



HAL
open science

Co-translational assembly pathways of nuclear multiprotein complexes involved in the regulation of gene transcription

Andrea Bernardini, László Tora

► **To cite this version:**

Andrea Bernardini, László Tora. Co-translational assembly pathways of nuclear multiprotein complexes involved in the regulation of gene transcription. *Journal of Molecular Biology*, 2024, 436 (4), pp.168382. 10.1016/j.jmb.2023.168382 . hal-04326481

HAL Id: hal-04326481

<https://hal.science/hal-04326481v1>

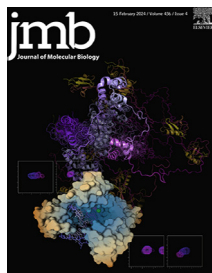
Submitted on 21 Aug 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License



Co-translational Assembly Pathways of Nuclear Multiprotein Complexes Involved in the Regulation of Gene Transcription

Andrea Bernardini* and László Tora*

Institut de Génétique et de Biologie Moléculaire et Cellulaire, Illkirch, France

Centre National de la Recherche Scientifique, UMR7104, Illkirch, France

Institut National de la Santé et de la Recherche Médicale, U964 Illkirch, France

Université de Strasbourg, Illkirch, France

Correspondence to Andrea Bernardini and László Tora: Institut de Génétique et de Biologie Moléculaire et Cellulaire, Illkirch, France. andrea.bernardini@igbmc.fr (A. Bernardini), laszlo@igbmc.fr (L. Tora)

<https://doi.org/10.1016/j.jmb.2023.168382>

Edited by Martine Collart

Abstract

Most factors that regulate gene transcription in eukaryotic cells are multimeric, often large, protein complexes. The understanding of the biogenesis pathways of such large and heterogeneous protein assemblies, as well as the dimerization partner choice among transcription factors, is crucial to interpret and control gene expression programs and consequent cell fate decisions. Co-translational assembly (Co-TA) is thought to play key roles in the biogenesis of protein complexes by directing complex formation during protein synthesis. In this review we discuss the principles of Co-TA with a special focus for the assembly of transcription regulatory complexes. We outline the expected molecular advantages of establishing co-translational interactions, pointing at the available, or missing, evidence for each of them. We hypothesize different molecular mechanisms based on Co-TA to explain the allocation “dilemma” of paralog proteins and subunits shared by different transcription complexes. By taking as a paradigm the different assembly pathways employed by three related transcription regulatory complexes (TFIID, SAGA and ATAC), we discuss alternative Co-TA strategies for nuclear multiprotein complexes and the widespread – yet specific – use of Co-TA for the formation of nuclear complexes involved in gene transcription. Ultimately, we outlined a series of open questions which demand well-defined lines of research to investigate the principles of gene regulation that rely on the coordinated assembly of protein complexes.

© 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Most components of the transcription machinery and co-factors are large heteromeric protein complexes

During the study of gene transcription, it quickly became evident that many components of the system were objects made of a large number of protein subunits, contributing to the inherent complexity of transcriptional regulation in eukaryotes. RNA polymerase II (Pol II) transcription initiation and its regulation rely on the

activity of an extremely diversified set of proteins, which can be classified in three broad functional groups¹: sequence-specific transcription factors (TFs), the basal (or general) transcription factors and transcription co-factors. Sequence specific TFs recognize DNA/chromatin-embedded sequence elements in regulatory regions at promoters and enhancers, instructing the other two groups of proteins on “where” and “when” to initiate transcription. TFs are a very diverse group of specialized proteins which are usually classified based on their DNA-binding domain (DBD). Importantly, most

of the DBDs in TFs work by combining in multimeric complexes, most often homo- or heterodimers.² The basal transcription machinery includes Pol II and a set of general transcription factors (GTFs) responsible for recruiting, positioning and regulating the activity of Pol II itself at the sites of transcription. Of the canonical Pol II GTFs (TFIID, TFIIA, TFIIB, TFIIF, TFIIE, TFIIH), all except TFIIB are heteromeric protein complexes, with TFIIH and TFIID consisting of 10 and 14 different polypeptides, respectively. Finally, transcription co-factors (co-activators and co-repressors) comprise a very heterogeneous group of proteins and protein complexes that transduce TFs signals to influence the position and the activity of the basal transcription machinery, without directly recognizing specific DNA elements. Co-factors can work as chromatin modifying enzymes, as nucleosome remodeling factors or as dynamic molecular scaffolds that integrate multiple local signaling inputs.^{3–6} As for GTFs, most of the *bona fide* transcription co-factors are large multiprotein complexes, often composed of distinct structural lobes or functional modules.

The biogenesis pathway of multiprotein transcription complexes impacts on gene expression

Compared to the longstanding and crucial efforts to investigate the role of multiprotein complexes in gene transcription, the understanding of their biogenesis has been understudied. Yet, the assembly pathways of multiprotein transcription complexes and their regulation are expected to play a key role in shaping gene expression programs and cell fate.

The subcellular location of complex formation (i.e., where complex assembly starts and where it comes to completion) has not only profound effects on the activities of the holo-complexes themselves, but also on their assembly intermediates. In fact, partially assembled sub-complexes or isolated complex submodules might either carry out specialized functions in the cell and/or have detrimental effects on gene expression when misregulated. The order of addition of subunits along the assembly pathway has been shown to be under evolutionary selection and to dictate the composition of assembly intermediates.⁷ The subcellular compartment of complex assembly defines the initial location of such intermediates.

Given that a single complex often directly regulates hundreds to thousands of genes, misregulation or defects in its assembly pathway can have broad and pleiotropic effects on gene expression programs. Therefore, specific subunit deletions, truncations, mutations, or overexpression contribute to the outcome of assembly defects. Importantly, the same principle also applies to and is key in the interpretation of

experimental perturbations often employed by researchers to assign molecular functions (gene knockout, knockdown or overexpression).

Complex formation is a key regulatory step for multimeric TFs, where the “choice” of the partner may result into a different effector function or in a change of the preferred DNA motif.^{2,8,9} Moreover, the abundance and half-life of unpaired TF molecules is under active control, as it affects the binding equilibrium with the partners and, ultimately, gene regulation.¹⁰ Since many of the transcription-related complexes mediate aberrant gene expression in disease, including neurodevelopmental disorders and cancer,^{11–15} understanding their assembly steps will set a new playground to ultimately rationalize the effects of current drug targets and to develop new classes of molecules that interfere with the assembly process itself, rather than with the holo-complex activity.

Co-translational assembly couples mRNA translation with complex formation

A critical aspect of a multiprotein complex formation is to effectively control the stability of assembly intermediates starting by the fundamental units of assembly: the single subunits. In this regard, a major conceptual bottleneck is the synthesis of subunits characterized by very poor intrinsic solubility in isolation. How do cells cope with the production of poorly soluble and aggregation-prone subunits? Despite the crucial role of molecular chaperones in assisting – or even correcting – protein folding and buffering protein solubility,¹⁶ they do not explain the whole picture. Biochemists have plenty of examples where the expression of single recombinant proteins in heterologous systems results in insoluble products. Yet, these proteins often become soluble if co-expressed with their natural partner subunit in the same system.¹⁷ This partner “chaperoning” effect is evident also from documented cases where the depletion of a protein is matched by the degradation or destabilization of its natural partner subunit in the cell (see below).

In addition to molecular chaperones, that can assist protein folding of the newly synthesized polypeptide as it emerges from the ribosome channel,^{16,18} individual subunits of multiprotein complexes can engage with their nascent partner protein during translation. The formation of stable interactions among partner subunits during protein synthesis is named co-translational assembly, hereafter called Co-TA. Several reviews described extensively the mechanism of Co-TA,¹⁹ its potential advantages^{20–22} and drawbacks,²³ as well as evolutionary considerations.²⁴ In the following sections we will summarize the fundamental concepts of Co-TA, describe different strategies used by cells to assemble transcription-related complexes

making use of Co-TA, and summarize the evidence for its extensive use for the biogenesis of complexes involved in gene regulation.

The alternative routes of co-translational assembly

By the definition of Co-TA (see above), co-translational interaction events can be classified in distinct scenarios (Figure 1). Co-TA events can be either sequential, when the co-translational interaction is established between a mature protein and the nascent polypeptide chain of its partner (Figure 1A, B), or simultaneous, when both engaging subunits are nascent chains (Figure 1C, D).

Sequential Co-TA is also referred to as “Co-Post” assembly, while simultaneous Co-TA as “Co-Co” assembly.¹⁹ Sequential co-translational interactions can be directional, where only one of the two proteins is engaged co-translationally (Figure 1A), or symmetrical, where the event occurs also in the opposite direction (Figure 1B). On another note, simultaneous Co-TA can occur in *cis* between nascent chains emerging from adjacent ribosomes on the same mRNA, leading to the formation of homo-

meric complexes (Figure 1C), or in *trans*, when the two nascent chains belong to different polysomes (Figure 1D). The latter case also imposes spatial proximity of the two mRNAs (or more) involved in simultaneous co-translational interactions. Importantly, cases for each of the above scenarios have been described experimentally and will be commented in the following sections.

The position of interaction domains affects co-translational interactions

A key aspect of Co-TA is the importance of the position of interaction domains along the protein. For a co-translational engagement, the interaction domain of the nascent protein must be exposed out of the ribosome exit tunnel and it is, in general, in the N-terminal half of the nascent protein. On the contrary, an interaction domain located at the very C-terminal end of a protein would not be compatible as Co-TA “acceptor”. However, the C-terminal interaction domain can still undergo sequential Co-TA in the form of a mature protein with the nascent chain of its partner subunit. These principles have been demonstrated for the TAF10/TAF8 pair, two

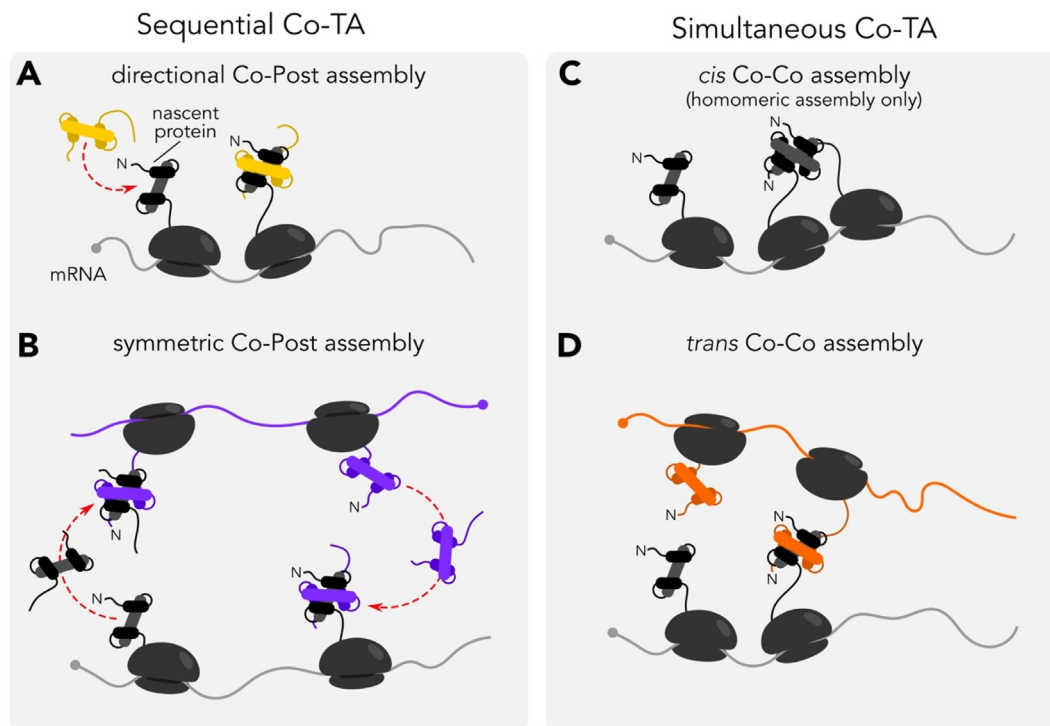


Figure 1. Co-translational assembly classification. Schematic representation of the possible scenarios of co-translational assembly (Co-TA) binding events. The left and right panels summarize sequential and simultaneous co-translational assembly events, respectively. Ribosomes are depicted in black. 5' cap is depicted as a circle on the mRNA. In (B) and (D) the identity of two distinct mRNAs is indicated with different colors. The interaction domain of the protein partners undergoing co-translational assembly is represented by thicker folded cylinders. The N-terminal of the nascent polypeptide chains is indicated. Red dashed arrows indicate the interactions between a fully synthesized protein with the partner's nascent chain (sequential Co-TA).

histone-fold (HF) subunits in TFIID, where the same interaction domain capable of sustaining Co-TA fails to do so if placed at the very C-terminus of the nascent polypeptide.²⁵ Another consequence of Co-TA is that the position of the interaction domain along the protein is expected to influence the time it becomes – and stays – available for a productive co-translational binding event, with the most N-terminal location having the longest dwell time. For the same reason, longer coding sequences (CDS) are expected to increase the dwell time of the corresponding nascent polypeptide on the ribosome, favoring co-translational interactions.

It is even conceivable that when ribosomes synthesize a N-terminal interaction domain and it becomes accessible and correctly folded by coming out from the ribosome exit channel, the ribosome would slow down considerably, waiting for the interaction partner to bind to the exposed interaction domain, similarly to what was proposed for co-translational folding of protein domains.²⁶ Once the interaction partners would have found each other the ribosome could speed up again. The process could be intrinsically mediated either by arrest peptides in the nascent chain,²⁷ or features at the messenger level, like codon usage or RNA secondary structures.²⁸ Future genome-wide ribosome footprinting experiments combined with protein interaction domain mapping and/or with sophisticated single molecule imaging experiments may decide whether such mechanisms would operate in general and be important for co-translational assembly.

The expected (and observed) advantages of co-translational assembly

As mentioned above, many of the Co-TA advantages have been explored and summarized in other reviews.^{20–22} Here we will highlight documented examples where the co-translational pathway was causally linked to a specific outcome in the assembly process. Conversely, we also discuss cases where the association between Co-TA and the outcome of the assembly is – by far – correlative rather than strictly causal.

Partner protein “chaperoning” effect

As mentioned, the synthesis of poorly soluble subunits can be overcome by Co-TA, where the aggregation-prone subunits would engage their partner protein co-translationally, before being released from the ribosome. For example, TAF6 and TAF9, two partner subunits in TFIID that undergo simultaneous Co-TA,²⁵ are insoluble when expressed separately, while they form a highly soluble heterodimer if co-expressed.²⁹ A related, frequently observed, phenomenon is the so-called

“orphan” protein degradation, where the depletion of a protein is paralleled by the destabilization and degradation of its direct partner. This orphan quality control is thought to be mediated by the exposure of interfaces (often hydrophobic) or termini, usually masked in the complex, that work as degrons by being recognized by broadly active cytoplasmic degradation cascades.^{30–32} As an example of transcription complexes, knockout of *TAF10* leads to the depletion of TAF8, its co-translational partner subunit in TFIID.^{25,33} The same behavior seems to be widespread among TFIID subunits also in *Drosophila*, as evidenced by the destabilization of several TAFs upon knockdown of subunits belonging to the same structural module.³⁴ Analogous observations were made for other transcription complexes as well, including the SAGA (Spt Ada Gcn5 acetyltransferase) and TREX-2 (transcription export complex 2) co-activator complexes,³⁵ the MLL lysine methyltransferase³⁶ and the SWI/SNF chromatin remodeling complexes.³⁷ Interestingly, the stability of mammalian SWI/SNF complexes can be regulated by the methylation-dependent proteolysis of the central SMARCC1/SMARCC2 subunits, that if degraded lead to the concurrent destabilization of the rest of the complex.³⁸ Moreover, recent evidence for the mechanism of orphan protein quality control has been gathered for several transcriptional regulators,¹⁰ even suggesting a potential application for therapy.³⁹ This coherent trend was also verified in a broader screening of multiple cytoplasmic complexes, where the subunits engaged in Co-TA as nascent proteins tend to be more susceptible to protein aggregation or degradation in absence of their partners.⁴⁰

The emerging picture is that the co-translational engagement of a nascent chain by its direct protein partner leads to a stabilization effect similar to the well-known role of molecular chaperones. Intriguingly, this co-translational partner “chaperoning” phenomenon might even substitute the function of classical chaperons for certain protein pairs. It is conceivable that chaperons might “accompany” a mature protein on its journey until it assembles with the nascent chain of its partner during translation. For instance, the CCT chaperonin complex was found associated with TAF5 (a core TFIID subunit) in the cytoplasm, where it works as an assembly checkpoint factor: CCT releases TAF5 only when it interacts with its direct partner in the complex, the TAF6/TAF9 heterodimer.⁴¹ We later found that TAF5 undergoes sequential Co-TA with the nascent TAF6 polypeptide, providing evidence for the contribution of molecular chaperons to certain Co-TA events.⁴² Whether molecular chaperons generally participate to direct Co-TA binding events has not been investigated systematically.

Overall, we consider the current evidence of the role of Co-TA in buffering protein solubility by facilitating complex assembly during protein

synthesis rather convincing. Still, we cannot formally exclude the possibility that post-translational assembly would also take place, and that the rescue in solubility/stability would be simply driven by the interaction between the two mature protein partners. In other terms, is Co-TA essential and/or is Co-TA an “obligate” pathway for all proteins, for certain proteins only, or eventually for none of them? Investigating these questions is as important as it could be challenging, since we currently lack the ability to quantitatively and reliably measure and control the occurrence of co-translational *versus* post-translational binding events in the cell.

Overcoming the formation of intricate protein–protein interfaces

Several protein partners are found intimately associated through elaborate structural interfaces, where the two polypeptide chains are deeply intertwined rather than using classical binding pockets, short linear motifs or extended flat surfaces. It is hard to conceive how these complex interfaces might form post-translationally, once the two proteins have already acquired a global fold. The emergence of a nascent polypeptide chain from the ribosome is a gradual phenomenon. As such, a co-translational interaction with the partner protein would allow a step-wise, rather than all-in-one, folding phase. Since translation rate is generally slower than protein folding, the occurrence of Co-TA would effectively “set the pace” for co-folding with an incoming protein partner.⁴³ In addition, since protein translation is directional, the co-folding order of the different structural subdomains with the partner protein is influenced by their position along the primary sequence. The combination of slower and sequential protein folding with the partner protein in Co-TA might be the only path to build very intricate protein–protein interfaces.

As an example, TAF1 and TAF7 are two direct protein partners in TFIID that undergo Co-TA⁴² and their dimeric interface is characterized by an interwoven triple β -barrel and contacts that extend into the complex core.⁴⁴ Accordingly, the TAF1 domain involved in the interaction was reported to be insoluble if not co-expressed with TAF7.⁴⁴ Similarly, Co-TA was shown to allow the association of protein partners through a specific site otherwise sterically inaccessible in the mature protein form.⁴⁵

Facilitation of homomeric assembly

Due to the close proximity forced by the co-existence of multiple nascent chains on the same polysome, simultaneous *cis* Co-TA is predicted to favor the formation of homomeric assemblies. Recently, the occurrence of co-translational homomeric assembly was found to be widespread in human cells,⁴⁶ and it did not require specialized

eukaryotic factors. In a related report, the biochemical benefits of Co-TA were studied for the homodimerization of lamin filaments,⁴⁷ showing that their co-translational interaction prevents misfolding of the individual subunits and drives the nucleation and growth of the native complex.

If *cis* Co-TA favors homomeric assembly, it would in turn disfavor heteromerization with potential proteins sharing compatible interaction domains, thus constraining the specificity of the assembly landscape. This principle has been proposed for the partner “choice” of a series of transcriptional regulators containing BTB dimerization domains,⁴⁸ also found to assemble co-translationally.⁴⁶ Yet, precocious interaction between neighboring nascent chains might also lead to misassembly when the interaction domains need to fold before dimerization. There is evidence that certain proteins might have evolved C-terminally positioned dimerization domains to prevent co-translational homodimerization.⁴⁹ While the authors suggested that this evolutionary constraint evolved to reduce misassembly defects sparked by *cis* Co-TA, it might well be an expedient to favor the formation of heterodimers by post-translational means.

Mitigation of dominant negative genetic disorders pathogenesis

In *cis* Co-TA, both assembling protein chains are translated by the same mRNA molecule, hence by the same allele. This simple fact has consequences in the mechanism of action of certain dominant negative genetic lesions, where the protein copies translated by the mutant allele impair their wild-type counterparts by assembling together in a non-functional complex.⁵⁰ It has been suggested that Co-TA would mitigate the effects of dominant negative mutations by promoting the assembly between proteins expressed by the same allele, while disfavoring the mixing of mutant copies with the wild-type ones.²³ An early example of co-translational dimerization associated with dominant negative effects in a TF has been reported for the homotetrameric tumor suppressor p53.⁵¹ Authors found evidence that p53 homodimerizes co-translationally, while tetramer formation would occur in solution (dimerization of dimers). Thus, the co-translational dimerization step would mitigate the effects of the dominant negative mutations of p53.⁵¹

Recently, a large survey of proteins undergoing homomeric Co-TA was put in relationship with known dominant or recessive disorders⁵²: the authors observed that subunits characterized by dominant negative genetic disorders are significantly less represented among proteins undergoing homomeric Co-TA; while subunits known to be involved in dominant negative effects tend to have their interactions domains shifted toward the C-terminus, thus disfavoring Co-TA. Although these theoretical and statistical evidences support the role

of Co-TA as an important tool to alleviate the burden of dominant negative mutations, it remains challenging to move from strong correlations to unequivocal causal effects. To this end, one would need to decouple complex assembly from protein synthesis, which remains a major experimental bottleneck in this field.

Localized complex assembly

The regulated intracellular trafficking of mRNAs leads to localized translation, providing a mechanism to control the spatiotemporal expression of proteins.⁵³ Precise subcellular mRNA transport and localization is thought to be mainly mediated by a diverse set of RNA-binding proteins. Mechanisms that directly target nascent protein chains have also been described, such as the well-known signal recognition particle for endoplasmic reticulum targeting.⁵⁴ Importantly, by anticipating complex assembly during protein synthesis, localized Co-TA would ensure the allocation of pre-assembled (sub)complexes directly at functional sites. In turn, this would increase the local complex concentration and it would reduce non-specific interactions in other subcellular districts.

While plenty examples exist for localized mRNA translation,^{55,56} the experimental evidence for localized Co-TA in cells is currently rather scarce. A remarkable example is the localized assembly of the nuclear pore complex (NPC), which in yeast employs extensive Co-TA among several of its subunits.^{57,58} Besides a wealth of binary co-translational interactions in the cytoplasm, the authors found that a fraction of mRNAs of two key NPC subunits, *NUP1* and *NUP2*, are physically associated with the nuclear pores. Moreover, this interaction was dependent on translation and on the integrity of N-terminal domains recognized by nuclear transport receptors (karyopherins), which co-translationally target the nascent chain and the corresponding mRNA to sites of NPC assembly. The nascent chains of the same two proteins (Nup1 and Nup2) were co-translationally associated with several other NPC subunits, suggesting that, once targeted to the nuclear envelope, they might deliver or interact with other components of the pore for localized assembly. A key point in the study was the experimental uncoupling of translation from assembly. The authors depleted the fraction of *NUP1* mRNA localized at NPCs by fusing it with *ASH1* mRNA 3' untranslated region (UTR), which anchors the transcript to the bud tip of diving cells. This led to the accumulation of Nup1 protein in cytoplasmic foci, presumably in aggregated form. To our knowledge, this is the only experimental approach that managed to physically decouple translation from complex assembly, taking advantage of the localized nature of NPC assembly in the cell.

Another form of localized Co-TA is related to the discovery of cytoplasmic foci where mRNAs

encoding for different subunits of a complex converge and exchange co-translational interactions. Examples of this are the cytoplasmic bodies where the proteasome subunits RPT1 and RPT2 interact by undergoing simultaneous Co-TA in *trans*.⁵⁹ These foci, referred by the authors as “assemblysomes”, also contain CNOT1 (Not1 in yeast), the scaffold subunit of the Ccr4-Not mRNA stability control complex. In yeast, the formation of these subcellular particles depends on the nascent protein chains harboring the respective interaction domains. Interestingly, in human cells the co-localization of *RPT1* and *RPT2* mRNAs was induced by proteotoxic stress and it was dependent on CNOT1. The same group previously reported a similar role of the Ccr4-Not complex in driving the Co-TA of specific subunits of the histone acetyltransferase (HAT) complex SAGA in yeast.⁶⁰ Although the contribution of the nascent chains and interaction domains remain to be assessed in the human system, these data suggest that localized Co-TA might depend on both intrinsic features (nascent interaction domains) and extrinsic factors (protein complexes controlling mRNA stability and fate). Indeed, examples exist where co-localized assemblies of functionally-related mRNA species do not necessarily undergo Co-TA,⁶¹ in line with the general concept of “RNA regulons”.⁶² Despite the well-documented role of UTRs in driving mRNA localization,⁵³ there is also evidence that 5' or 3'-UTRs might not be necessary for localized simultaneous Co-TA, since the expression from vectors devoid of native UTRs is sufficient to recapitulate co-translational interactions.²⁵ Clearly more efforts are needed to establish the dependence of localized Co-TA on intrinsic *versus* extrinsic cues.

Reduction of promiscuous interactions and increased complex assembly efficiency

A major challenge for the assembly of a multiprotein complex is the threat posed by non-specific promiscuous interactions with the innumerate components of the crowded cellular environment. The chance of establishing nonproductive interactions with off-target proteins increases with the number of complex subunits and with the number of potential spurious interaction partners.⁶³ In other terms, the number of possible assembly “dead-ends”, due to irreversible non-specific interactions, is predicted to severely impact large multiprotein complexes.

In this context, Co-TA is thought to play a role in promoting complex assembly efficiency, which can be interpreted as the probability of successful holo-complex formation per assembly round. A mathematical model set a theoretical framework for complex assembly efficiency, and it also predicts that Co-TA would support complex formation by intercepting and stabilizing aggregation-prone intermediates during their translation.⁶⁴ The reduction of exposed aggregation

moieties by anticipating partner binding would effectively reduce the loss of material from the assembly flux. On the same note, Co-TA is predicted to reduce the potential toxicity of free subunits, either driven by self-aggregation or complexation with other unwanted cellular components in off-pathway interactions. This is in line with the above-mentioned cytoplasmic Nup1 nucleoporin, which aggregates when localized Co-TA is impaired.⁵⁷

The allocation dilemma of paralog and moonlighting subunits

Proteins that are shared by multiple different protein complexes are known as ‘moonlighting’ proteins.⁶⁵ Conversely, protein paralogs (i.e. closely related homologous proteins generated by gene duplication) often participate in a mutually exclusive way to the composition of the same complex, giving rise to several different complex subtypes. This process, also known as “paralog switching”, can have profound effects in either providing robustness to perturbations for essential protein complexes or sub-specializing their functions.⁶⁶ Alternatively, paralogs might have diverged to a point at which they exert distinct functions.

All of the above mechanisms are extensively found across protein complexes involved in transcriptional regulation, including GTFs, chromatin architectural components, histone modifying and chromatin remodeling complexes.^{67–74} For example, paralog switching could create twelve (or more) different human SAGA complexes.⁷⁰ Yet, the “rules” that guide the relative abundance of different complex subtypes, or the allocation of a shared subunit between alternative complexes, are poorly understood. Many questions concerning the allocation of moonlighting and paralog proteins remain to be answered. For instance, is Co-TA playing a role?

Co-translational interactions differentiate assembly pathways

Studies on co-translationally assembled complexes provided a curious set of observations related to the allocation “dilemma”. A good example is the case of human TAF10, which heterodimerizes with TAF8 in the GTF TFIID, or with SUPT7L in the co-activator SAGA.⁷⁰ When investigating the assembly of mammalian TFIID and SAGA subunits we found that TAF10 undergoes sequential Co-TA with the nascent TAF8 protein, but not with its alternative partner SUPT7L.²⁵ Similar observations were made among paralogous and moonlighting proteins in NPCs shared with other unrelated complexes.⁶⁴

The above and related evidences opened the possibility that differential Co-TA might contribute

to determine the probability that a subunit ends up in one of the alternative complexes (moonlighting), or the probability that a subunit wins the competition to assemble in a complex to the detriment of its paralog (paralog switching). Note that it cannot be excluded that, in certain cases, differential Co-TA might result from binding events occurring with only a minor frequency, or that these binding events are poorly preserved by the experimental conditions. Keeping this in mind, here we suggest four molecular mechanisms that might explain differential Co-TA among related pairs of partner proteins. They would also represent new avenues of research, since, although plausible, they are still to be investigated by experimental means. Note that in this context we mainly discuss the mechanistic causes of this question, namely the molecular events leading to the phenomenon (“how?”), rather than its remote causes, such as the evolutionary track that fixed the phenomenon in the first place (“how did it evolve?”).

Affinity and concentration: as simple as that

The first potential mechanism to explain the observed differential Co-TA simply relies on two factors: affinities and concentrations. Consider a moonlighting protein that can dimerize with two alternative subunits with different affinities. This affinity difference and the concentrations of the three components in the system will be major determinants of subunit distribution at steady state.

While this simple principle is valid for all biomolecular interactions, we suggest that it may play a prominent role in determining Co-TA, due to the intrinsic properties of co-translational interactions. In Co-TA one of the two partners is a nascent polypeptide emerging from a mRNA being translated. For a given protein, the number of nascent chains in a cell is far lower than the number of fully synthesized molecules, even by taking into account a handful of nascent chains per mRNA molecule (in an actively translated polysome). On average, the number of protein molecules in a cell is three orders of magnitude higher than the number of mRNA molecules.^{75,76}

The consequence of the relative scarcity of nascent chains in the cell is that co-translational interactions with the low-affinity partner might result in a negligible pool of binding events, hard to detect and quantify. This can lead to the observed establishment of Co-TA with the first – high affinity – protein partner but not with the second one. It is also conceivable that the affinity of a moonlighting subunit for its two partners differ for the full-length proteins and their nascent counterparts. This would lead to Co-TA with the partner endowed with higher affinity in its nascent chain state. Assessing the relevance of this idea would demand the careful measurement of the binding affinities across protein partners and the relationship with their Co-TA behavior in cells.

Steric clashes and folding intermediate constraints

A second reason why the same protein manifests differential Co-TA with alternative partners can depend on differences in their folding behavior on the ribosome with respect to their free form. During translation, a nascent chain experiences a very different local environment given by the surrounding ribosome surface, its local charge and ribosome-associated factors,⁴³ which could either favor or impair Co-TA with a particular protein partner. For example, the study of the co-translational binding between domains of two human transcriptional co-factors, CBP (CEBP binding protein) and p160 (NCOA3, nuclear receptor coactivator 3), using an *in vitro* translation system,⁷⁷ showed that CBP can assemble co-translationally with the nascent polypeptide of p160, but not *vice versa*. This asymmetric Co-TA behavior was explained by the differential interactions between the nascent chains and the negatively charged ribosomal surface.

On the same line, steric clashes of certain interaction domains with the ribosome surface might explain differential Co-TA. Moreover, the intimate association with the ribosome imposes a constraint on the degrees of freedom for the nascent protein (and its interacting partner) in moving across conformational space. If on one side this constraint might funnel the correct sequential (co)folding path for certain protein pairs, it may as well impair it in other cases. For example, if certain assembly intermediates require large conformational changes that would be limited by tethering with the ribosome surface, Co-TA would be discouraged in favor of post-translational assembly.

Features affecting translation kinetics

Based on the in-depth analyses of translation kinetics,^{26–28,78,79} we emphasize that translation does not proceed at constant rate, and the modulation of translation speed affects the folding kinetics of the nascent protein and the dwell-time of interaction domains exposed outside of the ribosome exit tunnel. As a consequence, translation kinetics is predicted to affect co-translational binding events.

Apart from external cues, several features that modulate the speed at which an mRNA is translated are “encoded” in the transcript itself. These include intrinsic properties of certain aminoacidic sequences, some known as arrest peptides, which can drastically slow down or even temporarily stall translation by interacting with the ribosome exit tunnel and distorting the reactive center during elongation.⁸⁰ Although specific examples of these sequences with defined functions have been reported also in eukaryotic cells,⁸¹ we still lack a comprehensive understanding of the common properties of such peptides. Besides the effects mediated by the nascent polypeptide, the

codon usage of the mRNA itself, in relation with the abundance of isoacceptor tRNAs in the cell, affects elongation dynamics along the transcript.^{78,79}

These intrinsic properties of the transcripts leading to reversible pausing and translation slow-down at “strategic” positions, such as between protein domains or right after an interaction domain, would greatly favor co-translational binding events.^{82,83} For example, the above-mentioned co-translationally-assembled proteasome subunits Rpt1 and Rpt2 show a marked ribosome pause site downstream of the coiled coil domain that mediates their interaction.⁵⁹ An analogous example is the nucleoporin Nup1, which contains a ribosome pause site thought to buy time for its co-translational targeting to the nuclear envelope and local Co-TA with several other nucleoporins.⁵⁷

Localized assembly, mRNA proximity, translation factories

The processes of mRNA transport and localized translation might explain a subset of the differential Co-TA occurrence among paralogous proteins. Namely, the same protein might specifically undergo localized translation only with certain partners, thus promoting Co-TA with them. This does not exclusively rely on the physical connection of the corresponding mRNAs through the respective nascent chains (i.e., simultaneous Co-TA). Indeed, as long as the two mRNAs are localized in physical proximity, active translation could produce a localized hotspot of protein concentration, which would favor bimolecular interactions, including sequential Co-TA (in either directional or symmetric fashion). Moreover, in case of foci of multiple co-localized mRNAs (often called “assemblyosomes”, “microtranslatomes”, “translation factories” and others),^{56,59,61} also the local concentration of nascent chains would considerably increase, again favoring Co-TA. In the future it would need to be assessed whether the existence of translation foci between certain mRNA partners would favor Co-TA.

It is important to highlight that the molecular causes of differential Co-TA proposed above could also explain why certain proteins undergo Co-TA while others do not, regardless of whether they are paralogs or moonlighting proteins.

The use of co-translational assembly in the construction of transcription regulatory complexes

Despite their nuclear functions, complexes involved in the regulation of gene transcription can assemble in the cytoplasm, including RNA polymerase II itself.^{84,85} Early reports regarding the occurrence of Co-TA for nuclear proteins

described the homomeric dimerization of sequence-specific TFs, namely p53 and NF- κ B1.^{51,86}

The first example of Co-TA for a large nuclear multiprotein complex was reported by the Dichtl group, where the authors found that several subunits of the yeast COMPASS histone methyltransferase complex were associated with the *SET1* mRNA, which encodes the catalytic subunit of the complex.⁸⁷ They demonstrated that the *SET1* mRNA-associated subunits assembled co-translationally with the nascent Set1 polypeptide in the cytoplasm, effectively forming an assembly intermediate in the complex biogenesis during translation. Fueled by these and other observations of cytoskeletal proteins assembling co-translationally, Duncan and Mata performed a wider screening in fission yeast, by detecting mRNAs associated with 31 different proteins using RNA immunoprecipitation coupled to DNA microarrays (RIP-chip).⁸⁸ They found a widespread use of Co-TA across the sampled proteins, with almost 40% of the tested baits associated with mRNAs encoding partner proteins. Interestingly, among the tested proteins, several of them are subunits of well-known nuclear transcription regulatory complexes. Specifically, this approach revealed Co-TA occurrence for subunits of yeast TFs (Atf1/Pcr1), chromatin remodeling complexes (SWI/SNF, RSC, SWR1 and INO80), and the splicing-related exon junction complex (Mnh1, Mni1). This study suggested a widespread occurrence of Co-TA across diverse classes of yeast complexes and revealed an exquisite specificity of the partners interacting co-translationally.

More recently, our group investigated the occurrence of Co-TA in the construction of large multiprotein complexes involved in the regulation of Pol II transcription in mammalian cells.^{25,42,89} We focused on TFIID (14 subunits) and two histone acetyltransferase (HAT) complexes: SAGA (20 subunits) and ATAC (Ada two A containing; 10 subunits). As a common trait, Co-TA kickstarts the cytoplasmic formation of all three of these complexes, confirming its widespread use in the assembly of nuclear regulatory factors also in mammalian cells (Figure 2). Nonetheless, the three complexes show remarkable differences in other aspects of their biogenesis that we will overview in the following section.

Different assembly routes to a common destination: The nucleus

As a common ground of all co-translational interactions, the two interacting species are located – at least transiently – in the cytoplasm. Apart from this obligate property, the subsequent steps that define the assembly pathways of multiprotein complexes can differ remarkably, even for functionally related complexes. As a

paradigm for this concept, here we consider the assembly logics of three different nuclear protein complexes that sustain Pol II transcription: TFIID, SAGA and ATAC. We summarize the findings on their assembly pathways and highlight remaining gaps in our understanding of their biogenesis. We argue that the described blueprints for complex formation could apply to the biogenesis of many other large nuclear multiprotein complexes.

TFIID: cytoplasmic building blocks converge on a ‘driver’ subunit

Where and how the TFIID complex assembles was a longstanding question. Originally, the isolation of the holo-complex from nuclear extracts hinted that it formed in the nucleus starting from subunits or subcomplexes individually imported from the cytoplasm. It later became evident that the cytoplasm contained several defined TFIID subassemblies, such as TAF2/TAF8/TAF10,⁹⁰ TAF11/TAF13⁹¹ and TAF5/TAF6/TAF9.⁴¹ Our group first demonstrated the occurrence of Co-TA for specific pairs of TFIID subunits (TAF1/TBP, TAF8/TAF10, TAF6/TAF9) in mammalian cells, defining also the structural domains mediating the co-translational interactions.²⁵

Recently, we systematically probed the entire co-translational network within TFIID.⁴² We found an extensive use of Co-TA across the entire spectrum of TFIID subunits, including a surprising concentration of co-translational interactions on nascent TAF1 polypeptide. We summarize our findings on the assembly of TFIID in the following hierarchical co-translational assembly model (Figure 2B): first, all subunits participate to co-translational interactions with well-characterized direct partners in the complex, either by sequential or symmetrical Co-TA. This widespread use of Co-TA seeds the assembly pathway by providing preassembled simple components which efficiently combine post-translationally to form defined higher-order submodules that can be readily isolated from the cytoplasm at steady state. At this stage, the pathway converges in a higher order co-translational assembly phase: the different cytoplasmic building blocks combine on nascent TAF1 – the largest subunit of the complex – establishing a well-defined set of co-translational interactions with distinct anchor points along the nascent TAF1 polypeptide (Figure 2B). TAF1 is a modular protein that works as a flexible scaffold connecting the different structural lobes of TFIID, as evidenced by biochemical and structural observations.^{42,92–96} The TAF1-mediated co-translational recruitment of the pre-assembled building blocks leads to the release of a fully assembled complex, which is efficiently shuttled to the nucleus.

The nucleation of the late steps of TFIID assembly on nascent TAF1 protein, and the fact that it represents the main rate-limiting factor along the assembly line, define TAF1 as a ‘driver’

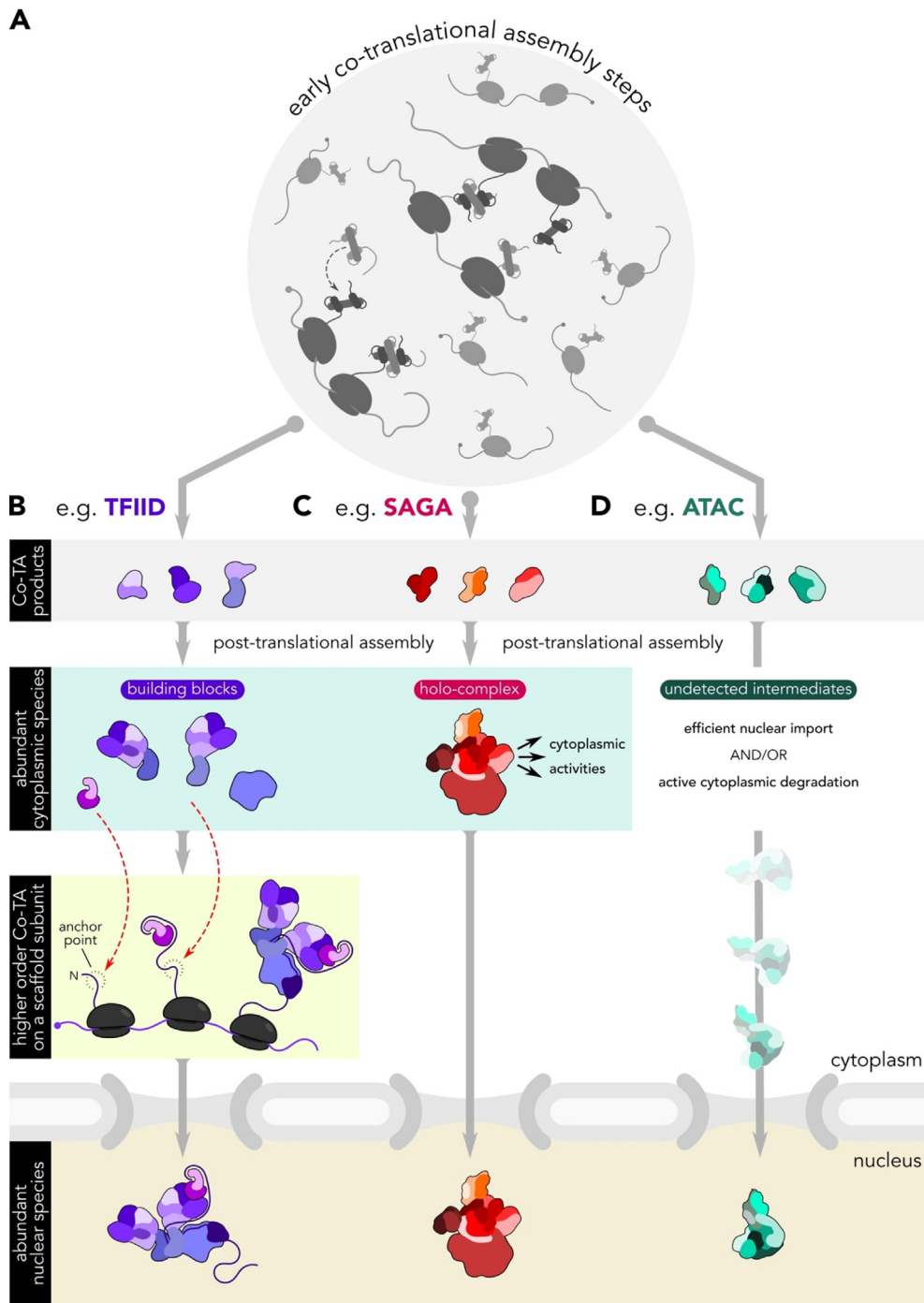


Figure 2. Distinct co-translational assembly strategies among transcription complexes. Schematic representation of alternative co-translational assembly pathways for distinct transcription-related regulatory protein complexes. (A) Early interactions between specific subunits are established co-translationally, giving rise to pre-assembled heterodimeric assemblies in the cytoplasm. The subsequent steps are exemplified by the distinct assembly strategies employed by three large multiprotein complexes involved in Pol II transcription: TFIID (B), SAGA (C) and ATAC (D). Note that the ATAC holo-complex formation might occur in the nucleus.

subunit that regulates the abundance of the holo-complex. On the same line, TAF1 depletion with siRNAs lead to the accumulation of TFIID building blocks in the cytoplasm.⁴² In this view, the

concerted Co-TA on the nascent driver subunit might represent a quality checkpoint for complex integrity that regulates nuclear import. Interestingly, in yeast, TAF1 mRNA was found physically

associated with the nuclear pore complex, opening the possibility of localized complex assembly followed by direct nuclear translocation.⁵⁷ Moreover, a systematic mapping of the interactome of human nuclear transport receptors by BioID suggested that all TFIID subunits share the same nuclear import systems, which is consistent with the translocation of a pre-assembled unit.⁹⁷

The strategy adopted for the assembly of TFIID shows several similarities with the co-translational complex formation of COMPASS in yeast.⁸⁷ Transcription-related complexes constructed around a scaffold subunit, like TFIID or COMPASS, are far from rare. Examples of subunits with a scaffolding function are Eaf1 in NuA4 complex,⁹⁸ Sin3 in SIN3-HDAC complex,⁹⁹ SUZ12 in PRC2 complex¹⁰⁰ or CNOT1 in CCR4-NOT complex.¹⁰¹ Further dedicated studies are needed to address the generality of the Co-TA strategy centered around a central hub, such as TAF1 or Set1.

SAGA: co-translational heterodimerization events, fully assembled cytoplasmic holo-complex and cytoplasmic activity

A second, distinct strategy for complex assembly is exemplified by the biogenesis of the multifunctional SAGA co-activator complex. The first suggestion for Co-TA in SAGA came from yeast proteins, where Sus1 (the homolog of human ENY2) efficiently solubilized Sgf11 (the human homolog of ATXN7L3) when co-expressed in bacteria, suggesting a partner chaperoning effect mediated by Co-TA.¹⁷ Similarly, knockdown of human ENY2, a subunit of the SAGA deubiquitylation (DUB) module, also decreased the expression of its interaction partner, ATXN7L3, in human cells.³⁵ More direct evidence for co-translational interactions in SAGA came from experiments in yeast, where the HAT module subunit Ada2 (human TADA2B) was found to interact with nascent Spt20 polypeptide (human SUPT20H).⁶⁰ Later we confirmed the direct Co-TA between human ENY2 and nascent ATXN7L3.²⁵

We recently expanded SAGA co-translational interaction network to its core structural module.⁸⁹ As for TFIID, the tested subunits established specific co-translational binding events with their most direct partners in the complex (i.e., TAF9/TAF6L, TAF12/TADA1, TAF5L/SUPT20H). Note however that we did not (yet) identify any higher-order Co-TA on a single nascent scaffold subunit in human SAGA. Yet, based on the cryo-EM structure of SAGA, SUPT20H or ATXN7 could be candidates that may fulfil such a role.¹⁰² Nevertheless, it is also possible that different higher order assembly logic operates in SAGA assembly.

A striking observation was that the entire SAGA complex, equipped with all of its functional modules, can be isolated from cytoplasmic extracts of different cell lines, regardless of the protein used as bait in the immunoprecipitation

(Figure 2C).⁸⁹ The relative subunit stoichiometries estimated for the cytoplasmic complex were analogous to those found for the nuclear complex. The detection of fully assembled SAGA in the cytoplasm suggests that the co-translational heterodimerization steps detected for specific subunits are followed by an efficient post-translational assembly phase, resulting in the formation of the holo-complex. In this regard, experiments in yeast evidenced an ordered assembly pathway among SAGA functional modules, where the incorporation of the large subunit Tra1 (human TRRAP) promotes the association of the DUB module with SAGA.¹⁰³ The same study also highlighted the role of molecular chaperones in the assembly process, but suggested that Tra1 does not use Co-TA to incorporate in SAGA.

The presence of a pool of SAGA complexes in the cytoplasm raised the question on a potential function of SAGA in this cellular compartment. Indeed, knockdown of SAGA HAT module specific subunits resulted in a substantial decrease in acetylated peptides in the cytoplasmic fraction.⁸⁹ Thus, cytoplasmic SAGA maintains its acetyltransferase activity and targets dozens of cytoplasmic proteins (Figure 2C). Whether well-defined cellular functions depend on this cytoplasmic SAGA population remains currently unknown. But what would cause the accumulation of active SAGA in the cytoplasm? Apart from hypothesizing suboptimal nuclear import rates, it is also conceivable that the pool of holo-SAGA might be imported to the nucleus upon (stress) stimuli or actively exported from the nucleus back into the cytoplasm thanks to – yet uncharacterized – nuclear export signals among its subunits. Still, we lack a precise understanding of the molecular features that determine the distribution of macromolecular complexes in the subcellular environments.

ATAC: co-translational assembly network and nucleus-restricted activity

In the above-mentioned study on SAGA assembly, we also investigated the biogenesis of the related ATAC complex, which shares with SAGA a similar HAT module, differing only by one subunit.^{104,105} The shared subunits are limited to the HAT module, while the structural core modules of the two complexes differ.

Similarly to TFIID, all subunits of ATAC core participate in co-translational interactions, either through the sequential or simultaneous pathways. The ATAC Co-TA network looks rather intricate, with specific nascent chains engaged co-translationally by multiple partners, and subunits reciprocally associated with partners' mRNAs, hinting at simultaneous Co-TA.⁸⁹ For one such a pair, YEATS2 and ZZZ3, we found that a fraction of the respective mRNAs reciprocally co-localize in the cytoplasm of HeLa cells. Nevertheless, the current lack of an experimental structure for the

ATAC complex limits the rationalization of its Co-TA network. The fact that, when available, the high-resolution structural observations in multiprotein complexes perfectly agree with the Co-TA data argues that new pair-wise interactions revealed by studying Co-TA represent a valuable source of information to set the basis for mapping interaction domains and guiding the structural analyses.

ATAC biogenesis strikingly differs from those described for TFIID or SAGA in that neither ATAC free subunits nor building blocks could be detected in the cytoplasm (Figure 2D). Despite the robust detection of co-translational interactions among ATAC subunits in this compartment, the fully synthesized components and the holo-complex were exclusively found in the nuclear fraction.⁸⁹ The cause for ATAC extreme tropism for the nuclear environment is currently unknown. Fast nuclear import kinetics coupled with the absence of nuclear export signals would result in a unidirectional flux of ATAC complex towards the nucleus. Otherwise, targeted degradation mechanisms could act in the cytoplasm to wipe out unassembled subunits or even the entire complex (Figure 2D). A third hypothesis would be that ATAC subunits are translated and assembled co-translationally in proximity to the nuclear pore complexes and directly delivered to the nucleus. Nevertheless, it is likely that multiple mechanisms contribute to the nuclear-restricted localization of ATAC. This is in striking opposition to the related SAGA complex, which lingers in the cytoplasm where it acetylates several non-histone proteins (Figure 2C). Thus, the determinants for the different subcellular distribution of the two complexes must not reside in the shared HAT module. While the early assembly steps for ATAC happen in the cytoplasm through Co-TA, the question on where the holo-complex is finally formed remains open.

Transcriptome-wide evidence for Co-TA in the assembly of other transcription complexes

Recently, a high-throughput method named Disome Selective Profiling (DiSP) opened the possibility to detect certain co-translational interactions transcriptome-wide.⁴⁶ The method is a variant of ribosome profiling, where the disome (pairs of ribosomes) fraction is isolated and the corresponding mRNA fragments protected from RNase digestion are sequenced (Figure 3A). The ribosome footprints derived from the disome fraction thus identify transcripts connected to a second ribosome through the respective nascent chains that undergo simultaneous Co-TA (Co-Co assembly), either in *cis* or in *trans* (Figure 3A, see also Figure 1C, D). Note that this approach cannot detect sequential assembly, where distinct ribosomes are not connected by the nascent chains. Moreover, DiSP is agnostic on the identity of the pairing

between partners undergoing Co-Co assembly, since it simply detects transcripts enriched in the disome fraction. Thus, we know that a certain protein undergoes Co-Co assembly, but we ignore the identity of the partner.

By applying a stringent set of criteria, the authors provided a list of high- and lower-confidence candidate proteins undergoing simultaneous Co-TA in human cells. The high-confidence list was depleted of monomeric proteins, while highly enriched with proteins annotated to form homodimers, pointing at Co-Co assembly in *cis*. Yet, the dataset was also enriched with proteins part of heteromeric assemblies.⁴⁶ Thus, this technique allowed to detect also co-translational heteromeric interactions established in *trans*, between nascent chains encoded by different transcripts.

Interestingly, Co-Co assembly candidates annotated as nuclear proteins (~34%) are slightly, but significantly enriched in the DiSP dataset (Figure 3B). Among them, the number of Co-Co candidates annotated as TFs is lower than what it would be expected by chance, suggesting that simultaneous Co-TA is not prevalent among TFs (Figure 3B). Yet, the dataset contains 231 proteins annotated as TFs in The Human Transcription Factors database,¹⁰⁶ suggesting that ~14% of human TFs homo- or heterodimerize in the cytoplasm using simultaneous Co-TA (Figure 3C). Nuclear proteins without TFs (non-TF nuclear proteins) are found robustly enriched in the Co-Co assembly candidate list (Figure 3B). Indeed, the dataset is rich in subunits belonging to large nuclear multiprotein complexes. Figure 3D shows a non-comprehensive overview of the complexes involved in transcriptional regulation with subunits found to undergo simultaneous Co-TA by DiSP.⁴⁶ The dataset includes subunits from several functionally-distinct complexes, such as the Mediator complex, histone acetyltransferase (HAT) complexes (SAGA, ATAC, NuA4, MORF, MOZ, HBO1, MSL, NSL), histone methyltransferase (HMT) complexes (SET1/MLL, PRC2), histone deacetylase (HDAC) complexes (SIN3A, MIDAC, NuRD), nucleosome remodeling complexes (SWI/SNF, NuA4, NoRC, NuRD), GTFs (TFIID, TFIIF, TFIH), promoter pausing factors (Integrator, NELF), the TREX mRNA export complex, the cohesin complex, the spliceosome and Pol II itself (Figure 3D). This transcriptome-wide analysis provides strong evidence for the widespread use of simultaneous Co-TA as a basis for the assembly of a plethora of nuclear multisubunit protein complexes in human cells.

Perspectives

This review summarizes the current understanding of co-translational interactions, with a special focus on the assembly of complexes that regulate gene transcription. We outlined the

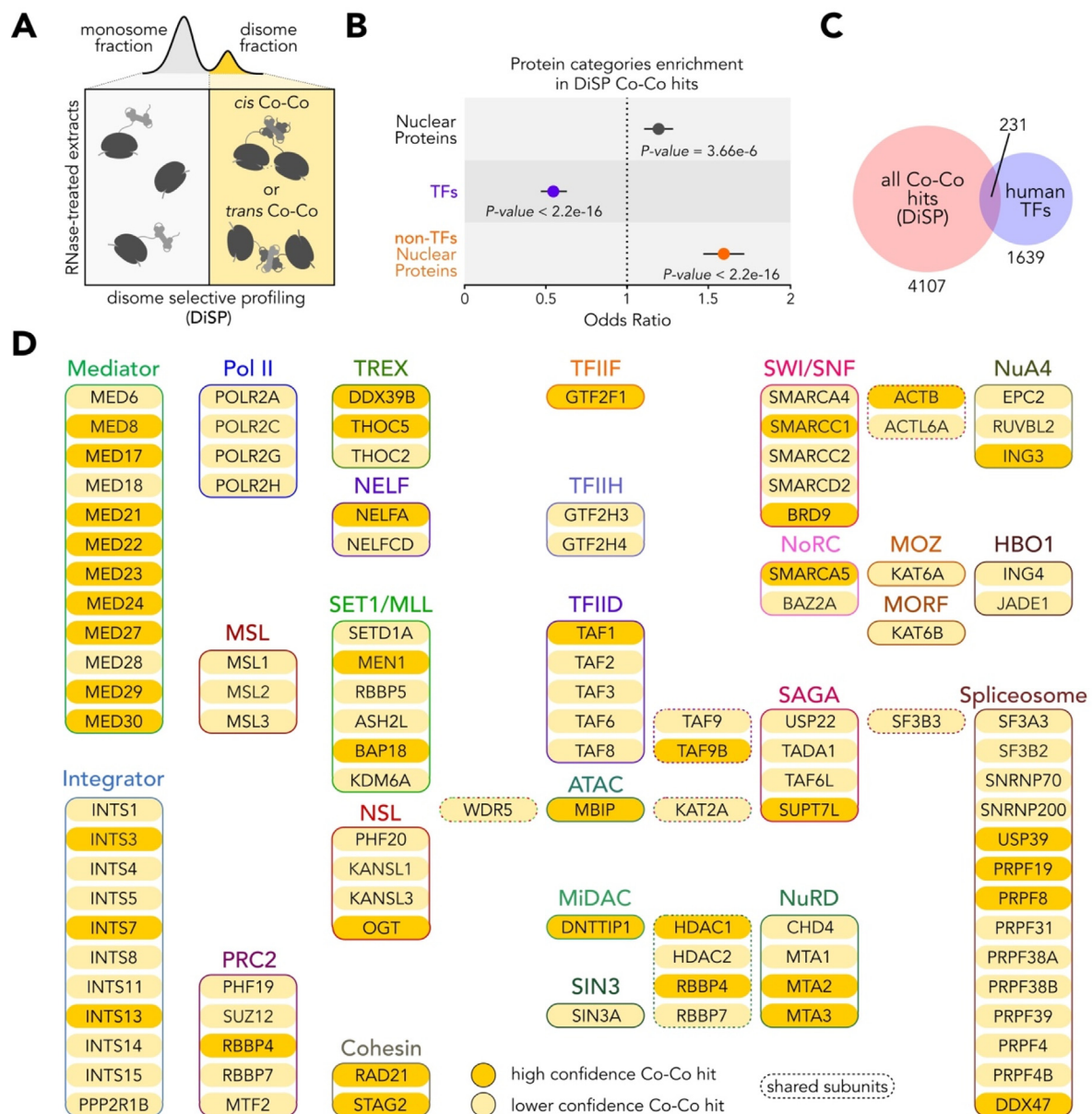


Figure 3. Transcriptome-wide evidence for Co-TA in the assembly of other transcription complexes. (A) Schematic outline of the DiSP method to detect transcriptome-wide simultaneous Co-TA events (Co-Co) by sequencing ribosome-protected mRNA fragments in the disome fraction. (B) The enrichment of nuclear proteins, TFs and TF-excluded nuclear proteins (non-TF nuclear proteins) in the Co-Co candidate list was evaluated by performing Fisher's exact test on the DiSP dataset.⁴⁶ The plot shows the resulting odds ratios and P-values for each category against the reference set. The reference set was built by the total number of expressed genes retrieved from the processed datasets in HEK293-T and U2OS cells (GEO ID: GSE151959). Nuclear protein annotation was retrieved from UniProt and TFs annotation was based on The Human Transcription Factor database.¹⁰⁶ (C) Venn diagram showing the fraction of DiSP hits annotated as TFs. (D) Non-comprehensive overview of the DiSP hits which are part of characterized multiprotein complexes involved in Pol II transcriptional regulation. All the depicted subunits are found either as high- (dark yellow) or lower-confidence (light yellow) hits in the published DiSP dataset.⁴⁶ The depicted identified subunits are represented in the corresponding protein complex box. Subunits shared with multiple complexes are shown with a dotted line and are positioned between the alternative complexes. The assignment of each protein to the corresponding human complex is based on the Complex Portal database.¹⁰⁷

expected advantages – and evidences – of Co-TA in guiding complex formation (1), we proposed a series of explanations for the allocation “dilemma”

of paralog and moonlighting subunits in the light of co-translational interactions (2), we described as a paradigm the distinct strategies employed for the

Co-TA of three key transcription regulatory complexes (TFIID, SAGA and ATAC) (3) and we summarized the current evidence for the widespread use of Co-TA for the assembly of nuclear regulatory complexes (4).

The many open questions disseminated throughout the manuscript make each of the above points also an important avenue for future research. Regarding point (1), a key bottleneck to experimentally investigate the predicted advantages of Co-TA is the lack of a way to control the occurrence of co-translational *versus* post-translational assembly events in the cell, since both rely on protein synthesis. The use of pure *in vitro* translation systems coupled with purified recombinant protein interaction assays or dedicated single molecule analyses might provide an expedient to compare complex assembly with or without concurrent protein synthesis. Yet, the overarching goal remains to generate a system to specifically inhibit or promote co-translational interactions in cells and test the predicted fitness advantages of Co-TA.

For the subunit allocation “dilemma” in point (2), we propose that the combination of interaction domain position, affinity of the bimolecular interaction, constraints in the conformational folding pathway on the ribosome, the transcript-specific translation kinetics, and the local concentration of the components in the cell could ultimately determine whether a protein engages in Co-TA. While additional mechanisms can be foreseen, our intent is to provide a series of relevant variables to be challenged by experimental means.

The assembly strategy of large multiprotein complexes involved in transcriptional regulation (point 3) has been largely overlooked. The strategies that we described for TFIID, SAGA and ATAC open new questions on the nuclear import/export pathways employed by the different complexes, and degradation sensitivity of the different subunits or submodules in the nuclear *versus* cytoplasmic environments. Ultimately, dedicated studies on the biogenesis of other nuclear multiprotein complexes will assess the generality of the assembly strategies described above (Figure 2) and will shed light on the underlying regulatory mechanisms.

Last, the handful of in-depth mechanistic studies describing the Co-TA of specific nuclear complexes involved in gene regulation have recently been flanked by a transcriptome-wide approach (DiSP) that provided a large-scale overview on the diffusion of Co-TA in the proteome (point 4). Since DiSP is limited to the detection of simultaneous Co-TA events only, and it has currently been performed on two cell lines, the resulting picture represents an underestimation of all assembly events that occur co-translationally. In the future, the development

of new methods that combine transcriptomics with proteomics might provide an overview of all co-translational interactions, including sequential Co-TA events. Ultimately, the knowledge obtained by pursuing these research avenues might provide novel opportunities to better understand and intervene on the expression of genes.

CRedit authorship contribution statement

Andrea Bernardini: Conceptualization, Visualization, Writing – original draft, Writing – review & editing, Project administration. **László Tora:** Supervision, Funding acquisition, Writing – review & editing, Project administration.

DATA AVAILABILITY

This is a review article, no new data has been generated.

DECLARATION OF COMPETING INTEREST

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

Acknowledgements

We thank the present and past members of the Tora lab and HT. M. Timmers for insightful discussions. This work was financially supported by ERC-Advanced grant (ERC-2013-340551, Birtaction), by Agence Nationale de la Recherche (ANR) ANR-PRCI-19-CE12-0029-01 (EpiCAST); ANR-20-CE12-0017-03, ANR-22-CE11-0013-01_ACT; Fondation pour la Recherche Médicale (EQU-2021-03012631); NIH MIRA (R35GM139564); and NSF (Award Number:1933344) grants. A.B. has been supported by the Fondation ARC pour la recherche sur le cancer (ARCPOST-DOC2021080004113). This work, as part of the ITI 2021-2028 program of the University of Strasbourg, was also supported by IdEx Unistra (ANR-10-IDEX-0002), and by SFRI-STRAT'US project (ANR 20-SFRI-0012) and EUR IMCBio (ANR-17-EURE-0023) under the framework of the French Investments for the Future Program.

Received 25 October 2023;
Accepted 1 December 2023;
Available online 5 December 2023

Keywords:

coactivator complexes;
mRNA;

cotranslational assembly;
moonlighting subunits;
genetic disorders

References

- Roeder, R.G., (2019). 50+ years of eukaryotic transcription: an expanding universe of factors and mechanisms. *Nature Struct. Mol. Biol.* **26**, 783–791. <https://doi.org/10.1038/s41594-019-0287-x>.
- Amoutzias, G.D., Robertson, D.L., de Peer, Y.V., Oliver, S.G., (2008). Choose your partners: dimerization in eukaryotic transcription factors. *Trends Biochem. Sci.* **33**, 220–229. <https://doi.org/10.1016/j.tibs.2008.02.002>.
- Malik, S., Roeder, R.G., (2023). Regulation of the RNA polymerase II pre-initiation complex by its associated coactivators. *Nature Rev. Genet.* **24**, 767–782. <https://doi.org/10.1038/s41576-023-00630-9>.
- Malik, S., Roeder, R.G., (2010). The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. *Nature Rev. Genet.* **11**, 761–772. <https://doi.org/10.1038/nrg2901>.
- Mannervik, M., (2014). Control of Drosophila embryo patterning by transcriptional co-regulators. *Exp. Cell Res.* **321**, 47–57. <https://doi.org/10.1016/j.yexcr.2013.10.010>.
- Millard, C.J., Watson, P.J., Fairall, L., Schwabe, J.W.R., (2013). An evolving understanding of nuclear receptor coregulator proteins. *J. Mol. Endocrinol.* **51**, T23–T36. <https://doi.org/10.1530/JME-13-0227>.
- Marsh, J.A., Hernández, H., Hall, Z., Ahnert, S.E., Perica, T., Robinson, C.V., Teichmann, S.A., (2013). Protein complexes are under evolutionary selection to assemble via ordered pathways. *Cell.* **153**, 461–470. <https://doi.org/10.1016/j.cell.2013.02.044>.
- de Martin, X., Sodaei, R., Santpere, G., (2021). Mechanisms of binding specificity among bHLH transcription factors. *Int. J. Mol. Sci.* **22**, 9150. <https://doi.org/10.3390/ijms22179150>.
- Rodríguez-Martínez, J.A., Reinke, A.W., Bhimsaria, D., Keating, A.E., Ansari, A.Z., (2017). Combinatorial bZIP dimers display complex DNA-binding specificity landscapes. *eLife.* **6**, e19272.
- Mark, K.G., Kolla, S., Aguirre, J.D., Garshott, D.M., Schmitt, S., Haakonsen, D.L., Xu, C., Kater, L., Kempf, G., Martínez-González, B., Akopian, D., See, S.K., Thomä, N.H., Rapé, M., (2023). Orphan quality control shapes network dynamics and gene expression. *Cell* **186**, 3460–3475.e23. <https://doi.org/10.1016/j.cell.2023.06.015>.
- Bradner, J.E., Hnisz, D., Young, R.A., (2017). Transcriptional addiction in cancer. *Cell* **168**, 629–643. <https://doi.org/10.1016/j.cell.2016.12.013>.
- Bushweller, J.H., (2019). Targeting transcription factors in cancer — from undruggable to reality. *Nature Rev. Cancer.* **19**, 611–624. <https://doi.org/10.1038/s41568-019-0196-7>.
- Helin, K., Dhanak, D., (2013). Chromatin proteins and modifications as drug targets. *Nature* **502**, 480–488. <https://doi.org/10.1038/nature12751>.
- Lee, T.I., Young, R.A., (2013). Transcriptional Regulation and Its Misregulation in Disease. *Cell* **152**, 1237–1251. <https://doi.org/10.1016/j.cell.2013.02.014>.
- Wang, Z., Wang, P., Li, Y., Peng, H., Zhu, Y., Mohandas, N., Liu, J., (2021). Interplay between cofactors and transcription factors in hematopoiesis and hematological malignancies. *Signal Transduct. Target. Ther.* **6**, 1–16. <https://doi.org/10.1038/s41392-020-00422-1>.
- Kim, Y.E., Hipp, M.S., Bracher, A., Hayer-Hartl, M., Ulrich Hartl, F., (2013). Molecular chaperone functions in protein folding and proteostasis. *Annu. Rev. Biochem.* **82**, 323–355. <https://doi.org/10.1146/annurev-biochem-060208-092442>.
- Diebold, M.-L., Fribourg, S., Koch, M., Metzger, T., Romier, C., (2011). Deciphering correct strategies for multiprotein complex assembly by co-expression: Application to complexes as large as the histone octamer. *J. Struct. Biol.* **175**, 178–188. <https://doi.org/10.1016/j.jsb.2011.02.001>.
- Pausch, P., Singh, U., Ahmed, Y.L., Pillet, B., Murat, G., Altegoer, F., Stier, G., Thoms, M., Hurt, E., Sinning, I., Bange, G., Kressler, D., (2015). Co-translational capturing of nascent ribosomal proteins by their dedicated chaperones. *Nature Commun.* **6**, 7494. <https://doi.org/10.1038/ncomms8494>.
- Morales-Polanco, F., Lee, J.H., Barbosa, N.M., Frydman, J., (2022). Cotranslational mechanisms of protein biogenesis and complex assembly in eukaryotes. *Annu. Rev. Biomed. Data Sci.* **5**, 67–94. <https://doi.org/10.1146/annurev-biodatasci-121721-095858>.
- Khan, K., Fox, P.L., (2023). Benefits of co-translational complex assembly for cellular fitness. *BioEssays* **45**, 2300024. <https://doi.org/10.1002/bies.202300024>.
- Schwarz, A., Beck, M., (2019). The benefits of cotranslational assembly: a structural perspective. *Trends Cell Biol.* **29**, 791–803. <https://doi.org/10.1016/j.tcb.2019.07.006>.
- Williams, N.K., Dichtl, B., (2018). Co-translational control of protein complex formation: a fundamental pathway of cellular organization? *Biochem. Soc. Trans.* **46**, 197–206. <https://doi.org/10.1042/BST20170451>.
- Natan, E., Wells, J.N., Teichmann, S.A., Marsh, J.A., (2017). Regulation, evolution and consequences of cotranslational protein complex assembly. *Curr. Opin. Struct. Biol.* **42**, 90–97. <https://doi.org/10.1016/j.sbi.2016.11.023>.
- Badonyi, M., Marsh, J.A., (2023). Hallmarks and evolutionary drivers of cotranslational protein complex assembly. *FEBS J.* <https://doi.org/10.1111/febs.16869>.
- Kamenova, I., Mukherjee, P., Conic, S., Mueller, F., El-Saafin, F., Bardot, P., Garnier, J.-M., Dembele, D., Capponi, S., Timmers, H.T.M., Vincent, S.D., Tora, L., (2019). Co-translational assembly of mammalian nuclear multisubunit complexes. *Nature Commun.* **10**, 1740. <https://doi.org/10.1038/s41467-019-09749-y>.
- Komar, A.A., (2009). A pause for thought along the co-translational folding pathway. *Trends Biochem. Sci.* **34**, 16–24. <https://doi.org/10.1016/j.tibs.2008.10.002>.
- Höpfler, M., Hegde, R.S., (2023). Control of mRNA fate by its encoded nascent polypeptide. *Mol. Cell.* **83**, 2840–2855. <https://doi.org/10.1016/j.molcel.2023.07.014>.
- Choi, J., Grosely, R., Prabhakar, A., Lapointe, C.P., Wang, J., Puglisi, J.D., (2018). How messenger RNA and nascent chain sequences regulate translation elongation. *Annu. Rev. Biochem.* **87**, 421–449. <https://doi.org/10.1146/annurev-biochem-060815-014818>.

29. Fribourg, S., Romier, C., Werten, S., Gangloff, Y.-G., Poterszman, A., Moras, D., (2001). Dissecting the interaction network of multiprotein complexes by pairwise coexpression of subunits in *E. coli*. Edited by K. Nagai. *J. Mol. Biol.* **306**, 363–373. <https://doi.org/10.1006/jmbi.2000.4376>.
30. Juszkievicz, S., Hegde, R.S., (2018). Quality control of orphaned proteins. *Mol. Cell.* **71**, 443–457. <https://doi.org/10.1016/j.molcel.2018.07.001>.
31. Pla-Prats, C., Thomä, N.H., (2022). Quality control of protein complex assembly by the ubiquitin–proteasome system. *Trends Cell Biol.* **32**, 696–706. <https://doi.org/10.1016/j.tcb.2022.02.005>.
32. Yagita, Y., Zavodszky, E., Peak-Chew, S.-Y., Hegde, R. S., (2023). Mechanism of orphan subunit recognition during assembly quality control. *Cell.* **186**, 3443–3459. e24. <https://doi.org/10.1016/j.cell.2023.06.016>.
33. Bardot, P., Vincent, S.D., Fournier, M., Hubaud, A., Joint, M., Tora, L., Pourquié, O., (2017). The TAF10-containing TFIID and SAGA transcriptional complexes are dispensable for early somitogenesis in the mouse embryo. *Dev. Camb. Engl.* **144**, 3808–3818. <https://doi.org/10.1242/dev.146902>.
34. Wright, K.J., Marr, M.T., Tjian, R., (2006). TAF4 nucleates a core subcomplex of TFIID and mediates activated transcription from a TATA-less promoter. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 12347–12352. <https://doi.org/10.1073/pnas.0605499103>.
35. Umlauf, D., Bonnet, J., Waharte, F., Fournier, M., Stierle, M., Fischer, B., Brino, L., Devys, D., Tora, L., (2013). The human TREX-2 complex is stably associated with the nuclear pore basket. *J. Cell Sci.* **126**, 2656–2667. <https://doi.org/10.1242/jcs.118000>.
36. Steward, M.M., Lee, J.-S., O'Donovan, A., Wyatt, M., Bernstein, B.E., Shilatifard, A., (2006). Molecular regulation of H3K4 trimethylation by ASH2L, a shared subunit of MLL complexes. *Nature Struct. Mol. Biol.* **13**, 852–854. <https://doi.org/10.1038/nsmb1131>.
37. Mashtalir, N., D'Avino, A.R., Michel, B.C., Luo, J., Pan, J., Otto, J.E., Zullo, H.J., McKenzie, Z.M., Kubiak, R.L., St. R., Pierre, A.M., Valencia, S.J., Poynter, S.H., Cassel, J. A., Ranish, C.K., (2018). Modular organization and assembly of SWI/SNF family chromatin remodeling complexes. *Cell* **175**, 1272–1288. e20. <https://doi.org/10.1016/j.cell.2018.09.032>.
38. Guo, P., Hoang, N., Sanchez, J., Zhang, E.H., Rajawasam, K., Trinidad, K., Sun, H., Zhang, H., (2022). The assembly of mammalian SWI/SNF chromatin remodeling complexes is regulated by lysine-methylation dependent proteolysis. *Nature Commun.* **13**, 6696. <https://doi.org/10.1038/s41467-022-34348-9>.
39. Taherbhoy, A.M., Daniels, D.L., (2023). Harnessing UBR5 for targeted protein degradation of key transcriptional regulators. *Trends Pharmacol. Sci.* **44**, 758–761. <https://doi.org/10.1016/j.tips.2023.09.001>.
40. Shiber, A., Döring, K., Friedrich, U., Klann, K., Merker, D., Zedan, M., Tippmann, F., Kramer, G., Bukau, B., (2018). Cotranslational assembly of protein complexes in eukaryotes revealed by ribosome profiling. *Nature* **561**, 268–272. <https://doi.org/10.1038/s41586-018-0462-y>.
41. Antonova, S.V., Haffke, M., Corradini, E., Mikuciunas, M., Low, T.Y., Signor, L., van Es, R.M., Gupta, K., Scheer, E., Vos, H.R., Tora, L., Heck, A.J.R., Timmers, H.T.M., Berger, I., (2018). Chaperonin CCT checkpoint function in basal transcription factor TFIID assembly. *Nature Struct. Mol. Biol.* **25**, 1119–1127. <https://doi.org/10.1038/s41594-018-0156-z>.
42. Bernardini, A., Mukherjee, P., Scheer, E., Kamenova, I., Antonova, S., Mendoza Sanchez, P.K., Yayli, G., Morlet, B., Timmers, H.T.M., Tora, L., (2023). Hierarchical TAF1-dependent co-translational assembly of the basal transcription factor TFIID. *Nature Struct. Mol. Biol.* **30**, 1141–1152. <https://doi.org/10.1038/s41594-023-01026-3>.
43. Waudby, C.A., Dobson, C.M., Christodoulou, J., (2019). Nature and regulation of protein folding on the ribosome. *Trends Biochem. Sci.* **44**, 914–926. <https://doi.org/10.1016/j.tibs.2019.06.008>.
44. Wang, H., Curran, E.C., Hinds, T.R., Wang, E.H., Zheng, N., (2014). Crystal structure of a TAF1-TAF7 complex in human transcription factor IID reveals a promoter binding module. *Cell Res.* **24**, 1433–1444. <https://doi.org/10.1038/cr.2014.148>.
45. Khan, K., Long, B., Baleanu-Gogonea, C., Gogonea, V., Deshpande, G.M., Vasu, K., Fox, P.L., (2022). Cotranslational interaction of human EBP50 and ezrin overcomes masked binding site during complex assembly. *Proc. Natl. Acad. Sci.* **119**, <https://doi.org/10.1073/pnas.2115799119> e2115799119.
46. 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**, 57–64. <https://doi.org/10.1126/science.abc7151>.
47. Wruck, F., Schmitt, J., Fenzl, K., Bertolini, M., Katranidis, A., Bukau, B., Kramer, G., Tans, S., (2023). Co-translational ribosome pairing enables native assembly of misfolding-prone subunits. *bioRxiv*. 2023.06.30.547139. <https://www.doi.org/10.1101/2023.06.30.547139>.
48. Barakat, S., Ezen, E., Devocioğlu, İ., Gezen, M., Piepoli, S., Erman, B., (2023). Dimerization choice and alternative functions of ZBTB transcription factors. *FEBS J.* <https://doi.org/10.1111/febs.16905>.
49. Natan, E., Endoh, T., Haim-Vilmovsky, L., Flock, T., Chalancon, G., Hopper, J.T.S., Kintszes, B., Horvath, P., Daruka, L., Fekete, G., Pál, C., Papp, B., Oszi, E., Magyar, Z., Marsh, J.A., Elcock, A.H., Babu, M.M., Robinson, C.V., Sugimoto, N., Teichmann, S.A., (2018). Cotranslational protein assembly imposes evolutionary constraints on homomeric proteins. *Nature Struct. Mol. Biol.* **25**, 279–288. <https://doi.org/10.1038/s41594-018-0029-5>.
50. Veitia, R.A., (2007). Exploring the molecular etiology of dominant-negative mutations. *Plant Cell.* **19**, 3843–3851. <https://doi.org/10.1105/tpc.107.055053>.
51. Nicholls, C.D., McLure, K.G., Shields, M.A., Lee, P.W.K., (2002). 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. <https://doi.org/10.1074/jbc.M108815200>.
52. Badonyi, M., Marsh, J.A., (2023). Buffering of genetic dominance by allele-specific protein complex assembly. *Sci. Adv.* **9**, eadf9845. <https://doi.org/10.1126/sciadv.adf9845>.
53. Martin, K.C., Ephrussi, A., (2009). mRNA Localization: Gene Expression in the Spatial Dimension. *Cell.* **136**, 719–730. <https://doi.org/10.1016/j.cell.2009.01.044>.

54. Das, S., Vera, M., Gandin, V., Singer, R.H., Tutucci, E., (2021). Intracellular mRNA transport and localized translation. *Nature Rev. Mol. Cell Biol.*, 1–22. <https://doi.org/10.1038/s41580-021-00356-8>.
55. Bourke, A.M., Schwarz, A., Schuman, E.M., (2023). Decentralizing the Central Dogma: mRNA translation in space and time. *Mol. Cell.* **83**, 452–468. <https://doi.org/10.1016/j.molcel.2022.12.030>.
56. Chouaib, R., Safieddine, A., Pichon, X., Imbert, A., Kwon, O.S., Samacoits, A., Traboulsi, A.-M., Robert, M.-C., Tsanov, N., Coleno, E., Poser, I., Zimmer, C., Hyman, A., Hir, H.L., Zibara, K., Peter, M., Mueller, F., Walter, T., Bertrand, E., (2020). A dual protein-mRNA localization screen reveals compartmentalized translation and widespread co-translational RNA targeting. *Dev. Cell.* **54**, 773–791.e5. <https://doi.org/10.1016/j.devcel.2020.07.010>.
57. Lautier, O., Penzo, A., Rouvière, J.O., Chevreux, G., Collet, L., Loïodice, I., Taddei, A., Devaux, F., Collart, M. A., Palancade, B., (2021). Co-translational assembly and localized translation of nucleoporins in nuclear pore complex biogenesis. *Mol. Cell.* **81**, 2417–2427.e5. <https://doi.org/10.1016/j.molcel.2021.03.030>.
58. Penzo, A., Palancade, B., (2023). Puzzling out nuclear pore complex assembly. *FEBS Lett* **597**, 2705–2727. <https://doi.org/10.1002/1873-3468.14713>.
59. Panasenko, O.O., Somasekharan, S.P., Villanyi, Z., Zagatti, M., Bezrukov, F., Rashpa, R., Cornut, J., Iqbal, J., Longis, M., Carl, S.H., Peña, C., Panse, V.G., Collart, M.A., (2019). Co-translational assembly of proteasome subunits in NOT1-containing assemblasomes. *Nature Struct. Mol. Biol.* **26**, 110–120. <https://doi.org/10.1038/s41594-018-0179-5>.
60. Kassem, S., Villanyi, Z., Collart, M.A., (2017). Not5-dependent co-translational assembly of Ada2 and Spt20 is essential for functional integrity of SAGA. *Nucleic Acids Res.* **45**, 1186–1199. <https://doi.org/10.1093/nar/gkw1059>.
61. Eichel, C.A., Ríos-Pérez, E.B., Liu, F., Jameson, M.B., Jones, D.K., Knickelbine, J.J., Robertson, G.A., (2019). A microtranslatome coordinately regulates sodium and potassium currents in the human heart. *eLife* **8**, e52654. <https://doi.org/10.7554/eLife.52654>.
62. Keene, J.D., (2007). RNA regulons: coordination of post-transcriptional events. *Nature Rev. Genet.* **8**, 533–543. <https://doi.org/10.1038/nrg2111>.
63. Johnson, M.E., Hummer, G., (2011). Nonspecific binding limits the number of proteins in a cell and shapes their interaction networks. *Proc. Natl. Acad. Sci.* **108**, 603–608. <https://doi.org/10.1073/pnas.1010954108>.
64. Seidel, M., Becker, A., Pereira, F., Landry, J.J.M., de Azevedo, N.T.D., Fusco, C.M., Kaindl, E., Romanov, N., Baumbach, J., Langer, J.D., Schuman, E.M., Patil, K.R., Hummer, G., Benes, V., Beck, M., (2022). Co-translational assembly orchestrates competing biogenesis pathways. *Nature Commun.* **13**, 1224. <https://doi.org/10.1038/s41467-022-28878-5>.
65. Singh, N., Bhalla, N., (2020). Moonlighting proteins. *Annu. Rev. Genet.* **54**, 265–285. <https://doi.org/10.1146/annurev-genet-030620-102906>.
66. Diss, G., Dubé, A.K., Boutin, J., Gagnon-Arsenault, I., Landry, C.R., (2013). A systematic approach for the genetic dissection of protein complexes in living cells. *Cell Rep.* **3**, 2155–2167. <https://doi.org/10.1016/j.celrep.2013.05.004>.
67. Adams, M.K., Banks, C.A.S., Thornton, J.L., Kempf, C.G., Zhang, Y., Miah, S., Hao, Y., Sardu, M.E., Killer, M., Hattem, G.L., Murray, A., Katt, M.L., Florens, L., Washburn, M.P., (2020). Differential complex formation via paralogs in the human Sin3 protein interaction network. *Mol. Cell. Proteomics* **19**, 1468–1484. <https://doi.org/10.1074/mcp.RA120.002078>.
68. Connelly, K.E., Dykhuizen, E.C., (2017). Compositional and functional diversity of canonical PRC1 complexes in mammals. *Biochim. Biophys. Acta BBA – Gene Regul. Mech.* **1860**, 233–245. <https://doi.org/10.1016/j.bbagr.2016.12.006>.
69. Fischer, S., Weber, L.M., Liefke, R., (2022). Evolutionary adaptation of the Polycomb repressive complex 2. *Epigenet. Chromatin* **15**, 7. <https://doi.org/10.1186/s13072-022-00439-6>.
70. Helmlinger, D., Tora, L., (2017). Sharing the SAGA. *Trends Biochem. Sci.* **42**, 850–861. <https://doi.org/10.1016/j.tibs.2017.09.001>.
71. Kadoch, C., Crabtree, G.R., (2015). Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics. *Sci. Adv.* **1**, e1500447.
72. Meeks, J.J., Shilatifard, A., (2017). Multiple roles for the MLL/COMPASS family in the epigenetic regulation of gene expression and in cancer. *Annu. Rev. Cancer Biol.* **1**, 425–446. <https://doi.org/10.1146/annurev-cancerbio-050216-034333>.
73. Pezic, D., Weeks, S.L., Hadjir, S., (2017). More to cohesin than meets the eye: complex diversity for fine-tuning of function. *Curr. Opin. Genet. Dev.* **43**, 93–100. <https://doi.org/10.1016/j.gde.2017.01.004>.
74. Reid, X.J., Low, J.K.K., Mackay, J.P., (2023). A NuRD for all seasons. *Trends Biochem. Sci.* **48**, 11–25. <https://doi.org/10.1016/j.tibs.2022.06.002>.
75. Csárdi, G., Franks, A., Choi, D.S., Airoidi, E.M., Drummond, D.A., (2015). Accounting for experimental noise reveals that mRNA levels, amplified by post-transcriptional processes, largely determine steady-state protein levels in yeast. *PLOS Genet.* **11**, e1005206.
76. Lahtvee, P.-J., Sánchez, B.J., Smialowska, A., Kasvandik, S., Elseman, I.E., Gatto, F., Nielsen, J., (2017). Absolute quantification of protein and mRNA abundances demonstrate variability in gene-specific translation efficiency in yeast. *Cell Syst.* **4**, 495–504.e5. <https://doi.org/10.1016/j.cels.2017.03.003>.
77. Marino, J., Buholzer, K.J., Zosel, F., Nettels, D., Schuler, B., (2018). Charge interactions can dominate coupled folding and binding on the ribosome. *Biophys. J.* **115**, 996–1006. <https://doi.org/10.1016/j.bpj.2018.07.037>.
78. Liu, Y., Yang, Q., Zhao, F., (2021). Synonymous but not silent: The codon usage code for gene expression and protein folding. *Annu. Rev. Biochem.* **90**, 375–401. <https://doi.org/10.1146/annurev-biochem-071320-112701>.
79. Novoa, E.M., de Pouplana, L.R., (2012). Speeding with control: codon usage, tRNAs, and ribosomes. *Trends Genet.* **28**, 574–581. <https://doi.org/10.1016/j.tig.2012.07.006>.
80. Ito, K., Chiba, S., (2013). Arrest peptides: Cis-acting modulators of translation. *Annu. Rev. Biochem.* **82**, 171–202. <https://doi.org/10.1146/annurev-biochem-080211-105026>.
81. Kanda, S., Yanagitani, K., Yokota, Y., Esaki, Y., Kohno, K., (2016). Autonomous translational pausing is required

- for XBP1u mRNA recruitment to the ER via the SRP pathway. *Proc. Natl. Acad. Sci.* **113**, E5886–E5895. <https://doi.org/10.1073/pnas.1604435113>.
82. Buskirk, A.R., Green, R., (2017). Ribosome pausing, arrest and rescue in bacteria and eukaryotes. *Philos. Trans. R. Soc. B Biol. Sci.* **372**, 20160183. <https://doi.org/10.1098/rstb.2016.0183>.
 83. Collart, M.A., Weiss, B., (2020). Ribosome pausing, a dangerous necessity for co-translational events. *Nucleic Acids Res.* **48**, 1043–1055. <https://doi.org/10.1093/nar/gkz763>.
 84. Boulon, S., Pradet-Balade, B., Verheggen, C., Molle, D., Boireau, S., Georgieva, M., Azzag, K., Robert, M.-C., Ahmad, Y., Neel, H., Lamond, A.I., Bertrand, E., (2010). HSP90 and its R2TP/prefoldin-like cochaperone are involved in the cytoplasmic assembly of RNA polymerase II. *Mol. Cell.* **39**, 912–924. <https://doi.org/10.1016/j.molcel.2010.08.023>.
 85. Wild, T., Cramer, P., (2012). Biogenesis of multisubunit RNA polymerases. *Trends Biochem. Sci.* **37**, 99–105. <https://doi.org/10.1016/j.tibs.2011.12.001>.
 86. Lin, L., DeMartino, G.N., Greene, W.C., (2000). Cotranslational dimerization of the Rel homology domain of NF- κ B1 generates p50–p105 heterodimers and is required for effective p50 production. *EMBO J.* **19**, 4712–4722. <https://doi.org/10.1093/emboj/19.17.4712>.
 87. Halbach, A., Zhang, H., Wengi, A., Jablonska, Z., Gruber, I.M.L., Halbeisen, R.E., Deh, P.-M., Kemmeren, P., Holstege, F., Géli, V., Gerber, A.P., Dichtl, B., (2009). Cotranslational assembly of the yeast SET1C histone methyltransferase complex. *EMBO J.* **28**, 2959–2970. <https://doi.org/10.1038/emboj.2009.240>.
 88. Duncan, C.D.S., Mata, J., (2011). Widespread cotranslational formation of protein complexes. *PLoS Genet.* **7**, e1002398. <https://doi.org/10.1371/journal.pgen.1002398>.
 89. Yayli, G., Bernardini, A., Mendoza Sanchez, P.K., Scheer, E., Damilot, M., Essabri, K., Morlet, B., Negroni, L., Vincent, S.D., Timmers, H.T.M., Tora, L., (2023). ATAC and SAGA co-activator complexes utilize co-translational assembly, but their cellular localization properties and functions are distinct. *Cell Rep.* **42**, <https://doi.org/10.1016/j.celrep.2023.113099> 113099.
 90. Trowitzsch, S., Viola, C., Scheer, E., Conic, S., Chavant, V., Fournier, M., Papai, G., Ebong, I.-O., Schaffitzel, C., Zou, J., Haffke, M., Rappsilber, J., Robinson, C.V., Schultz, P., Tora, L., Berger, I., (2015). Cytoplasmic TAF2–TAF8–TAF10 complex provides evidence for nuclear holo–TFIID assembly from preformed submodules. *Nature Commun.* **6**, 6011. <https://doi.org/10.1038/ncomms7011>.
 91. Gupta, K., Watson, A.A., Baptista, T., Scheer, E., Chambers, A.L., Koehler, C., Zou, J., Obong-Ebong, I., Kandiah, E., Temblador, A., Round, A., Forest, E., Man, P., Bieniossek, C., Laue, E.D., Lemke, E.A., Rappsilber, J., Robinson, C.V., Devys, D., Tora, L., Berger, I., (2017). Architecture of TAF11/TAF13/TBP complex suggests novel regulation properties of general transcription factor TFIID. *eLife* **6**, e30395. <https://doi.org/10.7554/eLife.30395>.
 92. Chen, J.-L., Attardi, L.D., Verrijzer, C.P., Yokomori, K., Tjian, R., (1994). Assembly of recombinant TFIID reveals differential coactivator requirements for distinct transcriptional activators. *Cell* **79**, 93–105. [https://doi.org/10.1016/0092-8674\(94\)90403-0](https://doi.org/10.1016/0092-8674(94)90403-0).
 93. Chen, X., Qi, Y., Wu, Z., Wang, X., Li, J., Zhao, D., Hou, H., Li, Y., Yu, Z., Liu, W., Wang, M., Ren, Y., Li, Z., Yang, H., Xu, Y., (2021). Structural insights into preinitiation complex assembly on core promoters. *Science* **372**, eaba8490. <https://doi.org/10.1126/science.aba8490>.
 94. Singh, M.V., Bland, C.E., Weil, P.A., (2004). Molecular and genetic characterization of a Taf1p domain essential for yeast TFIID assembly. *Mol. Cell. Biol.* **24**, 4929–4942. <https://doi.org/10.1128/MCB.24.11.4929-4942.2004>.
 95. Verrijzer, C.P., Chen, J.-L., Yokomori, K., Tjian, R., (1995). Binding of TAFs to core elements directs promoter selectivity by RNA polymerase II. *Cell* **81**, 1115–1125. [https://doi.org/10.1016/S0092-8674\(05\)80016-9](https://doi.org/10.1016/S0092-8674(05)80016-9).
 96. Weinzierl, R.O.J., Dynlacht, B.D., Tjian, R., (1993). Largest subunit of Drosophila transcription factor IID directs assembly of a complex containing TBP and a coactivator. *Nature* **362**, 511–517. <https://doi.org/10.1038/362511a0>.
 97. Mackmull, M.-T., Klaus, B., Heinze, I., Chokkalingam, M., Beyer, A., Russell, R.B., Ori, A., Beck, M., (2017). Landscape of nuclear transport receptor cargo specificity. *Mol. Syst. Biol.* **13**, 962. <https://doi.org/10.15252/msb.20177608>.
 98. Zukin, S.A., Marunde, M.R., Popova, I.K., Soczek, K.M., Nogales, E., Patel, A.B., (2022). Structure and flexibility of the yeast NuA4 histone acetyltransferase complex. *eLife* **11**, e81400.
 99. Guo, Z., Chu, C., Lu, Y., Zhang, X., Xiao, Y., Wu, M., Gao, S., Wong, C.C.L., Zhan, X., Wang, C., (2023). Structure of a SIN3–HDAC complex from budding yeast. *Nature Struct. Mol. Biol.* **30**, 753–760. <https://doi.org/10.1038/s41594-023-00975-z>.
 100. Grau, D., Zhang, Y., Lee, C.-H., Valencia-Sánchez, M., Zhang, J., Wang, M., Holder, M., Svetlov, V., Tan, D., Nudler, E., Reinberg, D., Walz, T., Armache, K.-J., (2021). Structures of monomeric and dimeric PRC2: EZH1 reveal flexible modules involved in chromatin compaction. *Nature Commun.* **12**, 714. <https://doi.org/10.1038/s41467-020-20775-z>.
 101. Raisch, T., Chang, C.-T., Levdansky, Y., Muthukumar, S., Raunser, S., Valkov, E., (2019). Reconstitution of recombinant human CCR4–NOT reveals molecular insights into regulated deadenylation. *Nature Commun.* **10**, 3173. <https://doi.org/10.1038/s41467-019-11094-z>.
 102. Herbst, D.A., Esbin, M.N., Louder, R.K., Dugast-Darzacq, C., Dailey, G.M., Fang, Q., Darzacq, X., Tjian, R., Nogales, E., (2021). Structure of the human SAGA coactivator complex. *Nature Struct. Mol. Biol.* **28**, 989–996. <https://doi.org/10.1038/s41594-021-00682-7>.
 103. Elías-Villalobos, A., Toullec, D., Faux, C., Séveno, M., Helmlinger, D., (2019). Chaperone-mediated ordered assembly of the SAGA and NuA4 transcription co-activator complexes in yeast. *Nature Commun.* **10**, 5237. <https://doi.org/10.1038/s41467-019-13243-w>.
 104. Kusch, T., Guelman, S., Abmayr, S.M., Workman, J.L., (2003). Two Drosophila Ada2 homologues function in different multiprotein complexes. *Mol. Cell. Biol.* **23**, 3305–3319. <https://doi.org/10.1128/MCB.23.9.3305-3319.2003>.

105. Muratoglu, S., Georgieva, S., Pápai, G., Scheer, E., Enünlü, I., Komonyi, O., Cserpán, I., Lebedeva, L., Nabirochkina, E., Udvardy, A., Tora, L., Boros, I., (2003). Two different *Drosophila* ADA2 homologues are present in distinct GCN5 histone acetyltransferase-containing complexes. *Mol. Cell. Biol.* **23**, 306–321. <https://doi.org/10.1128/MCB.23.1.306-321.2003>.
106. Lambert, S.A., Jolma, A., Campitelli, L.F., Das, P.K., Yin, Y., Albu, M., Chen, X., Taipale, J., Hughes, T.R., Weirauch, M.T., (2018). The human transcription factors. *Cell* **175**, 598–599. <https://doi.org/10.1016/j.cell.2018.09.045>.
107. Meldal, B.H.M., Bye-A-Jee, H., Gajdoš, L., Hammerová, Z., Horáčková, A., Melicher, F., Peretto, L., Pokorný, D., Lopez, M.R., Türková, A., Wong, E.D., Xie, Z., Casanova, E.B., del-Toro, N., Koch, M., Porras, P., Hermjakob, H., Orchard, S., (2018). Complex Portal 2018: extended content and enhanced visualization tools for macromolecular complexes. *Nucleic Acids Res.* **47** (2019), D550–D558. <https://doi.org/10.1093/nar/gky1001>.