

Structure of the 30S translation initiation complex coupled to paused RNA polymerase and its potential for riboregulation

Received: 20 January 2025

Accepted: 27 November 2025

Published online: 13 December 2025

 Check for updates

Johann J. Roske¹, Giulia Paris¹, Akanksha Goyal², Marina Rodnina^{1,2}, Nikolay Zenkin^{1,3}, Katarzyna J. Bandyra^{1,4}✉ & Ben F. Luisi¹✉

In many bacterial species, transcription and translation can be coupled physically, with potential impact on the rates and efficiency of gene expression. Here, we present structural evidence from cryo-EM demonstrating that a bacterial RNA polymerase that is paused proximally to the promoter can associate with the pioneering 30S translation initiation complex (30S IC). These findings suggest that the physical link between transcription and translation can be established prior to commitment to protein synthesis. Although the mRNA is embedded in this 'early expressome' complex, it can nonetheless interact with small regulatory RNA (sRNA) and be targeted for cleavage in the protein-coding region by the RNA degradosome assembly in vitro. The potential tagging of transcripts with sRNA during pioneering and subsequent stages of translation initiation, when the 30S IC is at the 5' end of a polyribosome, may in principle contribute to efficient and rapid termination of gene expression in response to regulatory signals.

In prokaryotic cells, the cytoplasmic milieu and chromatin intermingle. Consequently, the machinery of transcription and translation can interact and operate concurrently on the same nascent mRNA molecule^{1–4}. Transcription-translation coupling (TTC) appears to occur extensively in representative gram-negative bacterial species, but its extent can be affected by growth conditions and varies in different bacterial lineages. For example, in some firmicutes, there may be comparatively fewer encounters between RNA polymerase (RNAP) with the ribosome and TTC may be infrequent^{5–7}. The physical and kinetic basis of TTC has been studied extensively in *Escherichia coli*, and structural data from cryo-electron microscopy (cryo-EM) has provided details of the component interactions in complexes formed between RNAP, the 30S small ribosomal subunit, and the fully assembled 70S ribosome particle^{8–13}. Interactions between the elongating RNAP and 70S ribosomes have been visualised *in situ* in *Mycoplasma pneumoniae* under conditions of antibiotic stress¹⁴, implicating the *in vivo* relevance of that transient assembly. *In vitro* structural and

single-molecule studies of the transcription-translation complexes, referred to as the "expressome", reveal different configurations between RNAP and 70S are possible, e.g., in collided (TTC-A) or coupled (TTC-B) modes, that depend on the spacer length of the linking mRNA transcript and the recruitment of transcription factors NusG and NusA that can bridge the interface of the RNAP-70S to form an assembly^{9–11,13,15,16}. A recent study systematically increased the mRNA spacer length up to 20 codons and described the possibility for the spacer to loop out between the physically coupled transcription elongation complexes (TEC) and 70S ribosome in states of long-range coupling or loose coupling (TTC-LC)¹⁶.

The occurrence of TTC likely depends on relative rates of transcription and translation, and the frequency of transcription pausing by the RNAP. After transcription initiation, mRNA transcript elongation is highly processive but discontinuous, with a strictly regulated balance between elongation and transcription pausing¹⁷. The nascent mRNA itself can directly regulate sequence-dependent transcriptional

¹Department of Biochemistry, Sanger Building, University of Cambridge, Cambridge, UK. ²Max Planck Institute for Multidisciplinary Sciences, Göttingen, Germany. ³Centre for Bacterial Cell Biology, Biosciences Institute, Faculty of Medical Sciences, Newcastle University, Newcastle Upon Tyne, UK.

⁴Present address: Faculty of Chemistry, Biological and Chemical Research Centre, University of Warsaw, Warsaw, Poland. ✉e-mail: k.bandyra@uw.edu.pl; bfl20@cam.ac.uk

pausing without requirement for any co-factors^{18,19}. Transcript elongation of many genes is paused within the first 200 nts, which serves early regulation mechanisms, including the recruitment of NusG to facilitate TTC formation^{20–23}. Promotor-proximal pausing could provide the window of opportunity required to complete translation initiation, which occurs on a timescale of 10–30 s^{24–26}. However, it remains unclear whether the establishment of TTC occurs only once the pioneering 70S ribosome catches up and collides with the paused RNAP, which may have detrimental effects on transcription fidelity²⁷. Recent structural evidence shows that RNAP can recruit the 30S ribosomal subunit through NusG, potentially as an early step of translation initiation²⁸. Translation initiation and ribosome catch-up are proposed to involve the TTC-LC mentioned above, with a long, >11 codons, looped-out mRNA spacer between TEC and ribosome, and with accessibility of the mRNA spacer to regulatory factors¹⁶. Subsequent coupled transcription and translation involves a short-range, tightly coupled transcription-translation complex (TTC-B), with a different orientation of ribosome relative to TEC, with short, 7–11 codons, mRNA spacer and with little or no accessibility of the mRNA spacer to regulatory factors¹⁶.

Here, we have prepared *E. coli* 30S translation initiation complexes (30S ICs), including initiation factors IF1, IF2 and IF3 and initiator formylmethionyl-tRNA (fMet-tRNA^{fMet}), and solved cryo-EM structures of the 30S IC in two states. We further assembled 30S ICs on reconstituted TEC with a long nascent mRNA that mimics a state of promoter-proximal transcription pausing. Cryo-EM single-particle analysis shows that NusG establishes a physical link between the 30S IC and the paused RNAP, despite the long (>70 nts) mRNA spacer and at in vitro concentrations far below the proposed dissociation constant of the NusE-NusG interaction²⁹. The obtained complex displays a large degree of flexibility between the TEC and the 30S subunit, similar to previous observations in the absence of NusA and the 50S subunit²⁸. The physically coupled complex of paused transcription and translation initiation, or “early expressome”, may represent a transient precursor to mature expressomes *in vivo*.

We explore if the early expressome can be the target for RNA-mediated regulation by testing if the 30S IC-recruited transcript can be accessed by small regulatory RNA (sRNA) that represses the expression of the target gene through base-pairing complementarity with the mRNA. We used the sRNA MicC that directs cleavage and turnover of the *ompD* transcript encoding the outer membrane protein OmpD in *Salmonella Typhimurium*. Assisted by the RNA chaperone Hfq, MicC recognizes *ompD* and induces downstream *ompD*-cleavage by the endoribonuclease RNase E^{30,31}. Considering that *ompD*-targeting occurs in the coding region (at codons 23–26) and that the ternary complex of MicC and Hfq on *ompD* does not prevent ribosome engagement or translocation³⁰, the 30S IC and a TEC positioned downstream of the recognition sequence for MicC could potentially be targets for riboregulation. We observe that the *ompD* transcript in the early expressome can be cleaved by the endoribonuclease RNase E within the multi-enzyme RNA degradosome at a defined site that is directed by the MicC sRNA. Although the rate of mRNA degradation is slower when the target transcript lies within the early expressome, the preference of RNase E for the sRNA-guided cleavage site over alternative, unspecific cleavage positions is increased, suggesting that the early expressome may help expose the seed pairing region in the mRNA for recognition by sRNA and Hfq. Our cryo-EM data indicates that the flexible tether between RNAP and 30S ribosomal subunit can allow the early expressome to accommodate transcript recognition and degradation close to the RNAP. Thus, based on evidence presented here and in other studies, we propose that the establishment of TTC may serve as a transient checkpoint for the nascent mRNA intermediate, which once tagged by sRNA, becomes the target for co-transcriptional degradation.

Results

Cryo-EM structure of *E. coli* 30S translation initiation complexes on the *ompD* mRNA

We had previously prepared 30S IC on a 187 nts long fragment of the *ompD* transcript encompassing the 5'-UTR from position -69 up to +118 (with +1 being A in the AUG start codon) and showed that this *ompD*-30S IC can be recognised by MicC-Hfq and targeted for deactivating cleavage at position +83 by the RNA degradosome³². Here, we again reconstituted 30S IC on *ompD* together with fMet-tRNA^{fMet} and initiation factors IF1, IF2 and IF3, and determined their structure by cryo-EM single-particle analysis. The image processing resulted in two distinct structures of the *ompD*-30S IC that have IF2 and either IF3 or initiator tRNA and IF2 recruited (Fig. 1, “Methods”, Suppl Fig. 1 and Table 1). Both reconstructions contain well-defined density for the 30S ribosomal subunit, as well as IF1 and the mRNA, which are recruited in previously described positions³³. The mRNA is bound to the cavity that is formed between head and body regions of the 30S subunit, and the Shine-Dalgarno sequence (SD) in the 5'-UTR forms a duplex with the anti-SD of the 16S rRNA. IF3 is bound at the 30S near IF1 and mRNA (Fig. 1A, B) in an orientation corresponding to a previously reported model of the *T. thermophilus* 30S IC in “state 1”³³. This arrangement also agrees with a recent report of *E. coli* IF3-bound 30S in which the IF3-CTD is positioned at the P-site³⁴.

Our second structure, which originates from the same specimen but was reconstructed from a separate particle subset, lacks IF3 but contains IF2 and fMet-tRNA^{fMet} in addition to IF1 and the 24 nts segment of the *ompD* mRNA (Fig. 1C). The start codon of the mRNA forms triplet base pairs with the anticodon of fMet-tRNA^{fMet} that is bound in the P-site (Fig. 1D), confirming that the 30S IC is assembled at the AUG of the *ompD* mRNA fragment. The arrangement of the IC components on the 30S subunit is consistent with previously reported models of the *T. thermophilus* 30S IC in “state 4”³³. Additionally, our cryo-EM density resolves the N2 subdomain of the *E. coli* IF2 (residues 295–370) which attach to helix 16 of the 16S rRNA, resembling reported structures of *P. aeruginosa* IF2 at the 70S-IC³⁵ (Fig. 1E). 3D variability analysis indicates that the head region of the 30S subunit has a lower degree of mobility with respect to the 30S body, compared to the structure that contains IF3 but lacks fMet-tRNA^{fMet} (Supplementary Movie 1).

Ribosomal bS1 protein contacts the 5'-UTR at the 30S shoulder

We observe additional cryo-EM density at the 5' mRNA exit site that allows unambiguous placement of the first two domains of bS1 (residues 5–176) in an orientation that is in agreement with previous observations^{28,36} (Fig. 1C right, 1F and Suppl Fig. 2A). Notably, the relative positions of the first two OB domains of bS1 domains differs most from available 70S ribosome and expressome structures^{10,11,37,38}, but agrees with reports of the 30S-tRNA-RNAP complex²⁸, suggesting that bS1 can undergo large conformational changes (Suppl. Fig. 2A).

At lower map-thresholds, the density for bS1 in our *ompD*-30S IC structure further extends from the described binding site along the shoulder of the 30S body towards proteins bS6 and uS11 and to the binding site of IF3 (Suppl. Fig. 2B). The predicted model for the bS1 protein in isolation³⁹ shows an array of six OB fold-like S1-domains. The observed density in our 30S IC can accommodate two additional S1 motifs (Suppl. Fig. 2B, “OB3” and “OB4”) in an elongated arrangement and along with potentially bound 5'-UTR of the mRNA²⁸. This is consistent with other reports of bS1 interacting with upstream portions of mRNA on the 30S surface⁴⁰ and/or binding of A/U-rich segments that precede the Shine-Dalgarno sequence⁴¹.

The *ompD* region for seed pairing and deactivating cleavage is accessible in the paused transcription elongation complex

The sRNA MicC recognises a segment in the *ompD* coding region (codons 23–26, nucleotide positions +67 to +78), referred to as the

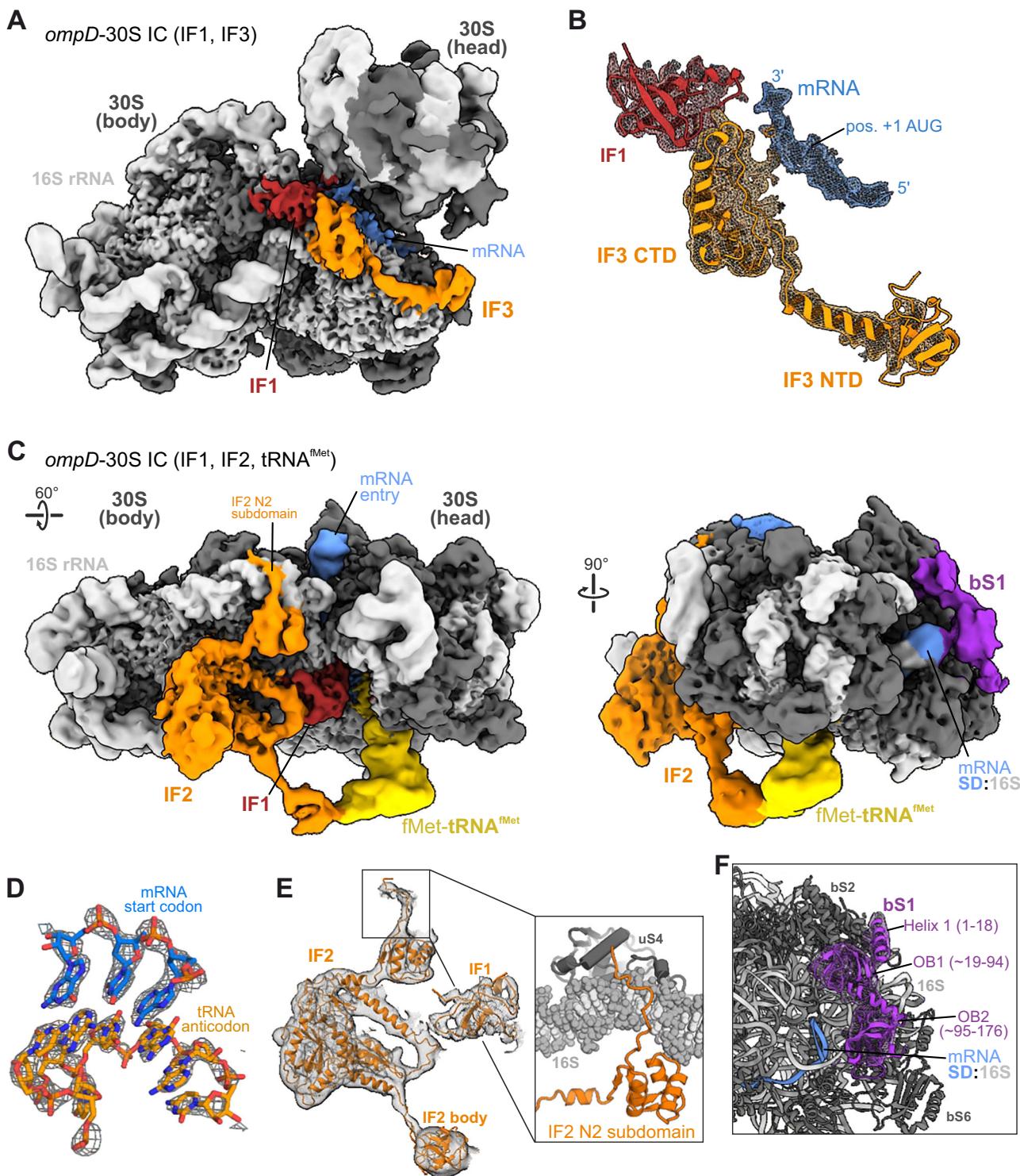


Fig. 1 | Structure of the *E. coli* translation initiation complex (30S IC) on the *ompD* transcript. **A** Cryo-EM map of the 30S IC on the *ompD* mRNA with IF1 and IF3. **B** View of the isolated cryo-EM density around the *ompD* mRNA and IF1/IF3 which are shown as ribbon models. **C** Views of the cryo-EM map of the 30S IC on the *ompD* mRNA with IF1, IF2 and fMet-tRNA^{fMet}. The rotation symbol on the left indicates the view in respect to (A). **D** Close view of the cryo-EM density of the codon-

anticodon pair between the *ompD* mRNA and the initiator tRNA. **E** View of the isolated cryo-EM density for IF1 and IF2. Inset: Molecular model for the N2 subdomain of IF2 with the contacting region of the 16S rRNA near ribosomal protein uS4. **F** Molecular model of the first two OB-domains of the ribosomal bS1 protein. The cryo-EM density is shown as a mesh around bS1.

seed pairing region, which leads to *ompD*-cleavage by the endoribonuclease RNase E^{30,31}. Since MicC and Hfq do not prevent ribosome engagement on *ompD*³⁰, it is likely that the 30S IC can be formed on *ompD* transcripts that are targeted for degradation. We therefore sought to determine at which stage of gene expression the sRNA can

still exert its regulatory effect. Intriguingly, the ability of RNase E to act on MicC-tagged *ompD* after formation of the 30S IC is dependent on the RNA degradosome assembly, which entails recruitment of DEAD-box RNA helicase RhlB, the glycolytic enzyme enolase and the exoribonuclease polynucleotide phosphorylase to RNase E³². The 30S IC can

exist at the tail of a polyribosome in the context of mature *ompD* transcripts, but also as part of the pioneering ribosome initiating on nascent *ompD* mRNA, upstream of the transcribing RNAP²⁸.

RNAP can transiently pause during transcription, particularly in promoter-proximal regions, which could help establish the TTC. Genome-wide studies of *in vivo* transcriptional pausing in *E. coli* have identified a gene-encoded signal element with the consensus motif G₋₁₀Y₋₁G₋₁ (where -1 corresponds to the 3'-end of the nascent RNA, with Y representing a pyrimidine base, Suppl. Fig. 3A) at which transcriptional pausing occurs with increased likelihood^{21,42,43}. Adjacent guanine bases at the upstream end (positions G₋₉ and G₋₁₁ within the motif) could stabilise the DNA-RNA hybrid of a paused TEC in pre- and post-translocated states²¹. Intriguingly, the *ompD* gene harbours this conserved consensus motif with G₋₁₀G₋₉C₋₁G₋₁ in its coding region, 28 nts downstream of the recognition sequence for the sRNA MicC, which might expose *ompD* to recognition and deactivating cleavage on the paused TEC (Suppl. Fig. 3B). In *S. Typhimurium* *ompD*, the G₋₁ nucleotide of the pause motif lies at position +106 in the coding region and is consistent with a putative paused transcript species of corresponding length that we observed when transcribing *ompD* *in vitro* using *E. coli* RNAP (Suppl. Fig. 3C).

To investigate how RNAP might impact *ompD* recognition by Hfq/MicC via the seed region at positions +67 to +78 in the *ompD* ORF, and whether a transiently paused TEC could be a potential target for riboregulation via deactivating cleavage by the RNA degradosome at position +83, we reconstituted TECs on different *ompD* segments of increasing length that position RNAP either upstream, at, or downstream of the regulatory region (Fig. 2A and Suppl. Fig. 3D–J). Reconstituted *ompD*-TECs encompass the 5'-UTR (starting from position -69 at the 5'-end) and expose the *ompD* mRNA up to positions +59/+83/+89/+107. The 9 terminal nucleotides at the mRNA 3'-ends were altered to base pair with the same template DNA strand of a short, artificial transcription bubble scaffold for TEC reconstitution (Suppl. Fig. 4A). This approach allowed us to obtain stable complexes suitable for degradation reactions with RNase E as well as cryo-EM studies, which validated *ompD*-TEC assembly on the artificial transcription bubble scaffold (see below).

Cleavage assays of the reconstituted TEC (or the free *ompD* fragment for comparison) with RNase E catalytic domain or full RNA degradosome were undertaken in the presence of MicC and Hfq. RNA fragments were quantified as fractions of respective starting material at time point 0 min. Rates for mRNA degradation and cleavage at position +83 were quantified as described in “Material and Methods”. The +72-cleavage product is an off-target species that is only observed *in vitro* and was disregarded for this analysis^{30–32}. We first positioned the *ompD*-TEC upstream of the MicC recognition site, exposing the *ompD* ORF up to position +59, and observed that mRNA and sRNA degradation rates are not affected by the presence of a reconstituted TEC (Suppl. Fig. 3D–F). Next, we placed the TEC so that the *ompD* ORF is exposed up to position +89, mimicking the approximate position of the above-described putative pause element immediately downstream of the MicC hybridisation site. Based on *in silico* docking with available molecular models of TEC (PDB: 2 ppb) and RNase E (PDB: 2c0b)^{44,45}, such a reconstituted *ompD*₈₉-TEC is positioned to expose the cleavage site at position +83 with six downstream nucleotides outside the RNAP exit channel and should thus be accessible to the active centre of RNase E. Indeed, we observe the *ompD* intermediate that is produced by MicC-guided cleavage at position +83, as expected (Fig. 2B). In the absence of Hfq-MicC, the +83-cleavage product is not observed and instead the *ompD* fragment is cleaved at off-target position +72. Concurrent to *ompD*-cleavage, MicC is also deactivated by endonucleolytic cleavage at position +9³¹. Thus, the *ompD*₈₉-TEC is accessible to MicC-induced cleavage at position +83 by RNase E as well as the full RNA degradosome complex (Fig. 2D). Moreover, cleavage at position +83 appears more efficient when the *ompD* segment is a part of the

elongation complex, indicated by an increased rate of accumulation of the corresponding intermediate (Fig. 2C and E).

A six nucleotides shorter *ompD*₈₃-TEC should still allow the MicC seed region to hybridize to the recognition sequence in *ompD*, but the +83-cleavage site would be masked by the RNAP exit channel and therefore sterically protected from RNase E cleavage. Indeed, RNase E activity on *ompD*₈₃-TEC at the MicC-induced +83-cleavage site was much less efficient than on free *ompD* in the absence of RNAP (Suppl. Fig. S3G–I), validating correct TEC assembly at the 3'-end of the mRNA and the stability of the complex throughout exposure to regulatory RNA, Hfq and RNase E. However, the deactivation of MicC is still stimulated by the presence of the TEC and the off-target cleavage at *ompD* position +72 is protected by MicC presence, indicating that MicC can hybridise with the seed region in the *ompD*₈₃-TEC. Positioning the TEC further downstream to expose *ompD* up to position +107 does not impact cleavage efficiency at position +83 compared to the same *ompD* fragment in the absence of RNAP (Suppl. Fig. S3J), implying that proximity of the RNAP to the regulatory region in *ompD* is a determinant for observed stimulation.

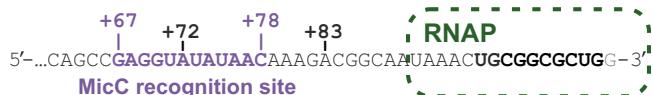
To explore potential direct interactions between RNAP and regulatory factors MicC and Hfq, we determined cryo-EM structures of the *ompD*₈₃-TEC in the presence of Hfq and MicC (Fig. 2F, G and Suppl. Fig. 4). The 3D reconstruction shows, at 3.3 Å global resolution, clear density for two α , β and β' subunits of the RNAP core enzyme, but none for the ω subunit, Hfq or MicC, which may have dissociated from the complex during gel filtration or vitrification. The substrate binding site of the RNAP harbours the duplex DNA construct, which can be continuously traced through the cryo-EM density. The transcription bubble displays the melted base DNA strands from which 10 nucleotides in the template DNA are hybridised to the mRNA and the active site harbours an unpaired template nucleobase with RNAP in the post-translocated state (Fig. 2F). We changed TEC reconstitution and vitrification parameters (see “Material and Methods”) and were able to validate the presence of the ω subunit in our RNAP preparation. The resulting cryo-EM reconstructions also display additional density in the vicinity of the mRNA exit channel of the RNAP, likely signal for MicC/Hfq hybridised to the *ompD* recognition site but not adopt in a stable position with respect to the TEC (Fig. 2G). The TEC in these particle classes does not display an altered conformation or additional density on its surface, indicating no direct interaction with the regulatory factors. Nevertheless, the structure shows that the *ompD*-TEC assembles as anticipated and demonstrates that the presence of MicC-Hfq does not disturb the complex.

Our data show that the nascent mRNA becomes accessible to recognition by MicC-Hfq immediately after exiting the RNAP. This is also consistent with the observation that MicC/Hfq protects off-target cleavage of *ompD*₈₃-TEC at position +72, which lies within the seed pairing region and becomes inaccessible to RNase E in the double-stranded mRNA:sRNA hybrid. Furthermore, we show that *ompD* becomes accessible to MicC-induced cleavage by RNase E once the cleavage site exits the TEC. We also observe that a proximally situated TEC increases the efficiency of sRNA-induced RNase E activity and that this effect vanishes when the RNAP is placed further downstream. The putative pause element in *ompD* could also allow the TEC to reside long enough for the pioneering ribosome to initiate translation and possibly establish early TTC. We explore combined effects of paused TEC and initiating 30S ribosome below.

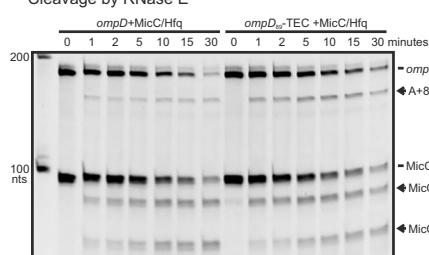
Cryo-EM structure of a coupled transcription elongation-translation initiation complex

Promoter-proximal pausing of transcription in the coding region of genes after synthesis of the ribosome binding site could allow the translational and transcriptional machineries to synchronise. We explored whether the paused *ompD*-TEC could engage in physical coupling with the pioneering 30S ribosomal subunit during

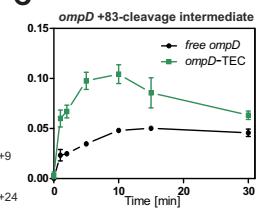
A reconstituted *ompD*₈₈-TEC



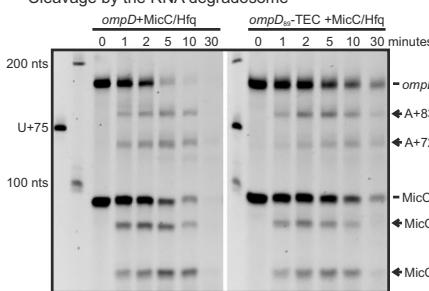
B Cleavage by RNase E



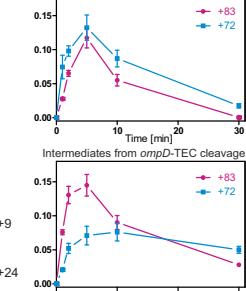
C



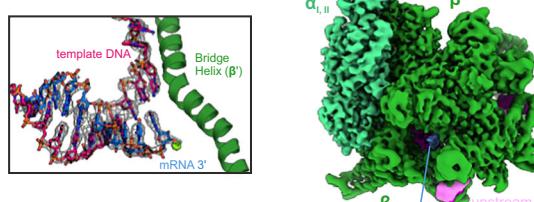
D Cleavage by the RNA degradosome



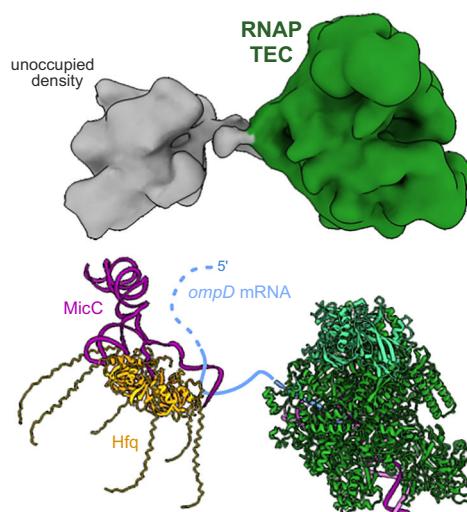
E Intermediates from free *ompD* cleavage



F Cryo-EM structure of reconstituted *ompD*-TEC



G



recruitment of translation initiation factors. As described above, we have successfully reconstituted TECs and 30S ICs separately and verified that the *ompD*-30S IC assembles at the start codon and demonstrated that the *ompD*-TEC assembles on the artificial transcription bubble at the 3'-end of the *ompD* mRNA fragment, mimicking paused TEC. The TEC has not been assembled on the natural pause-inducing sequence but instead on a scaffold with a modified sequence to

optimise stability for structural studies. We then reconstituted both assemblies together on the *ompD*₈₈ transcript fragment in the presence of NusA, NusG, MicC and Hfq, and performed cryo-EM analysis (Fig. 3A and Suppl Fig. 5). Through iterative rounds of 3D classification, we selected 30S particles that showed clear signal for the initiation factors and initiator tRNA (Suppl Fig. 5). The resulting map can accommodate the 30S IC, mRNA, fMet-tRNA^{fMet} and IF1-3 in the

Fig. 2 | The paused *ompD*-transcription elongation complex (TEC) can be directed by small regulatory RNA for cleavage by RNase E. **A** Schematic and 3'-sequence of the reconstituted *ompD₈₉*-TEC. The 3'-terminal nucleotides that hybridise with the template DNA strand within the TEC are shown in bold black letters. A green oval delineates the 3'-terminal portion of the mRNA that is masked by RNAP in the reconstituted TEC. The recognition sequence for MicC is marked in purple. RNase E cleavage sites observed during *in vitro* degradation reactions are annotated at positions +72, +83. **B–E** Time course reactions of the *ompD₈₉*-TEC (or free *ompD* mRNA fragment for comparison) incubated with MicC/Hfq and either RNase E (**B**, **C**) or the full RNA degradosome (**D**, **E**). Cleavage intermediates of *ompD*

and MicC are indicated with arrows. The relative abundances of the transient *ompD* cleavage intermediates were quantified (mean \pm SD from three independent reactions) and plotted over time in **C** (reactions with RNase E) and **E** (reactions with full RNA degradosome). **F** 3D-refined and locally filtered cryo-EM map of the *ompD*-TEC structure. Subunits of RNAP and bound DNA and mRNA are indicated. The inset shows an isolated view on the RNAP active site. The unwound template DNA strand is shown in red, the nascent mRNA in blue. The bridge helix feature of the RNAP β' -subunit at the active site is depicted in ribbon representation. **G** Structure of *ompD*-TEC with unoccupied density around the mRNA exit site where MicC and Hfq are expected to recognise *ompD*.

positions that we observe in our 30S IC data, and we again observe the elongated density near the mRNA exit site that can be accounted for by the ribosomal bS1 protein (Fig. 3A).

Additionally, our map displays a large area of unoccupied density on top of the 30S head, which is not present in the *ompD*-30S IC structures. To improve the map in this area, we subtracted the signal for the 30S IC from the cryo-EM particles and performed iterative rounds of 3D refinement and classification, which enabled the isolation of a subset of particles that refined to fit the TEC portion of the coupled complex to \sim 6 Å resolution (Suppl Figs. 5 and 6). The 30S-coupled TEC resembles the above-described structure of the *ompD*-TEC that was solved in the absence of the 30S IC but displays additional density for the ω subunit of RNAP as well as for the N-terminal domain (NTD) of transcription factor NusG. The NusG NTD is bound at the previously described binding site between β and β' subunits near the unwound non-template DNA strand⁴⁶ (Fig. 3A). We also observe well-defined density for the NusG-CTD on the 30S ribosomal subunit near uS10 (NusE)²⁹, and this density is absent in our structures of the 30S IC that was prepared in the absence of NusG. We observe NusA-NTD contacting the RNAP-portion of the coupled TEC-30S IC complex in a manner similar to previously reported TEC structures containing NusA⁴⁷. A structure of the uncoupled TEC originating from the 30S IC-containing dataset contains additional density for the NusA S1 motif that ensues from the NusA-NTD (Fig. 3A right). We do not observe density for NusA on the surface of the 30S IC, suggesting that this interaction is dispensable for physical coupling between TEC and 30S IC.

Notably, density for the TEC appeared only in 3D-reconstructions generated from the subset of 30S particles that also contained signal for the IFs and fMet-tRNA^{Met}. Additionally, *ompD₈₉*-TEC-IC samples prepared in the absence of NusG did not display physical coupling between TEC and 30S IC (Suppl. Fig. 7), suggesting that the physical coupling between RNAP and 30S IC is dependent both on NusG and mRNA-engagement by the small ribosomal subunit.

When the cryo-EM particles of the final subset are aligned on either 30S IC or RNAP, the respective other component remains only visible as noisy density (Fig. 3B and Suppl. Fig. 5). Multi-body Refinement in RELION allows the analysis of the flexibility between TEC and 30S⁴⁸. The first three eigenvectors account for 50.8 % of the variance in the data (19.8, 16.1 and 14.9 % for components 1, 2 and 3, respectively), which is visualised Fig. 3C and Suppl. Movie 2. This high mobility (up to 42° rotation and 87 Å change in distance between mass centres along individual eigenvectors) between TEC and 30S IC in the early expressome likely results from the long mRNA linker in which the start codon at the P-site of the 30S IC and the 3'-end at the TEC lie almost 90 nucleotides apart. In this regard, the orientation between 30S subunit and TEC is different from that in 70S-containing expressomes with increasing lengths of mRNA linker TTC-A and TTC-B^{9–11}. However, it resembles the orientation reported for a TTC-LC assembly containing an extended mRNA linker exceeding the length required for TTC-B formation¹⁶.

Overall, our structural data demonstrates that a putative early expressome can be formed *in vitro* from a TEC and a 30S IC, and the physical coupling of the two subcomplexes relies on NusG which forms a link between mRNA-engaged 30S and RNAP.

The coupled transcription elongation-translation initiation complex can be accessed at a target site by sRNA for effector recruitment

The linker mRNA in the early expressome model loops out between the TEC and 30S IC, which may provide an opportunity for sRNA pairing to a target site. To test whether the *ompD* mRNA, when embedded in the TEC-IC, is still accessible for recognition by MicC and deactivating cleavage by the RNA degradosome, we performed degradation reactions with the RNA degradosome in the presence of MicC and Hfq (Fig. 4 and Suppl. Fig. 8). We also subjected to degradation reactions the free *ompD* segment as well as the individual sub-assemblies *ompD₈₉*-TEC and *ompD₈₉*-30S IC. The degradation rates of *ompD₈₉* decreased with rising complexity of the assembly, down to a fourfold slower degradation of NusG-containing *ompD*-TEC-IC compared to free *ompD* (Fig. 4B, C). Double-exponential equation fitting of the formation of the +83-cleavage intermediate and degradation of *ompD* is shown in Fig. 4B, C. The formation rate allows comparison of cleavage specificity for the +83 site which is expressed as the formation rate of the +83 intermediate over the degradation rate of the *ompD₈₉* starting material (from Fig. 4C). Cleavage specificities for the +83 site in the different *ompD₈₉* assemblies are shown in Fig. 4D. The RNA degradosome shows the lowest preference for +83-cleavage on free *ompD* fragment and the highest on the *ompD₈₉*-TEC-IC substrate. The fraction of the unspecific +72-cleavage intermediates on the other hand is highest in the case of free *ompD* fragment, reduced for the *ompD₈₉*-TEC and -IC and was unquantifiable in the early expressome (Suppl. Fig. 8). The observation that MicC-guided processing by RNase E occurs with higher specificity when RNAP and/or 30S IC are engaged with the *ompD* mRNA suggests that the gene expression machineries might facilitate the recognition of *ompD* by the sRNA and/or present the MicC-*ompD* hybrid in a way that increases susceptibility to attacks by the RNA degradosome at the +83-cleavage site.

The cryo-EM maps of the *ompD₈₉*-TEC-IC complex reveal conformational sub-states that may facilitate regulatory access of the sRNA to the seed pairing site in the transcript (Fig. 4E, F). One of the modes of conformational variation is in a relative tilting of the RNAP with respect the 30S IC, and this is correlated with additional density on the exposed segment of the mRNA (right panel, Fig. 4E). The density is not sufficiently resolved to build a detailed model, but the shape of the envelope can accommodate Hfq and may represent a transient species in which the sRNA/Hfq interacts with the exposed seed region. The schematic in Fig. 4F illustrates a potential model for this transient access.

Discussion

In this study, we have explored the accessibility of a small RNA pairing site in a target transcript at early stages of transcription elongation and translation initiation. We reconstituted TEC mimicking a state of promoter-proximal transcription pausing of the target mRNA and assembled the 30S translation initiation complex (30S IC) on those mRNAs. Cryo-EM single-particle analysis shows that the transcription factor NusG can physically link the 30S IC and the paused TEC, despite the lengthy mRNA linker (>70 nts) and the *in vitro* concentrations used that are far below the proposed dissociation constant of the

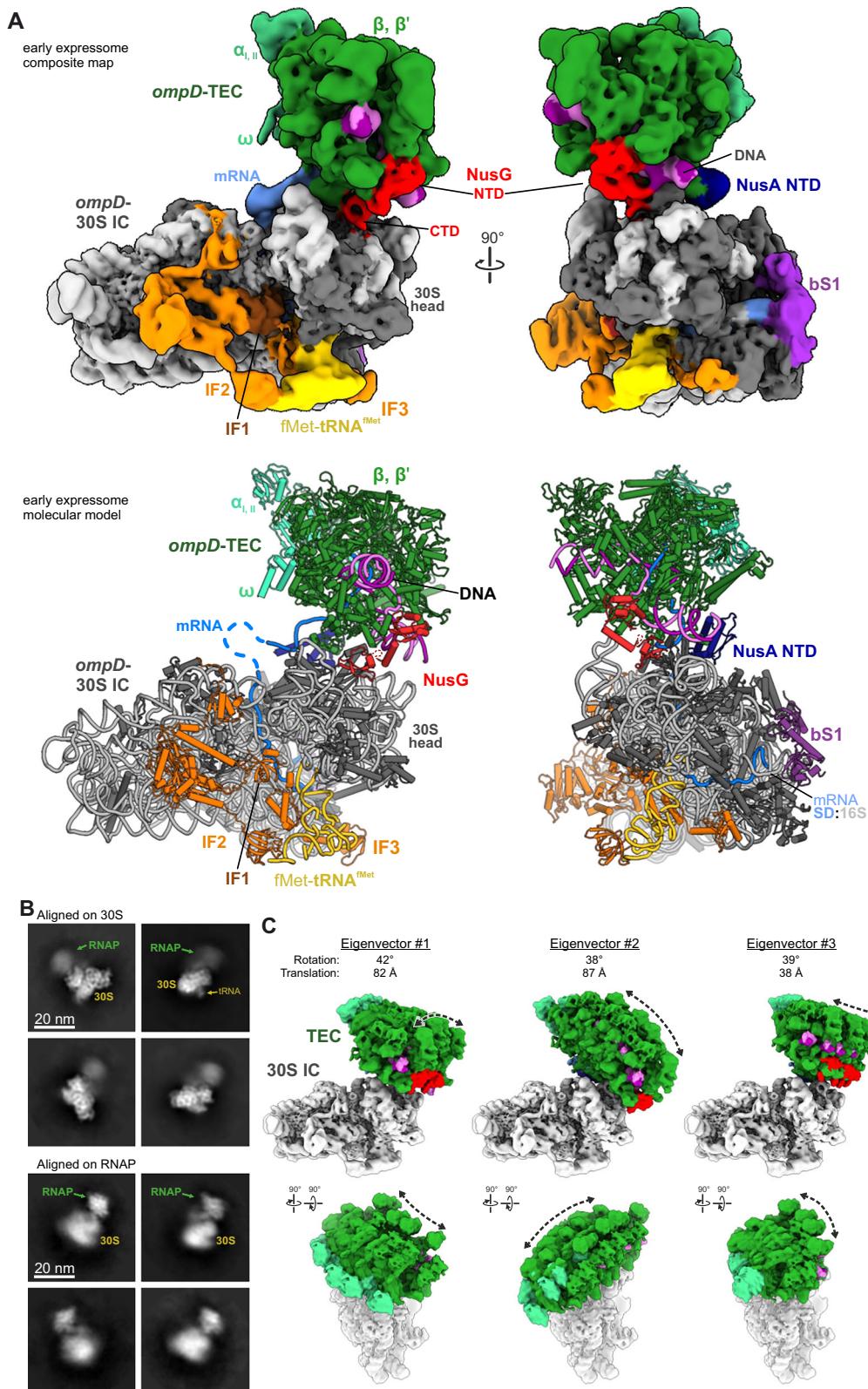
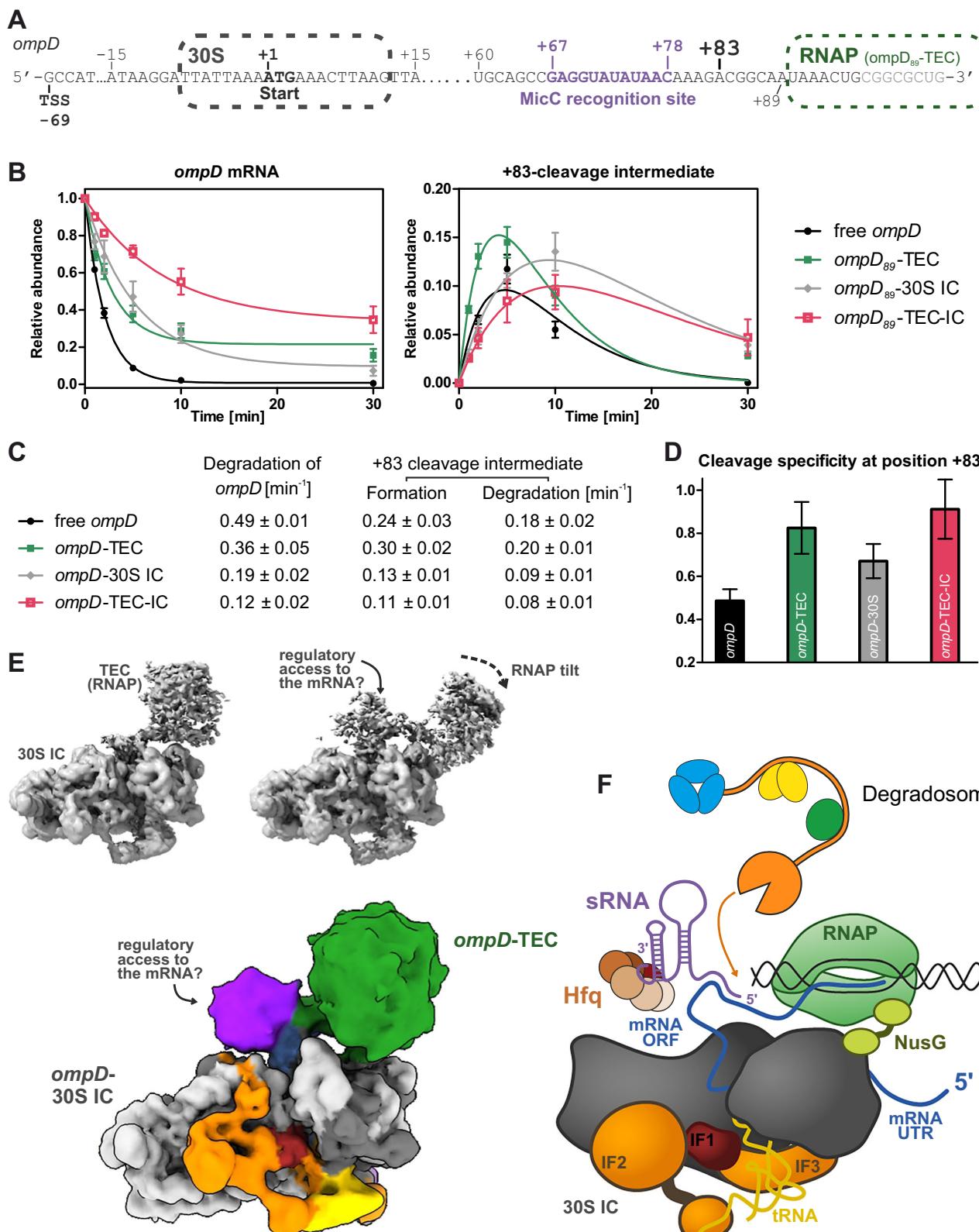


Fig. 3 | The structure of the transcription elongation-translation initiation complex with regulatory element in the bridging mRNA. A Composite cryo-EM map (top) and molecular model (bottom) of the “early expressome”, comprising NusG-coupled transcription elongation (*ompD*-TEC) and translation initiation (*ompD*-30S IC) complexes. Components that correspond to segments of the map are indicated. The lower panel shows the molecular model interpretation of the cryo-EM structure based on reported structures for the 30S translation initiation

complex (PDB: 5lmv) and the NusG-coupled expressome (PDB: 6xgf)^{10,33}. **B** 2D class averages of cryo-EM particles that were aligned on either 30S IC or TEC components of the early expressome assembly. **C** Multi-body motion between TEC and 30S IC is visualized by 4 superimposed maps from each of the three main eigenvectors. Maximum changes in angles of rotation and distance between mass centers are indicated.



NusE-NusG interaction²⁹. The physically coupled assembly, which we refer to as an “early expressome” may represent a transient precursor to mature expressomes in vivo. The complex is highly dynamic but has similar orientation between the 30S subunit and TEC as seen in the TTC with the long-range coupling (TTC-LC)¹⁶.

During the preparation of this manuscript, an independent study reported the in vitro reconstitution of transcription elongation-

translation pre-initiation complexes²⁸. Webster et al. present structures of mRNA delivery complexes between a paused TEC and 30S ribosomal subunit in the pre-active state and propose two pathways that initiate the coupling between transcription and translation through transcription-assisted recruitment of mRNA to the ribosome²⁸. One pathway involves ribosomal protein bS1 of the 30S subunit interacting with the TEC and binding the nascent mRNA to

Fig. 4 | The bridging element in the TEC-IC complex is accessible to the RNA degradosome. **A** Schematic of the *ompD* transcript and MicC recognition site. **B, C** Cleavage time courses of *ompD* mRNA fragment and +83-cleavage intermediate species (see Fig. 2D and Suppl. Fig. 8 for representative raw data). Free *ompD* mRNA fragment (black), *ompD₈₉*-TEC (green), *ompD₈₉*-30S IC (grey) and *ompD₈₉*-TEC-IC (red) were incubated with MicC/Hfq and then subjected to cleavage by the RNA degradosome. Reactions were stopped at indicated time points and analysed by denaturing PAGE. Data are mean \pm SD from three independent reactions. Degradation trajectories of *ompD* were fitted to one phase exponential decay. Formation and decay of intermediates from *ompD*-cleavage at position +83 were

fitted to two-exponential equation [$I = \exp(-kt) - \exp(-jt)$] with degradation rate k and formation rate j . **D** Cleavage specificities for nucleolytic processing of site +83 in *ompD* in the different assembly states, expressed as the formation rate of the +83 intermediate over the degradation rate of the respective *ompD* fragment starting material. Data are mean \pm SD from three independent reactions. **E** Conformational substates of the TEC-30S IC assembly. **F** Substate of the early expressome proposed to facilitate the accommodation of Hfq/MicC on the bound *ompD* mRNA. The colours of the segmented map in (E) correspond to the components shown in the schematic.

form an intermediate or standby complex that directs the nascent transcript to the Shine-Dalgarno interaction site of the 30S subunit. The RNAP is transiently located near the cluster of bS1 OB-domains and mRNA exit site of the 30S and is subsequently repositioned to be near the mRNA entry site in the translational 30S IC. The second pathway involves RNAP directly binding near the mRNA entry site of the inactive 30S, which then transitions to the active form in the assembled 30S IC by folding the 16S rRNA helix 44 away from the mRNA exit channel into its active position on the 30S body²⁸. Our data include IF1, IF2 and IF3 and present the 30S IC in “accommodated” active states, which can be formed via both proposed pathways.

The structural data presented here also corroborate how the ribosomal bS1 might be poised to facilitate interactions with 5'-UTR elements. Webster et al. report that the bS1 protein not only facilitates mRNA delivery to selectively accelerate duplex formation with the anti-Shine-Dalgarno sequence but helps RNAP to stimulate translation initiation. bS1 assists the remodelling of mRNA secondary structures in vitro to aid the 30S with mRNA loading and 30S IC formation^{49,50}. When the *ompD* mRNA is recruited to the 30S IC, RNase E relies on the assembly of the RNA degradosome for access and efficient MicC-guided degradation of the *ompD* transcript³². Ribosomal bS1 has been shown to form direct interaction with several sRNAs from *E. coli*⁵¹, thereby possibly exerting the 30S IC's influence on Hfq-MicC-guided RNase E attack while located on the 30S ribosome. Further, direct interactions with the TEC⁵² might also facilitate sRNA-mediated regulation at the early stage of nascent transcript recruitment to the translation machinery. The paused RNAP may increase the recognition efficiency of the target element within *ompD* by Hfq-MicC, consistent with our observation that *ompD*-cleavage by RNase E is more site-specific at the paused TEC. This is supported by a recent single-molecule study showing that co-transcriptional target recognition by sRNA and Hfq close to the RNAP exit channel is more efficient than recognition post-transcription⁵³. The biogenesis of 3'-UTR-derived sRNAs that autoregulate their genes⁵⁴, along with kinetic modelling studies, provide further support for mechanisms of co-transcriptional regulation by sRNAs⁵⁵. Access to decay machinery during transcription is also supported by observations of co-transcriptional stabilisation of riboswitches against ribonuclease action⁵⁶. Moreover, a recent study presents a model of co-transcriptional endonucleolytic cleavage of RNA in archaea that leads to transcription termination⁵⁷, suggesting the possibility of a functionally analogous system in that domain of life. In an early expressome of the *ompD* transcript in which the RNAP is paused at position +106, the deactivating cleavage site at position +83 for RNase E that is preceded by the recognition site for the MicC sRNA, remains accessible to regulatory RNA and deactivating cleavage by the RNA degradosome in in vitro reactions. Nevertheless, the relationship between the kinetics of sRNA-mediated regulation and the rate of TTC remains unclear, leaving open the question of how much of a temporal window of opportunity is available for sRNAs to access their targets. In addition, the influence of spacer length in long-range coupling on target accessibility to sRNAs remains to be elucidated.

The cellular location of the degradosome must be considered to explain how access might be gained to the putative TEC-30S IC complex tagged with a small RNA (Fig. 5). The envisaged

recognition event, in which sRNA recruits degradosome to the 30S IC or TEC-IC complex, could occur in species such as *Caulobacter crescentus*, in which the RNA degradative machinery can be associated with the nucleoid or ribonucleoprotein condensates in the cell interior and could have access to nascent mRNA targets⁵⁸. However, in species such as *E. coli* and *Salmonella* where the degradosome is compartmentalized to the cytoplasmic membrane, and potentially distant from the nucleoid⁵⁹, the sRNA-tagged early expressome might be the target for other effectors, such as Rho RNA translocase that terminates transcription through allosteric manipulation of the elongation complex^{15,55,60,61}. Rho can act to trigger premature transcription termination in the absence of TTC^{58,62} and its activity on nascent mRNA can be modulated by sRNA⁶³. Transcription-terminated transcripts could then encounter the degradosome on the membrane upon their diffusion following the release from RNAP.

Another scenario in which membrane-bound RNase E can act early on nascent transcripts is if the transcriptional-translational machinery is brought close to the membrane, as occurs in the transertion mechanism which was described for *Vibrio parahaemolyticus* and suggested to occur also in *E. coli*⁶⁴⁻⁶⁹. Transertion, a coupling of transcription, translation and membrane insertion at the membrane component of the type III secretion system, is a process suggested to be common for bacterial membrane proteins. Evidence indicates that membrane-associated RNase E can act on transcripts encoding membrane proteins in *E. coli* to result in co-transcriptional degradation⁵⁸.

A fourth scenario in which we envisage RNase E to act on a mRNA-30S IC complex is if the leading edge of a polyribosome comes into proximity of the membrane. Earlier studies have shown that RNase E and the degradosome can interact with polyribosomes in vitro and potentially in vivo, and it was proposed that this forms a passive complex that does not cleave the RNA until activated by a signal, such as a cognate sRNA⁷⁰. In species without RNase E, similar processes are likely to occur, such as in the model firmicute species *Bacillus subtilis*, where the membrane-bound degradosome is based on the distinct ribonuclease RNase Y⁷¹. This mode of degradation would enable the co-translational decay of a transcript. It also represents a more economical means of exiting translation, because it avoids generating transcript fragments that lack stop codons and entail the hidden metabolic costs of using the tmRNA system for rescue and recovery. The above mechanisms could account for processes of co-translational decay^{72,73} analogous to the recruitment of the 5'-to-3' exoribonuclease Xrn1, which follows the terminal translating ribosome identified in yeast and other eukaryotic species^{74,75}.

Methods

Protein production

Wild-type 30S subunits were prepared from *E. coli* MRE600 strain using zonal centrifugation^{76,77}. Initiator tRNA (fMet-tRNAsfMet) and *E. coli* initiation factors IF1, IF2 and IF3 were purified according to published protocols⁷⁸. *E. coli* Hfq was purified as described⁷⁹. The RNA degradosome was prepared as described⁸⁰. *E. coli* core RNAP was expressed from pVS10 plasmid coding for all five subunits⁸¹ and purified as described⁸².

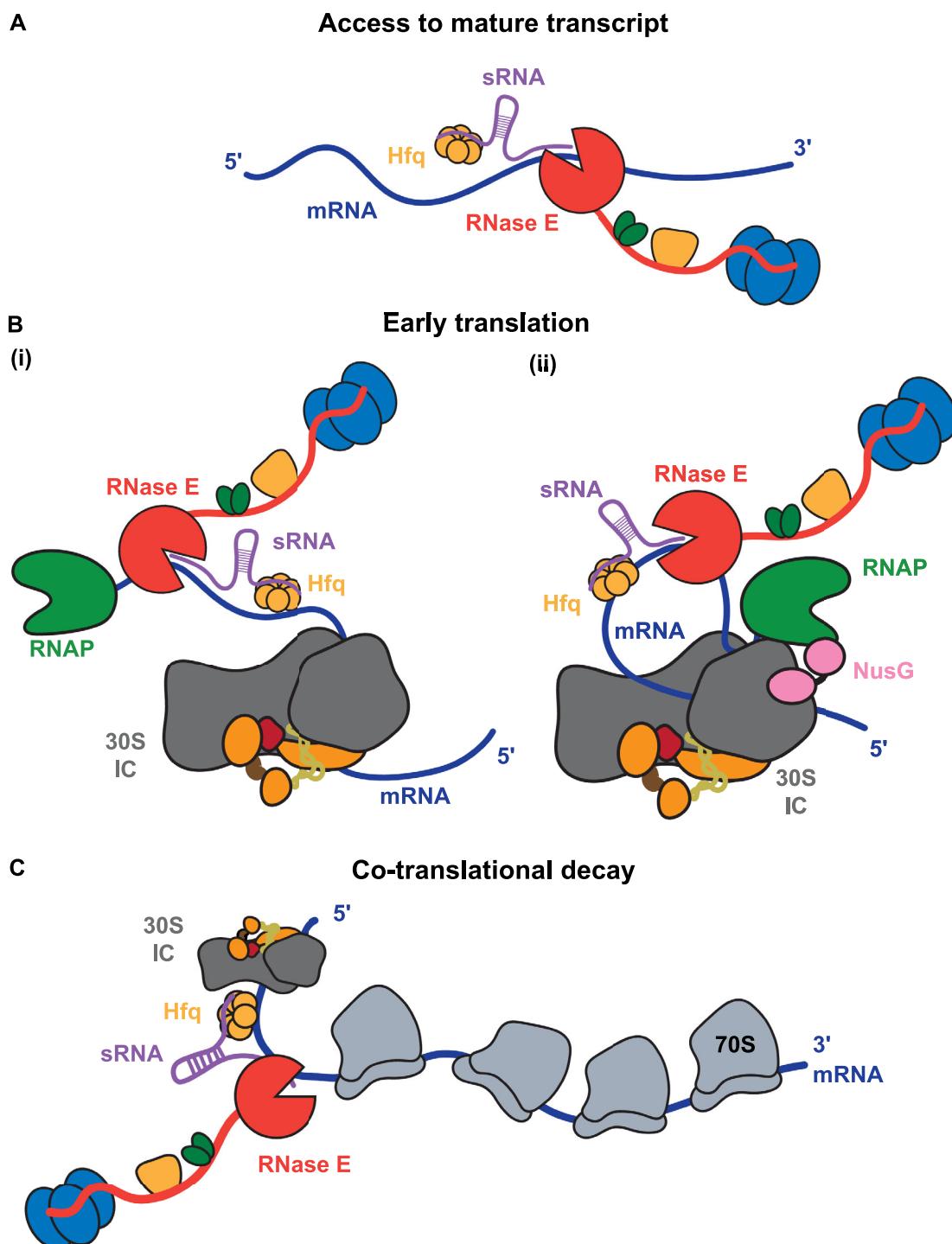


Fig. 5 | Schematic models for sRNA-mediated degradation of mRNAs during transcription and translation. Four modes are presented where the RNA degradosome could access transcripts tagged with sRNAs. **A** Access to the mature mRNA after transcription termination. **B** Access at early stages of translation explored in this study. Mode (i) depicts a scenario in which the degradosome can access the coding region following the formation of the pioneering 30S IC. Mode (ii) posits

access to the translation initiation complex in the vicinity of the paused transcription elongation complex. **C** Access for co-translation degradation. Here, the degradosome could be interacting with polysomes once the mRNA has begun to be translated. This mode is supported by the *in vivo* observation of RNA cluster formation by degradosomes in the presence of polysomes (Hamouche et al. 2021).

RNA production

RNAs were prepared by *in vitro* transcription (IVT) as described³¹. In brief, IVT templates were amplified by PCR from plasmid pVP042-3 that carries the *Salmonella Typhimurium* *ompD* gene³⁰. The forward primer introduces a T7 promoter sequence upstream of the transcription initiation site and, where indicated, the reverse primer

encodes a non-canonical 9 nts sequence (CGGCGCUGG) at the 3' end of the corresponding transcript. IVT reactions containing 3 µg of template DNA, 5 mM of each of ATP, UTP, GTP and CTP, 10 mM DTT and 0.5 U/µL RNaseOUT™ (Invitrogen) were incubated with recombinant T7 RNAP in 40 mM Tris, pH 8.0, 25 mM MgCl₂, 2 mM spermidine at 37 °C for 5 h. IVT products were DNase I-treated, resolved by 4%

denaturing PAGE and excised RNA bands were electroeluted using Elutrap™ Electroelution System Kit (Whatman), followed by clean-up using PureLink™ RNA Micro Kit (Invitrogen).

Assembly of transcription elongation complexes

TECs were assembled following a previously described strategy⁶⁰. To anneal nucleic acid scaffolds, 8 μM mRNA fragment was mixed with template DNA (tDNA) in a 1:1.25 molar ratio in 10 mM TRIS, pH 7.6, 40 mM KCl, 5 mM MgCl₂, heated to 95 °C for 2 min and slowly equilibrated to 37 °C. Two volumes TEC reconstitution buffer (20 mM TRIS, pH 7.6, 120 mM potassium acetate, 5 mM magnesium acetate, 1 mM TCEP) were pre-heated to 37 °C and added to annealed mRNA/tDNA. *E. coli* RNAP was added in 1:1 ratio to RNA, followed by 10 min incubation. Non-template DNA (ntDNA) was added in 1:1 ratio to tDNA and followed by incubation for 10 min at 37 °C. Assembled TECs were prepared fresh and used immediately for degradation assays or assembly of higher order complexes.

For cryo-EM specimen preparation (Sample 1), 5 μM TEC was mixed with equimolar amount of pre-incubated (32 °C for 30 min) MicC and Hfq injected onto a 3.2/300 Superose 6 size-exclusion chromatography column equilibrated with 20 mM TRIS-Cl, pH 7.6, 120 mM potassium acetate, 5 mM magnesium acetate, 2 mM DTT, 10 μM ZnCl₂. Fractions containing all components as determined by urea-PAGE and SDS-PAGE were pooled, concentrated to 40 μL and used directly for vitrification on graphene oxide coated grids as described below.

For Sample 2, 7.5 μM TEC was mixed with 1.4-fold molar excess of pre-incubated (32 °C for 10 min) MicC and Hfq and dialysed against TEC reconstitution buffer to remove residual glycerol. The reaction was recovered and concentrated to approx. 10 μM using a 30 MWCO spin concentrator. 8 mM CHAPSO detergent were added and 3 μL sample was applied to both faces of an R2/2 UltrAuFoil grids and vitrified after 2 s blotting on both sides.

Preparation of *ompD*–30S IC and early expressomes

30S IC was prepared following a previously described strategy⁸³. In brief, 30S subunits were incubated in buffer TAKM₂₀ (50 mM Tris-HCl [pH 7.5], 70 mM NH₄Cl, 30 mM KCl, 20 mM MgCl₂) for 30 min at 37 °C for reactivation. Reactivated 30S subunits were incubated with a 2.5-fold excess of mRNA, twofold excess of initiation factors 1–3 and a 2.5-fold excess of initiator fMet-tRNA^{fMet} in the presence of 250 μM GTP or GTPyS (Jena Bioscience) in TAKM₇ buffer⁸³. For transcription/translation-coupled complexes, the 30S IC was prepared as described above using pre-assembled *ompD*-TEC instead of free mRNA and omitting the chromatography step. Coupled TEC-IC for cryo-EM analysis was prepared with final concentrations 0.3 μM 30S, 1.2 μM TEC, 0.6 μM of IFs and tRNA, 10 μM NusA, 20 μM NusG and 5 μM MicC/Hfq and directly used for grid preparation on Quantifoil 1.2/1.3 grids with a 2 nm amorphous carbon layer as described below.

In vitro transcription

The first 306 bases of the *ompD* gene (−69 to +237) were inserted into the pMMB67HE vector between *tac* promoter and T1 terminator⁸⁴. For IVT assays, 15 μL reactions containing 5 nM *ompD*-pMMB67HE, 0.5 U/μL RNaseOUT™ and 0.5 U *E. coli* RNAP Holoenzyme (NEB, M0551S) were pre-incubated in reaction buffer (40 mM TRIS, pH 7.5, 150 mM KCl, 10 mM MgCl₂, 1 mM DTT, 0.01% Triton X-100™) at 37 °C for 5 min. Reactions were started by adding ribonucleotide mix to 5 mM, incubated at 37 °C for 1 h and quenched by addition of equal volume stop buffer (200 mM TRIS, pH 7.5, 25 mM EDTA, 300 mM NaCl, 2% SDS, 0.5 mg/mL Proteinase K). After proteolysis at 50 °C for 1 h, samples were mixed with 0.5 volumes 2× RNA Loading Dye (Thermo Scientific), heated to 95 °C for 2 min and separated by urea PAGE on 8 % polyacrylamide gels.

In vitro RNA degradation assays

Degradation assays were carried out following a protocol described in ref. 31. Degradation reactions were performed in degradation buffer (25 mM TRIS, pH 7.6, 50 mM NaCl, 50 mM KCl, 10 mM MgCl₂, 2 mM DTT) with 0.5 U/μL RNaseOUT™. 5'-monophosphorylated MicC sRNA was brought to 50 °C for 2 min and slowly cooled to 37 °C before use. Substrates (*ompD*-TEC, *ompD*-TEC-30S IC or *ompD*:ntDNA for control) were prepared as described above and pre-incubated with equimolar amounts of MicC and Hfq at 37 °C for 15 min at final concentrations of 0.2 μM. After withdrawal of reaction aliquots for *t* = 0 min, fresh RNase E^{1–598} or full-length RNA degradosome was diluted in 2× degradation buffer. Reactions were started by adding 1/10 volume of enzyme at final concentrations of 0.05 μM for RNase E^{1–598} or 0.01 μM degradosome. Reaction aliquots were withdrawn at indicated time points and quenched with equal volume stop buffer (see above under IVT) at 50 °C for 30–45 min. Samples were mixed with RNA loading dye II (ThermoFisher), heated to 95 °C for 2 min and resolved on 8% or 10% denaturing (7.5 M urea) polyacrylamide gels. Gels were stained in SYBR Gold (Invitrogen) and visualized under UV light. Bands were quantified with ImageJ and plotted with GraphPad PRISM as fraction of *t* = 0 min. Degradation rates of full-length RNAs was determined by fitting the data to the one phase decay fit equation in GraphPad PRISM, $I = (I_0 - P) * e^{-kt} + P$ with relative intensity, *I*; plateau, *P*; degradation rate, *k*; reaction time, *t*. Formation and decay rates of intermediates from *ompD*-cleavage at position +83 were determined by fit to two-exponential equation $[I = \exp(-kt) - \exp(-jt)]$ with degradation rate *k* and formation rate *j*. Cleavage site specificities were calculated by dividing intermediate production rates by the respective full-length RNA degradation rate.

Graphene oxide coating and grid preparation

Graphene oxide (GO) coating was prepared by an adaptation of the drop-casting method⁸⁵. Quantifoil Cu 300 1.2/1.3 grids were glow-discharged on the darker carbon side (PELCO easiGLOW: 15 mA, 0.28 mBar, 2 min). GO solution (Sigma-Aldrich 763705, 2 mg/mL dispersion) was diluted tenfold in water and centrifuged at 300 × g for 30 s to pellet insoluble GO flakes. The supernatant was further diluted tenfold to a working concentration of 0.02 mg/mL. 1 μL working solution was applied to the glow-discharged side of the Quantifoil grids and dried at room temperature for 10 min. GO-coated grids were kept at room temperature for at least 16 h and then directly used for sample vitrification. Quantifoil 1.2/1.3 grids with a 2 nm amorphous carbon support layer were briefly glow-discharged (PELCO easiGLOW: 10 s, 25 mA, 0.39 mBar) directly before sample application. For vitrification, 4 μL of sample was applied to grids and incubated for 30 s. Using a FEI Vitrobot, excess sample was blotted and grids were plunge-frozen in liquid ethane. Grids were subsequently stored in liquid nitrogen until screening and collection.

Cryo-EM data collection and processing

For structures of *ompD*-30S IC, 9104 multi-frame movies were collected on a 300 kV FEI Titan Krios equipped with a Gatan K3 detector. For the TEC structure, a dataset of 1619 multi-frame movies was collected on a 300 kV Titan Krios equipped with a Falcon 3 detector in Counting mode. For structures of different TEC-30S IC reconstitutions, small datasets were first collected on a 200 kV Talos Arctica equipped with a Falcon 3 detector in Counting mode and processed in Relion 3.1⁸⁶ as depicted in image processing workflows (Suppl. Fig. 7B, G, L). For the coupled TEC-30S IC structure, 12,011 multi-frame movies were then collected on a 300 kV FEI Titan Krios equipped with a Gatan K3 detector. Details for data collection are summarised in Supplementary Table 1. Multi-frame movies were motion-corrected in Relion 3.1⁸⁶ and the generated micrographs were imported in cryoSPARC v4⁸⁷ and further processed as follows and illustrated in Suppl. Figs. 1–3. Particles were picked using “Blob Picker”, extracted and downsampled fourfold.

For all described structures, an initial round of Heterogeneous Refinement was run on all extracted particles using 3D references generated by ab initio reconstruction from particles subsets selected after 2D classification to isolate 30S or TEC-containing particles from poor-quality “junk” particles.

For 30S structures, a subsequent 3D classification without refining particle alignments separated particles containing tRNA from particles without tRNA. tRNA-containing particles were further classified with focused masks around IFs and tRNA to isolate IF2-containing particles, which were re-extracted without downsampling, subjected to Homogenous Refinement, and further classified with a focused mask around IF2 to isolate particles that refine to highest resolution in this region. Separately, particles lacking tRNA were classified iteratively with focus masks around IFs and mRNA and classes with weak or no density for these subunits were discarded as illustrated in Suppl. Fig. 1B. The final particle set was re-extracted without downsampling.

For the *ompD*-TEC structure, particles isolated by heterogeneous refinement, as described above, were extracted without downsampling and subjected to global CTF refinement and one round of a 3D classification, from which the class with the particles that refined to the highest resolution was kept. Final particle sets were subjected to Homogenous Refinement with default settings and resulting reconstructions were subjected to Local Resolution Estimation and Local Filtering. Particle orientation plots and local resolution estimates are shown in Suppl. Fig. 6. 3D Variability Analysis (3DVA)⁸⁸ was performed on final particle sets of the 30S structures and visualised in “Simple” mode. For the TEC, 3DVA was performed on all TEC particles that refined to better than 4 Å resolution to better reflect the variability across the sample.

For the coupled TEC-30S IC structure, all non-junk particles that aligned on the 30S reference were subjected to a second round of Heterogeneous Refinement, and the particle class containing density for IFs and tRNA in the 30S IC was selected. A soft-edged mask around the 30S IC was used to subtract the signal of the 30S IC followed by 3D classification without re-alignment. The classes with the strongest density in the region of the TEC were selected and subjected to Heterogeneous Refinement providing low-pass filtered (20 Å) reference maps of the TEC to isolate particles that can be aligned on RNAP (these 34,106 particles were processed further; see below). After a Local Refinement of the best class using a Mask around the TEC, a final round of 3D classification was run, particles of the best class were re-extracted with twofold downsampling and locally refined (Suppl. Fig. 5C “TEC portion”) or subjected to Homogenous refinement providing the 30S IC reference (Suppl. Fig. 5C “Consensus Refinement”). Using the particle alignments of the local refinement around the TEC, the TEC signal was subtracted from the re-extracted particles and followed by a Homogenous Refinement (Suppl. Fig. 5C “30S IC portion”), which was combined with the TEC portion into a composite map inside ChimeraX. The final set of expressome particles was converted to a STAR file using UCSF PyEM (<https://doi.org/10.5281/zenodo.3576630>), imported into Relion and a subjected to a Multi-body Refinement⁴⁸ providing a local mask around 30S IC and TEC, respectively. Multi-body motion along the three main eigenvectors is visualised in Fig. 3C and Suppl. Movie 2. Angles of rotation and change of distance between mass centres were calculated inside PyMOL using RotationAxis (<http://pymolwiki.org/index.php/RotationAxis>).

Separately, the subset of 34,106 expressome particles (see above) was classified based on the position of the TEC in respect to the 30S IC by using the 30S-subtracted particles that contain signal for the TEC while providing alignments of the original particles that were refined on the 30S IC. The two classes that displayed additional density around the path that the mRNA would take between TEC and 30S IC were classified with a focus mask drawn around this additional density. The

two classes with the strongest density in this area were refined individually by Homogenous Refinement (Suppl. Fig. 5E).

Data availability

The data supporting the findings of this study are available from the corresponding authors upon request. Structure data are available from the EMDB.

EMD-55527 Early Expressome Composite Map of TEC and 30S IC

EMD-55528 Early Expressome Consensus Refinement

EMD-55529 Early Expressome RNAP/TEC body

EMD-55530 *E. coli* 30S IC containing mRNA, IF1 and IF3

EMD-55531 *E. coli* 30S IC containing mRNA, initiator tRNA, IF1 and IF2

References

1. Blaha, G. M. & Wade, J. T. Transcription-translation coupling in bacteria. *Annu. Rev. Genet.* **56**, 187–205 (2022).
2. Webster, M. W. & Weixlbaumer, A. Macromolecular assemblies supporting transcription-translation coupling. *Transcription* **12**, 103–125 (2021).
3. Woodgate, J. & Zenkin, N. Transcription-translation coupling: recent advances and future perspectives. *Mol. Microbiol.* **120**, 539–546 (2023).
4. Miller, O. L. Jr., Hamkalo, B. A. & Thomas, C. A. Jr. Visualization of bacterial genes in action. *Science* **169**, 392–395 (1970).
5. Johnson, G. E., Lalanne, J. B., Peters, M. L. & Li, G. W. Functionally uncoupled transcription-translation in *Bacillus subtilis*. *Nature* **585**, 124–128 (2020).
6. Zhu, M., Mu, H., Han, F., Wang, Q. & Dai, X. Quantitative analysis of asynchronous transcription-translation and transcription processivity in *Bacillus subtilis* under various growth conditions. *iScience* **24**, 103333 (2021).
7. Iyer, S., Le, D., Park, B. R. & Kim, M. Distinct mechanisms coordinate transcription and translation under carbon and nitrogen starvation in *Escherichia coli*. *Nat. Microbiol.* **3**, 741–748 (2018).
8. Derno, G. et al. Structure of RNA polymerase bound to ribosomal 30S subunit. *Elife* **6**, e28560 (2017).
9. Kohler, R., Mooney, R. A., Mills, D. J., Landick, R. & Cramer, P. Architecture of a transcribing-translating expressome. *Science* **356**, 194–197 (2017).
10. Wang, C. et al. Structural basis of transcription-translation coupling. *Science* **369**, 1359–1365 (2020).
11. Webster, M. W. et al. Structural basis of transcription-translation coupling and collision in bacteria. *Science* **369**, 1355–1359 (2020).
12. Fan, H. et al. Transcription-translation coupling: direct interactions of RNA polymerase with ribosomes and ribosomal subunits. *Nucleic Acids Res.* **45**, 11043–11055 (2017).
13. Qureshi, N. S. & Duss, O. Tracking transcription-translation coupling in real time. *Nature* **637**, 487–495 (2025).
14. O'Reilly, F. J. et al. In-cell architecture of an actively transcribing-translating expressome. *Science* **369**, 554–557 (2020).
15. Molodtsov, V., Wang, C., Firsov, E., Kaelber, J. T. & Ebright, R. H. Structural basis of Rho-dependent transcription termination. *Nature* **614**, 367–374 (2023).
16. Wang, C. et al. Structural basis of long-range transcription-translation coupling. Preprint at <https://www.biorxiv.org/content/10.1101/2024.07.20.604413> (2025).
17. Artsimovitch, I. & Landick, R. Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. *Proc. Natl. Acad. Sci. USA* **97**, 7090–7095 (2000).
18. Chan, C. L. & 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.* **233**, 25–42 (1993).
19. Dey, S. et al. Structural insights into RNA-mediated transcription regulation in bacteria. *Mol. Cell* **82**, 3885–3900.e10 (2022).

20. Mooney, R. A. et al. Regulator trafficking on bacterial transcription units in vivo. *Mol. Cell* **33**, 97–108 (2009).
21. Imashimizu, M. et al. Visualizing translocation dynamics and nascent transcript errors in paused RNA polymerases in vivo. *Genome Biol.* **16**, 98 (2015).
22. Hatoum, A. & Roberts, J. Prevalence of RNA polymerase stalling at *Escherichia coli* promoters after open complex formation. *Mol. Microbiol.* **68**, 17–28 (2008).
23. Washburn, R. S. et al. *Escherichia coli* NusG links the lead ribosome with the transcription elongation complex. *iScience* **23**, 101352 (2020).
24. Shaham, G. & Tuller, T. Genome scale analysis of *Escherichia coli* with a comprehensive prokaryotic sequence-based biophysical model of translation initiation and elongation. *DNA Res.* **25**, 195–205 (2018).
25. Irastortza-Olaziregi, M. & Amster-Choder, O. Coupled transcription-translation in prokaryotes: an old couple with new surprises. *Front. Microbiol.* **11**, 624830 (2020).
26. Zhu, M., Dai, X. & Wang, Y. P. Real time determination of bacterial in vivo ribosome translation elongation speed based on LacZalpha complementation system. *Nucleic Acids Res.* **44**, e155 (2016).
27. Wee, L. M. et al. A trailing ribosome speeds up RNA polymerase at the expense of transcript fidelity via force and allostery. *Cell* **186**, 1244–1262.e34 (2023).
28. Webster, M. W. et al. Molecular basis of mRNA delivery to the bacterial ribosome. *Science* **386**, eado8476 (2024).
29. Burmann, B. M. et al. A NusE:NusG complex links transcription and translation. *Science* **328**, 501–504 (2010).
30. Pfeiffer, V., Papenfort, K., Lucchini, S., Hinton, J. C. & Vogel, J. Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nat. Struct. Mol. Biol.* **16**, 840–846 (2009).
31. Bandyra, K. J. et al. The seed region of a small RNA drives the controlled destruction of the target mRNA by the endoribonuclease RNase E. *Mol. Cell* **47**, 943–953 (2012).
32. Bandyra, K. J. et al. Cooperation of regulatory RNA and the RNA degradosome in transcript surveillance. *Nucleic Acids Res.* **52**, 9161–9173 (2024).
33. Hussain, T., Llacer, J. L., Wimberly, B. T., Kieft, J. S. & Ramakrishnan, V. Large-scale movements of IF3 and tRNA during bacterial translation initiation. *Cell* **167**, 133–144.e13 (2016).
34. Uday, A. B., Mishra, R. K. & Hussain, T. Initiation factor 3 bound to the 30S ribosomal subunit in an initial step of translation. *Proteins* **93**, 279–286 (2023).
35. Basu, R. S., Sherman, M. B. & Gagnon, M. G. Compact IF2 allows initiator tRNA accommodation into the P site and gates the ribosome to elongation. *Nat. Commun.* **13**, 3388 (2022).
36. D'Urso, G., Chat, S., Gillet, R. & Giudice, E. Structural insights into the binding of bS1 to the ribosome. *Nucleic Acids Res.* **51**, 3410–3419 (2023).
37. Loveland, A. B. & Korostelev, A. A. Structural dynamics of protein S1 on the 70S ribosome visualized by ensemble cryo-EM. *Methods* **137**, 55–66 (2018).
38. Saito, K. et al. Ribosome collisions induce mRNA cleavage and ribosome rescue in bacteria. *Nature* **603**, 503–508 (2022).
39. Varadi, M. et al. AlphaFold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **50**, D439–D444 (2022).
40. Sengupta, J., Agrawal, R. K. & Frank, J. Visualization of protein S1 within the 30S ribosomal subunit and its interaction with messenger RNA. *Proc. Natl. Acad. Sci. USA* **98**, 11991–11996 (2001).
41. Komarova, A. V., Tchufistova, L. S., Supina, E. V. & Boni, I. V. Protein S1 counteracts the inhibitory effect of the extended Shine-Dalgarno sequence on translation. *RNA* **8**, 1137–1147 (2002).
42. Vvedenskaya, I. O. et al. Interactions between RNA polymerase and the “core recognition element” counteract pausing. *Science* **344**, 1285–1289 (2014).
43. Larson, M. H. et al. A pause sequence enriched at translation start sites drives transcription dynamics in vivo. *Science* **344**, 1042–1047 (2014).
44. Vassylyev, D. G. et al. Structural basis for substrate loading in bacterial RNA polymerase. *Nature* **448**, 163–168 (2007).
45. Callaghan, A. J. et al. Structure of *Escherichia coli* RNase E catalytic domain and implications for RNA turnover. *Nature* **437**, 1187–1191 (2005).
46. Kang, J. Y. et al. Structural basis for transcript elongation control by NusG family universal regulators. *Cell* **173**, 1650–1662.e14 (2018).
47. Guo, X. et al. Structural basis for NusA stabilized transcriptional pausing. *Mol. Cell* **69**, 816–827.e4 (2018).
48. Nakane, T. & Scheres, S. H. W. Multi-body refinement of Cryo-EM images in RELION. *Methods Mol. Biol.* **2215**, 145–160 (2021).
49. Duval, M. et al. *Escherichia coli* ribosomal protein S1 unfolds structured mRNAs onto the ribosome for active translation initiation. *PLoS Biol.* **11**, e1001731 (2013).
50. Kolb, A., Hermoso, J. M., Thomas, J. O. & Szer, W. Nucleic acid helix-unwinding properties of ribosomal protein S1 and the role of S1 in mRNA binding to ribosomes. *Proc. Natl. Acad. Sci. USA* **74**, 2379–2383 (1977).
51. Windbichler, N., von Pelchrzim, F., Mayer, O., Csaszar, E. & Schroeder, R. Isolation of small RNA-binding proteins from *E. coli*: evidence for frequent interaction of RNAs with RNA polymerase. *RNA Biol.* **5**, 30–40 (2008).
52. Sukhodolets, M. V. & Garges, S. Interaction of *Escherichia coli* RNA polymerase with the ribosomal protein S1 and the Sm-like ATPase Hfq. *Biochemistry* **42**, 8022–8034 (2003).
53. Rodgers, M. L., O'Brien, B. & Woodson, S. A. Small RNAs and Hfq capture unfolded RNA target sites during transcription. *Mol. Cell* **83**, 1489–1501.e5 (2023).
54. Hoyos, M., Huber, M., Forstner, K. U. & Papenfort, K. et al. Gene autoregulation by 3' UTR-derived bacterial small RNAs. *Elife* **9**, e58836 (2020).
55. Reyer, M. A. et al. Kinetic modeling reveals additional regulation at co-transcriptional level by post-transcriptional sRNA regulators. *Cell Rep.* **36**, 109764 (2021).
56. Lou, Y. & Woodson, S. A. Co-transcriptional folding of the glmS ribozyme enables a rapid response to metabolite. *Nucleic Acids Res.* **52**, 872–884 (2024).
57. You, L. et al. Structural basis of archaeal FttA-dependent transcription termination. *Nature* **635**, 229–236 (2024).
58. Kim, S., Wang, Y. H., Hassan, A. & Kim, S. Re-defining how mRNA degradation is coordinated with transcription and translation in bacteria. Preprint at <https://www.biorxiv.org/content/10.1101/2024.04.18.588412v1.full.pdf> (2024).
59. Carpousis, A. J., Campo, N., Hadjeras, L. & Hamouche, L. Compartmentalization of RNA degradosomes in bacteria controls accessibility to substrates and ensures concerted degradation of mRNA to nucleotides. *Annu Rev. Microbiol.* **76**, 533–552 (2022).
60. Said, N. et al. Steps toward translocation-independent RNA polymerase inactivation by terminator ATPase rho. *Science* **371**, eabd1673 (2021).
61. Hao, Z. et al. Pre-termination transcription complex: structure and function. *Mol. Cell* **81**, 281–292.e8 (2021).
62. Wang, B. & Artsimovitch, I. NusG, an ancient yet rapidly evolving transcription factor. *Front Microbiol* **11**, 619618 (2020).
63. Sedlyarova, N. et al. sRNA-mediated control of transcription termination in *E. coli*. *Cell* **167**, 111–121.e13 (2016).
64. Woldringh, C. L. The role of co-transcriptional translation and protein translocation (transertion) in bacterial chromosome segregation. *Mol. Microbiol.* **45**, 17–29 (2002).

65. Bakshi, S., Siryaporn, A., Goulian, M. & Weisshaar, J. C. Super-resolution imaging of ribosomes and RNA polymerase in live *Escherichia coli* cells. *Mol. Microbiol.* **85**, 21–38 (2012).
66. Fishov, I. & Norris, V. Membrane heterogeneity created by translocation is a global regulator in bacteria. *Curr. Opin. Microbiol.* **15**, 724–730 (2012).
67. Bakshi, S., Choi, H., Mondal, J. & Weisshaar, J. C. Time-dependent effects of transcription- and translation-halting drugs on the spatial distributions of the *Escherichia coli* chromosome and ribosomes. *Mol. Microbiol.* **94**, 871–887 (2014).
68. Roggiani, M. & Goulian, M. Chromosome-membrane interactions in bacteria. *Annu Rev. Genet.* **49**, 115–129 (2015).
69. Kaval, K. G. et al. Membrane-localized expression, production and assembly of *Vibrio parahaemolyticus* T3SS2 provides evidence for translocation. *Nat. Commun.* **14**, 1178 (2023).
70. Tsai, Y. C. et al. Recognition of the 70S ribosome and polysome by the RNA degradosome in *Escherichia coli*. *Nucleic Acids Res.* **40**, 10417–10431 (2012).
71. Laalami, S., Cavaiuolo, M., Oberto, J. & Putzer, H. Membrane localization of RNase Y is important for global gene expression in *Bacillus subtilis*. *Int. J. Mol. Sci.* **25**, 8537 (2024).
72. Huch, S. et al. Atlas of mRNA translation and decay for bacteria. *Nat. Microbiol.* **8**, 1123–1136 (2023).
73. Herzel, L., Stanley, J. A., Yao, C. C. & Li, G. W. Ubiquitous mRNA decay fragments in *E. coli* redefine the functional transcriptome. *Nucleic Acids Res.* **50**, 5029–5046 (2022).
74. Pelechano, V., Wei, W. & Steinmetz, L. M. Widespread Co-translational RNA decay reveals ribosome dynamics. *Cell* **161**, 1400–1412 (2015).
75. Tesina, P. et al. Structure of the 80S ribosome-Xrn1 nuclease complex. *Nat. Struct. Mol. Biol.* **26**, 275–280 (2019).
76. Rodnina, M. V., Fricke, R. & Wintermeyer, W. Transient conformational states of aminoacyl-tRNA during ribosome binding catalyzed by elongation factor Tu. *Biochemistry* **33**, 12267–12275 (1994).
77. Rodnina, M. V. & Wintermeyer, W. GTP consumption of elongation factor Tu during translation of heteropolymeric mRNAs. *Proc. Natl. Acad. Sci. USA* **92**, 1945–1949 (1995).
78. Milon, P. et al. Transient kinetics, fluorescence, and FRET in studies of initiation of translation in bacteria. *Methods Enzymol.* **430**, 1–30 (2007).
79. Dendooven, T. et al. A cooperative PNPase-Hfq-RNA carrier complex facilitates bacterial riboregulation. *Mol. Cell* **81**, 2901–2913 e5 (2021).
80. Dendooven, T. et al. Multi-scale ensemble properties of the *Escherichia coli* RNA degradosome. *Mol. Microbiol.* **117**, 102–120 (2022).
81. Svetlov, V. & Artsimovitch, I. Purification of bacterial RNA polymerase: tools and protocols. *Methods Mol. Biol.* **1276**, 13–29 (2015).
82. Castro-Roa, D. & Zenkin, N. Methodology for the analysis of transcription and translation in transcription-coupled-to-translation systems in vitro. *Methods* **86**, 51–59 (2015).
83. Goyal, A., Belardinelli, R. & Rodnina, M. V. Non-canonical binding site for bacterial initiation factor 3 on the large ribosomal subunit. *Cell Rep.* **20**, 3113–3122 (2017).
84. Furste, J. P. et al. Molecular cloning of the plasmid RP4 primase region in a multi-host-range tacP expression vector. *Gene* **48**, 119–131 (1986).
85. Pantelic, R. S., Meyer, J. C., Kaiser, U., Baumeister, W. & Plitzko, J. M. Graphene oxide: a substrate for optimizing preparations of frozen-hydrated samples. *J. Struct. Biol.* **170**, 152–156 (2010).
86. Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
87. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryo-SPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
88. Punjani, A. & Fleet, D. J. 3D variability analysis: resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. *J. Struct. Biol.* **213**, 107702 (2021).

Acknowledgements

We thank Andrzej Szewczak-Harris and the staff of the Cryo-EM Facility, Dimitri Y. Chirgadze, Steven Hardwick and Lee Cooper, for assistance with data collection. We thank the staff at eBIC for access to facilities and help with data collection. We thank Tom Dendooven, Kathi Frohlich and Joerg Vogel for helpful comments and suggestions. We thank Kai Katuya Gaviria for providing Hfq and helpful advice. K.J.B., B.F.L. and N.Z. are supported by Wellcome Trust Investigator Awards (222451/Z/21/Z, 200873/Z/16/Z, 217189/Z/19/Z). J.J.R. is supported by a Herchel Smith Studentship. G.P. was supported by a Benn W Levy–Vice Chancellor Award SBS DTP PhD studentship.

Author contributions

All authors contributed to data collection. CryoEM data and analysis were done by J.R. and G.P. All authors contributed to writing the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-67330-2>.

Correspondence and requests for materials should be addressed to Katarzyna J. Bandyra or Ben F. Luisi.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025