frontiers
in Microbiology

Coupling of Transcription and Translation in Archaea: Cues From the Bacterial World

Albert Weixlbaumer[1,2,3,4], Felix Grünberger[5], Finn Werner[6] and Dina Grohmann[5,7]*

[1]Department of Integrated Structural Biology, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Illkirch, France, [2]Université de Strasbourg, Strasbourg, France, [3]CNRS UMR7104, Illkirch, France, [4]INSERM U1258, Illkirch, France, [5]Institute of Microbiology and Archaea Centre, University of Regensburg, Regensburg, Germany, [6]RNAP Lab, Division of Biosciences, Institute for Structural and Molecular Biology, London, United Kingdom, [7]Regensburg Center for Biochemistry, University of Regensburg, Regensburg, Germany

The lack of a nucleus is the defining cellular feature of bacteria and archaea. Consequently, transcription and translation are occurring in the same compartment, proceed simultaneously and likely in a coupled fashion. Recent cryo-electron microscopy (cryo-EM) and tomography data, also combined with crosslinking-mass spectrometry experiments, have uncovered detailed structural features of the coupling between a transcribing bacterial RNA polymerase (RNAP) and the trailing translating ribosome in *Escherichia coli* and *Mycoplasma pneumoniae*. Formation of this supercomplex, called expressome, is mediated by physical interactions between the RNAP-bound transcription elongation factors NusG and/or NusA and the ribosomal proteins including uS10. Based on the structural conservation of the RNAP core enzyme, the ribosome, and the universally conserved elongation factors Spt5 (NusG) and NusA, we discuss requirements and functional implications of transcription-translation coupling in archaea. We furthermore consider additional RNA-mediated and co-transcriptional processes that potentially influence expressome formation in archaea.

Keywords: RNA polymerase, ribosome, archaea, expressome, Spt4/5, NusG, Nus

INTRODUCTION

The controlled and coordinated expression of genes plays a fundamental role in all cellular life forms and occurs in two steps: transcription of DNA to RNA by RNA polymerase (RNAP) and translation of RNA to protein by the ribosome. Cellular RNAPs share a conserved core architecture (Hirata et al., 2008; Korkhin et al., 2009; Werner and Grohmann, 2011; Jun et al., 2014; Griesenbeck et al., 2017). However, the archaeal RNAP structure, subunit composition, and use of basal transcription factors (TF) are more closely related to eukaryotic RNAP II than the bacterial counterpart. Ribosomes are large ribonucleoprotein particles that consist of two subunits that entail ribosomal proteins (r-proteins) and rRNAs. While the general organization and function of the ribosome is universally conserved, the complexity and protein content of ribosomes increases from bacteria to archaea to eukaryotes (Armache et al., 2013; Yusupova and Yusupov, 2014; Ferreira-Cerca, 2017). In fact, differences in the transcriptional and translational apparatus reflect the increase in complexity during evolution (Armache et al., 2013).

For example, major differences in ribosome subunit composition are already apparent in the four phylogenetically distinct (super-) phyla: Thaumarchaeota, Aigarchaeota, Crenarchaeota, and Korarchaeota (TACK); Euryachaeota; Diapherotrites, Parvarchaeota, Aenigmarchaeota, Nanoarchaeota, and Nanohaloarchaeota (DPANN); and Asgard archaea.

Prokaryotes lack a nucleus, so transcription and translation occur in the same cellular compartment, the cytoplasm. Biochemical evidence and electron micrographs of lysed bacteria led to the early proposal and realization that translation occurs co-transcriptionally (Byrne et al., 1964; Miller et al., 1970). This prompted the question whether coordination or coupling of elongating RNAP with the pioneering ribosome mutually influences transcription and translation. Data from bacteria provided direct evidence that rates of transcription and translation are interdependent, at least in some species and for some transcription units (Landick et al., 1985; Proshkin et al., 2010; Castro-Roa and Zenkin, 2012; Zhu et al., 2019; Johnson et al., 2020; Stevenson-Jones et al., 2020). However, recent work in *Bacillus subtilis* showed that coupling of transcription and translation is not conserved across all bacteria (Johnson et al., 2020). Recently, single-particle cryo-electron microscopy (cryo-EM) and cryo-electron tomography (cryo-ET) was used to elucidate structural details of the coupled bacterial RNAP and ribosome, a macromolecular assembly termed "expressome." It highlighted roles of transcription elongation factors NusG and/ or NusA that physically connect RNAP with the ribosome (Demo et al., 2017; Kohler et al., 2017; O'Reilly et al., 2020; Wang et al., 2020; Webster et al., 2020).

In contrast, little is known about the coupling of transcription and translation in archaea. It is unclear if direct interactions between RNAP and ribosomes occur or if their association is solely mediated by the shared mRNA. Likewise, the contribution and regulatory role of accessory transcription factors is unknown (McGary and Nudler, 2013; Artsimovitch, 2018). Based on the structural information of the bacterial expressome(s), we discuss whether a coupling between the archaeal RNAP and ribosome can take place in a comparable manner. While molecular structures often guide hypotheses about underlying molecular mechanisms, they rarely suffice to provide the complete picture. We discuss how additional functional evidence obtained *in vivo*, including reporter gene assays and systems biology data such as transcriptome analyses and ribosome profiling, can shed light on the coupled gene expression in archaea. Finally, gene expression takes place in the context of other essential physiological processes. Hence, events like RNA processing and degradation, and their impact on transcription, translation, and their coupling are important to consider.

STRUCTURAL INSIGHTS INTO THE BACTERIAL EXPRESSOME

Attempts to gain structural insights into bacterial expressomes were based on two approaches: (i) cryo-EM of samples formed by direct reconstitution of purified *Escherichia coli* components on mRNA substrates, which direct a precise spacing between RNAP and the 70S ribosome (Wang et al., 2020; Webster et al., 2020), or (ii) direct visualization using cryo-ET in combination with in-cell cross-linking mass spectrometry in *Mycoplasma pneumoniae* (O'Reilly et al., 2020). With sufficient mRNA separating the two machineries, RNAP adopts a wide range of orientations, the assembly is highly flexible, and the mRNA is the only consistent connection (**Figure 1A**). In *E. coli*, adding NusG restrains RNAP and aligns the mRNA with the ribosomal helicase (**Figure 1B**), proposed to prevent secondary structure formation in the transcript (Webster et al., 2020). Addition of the TF NusA stabilizes the NusG-coupled expressome (Wang et al., 2020; **Figures 1D,E**). In contrast, in *M. pneumoniae*, NusA alone appears to couple the two machineries without a role for NusG, albeit in a different relative orientation (O'Reilly et al., 2020; **Figure 1F**). This is consistent with the weak sequence conservation in the NusG KOW domain of *E. coli* and *M. pneumoniae* and suggests that a different mechanism for coupling evolved in this minimal genome species.

All three studies concluded that short spacings between RNAP and the ribosome, either directed by the mRNA or by adding a drug to halt RNAP, form expressomes that resemble an earlier lower-resolution reconstruction formed by collision of a translating ribosome with a stalled RNAP (**Figure 1C**; Kohler et al., 2017). Importantly, while RNAP is still mobile in this collided conformation, NusG cannot simultaneously bind RNAP and the ribosome and therefore cannot form a physical link.

While it is tempting to suggest uncoupled, NusG coupled, and collided expressome structures represent a ribosome approaching RNAP (in agreement with a reduction in RNA separating the two machineries), there is no other experimental evidence to support this chronological order of events, and this remains subject for further research.

IS TRANSCRIPTION COUPLED TO TRANSLATION IN ARCHAEA?

For the archaeon *Thermococcus kodakarensis*, DNA-attached polysomes have been visualized by electron microscopy (French et al., 2007) suggesting that transcription translation coupling (TTC) occurs in archaea. Given the bacterial expressome, the question arises whether the archaeal machineries are compatible with this architecture. To answer this question, the bacterial transcription and translation apparatus has to be compared concerning (i) the overall RNAP architecture, (ii) the RNA length bridging the RNAP active site with the ribosomal P-site (carrying the peptidyl-tRNA), (iii) the presence of NusG or NusA-like factors, and (iv) the conservation of interaction surfaces.

In contrast to bacterial RNAPs, archaeal-eukaryotic RNAPs contain subunits Rpo4/7 (the stalk domain), which binds nascent RNA (Todone et al., 2001; Meka et al., 2005) and stimulate RNAP processivity (Hirtreiter et al., 2009) suggesting the stalk guides the RNA away from RNAP once it emerges from the RNA exit channel. Complexes between bacterial RNAP and 70S ribosomes could be observed for RNA spacers as short
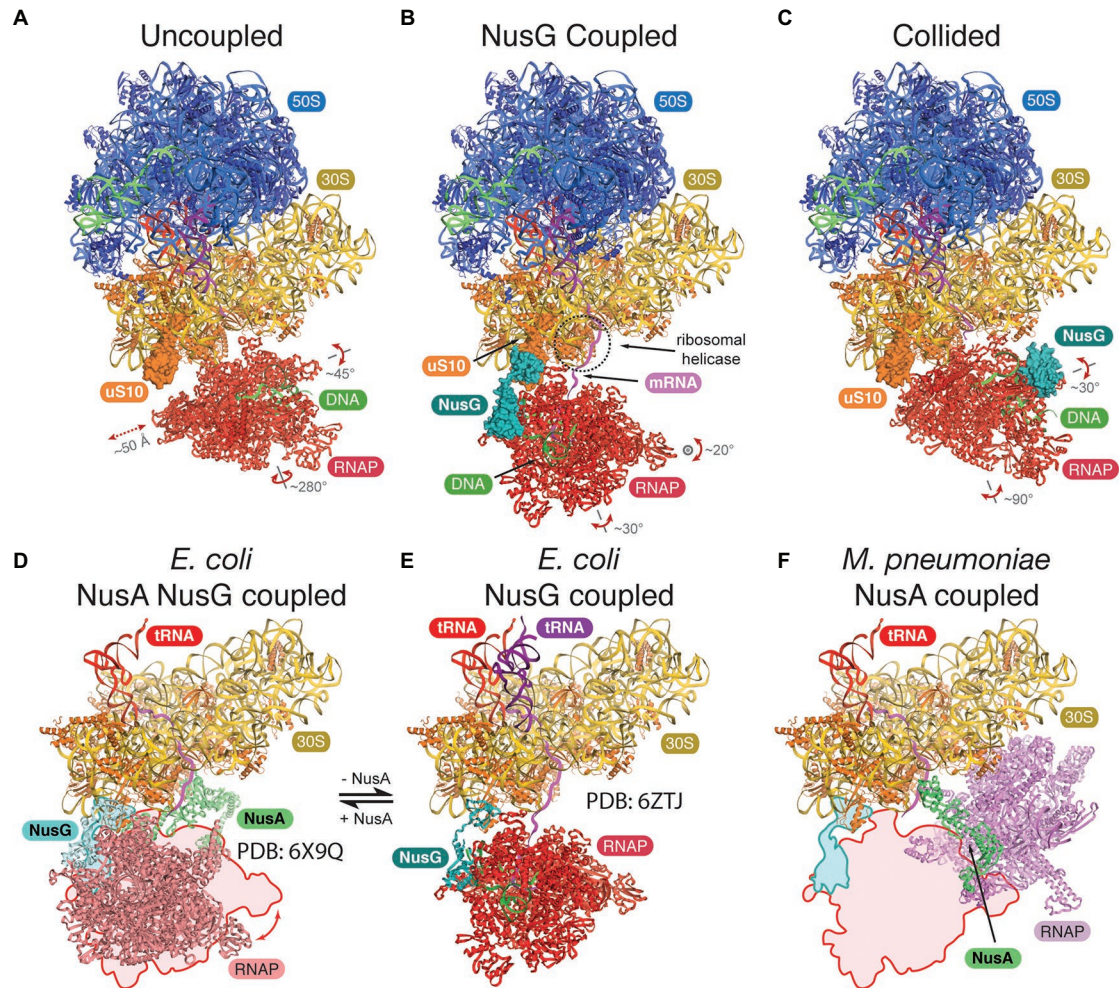
FIGURE 1 | Structures of the bacterial expressomes. **(A)** At mRNA spacings separating the RNA polymerase (RNAP) active site by more than ~35 nucleotides from the ribosomal P-site, and in absence of any coupling factor, RNAP adopts a wide range of orientations relative to the ribosome (uncoupled state; compare extent of translations and rotations). **(B)** Addition of the coupling factor NusG (turquoise) results in the formation of a physical link between RNAP and the ribosome (through uS10, orange) and restrains the rotational freedom. The emerging mRNA transcript (magenta) aligns with the ribosomal helicase (r-proteins uS3 and uS4, dashed circle; NusG coupled state). **(C)** Once the ribosome approaches RNAP further (mRNA spacings less than ~34 nucleotides), the expressome adopts a conformation similar to a previously observed collided state, where NusG can no longer form a physical bridge (collided state). **(D–F)** Comparison of transcription factor coupled states in *Escherichia coli* and *Mycoplasma pneumoniae*. The NusG-coupled state **(E)** gets stabilized by NusA in a similar relative orientation [**D**, compare outline with model, differences are within rotational freedom indicated in panel **(A)**]. In contrast, the NusA-coupled state observed in *M. pneumoniae* requires a major change in position and orientation of RNAP **(F)**, compare outline and model, *E. coli* RNAP and ribosome were docked into deposited cryo-electron tomography (cryo-ET) map (O'Reilly et al., 2020).

as 29 nt separating the RNAP active site from the ribosomal P-site (Wang et al., 2020). However, NusG-mediated coupling appears to be compatible only with spacer lengths greater than at least ~34 nt (Webster et al., 2020). Cryo-EM reconstructions (Bernecky et al., 2016; Ehara et al., 2017) and single-molecule FRET studies (Andrecka et al., 2008) of eukaryotic elongation complexes showed that transcripts of 14–29 nt reach the stalk base. Longer RNAs could not be mapped and appeared to be flexible. This suggests the attachment of longer RNAs to the stalk is transient or they are no longer associated with the stalk. In the context of the archaeal RNAP and assuming that the nascent RNA binds the stalk, for TTC to occur, a longer mRNA segment is required that can traverse the stalk

before being fed into the ribosome in contrast to the bacterial situation. Alternatively, the mRNA might be detached from Rpo4/7 and directly enter the ribosome.

In the *E. coli* expressome, RNA-dependent TTC is further mediated by NusG, which is the only universally conserved TF (Werner, 2012). In archaea and eukaryotes, the NusG homolog is called Spt5 and forms a heterodimer with Spt4. NusG/Spt5 has an N-terminal NGN domain and a C-terminal KOW domain, which bind the RNAP clamp domain and the r-protein uS10, respectively (**Figure 2A**). At the majority of genes, archaeal Spt4/5 associates with the elongation complex proximal to the promoter and reflects the RNAP association pattern (Smollett et al., 2017). This suggests early Spt4/5

recruitment to RNAP even for short transcripts and thus a coupling function may also occur early in transcription. NusG/Spt5 are structurally conserved (**Figures 2B,C**; Hirtreiter et al., 2010; Martinez-Rucobo et al., 2011; Liu and Steitz, 2017). Hence, the interaction interfaces between NusG/Spt5 and RNAP and/or the ribosome might also be conserved.

First, we focus on the NusG-mediated contact between RNAP and ribosome because biochemical data suggest this to be the prevalent arrangement of the expressome *in vivo* (Saxena et al., 2018; Washburn et al., 2020). The binding site of NusG/Spt5 on RNAP is conserved according to structural data in all three kingdoms of life (**Figures 2B,C**; Klein et al., 2011; Martinez-Rucobo et al., 2011; Ehara et al., 2017; Kang et al., 2018). While structural data on archaeal Spt5 (aSpt5) interacting with the archaeal ribosome are missing, the length and mobility of the linker connecting the NGN and KOW domain in aSpt5 resembles NusG. Thus, a similar interaction as observed for the bacterial NusG-coupled expressome is feasible. Furthermore, bacterial RNAP exhibits substantial rotational and translational freedom with respect to the ribosoume even in the NusG-coupled expressome. Modeling of an archaeal expressome based on bacterial RNAP orientations (Webster et al., 2020) shows that most orientations would require a different stalk orientation to avoid steric overlap with the 30S subunit (Coureux et al., 2020). Archaeal RNAP might either be more restricted in its orientation relative to the ribosome or adopt different orientations compatible with the stalk that have been modeled to be possible without steric clashes between RNAP and the ribosome (Kohler et al., 2017).

Bacterial uS10 provides a hydrophobic pocket for the KOW domain of NusG to insert several hydrophobic residues (Webster et al., 2020). Residues V84 and M88 in uS10 form one edge of the hydrophobic pocket in close proximity to F141, F144, and I164 in the NusG-KOW domain (Burmann et al., 2010; Webster et al., 2020). V84 and M88 in uS10 and F141 and I164 in the KOW domain are conserved among bacterial and archaeal proteins (**Figures 2B,C**; Melnikov et al., 2018) suggesting the hydrophobic interaction between NusG/Spt5-KOW and uS10 might be conserved. Moreover, the structure of bacterial and archaeal uS10 is conserved (**Figure 2C**) and residues in the putative interaction surface (β-strand 1 and 4, α-helix 2) of archaeal uS10 with Spt5 are highly conserved among archaeal uS10 proteins suggesting that the amino acid identity might play a role for the function and interaction of archaeal uS10 (Coureux et al., 2020; **Figure 2B**). Despite the conserved phenylalanine residues and overall sequence conservation of aSpt5, organisms of the euryarchaeal and crenarchaeal phylum do not share a high sequence conservation with bacterial KOW sequences.

The interactions in the expressome are not conserved across all bacteria and alternative coupling mechanisms have evolved. In *M. pneumoniae*, the bacterial elongation and termination factor NusA couples RNAP and the ribosome (O'Reilly et al., 2020). Commonly, bacterial NusA proteins contain an N-terminal domain (binds RNAP), and a S1 and two KH domains (bind RNA). *Mycoplasma pneumoniae* NusA contains an additional flexible C-terminal extension not found in *E. coli* or *B. subtilis*,

which contacts multiple r-proteins on the ribosome (**Figure 1D**). Consequently, the relative orientation of the ribosome to RNAP differs significantly from the *E. coli* expressome architecture (**Figure 1D**). NusA is also able to stabilize NusG-coupled expressomes in *E. coli* mediated by one of the KH domains (Wang et al., 2020). NusA-like homologs can be found in all archaeal phyla indicating a widespread distribution of this transcription factor but its function is unclear (Shibata et al., 2007). The domain organization differs significantly from bacterial NusA because archaeal NusA (aNusA) only contains KH domains but lacks the NTD, S1 domain and C-terminal extension that interacts with the ribosome in *M. pneumoniae*. Nevertheless, the structure of the bacterial and archaeal KH domains in NusA are highly conserved (**Figures 2D,E**) and aNusA also binds RNA (Shibata et al., 2007). It has been suggested that the RNAP interaction platform and S1 domain of Rpo7 in conjunction with the two KH domains of aNusA form the domain complement of bacterial NusA (**Figure 2F**; Belogurov and Artsimovitch, 2015; Fouqueau et al., 2018).

It is noteworthy that the archaeal domain of life encompasses highly diverse organisms, of which only a few model organisms have been studied so far. As documented for the bacterial world (Irastortza-Olaziregi and Amster-Choder, 2020; Johnson et al., 2020), expressome formation might occur in some archaeal species but not in others.

CO-TRANSCRIPTIONAL PROCESSES AND TRANSCRIPTOMIC FEATURES AFFECTING TRANSCRIPTION-TRANSLATION COUPLING IN ARCHAEA

The expressome structures illustrate the highly coordinated interplay of two molecular machineries. However, the expressome is not an isolated complex but operates with high specificity in a crowded cytoplasm where myriads of molecular processes occur simultaneously. In archaea, a number of transcriptional and co-transcriptional steps have been identified that might prevent the immediate loading of the ribosome onto the mRNA. Among others, processes like co-transcriptional RNA processing, binding of non-coding RNAs (ncRNAs) to and association of RNA chaperones and transcription termination factors with the RNA may influence expressome formation and will be shortly discussed in this section (compare **Figures 2G–I**).

Coupling of the ribosome to RNAP requires the mRNA to span the distance between the RNAP active site and the ribosomal P-site to provide enough space for both machineries (**Figure 2G**). Typically, regulatory sequences that confer translation initiation are encoded in the 5' untranslated region (5'-UTR). Some archaeal mRNAs have a short 5' UTR or none at all (analyzed for *Haloferax volcanii*, *Thermococcus onnurineus*, *Pyrococcus abyssi*, *Saccharolobus solfataricus*, Heyer et al., 2012; Xu et al., 2012; Beck and Moll, 2018). The relative number of leaderless mRNAs ranges between 1.4 and 72%. The mechanism of mRNA recognition and ribosome association appears to be highly diverse in prokaryotes, and we do not know whether the initiation mechanism influences and correlates with
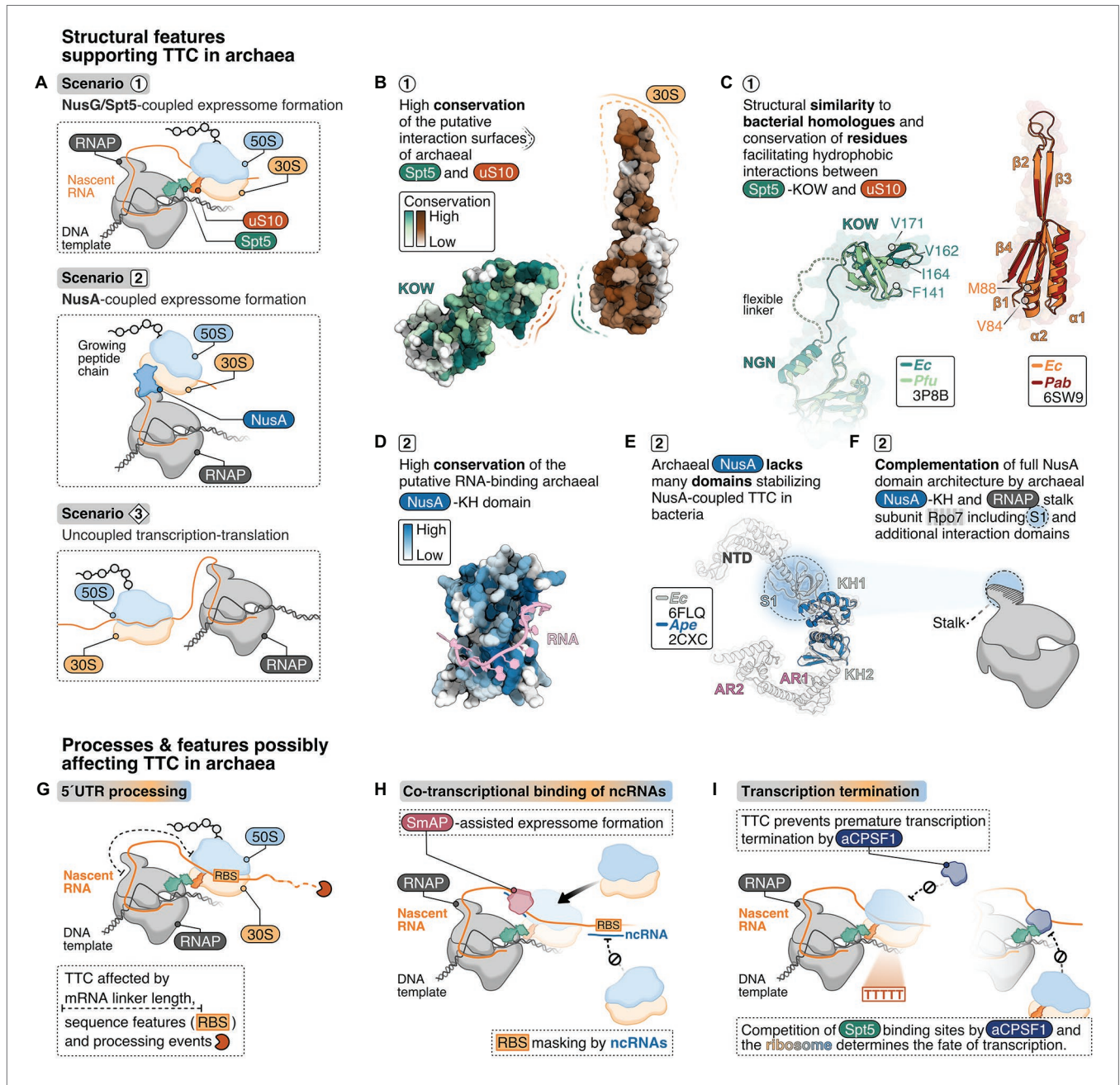
FIGURE 2 | Structural criteria and cellular processes that might mediate and influence transcription-translation coupling in archaea. **(A)** Possible scenarios for transcription translation coupling (TTC) in archaea derived from structural insights into expressome formation in bacteria: NusG/Spt5-coupled expressome formation is a possible scenario 1 representing the archaeal equivalent to the NusG-coupled state in *E. coli* (Webster et al., 2020). In addition, one could also imagine NusA-coupling (scenario 2), similar to what has been observed in *M. pneumoniae* (O'Reilly et al., 2020) or no coupling at all (scenario 3). **(B)** Analysis of conserved regions in archaeal Spt5 and uS10 sequences were aligned, and their conservation score projected color-coded from white (0, not conserved) to dark green or dark-red (9, highly conserved), respectively, on the surface of *Pyrococcus furiosus* Spt5 and *Pyrococcus abyssi* uS10 structure. **(C)** Superimposition of archaeal and bacterial Spt5/NusG and (bacteria: dark-green, PDB: 6ZTJ, *E. coli*; archaea: light-green, PDB: 3P8B, *P. furiosus*) and uS10 (bacteria: orange, PDB: 6ZTJ, *E. coli*; archaea: red, PDB: 6SW9, *P. abyssi*). Residues important for the interaction with NusG are highlighted. **(D)** Conservation analysis of archaeal NusA proteins. Conservation scores from white (0, not conserved) to dark-blue (0, highly conserved) were calculated based on the comparison of archaeal NusA proteins and projected on the surface of *Aeropyrum pernix* NusA. A superimposed model of bacterial RNA (Beuth et al., 2005) is shown in pink. **(E)** The *E. coli* NusA structure (gray, PDB: 6FLQ) overlayed with *A. pernix* NusA (Shibata et al., 2007) (blue, PDB: 2CXC). **(F)** Cartoon depiction of the archaeal RNAP highlighting the hypothesis that the S1 domain of the stalk-forming subunit Rpo7 and the archaeal NusA form a homologue of bacterial NusA. TTC in archaea may be affected by co-transcriptional processes and features depicted in **(G–I)**, including 5'UTR length and processing **(G)**, co-transcriptional binding of non-coding RNAs (ncRNAs) **(H)** and the transcription termination pathway **(I)**.

expressome formation (Wen et al., 2020). mRNAs that lack a ribosomal binding site (RBS) can also emerge from RNA processing events at the 5'-end that lead to cleavage of the 5'-UTR (Qi et al., 2017; **Figure 2G**). As shown for several bacterial (Mäder et al., 2004; Ramirez-Peña et al., 2010; Lioliou et al., 2012) and for the archaeal organisms *Methanocaldococcus jannaschii* and *Marinobacter psychrophilus* (Zhang and Olsen, 2009; Qi et al., 2017), processing of the mRNAs can stabilize transcripts and regulate translation of r-proteins (Qi et al., 2017) and mRNAs from multicistronic operons. In this case, the timing of mRNA processing and translation seems important to avoid conflicts between these two processes.

Co-transcriptional binding of a small regulatory ncRNA to an mRNA is a common posttranscriptional regulation mechanism in prokaryotes that influences RNA stability and translational efficiency of mRNAs in response to changing environmental conditions (Babski et al., 2014; Hör et al., 2018). For *H. volcanii* and *Methanosarcina mazei* small ncRNAs have been detected that can potentially bind to the 5' UTR thereby potentially masking the RBS (Jäger et al., 2009; Soppa et al., 2009; Heyer et al., 2012; Gelsinger and DiRuggiero, 2018; **Figure 2H**). For example, the small RNA_{41} in *M. mazei* binds multiple RBS in a polycistronic mRNA and decouples transcription and translation (Buddeweg et al., 2018).

In bacteria, ncRNA-mRNA hybridization is often mediated by the RNA chaperone Hfq, which belongs to the Sm protein family (Vogel and Luisi, 2011). Hfq can bind RNA co-transcriptionally (Kambara et al., 2018) and plays a role in transcription termination/antitermination (Rabhi et al., 2011; Sedlyarova et al., 2016), ribosome biogenesis (Andrade et al., 2018) and ribosome association with the mRNA in bacteria (Chen et al., 2019). In archaea, a *bona fide* Hfq protein is rarely encoded. More often, single or multiple genes encode an archaeal Sm-like protein (SmAP; Reichelt et al., 2018). Similar to bacterial Hfq, archaeal SmAPs were shown to bind RNAs (Nielsen et al., 2007; Fischer et al., 2010; Märtens et al., 2015). Hence, co-transcriptional association of a ncRNA at the 5' UTR (potentially supported by a SmAP) would prevent ribosome association with the 5' UTR (**Figure 2H**). Co-immunoprecipitation experiments showed that SmAPs not only bind RNAs but also r-proteins (Fischer et al., 2010). It is conceivable that SmAPs participate in posttranscriptional regulation, translation, or act as a bridging factor to recruit ribosomes to the mRNA (**Figure 2H**).

Lastly, the transcription termination pathway might be decisive whether TTC can occur, or vice versa (**Figure 2I**). In archaea, transcription terminates *via* two mechanisms that are not necessarily mutually exclusive: (i) intrinsic termination at poly(U) stretches (Santangelo and Reeve, 2006; Hirtreiter et al., 2009, 2010; Santangelo et al., 2009; Dar et al., 2016; Berkemer et al., 2020) or (ii) factor-dependent termination assisted by the archaeal termination factor aCPSF1/FttA that binds the nascent RNA (Sanders et al., 2020; Yue et al., 2020). Importantly, aCPSF1 also enhances termination at poly(U) stretches. Termination *via* aCSPF1 involves cleavage of the transcript at the 3'-end. In *Methanococcus maripaludis* deletion of aCPSF1 resulted in altered expression levels for the

majority of genes (Yue et al., 2020). Furthermore, aCPSF1-dependent termination gets stimulated by the presence of the stalk domain and Spt4/5 (Sanders et al., 2020). Even though a direct interaction between aCPSF1 and the stalk or Spt4/5 has not yet been experimentally verified, a physical interaction is likely and would be consistent with the observed increased termination efficiency. It is tempting to speculate that aCPSF1 and the ribosome interact with RNAP-bound Spt4/5 in a mutually exclusive fashion similar to Rho and the ribosome with RNAP-bound NusG in bacteria. As a consequence, transcription termination and ribosome coupling might be mutually exclusive. Ribosomes coupled to RNAP *via* Spt4/5 would prevent aCPSF1 interactions with the nascent RNA and prevent premature termination (**Figure 2H**). Alternatively, once aCPSF1 gains access to Spt4/5 it may interfere with TTC (**Figure 2H**). This would be reminiscent of the recruitment of Rho by NusG-KOW to RNAP leading to transcription termination of non-coding/untranslated RNA transcripts (Washburn et al., 2020). Whether TTC or termination prevails could be gene- or operon-specific, could be a target for regulation, and may vary from species to species. Direct, mRNA-independent interactions between the bacterial RNAP and ribosome have been shown. It is possible that in some instances, e.g., during transcription of short mRNAs, the archaeal ribosome might bind the mRNA close to the RNAP exit channel and direct contacts between the elongating RNAP and the ribosome (Wang et al., 2020; Webster et al., 2020).

FUTURE PERSPECTIVES

Are transcription and translation coupled in archaea similar to bacteria? We propose that this is likely, but definitive proof is still lacking. This problem can only be solved by a multidisciplinary effort that reaches beyond a molecular-structural analysis *in vitro*. In order to rationalize the underlying molecular mechanisms, we do need to understand the structural determinants of the RNAP-ribosome interactions and the potential role of general regulatory factors including NusG-Spt4/5 and NusA, as well as ribosomal proteins. A crucial question to be solved in the future is how co-transcriptional processes like SmAP binding, transcription termination, or RNA processing are coordinated with ribosome coupling in space and time. This also includes coordination of translation initiation and TTC. In bacteria, translation initiation was delineated in great detail showing that the 30S subunit is recruited to the mRNA with the help of the initiator tRNA and initiation factors before the 50S subunit joins to form the translation-competent ribosome (see for example, Milon et al., 2010; Tsai et al., 2012). In archaea, the situation is more complex as additional (eukaryotic-like) initiation factors are involved (Benelli et al., 2016). Nonetheless, even in the archaeal initiation complex, uS10 remains exposed and might be available for coupling to Spt5 (Coureux et al., 2020). Consequently, the 30S subunit is involved in translation initiation and coupling to RNAP, and it has to be seen whether these processes are compatible or mutually exclusive.

To elaborate on the finer points of biologically relevant interaction networks a combination of cross-linking/mass spectrometry experiments like the recent elegant study of Rappsilber and colleagues are necessary (O'Reilly et al., 2020). Complementary to these efforts are structural biology, functional genomics, and systems biology approaches that hold great promise to ascertain (i) to which extent the coupling applies to all transcription units or whether it is limited to specific subset or classes of operons, and (ii) whether the coupling-uncoupling is a dynamic process and dependent on environmental cues and stresses, i.e., whether it is subject to regulation. Key to this approach are experiments that monitor changes in the global characteristics of transcription, such as genome-wide RNAP occupancy profiles and transcriptome analyses, in response to perturbations of translation by using ribosome inhibitors/antibiotics or ribosome variants. We have to develop high-resolution methods that combine ribo-seq/proteomics and RNAP NET-seq or ChIP-exo/transcriptomics and integrate the data to obtain a complete view of the interdependence of transcription and translation. Finally, it is important to note that archaea are evolutionary diverse and tractable archaeal model organisms are scarce. Despite the conservation of NusG, the molecular mechanisms of transcription that were revealed for Crenarchaea and Euryarchaea are distinct in many ways including the RNAP subunit composition and chromatin structure. Likewise, we only know little about the mechanisms of translation across the archaeal phyla. The properties of their ribosomes are distinct including the molecular mechanisms of translation initiation, which might have an impact on the coupling of the leading ribosome to the RNAP.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Andrade, J. M., Dos Santos, R. F., Chelysheva, I., Ignatova, Z., and Arraiano, C. M. (2018). The RNA-binding protein Hfq is important for ribosome biogenesis and affects translation fidelity. *EMBO J.* 37:e97631. doi: 10.15252/embj.201797631

Andrecka, J., Lewis, R., Brückner, F., Lehmann, E., Cramer, P., and Michaelis, J. (2008). Single-molecule tracking of mRNA exiting from RNA polymerase II. *Proc. Natl. Acad. Sci. U. S. A.* 105, 135–140. doi: 10.1073/pnas.0703815105

Armache, J.-P., Anger, A. M., Márquez, V., Franckenberg, S., Fröhlich, T., Villa, E., et al. (2013). Promiscuous behaviour of archaeal ribosomal proteins: implications for eukaryotic ribosome evolution. *Nucleic Acids Res.* 41, 1284–1293. doi: 10.1093/nar/gks1259

Artsimovitch, I. (2018). Rebuilding the bridge between transcription and translation. *Mol. Microbiol.* 108, 467–472. doi: 10.1111/mmi.13964

Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., et al. (2016). ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* 44, W344–W350. doi: 10.1093/nar/gkw408

Babski, J., Maier, L.-K., Heyer, R., Jaschinski, K., Prasse, D., Jäger, D., et al. (2014). Small regulatory RNAs in archaea. *RNA Biol.* 11, 484–493. doi: 10.4161/rna.28452

Beck, H. J., and Moll, I. (2018). Leaderless mRNAs in the spotlight: ancient but not outdated! *Microbiol. Spectr.* 6, 155–170. doi: 10.1128/microbiolspec. RWR-0016-2017

Belogurov, G. A., and Artsimovitch, I. (2015). Regulation of transcript elongation. *Annu. Rev. Microbiol.* 69, 49–69. doi: 10.1146/annurev-micro-091014-104047

Benelli, D., La Teana, A., and Londei, P. (2016). "Evolution of translational initiation: from archaea to eukarya" in *Evolution of the Protein Synthesis Machinery and Its Regulation*. eds. G. Hernández and R. Jagus (Cham: Springer), 61–79.

Berkemer, S. J., Maier, L.-K., Amman, F., Bernhart, S. H., Wörtz, J., Märkle, P., et al. (2020). Identification of RNA 3' ends and termination sites in Haloferax volcanii. *RNA Biol.* 17, 663–676. doi: 10.1080/15476286.2020.1723328

Bernecky, C., Herzog, F., Baumeister, W., Plitzko, J. M., and Cramer, P. (2016). Structure of transcribing mammalian RNA polymerase II. *Nature* 529, 551–554. doi: 10.1038/nature16482

Beuth, B., Pennell, S., Arnvig, K. B., Martin, S. R., and Taylor, I. A. (2005). Structure of a *Mycobacterium tuberculosis* NusA-RNA complex. *EMBO J.* 24, 3576–3587. doi: 10.1038/sj.emboj.7600829

Buddeweg, A., Sharma, K., Urlaub, H., and Schmitz, R. A. (2018). sRNA41 affects ribosome binding sites within polycistronic mRNAs in Methanosarcina mazei Gö1. *Mol. Microbiol.* 107, 595–609. doi: 10.1111/mmi.13900

Burmann, B. M., Schweimer, K., Luo, X., Wahl, M. C., Stitt, B. L., Gottesman, M. E., et al. (2010). A NusE:NusG complex links transcription and translation. *Science* 328, 501–504. doi: 10.1126/science.1184953

Byrne, R., Levin, J. G., Bladen, H. A., and Nirenberg, M. W. (1964). The in vitro formation of a DNA-ribosome complex. *Proc. Natl. Acad. Sci. U. S. A.* 52, 140–148. doi: 10.1073/pnas.52.1.140

Castro-Roa, D., and Zenkin, N. (2012). In vitro experimental system for analysis of transcription-translation coupling. *Nucleic Acids Res.* 40:e45. doi: 10.1093/nar/gkr1262

Chen, J., Morita, T., and Gottesman, S. (2019). Regulation of transcription termination of small RNAs and by small RNAs: molecular mechanisms and biological functions. *Front. Cell. Infect. Microbiol.* 9:201. doi: 10.3389/fcimb.2019.00201

Coureux, P.-D., Lazennec-Schurdevin, C., Bourcier, S., Mechulam, Y., and Schmitt, E. (2020). Cryo-EM study of an archaeal 30S initiation complex

gives insights into evolution of translation initiation. *Commun. Biol.* 3:58. doi: 10.1038/s42003-020-0780-0

Dar, D., Prasse, D., Schmitz, R. A., and Sorek, R. (2016). Widespread formation of alternative 3′ UTR isoforms via transcription termination in archaea. *Nat. Microbiol.* 1:16143. doi: 10.1038/nmicrobiol.2016.143

Demo, G., Rasouly, A., Vasilyev, N., Svetlov, V., Loveland, A. B., Diaz-Avalos, R., et al. (2017). Structure of RNA polymerase bound to ribosomal 30S subunit. *Elife* 6:e28560. doi: 10.7554/eLife.28560

Ehara, H., Yokoyama, T., Shigematsu, H., Yokoyama, S., Shirouzu, M., and Sekine, S.-I. (2017). Structure of the complete elongation complex of RNA polymerase II with basal factors. *Science* 357, 921–924. doi: 10.1126/science.aan8552

Ferreira-Cerca, S. (2017). "Life and death of ribosomes in archaea" in *RNA Metabolism and Gene Expression in Archaea*. ed. B. Clouet-d'Orval (Cham: Springer), 129–158.

Fischer, S., Benz, J., Späth, B., Maier, L.-K., Straub, J., Granzow, M., et al. (2010). The archaeal Lsm protein binds to small RNAs. *J. Biol. Chem.* 285, 34429–34438. doi: 10.1074/jbc.M110.118950

Fouqueau, T., Blombach, F., Cackett, G., Carty, A. E., Matelska, D. M., Ofer, S., et al. (2018). The cutting edge of archaeal transcription. *Emerg. Top. Life Sci.* 2, 517–533. doi: 10.1042/ETLS20180014

French, S. L., Santangelo, T. J., Beyer, A. L., and Reeve, J. N. (2007). Transcription and translation are coupled in archaea. *Mol. Biol. Evol.* 24, 893–895. doi: 10.1093/molbev/msm007

Gelsinger, D. R., and DiRuggiero, J. (2018). The non-coding regulatory RNA revolution in archaea. *Genes* 9:141. doi: 10.3390/genes9030141

Griesenbeck, J., Tschochner, H., and Grohmann, D. (2017). Structure and function of RNA polymerases and the transcription machineries. *Subcell. Biochem.* 83, 225–270. doi: 10.1007/978-3-319-46503-6_9

Heyer, R., Dörr, M., Jellen-Ritter, A., Späth, B., Babski, J., Jaschinski, K., et al. (2012). High throughput sequencing reveals a plethora of small RNAs including tRNA derived fragments in Haloferax volcanii. *RNA Biol.* 9, 1011–1018. doi: 10.4161/rna.20826

Hirata, A., Klein, B. J., and Murakami, K. S. (2008). The X-ray crystal structure of RNA polymerase from archaea. *Nature* 451, 851–854. doi: 10.1038/nature06530

Hirtreiter, A., Damsma, G. E., Cheung, A. C. M., Klose, D., Grohmann, D., Vojnic, E., et al. (2010). Spt4/5 stimulates transcription elongation through the RNA polymerase clamp coiled-coil motif. *Nucleic Acids Res.* 38, 4040–4051. doi: 10.1093/nar/gkq135

Hirtreiter, A., Grohmann, D., and Werner, F. (2009). Molecular mechanisms of RNA polymerase—the F/E (RPB4/7) complex is required for high processivity in vitro. *Nucleic Acids Res.* 38, 585–596. doi: 10.1093/nar/gkp928

Hör, J., Gorski, S. A., and Vogel, J. (2018). Bacterial RNA biology on a genome scale. *Mol. Cell* 70, 785–799. doi: 10.1016/j.molcel.2017.12.023

Irastortza-Olaziregi, M., and Amster-Choder, O. (2020). Coupled transcription-translation in prokaryotes: an old couple with new surprises. *Front. Microbiol.* 11:624830. doi: 10.3389/fmicb.2020.624830

Jäger, D., Sharma, C. M., Thomsen, J., Ehlers, C., Vogel, J., and Schmitz, R. A. (2009). Deep sequencing analysis of the Methanosarcina mazei Gö1 transcriptome in response to nitrogen availability. *Proc. Natl. Acad. Sci. U. S. A.* 106, 21878–21882. doi: 10.1073/pnas.0909051106

Johnson, G. E., Lalanne, J.-B., Peters, M. L., and Li, G.-W. (2020). Functionally uncoupled transcription-translation in Bacillus subtilis. *Nature* 585, 124–128. doi: 10.1038/s41586-020-2638-5

Jun, S.-H., Hirata, A., Kanai, T., Santangelo, T. J., Imanaka, T., and Murakami, K. S. (2014). The X-ray crystal structure of the euryarchaeal RNA polymerase in an open-clamp configuration. *Nat. Commun.* 5:5132. doi: 10.1038/ncomms6132

Kambara, T. K., Ramsey, K. M., and Dove, S. L. (2018). Pervasive targeting of nascent transcripts by Hfq. *Cell Rep.* 23, 1543–1552. doi: 10.1016/j.celrep.2018.03.134

Kang, J. Y., Mooney, R. A., Nedialkov, Y., Saba, J., Mishanina, T. V., Artsimovitch, I., et al. (2018). Structural basis for transcript elongation control by NusG family universal regulators. *Cell* 173, 1650.e14–1662.e14. doi: 10.1016/j.cell.2018.05.017

Klein, B. J., Bose, D., Baker, K. J., Yusoff, Z. M., Zhang, X., and Murakami, K. S. (2011). RNA polymerase and transcription elongation factor Spt4/5 complex structure. *Proc. Natl. Acad. Sci. U. S. A.* 108, 546–550. doi: 10.1073/pnas.1013828108

Kohler, R., Mooney, R. A., Mills, D. J., Landick, R., and Cramer, P. (2017). Architecture of a transcribing-translating expressome. *Science* 356, 194–197. doi: 10.1126/science.aal3059

Korkhin, Y., Unligil, U. M., Littlefield, O., Nelson, P. J., Stuart, D. I., Sigler, P. B., et al. (2009). Evolution of complex RNA polymerases: the complete archaeal RNA polymerase structure. *PLoS Biol.* 7:e1000102. doi: 10.1371/journal.pbio.1000102

Landick, R., Carey, J., and Yanofsky, C. (1985). Translation activates the paused transcription complex and restores transcription of the trp operon leader region. *Proc. Natl. Acad. Sci. U. S. A.* 82, 4663–4667. doi: 10.1073/pnas.82.14.4663

Lioliou, E., Sharma, C. M., Caldelari, I., Helfer, A.-C., Fechter, P., Vandenesch, F., et al. (2012). Global regulatory functions of the *Staphylococcus aureus* endoribonuclease III in gene expression. *PLoS Genet.* 8:e1002782. doi: 10.1371/journal.pgen.1002782

Liu, B., and Steitz, T. A. (2017). Structural insights into NusG regulating transcription elongation. *Nucleic Acids Res.* 45, 968–974. doi: 10.1093/nar/gkw1159

Mäder, U., Hennig, S., Hecker, M., and Homuth, G. (2004). Transcriptional organization and posttranscriptional regulation of the Bacillus subtilis branched-chain amino acid biosynthesis genes. *J. Bacteriol.* 186, 2240–2252. doi: 10.1128/JB.186.8.2240-2252.2004

Märtens, B., Bezerra, G. A., Kreuter, M. J., Grishkovskaya, I., Manica, A., Arkhipova, V., et al. (2015). The heptameric SmAP1 and SmAP2 proteins of the crenarchaeon sulfolobus solfataricus bind to common and distinct RNA targets. *Life* 5, 1264–1281. doi: 10.3390/life5021264

Martinez-Rucobo, F. W., Sainsbury, S., Cheung, A. C. M., and Cramer, P. (2011). Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity. *EMBO J.* 30, 1302–1310. doi: 10.1038/emboj.2011.64

McGary, K., and Nudler, E. (2013). RNA polymerase and the ribosome: the close relationship. *Curr. Opin. Microbiol.* 16, 112–117. doi: 10.1016/j.mib.2013.01.010

Meka, H., Werner, F., Cordell, S. C., Onesti, S., and Brick, P. (2005). Crystal structure and RNA binding of the Rpb4/Rpb7 subunits of human RNA polymerase II. *Nucleic Acids Res.* 33, 6435–6444. doi: 10.1093/nar/gki945

Melnikov, S., Manakongtreecheep, K., and Söll, D. (2018). Revising the structural diversity of ribosomal proteins across the three domains of life. *Mol. Biol. Evol.* 35, 1588–1598. doi: 10.1093/molbev/msy021

Miller, O. L., Hamkalo, B. A., and Thomas, C. A. (1970). Visualization of bacterial genes in action. *Science* 169, 392–395. doi: 10.1126/science.169.3943.392

Milon, P., Carotti, M., Konevega, A. L., Wintermeyer, W., Rodnina, M. V., and Gualerzi, C. O. (2010). The ribosome-bound initiation factor 2 recruits initiator tRNA to the 30S initiation complex. *EMBO Rep.* 11, 312–316. doi: 10.1038/embor.2010.12

Nielsen, J. S., Bøggild, A., Andersen, C. B. F., Nielsen, G., Boysen, A., Brodersen, D. E., et al. (2007). An Hfq-like protein in archaea: crystal structure and functional characterization of the Sm protein from Methanococcus jannaschii. *RNA* 13, 2213–2223. doi: 10.1261/rna.689007

O'Reilly, F. J., Xue, L., Graziadei, A., Sinn, L., Lenz, S., Tegunov, D., et al. (2020). In-cell architecture of an actively transcribing-translating expressome. *Science* 369, 554–557. doi: 10.1126/science.abb3758

Proshkin, S., Rahmouni, A. R., Mironov, A., and Nudler, E. (2010). Cooperation between translating ribosomes and RNA polymerase in transcription elongation. *Science* 328, 504–508. doi: 10.1126/science.1184939

Qi, L., Yue, L., Feng, D., Qi, F., Li, J., and Dong, X. (2017). Genome-wide mRNA processing in methanogenic archaea reveals post-transcriptional regulation of ribosomal protein synthesis. *Nucleic Acids Res.* 45, 7285–7298. doi: 10.1093/nar/gkx454

Rabhi, M., Espéli, O., Schwartz, A., Cayrol, B., Rahmouni, A. R., Arluison, V., et al. (2011). The Sm-like RNA chaperone Hfq mediates transcription antitermination at Rho-dependent terminators. *EMBO J.* 30, 2805–2816. doi: 10.1038/emboj.2011.192

Ramirez-Peña, E., Treviño, J., Liu, Z., Perez, N., and Sumby, P. (2010). The group A Streptococcus small regulatory RNA FasX enhances streptokinase activity by increasing the stability of the ska mRNA transcript. *Mol. Microbiol.* 78, 1332–1347. doi: 10.1111/j.1365-2958.2010.07427.x

Reichelt, R., Grohmann, D., and Willkomm, S. (2018). A journey through the evolutionary diversification of archaeal Lsm and Hfq proteins. *Emerg. Top. Life Sci.* 2, 647–657. doi: 10.1042/ETLS20180034

Sanders, T. J., Wenck, B. R., Selan, J. N., Barker, M. P., Trimmer, S. A., Walker, J. E., et al. (2020). FttA is a CPSF73 homologue that terminates transcription in archaea. *Nat. Microbiol.* 5, 545–553. doi: 10.1038/s41564-020-0667-3

Santangelo, T. J., Cubonová, L. U., Skinner, K. M., and Reeve, J. N. (2009). Archaeal intrinsic transcription termination in vivo. *J. Bacteriol.* 191, 7102–7108. doi: 10.1128/JB.00982-09

Santangelo, T. J., and Reeve, J. N. (2006). Archaeal RNA polymerase is sensitive to intrinsic termination directed by transcribed and remote sequences. *J. Mol. Biol.* 355, 196–210. doi: 10.1016/j.jmb.2005.10.062

Saxena, S., Myka, K. K., Washburn, R., Costantino, N., Court, D. L., and Gottesman, M. E. (2018). *Escherichia coli* transcription factor NusG binds to 70S ribosomes. *Mol. Microbiol.* 108, 495–504. doi: 10.1111/mmi.13953

Sedlyarova, N., Shamovsky, I., Bharati, B. K., Epshtein, V., Chen, J., Gottesman, S., et al. (2016). sRNA-mediated control of transcription termination in *E. coli*. *Cell* 167, 111.e13–121.e13. doi: 10.1016/j.cell.2016.09.004

Shibata, R., Bessho, Y., Shinkai, A., Nishimoto, M., Fusatomi, E., Terada, T., et al. (2007). Crystal structure and RNA-binding analysis of the archaeal transcription factor NusA. *Biochem. Biophys. Res. Commun.* 355, 122–128. doi: 10.1016/j.bbrc.2007.01.119

Smollett, K., Blombach, F., Reichelt, R., Thomm, M., and Werner, F. (2017). A global analysis of transcription reveals two modes of Spt4/5 recruitment to archaeal RNA polymerase. *Nat. Microbiol.* 2:17021. doi: 10.1038/nmicrobiol.2017.21

Soppa, J., Straub, J., Brenneis, M., Jellen-Ritter, A., Heyer, R., Fischer, S., et al. (2009). Small RNAs of the halophilic archaeon Haloferax volcanii. *Biochem. Soc. Trans.* 37, 133–136. doi: 10.1042/BST0370133

Stevenson-Jones, F., Woodgate, J., Castro-Roa, D., and Zenkin, N. (2020). Ribosome reactivates transcription by physically pushing RNA polymerase out of transcription arrest. *Proc. Natl. Acad. Sci. U. S. A.* 117, 8462–8467. doi: 10.1073/pnas.1919985117

Todone, F., Brick, P., Werner, F., Weinzierl, R. O. J., and Onesti, S. (2001). Structure of an archaeal homolog of the eukaryotic RNA polymerase II RPB4/RPB7 complex. *Mol. Cell* 8, 1137–1143. doi: 10.1016/S1097-2765(01)00379-3

Tsai, A., Petrov, A., Marshall, R. A., Korlach, J., Uemura, S., and Puglisi, J. D. (2012). Heterogeneous pathways and timing of factor departure during translation initiation. *Nature* 487, 390–393. doi: 10.1038/nature11172

Vogel, J., and Luisi, B. F. (2011). Hfq and its constellation of RNA. *Nat. Rev. Microbiol.* 9, 578–589. doi: 10.1038/nrmicro2615

Wang, C., Molodtsov, V., Firlar, E., Kaelber, J., Blaha, G., Su, M., et al. (2020). Structural basis of transcription-translation coupling. *Science* 369, 1359–1365. doi: 10.1126/science.abb5317

Washburn, R. S., Zuber, P. K., Sun, M., Hashem, Y., Shen, B., Li, W., et al. (2020). *Escherichia coli* NusG links the lead ribosome with the transcription elongation complex. *iScience* 23:101352. doi: 10.1016/j.isci.2020.101352

Webster, M. W., Takacs, M., Zhu, C., Vidmar, V., Eduljee, A., Abdelkareem, M., et al. (2020). Structural basis of transcription-translation coupling and collision in bacteria. *Science* 369, 1355–1359. doi: 10.1126/science.abb5036

Wen, J.-D., Kuo, S.-T., and Chou, H.-H. D. (2020). The diversity of Shine-Dalgarno sequences sheds light on the evolution of translation initiation. *RNA Biol.* 1–12. doi: 10.1080/15476286.2020.1861406

Werner, F. (2012). A nexus for gene expression-molecular mechanisms of Spt5 and NusG in the three domains of life. *J. Mol. Biol.* 417, 13–27. doi: 10.1016/j.jmb.2012.01.031

Werner, F., and Grohmann, D. (2011). Evolution of multisubunit RNA polymerases in the three domains of life. *Nat. Rev. Microbiol.* 9, 85–98. doi: 10.1038/nrmicro2507

Xu, N., Li, Y., Zhao, Y.-T., Guo, L., Fang, Y.-Y., Zhao, J.-H., et al. (2012). Identification and characterization of small RNAs in the hyperthermophilic archaeon Sulfolobus solfataricus. *PLoS One* 7:e35306. doi: 10.1371/journal.pone.0053387

Yue, L., Li, J., Zhang, B., Qi, L., Li, Z., Zhao, F., et al. (2020). The conserved ribonuclease aCPSF1 triggers genome-wide transcription termination of archaea via a 3'-end cleavage mode. *Nucleic Acids Res.* 48, 9589–9605. doi: 10.1093/nar/gkaa702

Yusupova, G., and Yusupov, M. (2014). High-resolution structure of the eukaryotic 80S ribosome. *Annu. Rev. Biochem.* 83, 467–486. doi: 10.1146/annurev-biochem-060713-035445

Zhang, J., and Olsen, G. J. (2009). Messenger RNA processing in Methanocaldococcus (Methanococcus) jannaschii. *RNA* 15, 1909–1916. doi: 10.1261/rna.1715209

Zhu, M., Mori, M., Hwa, T., and Dai, X. (2019). Disruption of transcription-translation coordination in *Escherichia coli* leads to premature transcriptional termination. *Nat. Microbiol.* 4, 2347–2356. doi: 10.1038/s41564-019-0543-1