Taylor & Francis
Taylor & Francis Group

REVIEW

🔓 OPEN ACCESS | Check for updates

Transcription complexes as RNA chaperones

Nelly Said [ID][a] and Markus C. Wahl [ID][a,b]

[a]Freie Universität Berlin, Department Biology, Chemistry, Pharmacy, Institute of Chemistry and Biochemistry, Laboratory of Structural Biochemistry, Berlin, Germany; [b]Helmholtz-Zentrum Berlin Für Materialien Und Energie, Macromolecular Crystallography, Berlin, Germany

ABSTRACT

To exert their functions, RNAs adopt diverse structures, ranging from simple secondary to complex tertiary and quaternary folds. *In vivo*, RNA folding starts with RNA transcription, and a wide variety of processes are coupled to co-transcriptional RNA folding events, including the regulation of fundamental transcription dynamics, gene regulation by mechanisms like attenuation, RNA processing or ribonucleoprotein particle formation. While co-transcriptional RNA folding and associated co-transcriptional processes are by now well accepted as pervasive regulatory principles in all organisms, investigations into the role of the transcription machinery in co-transcriptional folding processes have so far largely focused on effects of the order in which RNA regions are produced and of transcription kinetics. Recent structural and structure-guided functional analyses of bacterial transcription complexes increasingly point to an additional role of RNA polymerase and associated transcription factors in supporting co-transcriptional RNA folding by fostering or preventing strategic contacts to the nascent transcripts. In general, the results support the view that transcription complexes can act as RNA chaperones, a function that has been suggested over 30 years ago. Here, we discuss transcription complexes as RNA chaperones based on recent examples from bacterial transcription.

Introduction

Transcription in bacteria is mediated by a multi-subunit RNA polymerase (RNAP), with minimal core subunit composition $\alpha_2\beta\beta'\omega$, and accessory factors that serve regulatory functions. During transcription, an RNA chain is synthesized 5′-to-3′ in a stepwise manner. In general, RNA synthesis is highly processive with transcription rates of 20 to 80 nucleotides per second [1]. However, RNAP frequently pauses along the DNA template [2], and the average rate of progression can, therefore, vary by several orders of magnitude [3]. At most pauses, RNAP initially adopts a transient, half-translocated, elemental paused state due to sequence-dependent interactions with the DNA and RNA [4–6]. A consensus elemental pause sequence has been delineated [7,8]. An elemental paused elongation complex (EC) can rearrange into a long-lived paused EC when an RNA hairpin invades the RNAP RNA exit channel (class I pause), leading to additional conformational rearrangements that hinder nucleotide addition

[4,5], or when RNAP backtracks on the template (class II pause), extruding the RNA 3′-end into the secondary channel [9]. Recently, a pseudoknot-stabilized pause (*que*/class III pause) in the *Bacillus subtilis* 7-methylamino-7-deazaguanine (preQ$_1$) transcriptional riboswitch that is converted to a pseudoknot-inhibited pause by effector binding has been identified, which shares certain characteristics with hairpin-stabilized pauses [10]. Here, a pseudoknot structure in the nascent RNA stabilizes the paused state and creates a time window for the preQ$_1$ ligand to bind, illustrating that pauses frequently represent sites where regulatory decisions are made. An additional, well-studied example is the bacterial class II *operon polarity suppressor* (*ops*) pause, which enables anti-termination factor RfaH to bind and stabilize the paused EC and to subsequently remodel it into a pause- and termination-resistant EC [11,12].

As in proteins, RNA function hinges on RNAs adopting diverse, and in part complex, 3D structures. It has long been noted that RNAs

CONTACT Nelly Said ✉ nellysaid@zedat.fu-berlin.de; Markus C. Wahl ✉ markus.wahl@fu-berlin.de 🏛 Freie Universität Berlin, Department Biology, Chemistry, Pharmacy, Institute of Chemistry and Biochemistry, Laboratory of Structural Biochemistry , Berlin, Germany
ⓘ Supplemental data for this article can be accessed here

start to fold during transcription. First evidence for co-transcriptional RNA folding was obtained in the early 1980s, when Boyle et al., observed that tRNA folding occurs sequentially 5′-to-3′, and starts before the entire RNA is transcribed [13]. Shortly thereafter, Kramer and Millis discovered the formation of initial meta-stable structures that rearrange into more stable structures while RNA synthesis proceeds [14]. Both the 5′-to-3′ directionality of RNA synthesis, due to the order in which RNA regions are produced, and transcription rates can strongly affect folding of the nascent RNAs. For example, the 5′-end of a nascent RNA can readily adopt local secondary structures before the 3′-end has been synthesized [15]. In addition, RNAP pausing at key positions creates time windows of opportunity for slow-folding RNA structures to form, for proteins to associate with the nascent transcript or for the binding of other ligands [16]. In pioneering studies, folding rates and pathways of the Tetrahymena group I intron were found to strongly depend on the order in which RNA regions are being transcribed [17,18]. While during transcription of wild-type (wt) B. subtilis RNase P RNA and circularly permuted variants by unmodified ECs, the catalytic domain invariably folded at a faster rate than the specificity domain, and folding was not affected by the transcription speed (modulated via the transcribing polymerase or the nucleotide tri-phosphate [NTP] concentrations), the presence of the general elongation factor, N-utilization substance (Nus) A, on Escherichia coli RNAP changed the folding pathway of a circularly permuted RNA variant by accelerating the formation of the specificity domain [19,20]. NusA strongly increased pausing at a specific site located 3′ of the catalytic domain, and the effect was abrogated by employing an RNAP variant that was deficient in pausing [20]. Native RNase P RNA exhibits several long-range helices, and structure probing of the RNA intermediate formed at the pause site revealed the formation of non-native structures that sequester the 5′-portions of these helices [21]. Similar intermediates were detected during folding of E. coli signal recognition particle RNA and transfer-messenger RNA, which likewise adopt long-range helices in

their native structures. Thus, the strategic position of a pause site located between the upstream and downstream portions of long-range helices can foster the formation of meta-stable, non-native structures, that protect the 5′-portions of the helices from being trapped within stable, nonproductive structures [21,22].

Co-transcriptional folding of RNAs can be monitored, e.g., by the emergence of catalytic activities in the case of ribozymes, or via oligonucleotide hybridization in combination with RNase H digestion [20,21,23]. In addition, elaborate techniques have by now been developed to study RNA folding in more detail, and to bridge between in vitro and in vivo RNA folding mechanisms [24–26]. Structure probing in combination with high-throughput RNA sequencing has been used to resolve RNA folding at nucleotide resolution [27,28]. E.g., selective 2′-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-seq) combines in vitro chemical RNA structure probing, in which modifications occur at flexible nucleotides present mostly in unpaired RNA regions, with high-throughput sequencing that detects the precise sites of modification. A recent study combined datasets obtained from SHAPE-seq with computational structure prediction algorithms, which allowed the simulation of co-transcriptional folding pathways [29]. Structural probing of elongating transcripts (SPET-seq) represents a similar approach in vivo [30], complementing other in vivo and genome-wide techniques [31,32].

Single-molecule techniques, such as optical-trapping assays and single-molecule fluorescence resonance energy transfer (smFRET) approaches monitor co-transcriptional RNA folding events in real time and provide information about the formation of native and nonproductive meta-stable RNA intermediate structures and their dependence on transcription rates and pause sites, which previously have escaped detection in folding studies based on pre-formed, full-length transcripts [22,33–35]. Kinetic studies of RNA folding have been primarily performed using T7 RNAP or E. coli RNAP. An alternative approach employs the helicase Rep-X that unwinds an RNA/DNA

hetero-duplex in a kinetically controlled manner, mimicking the stepwise release of RNA from a transcription elongation complex (EC) [36,37].

Based on such investigations, RNA structure formation has been described within a rugged free energy landscape, which allows regions of RNA molecules to rapidly adopt their native structures, while other portions can become kinetically trapped in mis-folded conformations [38]. A population of folding RNA molecules can thus be envisioned as adopting an ensemble of structures located in the minima of the folding free energy landscape that determine their relative abundances [25]. Transcription can strongly influence both the shape of the free energy landscape of RNA folding, and the RNA ensemble at any time of the transcription process, depending on the order in which RNA regions are synthesized, the transcription kinetics and the co-transcriptional interaction of the nascent RNA with proteins and other ligands [25].

While a profound influence of transcription on RNA folding is by now amply documented, co-transcriptionally folding RNAs in turn also signal back to transcription [39]. Co-transcriptional RNA folding can profoundly impact transcription rates, as illustrated by hairpin-stabilized pausing [16]. Likewise, co-transcriptional RNA folding can influence the transcription outcome, e.g., through mechanisms such as attenuation [40] or anti-termination [41]. Furthermore, the ability of all classes of RNAs to fold co-transcriptionally is interwoven with co-transcriptional RNA processing and the assembly of RNA-protein complexes (RNPs) [42]. A number of recent reviews cover the aspects of transcription directionality and kinetics on nascent RNA folding, the impact of co-transcriptional RNA folding on transcription and co-transcriptional processes as well as methods to study these mechanisms [16,42–45].

RNA folding can also be strongly influenced by proteins that transiently bind RNAs [38]. Such RNA chaperones were initially defined as proteins that resolve mis-folded RNA structures in an ATP-independent manner [46,47], but the term has been expanded to also include proteins that modulate RNA folding, e.g., by binding to an RNA and preventing the formation of nonproductive structures, by presenting RNA regions for secondary or higher-order structure formation (annealing) or by remodeling RNA structures in an ATP-dependent manner (unwinding) [48]. RNA chaperones can also guide the assembly of RNA-protein complexes (RNPs) [49]. Reflecting their diverse modes of action, RNA chaperones belong to diverse protein families, including cold shock domain proteins, such as bacterial CspA, that can prevent mis-folding [50]; Sm family proteins, such as bacterial Hfq, that can promote RNA annealing [51]; histone-like proteins, such as bacterial StpA, that can mediate annealing and strand displacement [51]; nucleic acid-dependent NTPases/helicases, such as DEAD-box proteins, that can unwind RNA duplexes [52]; or ribosomal proteins that can guide assembly of ribosomal subunits [49], to name just a few.

Similar to protein chaperones, RNA chaperones are thought to iteratively cycle on and off a client RNA, thereby biasing the folding landscape toward a productive or native fold [48,53]. Various components of transcription complexes (TCs) can establish such transient RNA interactions, and the idea that TCs can act as RNA folding or RNP assembly chaperones has been proposed more than three decades ago [54]. Recent structural and structure-guided functional analyses emphasize such RNA chaperone functions of TCs and suggest specific molecular mechanisms underlying these activities, as we will discuss in the following, resorting to selected examples.

RNA polymerase as an RNA chaperone

Control of RNA entry into the RNA exit channel of RNA polymerase during transcription initiation

RNAP is the first protein to contact the nascent RNA chain, and thus the earliest determinant of the RNA's co-transcriptional folding pathway. A narrow, constrictive RNA exit channel of RNAP is formed by its two largest subunits, β and β'. Structural elements forming the channel include the β flap, β C-terminal domain (CTD) and β' clamp, the latter harboring the RNA-binding modules β' lid, β' zipper, β' zinc-binding domain (ZBD) and β dock (Figure 1). The β flap tip helix (FTH) is located at the rim and constitutes a key regulatory element that depending on its
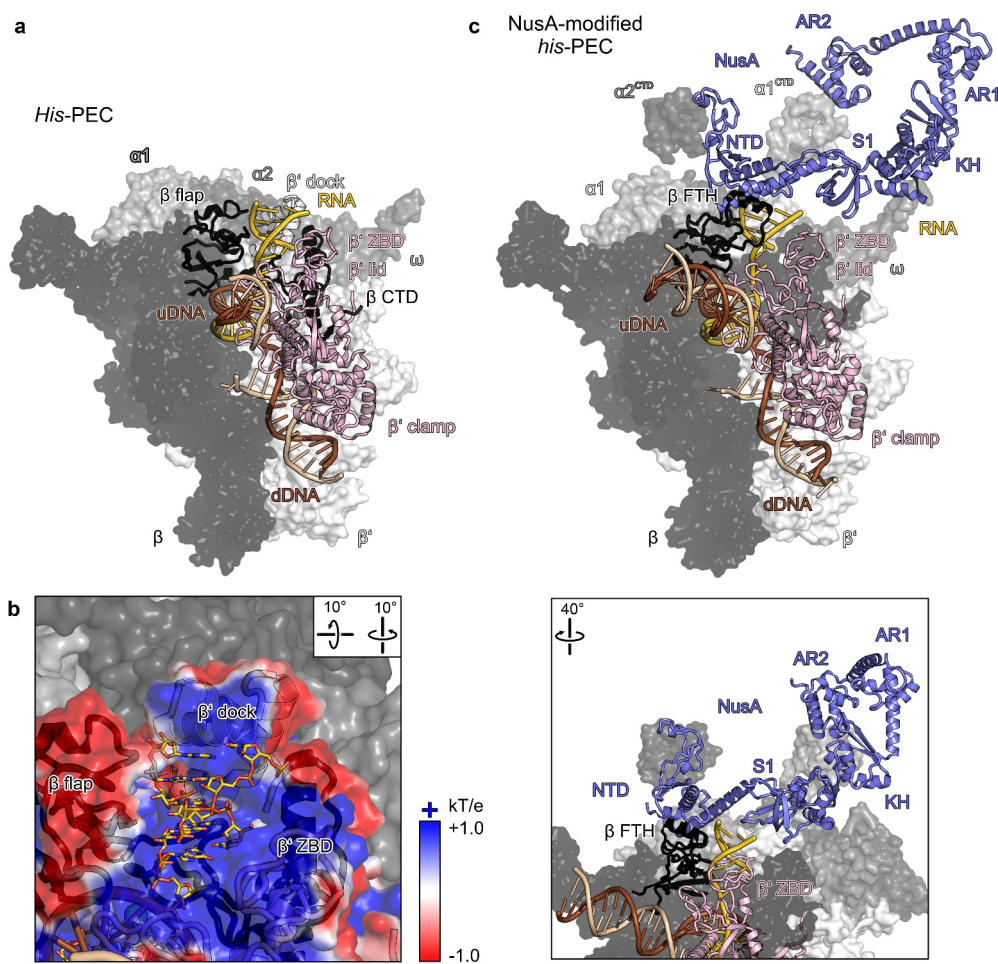
Figure 1. Semi-transparent surface views of the *his*-PEC (PDB ID 6ASX) and of the NusA-modified *his*-PEC (PDB ID 6FLQ). Nucleic acids, NusA and elements of the RNA exit channel are shown as cartoon. Color-coding in this and the following figures: RNAP subunits, different shades of gray; β' clamp, pink; NusA, slate blue; template DNA, brown; non-template DNA, beige; RNA, gold. dDNA, downstream DNA; uDNA, upstream DNA. Key elements are labeled. All subsequent complexes were superimposed based on the β subunits. In this and the following figures, rotation symbols indicate the orientation relative to Figure 1a if not indicated otherwise. **a**. *His*-PEC structure. The RNA exit channel is formed by the β flap, β CTD and β' clamp harboring the β' lid, β' zipper, β' ZBD and β' dock. The *his*-pause hairpin forms an A-form RNA stem along the β flap, β' dock and β CTD, which provide a positively charged surface that complements the negatively charged backbone in the RNA. **b**. Electrostatic surface view of the β and β' subunits of the *his*-PEC structure. Positive potential, blue; negative potential, red. RNA residues are shown as sticks and colored by atom type. Color code: RNA carbon, gold; oxygen, light red; nitrogen, blue; phosphorus, Orange. **c**. NusA-modified *his*-PEC. NusA binds to one of the αCTDs and the β FTH *via* its NTD. A three-helix bundle of NusA NTD and the linker helix connecting the NTD with the S1 domain form direct contacts with the β FTH, thereby stabilizing the β FTH above the RNA exit channel and widening the channel. The S1 domain contacts the β' ZBD, and NusA is further stabilized by interactions of the KH domains with the ω subunit. The AR2 domain interacts with the other αCTD. PyMOL (Schrödinger, LLC) sessions of structure figures are available as supplemental files 1, 2 and 3.

position can modulate the width of the mouth of the RNA exit channel [4,5,55–59]. The RNA exit channel can accommodate both single stranded (ss) and double-stranded (ds) RNA, with ~10 nucleotides of ssRNA and five base-paired nucleotides of dsRNA following the RNA-DNA hybrid [4,5,60]. We will discuss below how the exit channel acts as an RNAP-intrinsic RNA chaperone.

Transcription is initiated by an RNAP holoenzyme formed by association of core RNAP with a σ factor that facilitates promotor-specific DNA binding and unwinding. Initiation is a multi-step process comprising promoter recognition (closed complex formation), initial promoter melting, open complex formation, start of *de novo* RNA synthesis and promoter escape [61]. These steps

are accompanied by structural reorganization within the holo-enzyme leading to the stepwise release of σ, disruption of RNAP interactions with promoter DNA and entry into the elongation phase. Some initial transcribing complexes are subject to abortive initiation, during which RNAP releases the short, initial RNAs and cycles back to the open complex [62]. A pause site encountered after about six nucleotide addition steps marks a branch point for the initial transcribing complex to either escape the promoter or to release the initial RNA [63], and different factors influence abortive initiation, including the promoter sequence, interactions of RNAP with σ and the initial RNA [64]. To allow further RNA synthesis, σ domain (σD) 3.2 has to be displaced as it otherwise blocks RNA entry into the RNA exit channel [61]. Displacement of σD3.2 is driven by collision with the 5′-end of the growing RNA chain, accompanied by expansion of the RNA:DNA hybrid and accommodation of downstream DNA in the RNAP active site cleft (scrunching) [65]. RNA interactions at the exit channel confer stability to the EC and are first established during the transition from transcription initiation to elongation, when the transcript reaches a length of about ten nucleotides [66].

At certain promoters, some ECs are subject to promoter-proximal pausing. Here, σ fails to disengage from the EC after synthesis of about 15–25 nucleotides of RNA due to interactions of σD2 and σD4 with −10-like and −35-like sequences downstream of the promoter, respectively [67–70]. These paused ECs are prone to backtracking and constitute important regulatory intermediates [71,72]. Apart from binding the DNA template, σD4 also binds the β flap at the outer rim of the RNA exit channel in the RNAP holo-enzyme [73] and must be displaced after synthesis of approximately 16 nucleotides during early elongation when the RNA emerges from the RNA exit channel [74,75]. Therefore, promoter-proximal pausing might not only be regulated by σ-DNA interactions but also by σD4-nascent RNA interactions at the RNA exit channel. Taken together, during transcription initiation and early elongation, elements of σ can be considered gate-keepers that control entry of RNA into RNAP's intrinsic RNA chaperone, the exit channel.

RNA polymerase chaperoning simple RNA hairpin structures

The RNA exit channel is allosterically coupled to the RNAP active site, presumably *via* the β connector, a long, two-stranded β-sheet that extends from the β flap to the active site [58]. Thus, RNAP can "sense" nascent RNA structures *via* the exit channel and respond with altered activity. Prime examples for RNA secondary structures that modulate RNAP activity during elongation *via* exit channel-to-active site signaling are hairpin structures on the nascent transcript that can invade the RNA exit channel and stabilize RNAP pausing or lead to intrinsic termination. Hairpin-stabilized pausing has been characterized in diverse bacteria [76–78]. Invasion of the RNAP RNA exit channel by an RNA hairpin formed eleven nucleotides upstream of an elemental pause site in the nascent transcript has been shown to increase pause lifetimes around 10 to 20-fold [58]. During intrinsic termination, RNAP pauses when transcribing a U-rich region; a preceding stable hairpin structure can lead to transcription termination upon invasion of the RNAP RNA exit tunnel [79]. Despite different transcriptional outcomes, hairpin-stabilized pausing and intrinsic termination bear mechanistic similarities. In both cases, effects on RNAP are elicited *via* induced allosteric changes [58] and both processes are enhanced by the general elongation factor, NusA [80,81]. As opposed to hairpin-stabilized pausing [82], formation of the terminator hairpin leads to partial disruption of the DNA:RNA hybrid [83,84].

Monitoring the rate of binding and dissociation of anti-sense RNA oligos complementary to nascent RNA regions located in the exit channel revealed that, despite the steric constraints imposed by the channel, RNA duplex formation in the channel was only moderately reduced as compared to annealing free in solution [85]. Perhaps even more surprisingly, dissociation of an anti-sense oligo annealed within the exit channel was not affected at all [85]. These findings suggest that the exit channel has fine-tuned RNA chaperone activities that allow efficient duplex formation in the exit channel without significantly

stabilizing the channel-embedded duplexes. A recent cryogenic electron microscopy (cryoEM) structure illuminated the configuration of a hairpin-stabilized paused EC derived from the leader sequence in the *his* biosynthetic operon (*his*-PEC) [4] (Figure 1a). The structure revealed molecular details of the conformational rearrangements of RNAP associated with its transition into an off-state, which include swiveling of a group of conserved structural elements (the "swivel module"). As a consequence, nucleotide addition is sterically prevented and the RNA-DNA hybrid is maintained in a half-translocated state, a structural feature of an elemental pause site [6]. The structure also revealed how RNA secondary structures can form within the RNA exit channel without perturbing the boundaries of the channel and supported by the configuration of the channel, contrary to a prior view of RNAP as a rigid body that cannot adopt RNA secondary structures within the RNA exit channel [84]. An A-form RNA duplex can be accommodated based on the arrangement of basic amino acid residues within and right outside the RNA exit channel along the β flap, β CTD and β' dock, which provide positively charged surface patches complementary to the negatively charged backbone of the RNA [4] (Figure 1b). Additionally, the loop of the hairpin and the β FTH might interact outside of the channel, but both structures were not resolved.

Thus, the properties of the RNA exit channel resemble those described for certain RNA chaperone proteins, which are rich in basic amino acid residues and due to electrostatic interactions can modulate the stability and dynamics of RNA structures [38,86,87]. In addition, longer RNA helixes could be extended co-transcriptionally by the emerging duplex corkscrewing along the positive path outside the exit channel, which would be required for the formation of long terminator stem-loops [4]. As the positive charges that line the RNA exit channel and surroundings are conserved among bacterial RNAPs [88] it is likely that all these enzymes can chaperone formation of simple RNA secondary structures within and close to their RNA exit channels.

While a structure of an EC at the verge of intrinsic termination has not yet been obtained, it can be assumed that initial formation of the terminator hairpin in the exit channel is supported by the same RNA chaperoning functions of RNAP as in the case of pause hairpins. However, biochemical studies have suggested that hairpin invasion is followed by conformational changes in RNAP, during which the hairpin visits also other sites within RNAP, eventually leading to EC disruption [89,90]. Thus, additional RNAP elements likely guide the intrinsic terminators during these late steps. Furthermore, both during hairpin stabilized-pausing and intrinsic termination, a third strand (region 5′ of the hairpin formed) has to also be accommodated in the exit channel, a situation that has not yet been structurally characterized in detail. While in cryoEM structures of an unmodified [4] or NusA-modified [5] (see also below) *his*-PEC space for this third strand appears to be available between the β ZBD, C-terminal clamp and N-terminal β' clamp, accommodation of the additional strand may also lead to some opening of the exit channel, possibly depending on how deep the pause/termination hairpin invades the exit channel. Exit channel opening has been observed, e.g., in recent structures of *B. subtilis* and *Mycobacterium smegmatis* RNAPs in complex with the recycling NTPase, HelD [91–93]. HelD does not directly pry open the RNA exit channel, but widens it from a distance by inserting a massive protrusion into the primary channel. Thus, an opening of the RNA exit channel by invading pause or terminator hairpins may likewise affect RNAP allosterically, possibly further supporting adoption of a transcriptional off-state.

RNA polymerase chaperoning more complex RNA structures

Apart from simple hairpins, more complex structures emerging co-transcriptionally in nascent RNA can modulate RNAP transcriptional activity. A prominent example is an RNA-based antitermination system of the lambdoid bacteriophage, HK022. In HK022-infected cells, early gene expression during the lytic life cycle is mediated by ~65-nt RNA elements called polymerase utilization (*put*) sites (Figure 2a) [94]. After its synthesis, *put* modifies RNAP so that the enzyme can read through intrinsic, ρ-dependent and Nun-dependent termination signals [95–97]. The *put*
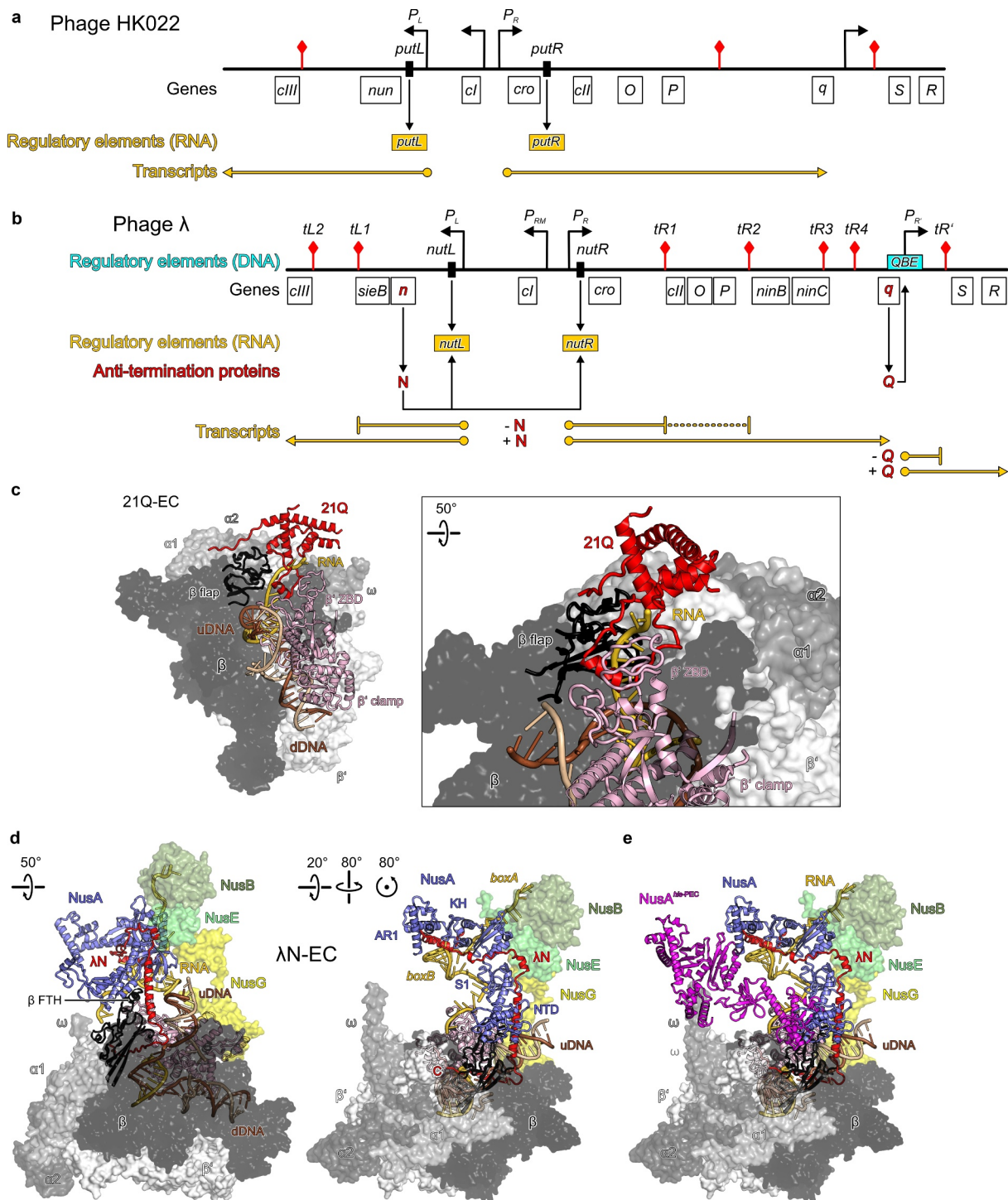
Figure 2. a, b. Schemes of parts of the HK022 (a) and phage λ (b) genomes (black lines) illustrating early and late control regions. Protein-coding genes are indicated in white boxes. Angled arrows along the genomes, promoters; red rhomboid icons, terminators; black boxes, positions of *put* or *nut* sites; yellow boxes, RNA regulatory elements; cyan box, *QBE* regulatory region; N and Q, anti-termination proteins; angled arrows connecting anti-termination proteins and regulatory DNA or RNA elements, recruitment of anti-termination proteins; yellow lines, transcripts. Schemes adapted from [129,130] with changes. **c**. Semi-transparent surface view of the 21Q-EC (PDB ID 6P19). 21Q is shown in cartoon and colored in red. It binds along the outer rim and inside the RNA exit channel (left). The 21Q NTD forms a lasso-like structure that encircles the nascent RNA transcript, sterically preventing secondary structure formation in the exit channel (right). **d**. Semi-transparent surface view of the λN-EC (PDB ID 6GOV). λN is shown in cartoon and colored in red. In this and the following figures: NusB, smudge green; NusE, lime green; NusG, yellow. Upon transcription of a *nut* site, the intrinsically unstructured λN protein builds up a modifying RNP on the surface of RNAP, also comprising NusA, B, E and G. The λN protein directly contacts the *boxB* stem-loop of *nut* RNA *via* its N-terminal ARM domain, meanders around the NusA KH domains and along the NusA NTD-S1 linker helix, folding locally into α-helical structures, to finally enter the RNAP catalytic cavity

next to upstream DNA. The C-terminal portion of λN runs along the DNA:RNA hybrid into the RNA exit channel, thereby restricting the diameter of the channel and altering its electrostatic potential to counteract the formation of RNA secondary structures in the channel. **e**. By acting as a molecular glue, λN repositions NusA from the RNA exit channel through cross strutting with NusB-NusE on the *boxA* element of *nut* RNA, and with the NusG CTD that is flexibly linked to the NusG NTD, located across the RNAP active site cleft. NusA position as observed in a NusA-modified *his*-PEC structure, magenta.

element folds into a double-hairpin structure, which is required for its anti-termination activity [94,98]. Moreover, *put* exerts a local anti-pausing effect at a backtracked pause signal [99]. Anti-termination and anti-pausing activities rely on *put* being anchored to RNAP, most likely involving the β' ZBD [99–101]. How the *put* element exerts its anti-termination and anti-pausing effects is presently unclear, but it has been suggested that both effects rely on distinct mechanisms. The anti-pausing effect is limited to a region in immediate vicinity of the *put*-encoding site, whereas anti-termination remains active distal to the site where *put* is transcribed [99]. Nascent *put* might suppress backtracking by preventing reentry of the nascent RNA due to secondary structure formation close to the RNA exit channel and anchoring to RNAP, and some interactions may be released upon further transcription. To exert its anti-termination effect, *put* needs to stay associated with transcribing RNAP [94,98]. As part of its anti-termination activity, *put* may be bound in vicinity of the RNA exit channel in a way that sterically interferes with the formation of a terminator hairpin [99]. Irrespective of the precise mechanism of *put*-mediated anti-pausing and anti-termination, co-transcriptional folding is a requirement for *put* function [94,98]. It will be interesting to see, e.g., by structural analyses and structure-informed mutagenesis in combination with RNA structure probing, which RNAP regions beyond the RNA exit channel elements may help chaperone *put* into its functional structure(s).

Additional examples of anti-termination systems based on complex RNA structures are afforded by the *B. subtilis EAR* RNA [102] and a recently described ρ-antagonizing RNA element (*RARE*) in *E. coli* [103]. How these RNAs act in detail again remains to be tested. It will be interesting to compare these different RNA-based anti-termination systems not only in terms of the mechanisms that they employ to modulate

RNAP, but also in terms of which RNAP features might aid them in adopting their functional structures co-transcriptionally.

RNAP-based chaperoning of co-transcriptional RNA folding is likely a pervasive principle during riboswitch-dependent transcription control. E.g., using single-molecule FRET in combination with biochemical and simulation approaches, a recent study revealed a crosstalk between the folding of the $preQ_1$ riboswitch and pausing of RNAP (Figure 3) [10]. The aptamer domain of the *B. subtillis* $preQ_1$ riboswitch adopts a pseudoknot structure, and co-transcriptional folding is guided by the *que* (class III) pause located downstream of the aptamer domain. Cross-linking studies identified direct interactions of the pseudoknot with elements of the RNAP RNA exit channel, which not only stabilize the aptamer conformation but also increase the pause lifetime of RNAP, an effect comparable to hairpin-stabilized class I pause sites. In contrast, binding of the ligand results in pause escape despite the stabilizing effect of the ligand on the pseudoknot structure. How a pseudoknot-stabilized pause can be rearranged into a pseudoknot-inhibited pause remains to be shown, but these results afford another intricate example of the transcription machinery cross-coupling the regulation of RNA folding and transcription kinetics.

Modulated RNA chaperoning functions in factor-modified transcription complexes

NusA as an RNAP-associated RNA chaperone

During transcription elongation, RNAP can associate with several elongation factors. Some of these factors expand the RNA-binding potential of ECs and thereby their potential RNA chaperone functions. Transcription kinetics provide time windows of opportunity for such proteins to associate with elongating RNAP and the nascent
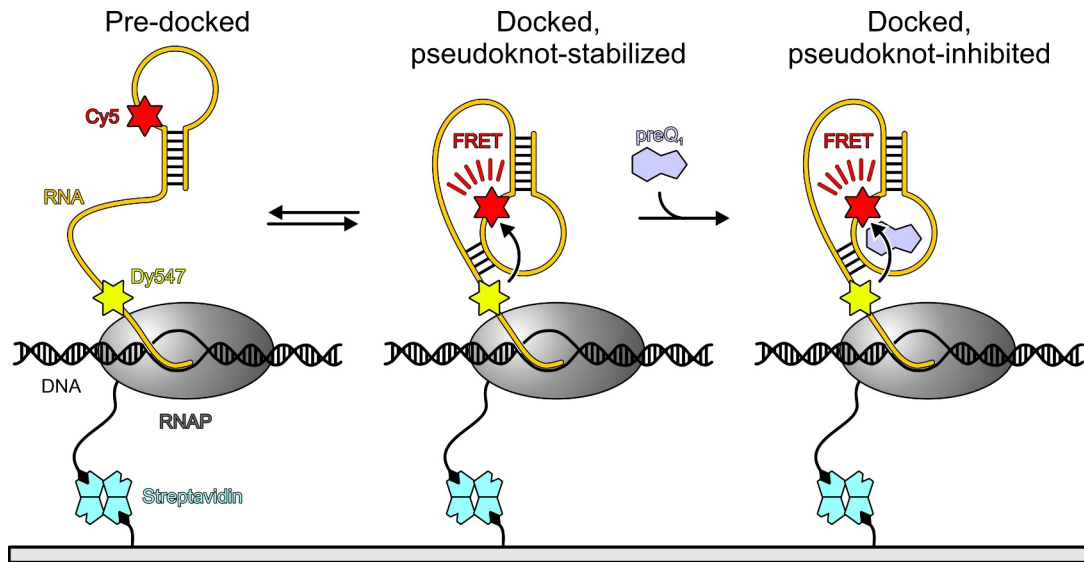
Figure 3. Scheme illustrating the folding of the *B. subtilis* preQ$_1$ riboswitch as revealed by smFRET studies, using prism-based total internal reflection fluorescence microscopy [10]. Active ECs were assembled on artificial bubble scaffolds, using RNAs that contained the preQ$_1$ RNA aptamer and different portions of the expression platform. ECs were immobilized on polyethylene glycol-passivated, streptavidin-coated quartz slides *via* biotinylated RNAP. The RNA contained the donor fluorophore (Dy547) at the 3'-end of the aptamer domain and the acceptor fluorophore (Cy5) within a loop region. High FRET signals were obtained upon formation of a docked state from a pre-docked state within the aptamer domain. The docked conformation is further stabilized by binding of the preQ$_1$ ligand. The ligand-free but pre-folded riboswitch pseudoknot that stabilizes the paused state through interactions with RNAP is a hallmark of the *que* (class III) pause. Binding of the preQ$_1$ ligand stabilizes a docked conformation different from the *que*-paused state and counteracts pausing.

RNA, while the ensuing RNA–protein interactions themselves may exert a strong influence on co-transcriptional RNA folding.

NusA is a highly conserved transcription elongation factor in bacteria and archaea [41] and is associated with RNAP during transcription of many, if not all, transcription units [104]. NusA joins the EC shortly after initiation and stays attached to RNAP throughout the transcription cycle [104]. Thus, NusA can be envisioned as an additional regulatory subunit of RNAP during the elongation phase. Traveling with the transcriptional machinery, it directly affects transcription kinetics by regulating pausing [2,5,105] and intrinsic as well as factor-dependent termination [57,106,107]. NusA can interact with a plethora of other regulatory factors that associate with the transcription apparatus [57,106–115], constituting a global gene regulator during transcription and transcription-related processes [116].

NusA is composed of six domains, an N-terminal domain (NTD), an S1 and two KH RNA-binding domains (the SKK module) [117] and two acid repeat domains (AR1 and AR2) [111,118]. The S1 domain is also found in cold shock domain proteins that act as RNA chaperones [46]. The NTD binds to the β FTH and to the C-terminal region of one RNAP α subunit (αCTD), and is flexibly linked to the SKK module that in turn can bind the nascent transcript [5,57,106,107,113]. The C-terminal AR domains are only conserved in γ-proteobacteria, with AR2 binding to the other αCTD of RNAP [5,119].

Via its pause-modulating properties, NusA can have a profound influence on RNA folding [19]. However, NusA is also a versatile RNA-binding protein, suggesting that its RNA binding activities *per se* might further modulate co-transcriptional RNA folding. Indeed, recent studies portray NusA also as a co-transcriptional RNA chaperone, a function perfectly consistent with its structure and location at the mouth of the RNA exit channel [5,16].

NusA supports hairpin-stabilized RNAP pausing and intrinsic termination [80,81]. A recent cryoEM structure of a NusA-modified *his*-PEC provided insights into the molecular basis by

which NusA can affect hairpin-stabilized pausing by resorting to its RNA chaperoning activity [5], which is presumably also the mechanism by which NusA can stimulate intrinsic termination. In the NusA-*his*-PEC structure, the NusA NTD binds to one αCTD and to the β flap domain, directly contacting the β FTH with a three-helix bundle of the NTD and the flexible linker helix connecting the NTD to the S1 domain (Figure 1c). NusA stabilizes the β FTH in a position above the RNA exit channel, thereby widening the channel and extending one wall together with the β' zipper and β' ZBD. On the other side, the channel is further elongated by the NusA S1 domain, which sits on top of the β' ZBD, and is additionally stabilized by interaction of the adjacent KH domains with the ω subunit. The AR2 domain interacts with the other αCTD. Therefore, NusA NTD-S1 together with the β' dock and β CTD provide an extended path for the nascent transcript emerging from the mouth of the RNA exit channel. Although the *his*-pause hairpin RNA folds similarly in the presence and absence of NusA [4,5], the structural organization of NusA around the RNA exit channel reveals key contacts important for the stimulatory effect of NusA on RNA folding [85]. First, the NusA-β FTH interaction has been shown to suppress the effect of the β FTH in delaying duplex formation, and positioning the FTH distal to the site of hairpin formation probably relieves steric interference. Second, NusA provides a surface of conserved positively charged residues thought to bind the nascent RNA [57,106,107,117]. The NusA S1 domain, the β' dock and the β' ZBD form a positively charged pore that might provide a pathway for the RNA toward the KH domains, while the NusA N-terminal linker helix and S1 domain form a concave, positively charged cradle appropriately positioned to bind the RNA hairpin loop after hairpin invasion [5]. While the RNA loop is again not defined in the structure [5], suggesting that it remains dynamic in the NusA-modified *his*-PEC, the structure suggests that NusA assists RNA folding by guiding the RNA along its S1 domain and stabilizes the RNA hairpin in the RNA exit channel.

Several conformational states observed for the NusA-modified *his*-PEC [5] revealed a large degree of conformational freedom, which allows NusA to adopt multiple positions relative to RNAP. Cavities formed between NusA and RNAP are wide enough to accommodate small structured RNAs. The flexibly linked domains/modules in NusA and the flexible anchoring points on RNAP (β FTH, αCTD) might allow NusA not only to adapt to different RNA hairpin lengths during hairpin-stabilized pausing and intrinsic termination, but also to more generally support co-transcriptional RNA folding. NusA might be dynamically repositioned while growing RNA structures accumulate outside of the RNA exit tunnel, which may explain at least in part how NusA can stabilize RNA secondary structures and facilitate RNA folding [19,81].

λQ as a fold-delaying co-transcriptional RNA chaperone

Investigations of bacteriophages have afforded examples of many phage-encoded transcription regulators [120], whose further studies have revealed basic principles of bacterial transcription regulation. For example, lambdoid phages regulate their gene expression programs by expressing proteins that modulate the pausing and termination behavior of RNAP transcribing phage genes (Figure 2b). A prime example is the family of Q proteins that mediate a switch from middle to late gene expression during the lytic life cycle [121,122]. Q loads onto RNAP paused at a σ-dependent pause element of the transcription unit following the Q-coding region. It then accompanies RNAP during further transcription elongation, rendering it resistant to pausing and termination in a highly processive manner. Thus, Q represents a textbook example of a transcription anti-termination factor.

Recent structural analyses have revealed how Q of phage 21 (21Q) loads onto TCs as well as the configuration of a 21Q-modified EC (21Q-EC) [123,124]. While two 21Q protomers associate with paused RNAP and a direct repeat in a DNA Q-binding element (*QBE*) between −35 and −10 promoter regions in a σ-dependent manner [123,124], only the upstream-bound 21Q protomer (21Q^u) remains associated with the EC after pause escape and dissociation of σ [124]. Both in the

loading complex and in the 21Q-EC, 21Q^u modulates RNAP in a remarkable manner that profoundly alters RNAP's ability to chaperone the formation of RNA hairpins in the RNA exit channel. An N-terminal region of 21Q^u forms a lasso-like structure that binds along the outer rim and inside the upper part of the RNA exit channel, displacing the β FTH from the rim of the channel (Figure 2c). It thereby extends the channel and restricts its diameter to below 10 Å, which is insufficient to adopt double-stranded RNA. The 21Q-EC structure clearly revealed that nascent RNA is threaded in single-stranded form through the 21Q^u lasso, which thereby effectively prevents the invasion of the RNAP RNA exit channel by RNA hairpins, counteracting hairpin-stabilized pausing and intrinsic termination. The lasso-like structure of 21Q^u bears a prominence of positive charge and interacts sequence-independently with the RNA backbone. A serine and threonine residue at the narrowest constriction may act as a "molecular bearing" that supports continuous threading of the transcript upon further elongation. Thus, 21Q^u can be regarded as an RNA chaperone that acts by delaying the folding of the nascent transcript and by preventing RNA chaperoning functions of the RNAP exit channel to take effect.

Q proteins have been classified into different sub-families. Members of the 82Q family can load onto ECs that contain long transcripts [125], a phenomenon seemingly incompatible with end-on threading of the transcript through a molecular lasso. It will be interesting to see in the future whether 82Q-like proteins act by mechanisms distinct from those observed for 21Q. Furthermore, it has been shown that λQ directly interacts with NusA [110], and it remains to be seen whether and how this interaction on a λQ-EC modulates NusA's co-transcriptional RNA chaperoning activity.

Modulation of the RNA polymerase RNA chaperoning activities by λN

Similar to the 21Q-EC example provided above, additional structures of differentially modified ECs recently revealed how also other factors can modulate RNAP elements that form part of the RNA exit channel and affect NusA [5,57,107,113–115]. Thereby, the RNA chaperoning activities of RNAP and NusA appear to be strategically modulated to alter the effects of regulatory RNA elements on RNAP.

One example of these principles is afforded by a processive anti-termination complex based on the N protein of phage λ (λN). Lambdoid phage N proteins constitute a second textbook example of transcription anti-pausing/anti-termination factors. They are employed by many lambdoid phages to switch from early to middle gene expression during their lytic life cycles. N proteins are small (about 110 residues) and intrinsically unstructured (Figure 2b). Using an N-terminal arginine-rich motif (ARM), they recognize a so-called *boxB* hairpin within an N-utilization (*nut*) signal element in nascent phage RNA encoded proximal to middle gene promoters (Figure 2b) [126,127]. Additionally, they contact RNAP and string Nus factors A, B, E (equivalent to ribosomal protein S10) and G, as well as *nut*-site RNA, into a higher-order RNA-protein complex (RNP) on the surface of RNAP that stably modifies RNAP and renders the enzyme pause and termination-resistant [128].

The molecular organization and mechanistic principles underlying λN-mediated anti-pausing and anti-termination have been recently described [57,106]. Within the λN-EC, λN remains highly elongated, folding locally into α-helical structures and interconnecting RNA elements, Nus-factors and RNAP (Figure 2d). The NusA KH domains additionally contact *boxB*, a NusB-NusE dimer engages a linear *boxA* motif of the *nut* site located 5′ to the *boxB* stem-loop as suggested previously [131] and a λN-stabilized NusA-NusG interface is built up [57,106].

Among the multi-pronged molecular strategy that it installs to achieve anti-pausing and anti-termination, λN directly modulates elements of RNAP that mediate contacts with the nascent transcript in other ECs [4,5,16], thereby altering RNAP's ability to chaperone RNA hairpins in the RNA exit channel. Running along the modifying RNAP, λN enters the RNAP catalytic cavity next to the upstream DNA, contacting the β flap, β' ZNB and β' dock, all of which are involved in RNA secondary structure accommodation and/or formation [4,5,99]. In particular, together with

N-terminal elements of NusA, it restructures the β flap tip, such that this element is displaced into the RNA exit channel. Moreover, the C-terminal portions of λN run along the DNA-RNA hybrid and into the RNA exit channel (Figure 2d, right). *Via* these molecular principles, λN leads to restriction of the channel diameter and to an alteration of the electrostatic potential of the inner channel surface, sterically and electrostatically hindering the formation of RNA helixes inside the channel. Thus, λN modulates RNAP RNA chaperoning activities in a manner similar to 21Q, but by resorting to different molecular principles. In addition, it seems to stabilize RNAP in an elongation competent conformation [57], possibly also affecting co-transcriptional RNA folding by modulating transcription kinetics.

Modulation of the NusA RNA chaperoning activities by λN

In addition to direct modulation of RNAP elements, λN-mediated buildup of a modifying RNP on RNAP leads to a large-scale repositioning of NusA on RNAP (Figure 2e). As part of this repositioning, contacts of the N- and C-termini of NusA to αCTDs of RNAP, which help position NusA in regular ECs [5], appear to be broken. As a consequence, the RNA-binding SKK module of NusA is relocated to a position remote from the mouth of the RNA exit channel, thus undermining NusA's ability to support formation of pause and termination hairpins and stabilize them in the RNA exit channel [5,57]. Instead, the repositioned NusA RNA-binding domains may effectively guide nascent RNA regions away from the RNA exit channel, further counteracting their potential pairing with complementary regions in the exit channel, as previously suggested as an operative principle in the λN-EC [81]. Thus, simply by strategically repositioning the RNA-binding domains of NusA, NusA seems to be converted from an RNA chaperone that supports RNA hairpin invasion of the RNAP RNA exit tunnel to a fold-suppressing RNA chaperone that elicits precisely the opposite effect.

λN achieves this large-scale repositioning of NusA by acting as a molecular glue that rewires Nus-factor interactions, thereby concomitantly subverting regular activities of the other Nus factors as well. NusG and relatives represent the only transcription factor family that is universally conserved across bacteria, archaea and eukaryotes [132]. *E. coli* NusG is a two-domain protein [133]. In NusG-modified ECs, including the λN-EC, the NusG NTD binds RNAP across the main channel and can contact upstream DNA (Figure 2d), increasing RNAP processivity and counteracting RNAP backtracking [12,57,134]. *Via* its CTD, NusG can interact with NusE [135] or transcription termination factor ρ [136] in a mutually exclusive manner. The single-domain NusB protein can form a stable complex with NusE [137] that is compatible with concomitant binding of the NusG CTD to NusE [135]. However, while NusG can interact with NusE that is part of a ribosome [135,138], the NusB-binding surface on NusE is occluded in the small ribosomal subunit [137]. The interactions among the Nus factors, RNAP, ρ and the ribosome are thought to regulate transcription-translation coupling and translational polarity [113,135,138]. λN anchors NusA to a NusB-NusE dimer on *boxA* and the C-terminal domain of NusG (Figure 2d). It seems to enhance NusG-dependent anti-backtracking by running along the opposite surface of upstream DNA, and it sequesters the NusG CTD from ρ or a trailing ribosome by reinforcing its interaction with NusB/E and NusA [57,106]. It may thereby further affect co-transcriptional RNA folding by modulating NusG-dependent transcription kinetics and/or transcription-translation coupling (see below).

As in the Q superfamily, N proteins of lambdoid phages exhibit significant sequence diversity, in particular in C-terminal regions. For example, the C-terminal region of the phage H19B N protein diverges from that of λN, and biochemical structure probing experiments suggested that it might closely approach the RNAP active site [139]. Thus, in the future it will be of interest to conduct comparative structural analyses of diverse N-ECs to explore the possible diversity in molecular principles, by which these proteins may modulate RNA chaperoning by RNAP and factors.

A complex RNAP-associated RNA chaperone that supports co-transcriptional folding of ribosomal RNA

Co-transcriptional RNA folding linked to co-transcriptional RNP assembly and RNA processing

is a fundamental aspect of the biosynthesis of ribosomal (r) subunits [140]. In bacteria, one or several rRNA operons encode primary transcripts of concatenated 16S, 23S and 5S rRNAs, with intervening tRNAs, which are processed by several nucleases to yield mature 16S, 23S, 5S rRNAs and tRNAs [141,142]. In the late 1980s, Miller and colleagues studied the synthesis of rRNA using negative stain electron microscopy of *E. coli* chromatin [143,144]. On so-called "Miller spread" electron micrographs, they observed densely packed trains of RNAP on rRNA operons. The nascent rRNAs were associated with compact structures that were interpreted as RNA-protein complexes building up prior to completion of transcription. Moreover, they observed discrete transitions from long to short transcripts, strongly indicating that processing of rRNAs starts before RNAP reaches the 3′-end of the operon. A major transition point was interpreted to mark cleavage by RNase III, generating pre-16S and pre-23S rRNAs [145]. Indeed, full-length primary rRNA transcripts can only be detected when RNase III is inactivated [146]. A major impact of co-transcriptional rRNA folding on rRNP assembly and rRNA processing is also suggested by the observation that *in vitro* subunit assembly based on full-length rRNAs is several orders of magnitude slower than ribosome biogenesis *in vivo* probably due to rRNA folding problems [147,148].

At about the same time, the composition of the transcriptional apparatus responsible for the synthesis of rRNAs was also elucidated. In these *rrn*ECs, RNAP is distinctly modified compared to ECs transcribing other RNAs, in a manner reminiscent of N-ECs [149]. Assembly of *rrn*ECs is nucleated by a *nut*-like RNA element encoded in the 16S rRNA leader and 16S-23S rRNA spacer regions. In the *rrn nut*-like sites the order of *boxA* and *boxB* elements is reversed compared to the phage λ *nut* sites, and an additional, linear motif, *boxC*, is present at the 3′-end [150,151]. In addition, the formation of an *rrn*EC depends predominantly on the *boxA* element, with *boxB* being dispensable under most conditions [151]. As the λN-EC, *rrn*ECs encompass Nus-factors A, B, E and G. However, instead of an N protein, several r-proteins have been shown to be part of *rrn*ECs [152]. In

particular, r-protein S4 has been identified as a key factor that directly interacts with RNAP and elicits NusA-like effects, e.g., shifting ρ-dependent termination windows downstream [153]. More recently, the inositol monophosphatase, SuhB, was discovered as another essential component of *rrn*ECs. *SuhB* and *nusB* mutants showed similar phenotypes *in vivo* [154], SuhB can directly interact with RNAP [155] and SuhB can modulate the transcription behavior of RNAP in dependence of an *rrn nut*-like element and Nus-factors [108].

Further in analogy to the λN-EC, an *rrn*EC transcribes twice as fast as regular ECs, and it suppresses RNAP pausing and ρ-dependent termination [149,156,157]. *A priori*, the fast, pause-suppressed rRNA synthesis by *rrn*ECs appears to conflict with efficient co-transcriptional rRNA folding, as co-transcriptional folding of many other RNAs requires strategic pausing of RNAP (see above). However, reduced rRNA synthesis rates have little or even negative influence on the proper folding of the mature rRNAs [158,159], suggesting that other mechanisms supporting co-transcriptional folding must be at work. The observation that rRNA synthesis based on a phage RNA polymerase leads to largely inactive ribosomes [160] suggests that these mechanisms are specific to the endogenous *rrn*ECs.

Nus factors have been proposed to act as rRNA chaperones [154,161]. Indeed, recent structural and structure-based functional analyses suggested that RNA chaperoning by components of *rrn*ECs substitute for pausing-mediated co-transcriptional folding during rRNA synthesis [107]. In the *in vitro*-assembled *rrn*ECs, Nus-factors and a SuhB dimer form a ring-like structure around the mouth of the RNA exit channel (Figure 4a). Apart from protein–protein contacts, the modifying factors are further inter-connected by their interactions with *boxA* (NusB-NusE) and the *boxA-boxC* linker (NusA, SuhB). These protein–RNA interactions also fix the 5′-end of the nascent rRNA at a channel built up by NusB, NusE, NusA and SuhB. When present, r-protein S4 seems to cover the channel opposite of NusA and SuhB and to contact the RNA at the 3′-end of the *nut*-like site in concert with the NusA S1 domain, probably forming a flexible lid (Figure 4b); however, the
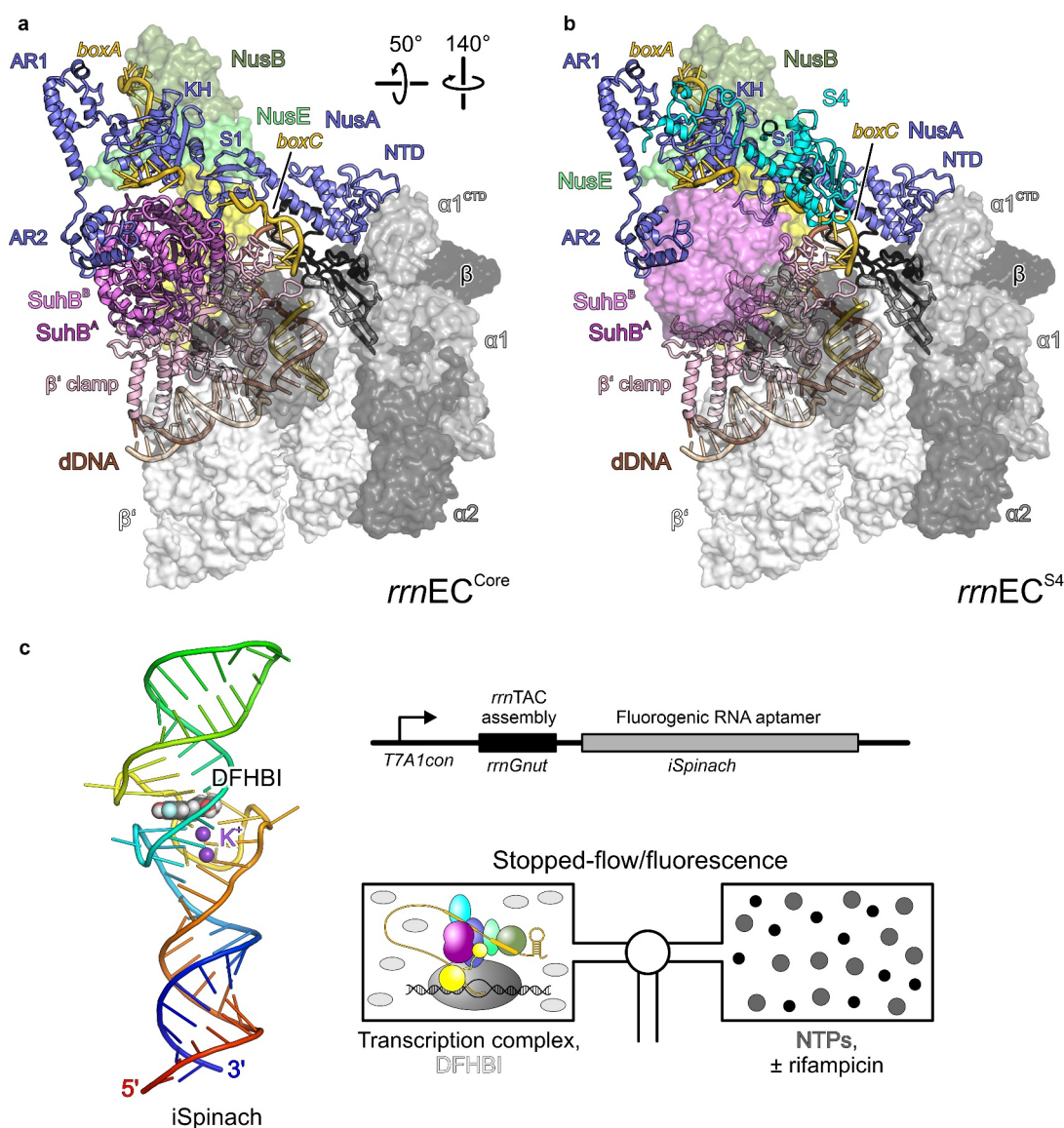
Figure 4. Semi-transparent surface views of an *rrn*EC lacking S4 (PDB ID 6TQN) and an *rrn*EC containing S4 (PDB ID 6TQO). SuhB[A], purple; SuhB[B], violet; r-protein S4, cyan. **a**. *rrn*EC lacking S4. SuhB[A] and SuhB[B] are shown as cartoon. Nus factors and a SuhB dimer build up a composite RNA chaperone ring around the RNA exit channel. Proteins interact with *boxA* (NusB-NusE) and the *boxA-boxC* linker (NusA, SuhB). The network of protein-protein and protein–RNA interactions fixes the 5′-*nut*-like element of the nascent RNA next to and within the chaperone ring. **b**. *rrn*EC containing S4. SuhB[A] and S4 are shown as cartoon. S4 adds to one wall of the chaperone ring opposite of NusA and SuhB, probably forming a flexible lid that contacts the 3′-end of the *nut* RNA and sensing RNA that loops out from the RNA exit channel. **c**. Left, structure of the *iSpinach* RNA aptamer, forming a binding platform for the pro-fluorophore, DFHBI (PDB ID 5OB3) [162]. Right, scheme of the DNA template used to study effects of *rrn*ECs on co-transcriptional folding [107] (top) and scheme illustrating the experimental setup using a stopped-flow/fluorescence device (bottom).

precise configuration of S4 in the *rrn*EC should be considered tentative due to weak cryoEM density in the corresponding regions [107].

Density for sequences of RNA regions between the 3′-end of the *nut*-like site and the RNA region located in the RNA exit channel was missing, and this part of the nascent transcript is expected to loop out during ongoing transcription. In contrast to the λN-EC, secondary structure formation within the RNA exit channel is not restricted, leaving enough space for pause hairpins or other secondary structures to form. However, as in the λN-EC, NusA is repositioned relative to RNAP, preventing hairpin stabilization by NusA important for both, pausing and intrinsic termination [5]. Indeed,

NusA-mediated stabilization of hairpin-dependent pausing is suppressed in the *rrn*EC [107].

The Nus-factors and SuhB generate a partially positively charged crevice, essentially extending the RNA exit channel. rRNA sequences downstream of the *nut*-like site that fold into diverse local structures will thus be guided along elements of the RNA exit channel (β' ZNB, β' dock, β flap and β NTD), NusA-SKK, SuhB and possibly flexibly associated S4. This molecular organization is consistent with the factors forming a versatile, composite RNA chaperone at the mouth of the RNAP RNA exit channel. By transiently binding RNA regions that emerge from RNAP, the factors could prevent such regions from becoming kinetically trapped in nonproductive folds. Alternatively or in addition, they may present regions of RNA to subsequently produced regions, or bind different portions of the nascent RNA concomitantly, thereby supporting RNA duplex formation, as required for efficient rRNA folding [46,48].

Fluorescence-based co-transcriptional RNA folding assays based on the fluorogenic RNA aptamer, *iSpinach*, in conjunction with structure-informed mutagenesis provided direct support for this model (Figure 4c). In these assays, real-time monitoring of emerging fluorescence served as a readout for co-transcriptional folding, as only properly folded *iSpinach* provides a binding platform for the pro-fluorophore, DFHBI, which then becomes fluorescent. Folding rates were significantly increased in *rrn*ECs containing or lacking S4 compared to unmodified RNAP, and altering positively charged or aromatic residues on NusA or SuhB exposed to the nascent RNA reduced folding efficiency [107].

Thus, transcriptional pausing is not a *sine qua non* for co-transcriptional RNA folding. The modifying factors of *rrn*ECs seem to directly support folding by well-known RNA chaperoning mechanisms. As large structures cannot be accommodated within the composite chaperone, it presumably supports the nucleation of structures that then might be quickly released or "grow" out of the folding chamber. Molecular crowding effects that increase the local concentration of interacting RNA regions might additionally support co-transcriptional folding at *rrn*ECs [163]. In any case, an RNA folding landscape around the mouth of the RNAP RNA exit channel in *rrn*ECs is reminiscent of protein biogenesis factors that bind the nascent polypeptide at the polypeptide tunnel exit of the ribosome, where they mediate multiple co-translational processes, including protein modifications, folding, targeting and degradation [164].

The ribosome as a co-transcriptional RNA chaperone

In bacteria, transcription is often coupled to translation, as a ribosome can initiate translation while the mRNA is still being transcribed. A specific role during transcription-translation coupling is assigned to the lead ribosome that performs the first round of translation and that can physically interact with the EC. The lead ribosome apparently synchronizes transcription and translation rates, can prevent transcriptional arrest and can protect the nascent RNA from ρ-dependent termination [165,166]. In contrast to the lead ribosome, all following ribosomes, which together with the lead ribosome build up a polysome, lack a direct interaction with the EC.

Apart from coordinating the molecular machineries during bulk transcription/translation, transcription-translation coupling can install specific regulatory mechanisms, which additionally depend on co-transcriptional RNA folding. Transcription attenuation constitutes one of the most important mechanisms of gene expression control in bacteria [40]. Attenuation occurs in the mRNA 5'-leader regions of bacterial operons and relies on RNA regions that can form an effector-binding domain, the target for a metabolite, a protein, an RNA or a translating ribosome, and a transcription terminator. The presence or absence of the effector determines the RNA structure in the effector domain, which in turn influences the formation of the downstream terminator. Pause-guided co-transcriptional folding of the regulatory RNA sequences within the leader region was described for a variety of transcription units controlled by transcription attenuation mediated by ribosomes, RNA binding proteins or riboswitches [16], and in many cases is supported by NusA. Prominent examples include amino-acid biosynthetic operons

in enterobacteria, such as the *his* and *trp* operons [40,77,167]. Leader regions of these *his* and *trp* operons encode pause hairpins, and RNAP pausing is envisioned to provide a time window for the ribosome to start translation [16]. Following translation initiation, the ribosome is thought to physically disrupt the pause hairpin and release RNAP to continue transcription. Subsequently, depending on where the ribosome itself will pause on the nascent transcript while translating a His-rich or Trp-rich leader peptide, which depends on the available levels of amino acids, two mutually exclusive RNA structures can form, a transcription anti-terminator or a terminator hairpin [40].

Recent cryoEM structures of *E. coli* transcription-translation coupled complexes (TTCs) *in vitro* [115,168] and *in vivo* [169] not only revealed that NusG can physically bridge the transcription and translation machineries, as previously suggested [135,138], but that in particular NusA seems to form a "coupling pantograph" between RNAP and the 30S ribosomal subunit that can maintain contacts between the transcription and translation machineries, while simultaneously providing enough flexibility in the connections to allow different orientations of RNAP relative to the lead ribosome (Figure 5). Different lengths of the connector RNA between the Shine-Dalgarno sequence and the 3′-end of the mRNA induced distinct relative orientations between the ribosome and RNAP [115]. Furthermore, the NusA NTD and S1 domain in the coupled machineries could still engage in electrostatic interactions with the nascent transcript and thereby assist formation of secondary structures, such as pause and terminator hairpins or anti-terminators. Together, these observations suggest that the lead ribosome may influence co-transcriptional RNA folding, and thus the outcome of certain transcription regulatory events, not only by occupying or leaving available RNA regions for RNA folding, but also by influencing the positioning of RNA chaperones, such as NusA, or by otherwise modulating the RNA folding landscape between the coupled transcription and translation machineries.

In turn, co-transcriptional RNA folding can exert an effect on the recruitment of the lead
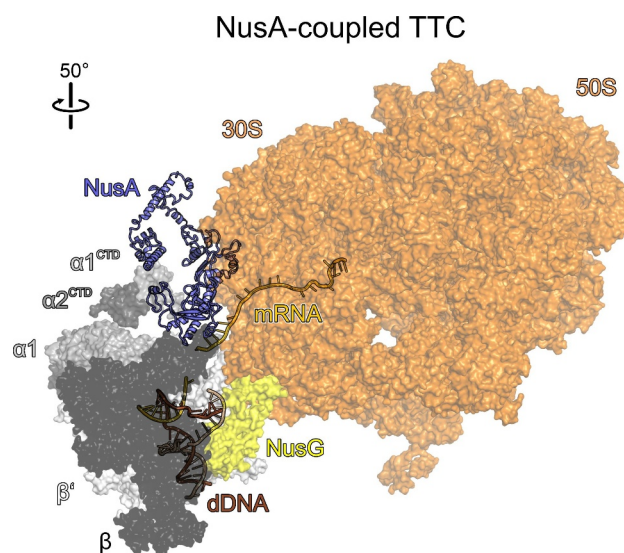

Figure 5. Semi-transparent surface representation of a NusA-coupled TTC (PDB ID 6X7F). Ribosome subunits are shown in Orange. NusA act as a "coupling-panthograph" between RNAP and the 30S ribosomal subunit. NusA is constantly but flexibly linked to RNAP *via* its NTD and AR2 domains, providing NusA with a high degree of conformational freedom between RNAP and the S30 subunit, which could facilitate NusA to make favorable contacts to the nascent RNA transcript.

ribosome and the establishment of transcription-translation coupling, as revealed in a recent study of the $preQ_1$-sensing translational riboswitch of *Bacillus anthracis* [170]. This riboswitch determines the rate of translation initiation by controlling the accessibility to the Shine-Dalgarno sequence. In the absence of the $preQ_1$ ligand, the Shine-Dalgarno sequence is accessible, allowing ribosome binding. $PreQ_1$ binding to the riboswitch induces structural rearrangements in the RNA that lead to sequestration of the ribosome-binding site and decreased 30S subunit binding. Interestingly, binding of the 30S subunit was promoted in the presence of an EC and further facilitated or consolidated by the transcription factor paralogs, NusG and RfaH, respectively, indicating that co-transcriptional folding of the riboswitch and the possibility for formation of a physical bridge between the EC and the 30S subunit are important for establishing a coupled transcription-translation complex. While NusG supported 30S recruitment in the absence but not in the presence of the $preQ_1$ ligand, RfaH circumvented the need of a Shine-Dalgarno sequence for efficient ribosome recruitment, as in the presence of RfaH binding of the 30S subunit was unaffected by the addition of the

preQ$_1$ ligand [170]. These findings not only reveal that co-transcriptional RNA folding can elicit differential effects on the establishment of coupled transcription-translation complexes, depending on which coupling factors are present, but also that co-transcriptional RNA folding can impact several steps along the gene expression process, affecting transcription, translation initiation and transcription-translation coupling.

Interplay of co-transcriptional RNA folding and other co-transcriptional processes

Co-transcriptional RNA chaperoning in rrnECs may be linked to rRNA processing and rRNP assembly

During transcription of 16S rRNA, the 5′-end of the nascent transcript folds first, followed by the central and 3′-domains [171]. Likewise, *in vitro* reconstruction experiments established that ribosomal subunit assembly proceeds in a sequential and cooperative manner [147,172–176]. Although direct links between rRNA transcription and ribosome assembly have been noticed decades ago [143,177], molecular mechanisms underlying the elaborate interplay between co-transcriptional rRNA folding, rRNP assembly and rRNA processing are still being explored, e.g., by genetics, single-molecule fluorescence and structural approaches [107,158,161,178].

Pioneering studies by the Williamson and Woodson groups employed fluorescence-based single-molecule techniques, in which they monitored co-transcriptional rRNA synthesis and folding with simultaneous r-protein recruitment (Figure 6a,b) [158,159,178]. By following the assembly of the 5′ and 3′-domains of 16S rRNA in real time, they observed that primary-binding r-proteins S4 and S7 initially bind transiently to nascent rRNA and become more stably associated in the course of transcription, promoting correct rRNA folding [158,159]. Furthermore, formation of these primary interactions is chaperoned by additional r-proteins originally thought to be recruited later in the subunit assembly process, which is in contrast to the classical model of a strict r-protein hierarchy in subunit assembly.

These single-molecule studies relied on transcription by unmodified RNAP or heterologous phage RNAP, posing the question of how fully modified *rrn*ECs might support further links between co-transcriptional rRNA folding, rRNP assembly and rRNA processing. A possible impact of the modifying factors on rRNA processing was suggested by FRET-based RNA annealing assays, using oligos complementary to RNA regions located at the *rrn*EC's composite RNA chaperone. These analyses showed that S4 serves as an RNA annealing factor, promoting the formation of double-stranded RNA from regions that are remote from each other within the Nus-factor/SuhB folding landscape [107]. Thus, *rrn*ECs might not only chaperone local RNA structure formation but also some long-range RNA interactions, such as the formation of the primary RNase III target site formed from complementary regions upstream and downstream of 16S rRNA, involving *boxC* [179]. The latter long-range interaction is additionally supported by the topological restraints imposed on the nascent rRNA in *rrn*ECs, with the *nut*-like element preceding rRNA regions being tightly fastened within the modifying RNP [107]. Thus, the *boxC* region can be efficiently "delivered" to its complement downstream of 16S rRNA as proposed already several decades ago [151].

In further support of a possible functional connection between rRNA co-transcriptional folding and processing, Bubunenko *et al.* showed that altered translation and accumulation of 30S precursors caused by cold-sensitive mutants of *nusA* and *nusB* [177] can be suppressed by deletion of *rnc* [161], the gene encoding RNase III. As RNase III cleaves dsRNA within an RNA duplex formed from regions upstream and downstream of 16S rRNA [179], defects in RNA folding caused by the *nus* mutants were suggested to be suppressed in a Δ*rnc* strain by artificially stabilizing this RNA duplex structure [161].

In the traditional model of rRNP assembly, initial RNA folding is chaperoned by primary-binding proteins that create the binding site for secondary binding proteins. S4 is one of the 16S rRNA primary-binding proteins that nucleates correct folding and assembly of the 30S ribosomal
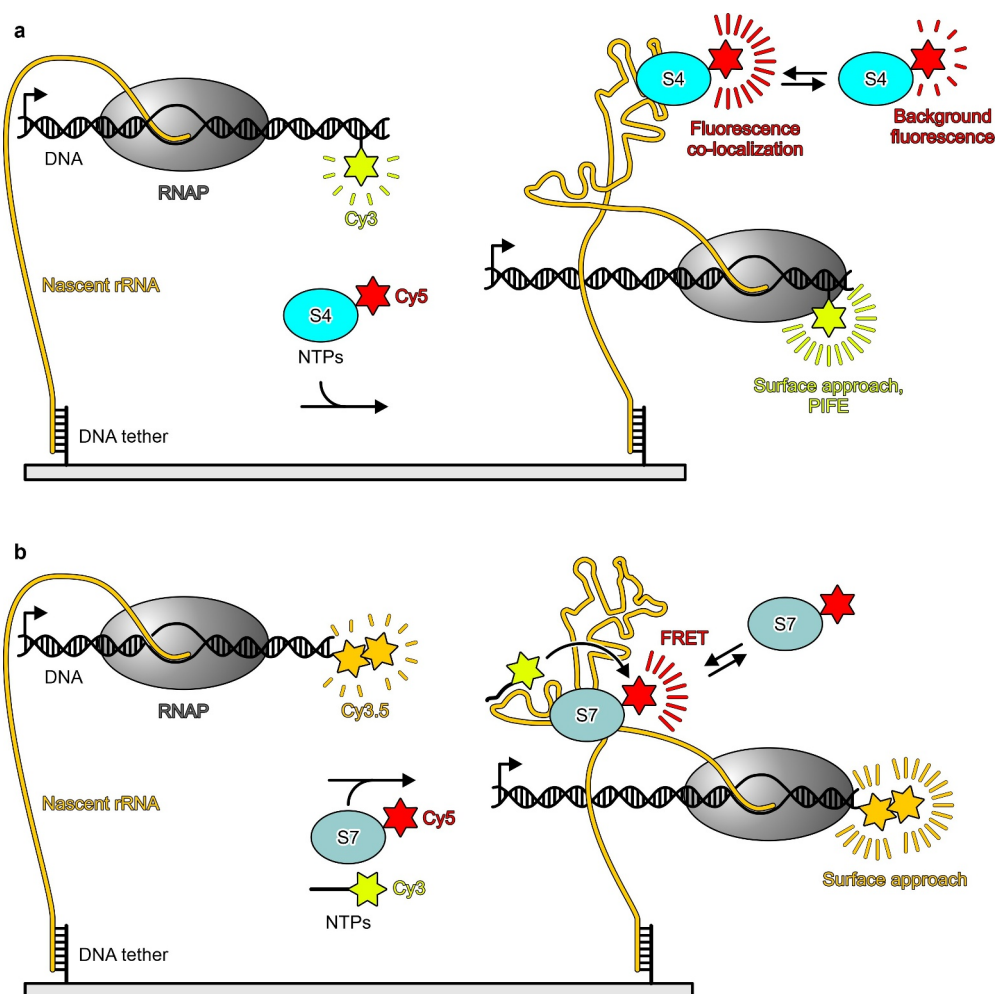
Figure 6. Schemes illustrating single-molecule fluorescence approaches to study co-transcriptional rRNA folding and rRNP assembly. **a**. Single-molecule co-localization co-transcriptional assembly (smCoCoA) studies of r-protein S4 during 16S rRNA transcription using total internal reflection fluorescence microscopy [158]. Stalled ECs were formed on a DNA template containing a Cy3 fluorophore upstream of a transcription terminator and were immobilized on a slide surface through a biotinylated DNA tether complementary to the 5′-end of the nascent RNA. Transcription elongation by T7 RNAP was started by the addition of NTPs and Cy5-labeled r-protein S4 to simultaneously monitor transcription (indicated by a gradual increase in Cy3 fluorescence as the 3′-end of the template approaches the surface) and S4 binding. A spike in Cy3 fluorescence intensity marked the end of transcription due to protein-induced fluorescence enhancement (PIFE) when T7 RNAP approaches the transcription termination site. S4 binding events were detected by co-localization of the Cy5 signal with active Cy3-labeled ECs. **b**. Single-molecule fluorescence microscopy studies monitoring co-transcriptional rRNA folding and assembly of r-protein S7 during transcription of the 3′-domain of 16S rRNA by using zero-mode waveguide (ZMW) technology [159]. Stalled ECs were formed on a DNA template containing two Cy3.5 fluorophores at the 3′-end and were immobilized on the ZMW surface *via* a biotinylated DNA tether complementary to the 5′-end of the nascent RNA. Transcription by *E. coli* RNAP was initiated by addition of NTPs, Cy5-labeled r-protein S7 and a Cy3-labeled oligo that can hybridize to either terminus of the nascent RNA. Transcription was monitored by the increase in fluorescence intensity while the Cy3.5 fluorophores in the DNA template approached the surface of the ZMW. Simultaneous annealing of the Cy3-oligo and binding of Cy5-S7 at or close to its specific binding site led to fluorescence resonance energy transfer (FRET) to Cy5-S7.

subunit. S4 binds to a five-way helix junction (5WJ) that is formed by a long-range interaction between the beginning and end of the 16S 5′-domain [180,181]. In that respect, the apparent activity of S4 as an RNA annealing factor in the *rrn*EC [107] may again be relevant. However, apart from exerting co-transcriptional RNA chaperone activity by supporting local rRNA folding and long-range RNA interactions, *rrn*ECs may also have an intriguing, "indirect" chaperoning function, by acting as reservoirs of r-proteins and other proteins that may be handed off to nascent rRNPs, where they might then chaperone rRNA folding and rRNP assembly independent of the *rrn*EC.

Besides S4, NusE/S10 is a long-recognized component of *rrn*ECs. In addition, r-proteins L3, L4 and L13 [153] as well as the heat shock protein YbeY [182] have been suggested to constitute additional components of *rrn*ECs. The hypothesis that the very same proteins may initially modulate rRNA transcription as subunits of *rrn*ECs and then are transferred to nascent rRNPs to exert additional rRNA/rRNP chaperoning functions will require the development of elaborate co-transcriptional rRNA folding and rRNP assembly assays based on authentic *rrn*ECs.

Modulations of the rrnEC theme

Topological restraints that might guide co-transcriptional RNA folding as observed in *rrn*ECs are also implemented in the λN-EC, in which the modifying RNP keeps a tight grip on the 5′-proximal *nut* site RNA during further transcription. As in the *rrn*EC, the λN-dependent fixation of the *nut* site at the 5′-end of the nascent transcript may have a profound effect on nascent RNA folding by forcing subsequent RNA regions to loop out and by presenting 5′-terminal regions to regions exiting the RNAP RNA exit channel later on. In contrast to the other Nus-factors, SuhB is excluded from the λN-EC, as it would clash with the phage *boxB* element [57,107]. However, in principle N proteins could replace SuhB in *rrn*ECs. It would thus be interesting to test if N proteins could organize Nus-factors into an *rrn*-like EC based on the *rrn nut*-like elements, and if the resulting N-modified complexes could also support co-transcriptional rRNA folding, rRNP assembly and rRNA processing. Such a scenario might ensue during the lytic life cycle of the phages and maintain efficient ribosome production when much of the cell's transcriptional resources are diverted to transcription of the phage genome.

RNA chaperoning by NusA may modulate ρ-dependent termination

ρ is an NTP-dependent RecA-like hexameric RNA translocase/helicase implicated in diverse regulatory processes that depend on its best known function as a transcription termination factor [183–185]. ρ can exist in an open conformation, in which it can bind RNA at the center of a ρ spiral, and a closed conformation, in which it clamps down on the entrapped RNA and can translocate 5′-to-3′ on this RNA in an NTPase-dependent manner. Although NusA-mediated contacts between ρ and RNAP have been noted long ago [186], the question of whether ρ can engage ECs without terminating transcription has been debated, as the classical model of ρ-dependent termination implies that ρ loads onto a ρ-utilization (*rut*) site in the nascent transcript independent of an EC, subsequently tracks down the EC by virtue of its RNA translocase activity and, upon encounter, leads to termination by resorting to its powerful motor activity [183–185]. Recent cryoEM structures in conjunction with structure-informed functional analyses strongly support an alternative model, in which ρ initially travels on ECs independent of RNA contacts and without leading to termination, and traps ECs in a pre-termination state only at RNAP pause sites [113,114]. This model is consistent with observations that ρ seems to traffic on NusA/NusG-modified ECs throughout most of the transcription cycle on almost every transcription unit [104,187,188].

ρ binds ECs through extensive contacts to RNAP, NusA and NusG in an open ring conformation placing two of its subunits around the RNA exit channel (Figure 7). NusA seems to initially inhibit ρ-dependent termination by inserting between two ρ subunits, thereby keeping ρ in an open conformation, consistent with its effects in biochemical assays [113]. In addition, in an initial ρ engagement complex, NusA seems to guide exiting RNA away from ρ [113], consistent with a previous suggestion based on biochemical analyses [189]. Stepwise remodeling, including NusA displacement, seems to turn an initial ρ trafficking complex [114] into a pre-termination complex, in which ρ gains access to the nascent transcript, without resorting to its NTP-dependent motor function [113]. Thus, NusA seems to chaperone RNA engagement by ρ to modulate transcription termination.

ρ has been shown to strongly influence transcriptional pausing *in vitro* [187]. Furthermore, a *RARE* [103] and inhibitory RNAP-binding aptamers (*iRAP*s) [190] have been identified that interfere with and promote ρ-dependent termination, respectively. Also, recent studies revealed
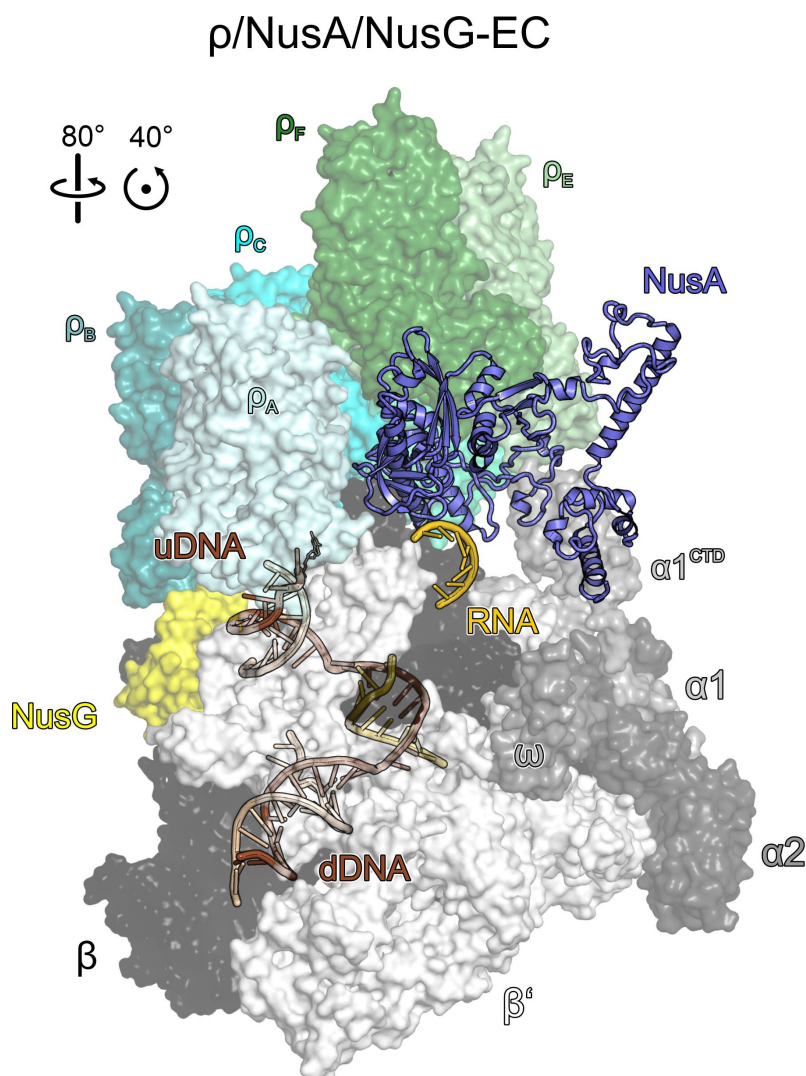
ρ/NusA/NusG-EC

Figure 7. ρ/NusA/NusG-modified EC. Semi-transparent surface representation of the engagement complex (PDB ID 6Z9P). ρ subunits, different shades of green/cyan. ρ is recruited to RNAP at the upstream face of the active site, making direct contacts to RNAP, NusA and NusG. NusA keeps ρ in an open ring conformation by inserting into the opening of the ρ ring. In the engagement complex, NusA is located proximal to the RNA exit channel and seemingly guides the nascent RNA along its RNA binding domains, preventing ρ from engaging the transcript.

a hitherto underestimated interplay between ρ and small regulatory RNAs (sRNAs) whose functions depend on RNA chaperones, such as Hfq (see below) [191–193]. Thus, ρ might function as a general transcription factor that resorts to RNA chaperoning activities in various scenarios [114].

Additional co-transcriptional RNA chaperones

Ribosomal protein S1

S1 is the largest ribosomal protein with a molecular mass of around 60 kDa and is composed of six S1 domains, as also found in NusA [194,195]. S1 exhibits RNA chaperone activity, e.g., during translation initiation, where it mediates mRNA binding to the small ribosomal subunit and resolves bulky secondary structures within the mRNA for the ribosome to start protein synthesis [46,196]. Observations that r-protein S1 co-purifies with RNAP [197,198] hint at the possibility that it might also act as an RNA chaperone during transcription. S1 has been observed to promote transcriptional cycling *in vitro* [199], has been suggested to play a role during rRNA anti-termination [200] and can bridge RNAP and the

30S ribosomal subunit [201]. In addition, S1 is an integral part of the bacteriophage Qβ replication complex [202] that interacts with NusA [203].

S1 action during transcriptional cycling has been proposed to depend on S1's interaction with both RNAP and the nascent transcript [199]. The molecular architecture of S1 with its string of RNA-binding modules would allow the protein to bind to the nascent transcript at several sites simultaneously and potentially prevent inhibitory interactions to RNAP. Moreover, by binding to 5′-untranslated regions (UTRs) of mRNAs, it could recruit S1-free ribosomes by forming a bridge between RNAP and the translational machinery. This model is consistent with S1 being the only r-protein that is not stably associated with one of the ribosomal subunits. A recent cryoEM structure indeed places S1 at the interface between core RNAP and the 30S subunit, suggesting that S1 could guide the mRNA into the entry channel of the ribosome, consistent with its role during translation initiation [201]. However, S1 is absent from cryoEM structures of active transcription-translation coupled complexes [115,168,169].

A role during rRNA transcription and/or processing has been contemplated based on S1's affinity to the *boxA* sequence located in the rRNA operon leader and spacer regions within the *rrn nut*-like elements [200]. However, S1 would have to replace the NusB-NusE dimer in *rrn*ECs [107] to exert such a role. As S1 harbors well-recognized RNA chaperone activity [48], and as even formally late-binding ribosomal proteins can modulate the initial steps of co-transcriptional rRNA folding and rRNP assembly [158,178], a more sophisticated role for S1 as a co-transcriptional RNA chaperone appears possible, but needs further experimental testing.

Ribosomal protein L4

In *E. coli*, the synthesis of r-proteins is often autogenously regulated by one of the products of r-protein operons at the level of translation initiation [204,205]. However, the S10 operon, which encodes eleven r-proteins including L4, is unique, as r-protein L4 employs two distinct and independent mechanisms to control gene expression, transcriptional attenuation and translational inhibition

[206,207]. When free L4 protein accumulates in the cell, it stimulates premature termination of transcription at a specific site in the S10 operon leader region, upstream of the first gene in the operon [206,208]. L4-mediated transcription termination is strictly dependent on NusA, which stabilizes the paused EC at the site of termination [209,210]. NusA-stabilized pausing depends on the formation of a pause-hairpin at the site of termination, whereas L4 regulatory sequences are located within the pause-hairpin as well as in a hairpin upstream of the pause site [209–211]. A NusA-stabilized hairpin pause is a requirement for subsequent termination, and it has been suggested that L4-induced structural changes in the leader region create a "super-paused" complex that leads to termination of transcription [211,212].

As the structure of the NusA-modified *his*-PEC revealed enough space for a terminator hairpin to form [5], it is intriguing to speculate that L4 might be part of the EC during S10 operon transcription, and together with NusA chaperones folding of the leader region. While NusA is known to support intrinsic termination [105], such a mechanism could explain how attenuation of S10 operon transcription is rendered dependent on L4.

Hfq

The RNA chaperone Hfq is widely conserved in bacteria, in which it acts as a global regulator of cell physiology. It is best characterized as a post-transcriptional regulator of gene expression during stress responses in bacteria [213,214]. As an RNA chaperone, Hfq promotes base-pairing between small non-coding RNAs (sRNAs) and their mRNA targets, acting as an RNA matchmaker that controls expression levels of the proteins encoded on the target mRNAs. The functions of Hfq in post-transcriptional gene regulation and the molecular basis of its chaperone activities in these contexts have been reviewed extensively [48,213–215].

Hfq has also been shown to be associated with ECs [197,198], possibly dependent on r-protein S1. The structural basis for Hfq-EC interactions is not yet known, but a role of Hfq in affecting mRNA levels during transcription has been proposed [216]. Hfq might promote co-transcriptional

RNA folding to counteract transcription pausing or arrest, thereby modulating attenuation in specific operons [217]. Hfq has also been shown to mediate transcription anti-termination at ρ-dependent terminators by directly interacting with termination factor ρ [218]. Although the functional targets of Hfq-mediated anti-termination still need to be identified, the potential of ρ to be globally associated with RNAP during transcription [104] and to directly bind ECs [113,114] opens up new mechanistic pathways for regulation, in which both, Hfq and ρ, might be involved. Interestingly, recent studies indicate an effect of sRNAs in the regulation of ρ-dependent termination, which might again rely on Hfq's RNA matchmaking activities [193]. In addition, Hfq from the opportunistic pathogen *Pseudomonas aeruginosa* has been shown to bind co-transcriptionally to hundreds of nascent transcripts [219] as does the post-transcriptional regulator RsmA [220]. These observations support the notion that the function of nominal post-transcriptional regulators as co-transcriptional RNA chaperones may be commonplace in bacteria.

Conclusion

Based on major advances in cryoEM-based structural analysis, single-molecule approaches and RNA/RNP structure probing techniques in combination with RNA sequencing we have deepened our insights into how RNAP simultaneously coordinates RNA synthesis, folding and processing. RNAP provides a chamber, the RNA exit tunnel, and most likely additional surfaces on which nascent transcripts start folding into secondary structures, co-transcriptionally guiding the RNAs toward their functional conformations. An intricate interplay is emerging between RNA chaperoning functions of RNAP, the order in which RNA regions are synthesized and transcription kinetics. RNA chaperoning by RNAP and transcription kinetics can be modulated by additional factors that interact with the nascent RNA, with RNAP or both, also linking co-transcriptional RNA folding to RNA processing and/or assembly of RNPs.

The core architecture of multi-subunit RNAPs, and in particular the structural elements that make up the pores and channels important for nucleic acid guidance or for granting substrates and regulators access to the active site, are conserved throughout all kingdoms of life [88,221], suggesting that basic RNA chaperoning functions are phylogenetically conserved. However, not only do the common subunits of RNAPs from different kingdoms show increasing sequence divergence in more peripheral regions, there are also species- or kingdom-specific subunits and sets of regulatory factors, suggesting that there will likewise be species- and kingdom-specific RNA chaperoning in TCs. E.g., within the bacterial kingdom, firmicutes contain additional small RNAP subunits, δ and ε, implicated in transcriptional recycling and structural integrity of RNAP, respectively [93], the precise functions of general transcription factors can differ between phyla [222–225], ρ-dependent termination is not essential in all bacteria [222,226] and there are different degrees of transcription-translation coupling [227].

Some gaps remain in our understanding of some already well-studied transcription regulatory mechanisms that involve co-transcriptional RNA folding, such as the structural consequences of a third strand in the RNA exit tunnel during hairpin-stabilized pausing or the precise structural basis of intrinsic termination. Likewise, it will be interesting to further decipher variations of recognized themes, e.g., during transcription anti-pausing/anti-termination by different members of the families of Q and N anti-termination factors. An important task will be the experimental delineation of RNAP regions beyond the exit channel, as well as of additional TFs and specific TF surfaces, involved in co-transcriptional RNA folding, processing and RNP assembly. Such surfaces may be involved in the co-transcriptional folding of *put, EAR, RARE*, or *iRAP* elements as well as of riboswitches. In addition, the role of RNAP and or TFs in the functional switches observed for some of these elements that depend, e.g., on transcriptional progression or ligand binding remain to be explored.

Ribosomes will remain major model systems to study co-transcriptional RNA folding, processing and RNP assembly mechanisms. Regarding rRNA folding and rRNP assembly in bacteria, it will be important to clarify the molecular basis for the functions of certain r-proteins as subunits

of *rrn*ECs, and whether precisely the same molecules are also used as rRNP building blocks that are co-transcriptionally incorporated into nascent subunits. To this end, single-molecule approaches that make use of authentically modified *rrn*ECs have to be developed. Regarding the analysis of the structural basis of co-transcriptional rRNP assembly, it would be most desirable to investigate endogenous assembly intermediates, e.g., by employing cryogenic electron tomography on small bacteria or in combination with focused ion beam milling. In principle, advanced cross-linking/mass spectrometry approaches also offer an opportunity for *in situ* structural analyses, but a challenge will be to attribute observed cross-links to specific assembly intermediates. Knock-out/knock-down of assembly factors that lead to enrichment of certain intermediates may offer a solution, and could also be helpful in their *ex vivo* structural analysis by single-particle cryoEM.

Further studies are also needed to unveil how the complex functional interplay between transcription and translation machineries might be influenced by co-transcriptional structure formation and how, in turn, transcription-translation coupling might impact co-transcriptional RNA folding. Given the observation that r-proteins also moonlight as transcription factors in other scenarios, such as L4 during transcriptional attenuation in the S10 operon, a particular challenge will be to disentangle the precise roles of r-proteins when conducting genome-wide or transcriptome-wide studies using, e.g., chromatin immunoprecipitation or cross-linking/immunoprecipitation-based technologies that target r-proteins. While certainly challenging, insights into such mechanisms may also open up new avenues in the design of transcription-modulatory drugs, including novel antibiotics.

Acknowledgments

Disclosure statement

Funding

ORCID

Nelly Said 🔟 http://orcid.org/0000-0001-8352-5555
Markus C. Wahl 🔟 http://orcid.org/0000-0002-2811-5307

References

[1] Pan T, Sosnick T. RNA folding during transcription. Annu Rev Biophys Biomol Struct. 2006;35(1):161–175.

[2] Kang JY, Mishanina TV, Landick R, et al. Mechanisms of transcriptional pausing in bacteria. J Mol Biol. 2019;431(20):4007–4029.

[3] Herbert KM, La Porta A, Wong BJ, et al. Sequence-resolved detection of pausing by single RNA polymerase molecules. Cell. 2006;125(6):1083–1094.

[4] Kang JY, Mishanina TV, Bellecourt MJ, et al. RNA polymerase accommodates a pause RNA hairpin by global conformational rearrangements that prolong pausing. Mol Cell. 2018;69(5):802–815 e805.

[5] Guo X, Myasnikov AG, Chen J, et al. Structural Basis for NusA Stabilized Transcriptional Pausing. Mol Cell. 2018;69(5):816–827 e814.

[6] Saba J, Chua XY, and Mishanina TV, et al. The elemental mechanism of transcriptional pausing. Elife. 2019 ;8:e40981.

[7] Larson MH, Mooney RA, Peters JM, et al. A pause sequence enriched at translation start sites drives transcription dynamics in vivo. Science. 2014;344(6187):1042–1047.

[8] Vvedenskaya IO, Vahedian-Movahed H, Bird JG, et al. Interactions between RNA polymerase and the "core recognition element" counteract pausing. Science. 2014;344(6189):1285–1289.

[9] Abdelkareem M, Saint-Andre C, Takacs M, et al. Structural basis of transcription: RNA polymerase backtracking and its reactivation. Mol Cell. 2019;75(2):298–309 e294.

[10] Widom JR, Nedialkov YA, Rai V, et al. Ligand modulates cross-coupling between riboswitch folding and

transcriptional pausing. Mol Cell. 2018;72(3):541–552 e546.

[11] Artsimovitch I, Landick R. The transcriptional regulator RfaH stimulates RNA chain synthesis after recruitment to elongation complexes by the exposed nontemplate DNA strand. Cell. 2002;109 (2):193–203.

[12] Kang JY, Mooney RA, Nedialkov Y, et al. Structural basis for transcript elongation control by nusg family universal regulators. Cell. 2018;173(7):1650–1662 e1614.

[13] Boyle J, Robillard GT, Kim SH. Sequential folding of transfer RNA. A nuclear magnetic resonance study of successively longer tRNA fragments with a common 5′ end. J Mol Biol. 1980;139(4):601–625.

[14] Kramer FR, Mills DR. Secondary structure formation during RNA synthesis. Nucleic Acids Res. 1981;9 (19):5109–5124.

[15] Lai D, Proctor JR, Meyer IM. On the importance of cotranscriptional RNA structure formation. RNA. 2013;19(11):1461–1473.

[16] Zhang J, Landick R, Two-Way Street: A. Regulatory interplay between RNA polymerase and nascent rna structure. Trends Biochem Sci. 2016;41(4):293–310.

[17] Brehm SL, Cech TR. Fate of an intervening sequence ribonucleic acid: excision and cyclization of the tetrahymena ribosomal ribonucleic acid intervening sequence in vivo. Biochemistry. 1983;22(10):2390–2397.

[18] Heilman-Miller SL, Woodson SA. Effect of transcription on folding of the Tetrahymena ribozyme. RNA. 2003;9(6):722–733.

[19] Pan T, Artsimovitch I, Fang XW, et al. Folding of a large ribozyme during transcription and the effect of the elongation factor NusA. Proc Natl Acad Sci U S A. 1999;96(17):9545–9550.

[20] Wong T, Sosnick TR, Pan T. Mechanistic insights on the folding of a large ribozyme during transcription. Biochemistry. 2005;44(20):7535–7542.

[21] Wong TN, Sosnick TR, Pan T. Folding of noncoding RNAs during transcription facilitated by pausing-induced nonnative structures. Proc Natl Acad Sci U S A. 2007;104(46):17995–18000.

[22] Fukuda S, Yan S, Komi Y, et al. The biogenesis of SRP RNA is modulated by an RNA folding intermediate attained during transcription. Mol Cell. 2020;77 (2):241–250 e248.

[23] Perdrizet GA 2nd, Artsimovitch I, Furman R, et al. Transcriptional pausing coordinates folding of the aptamer domain and the expression platform of a riboswitch. Proc Natl Acad Sci U S A. 2012;109 (9):3323–3328.

[24] Leamy KA, Assmann SM, Mathews DH, et al. Bridging the gap between in vitro and in vivo RNA folding. Q Rev Biophys. 2016;49:e10.

[25] Ganser LR, Kelly ML, Herschlag D, et al. The roles of structural dynamics in the cellular functions of RNAs. Nat Rev Mol Cell Biol. 2019;20(8):474–489.

[26] Licatalosi DD, Ye X, Jankowsky E. Approaches for measuring the dynamics of RNA-protein interactions. Wiley Interdiscip Rev RNA. 2020;11(1):e1565.

[27] Watters KE, Strobel EJ, Yu AM, et al. Cotranscriptional folding of a riboswitch at nucleotide resolution. Nat Struct Mol Biol. 2016;23(12):1124–1131.

[28] Strobel EJ, Watters KE, Nedialkov Y, et al. Distributed biotin-streptavidin transcription roadblocks for mapping cotranscriptional RNA folding. Nucleic Acids Res. 2017;45(12):e109.

[29] Yu AM, Gasper PM, Cheng L, et al. Computationally reconstructing cotranscriptional RNA folding from experimental data reveals rearrangement of non-native folding intermediates. Mol Cell. 2021;81 (4):870–883 e810.

[30] Incarnato D, Morandi E, Anselmi F, et al. In vivo probing of nascent RNA structures reveals principles of cotranscriptional folding. Nucleic Acids Res. 2017;45 (16):9716–9725.

[31] Bevilacqua PC, Assmann SM. Technique development for probing RNA structure in vivo and genome-wide. Cold Spring Harb Perspect Biol. 2018;10(10):10.

[32] Masachis S, Tourasse NJ, and Lays C, Elife. 2019;8: e47549.

[33] Frieda KL, Block SM. Direct observation of cotranscriptional folding in an adenine riboswitch. Science. 2012;338(6105):397–400.

[34] Uhm H, Kang W, Ha KS, et al. FRET studies on the cotranscriptional folding of a thiamine pyrophosphate riboswitch. Proc Natl Acad Sci U S A. 2018;115 (2):331–336.

[35] Dangkulwanich M, Ishibashi T, Bintu L, et al. Molecular mechanisms of transcription through single-molecule experiments. Chem Rev. 2014;114 (6):3203–3223.

[36] Hua B, Panja S, Wang Y, et al. RNA folding using a superhelicase. J Am Chem Soc. 2018;140 (32):10067–10070.

[37] Hua B, Jones CP, Mitra J, et al. Real-time monitoring of single ZTP riboswitches reveals a complex and kinetically controlled decision landscape. Nat Commun. 2020;11(1):4531.

[38] Woodson SA. Taming free energy landscapes with RNA chaperones. RNA Biol. 2010;7(6):677–686.

[39] Turowski TW, Petfalski E, Goddard BD, et al. Nascent transcript folding plays a major role in determining RNA polymerase elongation rates. Mol Cell. 2020;79 (3):488–503 e411.

[40] Turnbough CL Jr. Regulation of bacterial gene expression by transcription attenuation. Microbiol Mol Biol Rev. 2019;83:3.

[41] Nudler E, Gottesman ME. Transcription termination and anti-termination in E coli. Genes Cells. 2002;7 (8):755–768.

[42] Scharfen L, Neugebauer KM. Transcription regulation through nascent RNA folding. J Mol Biol. 2021;433 (14):166975.

[43] Liu SR, Hu CG, Zhang JZ. Regulatory effects of cotranscriptional RNA structure formation and transitions. Wiley Interdiscip Rev RNA. 2016;7(5):562–574.

[44] Gong S, Wang Y, Wang Z, et al. Co-transcriptional folding and regulation mechanisms of riboswitches. Molecules. 2017;22(7):7.

[45] Gong S, Wang Y, Wang Z, et al. Computational methods for modeling aptamers and designing riboswitches. Int J Mol Sci. 2017;18(11):11.

[46] Rajkowitsch L, Chen D, Stampfl S, et al. RNA chaperones, RNA annealers and RNA helicases. RNA Biol. 2007;4(3):118–130.

[47] Semrad K. Proteins with RNA chaperone activity: a world of diverse proteins with a common task-impediment of RNA misfolding. Biochem Res Int. 2011;2011:532908.

[48] Woodson SA, Panja S. Proteins that S-FA. Chaperone RNA regulation. Microbiol Spectr. 2018;6(4):4.

[49] Bohnsack KE, Mt B. RNA-binding proteins chaperone ribonucleoprotein complex Assembly to solve the RNA-folding problem. Cell. 2019;179(6):1248–1250.

[50] Jiang W, Hou Y, Inouye M. CspA, the major cold-shock protein of Escherichia coli, is an RNA chaperone. J Biol Chem. 1997;272(1):196–202.

[51] Rajkowitsch L, Dissecting SR. RNA chaperone activity. RNA. 2007;13(12):2053–2060.

[52] Iost I, Bizebard T, Dreyfus M. Functions of DEAD-box proteins in bacteria: current knowledge and pending questions. Biochim Biophys Acta. 2013;1829 (8):866–877.

[53] Hyeon C, Thirumalai D. Generalized iterative annealing model for the action of RNA chaperones. J Chem Phys. 2013;139(12):121924.

[54] Morgan EA. Antitermination mechanisms in rRNA operons of Escherichia coli. J Bacteriol. 1986;168 (1):1–5.

[55] Ha KS, Toulokhonov I, Vassylyev DG, et al. The NusA N-terminal domain is necessary and sufficient for enhancement of transcriptional pausing via interaction with the RNA exit channel of RNA polymerase. J Mol Biol. 2010;401(5):708–725.

[56] Ma C, Mobli M, Yang X, Ma C, Mobli M, Yang X, Keller AN, King GF, Lewis PJ. RNA polymerase-induced remodelling of NusA produces a pause enhancement complex. Nucleic Acids Res. 2015;43(5):2829–2840.

[57] Krupp F, Said N, Huang YH, et al. Structural basis for the action of an all-purpose transcription anti-termination factor. Mol Cell. 2019;74(1):143–157 e145.

[58] Toulokhonov I, Artsimovitch I, Landick R. Allosteric control of RNA polymerase by a site that contacts nascent RNA hairpins. Science. 2001;292 (5517):730–733.

[59] Toulokhonov I, Landick R. The flap domain is required for pause RNA hairpin inhibition of catalysis by RNA polymerase and can modulate intrinsic termination. Mol Cell. 2003;12(5):1125–1136.

[60] Kang JY, Olinares PD, and Chen J, et al. Structural basis of transcription arrest by coliphage HK022 Nun in an Escherichia coli RNA polymerase elongation complex. Elife. 2017;6:e25478.

[61] Chen J, Boyaci H, Campbell EA. Diverse and unified mechanisms of transcription initiation in bacteria. Nat Rev Microbiol. 2021;19(2):95–109.

[62] Goldman SR, Ebright RH, Nickels BE. Direct detection of abortive RNA transcripts in vivo. Science. 2009;324 (5929):927–928.

[63] Dulin D, Bauer DLV, Malinen AM, et al. Pausing controls branching between productive and non-productive pathways during initial transcription in bacteria. Nat Commun. 2018;9(1):1478.

[64] Hsu LM. Monitoring abortive initiation. Methods. 2009;47(1):25–36.

[65] Revyakin A, Liu C, Ebright RH, et al. Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. Science. 2006;314 (5802):1139–1143.

[66] Vassylyev DG, Vassylyeva MN, Perederina A, et al. Structural basis for transcription elongation by bacterial RNA polymerase. Nature. 2007;448(7150):157–162.

[67] BZ R, WS Y, JW R. Function of E. coli RNA polymerase sigma factor sigma 70 in promoter-proximal pausing. Cell. 1996;86(3):485–493.

[68] Mooney RA, Landick R. Tethering sigma70 to RNA polymerase reveals high in vivo activity of sigma factors and sigma70-dependent pausing at promoter-distal locations. Genes Dev. 2003;17 (22):2839–2851.

[69] Hatoum A, Roberts J. Prevalence of RNA polymerase stalling at Escherichia coli promoters after open complex formation. Mol Microbiol. 2008;68(1):17–28.

[70] Perdue SA, Roberts JW. A backtrack-inducing sequence is an essential component of Escherichia coli sigma(70)-dependent promoter-proximal pausing. Mol Microbiol. 2010;78(3):636–650.

[71] Marr MT, Roberts JW. Function of transcription cleavage factors GreA and GreB at a regulatory pause site. Mol Cell. 2000;6(6):1275–1285.

[72] Shankar S, Hatoum A, Roberts JW. A transcription antiterminator constructs a NusA-dependent shield to the emerging transcript. Mol Cell. 2007;27(6):914–927.

[73] Kuznedelov K, Minakhin L, Niedziela-Majka A, et al. A role for interaction of the RNA polymerase flap domain with the sigma subunit in promoter recognition. Science. 2002;295(5556):855–857.

[74] Vassylyev DG, Sekine S, Laptenko O, et al. Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 A resolution. Nature. 2002;417(6890):712–719.

[75] Nickels BE, Garrity SJ, Mekler V, et al. The interaction between sigma70 and the beta-flap of Escherichia coli RNA polymerase inhibits extension of nascent RNA during early elongation. Proc Natl Acad Sci U S A. 2005;102(12):4488–4493.

[76] Hollands K, Sevostiyanova A, Groisman EA. Unusually long-lived pause required for regulation of a Rho-dependent transcription terminator. Proc Natl Acad Sci U S A. 2014;111(19):E1999–2007.

[77] Landick R, Yanofsky C. Stability of an RNA secondary structure affects in vitro transcription pausing in the trp operon leader region. J Biol Chem. 1984;259 (18):11550–11555.

[78] Yakhnin AV, Yakhnin H, Babitzke P. RNA polymerase pausing regulates translation initiation by providing additional time for TRAP-RNA interaction. Mol Cell. 2006;24(4):547–557.

[79] Ray-Soni A, Bellecourt MJ, Landick R. Mechanisms of Bacterial Transcription Termination: All good things must end. Annu Rev Biochem. 2016;85(1):319–347.

[80] Sigmund CD, Morgan EA. Nus A protein affects transcriptional pausing and termination in vitro by binding to different sites on the transcription complex. Biochemistry. 1988;27(15):5622–5627.

[81] Gusarov I, Nudler E. Control of intrinsic transcription termination by N and NusA: the basic mechanisms. Cell. 2001;107(4):437–449.

[82] Chan CL, Landick R. Dissection of the his leader pause site by base substitution reveals a multipartite signal that includes a pause rna hairpin. J Mol Biol. 1993;233 (1):25–42.

[83] Artsimovitch I, Landick R. Interaction of a nascent RNA structure with RNA polymerase is required for hairpin-dependent transcriptional pausing but not for transcript release. Genes Dev. 1998;12 (19):3110–3122.

[84] Yarnell WS, Roberts JW. Mechanism of intrinsic transcription termination and antitermination. Science. 1999;284(5414):611–615.

[85] Hein PP, Kolb KE, Windgassen T, et al. RNA polymerase pausing and nascent-RNA structure formation are linked through clamp-domain movement. Nat Struct Mol Biol. 2014;21(9):794–802.

[86] Chen SJ. RNA folding: conformational statistics, folding kinetics, and ion electrostatics. Annu Rev Biophys. 2008;37(1):197–214.

[87] Brion P, Westhof E. Hierarchy and dynamics of RNA folding. Annu Rev Biophys Biomol Struct. 1997;26 (1):113–137.

[88] Lane WJ, Darst SA. Molecular evolution of multisubunit RNA polymerases: sequence analysis. J Mol Biol. 2010;395(4):671–685.

[89] Epshtein V, Cardinale CJ, Ruckenstein AE, et al. An allosteric path to transcription termination. Mol Cell. 2007;28(6):991–1001.

[90] Bellecourt MJ, Ray-Soni A, Harwig A, et al. RNA polymerase clamp movement Aids dissociation from DNA but is not required for RNA release at intrinsic terminators. J Mol Biol. 2019;431 (4):696–713.

[91] Kouba T, Koval T, Sudzinova P, et al. Mycobacterial HelD is a nucleic acids-clearing factor for RNA polymerase. Nat Commun. 2020;11(1):6419.

[92] Newing TP, Oakley AJ, Miller M, et al. Molecular basis for RNA polymerase-dependent transcription complex recycling by the helicase-like motor protein HelD. Nat Commun. 2020;11(1):6420.

[93] Pei HH, Hilal T, Chen ZA, et al. The delta subunit and NTPase HelD institute a two-pronged mechanism for RNA polymerase recycling. Nat Commun. 2020;11(1):6418.

[94] King RA, Banik-Maiti S, Jin DJ, King RA, Banik-Maiti S, Jin DJ, Weisberg RA. Transcripts that increase the processivity and elongation rate of RNA polymerase. Cell. 1996;87(5):893–903.

[95] Ra K, Ra W. Suppression of factor-dependent transcription termination by antiterminator RNA. J Bacteriol. 2003;185(24):7085–7091.

[96] Vitiello CL, Kireeva ML, Lubkowska L, et al. HK022 Nun protein inhibits RNA polymerase translocation. Proc Natl Acad Sci U S A. 2014;111(23):E2368–2375.

[97] Oberto J, Clerget M, Ditto M, et al. Antitermination of early transcription in phage HK022. absence of a phage-encoded antitermination factor. J Mol Biol. 1993;229(2):368–381.

[98] Banik-Maiti S, King RA, Weisberg RA, Banik-Maiti S, King RA, Weisberg RA. The antiterminator RNA of phage HK022. J Mol Biol. 1997;272(5):677–687.

[99] Komissarova N, Velikodvorskaya T, Sen R, Komissarova N, Velikodvorskaya T, Sen R, King RA, Banik-Maiti S, Weisberg RA. Inhibition of a transcriptional pause by RNA anchoring to RNA polymerase. Mol Cell. 2008;31(5):683–694.

[100] Clerget M, Jin DJ, Weisberg RA. A zinc-binding region in the beta' subunit of RNA polymerase is involved in antitermination of early transcription of phage HK022. J Mol Biol. 1995;248(4):768–780.

[101] Sen R, King RA, Weisberg RA, Sen R, King RA, Weisberg RA. Modification of the properties of elongating RNA polymerase by persistent association with nascent antiterminator RNA. Mol Cell. 2001;7(5):993–1001.

[102] Irnov I, Winkler WC. A regulatory RNA required for antitermination of biofilm and capsular polysaccharide operons in Bacillales. Mol Microbiol. 2010;76(3):559–575.

[103] Sevostyanova A, Groisman EA. An RNA motif advances transcription by preventing Rho-dependent termination. Proc Natl Acad Sci U S A. 2015;112(50): E6835–6843.

[104] Mooney RA, Davis SE, Peters JM, et al. Regulator trafficking on bacterial transcription units in vivo. Mol Cell. 2009;33(1):97–108.

[105] Artsimovitch I, Landick R. Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. Proc Natl Acad Sci U S A. 2000;97(13):7090–7095.

[106] Said N, Krupp F, Anedchenko E, et al. Structural basis for lambdaN-dependent processive transcription antitermination. Nat Microbiol. 2017;2(7):17062.

[107] Huang YH, Hilal T, Loll B, et al. Structure-based mechanisms of a molecular RNA polymerase/chaperone machine required for ribosome biosynthesis. Mol Cell. 2020;79(6):1024–1036 e1025.

[108] Huang YH, Said N, Loll B, et al. Structural basis for the function of SuhB as a transcription factor in ribosomal RNA synthesis. Nucleic Acids Res. 2019;47(12):6488–6503.

[109] Dudenhoeffer BR, Schneider H, Schweimer K, et al. SuhB is an integral part of the ribosomal antitermination complex and interacts with NusA. Nucleic Acids Res. 2019;47(12):6504–6518.

[110] Dudenhoeffer BR, Borggraefe J, Schweimer K, et al. NusA directly interacts with antitermination factor Q from phage lambda. Sci Rep. 2020;10(1):6607.

[111] Bonin I, Muhlberger R, Bourenkov GP, et al. Structural basis for the interaction of Escherichia coli NusA with protein N of phage lambda. Proc Natl Acad Sci U S A. 2004;101(38):13762–13767.

[112] Strauss M, Vitiello C, Schweimer K, et al. Transcription is regulated by NusA:NusG interaction. Nucleic Acids Res. 2016;44(12):5971–5982.

[113] Said N, Hilal T, Sunday ND, et al. Steps toward translocation-independent RNA polymerase inactivation by terminator ATPase rho. Science. 2021;371(6524):6524.

[114] Hao Z, Epshtein V, and Kim KH, et al. Pre-termination Transcription Complex: structure and function. Mol Cell. 2021;81(2):281–292 e288.

[115] Wang C, Molodtsov V, Firlar E, et al. Structural basis of transcription-translation coupling. Science. 2020;369(6509):1359–1365.

[116] Belogurov GA, Artsimovitch I. Regulation of transcript Elongation. Annu Rev Microbiol. 2015;69(1):49–69.

[117] Worbs M, Bourenkov GP, Bartunik HD, et al. An extended RNA binding surface through arrayed S1 and KH domains in transcription factor NusA. Mol Cell. 2001;7(6):1177–1189.

[118] Eisenmann A, Schwarz S, Prasch S, et al. The E. coli NusA carboxy-terminal domains are structurally similar and show specific RNAP- and lambdaN interaction. Protein Sci. 2005;14(8):2018–2029.

[119] Schweimer K, Prasch S, Sujatha PS, et al. NusA interaction with the alpha subunit of E. coli RNA polymerase is via the UP element site and releases autoinhibition. Structure. 2011;19(7):945–954.

[120] Tabib-Salazar A, Mulvenna N, Severinov K, et al. Xenogeneic regulation of the bacterial transcription machinery. J Mol Biol. 2019;431(20):4078–4092.

[121] Roberts JW, Yarnell W, Bartlett E, et al. Antitermination by bacteriophage lambda Q protein. Cold Spring Harb Symp Quant Biol. 1998;63 319–325.

[122] Santangelo TJ, Artsimovitch I. Termination and antitermination: RNA polymerase runs a stop sign. Nat Rev Microbiol. 2011;9(5):319–329.

[123] Shi J, Gao X, Tian T, et al. Structural basis of Q-dependent transcription antitermination. Nat Commun. 2019;10(1):2925.

[124] Yin Z, Kaelber JT, Ebright RH. Structural basis of Q-dependent antitermination. Proc Natl Acad Sci U S A. 2019;116(37):18384–18390.

[125] Wells CD, Deighan P, Brigham M, et al. RNA length dictates opposing effects of NusA on antitermination. Nucleic Acids Res. 2016;44(11):5378–5389.

[126] Nodwell JR, Greenblatt J. The nut site of bacteriophage lambda is made of RNA and is bound by transcription antitermination factors on the surface of RNA polymerase. Genes Dev. 1991;5(11):2141–2151.

[127] Legault P, Li J, Mogridge J, et al. NMR structure of the bacteriophage lambda N peptide/boxB RNA complex: recognition of a GNRA fold by an arginine-rich motif. Cell. 1998;93(2):289–299.

[128] Greenblatt J, Nodwell JR, Mason SW. Transcriptional antitermination. Nature. 1993;364(6436):401–406.

[129] Goodson JR, Winkler WC, Storz G. Processive antitermination. Microbiol Spectr. 2018;6(5). DOI:10.1128/microbiolspec.RWR-0031-2018

[130] Wahl MC, Sen R. Exploiting phage strategies to modulate bacterial transcription. Transcription. 2019;10(4–5):222–230.

[131] Nodwell JR, Greenblatt J. Recognition of boxA antiterminator RNA by the E. coli antitermination factors NusB and ribosomal protein S10. Cell. 1993;72(2):261–268.

[132] Wang B, Artsimovitch I. NusG, an ancient yet rapidly evolving transcription factor. Front Microbiol. 2020;11:619618.

[133] Steiner T, Kaiser JT, Marinkovic S, et al. Crystal structures of transcription factor NusG in light of its nucleic acid- and protein-binding activities. EMBO J. 2002;21(17):4641–4653.

[134] Burova E, Hung SC, Sagitov V, et al. Escherichia coli NusG protein stimulates transcription elongation rates in vivo and in vitro. J Bacteriol. 1995;177(5):1388–1392.

[135] Burmann BM, Schweimer K, Luo X, et al. A NusE: NusG complex links transcription and translation. Science. 2010;328(5977):501–504.

[136] Lawson MR, Ma W, Bellecourt MJ, et al. Mechanism for the regulated control of bacterial transcription termination by a universal adaptor protein. Mol Cell. 2018;71(6):911–922 e914.

[137] Luo X, Hsiao HH, Bubunenko M, et al. Structural and functional analysis of the E. coli NusB-S10 transcription antitermination complex. Mol Cell. 2008;32 (6):791–802.

[138] Saxena S, Myka KK, Washburn R, et al. Escherichia coli transcription factor NusG binds to 70S ribosomes. Mol Microbiol. 2018;108(5):495–504.

[139] Cheeran A, Kolli NR, Sen R. The site of action of the antiterminator protein N from the lambdoid phage H-19B. J Biol Chem. 2007;282(42):30997–31007.

[140] Davis JH, Williamson JR. Structure and dynamics of bacterial ribosome biogenesis. Philos Trans R Soc Lond B Biol Sci. 2017;372(1716):1716.

[141] Ginsburg D, Steitz JA. The 30 S ribosomal precursor RNA from Escherichia coli. A primary transcript containing 23 S, 16 S, and 5 S sequences. J Biol Chem. 1975;250(14):5647–5654.

[142] Kaczanowska M, Ryden-Aulin M. Ribosome biogenesis and the translation process in Escherichia coli. Microbiol Mol Biol Rev. 2007;71(3):477–494.

[143] French SL, Miller OL Jr. Transcription mapping of the Escherichia coli chromosome by electron microscopy. J Bacteriol. 1989;171(8):4207–4216.

[144] Gotta SL, Miller OL Jr., French SL. rRNA transcription rate in Escherichia coli. J Bacteriol. 1991;173 (20):6647–6649.

[145] Wilson DN, Nierhaus KH. The weird and wonderful world of bacterial ribosome regulation. Crit Rev Biochem Mol Biol. 2007;42(3):187–219.

[146] Gegenheimer P, Apirion D. Escherichia coli ribosomal ribonucleic acids are not cut from an intact precursor molecule. J Biol Chem. 1975;250(6):2407–2409.

[147] Traub P, Nomura M. Structure and function of E. coli ribosomes. V. Reconstitution of functionally active 30S ribosomal particles from RNA and proteins. Proc Natl Acad Sci U S A. 1968;59(3):777–784.

[148] Nierhaus KH, Dohme F. Total reconstitution of functionally active 50S ribosomal subunits from Escherichia coli. Proc Natl Acad Sci U S A. 1974;71(12):4713–4717.

[149] Aksoy S, Squires CL, Squires C. Evidence for antitermination in Escherichia coli RRNA transcription. J Bacteriol. 1984;159(1):260–264.

[150] Berg KL, Squires C, Squires CL. Ribosomal RNA operon anti-termination. function of leader and spacer region box B-box A sequences and their conservation in diverse micro-organisms. J Mol Biol. 1989;209 (3):345–358.

[151] Condon C, Squires C, Squires CL. Control of rRNA transcription in Escherichia coli. Microbiol Rev. 1995;59(4):623–645.

[152] Squires CL, Greenblatt J, Li J, et al. Ribosomal RNA antitermination in vitro: requirement for Nus factors and one or more unidentified cellular components. Proc Natl Acad Sci U S A. 1993;90(3):970–974.

[153] Torres M, Condon C, Balada JM, et al. Ribosomal protein S4 is a transcription factor with properties remarkably similar to NusA, a protein involved in both non-ribosomal and ribosomal RNA antitermination. EMBO J. 2001;20(14):3811–3820.

[154] Singh N, Bubunenko M, Smith C, et al. SuhB associates with Nus factors to facilitate 30S ribosome biogenesis in Escherichia coli. mBio. 2016;7(2):e00114.

[155] Wang Y, Stieglitz KA, Bubunenko M, et al. The structure of the R184A mutant of the inositol monophosphatase encoded by suhB and implications for its functional interactions in Escherichia coli. J Biol Chem. 2007;282(37):26989–26996.

[156] Vogel U, Jensen KF. Effects of the antiterminator BoxA on transcription elongation kinetics and ppGpp inhibition of transcription elongation in Escherichia coli. J Biol Chem. 1995;270(31):18335–18340.

[157] Klumpp S, Hwa T. Stochasticity and traffic jams in the transcription of ribosomal RNA: intriguing role of termination and antitermination. Proc Natl Acad Sci U S A. 2008;105(47):18159–18164.

[158] Rodgers ML, Woodson SA. Transcription increases the cooperativity of ribonucleoprotein assembly. Cell. 2019;179(6):1370–1381 e1312.

[159] Duss O, Stepanyuk GA, Puglisi JD, et al. Transient protein-RNA interactions guide nascent ribosomal RNA folding. Cell. 2019;179(6):1357–1369 e1316.

[160] Lewicki BT, Margus T, Remme J, et al. Coupling of rRNA transcription and ribosomal assembly in vivo. formation of active ribosomal subunits in Escherichia coli requires transcription of rRNA genes by host RNA polymerase which cannot be replaced by bacteriophage T7 RNA polymerase. J Mol Biol. 1993;231 (3):581–593.

[161] Bubunenko M, Court DL, Al Refaii A, et al. Nus transcription elongation factors and RNase III modulate small ribosome subunit biogenesis in Escherichia coli. Mol Microbiol. 2013;87(2):382–393.

[162] Fernandez-Millan P, Autour A, Ennifar E, et al. Crystal structure and fluorescence properties of the iSpinach aptamer in complex with DFHBI. RNA. 2017;23 (12):1788–1795.

[163] Desai R, Kilburn D, Lee HT, et al. Increased ribozyme activity in crowded solutions. J Biol Chem. 2014;289 (5):2972–2977.

[164] Gamerdinger M. Protein quality control at the ribosome: focus on RAC, NAC and RQC. Essays Biochem. 2016;60(2):203–212.

[165] Proshkin S, Rahmouni AR, Mironov A, et al. Cooperation between translating ribosomes and RNA polymerase in transcription elongation. Science. 2010;328(5977):504–508.

[166] Ab C, Diggs S, TK T, et al. Two old dogs, one new trick: a review of rna polymerase and ribosome interactions during transcription-translation coupling. Int J Mol Sci. 2019;20:10.

[167] Henkin TM, Yanofsky C. Regulation by transcription attenuation in bacteria: how RNA provides instructions for transcription termination/antitermination decisions. Bioessays. 2002;24(8):700–707.

[168] Webster MW, Takacs M, Zhu C, et al. Structural basis of transcription-translation coupling and collision in bacteria. Science. 2020;369(6509):1355–1359.

[169] O'Reilly FJ, Xue L, Graziadei A, et al. In-cell architecture of an actively transcribing-translating expressome. Science. 2020;369(6503):554–557.

[170] Chatterjee S, Chauvier A, Dandpat SS, et al. A translational riboswitch coordinates nascent transcription-translation coupling. Proc Natl Acad Sci U S A. 2021;118(16):16.

[171] Nierhaus KH. The assembly of prokaryotic ribosomes. Biochimie. 1991;73(6):739–755.

[172] Held WA, Ballou B, Mizushima S, et al. Assembly mapping of 30 S ribosomal proteins from Escherichia coli. further studies. J Biol Chem. 1974;249(10):3103–3111.

[173] Woodson SA. RNA folding and ribosome assembly. Curr Opin Chem Biol. 2008;12(6):667–673.

[174] Sykes MT, Williamson JR. A complex assembly landscape for the 30S ribosomal subunit. Annu Rev Biophys. 2009;38(1):197–215.

[175] Shajani Z, Sykes MT, Williamson JR. Assembly of bacterial ribosomes. Annu Rev Biochem. 2011;80(1):501–526.

[176] Woodson SA. RNA folding pathways and the self-assembly of ribosomes. Acc Chem Res. 2011;44(12):1312–1319.

[177] Shiba K, Ito K, Nakamura Y, et al. Altered translation initiation factor 2 in the cold-sensitive ssyG mutant affects protein export in Escherichia coli. EMBO J. 1986;5(11):3001–3006.

[178] Duss O, Stepanyuk GA, Grot A, et al. Real-time assembly of ribonucleoprotein complexes on nascent RNA transcripts. Nat Commun. 2018;9(1):5087.

[179] Young RA, Steitz JA. Complementary sequences 1700 nucleotides apart form a ribonuclease III cleavage site in Escherichia coli ribosomal precursor RNA. Proc Natl Acad Sci U S A. 1978;75(8):3593–3597.

[180] Nowotny V, Nierhaus KH. Assembly of the 30S subunit from Escherichia coli ribosomes occurs via two assembly domains which are initiated by S4 and S7. Biochemistry. 1988;27(18):7051–7055.

[181] Sapag A, Vartikar JV, Draper DE. Dissection of the 16S rRNA binding site for ribosomal protein S4. Biochim Biophys Acta. 1990;1050(1–3):34–37.

[182] Grinwald M, Ron EZ, Sugiura R. The Escherichia coli translation-associated heat shock protein YbeY is involved in rRNA transcription antitermination. PLoS One. 2013;8(4):e62297.

[183] Peters JM, Vangeloff AD, Landick R. Bacterial transcription terminators: the RNA 3′-end chronicles. J Mol Biol. 2011;412(5):793–813.

[184] Mitra P, Ghosh G, Hafeezunnisa M, et al. Rho protein: Roles and mechanisms. Annu Rev Microbiol. 2017;71(1):687–709.

[185] Roberts JW. Mechanisms of bacterial transcription termination. J Mol Biol. 2019;431(20):4030–4039.

[186] Schmidt MC, Chamberlin MJ. Binding of rho factor to Escherichia coli RNA polymerase mediated by nusA protein. J Biol Chem. 1984;259(24):15000–15002.

[187] Epshtein V, Dutta D, Wade J, et al. An allosteric mechanism of Rho-dependent transcription termination. Nature. 2010;463(7278):245–249.

[188] Peters JM, Mooney RA, Kuan PF, et al. Rho directs widespread termination of intragenic and stable RNA transcription. Proc Natl Acad Sci U S A. 2009;106(36):15406–15411.

[189] Qayyum MZ, Dey D, Sen R. Transcription Elongation Factor NusA Is a General Antagonist of Rho-dependent Termination in Escherichia coli. J Biol Chem. 2016;291(15):8090–8108.

[190] Sedlyarova N, Rescheneder P, Magan A, et al. RNA polymerase aptamers regulate transcription in E coli. Mol Cell. 2017;67(1):30–43 e36.

[191] Sedlyarova N, Shamovsky I, Bharati BK, et al. sRNA-mediated control of transcription termination in E coli. Cell. 2016;167(1):111–121 e113.

[192] Chen J, Morita T, Gottesman S. Regulation of transcription termination of small rnas and by small RNAs: Molecular mechanisms and biological functions. Front Cell Infect Microbiol. 2019;9:201.

[193] Bossi L, Figueroa-Bossi N, Bouloc P, et al. Regulatory interplay between small RNAs and transcription termination factor Rho. Biochim Biophys Acta Gene Regul Mech. 2020;1863(7):194546.

[194] Subramanian AR. Structure and functions of ribosomal protein S1. Prog Nucleic Acid Res Mol Biol. 1983;28:101–142.

[195] Bycroft M, Hubbard TJ, Proctor M, et al. The solution structure of the S1 RNA binding domain: a member of an ancient nucleic acid-binding fold. Cell. 1997;88(2):235–242.

[196] Duval M, Korepanov A, Fuchsbauer O, et al. Escherichia coli ribosomal protein S1 unfolds structured mRNAs onto the ribosome for active translation initiation. PLoS Biol. 2013;11(12):e1001731.

[197] Sukhodolets MV, Garges S. Interaction of Escherichia coli RNA polymerase with the ribosomal protein S1 and the Sm-like ATPase Hfq. Biochemistry. 2003;42(26):8022–8034.

[198] Eason IR, Kaur HP, Alexander KA, et al. Growth phase-specific changes in the composition of E. coli transcription complexes. J Chromatogr B Analyt Technol Biomed Life Sci. 2019;1109:155–165.

[199] Sukhodolets MV, Garges S, Adhya S. Ribosomal protein S1 promotes transcriptional cycling. RNA. 2006;12(8):1505–1513.

[200] Mogridge J, Greenblatt J. Specific binding of Escherichia coli ribosomal protein S1 to boxA transcriptional antiterminator RNA. J Bacteriol. 1998;180(8):2248–2252.

[201] Demo G, Rasouly A, and Vasilyev N, et al. Structure of RNA polymerase bound to ribosomal 30S subunit. Elife. 2017;6:e28560.

[202] Takeshita D, Yamashita S, Tomita K. Molecular insights into replication initiation by Qbeta replicase using ribosomal protein S1. Nucleic Acids Res. 2014;42 (16):10809–10822.

[203] Tv V, Cm R. Ribosomal protein S1 and NusA protein complexed to recombination protein beta of phage lambda. J Bacteriol. 1993;175(6):1844–1846.

[204] Zengel JM, Lindahl L. Diverse mechanisms for regulating ribosomal protein synthesis in Escherichia coli. Prog Nucleic Acid Res Mol Biol. 1994;47:331–370.

[205] Meyer MM, Storz G, Papenfort K. rRNA mimicry in RNA Regulation of Gene Expression. Microbiol Spectr. 2018;6(2):2.

[206] Zengel JM, Lindahl L. Escherichia coli ribosomal protein L4 stimulates transcription termination at a specific site in the leader of the S10 operon independent of L4-mediated inhibition of translation. J Mol Biol. 1990;213(1):67–78.

[207] Lindahl L, Archer RH, McCormick JR, et al. Translational coupling of the two proximal genes in the S10 ribosomal protein operon of Escherichia coli. J Bacteriol. 1989;171(5):2639–2645.

[208] Zengel JM, Lindahl L. Ribosomal protein L4 stimulates in vitro termination of transcription at a NusA-dependent terminator in the S10 operon leader. Proc Natl Acad Sci U S A. 1990;87(7):2675–2679.

[209] Zengel JM, Lindahl L. A hairpin structure upstream of the terminator hairpin required for ribosomal protein L4-mediated attenuation control of the S10 operon of Escherichia coli. J Bacteriol. 1996;178(8):2383–2387.

[210] Sha Y, Lindahl L, Zengel JM. Role of NusA in L4-mediated attenuation control of the S10 r-protein operon of Escherichia coli. J Mol Biol. 1995;245(5):474–485.

[211] Stelzl U, Zengel JM, Tovbina M, et al. RNA-structural mimicry in Escherichia coli ribosomal protein L4-dependent regulation of the S10 operon. J Biol Chem. 2003;278(30):28237–28245.

[212] Zengel JM, Sha Y, Lindahl L. Surprising flexibility of leader RNA determinants for r-protein L4-mediated transcription termination in the Escherichia coil S10 operon. RNA. 2002;8(5):572–578.

[213] Vogel J, Luisi BF. Hfq and its constellation of RNA. Nat Rev Microbiol. 2011;9(8):578–589.

[214] Santiago-Frangos A, Woodson SA. Hfq chaperone brings speed dating to bacterial sRNA. Wiley Interdiscip Rev RNA. 2018;9(4):e1475.

[215] Waters LS, Regulatory SG. RNAs in bacteria. Cell. 2009;136(4):615–628.

[216] Le Derout J, Boni IV, Regnier P, et al. Hfq affects mRNA levels independently of degradation. BMC Mol Biol. 2010;11(1):17.

[217] Zhang A, Wassarman KM, Rosenow C, et al. Global analysis of small RNA and mRNA targets of Hfq. Mol Microbiol. 2003;50(4):1111–1124.

[218] Rabhi M, Espeli O, Schwartz A, et al. RNA chaperone Hfq mediates transcription antitermination at Rho-dependent terminators. EMBO J. 2011;30 (14):2805–2816.

[219] Kambara TK, Ramsey KM, Dove SL. Pervasive targeting of nascent transcripts by Hfq. Cell Rep. 2018;23 (5):1543–1552.

[220] Gebhardt MJ, Kambara TK, Ramsey KM, et al. Widespread targeting of nascent transcripts by RsmA in pseudomonas aeruginosa. Proc Natl Acad Sci U S A. 2020;117(19):10520–10529.

[221] Lane WJ, Darst SA. Molecular evolution of multisubunit RNA polymerases: structural analysis. J Mol Biol. 2010;395(4):686–704.

[222] Ingham CJ, Dennis J, Furneaux PA. Autogenous regulation of transcription termination factor Rho and the requirement for Nus factors in bacillus subtilis. Mol Microbiol. 1999;31(2):651–663.

[223] Yakhnin AV, NusA-stimulated BP. RNA polymerase pausing and termination participates in the bacillus subtilis trp operon attenuation mechanism invitro. Proc Natl Acad Sci U S A. 2002;99 (17):11067–11072.

[224] Arnvig KB, Zeng S, Quan S, et al. Evolutionary comparison of ribosomal operon antitermination function. J Bacteriol. 2008;190(21):7251–7257.

[225] Yakhnin AV, FitzGerald PC, McIntosh C, et al. NusG controls transcription pausing and RNA polymerase translocation throughout the bacillus subtilis genome. Proc Natl Acad Sci U S A. 2020;117 (35):21628–21636.

[226] Nicolas P, Mader U, Dervyn E, et al. Condition-dependent transcriptome reveals high-level regulatory architecture in bacillus subtilis. Science. 2012;335 (6072):1103–1106.

[227] Johnson GE, Lalanne JB, Peters ML, et al. Functionally uncoupled transcription-translation in bacillus subtilis. Nature. 2020;585(7823):124–128.