PheS and PheT as in *C*. NTD, N-terminal domain. (*F*) Anthranilate synthase complex [PDB: 1i1q (34)] from *Salmonella typhimurium* (*E. coli* TrpD sequence identity 97%). Note that the AnPRT domain of TrpD is not resolved in this structure. (*G*) SeRP data as in *E* for the TrpED complex. AS CoI, anthranilate synthase component I. GATase, glutamine amidotransferase. AnPRT, anthranilate phosphoribosyltransferase. Distinct graphs show biological duplicates. (*H*) Schematic of disome selective profiling (DiSP). Monosomes and disomes of an RNase-treated cell lysate are separated by sucrose gradient density ultracentrifugation. ~30 nucleotide-long ribosome-protected footprints from both fractions are sequenced separately and mapped onto ORFs. Nascent chain-mediated disome formation, the hallmark of co-co assembly, is detected by a shift of reads from the monosome to the disome sample. (*I*) DiSP data showing the distribution of 30 nt footprints from monosomes and disomes across the *trpD* ORF in wild-type cells (GSE291748). Domains indicated as in *G*. Width of the graph at a specific codon indicates the 95% CI. (*J*) Schematic model of the $(TrpED)_2$ cotranslational assembly network involving co-post and co-co assembly. FL, full-length. NC, nascent chain. (*K*) DiSP data showing the distribution of 30 nucleotide footprints from monosomes and disomes across the wild-type *pheT* ORF (*Left*) or the mEYFP-tagged *pheT* ORF (*Right*). (*L*) Gene fusions can extend the duration of cotranslational interface exposure and therefore bias the assembly pathway from post to co-post and from co-post to co-co assembly.

Downloaded from https://www.pnas.org by AMOLF LIBRARY SERVICES on November 18, 2026 from IP address 194.171.111.64.

the interconnected architecture of $(TrpED)_2$ raised questions concerning the potential contribution of the alternative co-co assembly pathway, involving two nascent chains. Disome selective profiling (DiSP) (6) was recently developed to detect co-co assembly in human cells by sequencing ~30 nt ribosomal footprints of monosomes and nascent chain-linked disomes isolated via sucrose gradient ultracentrifugation (Fig. 1*H*). Analysis of DiSP data from wild-type *E. coli* grown in medium lacking tryptophan to induce expression of the native *trp* operon revealed co-co assembly after partial exposure of TrpD's C-terminal AnPRT domain (Fig. 1*I*). More specifically, co-co assembly initiated immediately after exposure of helix 3 (residues 232 to 245), which forms the interaction hotspot in TrpD homodimers (35). Two residues within this helix were previously identified as key determinants of TrpD quaternary structure. Sequence comparison with characterized monomeric and dimeric TrpD homologs showed that the corresponding residues in *E. coli* TrpD, E233 and M242, match the consensus associated with homodimer formation (*SI Appendix*, Fig. S2). These results indicate a dynamic assembly network involving co-post as well as co-co assembly modes (Fig. 1*J*). The assembly pathway can be initiated by full-length TrpE binding nascent TrpD. In one assembly route, such a heterodimer may be joined by another preformed heterodimer thus completing $(TrpED)_2$ on one nascent TrpD polypeptide. Alternatively, two heterodimers, each formed by co-post assembly of TrpE on a nascent TrpD, might join and establish the tetrameric complex between two nascent TrpD polypeptides.

The requirement for complete emergence of the TrpD GATase domain from the ribosome before TrpE binding suggested an interesting functional consequence of its fusion to the AnPRT domain: the latter extends the ORF and hence prolongs ribosome association of the GATase domain, which may promote cotranslational assembly. We tested this idea using PheT, asking whether ORF extension with a structurally unrelated sequence could artificially impose such a bias. PheT forms homomeric contacts via its C terminus (Fig. 1*D* and *SI Appendix*, Fig. S1*C*) and normally does not show pronounced co-co assembly (Fig. 1*K*). We therefore analyzed co-co assembly of PheT carrying a C-terminal mEYFP-tag. DiSP revealed a pronounced increase in its co-co assembly propensity (Fig. 1 *K* and *L*). This demonstrates that assembly routes are not dictated solely by intrinsic structural features but can be redirected by mere ORF extension through gene fusion. In contrast, however, extending the PheS ORF did not markedly increase its co-co assembly (*SI Appendix*, Fig. S1*E*). This is consistent with the lack of homomeric PheS-PheS contacts and a general preference of the co-co assembly pathway for homomer formation.

Among the four complexes analyzed, the succinyl-CoA synthetase $(SucCD)_2$ represents the structurally simplest case, with two compact subunits that associate through relatively small and planar interfaces (*SI Appendix*, Fig. S1 *F* and *G*). In contrast to the other complexes, the dimerization interface of the succinyl-CoA synthetase subunits SucC and SucD is scattered across their N- and C-terminal domains. Consistent with the lack of continued translation after full exposure of the dimerization interface, we did not detect cotranslational assembly of the Succinyl-CoA synthetase complex (*SI Appendix*, Fig. S1*H*).

TF Promotes Co-post Assembly Efficiency. Whether TF plays a role in co-post assembly, for instance by delaying, accelerating, or enhancing subunit engagement, is an open question. To address this question, we first analyzed co-post assembly by subunit-specific SeRP in a Δ*tig* strain. The lack of TF caused decreased binding of the full-length subunits CarA, PheS, and TrpE to their respective nascent partners CarB, PheT, and TrpD by 3- to 10-fold (Fig. 2*A*), indicating that TF enhances the efficiency of co-post assembly. We furthermore observed decreased association of full-length PheT and TrpD with nascent PheT and TrpD, respectively (*SI Appendix*, Fig. S3). Notably, the onset of assembly during translation was unchanged for all subunits in all three complexes, showing that TF can promote co-post assembly without affecting its timing.

Next, we quantified cotranslational TF binding to the nascent subunits involved by TF-specific SeRP (TF-SeRP). TF robustly bound RNCs of all three protein complexes (Fig. 2*B*), engaging the heterodimerization interfaces during their synthesis, namely the CarB domains C-D, the PheT domain B7, and the TrpD

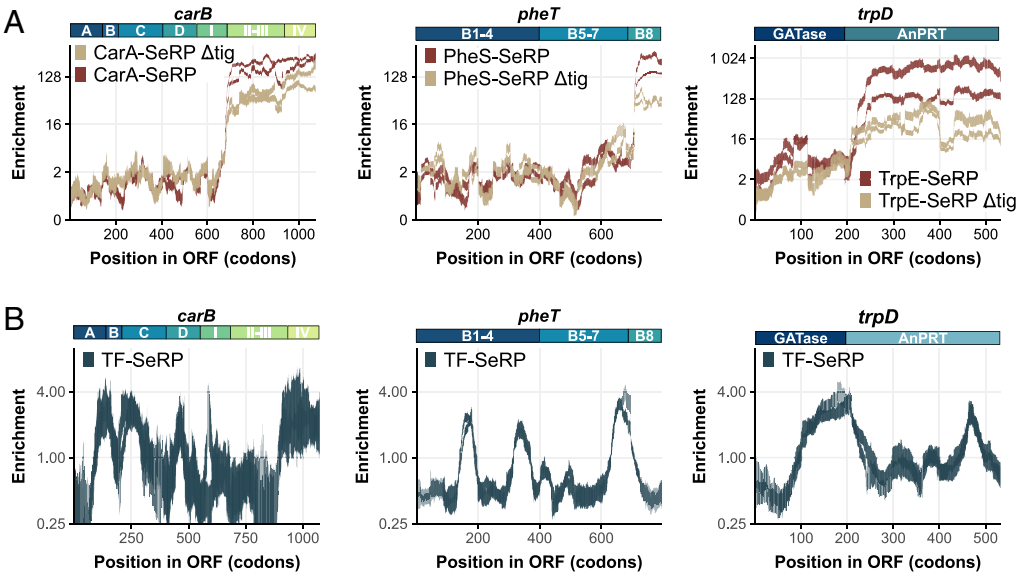


Fig. 2. TF enhances co-post assembly efficiency. (*A*) Subunit-specific SeRP data showing the effect of presence (MC4100 wild-type cells) or absence (MC4100 Δ*tig* mutant cells) of TF on the interactions of CarA with nascent CarB, PheS with nascent PheT, and TrpE with nascent TrpD. Depicted are the codon-resolved enrichment of subunit-bound RNCs over the total translatome. Enrichment plot y-axis in log2 scale. Color-coded rectangles indicate the sequence position of domains. Distinct graphs show biological duplicates. Width of the graph at a specific position indicates the 95% CI. (*B*) SeRP data showing the codon-resolved respective enrichment of TF-bound RNCs over the total translatome. Distinct graphs show biological replicates.

Downloaded from https://www.pnas.org by AMOLF LIBRARY SERVICES on November 18, 2026 from IP address 194.171.111.64.

GATase domain. We observed a direct transition from TF binding to co-post assembly for TrpD and PheT. TF binding to TrpD started when about 50 residues of the GATase domain had emerged, persisted throughout domain synthesis and dissociated immediately after its complete exposure (Fig. 2 *B*, *Right*). TF dissociation coincided precisely with the onset of TrpE engagement (Fig. 2 *A*, *Right*).

Within the highly intertwined heterodimerization interface of PheST, the B5 and B6/7 modules of PheT form a clamp-like structure, interconnected by a flexible interdomain linker. This clamp is structurally stabilized upon insertion of the globular PheS core (Fig. 1*D*). Notably, TF avoided the unstructured segment but engaged the structured part of domain B6/7 (Fig. 2 *B*, *Middle*, codons 400 to 600), consistent with TF stably binding partially folded domains (19). In the case of CarB, TF does not show prolonged binding just prior to CarA engagement, while TF binding peaks are observed as the C and D dimerization domains emerge (Fig. 2 *A* and *B*). The data are consistent with CarA binding being delayed by dependence on emergence of the domain I, which structurally supports the C and D domains, rather than by competitive TF binding. Together, these findings indicate that TF increases the proportion of nascent chains that achieve assembly competence cotranslationally, presumably by accelerating folding.

Nascent Chain Folding Coordinates TF Binding and Co-post Assembly. The observed timing of co-post assembly during translation suggests that it relies on the folding of the nascent oligomerization domain. To test this directly, we developed an in vitro assay for studying nascent chain folding and subunit binding. Optical tweezers were used to monitor folding of the single GATase domain of TrpD, while simultaneous two-color fluorescence microscopy detected TF and TrpE binding. Specifically, we synthesized SecM-stalled TrpD chains on ribosomes linked to polystyrene beads via a DNA handle. An N-terminal biotin was incorporated by amber stop codon suppression, enabling attachment to a second bead again using DNA handles (*SI Appendix*, Fig. S4*A*). This tethering allowed us to hold the stalled ribosome and nascent chain in the imaging plane for fluorescent detection of TrpE and TF binding, and to study nascent TrpD unfolding and refolding by moving the two beads with the optical traps. We chose translation stages immediately before, during, and after the onset of TrpE binding, encompassing completion of the GATase domain (Fig. 3*A*). TrpE and TF were labeled with different fluorophores, and added at 100 nM, which is the maximum concentration possible given the background fluorescence.

Single TrpE and TF molecules were observed binding and dissociating from the tethered TrpD-RNCs, as indicated by stable fluorescent spots appearing between the beads (Fig. 3*B*). These spots repeatedly disappeared and reappeared, reflecting multiple binding cycles (Fig. 3*B* and *SI Appendix*, Fig. S4*B*). The fraction of TrpD RNCs engaged by TrpE strongly depended on nascent chain length, increasing from negligible levels for the shortest fragment to ~50% for the longest (Fig. 3*C*). TrpE residence time also increased with nascent chain length, from a median of 7 s for the shortest fragment to 15 s for the longest, although the differences did not reach statistical significance (Fig. 3 *D*, *Right*). In contrast, TF engaged all four TrpD fragments at similar frequencies (Fig. 3*C*). Residence times, while broadly distributed, were significantly shorter for the longest TrpD fragment (Fig. 3 *D*, *Left*), consistent with the reduced SeRP enrichment at this chain length. These data agree with the handover from TF to TrpE during translation, while providing underlying kinetics, and more generally showed that no other cellular factors are required to enable co-post assembly.

Next, we examined how TrpE binding depended on the degree of compaction of nascent TrpD due to folding. Hence, we quantified the number of TrpD residues that were stably compacted (with the rest of the TrpD residues then being unfolded), using the measured forces and bead-to-bead distances (extensions) during stretching at low forces. Here, we focused on the first stretching experiment immediately after nascent chain synthesis and cotranslational folding in presence of TrpE and TF. The resulting aggregate histogram for all fragments showed a narrow peak corresponding to unfolded chains, and a broad peak corresponding to compacted states (Fig. 3 *E*, *Top*). When considering fragment differences, we found the unfolded state less frequently (and hence a folded state more frequently) when the fragment was longer (Fig. 3 *E*, *Top*, and *SI Appendix*, Fig. S4*C*). This observation is consistent, as smaller fragments lack residues that may be central to folding.

Fluorescence detection revealed that TrpE exclusively bound compacted nascent chains, and no binding was observed for unfolded chains (Fig. 3*E*). These results demonstrate that TrpD must be folded to stably bind TrpE. The most compacted TrpD chains beyond about 180 compacted residues showed limited TrpE binding, which could mean that continued translation at the end of the GATase domain and beyond results in the formation of closed structures that are not assembly competent. TF bound all TrpD conformational states, from fully unfolded to fully compacted (Fig. 3*E*). However, normalizing for the observed population frequency of each state indicated a preference for states of intermediate level of compaction (*SI Appendix*, Fig. S4 *D* and *E*). Given that TrpE only bound intermediately compacted states, these data support a TF to TrpE handover model mediated by TrpD structure formation.

Operon Structure Enhances Co-post Assembly Efficiency and Cellular Fitness. A key question is whether cellular fitness depends on the spatial organization of protein synthesis, as achieved here by subunits encoded within operons. We addressed this question using the CarAB complex as a model, because i) it follows a strict unidirectional co-post assembly pathway, ii) contains a large interlocked heteromeric interface that may rely on cotranslational assembly, and iii) is conditionally essential for growth. We deleted the *carAB* operon and reintegrated both genes chromosomally, either separately at distant loci (split strain) or together in an operon (operon strain, Fig. 4*A*). Transcription was driven by the IPTG-inducible Ptac promoter. To ensure consistent expression, we adjusted the strength of ribosome binding sites in all strains. This empirical optimization is critical because biased expression levels would distort the apparent assembly efficiencies measured by SeRP.

We first tested whether splitting the operon affected cotranslational assembly efficiency by SeRP upon induction with 10 μM IPTG. Indeed, the co-post assembly frequency was reduced approximately fourfold in the split strain (Fig. 4*B*). To examine physiological consequences, we monitored growth in minimal medium, where carbamoyl-phosphate synthetase activity is essential for pyrimidine and arginine biosynthesis (36) (*SI Appendix*, Fig. S5*A*). Demonstrating that growth is limited by CarAB synthesis rates, the generation time decreased with increasing IPTG concentrations in both strains (Fig. 4*C* and *SI Appendix*, Fig. S5*B*). Importantly, the split strain required four to six times higher IPTG concentrations than the operon strain to reach comparable growth rates. At 10 μM IPTG, where we detected a fourfold reduced co-post assembly frequency in the split strain, the operon strain displayed a median doubling time of 82 min compared to 183 min for the split strain. Because *carAB* expression in the operon strain

Downloaded from https://www.pnas.org by AMOLF LIBRARY SERVICES on November 18, 2026 from IP address 194.171.111.64.

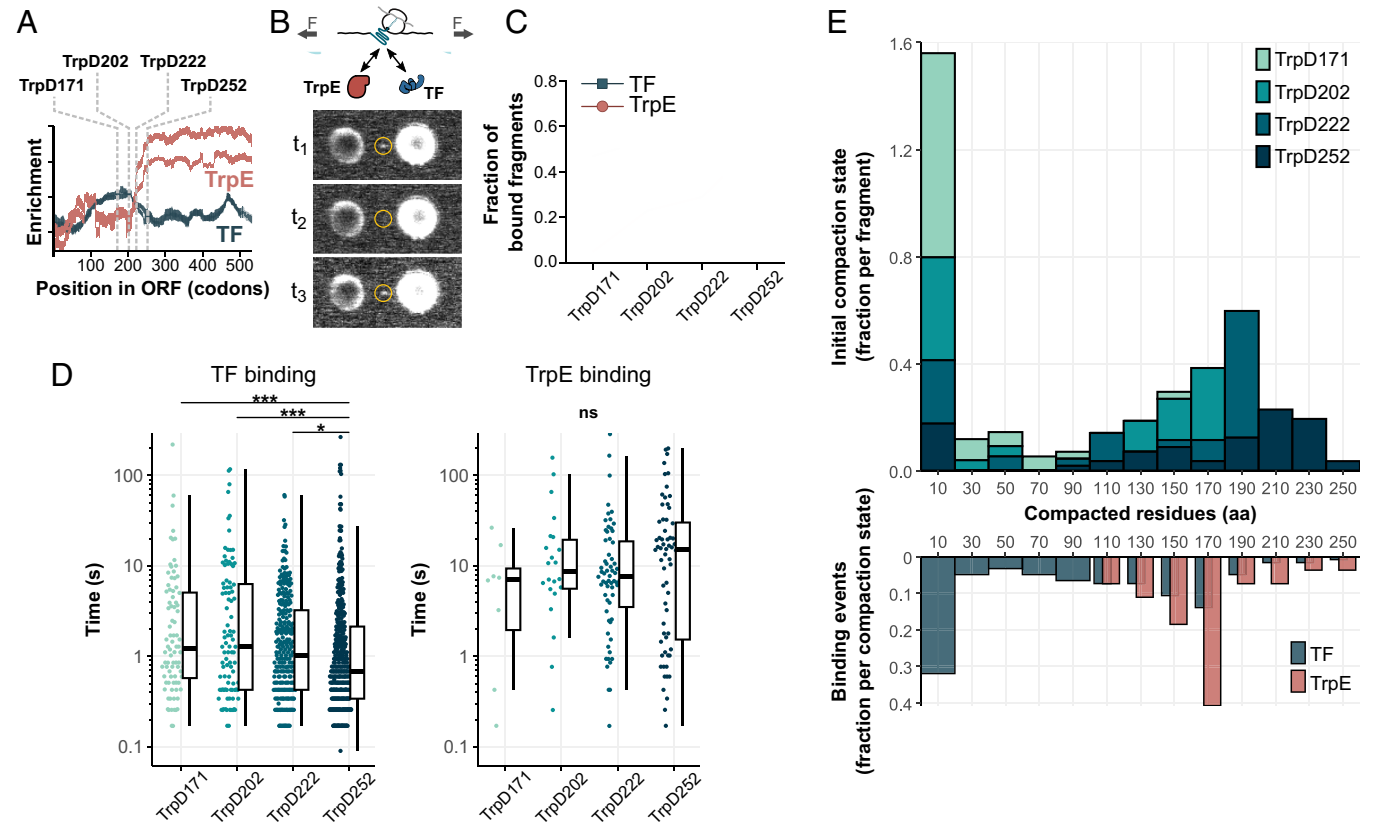


Fig. 3. Single molecule analysis links nascent chain folding to TF binding and co-post assembly. (*A*) Translation stages mimicked by SecMstr-stalled TrpD-RNCs. TrpD-RNCs contain 152, 183, 203, or 233 N-terminal TrpD residues, followed by a 19 amino acid-long SecMstr sequence, thus representing translation at codon 171, 202, 222, or 252. Traces illustrate TrpE (red) or TF (blue) binding to nascent TrpD determined by SeRP. (*B*) Experimental setup for *C*–*E*: SecMstr-stalled TrpD-RNCs are tethered via the nascent polypeptide's N terminus and the ribosome to two polystyrene beads, each held in an optical trap. Fluorescence microscopy records binding periods of fluorescently labeled TF (100 nM) and TrpE (100 nM) to TrpD-RNCs before unfolding. Multiple binding events of TF and TrpE may be measured per TrpD molecule before mechanical unfolding. t_1-t_3 depict representative snapshots of the optical trap with the fluorescent signal encircled in between two beads, visible due to their autofluorescence. Subsequently, information on the compaction state of the nascent TrpD polypeptide is obtained by measuring the force-extension profile upon gradually moving the optical traps further apart. (*C*) Binding preferences of TF or TrpE to specific TrpD-RNCs, prepared as described in B. Plotted values are the fraction of TrpD molecules of a given fragment length with detected factor engagement (28 TrpD molecules with detected binding of TrpE and 122 with detected binding of TF). (*D*) Recorded residence time of fluorescently labeled TF and TrpE on optically trapped TrpD-RNCs, prepared as described in *B*. Boxes indicate the interquartile range; center lines indicate the median; whiskers extend to the most extreme data point within 1.5× the interquartile range. Individual data points are overlaid. Kruskal–Wallis test: TF, P = 7.433e-05; TrpE, P = 0.505. Significant results of post hoc pairwise comparisons (two-sided Wilcoxon rank-sum, Holm correction) are shown for TF. *P < 0.05, ***P < 0.001. (*E*) TF and TrpE binding to single molecules of optically trapped TrpD-RNCs was recorded by fluorescence microscopy as described in panel *B*. Subsequently, the compaction state of the nascent TrpD polypeptide was measured by force-spectroscopy. *Top*: Stacked histograms showing the prevalence of compaction states among the different nascent TrpD polypeptide lengths, in the first stretching experiment for each molecule. Bin size = 20 amino acids (aa). Number of molecules analyzed: TrpD171 = 38, TrpD202 = 26, TrpD222 = 38, TrpD252 = 57. *Bottom*: fraction of total binding events per compaction state (28 TrpD molecules with detected binding of TrpE and 122 with detected binding of TF).

was slightly elevated (~1.3-fold) at 10 μM IPTG, we quantified expression at 2.5 μM IPTG. Despite expressing *carAB* at approximately half the level (*SI Appendix*, Table S1), the operon strain still grew significantly faster at 2.5 μM IPTG than the split strain at 10 μM IPTG (median 132 and 183 min, Fig. 4*C*). Thus, operonic gene arrangements can permit superior growth rates at lower metabolic costs. We also observed extended lag phases in the split strain, consistent with delayed accumulation of the functional CarAB complex (*SI Appendix*, Fig. S5*C*). Supporting this interpretation, modest induction of *carAB* expression already in the starter culture, prior to the shift to the CarAB-dependent growth medium, markedly reduced lag phases (*SI Appendix*, Fig. S5*D*).

Discussion

This study identifies principles that govern cotranslational assembly of bacterial protein complexes. We observe a pronounced bias in the directionality of co-post assembly, with the first subunit within an operon engaging the nascent polypeptide of the subunit encoded downstream. This finding is consistent with the general tendency of complexes to assemble via ordered pathways (37), which are imprinted in operonic gene arrangements (30, 31). We also observed a notable exception to this common rule, with the downstream TrpD assembling in fully synthesized form with the upstream nascent TrpE. We note that for the strictly directional CarA-CarB and PheS-PheT engagements, the upstream encoded protein is substantially smaller than the downstream encoded protein, in contrast to TrpE-TrpD. Such asymmetry may promote directionality because the downstream protein remains ribosome-tethered for longer. This operon-driven spatial coupling facilitates efficient cotranslational assembly, which provides a physiological advantage, as demonstrated by the fitness cost of disrupting the native polycistronic arrangement.

Our analysis further reveals structural constraints that determine the onset of co-post assembly. Across all three complexes undergoing co-post assembly, independent of whether they exhibit an intertwined architecture, stable partner binding occurs after the complete oligomerization interface of the nascent subunit has fully emerged from the ribosome. For CarB, CarA engagement

Downloaded from https://www.pnas.org by AMOLF LIBRARY SERVICES on November 18, 2026 from IP address 194.171.111.64.

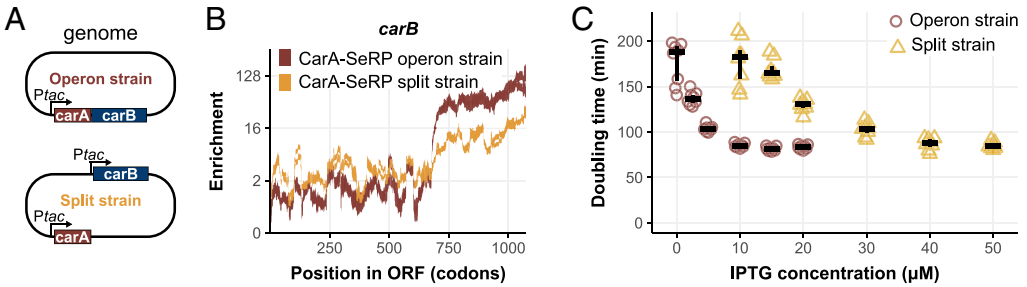


Fig. 4. Effect of operon structure on co-post assembly and cellular growth. (*A*) Genomic organization of *carA* and *carB* genes in the operon strain (MC4100 *ΔcarAB λ::carA-mEYFP_carB*) and the split strain (MC4100 *ΔcarAB λ::carA-mEYFP ϕ80::carB*). (*B*) CarA-SeRP in the operon strain (translation of bicistronic *carAB* mRNA) and split strain (translation of two monocistronic mRNAs). Depicted are the codon-resolved enrichment of subunit-bound RNCs over the total translatome. Y-axis in log2 scale. For each strain, data of biological duplicates are depicted as separate graphs. Width of the graph at a specific codon indicates the 95% CI. (*C*) Minimal doubling time of operon and split strains in dependence of *carAB* expression levels. Cells were grown overnight in M9 minimal medium supplemented with 0.4 mM uracil and 0.5 mM L-arginine and diluted into M9 minimal medium supplemented with the indicated IPTG concentrations to titrate expression of both genes. IPTG concentrations below 10 μM did not permit sufficient growth of the split strain within 24 h and were therefore omitted. Horizontal bars indicate the median and vertical lines the interquartile range (n = 8). Minimal doubling times were derived from the maximum specific growth rate of a reparameterized generalized logistic growth model fitted to the data.

requires exposure of the heterodimerization domains C and D together with subsequent synthesis of domain I that stabilizes their interdomain arrangement. For PheT, both PheS and PheT binding coincides with emergence of the B5 and B6/7 modules that form the central interaction surface of the heterodimer. For TrpD, stable TrpE binding begins after complete exposure of the GATase domain. These observations indicate that co-post assembly requires ribosomal exposure and productive folding of a structurally complete oligomerization interface.

Our data indicate notable higher-order cotranslational assembly pathways that integrate multiple co-post and co-co assembly processes. Indeed, a single nascent chain is often not engaged by just one subunit, but jointly by different subunits, and at the same moment of translation. The $(PheST)_2$ and $(TrpED)_2$ complexes that show this behavior also exhibit an interconnected topology, which include not only the elementary heteromeric contacts between up- and downstream subunits, but also homomeric contacts. In contrast, the $(CarAB)_4$ complex, which does not show such higher-order pathways, exhibits only limited homomeric contacts, many of which are also encoded C-terminally. Within this cotranslational assembly network, heteromeric interactions are preferably established via co-post assembly, while co-co assembly is biased toward homomerization and can be promoted by ORF extension. The latter was exemplified by fusing YFP sequences to the 3′ end of the *pheS* and *pheT* gene, respectively, triggering homomeric co-co assembly of PheT-YFP, while heteromeric co-co assembly with PheS-YFP remained limited. The lack of co-co interactions between RNCs translating neighboring cistrons suggests that, while intrinsic protein features determine the overall propensity for cotranslational assembly (5), the choice between co-co and co-post pathways is constrained by the spatial organization of ribosomes within polysomes. By extending polysome association, gene fusions such as the combination of TrpD GATase (ancestral *trpG*) and AnPRT (ancestral *trpD*) sequences (38), may confer advantages of cotranslational assembly pathways, including increased assembly efficiency and limited cytosolic diffusion and protein–protein aggregation. Hence, our results indicate how cotranslational assembly can control the formation of large higher-order protein complexes and might shape the evolution of protein and operon architectures.

Our experimental data partially aligned with cotranslational assembly probabilities predicted from the structural features of protein complexes (5). For TrpED, no prediction was available for comparison. For CarAB, the predicted unidirectional co-post assembly was consistent with our observations. Our PheST data, however, showed differences with the predicted assembly pathway. Both subunits were predicted to act as the nascent partner during interface formation, consistent with either heteromeric co-co assembly or bidirectional co-post assembly. Experimentally, however, nascent PheS showed neither substantial co-post nor co-co assembly. Even upon ORF extension, co-co assembly of PheS remained limited, and the markedly increased DiSP signal of the PheT-mEYFP fusion reflected homomeric co-co assembly between two PheT subunits rather than heteromeric assembly with PheS. This discrepancy between structure-based predictions and experimental data implies that cotranslational assembly is controlled by factors beyond interface geometry. Our data showed that fusing oligomerization-mediating domains to structurally unrelated sequences that prolongs their cotranslational exposure, can bias assembly pathways without altering the protein–protein interface. We speculate that the limited extent of heteromeric co-co assembly may also reflect spatial constraints imposed by polysome architecture, though other factors such as folding dynamics and chaperone engagement may also play a role. SucCD features a relatively small, planar interface and was predicted to assemble posttranslationally, in line with our results. Our data suggested, however, that the assembly pathway of SucCD is determined by the timing of interface exposure rather than by its size or geometry per se.

The coordination of cotranslational assembly with nascent chain folding is further reflected in the interplay with the ribosome-associated chaperone TF. We find that TF binds emerging oligomerization domains and enhances the efficiency of subsequent co-post assembly. The timing of the transition from chaperone engagement to complex formation is determined by the folding state of the nascent polypeptide. Partially folded polypeptides recruit TF via a high local concentration of unsatisfied contacts (19). Sequential addition of new residues enables continued cotranslational domain folding and compaction. This reduces the affinity for TF, allowing the folded oligomerization domain exposed at the ribosomal tunnel exit to be swiftly engaged by its partner subunit. In contrast to the endogenous complexes studied here, TF delayed the co-post assembly of heterologously expressed luciferase heterodimers, which originate from the marine bacterium *Vibrio harveyi* (17). Notably, the cotranslationally engaged LuxB subunit is a short (324 residues), single-domain protein and co-post assembly initiated before full domain exposure but reached only modest efficiency. We propose that TF fulfills a dual role in coordinating cotranslational complex assembly: it promotes proper domain folding to facilitate subsequent assembly

Downloaded from https://www.pnas.org by AMOLF LIBRARY SERVICES on November 18, 2026 from IP address 194.171.111.64.

and simultaneously prevents premature low-affinity, subunit interactions by binding partially folded intermediates.

Interestingly, although the lack of TF reduced the co-post assembly efficiency of all three protein complexes studied here, Δ*tig* cells generally do not exhibit reduced growth rates. One intriguing explanation is that given the redundancy of the *E. coli* chaperone network (19, 21, 26, 39–41), other chaperones like DnaK can compensate for the absence of TF. In contrast, no comparable mechanism may exist to offset defects in cotranslational assembly caused by disrupted operon order.

Materials and Methods

Plasmid Construction. All plasmids were cloned using standard procedures. CRIM plasmid derivatives encoding *carAB* genes were constructed as follows: For *carB* integration into Φ80 phage attachment-site, the $attP_{\lambda}$ attachment site together with the *cat* resistance cassette in pCAH56 were replaced by $att_{\Phi 80}$ and the *tetA* resistance marker provided by the CRIM (42) series (pAH162) using *SalI*/*Sph*I restriction sites, creating pTAH162. The *carA-yfp*, *carB,* and *carA-yfp_carB* sequences were PCR amplified from MC4100 Δ*thyA carA-yfp_carB* and cloned into pCAH56 (cm^R) (42) or the tetracycline resistant variant *pTAH56_162* (tet^R), generating plasmids *pCAH56_carA-yfp*, *pCAH56_carA-yfp_carB* and *pTAH56_162_carB*. To adjust relative gene expression, the Shine–Dalgarno sequence upstream of *carB* (AGGAGA) was replaced by AF1 (AACCGA) generating *pTAH56_162_AF1_carB*. AF1 was designed with EMOPEC (43).

Scar-Free Endogenous *yfp*-Tagging. For scar-free endogenous gene fusion with a *yfp* (mEYFP) affinity-tag, the thymidylate synthase encoding gene *thyA* was deleted by standard homologous recombination with subsequent removal of the kanamycin resistance marker (44). MC4100 Δ*thyA* cells were transformed with pKD46 (45) and electroporated subsequently with 100 ng purified PCR fragment coding the *yfp-thyA-yfp* sequence 5′- and 3′-flanked by 100 nt overhang sequences homologous to the integration sites of interest. Cells were recovered for 2 h 30 min in lysogeny broth (LB) supplemented with thymine (100 μg/mL), glycerol (0.2%), and arabinose (0.2%) for the expression of the λ-Red recombinase, plated on M9 minimal media plates supplemented with ampicillin ($M9_{Amp}$) and incubated for 2 d at 30 °C. Cells were replica plated on $M9_{Amp}$ plates and incubated overnight at 37 °C. Recombinant thymine prototrophic cells were restreaked onto $M9_{Amp}$ and incubated for 2 d at 37 °C. For the second homologous recombination step, 1 mL of LB_{Amp} supplemented with glycerol (0.2%) was inoculated with a single colony of thymine prototrophic cells. After incubating for 2 h at 37 °C, thymine (50 μg/mL) and arabinose (0.2%) were added to induce λ-Red recombinase, followed by further cultivation for 7 h at 37 °C. 100 μL cell suspension was plated on LB plates supplemented with thymine (100 μg/mL) and trimethoprim (20 μg/mL) and incubated overnight at 37 °C. Single colonies were tested for *yfp*-integration by colony PCR using gene of interest flanking primers and immunoblot analysis. All strains used in this study can be found in *SI Appendix*, Table S2.

Generation of *carAB* Operon and Split Strains. The *carAB* gene in MC4100 was deleted by standard P1 bacteriophage transduction (46) using *E. coli* strain MG1655 Δ*carAB::kan* as the donor strain. Gene deletions were controlled by colony PCR and immunoblot analysis. All integratable plasmids used in this study are CRIM plasmid derivatives (42) and were propagated in BW23472 (CGSC7837). Integration of CRIM plasmids into *E. coli* strain MC4100 Δ*carAB::kan* was performed and controlled as described before (42). The operon strain JA138 was created by integrating *pCAH56_carA-yfp_carB* into MC4100 Δ*carAB::kan*. The split strain JA153 was created by integrating *pCAH56_carA-yfp* and *pTAH56_162_AF1_carB* into the λ and Φ80 attachment site of MC4100 Δ*carAB::kan*, respectively.

For growth comparison, a spontaneous insertion of IS421 at the *lon* locus (*lon::IS421*) in the JA153 split strain was replaced by the wild-type *lon* locus by cotransduction with the *proC::kan* locus from wild-type MC4100 by P1vir transduction. Reintroduction of the wild-type *proC* locus was achieved by P1vir transduction using an MC4100 wild-type lysate and selected for on M9 minimal plates supplemented with 200 μM IPTG. All loci were confirmed by PCR.

SeRP of mEYFP-Tagged Subunits. Exponentially growing cells were harvested at OD_{600} = 0.45 to 0.5 by rapid-filtration and flash-freezing in liquid nitrogen. Expression of IPTG-controlled genes was induced by 10 μM IPTG for 15 to 20 min before cell harvest. Cell pellets were lysed under cryogenic conditions by mixer milling (2 min, 30 Hz) with SeRP lysis buffer (50 mM HEPES KOH pH 7.0, 100 mM NaCl, 10 mM $MgCl_2$, 5 mM $CaCl_2$, 50 μg/mL chloramphenicol, 1 mM PMSF, cOmplete EDTA-free protease inhibitor (Roche)). Upon thawing on ice, SeRP lysis buffer containing DNaseI was added to the lysate, and the lysate was cleared by centrifugation (20,000×*g*, 10 min, 4 °C). The supernatant was digested with self-made micrococcal nuclease (MNaseI) for 1 h at 25 °C, quenched by EGTA (6 mM). For the total translatome, ribosomes from one-fifth of the lysate were purified by sucrose density centrifugation (20 mM Tris-HCl pH 8.0, 100 mM NH_4Cl, 10 mM $MgCl_2$, 50 μg/mL chloramphenicol) and ribosome-protected mRNA-footprints were isolated by acid phenol extraction as previously described (47). The other four-fifths of the lysate was incubated with GFP-binder coupled sepharose beads (800 μl 1:1 suspension) for 1 h at 4 °C by end-to-end rotation. Beads were washed three times with wash buffer (50 mM HEPES KOH pH 7.0, 100 mM NaCl, 10 mM MgCl2, 50 μg/mL chloramphenicol, 1 mM PMSF, cOmplete EDTA-free protease inhibitor (Roche), 0.02% (v/v) Tween-20) for 5 min at 4 °C and sedimented at 500 rcf, 4 °C, for 1 min between washes. Ribosome-protected footprints were extracted by acid phenol extraction. cDNA libraries for deep sequencing analysis were generated as described previously (47) and sequenced on a HiSeq 2000 v4 (Illumina).

SeRP of TF Interactome. Nascent TF interactome was analyzed by SeRP as described previously (47), with three minor modifications. In brief, MG1655 tig::tig-TEV-avi cells were grown in EZ RDM medium lacking tryptophan, supplemented with 40 μg/mL biotin and lactose as a carbon source. Cells were harvested at OD_{600} = 0.45 to 0.5 by rapid-filtration and flash-freezing in liquid nitrogen. Cell pellets were lysed in frozen state by mixer milling (2 min, 30 Hz) with SeRP lysis buffer. Upon thawing on ice, SeRP lysis buffer containing DNaseI was added to the lysate and the lysate was cleared by centrifugation (20,000 rcf, 10 min, 4 °C). Supernatant was digested with self-made micrococcal nuclease (MNase, *nucB*) for 1 h at 25 °C, quenched by EGTA (6 mM). Ribosomes were purified by sucrose density centrifugation (20 mM Tris-HCl pH 8.0, 100 mM NH_4Cl, 10 mM $MgCl_2$, 50 μg/mL chloramphenicol) and resuspended in wash buffer [50 mM HEPES KOH pH 7.0, 100 mM NaCl, 10 mM $MgCl_2$, 50 μg/mL chloramphenicol, 1 mM PMSF, cOmplete EDTA-free protease inhibitor (Roche), 0.02% (v/v) Tween-20]. From one-fifth of the sample, ribosome-protected mRNA footprints were isolated by acid phenol extraction as previously described (47) for the total translatome. The remaining sample was mixed with StrepTactin-Sepharose resin that was pre-equilibrated in wash buffer and incubated on an overhead roller for 1 h at 4 °C. Beads were washed three times with wash buffer for 5 min at 4 °C, and sedimented at 500 rcf, 4 °C, for 1 min between washes. Ribosome-protected footprints were extracted by acid phenol extraction. cDNA libraries for deep sequencing analysis were generated as described previously (47) and sequenced on a HiSeq 2000 v4 (Illumina).

Disome Selective Profiling. *E. coli* MG1655 cells were grown at 37 °C in EZ RDM with lactose as the sole carbon source and harvested by rapid filtration and flash-freezing in liquid nitrogen at OD_{600} = 0.45 to 0.5. Cells were lysed in the presence of DiSP lysis buffer [50 mM HEPES KOH pH 7.0, 100 mM NaCl, 10 mM $MgCl_2$, 5 mM $CaCl_2$, 50 μg/mL chloramphenicol, 1 mM PMSF, cOmplete EDTA-free protease inhibitor (Roche), 0.4% (v/v) Triton X-100, 0.1% Nonidet P40 Substitute] under cryogenic conditions by mixer milling (2 min, 30 Hz). Upon thawing, 750 μL DiSP lysis buffer containing DNase I was added and the lysate was cleared by centrifugation (20,000 rcf, 10 min, 4 °C). The supernatant was digested with self-made micrococcal nuclease for 10 min at 25 °C, quenched by EGTA (6 mM). 250 μg total RNA was layered over a 5 to 45% sucrose gradient (50 mM HEPES-KOH pH 7.0, 500 mM KCl, 10 mM $MgCl_2$, 50 μg/mL chloramphenicol) and separated by ultracentrifugation for 3.5 h at 35,000 rpm, 4 °C. Monosome and disome fractions were isolated, and the RNA extracted by acid phenol. cDNA libraries for deep sequencing analysis were generated as described previously (48) and sequenced on a NextSeq 550 (Illumina).

Sequencing Data Analysis. SeRP libraries were sequenced on a HiSeq 2000 (Illumina). 3′ adapter sequences were trimmed from demultiplexed raw sequencing files using Cutadapt (49) with the following parameters: -q20 -m23

Downloaded from https://www.pnas.org by AMOLF LIBRARY SERVICES on November 18, 2026 from IP address 194.171.111.64.

--discard-untrimmed -O6 -a CTGTAGGCACCATCAATTCGTATGCCGTCTTCTGCTTG. Trimmed reads were aligned to the *E. coli* reference genome (NC_000913.3) using Bowtie1 (50) with the parameters -m1 -t -n2 --best --strata. Aligned reads were assigned to the P-site using a custom Julia script (Script2) published previously (6) with the parameters --length 15:45 --assignment_type 3.

DiSP libraries were sequenced on a NextSeq 550 (Illumina). 3′ adapter sequences were trimmed from demultiplexed raw sequencing files using Cutadapt with the parameters -q 20 -m23 --discard-untrimmed -O6 -a ATCGTAGATCGGAAGAGCACACGTCTGAACTCCAGTCAC. Unique molecular identifiers (UMI) were extracted from each read and added to the read identifier using a custom Julia script (Script1) published previously (6) with the parameters --umi3 5 --umi5 2.

Trimmed reads were aligned to the *E. coli* reference genome (NC_000913.3) using Bowtie1 with the parameters -m 1 -t -n 2 --best --strata. Aligned reads were deduplicated based on their UMI and assigned to the P-site using Script2 with the parameters --use_umis --length 15:45 --assignment_type 3.

The data were plotted using RiboSeqTools (available at https://github.com/ilia-kats/RiboSeqTools).

Growth Assays. Strains were grown overnight at 37 °C in M9 minimal medium (51) (33.7 mM Na_2HPO_4, 22 mM KH_2PO_4, 8.55 mM NaCl, 9.35 mM NH_4Cl, 1 mM $MgSO_4$, 0.3 mM $CaCl_2$, 0.134 mM EDTA, 31 μM $FeCl_3$, 6.2 μM $ZnCl_2$, 0.76 μM $CuCl_2$, 0.42 μM $CoCl_3$, 1.62 μM H_3BO_3, 81 nM $MnCl_2$) with 0.4% glucose as the carbon source and supplemented with either IPTG or uracil and arginine as indicated. Cells were pelleted (2 min at 2,000 rcf), washed, and resuspended in M9 minimal medium. Cell suspensions were then diluted into 750 μL M9 minimal medium supplemented with either IPTG or uracil and arginine as indicated, in a 24-well flat bottom plate (Greiner Cellstar 662160) covered with a breathable sealing membrane (Diversified Biotech BEM-1). Growth was followed automatically by measuring optical density (OD) at 600 nm using the SpectroStar Nano plate reader at 37 °C, 500 rpm double-orbital shaking. A reparameterized Richards (generalized logistic) model (52, 53) (Eq. **1**) was fitted to the logarithm of the normalized growth data ($\ln(OD_t/OD_0)$) using scipy.optimize.curve_fit (54) with u and A parameter bounds chosen to ensure numerical stability. In this formula $y(t)$ is the fitted optical density at time t, y_0 represents the baseline level of y, A its total amplitude, v controls the asymmetry of the sigmoid, u is the maximum specific growth rate defined as the tangent at the inflection point, and d is the lag-time defined as the x-axis intercept of this tangent.

$$\ln(OD_t/OD_0) = y(t) = y_0 + A\left(1 + v\,\exp\left[1 + v + \frac{u}{A}(1+v)^{1+\frac{1}{v}}\right](d-t)\right)^{-\frac{1}{v}}. \quad [\mathbf{1}]$$

The generation time was derived from the maximum specific growth rate (u) as

$$T_d = \frac{\ln(2)}{u}. \quad [\mathbf{2}]$$

Coupling of Neutravidin-DNA Handles to Beads. Neutravidin-coupled DNA handles were prepared and coupled to polystyrene beads (55). DIG- and biotin-modified primers (2DIGfw5kbp, 3BIOrev5kbp) were used to generate dsDNA handles by two-step Phire Green Hot Start II PCR (Thermo Scientific), with pOSIP-TT as template. The resulting 5 kb amplicon was purified with the QIAquick PCR Purification Kit (QIAGEN) and conjugated to neutravidin (Thermo Scientific) by overnight incubation of 10 nM DNA with 2 μM neutravidin in 10 mM PBS at 4 °C. In parallel, 2.1 μm carboxyl-functionalized polystyrene beads (Spherotech) were derivatized with anti-digoxigenin antibody (anti-DIG, Roche) via EDAC carbodiimide coupling, using the PolyLink protein coupling kit (Polysciences). The neutravidin-coupled 5 kb DNA handles (1.7 μm contour length) were bound to the anti-DIG beads at a ratio of approximately 10 DNA handles per bead (30 min, 4 °C). Bead-bound handles were washed several times in Tico buffer (20 mM HEPES-KOH pH 7.6, 10 mM (Ac)2 Mg, 30 mM $AcNH_4$, 4 mM 2-mercaptoethanol) and divided into two aliquots for ribosome coupling.

Coupling of Ribosomes to Beads with DNA Handles. Ribosomes biotinylated in vivo via an avi-tagged uL4 were purified from the RNase-deficient *E. coli* K-12 strain CAN20-12E (56). The purified ribosomes were incubated with one aliquot of the Tico-washed neutravidin–DNA beads at a final concentration of 350 nM in the presence of murine RNase inhibitor (New England Biolabs) for 30 min at 4 °C. Excess, unbound ribosomes were removed by pelleting, and the ribosome–bead conjugates were washed once in Tico buffer before direct resuspension in the cell-free transcription/translation reaction described below.

Cell-Free Protein Synthesis and Cotranslational Labeling. In vitro translation reactions were performed in a custom ribosome-free derivative of the bacterial PURE system (PURExpress Δribosomes, New England Biolabs), reconstituted with the bead-coupled *E. coli* ribosomes, generated as described above, and 5.5 nM linear DNA template. The reaction was supplemented with murine RNase inhibitor (New England Biolabs) and 125 nM TF and suppressor tRNA (10 μM Biotin-XX-AF_tRNA, amber, CloverDirect) for the incorporation of biotinylated-lysine at a 5′ amber (TAG) codon. Translation proceeded for 30 min at 37 °C, with the C-terminal SecMstr arrest peptide retaining the nascent chain on the ribosome (57, 58). After the reaction, bead-tethered, stalled ribosome-nascent chain complexes (RNCs) were resuspended in Tico buffer (20 mM HEPES-KOH pH 7.6, 10 mM (Ac)2 Mg, 30 mM $AcNH_4$, 4 mM 2-mercaptoethanol, 4 °C) and diluted to 300 μL in the same buffer prior to optical trapping. Oxygen was depleted using the P_2O scavenging system (3 U/mL pyranose oxidase, 90 U/mL catalase, 50 mM glucose; Sigma).

Optical Tweezers Assay and Single-Molecule Data Analysis. Single-molecule force spectroscopy combined with multicolor confocal fluorescence detection was performed on a C-trap instrument (Lumicks, Amsterdam). A 1,064 nm laser was split into two orthogonally polarized, independently steerable beams. Two confocal excitation lines (532 and 638 nm) and single-photon-sensitive avalanche photodiodes (APD) enabled dual-color fluorescence readout. All measurements were carried out in a laminar-flow cell under pressure-driven flow. Force and position data were collected at 50 kHz and down-sampled to 500 Hz. Trap calibration relied on the power-spectrum method, by fitting Lorentzian functions to the Brownian-motion power spectra of trapped beads. Laser intensity was constant throughout, giving a trap stiffness of 260 ± 50 pN/μm.

For tethering, an RNC-carrying bead held in one trap was brought near a second bead that carried free neutravidin-DNA handles. After a short period to allow contact formation between the N-terminal biotin-lysine and the neutravidin-DNA handle, the two beads were separated to an interbead distance of approximately 3 μm. Successful tether formation was indicated by a small force increase. Force-extension data were recorded in cycling mode, with the steerable trap moving at 0.2 μm/s between a 2 μm minimum bead separation and forces of up to 65 pN. Each tether's force-extension curve was described by two worm-like chains in series, an extensible WLC (eWLC) for the DNA handle and an inextensible WLC for the stalled nascent chain, yielding an average DNA persistence length of 43 ± 10 nm (SD 5 nm) and stretch modulus of 1,037 ± 388 pN/nm (SD 194 pN/nm).

Data, Materials, and Software Availability. The DiSP sequencing data pertaining to the *trpED* operon were previously published (16) and are accessible through GEO Series Accession No. GSE291748 (59). All other sequencing data (NGS data) are deposited in the European Nucleotide Archive (ENA) under Accession No. PRJEB108825 (60). All other data are included in the manuscript and/or *SI Appendix*.

ACKNOWLEDGMENTS. We acknowledge funding by the German-Israeli Foundation for Scientific Research and Development (I-1564-417.13/2023) to G.K., a European Research Council Synergy Grant (ERC-SyG-101072047-CoTransComplex) to B.B. and S.J.T., and a European Research Council Advanced Grant (743118) to B.B. Views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them. Work in the group of S.J.T. is supported by the Netherlands Organization for Scientific Research. F.W. received funding from the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie Grant Agreement No. 745798. J.L.S. was supported by a PhD fellowship from the Studienstiftung des deutschen Volkes (German Academic Scholarship Foundation) and a completion grant from the Baden-Württemberg State Graduate Funding. J.L.S. is a member of the Heidelberg Biosciences International Graduate School.

Author affiliations: [a]Center for Molecular Biology of Heidelberg University, German Cancer Research Center - Center for Molecular Biology of Heidelberg University Alliance, Heidelberg 69120, Germany; [b]Autonomous Matter Department, AMOLF institute, Amsterdam 1098 XG, The Netherlands; and [c]Bionanoscience Department of Delft University of Technology and Kavli Institute of Nanoscience Delft, Delft 2629 HZ, The Netherlands

Downloaded from https://www.pnas.org by AMOLF LIBRARY SERVICES on November 18, 2026 from IP address 194.171.111.64.

1. S. Juszkiewicz, R. S. Hegde, Quality control of orphaned proteins. *Mol. Cell* **71**, 443–457 (2018).
2. F. Morales-Polanco, J. H. Lee, N. M. Barbosa, J. Frydman, Cotranslational mechanisms of protein biogenesis and complex assembly in eukaryotes. *Annu. Rev. Biomed. Data Sci.* **5**, 67–94 (2022).
3. J. Koubek, J. Schmitt, C. V. Galmozzi, G. Kramer, Mechanisms of cotranslational protein maturation in bacteria. *Front. Mol. Biosci.* **8**, 689755 (2021).
4. M. Badonyi, J. A. Marsh, Large protein complex interfaces have evolved to promote cotranslational assembly. *eLife* **11**, e79602 (2022).
5. S. Mallik *et al.*, Structural determinants of co-translational protein complex assembly. *Cell* **188**, 764–777.e22 (2025).
6. M. Bertolini *et al.*, Interactions between nascent proteins translated by adjacent ribosomes drive homomer assembly. *Science* **371**, 57–64 (2021).
7. A. Shiber *et al.*, Cotranslational assembly of protein complexes in eukaryotes revealed by ribosome profiling. *Nature* **561**, 268–272 (2018).
8. I. Kamenova *et al.*, Co-translational assembly of mammalian nuclear multisubunit complexes. *Nat. Commun.* **10**, 1740 (2019).
9. M. Seidel *et al.*, Co-translational assembly orchestrates competing biogenesis pathways. *Nat. Commun.* **13**, 1224 (2022).
10. O. O. Panasenko *et al.*, Co-translational assembly of proteasome subunits in NOT1-containing assemblysomes. *Nat. Struct. Mol. Biol.* **26**, 110–120 (2019).
11. F. Morales-Polanco *et al.*, Core Fermentation (CoFe) granules focus coordinated glycolytic mRNA localization and translation to fuel glucose fermentation. *iScience* **24**, 102069 (2021).
12. F. Liu, D. K. Jones, W. J. de Lange, G. A. Robertson, Cotranslational association of mRNA encoding subunits of heteromeric ion channels. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 4859–4864 (2016).
13. O. Lautier *et al.*, Co-translational assembly and localized translation of nucleoporins in nuclear pore complex biogenesis. *Mol. Cell* **81**, 2417–2427.e5 (2021).
14. D. Zipser, Studies on the ribosome-bound beta-galactosidase of *Escherichia coli*. *J. Mol. Biol.* **7**, 739–751 (1963).
15. Y. Kiho, A. Rich, Induced enzyme formed on bacterial polyribosomes. *Proc. Natl. Acad. Sci. U.S.A.* **51**, 111–118 (1964).
16. A. Roeselová *et al.*, The ribosome synchronizes folding and assembly to promote oligomeric protein biogenesis. *Mol. Cell* **86**, 693–706.e7 (2026).
17. Y.-W. Shieh *et al.*, Operon structure and cotranslational subunit association direct protein assembly in bacteria. *Science* 350, 678–680 (2015).
18. E. Oh *et al.*, Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. *Cell* **147**, 1295–1308 (2011).
19. C. V. Galmozzi *et al.*, Proteome-wide determinants of co-translational chaperone binding in bacteria. *Nat. Commun.* **16**, 4361 (2025).
20. G. Kramer *et al.*, L23 protein functions as a chaperone docking site on the ribosome. *Nature* **419**, 171–174 (2002).
21. P. Genevaux *et al.*, In vivo analysis of the overlapping functions of DnaK and trigger factor. *EMBO Rep.* **5**, 195–200 (2004).
22. T. Saio, X. Guan, P. Rossi, A. Economou, C. G. Kalodimos, Structural basis for protein antiaggregation activity of the trigger factor chaperone. *Science* 344, 1250494 (2014).
23. C. M. Kaiser *et al.*, Real-time observation of trigger factor function on translating ribosomes. *Nature* **444**, 455–460 (2006).
24. V. R. Agashe *et al.*, Function of trigger factor and DnaK in multidomain protein folding: Increase in yield at the expense of folding speed. *Cell* **117**, 199–209 (2004).
25. G. Kramer *et al.*, Functional dissection of *Escherichia coli* trigger factor: Unraveling the function of individual domains. *J. Bacteriol.* **186**, 3777–3784 (2004).
26. E. Deuerling, A. Schulze-Specking, T. Tomoyasu, A. Mogk, B. Bukau, Trigger factor and DnaK cooperate in folding of newly synthesized proteins. *Nature* **400**, 693–696 (1999).
27. A. Mashaghi *et al.*, Reshaping of the conformational search of a protein by the chaperone trigger factor. *Nature* **500**, 98–101 (2013).
28. K. Till *et al.*, Trigger factor accelerates nascent chain compaction and folding. *Proc. Natl. Acad. Sci. U.S.A.* 122, e2422678122 (2025).
29. K. Singhal, J. Vreede, A. Mashaghi, S. J. Tans, P. G. Bolhuis, The trigger factor chaperone encapsulates and stabilizes partial folds of substrate proteins. *PLoS Comput. Biol.* **11**, e1004444 (2015).
30. J. N. Wells, L. T. Bergendahl, J. A. Marsh, Operon gene order is optimized for ordered protein complex assembly. *Cell Rep.* **14**, 679–685 (2016).
31. S. Bedi, S. M. Rose, S. Kaur, P. Negi, S. Sinha, Protein-protein interactions as determinants of operon architecture. *Biochim. Biophys. Acta (BBA)-Gen. Subj.* **1869**, 130794 (2025).
32. B. W. Miles, J. B. Thoden, H. M. Holden, F. M. Raushel, Inactivation of the amidotransferase activity of carbamoyl phosphate synthetase by the antibiotic acivicin. *J. Biol. Chem.* **277**, 4368–4373 (2002).
33. I. Mermershtain *et al.*, Idiosyncrasy and identity in the prokaryotic Phe-system: Crystal structure of *E. coli* phenylalanyl-tRNA synthetase complexed with phenylalanine and AMP. *Protein Sci.* **20**, 160–167 (2011).
34. A. A. Morollo, M. J. Eck, Structure of the cooperative allosteric anthranilate synthase from *Salmonella typhimurium*. *Nat. Struct. Biol.* **8**, 243–247 (2001).
35. S. Schlee *et al.*, Prediction of quaternary structure by analysis of hot spot residues in protein-protein interfaces: The case of anthranilate phosphoribosyltransferases. *Proteins* **87**, 815–825 (2019).
36. D. Charlier, P. Le Nguyen Minh, M. Roovers, Regulation of carbamoylphosphate synthesis in *Escherichia coli*: An amazing metabolite at the crossroad of arginine and pyrimidine biosynthesis. *Amino Acids* **50**, 1647–1661 (2018).
37. J. A. Marsh *et al.*, Protein complexes are under evolutionary selection to assemble via ordered pathways. *Cell* **153**, 461–470 (2013).
38. M. G. Plach *et al.*, Evolutionary diversification of protein-protein interactions by interface add-ons. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E8333–E8342 (2017).
39. E. Deuerling *et al.*, Trigger factor and DnaK possess overlapping substrate pools and binding specificities. *Mol. Microbiol.* **47**, 1317–1328 (2003).
40. S. Vorderwülbecke *et al.*, Low temperature or GroEL/ES overproduction permits growth of *Escherichia coli* cells lacking trigger factor and DnaK. *FEBS Lett.* **559**, 181–187 (2004).
41. G. Calloni *et al.*, DnaK functions as a central hub in the *E. coli* chaperone network. *Cell Rep.* **1**, 251–264 (2012).
42. A. Haldimann, B. L. Wanner, Conditional-replication, integration, excision, and retrieval plasmid-host systems for gene structure-function studies of bacteria. *J. Bacteriol.* **183**, 6384–6393 (2001).
43. M. T. Bonde *et al.*, Predictable tuning of protein expression in bacteria. *Nat. Methods* **13**, 233–236 (2016).
44. K. A. Datsenko, B. L. Wanner, One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6640–6645 (2000).
45. A. M. Stringer *et al.*, FRUIT, a scar-free system for targeted chromosomal mutagenesis, epitope tagging, and promoter replacement in *Escherichia coli* and *Salmonella enterica*. *PLoS ONE* **7**, e44841 (2012).
46. L. C. Thomason, N. Costantino, D. L. Court, *E. coli* genome manipulation by P1 transduction. *Curr. Protoc. Mol. Biol.* Chapter 1, 1.17.1–1.17.8 (2007).
47. A. H. Becker, E. Oh, J. S. Weissman, G. Kramer, B. Bukau, Selective ribosome profiling as a tool for studying the interaction of chaperones and targeting factors with nascent polypeptide chains and ribosomes. *Nat. Protoc.* **8**, 2212–2239 (2013).
48. C. V. Galmozzi, D. Merker, U. A. Friedrich, K. Döring, G. Kramer, Selective ribosome profiling to study interactions of translating ribosomes in yeast. *Nat. Protoc.* **14**, 2279–2317 (2019).
49. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10 (2011).
50. B. Langmead, C. Trapnell, M. Pop, S. L. Salzberg, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
51. J. H. Miller, *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, 1972).
52. F. J. Richards, A flexible growth function for empirical use. *J. Exp. Bot.* **10**, 290–301 (1959).
53. M. H. Zwietering, I. Jongenburger, F. M. Rombouts, K. van't Riet, Modeling of the bacterial growth curve. *Appl. Environ. Microbiol.* **56**, 1875–1881 (1990).
54. P. Virtanen *et al.*, SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).
55. V. S. Jadhav, D. Brüggemann, F. Wruck, M. Hegner, Single-molecule mechanics of protein-labelled DNA handles. *Beilstein J. Nanotechnol.* **7**, 138–148 (2016).
56. F. Wruck, A. Katranidis, K. H. Nierhaus, G. Büldt, M. Hegner, Translation and folding of single proteins in real time. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E4399–E4407 (2017).
57. F. Cymer, R. Hedman, N. Ismail, G. von Heijne, Exploration of the arrest peptide sequence space reveals arrest-enhanced variants. *J. Biol. Chem.* **290**, 10208–10215 (2015).
58. N. Kempf *et al.*, A novel method to evaluate ribosomal performance in cell-free protein synthesis systems. *Sci. Rep.* **7**, 46753 (2017).
59. A. Roeselová *et al.*, The ribosome synchronizes folding and assembly to promote oligomeric protein biogenesis. Gene Expression Omnibus. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE291748. Deposited 12 March 2025.
60. J. L. Schmitt *et al.*, Molecular drivers of cotranslational protein assembly in bacteria. European Nucleotide Archive. https://www.ebi.ac.uk/ena/browser/view/PRJEB108825. Deposited 26 February 2026.

Downloaded from https://www.pnas.org by AMOLF LIBRARY SERVICES on November 18, 2026 from IP address 194.171.111.64.