Interactions between nascent proteins translated by adjacent ribosomes drive homomer assembly

Bertolini, Matilde; Fenzl, Kai; Kats, Ilia; Wruck, Florian; Tippmann, Frank; Schmitt, Jaro; Auburger, Josef Johannes; Tans, Sander; Bukau, Bernd; Kramer, Günter

DOI
10.1126/science.abc7151

Publication date
2021

Document Version
Final published version

Published in
Science

Citation (APA)
Bertolini, M., Fenzl, K., Kats, I., Wruck, F., Tippmann, F., Schmitt, J., Auburger, J. J., Tans, S., Bukau, B., & Kramer, G. (2021). Interactions between nascent proteins translated by adjacent ribosomes drive homomer assembly. *Science*, *371*(6524), 57-64. Article eabc7151. https://doi.org/10.1126/science.abc7151

RESEARCH ARTICLE SUMMARY

PROTEIN SYNTHESIS

Interactions between nascent proteins translated by adjacent ribosomes drive homomer assembly

Matilde Bertolini*, Kai Fenzl*, Ilia Kats, Florian Wruck, Frank Tippmann, Jaro Schmitt, Josef Johannes Auburger, Sander Tans, Bernd Bukau†, Günter Kramer†

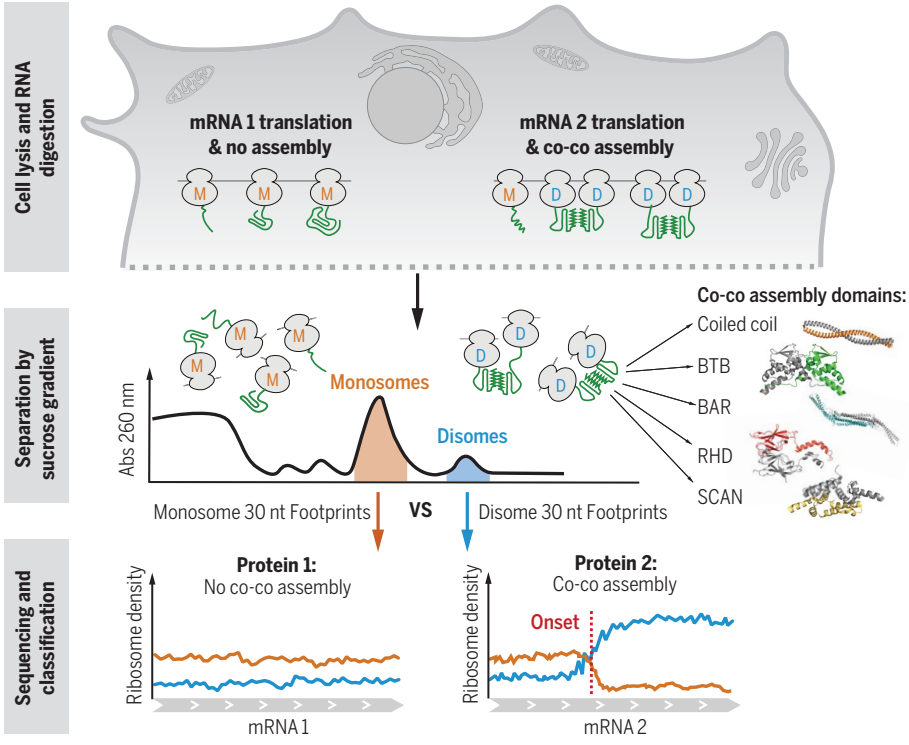
INTRODUCTION: Most newly synthesized proteins associate into macromolecular complexes to become functional. Complex formation requires that subunits find each other in the crowded cellular environment while avoiding unspecific interactions and aggregation.

Recent findings indicate that native complex formation is facilitated by coupling protein synthesis by ribosomes (translation) with folding and assembly. Studies analyzing formation of heteromeric complexes have elucidated the cotranslational engagement of nascent subunits by their fully translated, diffusing partner proteins (co-post assembly).

We considered an alternative assembly mechanism that involves the interaction of two nascent subunits during their concurrent translation (co-co assembly) and thereby uncouples assembly from subunit diffusion. Provided that the interacting subunits are synthetized on one polysome, co-co assembly would increase the fidelity of homomer formation, prevent nonspecific interactions with structural homologs and isoforms, and facilitate spatial and temporal coordination of the process. Whether cells employ co-co assembly as a general strategy for complex assembly, when and how efficiently nascent subunits interact, and what mechanisms are driving the process remain unclear.

RATIONALE: Upon co-co assembly, single translating ribosomes (monosomes) become connected via nascent proteins. These ribosome pairs (disomes) persist during nuclease treatment of cell lysates and protect mRNA fragments of 30 nucleotides in length (ribosome footprints).

Our approach relies on the different sucrose gradient sedimentation properties of disomes and monosomes. Sequencing of footprints isolated from monosome and disome fractions identifies co-co assembly candidates across the nascent proteome as the mRNAs on which ribosomes shift from the monosome to the disome fraction during translation [disome selective profiling (DiSP)]. The position of the shift defines the co-co assembly onset and reveals exposed nascent protein segments that mediate dimerization.



Disome selective profiling reveals proteome-wide interactions between nascent proteins. Ribonuclease treatment of human cell lysates generates monosomes (M) and nascent protein–connected disomes (D) that are purified by sucrose gradient centrifugation. Ribosome-protected footprints from both fractions are deep-sequenced. A shift of elongating ribosomes from the monosome to the disome fraction indicates co-co assembly. The mRNA position of the shift reveals the dimerization motif that mediates assembly. Abs, absorbance; nt, nucleotides.

RESULTS: We employed DiSP to reveal comprehensive information about the co-co assembling proteome of two human cell lines and mechanistic principles of the assembly process. Interactions between nascent subunits are highly prevalent, involving thousands of candidate proteins from different cellular compartments. Co-co assembly is mostly employed to form homomeric rather than heteromeric complexes and is generally correlated with the exposure of N-terminal dimerization interfaces. Five conserved structural motifs are the main drivers of co-co assembly; among these, coiled coils are most prevalent, followed by BTB, BAR, SCAN, and RHD domains.

Reconstitution in bacteria revealed that this process can occur independent of dedicated, eukaryote-specific assembly factors and minimally relies on the dimerization propensity of nascent protein N termini.

Finally, we monitored the composition of lamin dimers inside human cells and showed that homodimer-forming subunits are templated by one transcript. This observation implies that cells may generally employ co-co assembly on a polysome to avoid mixing isoforms that share identical dimerization domains.

CONCLUSION: Our study shows a previously unrecognized level of coupling of protein synthesis with complex assembly and provides direct evidence for the widespread occurrence of cotranslational interactions between nascent subunits in human cells.

We propose that the polysome constitutes the platform for most co-co assembly interactions. This enhances the efficiency and accuracy of homomer formation and enables cells to independently evolve functionally diverse homomeric protein complexes that use recurrent oligomerization domains. ■

The list of author affiliations is available in the full article.
*These authors contributed equally to this work.
†Corresponding author. Email: g.kramer@zmbh.uni-heidelberg.de (G.K.); bukau@zmbh.uni-heidelberg.de (B.B.)

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abc7151

RESEARCH ARTICLE

PROTEIN SYNTHESIS

Interactions between nascent proteins translated by adjacent ribosomes drive homomer assembly

Matilde Bertolini[1]*, Kai Fenzl[1]*, Ilia Kats[1]†, Florian Wruck[2], Frank Tippmann[1], Jaro Schmitt[1], Josef Johannes Auburger[1], Sander Tans[2,3], Bernd Bukau[1]‡, Günter Kramer[1]‡

Accurate assembly of newly synthesized proteins into functional oligomers is crucial for cell activity. In this study, we investigated whether direct interaction of two nascent proteins, emerging from nearby ribosomes (co-co assembly), constitutes a general mechanism for oligomer formation. We used proteome-wide screening to detect nascent chain–connected ribosome pairs and identified hundreds of homomer subunits that co-co assemble in human cells. Interactions are mediated by five major domain classes, among which N-terminal coiled coils are the most prevalent. We were able to reconstitute co-co assembly of nuclear lamin in *Escherichia coli*, demonstrating that dimer formation is independent of dedicated assembly machineries. Co-co assembly may thus represent an efficient way to limit protein aggregation risks posed by diffusion-driven assembly routes and ensure isoform-specific homomer formation.

Sophisticated mechanisms have evolved to ensure efficient and accurate protein complex biogenesis, including the fine-tuning of subunit expression to match complex stoichiometries (*1*), the employment of general or dedicated chaperones to guide oligomerization (*2*–*4*), the colocalization of subunit synthesis (*5*–*7*), and the timely oligomerization by coupling translation and subunit interactions (cotranslational assembly) (*3*, *8*, *9*). Selective ribosome profiling (SeRP) has provided mechanistic details of cotranslational assembly for *Vibrio harveyi* luciferase expressed in *Escherichia coli* (*3*) and several heteromeric complexes in yeast (*8*). In all cases studied, a freely diffusing, presumably folded protein engages its nascent partner subunit (co-post assembly).

In this study, we tested whether cotranslational assembly of protein complexes may also occur via association of two nascent subunits concurrently translated by two ribosomes (co-co assembly). A priori, co-co assembly may involve nascent chains synthesized on two different mRNAs (in trans) or, for homo-oligomer assembly, on the same mRNA (in cis). Notably, cis assembly does not require that distinct mRNA molecules colocalize in the cytosol and enables transcript-specific homomeric complex generation, avoiding undesired interactions between closely related proteins or wild-type and mutant alleles (*10*). Although co-co assembly has already been proposed for individual protein complexes in different organisms (*10*–*14*), direct experimental evidence that two ribosome–nascent chain complexes interact is still missing, and we lack any information on the prevalence, molecular mechanisms, and relevance of this proposed assembly process. We thus developed disome selective profiling (DiSP)—an unbiased, proteome-wide screening based on ribosome profiling (*15*)—to reveal the co-co assembly proteome in human cells.

[1]Center for Molecular Biology of Heidelberg University (ZMBH) and German Cancer Research Center (DKFZ), DKFZ-ZMBH Alliance, Im Neuenheimer Feld 282, Heidelberg D-69120, Germany. [2]AMOLF, Science Park 104, 1098 XG Amsterdam, Netherlands. [3]Department of Bionanoscience, Delft University of Technology and Kavli Institute of Nanoscience Delft, 2629HZ Delft, Netherlands.
*These authors contributed equally to this work.
†Present address: Computational Genomics and System Genetics, DKFZ, Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany.
‡Corresponding author. Email: g.kramer@zmbh.uni-heidelberg.de (G.K.); bukau@zmbh.uni-heidelberg.de (B.B.)

DiSP reveals widespread disome formation mediated by nascent chain interactions

To identify co-co assembling complexes across the proteome, we reasoned that ribosome pairs (disomes) connected by their exposed nascent chains will remain connected even upon mRNA digestion. Thus, it should be possible to detect co-co assembly candidates by ribonuclease (RNase) treatment of cell lysates, followed by separation of monosomes and disomes in sucrose gradients and deep sequencing of 30–nucleotide (nt) ribosomal footprints from both fractions (DiSP; Fig. 1A and fig. S1A). The disome fraction will also contain RNase-resistant disomes that form upon collision of ribosomes that translate the same mRNA; however, these disomes will protect double-length (60-nt) mRNA fragments (*16*) and are not analyzed by DiSP. Translating ribosomes engaged in co-co assembly will shift from the monosome to the disome fraction upon nascent chain dimerization, which could be detected by analyzing the relative footprint density of both samples (separately or as enrichment of disome over monosome) along a gene's coding sequence (Fig. 1A). In contrast to SeRP, which has been used to explore co-post assembly of selected protein complexes (*3*, *8*), DiSP can provide proteome-wide interaction profiles of all translating ribosomes.

We initially performed DiSP of human embryonic kidney 293-T (HEK293-T) cells. To identify co-co assembly candidates, we first

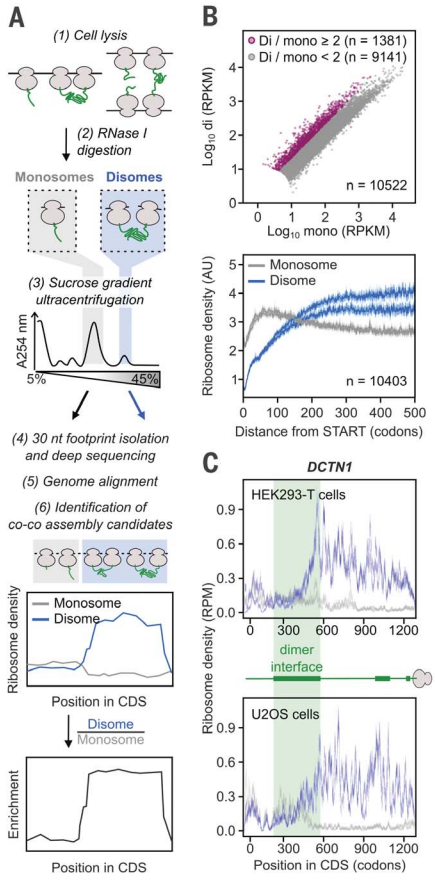


Fig. 1. Disome selective profiling (DiSP) reveals widespread disome formation. (**A**) Experimental procedure of DiSP. Cell lysates are treated with RNase (1 and 2); monosomes and disomes are separated by sucrose gradient ultracentrifugation (3); and ~30-nt-long ribosome footprints are extracted, converted into a DNA library, and sequenced (4). Co-co assembly candidates are identified by a shift of the footprint density from monosome to disome fraction, or by a disome-over-monosome enrichment profile (5 and 6). A254 nm, absorbance at 254 nm. (**B**) Comparison of disome (di) and monosome (mono) footprint density of all detected genes in HEK293-T cells (top; one replicate shown). Average footprint density along the coding sequence of all detected genes (metagene) aligned to the start of translation (bottom; two biological replicates). RPKM, reads per kilobase per million mapped reads. (**C**) Monosome (gray) and disome (blue) footprint density along the coding sequence (CDS) of *DCTN1*. The cartoon shows exposed nascent chain segments during translation; green bars indicate dimerization interfaces. DiSP data of HEK293-T cells (two biological replicates) and U2OS cells (two biological replicates) are compared. RPM, reads per million.

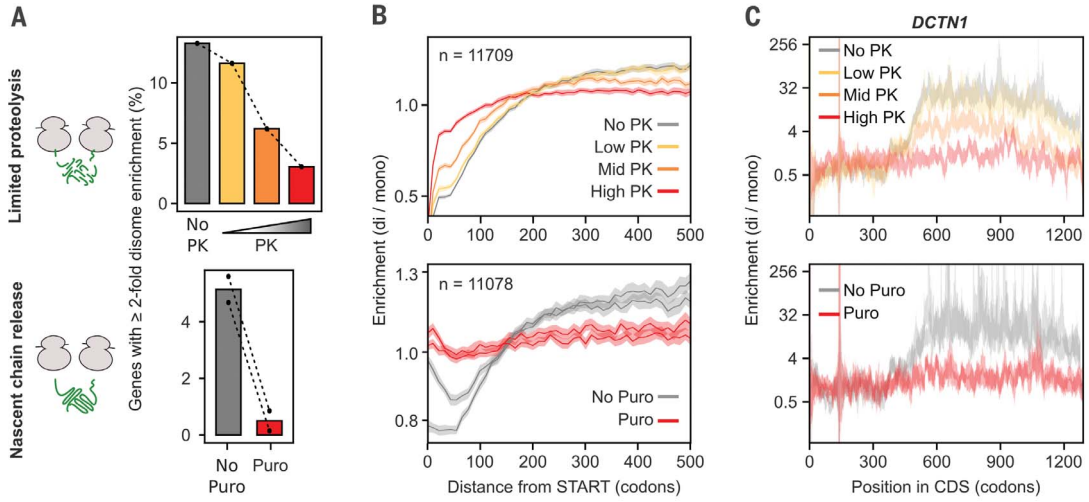


Fig. 2. Disome formation is nascent chain dependent. (**A**) DiSP was performed on lysates treated with increasing proteinase K (PK, one biological replicate) concentrations or with puromycin (Puro, two biological replicates) to degrade or release nascent chains. Both treatments resulted in a large depletion of genes with ≥twofold higher footprint density in the disome fraction than in the monosome fraction. (**B**) Metagene enrichment profiles (disomemonosome) aligned to translation start of all detected genes in PK (top) and Puro (bottom) DiSP experiments. (**C**) Enrichment profiles (disome/monosome) of *DCTN1* of untreated DiSP samples and samples treated with increasing concentrations of proteinase K (PK, top) or with puromycin (Puro, bottom).

compared gene-specific footprint densities in the disome and monosome fractions, revealing more than 1300 genes with a disome-over-monosome enrichment value ≥2 (Fig. 1B, top). A metagene profile of the averaged monosome and disome density along all coding sequences showed that early during translation, when nascent chains are short, ribosomes mostly migrated as monosomes, followed by a steady disome enrichment that leveled out at ~200 codons (Fig. 1B, bottom). The monosome-to-disome shift of translating ribosomes occurred only in a subset of genes, supporting the assumption that it depended on interaction properties of nascent chains (Fig. 1B, top, and fig. S1B). One example among the twofold disome enriched genes is *DCTN1*, which encodes $p150^{glued}$, a subunit of the dynactin motor complex. Ribosomes that translate *DCTN1* convert from monosomes to disomes near codon 430, when ~400 amino acids of nascent $p150^{glued}$ are exposed on the ribosomal surface. This N-terminal segment includes major parts of the coiled-coil dimerization domain, suggesting that the disome shift was caused by cotranslational homodimerization (Fig. 1C, top). Repeating DiSP in U2OS cells, we found a large overlap of disome-enriched genes and robustly correlated enrichment profiles (Fig. 1C and fig. S1, B and C), demonstrating that disome formation is a general feature of a specific subset of nascent proteins across different cell types.

To challenge our model that disome formation is mediated by nascent proteins, we explored whether disome shifts were sensitive to release or degradation of nascent chains. Treatment of lysates with puromycin (Puro) or increasing concentrations of proteinase K (PK) efficiently suppressed the shift of footprints from monosome to disome. This was apparent from a general reduction of the disome enrichment (Fig. 2A) and a flattening of enrichment profiles at the metagene level (Fig. 2B) and for individual genes (Fig. 2C and fig. S1, D to G). Thus, the stability of DiSP-detected disomes critically depends on the integrity of nascent chains, in agreement with the model of co-co assembly.

A high-confidence list of co-co assembly candidates enriched for homomers

We developed an unbiased bioinformatics selection regime to classify proteins on the basis of their proficiency to co-co assemble. Accordingly, a protein qualified as a high-confidence candidate if all of the following criteria were fulfilled: (i) The gene's enrichment profile had a sigmoidal shape, indicating that with progressing translation, ribosomes shifted from the monosome to the disome fraction. If one of the interacting ribosomes terminates earlier, the other ribosome in the pair will shift back to the monosome fraction before it reaches the end of the coding sequence, resulting in a double-sigmoidal shift (Fig. 3A). (ii) The enrichment profile becomes less sigmoidal upon treatment of the lysate with puromycin and (iii) similarly with PK. (iv) The mature protein localizes to either the cytoplasm or the nucleus. We decided to categorize translocated proteins as low-confidence candidates because we cannot formally exclude the possibility that these ribosomes interact with membrane components of the translocation machinery and therefore migrate in the disome fraction. In addition, our validation experiments focused on cytosolic and nuclear candidates (fig. S4), and poor structural annotation of membrane proteins complicates the downstream bioinformatics analysis. Out of 15,898 detected genes, 829 fulfilled all criteria and were classified as high-confidence co-co assembly candidates (table S1). A large number of genes (3301) fulfilled the important criterion (i) but not all of criteria [(ii) to (iv)] and were therefore categorized as low-confidence candidates (table S1). The low-confidence list included 1404 proteins that are translocated across or inserted into organelle membranes [mainly the endoplasmic reticulum (ER)]; of these, 443 fulfilled all other criteria. The latter fraction reflects the general frequency of ER-translocated proteins in the human proteome and indicates that co-co assembly may be an equally important mechanism for assembly of cytosolic or nuclear and ER complexes, in agreement with previous experimental indications (*17*–*19*). The disome shift of ribosomes that synthesize membrane proteins frequently occurs after exposure of the first transmembrane domain (TMD) (fig. S2A), which may suggest that co-co assembly involves interactions of two TMDs in the ER membrane.

Our next aim was to quantitatively assess what fraction of each high-confidence candidate assembles cotranslationally (hereafter termed "efficiency" of co-co assembly). The efficiency was estimated by determining the reduction of footprints in the monosome fraction after initiation of co-co assembly relative to those in the total translatome [including all translating ribosomes, determined by classical ribosome profiling (*15*, *20*)]. Metagene analyses of footprint densities of all high-confidence genes aligned to the onset of assembly revealed a reduction of footprints in the monosome fraction from a DiSP experiment but not in the total translatome (Fig. 3B, top). This result confirmed that the monosome depletion was caused by a shift of ribosomes to the disome fraction. The median monosome footprint reduction after the detected co-co assembly onset of high-confidence genes was ~40%, and for some genes even exceeded 90%, indicating that, in many cases, most nascent chains assembled cotranslationally (Fig. 3B, bottom). Monosome depletion was also observed (to a smaller extent) for many low-confidence candidates,

suggesting that this list includes additional proteins that employ co-co assembly as a main route for complex formation (fig. S2, B and C). Notably, the calculated depletion value most likely underestimates the in vivo co-co assembly efficiency because of (i) the inevitable slight cross-contamination between the monosome and disome fractions and (ii) the possibility of a partial loss of disomes, which are connected by comparably weak nascent chain interactions, during sucrose gradient centrifugation. Supporting this notion, the three proteins with the highest efficiency (≥90% depletion; namely, TPR, EEA1, and CLIP1) contained extremely long coiled-coil homodimerization

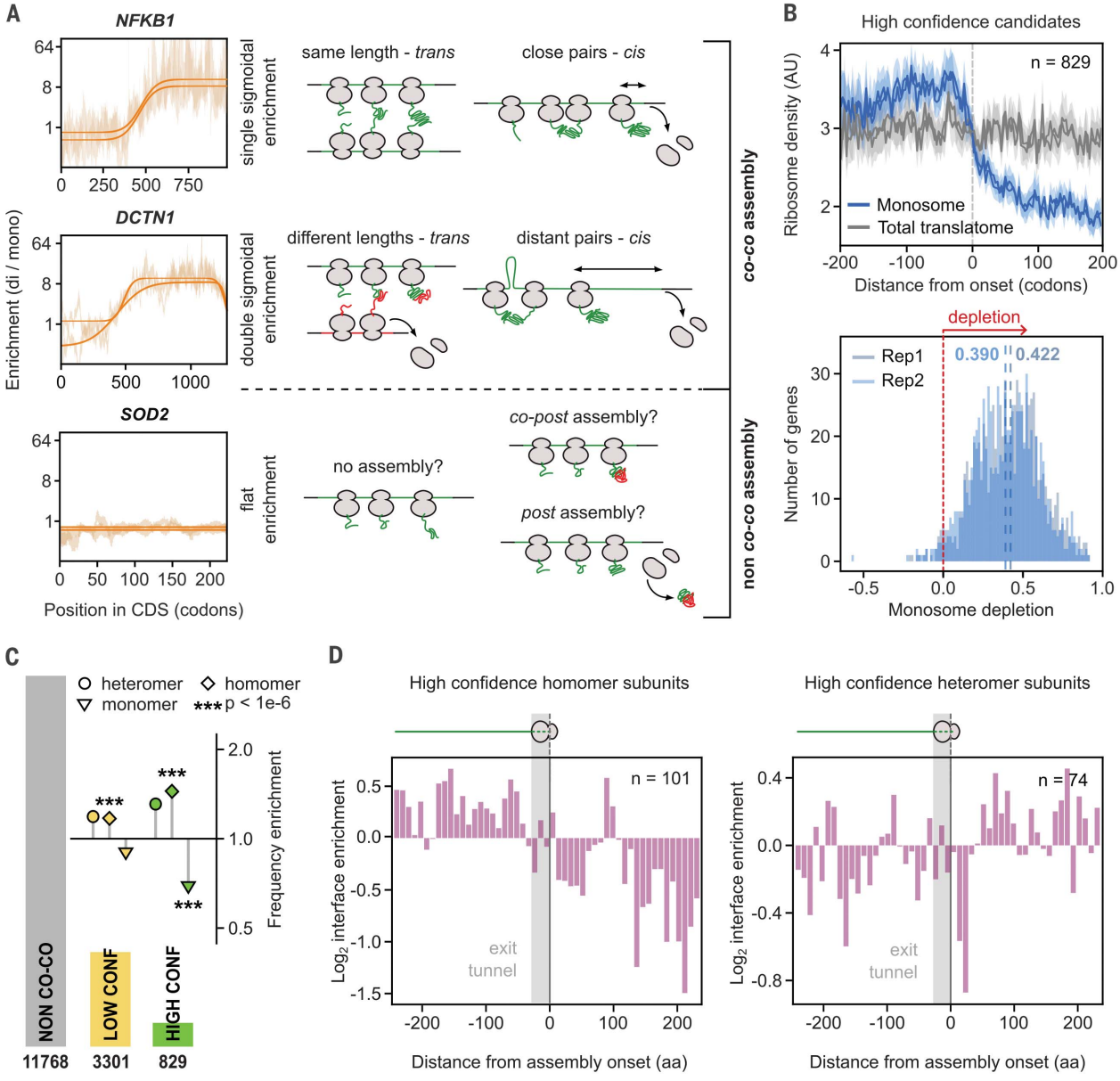


Fig. 3. High-confidence co-co assembly proteins are enriched in homo-oligomers. (**A**) Examples of gene-specific disome-over-monosome enrichment profiles (DiSP data, in the background; two biological replicates) and the corresponding fitting (solid lines) for each of the three possible shapes of DiSP enrichments. The single sigmoid is consistent with nascent chain-connected ribosomes that terminate translation simultaneously, either by co-co assembly in trans (if the mRNA segments translated by both ribosomes after co-co assembly have similar lengths) or in cis (with ribosomes that closely follow each other on the same mRNA) (top). The double sigmoid is consistent with co-co assembly involving two ribosomes that do not terminate at the same time; this may occur in trans (if the mRNA segments translated by both ribosomes after co-co assembly have different lengths) or in cis (if the leading ribosome is distant from the trailing one) (middle). Flat enrichment profiles indicate that nascent proteins do not co-co assemble. (**B**) (Top) Metagene profiles of all high-confidence candidates aligned to assembly onset. Footprint density in the monosome fraction and the total translatome are shown (two biological replicates). (Bottom) Gene-specific quantification of the efficiency of co-co assembly, calculated as the relative depletion of footprint density in the monosome relative to the total translatome after assembly onset. The median monosome depletion for each replicate is indicated by blue dashed lines. AU, arbitrary units. (**C**) Frequency enrichment of annotated subunits of protein complexes in high- and low-confidence lists relative to the whole proteome (absolute and relative numbers are provided in table S2) (*31*). The number of genes included in each assembly class is indicated in the bar plot. *P* values were calculated using an enrichment test adjusted for expression bias (*31*, *32*). (**D**) Distribution of residues forming the intersubunit interface of protein complexes determined from available crystal structures. The position of interface residues on the proteins' primary sequence is aligned to assembly onset of high-confidence homomers (left) or heteromers (right). aa, amino acids.

domains (between 1000 and 1500 amino acids, compared with a median coiled-coil length of 66 amino acids in the cellular proteome), suggesting high stability.

We went on to analyze the features of proteins included in the high- and low-confidence lists. Consistently, annotated monomeric proteins were depleted in both lists of co-co assembly candidates, most extensively among the high-confidence proteins (Fig. 3C and table S2). Both classes showed a significant enrichment of homomers, but heteromers were not significantly enriched. As our statistical analysis accounts for differences in expression levels in our datasets and the annotation database, the heteromer enrichment in the low-confidence class is statistically not significant, although it slightly exceeds the homomer enrichment. Furthermore, we often found only one subunit of a heterodimer in our candidate list, which suggests that this subunit formed a homo-oligomer or co-co assembled with an as-yet-unknown partner subunit.

We used available crystal structures of protein complexes to determine the position of residues involved in subunit interaction at the onset of the disome shift. This analysis showed that the onset of assembly often coincided with the emergence of nascent chain segments that form the interfaces for the homo-oligomers (Fig. 3D, left). This correlation was not detected for heteromeric high-confidence candidates (Fig. 3D, right). Although these findings do not exclude the possibility that individual heteromers co-co assemble, as previously reported (*13*, *14*, *19*), they suggest that co-co assembly is predominantly employed for the formation of homomeric protein complexes.

Co-co assembly is driven by exposure of conserved N-terminal homodimerization domains

Most detected co-co assembly interactions were established at early translation stages (fig. S3A). Consistently, homodimerization interfaces are enriched in the N-terminal halves of high-confidence candidates (fig. S3B, left). This is different in the majority of the human proteome, where homodimerization interfaces are more often located in the C-terminal half of the protein, as previously reported (*21*) (fig. S3B, right).

We next aimed to identify protein motifs or folds that mediate co-co assembly, by studying the enrichment of exposed domains at the onset of assembly. This analysis identified seven domain clusters mediating co-co assembly (color coded in Fig. 4A), of which five are established homodimerization units.

Among our high-confidence candidates, coiled coils were the most prevalent annotated domain class that is exposed on the ribosome surface at assembly onset (193 of 829 proteins according to UniprotKB; Fig. 4B, left). Furthermore, the DeepCoil prediction tool (*22*) identified coiled-coil segments on the exposed nascent chains in 408 genes (fig. S3C), suggesting that up to 50% of high-confidence candidates employ this fold for co-co assembly. In many cases, the coiled coil is only partially exposed at assembly onset (Fig. 4B, left). The number of exposed residues involved in coiled-coil formation varied (median of 111 residues in the high-confidence class; fig. S3D), which may indicate that different lengths of the coiled coil are needed to form a stable dimer.

We found seven additional domains that were generally positioned N-terminally to coiled-coil domains in myosins, kinesins, and AGC kinases (orange in Fig. 4A) and were therefore exposed at the onset of co-co assembly. However, disome enrichment generally required the partial or complete exposure of the coiled-coil segment, suggesting that these domains do not contribute to oligomerization.

A second domain class that was often only partially exposed at the onset of assembly is BAR domains (named after Bin, amphiphysin, and Rvs proteins; Fig. 4B, right). These conserved dimerization domains are found in many proteins that mediate membrane curvature. They consist of three (classical BAR) to five (F-BAR) bent antiparallel α helices. According to our dataset, co-co assembly generally required the exposure of the most N-terminal α helix (helix1; Fig. 4B, right), which interacts with its partner (helix1′) in an antiparallel fashion.

All other enriched domain classes—including BTB (Broad-Complex, Tramtrack, and Bric a brac), RHD (Rel homology domain), and SCAN (SRE-ZBP, CTfin51, AW-1, and Number 18 cDNA) domains (Fig. 4C)—were globular and fully exposed at assembly onset, implying that their cotranslational folding was required for assembly. BTBs are highly conserved globular dimerization domains located at the N termini of many transcription factors, ion channels, and E3 ligase subunits, and were found in 36 of our high-confidence candidates (Fig. 4C, left). The less abundant RHDs are found at the N terminus of proteins involved in nuclear factor κB (NF-κB) complex formation and create the interface of homo- and heteromeric interactions. According to our DiSP, all NF-κB homologs co-co assemble, confirming earlier indications that proteins encoded by *NFKB1* may cotranslationally assemble in cis and that early assembly is required for native biogenesis of the p50 transcription factor (*12*, *23*) (Fig. 4C, middle, and fig. S1B, right). This notion very likely also holds true for the *RELB*-encoded homolog; however, because *RELB* is poorly expressed in HEK293-T cells, we cannot make a definite statement.

The high-confidence list also included 12 transcription factors that employ SCAN domains for co-co assembly (Fig. 4C, right). SCAN domains are leucine-rich, N-terminal motifs composed of five packed α helices that mediate homo- and hetero-oligomerization of a large family of C2H2 zinc finger proteins by intercalating helix 2 of one monomer between helices 3 and 5 of the opposing monomer.

By comparing the co-co assembly efficiency of these five major dimerization domains, we found that coiled coils conferred the highest (yet very variable) stability to the nascent chain interactions, followed by BTB, BAR, RHD, and SCAN domains (fig. S3E).

Finally, our dataset included two less characterized domains that were significantly enriched (Fig. 4A). The first are STI1 repeats of ubiquilin proteins. This domain mediates homo- and heterodimerization of ubiquilin 1 and 2 (*24*), both of which were high-confidence candidates that fully exposed the second STI1 repeat (STI1 2) at the assembly onset (fig. S3F). The second, GBD/FH3, are conserved N-terminal regulatory elements in diaphanous-related formins, a protein class involved in nucleation and remodeling of the actin cytoskeleton. The FH3 domain has been implicated in dimerization of the mouse homolog of human *DIAPH1* (*25*). We found six human formins among our high-confidence proteins; in all cases, the FH3 domain was exposed at assembly onset, suggesting that formins may cotranslationally assemble via the FH3 domain (fig. S3G).

Co-co assembly is independent of eukaryotic assembly factors

We next examined whether ribosome exposure of co-co assembly–competent nascent chains suffices for disome formation, and whether it could occur outside the eukaryotic folding environment. To investigate these questions, we performed DiSP of *E. coli* that synthesize human lamin C (*LMNA*), one of the mammalian intermediate filaments that were all high-confidence candidates of our DiSP screening. Lamins form homodimers in the cytosol and assemble into higher-order polymers in the nucleus. Dimerization involves the N-terminal rod domain, a long, discontinuous coiled coil that includes three segments (coils 1A, 1B, and 2AB). *LMNA* overexpression generated a disome peak in the RNase-digested lysate (Fig. 5A). DiSP revealed that these disomes were enriched with ribosomes that translate *LMNA* (Fig. 5B), indicating that nascent lamin C can cotranslationally dimerize in bacteria. The minimal length of nascent lamin C mediating the disome shift in *E. coli* was close to that of the endogenously expressed lamin C in mammalian cells (Fig. 5B). Likewise, overexpression of *DCTN1* generated a disome peak that was enriched with ribosomes exposing the coiled coil of $p150^{glued}$, and the assembly onset was similar to that in human cells (fig. S4A). This observation indicates that co-co assembly of coiled coils is independent of eukaryote-specific assembly factors or mRNA subcellular localization.

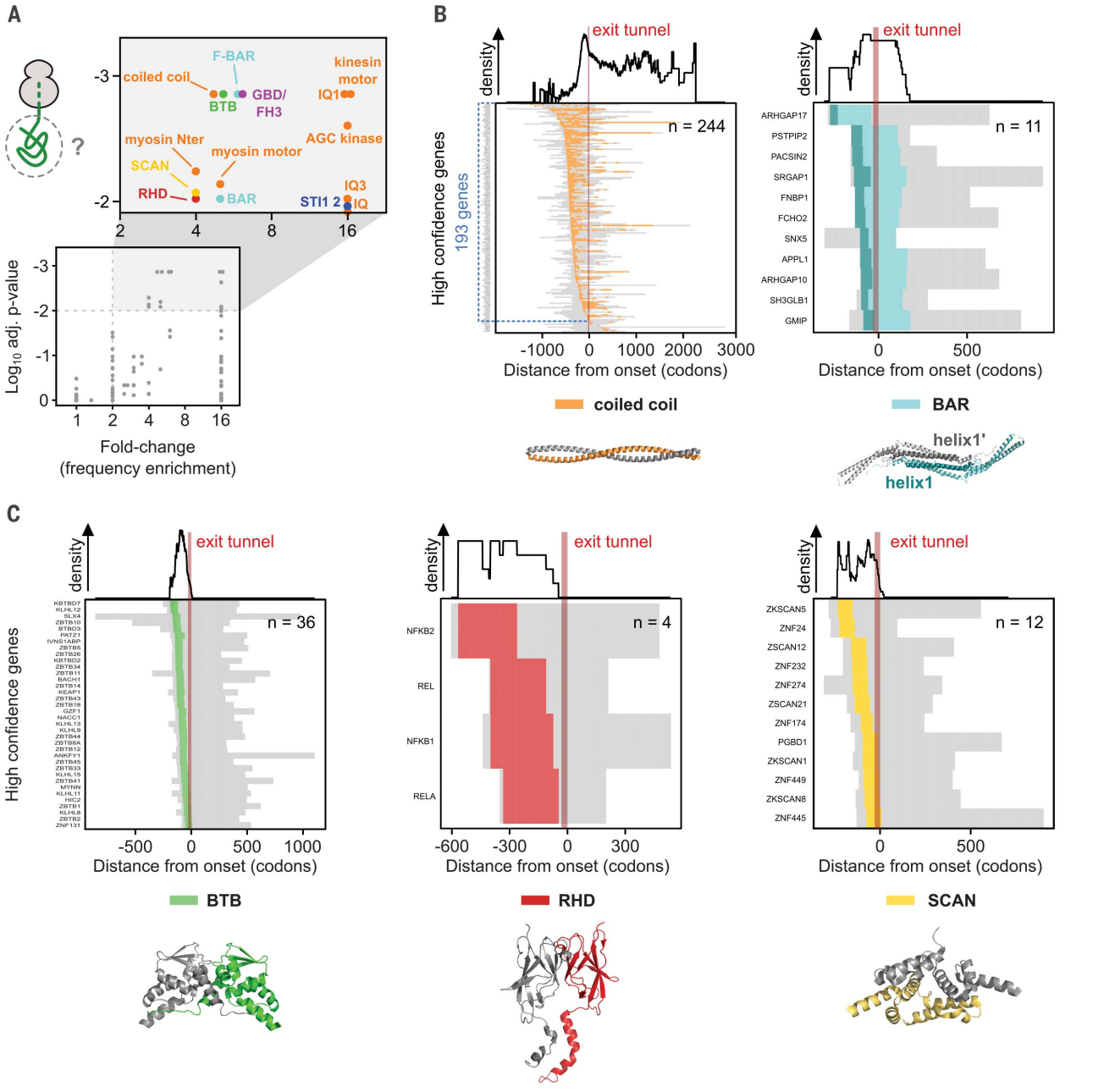


Fig. 4. Co-co assembly is coordinated with exposure of five major dimerization domain classes. (**A**) Analysis of protein domains on nascent chain segments exposed at assembly onset. The frequency of each domain in the high-confidence class is compared with their general frequency in the proteome (*31*). We used a Monte Carlo simulation of the null hypothesis to calculate the *P* value (*31*) and the Benjamini-Yekutieli procedure to correct for multiple testing. The adjusted *P* value is plotted against the respective fold change (frequency enrichment). Domains passing a significance (adjusted $P \geq 0.01$) and fold change (≥2) threshold are shown in the magnified rectangle and further analyzed. (**B**) Heatmaps of partially exposed domains: coiled coil (left) and BAR (right). In the heatmaps, nascent chain segments on the left side of the indicated ribosome exit tunnel (~30 codons, shown by a red bar) are exposed when assembly starts. The subset of genes exposing a coiled-coil segment on the nascent chain at the onset of assembly is highlighted in blue ($n = 193$). Residues forming helix1 of BAR domains are colored dark green in the heatmap and in the exemplary structure. Corresponding domain density profiles are shown atop the heatmaps. Protein Data Bank (PDB) IDs for representative structures: 1D7M (coiled coil) and 3Q0K (BAR). (**C**) Heatmaps of completely exposed domains: BTB (left), RHD (middle) and SCAN (right). Corresponding domain density profiles are shown atop the heatmaps. PDB IDs for representative structures: 1BUO (BTB), 1K3Z (RHD), and 3LHR (SCAN).

To test our hypothesis that the formation of a coiled coil between two nascent chains is minimally required and sufficient to induce disome shifts in bacteria, we used coil 1B of lamin C as a paradigm. First, we employed an established in vivo dimerization assay based on a λ repressor fusion system (*26*) to show that the isolated 1B efficiently dimerized in *E. coli* (fig. S4B). Second, we performed DiSP to verify that nascent 1B, N-terminally fused to mCherry, efficiently mediated co-co assembly (Fig. 5C, left). Third, we perturbed the periodicity of nonpolar and charged amino acids required for coiled-coil formation of 1B by swapping positions "a" and "e" of the coiled-coil heptameric repeats (1B*; Fig. 5C, middle). These swaps do not change the overall amino acid composition, the hydrophobicity, or the

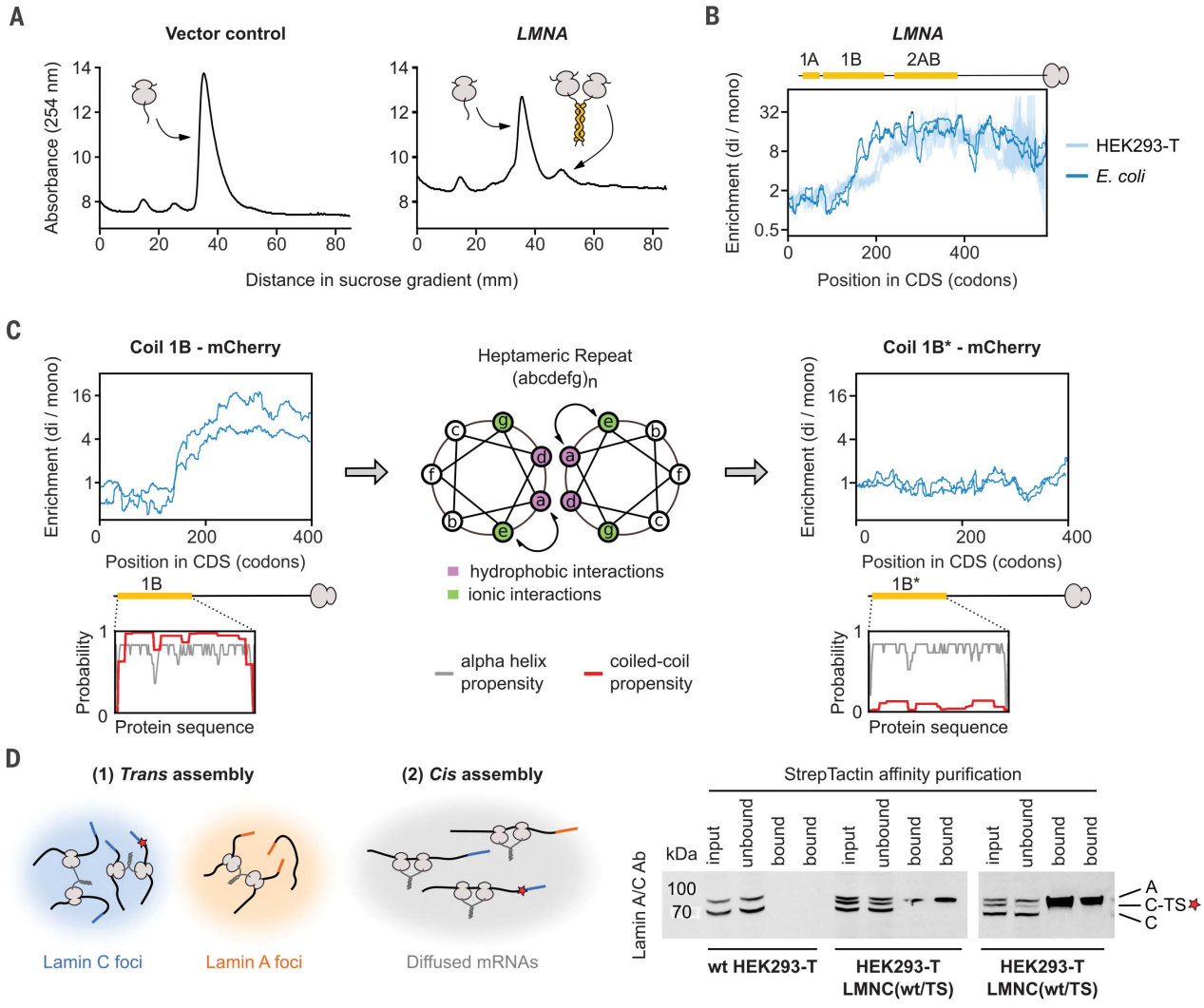


Fig. 5. Co-co assembly does not rely on eukaryote-specific factors and facilitates native biogenesis of lamin C homodimers. (**A**) Sucrose gradient sedimentation analysis of *E. coli* ribosomes from cells transformed with a control plasmid (left) or a plasmid that encloses human *LMNA* encoding lamin C (right), lacking the unstructured N-terminal head domain (*31*). (**B**) Disome-over-monosome enrichment profile of plasmid-encoded *LMNA* expressed in *E. coli* (dark blue; two biological replicates), and endogenously expressed *LMNA* in HEK293-T cells (light blue; two biological replicates). The ribosome-exposed coiled-coil interfaces are indicated by yellow bars. (**C**) Disome-over-monosome enrichment profiles of *LMNA* encoding lamin coil 1B (left) or the version of 1B with positions a and e swapped (1B*; right) fused N terminally to mCherry and expressed in *E. coli* (two biological replicates). The ribosome-exposed coiled-coil interfaces are indicated by yellow bars. A helical wheel projection shows residue arrangements (a to g) of the heptad repeat (middle). Coiled-coil (red) and α-helical (gray) probability predictions are shown for both wild-type and mutant 1B (insets). (**D**) (Left) Hypothetical models of co-co assembly supporting isoform-specific homodimerization. A red star represents the TwinStrep tag (TS). (Right) Affinity purification of tagged lamin C (C-TS) from wild-type (wt) or heterozygous [*LMNC*(wt/TS)] HEK293-T cells (technical replicates shown). Bands are labeled: A (lamin A), C (lamin C), and C-TS (lamin C–TwinStrep).

predicted propensity to form α helices, but they do eliminate the proficiency of 1B to form a coiled coil (Fig. 5C, insets). In contrast to 1B, the mutated 1B* did not confer cotranslational disome formation in *E. coli* (Fig. 5C, right), further indicating that DiSP detects productive, in vivo interactions between nascent chains that drive protein oligomer formation.

Co-co assembly in cis may ensure isoform-specific coiled-coil formation

Lamins A and C are isoforms encoded by the same gene but translated on two alternatively spliced transcripts. Although they share the same N-terminal rod dimerization domain, lamins A and C exclusively form homodimers in vivo (*27*). How this isoform specificity is achieved in the cellular environment is not known. Co-co assembly may provide a simple answer to this conundrum: Isoform-specific assembly may be achieved by co-co assembly in trans on colocalized mRNAs of the same type [which might segregate in the cytosol, owing to their distinct 3′ untranslated regions (UTRs)], or in cis, facilitated by interaction of nascent proteins synthesized by neighboring ribosomes organized in a polysome (Fig. 5D, left).

To distinguish between these possibilities, we generated a heterozygous HEK293-T cell line, in which one *LMNA* allele encodes a C-terminally TwinStrep-tagged lamin C. We performed a series of affinity purification experiments, which revealed that tagged lamin C never copurified the untagged counterpart, even though both proteins are derived from identically spliced mRNAs with identical UTRs (Fig. 5D, right). This result supports the model that co-co assembly in cis facilitates

isoform-specific lamin dimerization in human cells.

Discussion

In this paper, we provide a comprehensive analysis of cotranslational protein complex assembly mediated by two nascent subunits. The ribosome profiling–based approach that we developed (DiSP) allowed us to identify hundreds of high-confidence candidates and thousands of low-confidence candidates in human cells, revealing co-co assembly as a major route to complex formation.

We decided to include all translocated proteins in the low-confidence list. Many of them are membrane proteins that are often partially or fully resistant to PK but sensitive to puromycin—in particular, small proteins (up to 35 kDa) with multiple annotated TMDs. PK resistance may be conferred by ribosome docking to the translocon that limits the access of PK to the nascent protein. We speculate that docking of ribosomes that closely follow each other in a polysome may spatially organize translocons in the membrane and facilitate homomer assembly.

Our data show that predominantly homodimers co-co assemble. We did not find clear evidence that heteromers co-co assemble in trans, because our high-confidence list, in most cases, contained only one subunit of an annotated heteromer. The absence of a known partner subunit may be caused by the less complete structural characterization of heteromeric complexes.

We also did not find clear evidence that the recently described assembly of the TAF6-TAF9 nuclear complex includes nascent chain interactions (*14*). Both subunits are included in the low-confidence list, but the length of the disome shift and the enrichment efficiency is very different between the two proteins, which is not consistent with a model of co-co assembly in trans.

Co-co assembly of homomers in cis may be facilitated by a generally high ribosome occupancy to ensure close proximity of the interacting nascent chains. In addition, both heteromer assembly (in trans) and homomer assembly (in cis or in trans) may benefit from the slowdown of ribosomes at the onset of assembly, to allow the trailing ribosome translating the same mRNA to catch up or to provide an extended time frame to establish the interaction with another nascent chain translated on a distinct mRNA (*13*).

We discovered two different types of nascent chain dimerization. The first is a zipper-like formation of coiled coils and BAR domains. For this type, the interaction strength may gradually increase as both nascent chains grow, until enough residues involved in dimerization are ribosome exposed to drive the co-co assembly of stable dimers. The second type of nascent chain dimerization may require the prior folding of a fully emerged, globular interaction domain (a BTB, RHD, or SCAN domain), a feature already reported for co-post assembly (*3*, *8*).

Homodimerization contact regions are evolutionarily selected to be enriched in C-terminal halves of proteins, supposedly to ensure that folding is not disturbed by the proximity of another identical, incompletely folded subunit (*21*). Our analysis supports this C-terminal enrichment for most of the human proteome, except for the proteins in our high-confidence list. For the latter proteins, the selective pressure to assemble early apparently outweighs the risk for misfolding of yet-to-be-synthesized C-terminal domains. We speculate that productive folding of the native dimer, beyond co-co assembly, is likely supported by extensive, finely tuned intervention of molecular chaperones.

Multiple factors may create selective pressure against diffusion-driven assembly and favor co-co assembly: (i) Co-co assembly may increase the efficiency and rate of complex formation. This advantage is most evident for the cis assembly mode in which dimerizing nascent chains are already adjacent within polysomes. (ii) Synthesis-coupled assembly may suppress unproductive interactions and facilitate native folding by limiting the exposure of aggregation-prone dimerization interfaces to the crowded cellular environment. (iii) Cis assembly creates mRNA-specific homomers. Coiled coils and BTB domains are recurrent dimerization modules in the human proteome, with high potential for non-specific, potentially deleterious heteromeric interactions (*28*, *29*). Such interactions—including those among splicing-derived isoforms that share identical dimerization domains, as in the case of human lamin A and C (*27*, *28*)—would be efficiently prevented in cis assembly. Misassembled subunits that failed to co-co assemble may be recognized by a recently described pathway that specifically detects and eliminates complexes of aberrant composition [dimerization quality control (DQC) (*30*)]. Notably, DQC has been reported as a surveillance mechanism for BTB complexes, but a similar molecular machinery that monitors the composition of other complexes, including coiled coils, may exist. Our proteome-wide study reveals that cotranslational interactions between nascent subunits are a general and efficient strategy to guide the isoform-specific formation of protein complexes.

Materials and methods summary

Detailed materials and methods can be found in the supplementary materials.

Human osteosarcoma U2OS (ATCC catalog no. HTB-96), human embryonal kidney HEK293-T (DSMZ catalog no. ACC 635), and *E. coli* Rosetta cells (Novagene) were employed for DiSP experiments.

All ribosome profiling libraries were prepared as described in (*20*) and sequenced on a NextSeq550 (Illumina) according to the manufacturer's protocol, except for libraries of U2OS samples, which were prepared as described in (*8*) and sequenced on a HiSeq 2000 (Illumina).

DiSP with PK treatment included incubation of the cell lysates for 30 min at 4°C with the following ratios of PK to total protein: (i) low PK = 1:20,000; (ii) mid PK = 1:6000; (iii) high PK = 1:2000; and (iv) very high PK = 1:200.

DiSP with puromycin omitted cycloheximide from all buffers; cell lysates were incubated for 25 min with 2 mM of puromycin and cross-linked with 0.5% formaldehyde.

REFERENCES AND NOTES

1. G.-W. Li, D. Burkhardt, C. Gross, J. S. Weissman, Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell* **157**, 624–635 (2014). doi: 10.1016/j.cell.2014.02.033; pmid: 24766808
2. G. Tian *et al*., Tubulin subunits exist in an activated conformational state generated and maintained by protein cofactors. *J. Cell Biol.* **138**, 821–832 (1997). doi: 10.1083/jcb.138.4.821; pmid: 9265649
3. Y.-W. Shieh *et al*., Operon structure and cotranslational subunit association direct protein assembly in bacteria. *Science* **350**, 678–680 (2015). doi: 10.1126/science.aac8171; pmid: 26405228
4. A. Rousseau, A. Bertolotti, Regulation of proteasome assembly and activity in health and disease. *Nat. Rev. Mol. Cell Biol.* **19**, 697–712 (2018). doi: 10.1038/s41580-018-0040-z; pmid: 30065390
5. L. A. Mingle *et al*., Localization of all seven messenger RNAs for the actin-polymerization nucleator Arp2/3 complex in the protrusions of fibroblasts. *J. Cell Sci.* **118**, 2425–2433 (2005). doi: 10.1242/jcs.02371; pmid: 15923655
6. M. Pizzinga *et al*., Translation factor mRNA granules direct protein synthetic capacity to regions of polarized growth. *J. Cell Biol.* **218**, 1564–1581 (2019). doi: 10.1083/jcb.201704019; pmid: 30877141
7. B. Hampoelz *et al*., Nuclear Pores Assemble from Nucleoporin Condensates During Oogenesis. *Cell* **179**, 671–686.e17 (2019). doi: 10.1016/j.cell.2019.09.022; pmid: 31626769
8. A. Shiber *et al*., Cotranslational assembly of protein complexes in eukaryotes revealed by ribosome profiling. *Nature* **561**, 268–272 (2018). doi: 10.1038/s41586-018-0462-y; pmid: 30158700
9. G. Kramer, A. Shiber, B. Bukau, Mechanisms of Cotranslational Maturation of Newly Synthesized Proteins. *Annu. Rev. Biochem.* **88**, 337–364 (2019). doi: 10.1146/annurev-biochem-013118-111717; pmid: 30508494
10. C. D. Nicholls, K. G. McLure, M. A. Shields, P. W. K. Lee, Biogenesis of p53 involves cotranslational dimerization of monomers and posttranslational dimerization of dimers. Implications on the dominant negative effect. *J. Biol. Chem.* **277**, 12937–12945 (2002). doi: 10.1074/jbc.M108815200; pmid: 11805092
11. R. Gilmore, M. C. Coffey, G. Leone, K. McLure, P. W. Lee, Co-translational trimerization of the reovirus cell attachment protein. *EMBO J.* **15**, 2651–2658 (1996). doi: 10.1002/j.1460-2075.1996.tb00625.x; pmid: 8654362
12. L. Lin, G. N. DeMartino, W. C. Greene, Cotranslational dimerization of the Rel homology domain of NF-kappaB1 generates p50-p105 heterodimers and is required for effective p50 production. *EMBO J.* **19**, 4712–4722 (2000). doi: 10.1093/emboj/19.17.4712; pmid: 10970863
13. O. O. Panasenko *et al*., Co-translational assembly of proteasome subunits in NOT1-containing assemblysomes. *Nat. Struct. Mol. Biol.* **26**, 110–120 (2019). doi: 10.1038/s41594-018-0179-5; pmid: 30692646
14. I. Kamenova *et al*., Co-translational assembly of mammalian nuclear multisubunit complexes. *Nat. Commun.* **10**, 1740 (2019). doi: 10.1038/s41467-019-09749-y; pmid: 30988355
15. N. T. Ingolia, S. Ghaemmaghami, J. R. S. Newman, J. S. Weissman, Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* **324**, 218–223 (2009). doi: 10.1126/science.1168978; pmid: 19213877
16. P. Han *et al*., Genome-wide Survey of Ribosome Collision. *Cell Rep.* **31**, 107610 (2020). doi: 10.1016/j.celrep.2020.107610; pmid: 32375038

17. S. D. Redick, J. E. Schwarzbauer, Rapid intracellular assembly of tenascin hexabrachions suggests a novel cotranslational process. *J. Cell Sci.* **108**, 1761–1769 (1995). pmid: 7542260
18. J. Lu, J. M. Robinson, D. Edwards, C. Deutsch, T1-T1 interactions occur in ER membranes while nascent Kv peptides are still attached to ribosomes. *Biochemistry* **40**, 10934–10946 (2001). doi: 10.1021/bi010763e; pmid: 11551188
19. 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). doi: 10.1073/pnas.1521577113; pmid: 27078096
20. N. J. McGlincy, N. T. Ingolia, Transcriptome-wide measurement of translation by ribosome profiling. *Methods* **126**, 112–129 (2017). doi: 10.1016/j.ymeth.2017.05.028; pmid: 28579404
21. E. Natan *et al*., Cotranslational protein assembly imposes evolutionary constraints on homomeric proteins. *Nat. Struct. Mol. Biol.* **25**, 279–288 (2018). doi: 10.1038/s41594-018-0029-5; pmid: 29434345
22. J. Ludwiczak, A. Winski, K. Szczepaniak, V. Alva, S. Dunin-Horkawicz, DeepCoil-a fast and accurate prediction of coiled-coil domains in protein sequences. *Bioinformatics* **35**, 2790–2795 (2019). doi: 10.1093/bioinformatics/bty1062; pmid: 30601942
23. L. Lin, G. N. DeMartino, W. C. Greene, Cotranslational biogenesis of NF-kappaB p50 by the 26S proteasome. *Cell* **92**, 819–828 (1998). doi: 10.1016/S0092-8674(00)81409-9; pmid: 9529257
24. D. L. Ford, M. J. Monteiro, Dimerization of ubiquilin is dependent upon the central region of the protein: Evidence that the monomer, but not the dimer, is involved in binding presenilins. *Biochem. J.* **399**, 397–404 (2006). doi: 10.1042/BJ20060441; pmid: 16813565
25. R. Rose *et al*., Structural and mechanistic insights into the interaction between Rho and mammalian Dia. *Nature* **435**, 513–518 (2005). doi: 10.1038/nature03604; pmid: 15864301
26. J. C. Hu, E. K. O'Shea, P. S. Kim, R. T. Sauer, Sequence requirements for coiled-coils: Analysis with λ repressor-GCN4 leucine zipper fusions. *Science* **250**, 1400–1403 (1990). doi: 10.1126/science.2147779; pmid: 2147779
27. T. Kolb, K. Maass, M. Hergt, U. Aebi, H. Herrmann, Lamin A and lamin C form homodimers and coexist in higher complex forms both in the nucleoplasmic fraction and in the lamina of cultured human cells. *Nucleus* **2**, 425–433 (2011). doi: 10.4161/nucl.2.5.17765; pmid: 22033280
28. Q. Ye, H. J. Worman, Protein-protein interactions between human nuclear lamins expressed in yeast. *Exp. Cell Res.* **219**, 292–298 (1995). doi: 10.1006/excr.1995.1230; pmid: 7628545
29. G. Schreiber, A. E. Keating, Protein binding specificity versus promiscuity. *Curr. Opin. Struct. Biol.* **21**, 50–61 (2011). doi: 10.1016/j.sbi.2010.10.002; pmid: 21071205
30. E. L. Mena *et al*., Dimerization quality control ensures neuronal development and survival. *Science* **362**, eaap8236 (2018). doi: 10.1126/science.aap8236; pmid: 30190310
31. Materials and methods are available as supplementary materials.
32. M. D. Young, M. J. Wakefield, G. K. Smyth, A. Oshlack, Gene ontology analysis for RNA-seq: Accounting for selection bias. *Genome Biol.* **11**, R14 (2010). doi: 10.1186/gb-2010-11-2-r14; pmid: 20132535
33. ilia-kats, ilia-kats/RiboSeqTools: v0.1, Zenodo (2020); doi: 10.5281/ZENODO.4016066.

ACKNOWLEDGMENTS

We thank all members of B.B.'s laboratory for discussions and advice; D. Coombs for help with optimization of ribosome separation on sucrose gradients; U. Friedrich for help with establishing pipelines for processing ribosome profiling data; S. Anders for valuable advice concerning development of DiSP data analysis tools; the ZMBH Flow Cytometry & FACS Core Facility, the DKFZ Sequencing Core Facility, and the DKFZ Vector and Clone Repository for support of experimental work. M.B., K.F., and J.S. are members of the Heidelberg Biosciences International Graduate School (HBIGS). **Funding:** M.B. and K.F. were supported by a HBIGS Ph.D. fellowship. M.B. was additionally supported by a Boehringer Ingelheim Fonds (BIF) Ph.D. fellowship. F.W. received funding from the European Union's Horizon 2020 research and innovation program under Marie Skłodowska-Curie grant 745798. This work was supported by the Helmholtz-Gemeinschaft [DKFZ NCT3.0 Integrative Project in Cancer Research (DysregPT_Bukau 1030000008 G783)], the Deutsche Forschungsgemeinschaft (SFB 1036), the European Research Council [ERC Advanced Grant (743118)], and the Klaus Tschira Foundation. Work in the Tans laboratory was supported by the Netherlands Organization for Scientific Research (NWO). **Authors contributions:** Conceptualization: M.B., K.F., F.W., J.S., S.T., B.B., and G.K. Methodology: M.B., K.F., J.J.A., B.B., and G.K. Investigation: M.B. and K.F. Software: I.K., F.T., M.B., and K.F. Formal analysis, Data curation, and Visualization: M.B., K.F., I.K., F.T., B.B., and G.K. Writing – original draft: M.B., K.F., I.K., and G.K. Writing – review & editing: all authors. Supervision: S.T., B.B., and G.K. **Competing interests:** All authors declare no competing interests. **Data and materials availability:** All sequencing data reported in this study are available at GEO under accession number GSE151959. Explicit Julia code is available as supplementary material; explicit R code will be made available upon request. Data analysis of ribosome profiling datasets was performed with RiboSeqTools [available at https://github.com/ilia-kats/RiboSeqTools and Zenodo (*33*)].

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/371/6524/57/suppl/DC1
Materials and Methods
Figs. S1 to S4
Tables S1 to S5
References (*34*–*56*)
MDAR Reproducibility Checklist
Custom Julia Scripts 1 to 3

View/request a protocol for this paper from *Bio-protocol*.

16 June 2020; accepted 27 October 2020
10.1126/science.abc7151

Interactions between nascent proteins translated by adjacent ribosomes drive homomer assembly

Matilde Bertolini, Kai Fenzl, Ilia Kats, Florian Wruck, Frank Tippmann, Jaro Schmitt, Josef Johannes Auburger, Sander Tans, Bernd Bukau and Günter Kramer

Science **371** (6524), 57-64.
DOI: 10.1126/science.abc7151

Co-co assembly for oligomers

Most of the human proteome forms oligomeric protein complexes, but how they assemble is poorly understood. Bertolini *et al.* used a ribosome-profiling approach to explore the existence of a cotranslational assembly mode based on the interaction of two nascent polypeptides, which they call the "co-co" assembly. Proteome-wide data were used to show whether, when, and how efficiently nascent complex subunits interact. The findings also show that human cells use co-co assembly to produce hundreds of different homo-oligomers. Co-co assembly involving ribosomes translating one messenger RNA may resolve the longstanding question of how cells prevent unwanted interactions between different protein isoforms to efficiently produce functional homo-oligomers.

Science, this issue p. 57