Title:

Structural basis of transcription-translation coupling

Authors:

Chengyuan Wang[1,*], Vadim Molodtsov[1,*], Emre Firlar[2], Jason Kaelber[2], Gregor Blaha[3], Min Su[4,**], and Richard H. Ebright[1,**]

Affiliations:

[1]Waksman Institute and Department of Chemistry and Chemical Biology, Rutgers University, Piscataway NJ 08854, USA

[2]Cryo-EM Core Facility, Rutgers University, Piscataway NJ 08854, USA

[3]Departent of Biochemistry, University of California, Riverside CA 92521, USA

[4]Life Sciences Institute, University of Michigan, Ann Arbor MI,48109, USA

[*]Contributed equally

[**]Correspondence to minsu@umich.edu; ebright@waksman.rutgers.edu

One Sentence Summary:

Cryo-EM defines states that mediate NusG-dependent and NusG-independent transcription-translation coupling in bacteria

Abstract:

In bacteria, transcription and translation are coupled processes, in which movement of RNA polymerase (RNAP) synthesizing mRNA is coordinated with movement of the first ribosome translating mRNA. Coupling is modulated by the transcription factors NusG--which is thought to bridge RNAP and ribosome--and NusA. Here, we report cryo-EM structures of *Escherichia coli* transcription-translation complexes (TTCs) containing different-length mRNA spacers between RNAP and the ribosome active-center P-site. Structures of TTCs containing short spacers show a state incompatible with NusG bridging and NusA binding (TTC-A; previously termed "expressome"). Structures of TTCs containing longer spacers reveal two new states, one compatible with NusG bridging and both compatible with NusA binding (TTC-B and TTC-C). We propose that TTC-B mediates NusG-dependent transcription-translation coupling and TTC-C mediates NusG-independent transcription-translation coupling.

Main Text:

Bacterial transcription and bacterial translation occur in the same cellular compartment, occur at the same time, and are coordinated processes, in which the rate of transcription by the RNA polymerase (RNAP) molecule synthesizing an mRNA is coordinated with the rate of translation by the first ribosome ("lead ribosome") translating the mRNA (*1-8*; see, however, *9*). Data indicate that the coordination is mediated by transcription elongation factors of the NusG/RfaH family, which contain an N-terminal domain (N) that interacts with RNAP β' and β subunits and a flexibly tethered C-terminal domain (C) that interacts with ribosomal protein S10, and which are thought to bridge, and thereby connect, the RNAP molecule and the lead ribosome (*2*, *5-8*). Further data indicate that the coordination is modulated by the transcription elongation factor NusA (*10*).

Cramer and co-workers recently reported a 7.6 Å resolution cryo-EM structure of an *Escherichia coli* transcription-translation complex (TTC) termed the "expressome," obtained by halting a transcription elongation complex (TEC) and allowing a translating ribosome to collide with the halted TEC (*11*). However, the mRNA molecule in the structure was not fully resolved, precluding determination of the

number of mRNA nucleotides between the TEC and the ribosome active center in the structure (*11*), and the functional relevance of the structure has been challenged, due to its genesis as a collision complex, and due to its incompatibility with simultaneous interaction of NusG-N with RNAP and NusG-C with the ribosome (*6-8*). Demo *et al*. recently reported a ~7 Å resolution cryo-EM structure of a complex of *E. coli* RNAP and a ribosome 30S subunit (*12*). However, the structure did not contain mRNA, did not position RNAP close to the 30S mRNA-entrance portal, and was incompatible with simultaneous interaction of NusG-N with RNAP and NusG-C with the ribosome (*12*).

Here, we report cryo-EM structures of *E. coli* TTCs containing defined-length mRNA spacers between the TEC and the ribosome active-center product site (P site), both in the presence of NusG and in the absence of NusG (Figs. 1, S1-S5; Tables S1-S2). We prepared synthetic nucleic-acid scaffolds that contained (i) DNA and mRNA determinants that direct formation of a TEC upon interaction with RNAP, (ii) an mRNA AUG codon that enables formation of a translation complex having the AUG codon positioned in the ribosome active-center P site upon interaction with a ribosome and $tRNA^{fMet}$, and (iii) an mRNA spacer having a length, *n*, of 4, 5, 6, 7, 8, or 9 codons (12, 15, 18, 21, 24, or 27 nt) between (i) and (ii) (Fig. 1A). We then reacted the nucleic-acid scaffolds with RNAP, with ribosome and $tRNA^{fMet}$, and optionally with NusG, and we determined structures by single-particle-reconstruction cryo-EM (see Supporting Information, Materials and Methods). With nucleic-acid scaffolds having short spacers (n = 4, 5, 6, or 7), we obtained structures matching the "expressome" of *11* (TTC-A; Figs. 1B-left, S1-S3; Tables S1-S2). However, with nucleic-acid scaffolds having longer mRNA spacers (n = 7, 8, or 9), we obtained structures of two new molecular assemblies--one with features suggesting it is the molecular assembly that mediates NusG-dependent transcription-translation coupling in cells, and one with features suggesting it is the molecular assembly that mediates NusG-independent transcription-translation coupling in cells (TTC-B and TTC-C; Figs. 1B-center, 1B-right, S4-S5; Table S2).

TTC-A was obtained with nucleic-acid scaffolds having mRNA spacers of 4, 5, 6, or 7 codons--but not with longer mRNA spacers (Figs. 1B-left, 2, S1-S4; Tables S1-S2). In the presence of the detergent CHAPSO, 3.7-4.1 Å resolution EM density maps were obtained (~7 Å and ~3.5 Å local

resolution for TEC and ribosome, respectively), enabling unambiguous rigid-body docking of atomic structures of TEC, ribosome 30S subunit, ribosome 50S subunit with tRNA in active-center P site and exit site (E site), and NusG-N, followed by manual fitting of residues in the RNAP-ribosome interface and in DNA and mRNA (Figs. 1B-left, 2, S1-S2; Table S1). In the absence of CHAPSO, 11-19 Å resolution EM density maps were obtained (~14-20 Å and ~10-18 Å local resolution for TEC and ribosome, respectively), enabling unambiguous rigid-body docking of atomic structures of TEC, 30S, 50S with tRNAs, and NusG-N (Figs. S3-S4; Table S2). Two TTC-A subclasses were obtained; one corresponds to the previously reported "expressome" of *11* (TTC-A1, Figs. 1B-left, 2, S1-S3), and the other differs by a ~60° rotation of RNAP relative to the ribosome (TTC-A2; Fig. S3; movie S1).

Remarkably, in TTC-A, the spatial relationship of RNAP relative to the ribosome is identical in structures obtained with nucleic-acid scaffolds having mRNA spacer lengths of 4, 5, 6, and 7 codons (Figs. S1F, S3F, S4F). High-resolution data for TTC-A1 reveal that differences in mRNA spacer length are accommodated through differences in extents of compaction of mRNA in the RNAP RNA-exit channel and in RNAP-ribosome interface (Fig. 2B). As mRNA spacer length increases from 4 codons to 5 codons to 6 codons, the number of mRNA nucleotides in the RNAP RNA-exit channel and RNAP-ribosome interface increases from 7 nt (5 nt in exit channel; 2 nt in interface) to 10 nt (7 nt in exit channel; 3 nt in interface) to 13 nt (9 nt in exit channel; 4 nt in interface) (Fig. 2B, subpanels 1-3). When the mRNA spacer length increases to 7 codons, 16 nt of mRNA are accommodated in the RNAP RNA-exit channel and RNAP-ribosome interface, and the 16 nt of mRNA exhibit segmental disorder, indicating they adopt an ensemble of different conformations (Fig. 2B, subpanel 4). We point out that the volume of the RNAP RNA-exit channel and RNAP-ribosome interface cannot accommodate more than ~12 nt of mRNA without changing the conformation of the RNAP RNA-exit channel or disrupting the RNAP-ribosome interface, and we suggest that this accounts for our finding that TTC-A is not obtained with nucleic-acid scaffolds having mRNA spacer lengths >7 codons.

In TTC-A, the RNAP-ribosome interface is extensive (1,762 $Å^2$ buried surface area in TTC-A1) and involves contacts of RNAP β' zinc-binding domain (ZBD), RNAP β flap, and RNAP α^I with ribosomal proteins S4, S3, and S10, respectively (Figs. 2C-D).

In EM density maps of TTC-A, density is absent for RNAP ω subunit, indicating--particularly in high-resolution maps for TTC-A1--that RNAP ω subunit is either absent, or at a low occupancy level, or segmentally disordered (Figs. 2E, S3I). Molecular modelling indicates that, if RNAP ω were present and fully folded, the C-terminal α-helix of ω would clash with the ribosome (Figs. 2E, S3I).

In EM density maps of TTC-A obtained in the presence of NusG, EM density is present for NusG-N (residues 1-118) at its expected binding location on RNAP (*13*; Figs. 1B-left, S1H, S3F), but is absent--even in high-resolution maps of TTC-A1--for the NusG-N/NusG-C linker and NusG-C, consistent with unrestricted motion of the linker and NusG-C relative to NusG-N (*13*; Figs. 1B-left, S3F). Density maps for TTC-A obtained in the absence of NusG are identical to those obtained in the presence of NusG, except that density for NusG is missing (Fig. S2). Model building indicates that the shortest sterically allowed distance between NusG-N bound to RNAP and NusG-C modelled as bound to its molecular target on the ribosome, ribosomal protein S10 (*2*, *5-8*), is 160 Å in TTC-A1 and 190 Å in TTC-A2--which are 1.9 times and 2.2 times the length of the NusG-N/NusG-C linker--indicating that TTC-A1 and TTC-A2 are incompatible with NusG bridging of RNAP and S10 (Fig. S6).

Molecular modelling indicates that TTC-A1 and TTC-A2 also are incompatible with other known functional properties of transcription elongation, pausing, and termination in *E. coli*. TTC-A1 is sterically incompatible with binding of NusA (*14-15*; Fig. S7A-B), formation of a λ N antitermination complex (*15*; Fig. S7B), formation of an 21 Q antitermination complex (*16-17*; Fig. 7C), binding of σ region 2 (σR2; *16-19*; Fig. S7D), and formation of pause and termination hairpins having stem length ≥4 bp (*14*,*20-21*; Fig. S7E). TTC-A2 is sterically compatible with binding of NusA in the binding mode of *15* and with formation of pause and termination hairpins (Fig. S7B,E), but is incompatible with binding of NusA in the binding mode of *14,* and with formation of a λ N antitermination complex, formation of a 21 Q antitermination complex, and biding of σR2 (Figs 7A,C-D), The finding that TTC-A1--the "expressome"

of *11*--and TTC-A2 lack RNAP ω subunit, are incompatible with NusG bridging, and are incompatible with known functional properties of transcription and transcriptional regulation in *E. coli* indicates that TTC-A1 and TTC-A2 are unlikely to be functionally relevant to transcription-translation coupling under most conditions in *E. coli*. We propose that TTC-A1 and TTC-A2 are either: (i) specialized complexes that mediate transcription-translation coupling under specialized circumstances (e.g., transcription-translation coupling by ω-lacking RNAP), or (ii) anomalous complexes formed when the length of mRNA spacer between RNAP and ribosome is anomalously low (e.g. "collision-omes" or "crash-omes").

TTC-B was obtained with nucleic-acid scaffolds having mRNA spacer lengths of 7, 8, or 9 codons--but not with shorter mRNA spacers (Figs. 1B-center, 3, S4-S5; Table S2). TTC-B differs from TTC-A by rotation of RNAP relative to the ribosome by ~90° on one axis and ~180° on an orthogonal axis (Fig. 1B; movie S2). EM density maps at 7.0-9.9 Å resolution were obtained, (~10 Å and ~7 local resolution for TEC and ribosome, respectively), enabling unambiguous rigid-body docking (Figs. 3, S4-S5). Six TTC-B subclasses were obtained: TTC-B1 through TTC-B6, differing slightly in position of RNAP relative to ribosome and in rotational orientation of RNAP relative to ribosome (Figs. 3B, S4-S5; movie S3).

For TTC-B, differences in mRNA spacer length, from 7 codons to 8 codons to 9 codons, are accommodated through differences in the distance between RNAP and ribosome (smaller with shorter spacers; and larger with longer spacers; Fig. 3C, bottom three subpanels *vs.* top three subpanels) and differences in the rotational orientation of RNAP relative to ribosome (smaller rotation with shorter mRNA spacer and larger rotation with longer RNA spacer; Figs. 3C, S3F, S4F).

In TTC-B, the main direct interaction between RNAP and ribosome is mediated by RNAP β sequence insertion 2 (βSI2; also known as βi9 and "hook-and-ladder"; *22-24*), a long (60 Å) α-helical antiparallel coiled-coil that is flexibly tethered to the rest of RNAP through a "molecular swivel" that enables ~360° rotation on one axis and ~90° rotation on an orthogonal axis, analogous to a fire-truck hook-and-ladder (*24*) (Figs. 3A,C-D). The different positions and rotational orientations of RNAP relative to the ribosome in TTC-B structures having different mRNA spacer lengths are accommodated through

the flexibility of the molecular-swivel attachment of βSI2 to the rest of RNAP, with differences in the orientation of βSI2 relative to the rest of RNAP compensating for differences in positions and orientations of RNAP relative to the ribosome (Fig. 3C). In TTC-B1 through TTC-B3, βSI2 interacts--through the RNAP-distal parts of its coiled-coil helices, and through the loop connecting its coiled-coil helices--with ribosomal proteins S2, S3, and S5 (Figs. 3C, subpanels 1-3, D). In particular, residues 1003-1006 of the βSI2 loop, which have high net negative charge (-3), are positioned to make favorable electrostatic interactions with residues 105-111 of S2, which have high net positive charge (+4). In TTC-B4 through TTC-B6, EM density for βSI2 is not observed (Fig. 3C, subpanels 4-6), but, at the resolution of the maps for TTC-B, strong conclusions cannot be drawn from unobserved density, and molecular modeling indicates that, in TTC-B4 through TTC-B6, βSI2 would be positioned in the RNAP-ribosome interface adjacent to S2, S3, and S5 (Fig. 3C, subpanels 4-6).

In addition to the main direct RNAP-ribosome interaction involving RNAP βSI2 in TTC-B, a second direct RNAP-ribosome interaction occurs in TTC-B1 and TTC-B4, involving contact between the RNAP β pincer tip (βPT), which forms part of the binding site on RNAP for NusG-N (*13*), and ribosomal protein S10 (Fig. 3E).

In EM density maps of TTC-B, EM density is present for the NusG-N globular region (residues 1-104) at its expected binding location on RNAP (*13*; Figs. 3A, S4F, S5F). EM density is not observed for the NusG-N/NusG-C linker and NusG (Figs. 3A; S4F, S5F), but, as noted above, at the resolution of the maps for TTC-B, strong conclusions cannot be drawn from unobserved density. Molecular modelling of putative TTC-B1 through TTC-B6 complexes in which NusG-C interacts with ribosomal protein S10 as in atomic structures of the NusG-C-S10 and NusG-C-ribosome complexes (*2*,*8*), indicates that the distances between the C-terminal ordered residue of NusG-N (residue 104) and the first ordered residue of NusG-C (residue 125) would be 63-81 Å, all of which easily could be spanned by a 22-residue linker (22 residues x 3.9 Å/residue = 86 Å; Figs. 3D-E, S6). We propose that, in TTC-B, NusG bridges RNAP and ribosome in *E. coli*.

TTC-B is fully compatible with known functional aspects of transcription elongation and transcriptional regulation in *E. coli*, including compatibility with binding of NusA, formation of the λ N antitermination complex, formation of the 21 Q antitermination complex, binding of σR2, and formation of pause and termination hairpins (Fig. S7). Based on the observation that TTC-B is compatible with NusG bridging and with known functional aspects of transcription elongation, we propose that TTC-B is the state that mediates NusG-dependent transcription-translation coupling in *E. coli*.

TTC-C was obtained with nucleic-acid scaffolds having mRNA spacers of 7, 8, or 9 codons--but not with shorter mRNA spacers (Figs. 1B-right, 4, S4-S5; Table S2). TTC-C differs from TTC-A by a ~45° rotation of RNAP relative to the ribosome (Fig. 1B; movie S4) and differs from TTC-B by a ~180° rotation of RNAP relative to the ribosome (Fig. 1B; movie S5). EM density maps at 8.9-10 Å resolution were obtained (~12 Å and ~7 Å local resolution for TEC and ribosome, respectively), enabling unambiguous rigid-body docking (Figs. 4, S4F, S5F). Three TTC-C subclasses were obtained; TTC-C1 through TTC-C3, differing in distance from RNAP to ribosome and rotational orientation of RNAP relative to the ribosome (Fig. 4B, S4F, S5F; movie S6). For TTC-C, as with TTC-B, the different mRNA spacer lengths--7 , 8, or 9 codons--are accommodated by differences in the distance between RNAP and ribosome (smaller with shorter spacers and larger with longer spacers; Fig. 4C) and differences in rotational orientation of RNAP relative to the ribosome (smaller rotation with shorter mRNA spacers and larger rotation with longer RNA spacers; Fig 4C).

For TTC-C, as with TTC-B, the main direct interaction between RNAP and ribosome is mediated by RNAP βSI2, and the different distances and rotational orientations of RNAP relative to ribosome in complexes containing different mRNA spacer lengths are accommodated through the flexible molecular-swivel attachment of βSI2 to the remainder of RNAP (Fig. 4C). In TTC-C1 and TTC-C3, the RNAP-distal end of βSI2, including the loop, interacts with ribosomal proteins S2, S3, and S5 (Figs. 4C-D). In TTC-C2, βSI2 is disordered (Fig. 4C), but, at the resolution of the maps for TTC-C, no strong conclusions can be drawn, and molecular modeling indicates that βSI2 would be located in the RNAP-ribosome interface, potentially positioned to make similar interactions (Fig. 4C).

In EM density maps of TTC-C, EM density is present for NusG-N at the expected binding location on RNAP, and no density is observed for NusG-N/NusG-C linker and NusG-C (*13*; Figs. 4A, S4F, S5F). In TTC-C, the shortest sterically allowed distance between NusG-N and NusG-C modelled as bound to its molecular target on the ribosome, ribosomal protein S10 (*2*, *5*-*8*), is >180 Å, which is at least 2 times the length that could be spanned by the NusG-N/NusG-C linker, indicating that TTC-C is incompatible NusG bridging (Fig. S6).

TTC-C, like TTC-B, is compatible with all other known functional aspects of transcription elongation and transcriptional regulation in *E. coli*, including compatibility with binding of NusA, formation of the λ N antitermination complex, formation of the 21 Q antitermination complex, binding of σR2, and formation of pause and termination hairpins (Fig. S7). We propose that TTC-C mediates NusG-independent transcription-translation coupling in *E. coli*.

The results presented define three structural classes of TTCs--TTC-A (the previously reported "expressome"; *11*), TTC-B, and TTC-C--and show that TTC-B and TTC-C have structural properties suggesting they mediate NusG-dependent and NusG-independent transcription-translation coupling in *E. coli*.

The proposed structural basis of transcription-translation coupling in TTC-B--a flexible rod (RNAP βSI2) and a flexible tether (NusG)--bears a notable resemblance to the "buffer-and-chain coupler" used in locomotive-locomotive coupling (*25*; Figs. 3C, 4C). A "buffer-and-chain coupler" comprises a flexible rod ("buffer") that enables the rear locomotive to push the forward locomotive and a flexible tether ("chain") that enables the forward locomotive to pull the rear locomotive (*25*; Figs. 3C, 4C). The flexibility in both the "buffer" and the "chain" accommodates differences in spacing between the forward and rear locomotives and accommodates curves in the track on which the locomotives move. Extending this analogy, we hypothesize that, in TTC-B, the flexible rod βSI2 ("buffer") enables the ribosome to push RNAP, and thereby assist RNAP in overcoming transcriptional pauses, and the flexible tether NusG ("chain") enables RNAP to pull the ribosome, and thereby to assist the ribosome in overcoming translational pauses. Further extending this analogy, the flexibility of both flexible rod βSI2 and the

flexible tether NusG enables TTC-B to accommodate differences in spacing between RNAP and ribosome--which is crucial for coupled molecular motors where one molecular motor, RNAP, moves in 1-nt steps, and the other molecular motor, the ribosome, moves in 3-nt steps--and enables TTC-B to accommodate local conformational differences in the mRNA track on which the two motors run.

The proposed structural basis for transcription-translation coupling in TTC-C involves only a single coupler: flexible rod βSI2. Extending the analogy of the previous paragraph, we hypothesize that, in TCC-C, the ribosome can push RNAP, assisting RNAP in overcoming transcriptional pausing, but RNAP cannot pull the ribosome, and this cannot assist the ribosome in overcoming translational pausing.

The structural module of RNAP that is proposed to mediate coupling in both TTC-B and TTC-C, flexible rod βSI2, is a lineage-specific sequence insertion that is present in RNAP from Gram-negative bacteria and *Chlamydia*, but that is absent in RNAP from Gram-positive bacteria, *Deinococcus*/*Thermus*-clade bacteria, and *Mycoplasma* (Fig. S8; *22-24*). We infer that Gram-positive bacteria, *Deinococcus*/*Thermus*-clade bacteria, and *Mycoplasma*--with RNAP lacking βSI2--either do not perform tight transcription-translation coupling in which the ribosome assists RNAP in overcoming transcriptional pausing, or do so using a different structural module of RNAP, or do so using an additional factor.

Precise deletion of βSI2 of *E. coli* RNAP results in an RNAP derivative unable to support growth on minimum media (*22*.). We hypothesize that a defect in transcription-translation coupling--specifically, a defect in ability of ribosomes to assist RNAP in overcoming transcriptional pausing, accounts for the observed auxotrophy. The observation that βSI2 is essential for growth on minimal media, but not for growth on rich media (*22*), is consistent with the proposal that different transcription-translation coupling mechanisms function in different transcription units and in different regulatory contexts (*6*).

The results presented reframe our understanding of the structural and mechanistic basis of transcription-translation coupling, providing high-resolution structures of the previously described "expressome" (*11*; TTC-A1), demonstrating the incompatibility of the previously described "expressome" with general transcription-translation coupling, and defining two new structural states, TTC-B and TTC-C, with properties assignable to, respectively, NusG-dependent and NusG-independent

transcription-translation coupling. In addition, the results provide testable new hypotheses regarding the identity of the RNAP structural module crucial for transcription-translation coupling (βSI2), regarding the interactions made by that structural module (interactions with ribosomal proteins S2, S3, and S5), regarding the functional role of those interactions in transcription-translation coupling (RNAP-ribosome cooperativity in overcoming transcriptional pausing), and regarding lineage-specific differences in transcription-translation coupling in Gram-negative bacteria and *Chlamydia* vs. in Gram-positive bacteria and *Mycoplasma* (presence of βSI2 vs. absence of βSI2).

References:

1. O. Miller, B. Hamkalo, C. Thomas, Visualization of bacterial genes in action, *Science* **169**, 392-395 (1970).
2. B. Burmann, K. Schweimer, X. Luo, M. Wahl, B. Stitt, M. Gottesman, P. Rösch, A NusE-NusG complex links transcription and translation., *Science* **328**, 501-504 (2010).
3. S. Proshkin, A. Rahmouni, A. Mironov, E. Nudler, Cooperation between translating ribosomes and RNA polymerase in transcription elongation., *Science* **328**, 504-508 (2010).
4. D. Castro-Roa, N. Zenkin, *In vitro* experimental system for analysis of transcription-translation coupling, *Nucl. Acids Res.* **40**, e45 (2012).
5. K. McGary, E. Nudler, RNA polymerase and the ribosome: the close relationship, *Curr. Opin. Microbiol.* **16**, 112-117 (2013).
6. I. Artsimovitch, Rebuilding the bridge between transcription and translation, *Mol. Microbiol.* **108**, 467-472 (2018).
7. S. Saxena, K. Myka, R. Washburn, N. Costantino, L. Court, E. Gottesman, *Escherichia coli* transcription factor NusG binds to 70S ribosomes, *Mol. Microbiol.* **108**, 495-504 (2018).
8. R. Washburn, P. Zuber, M. Sun, Y. Hashem, B. Shen, W. Li, S. Harvey, S. Knauer, J. Frank, M. Gottesman, *Escherichia coli* NusG links the lead ribosome with the transcription elongation complex, *bioRxiv*, www.biorxiv.org/content/10.1101/871962v1 (2019).
9. M. Chen, K. Fredrick, Measures of single- versus multiple-round translation argue against a mechanism to ensure coupling of transcription and translation, *Proc. Natl. Acad. Sci. USA.* **115**, 10774-10779 (2018).
10. M. Strauss, C. Vitiello, K. Schweimer, M. Gottesman, P. Rosch, S. Knauer, Transcription is regulated by NusA:NusG interaction, *Nucl. Acids Res.* **44**, 5971-5982 (2016).
11. R. Kohler, A. Mooney, J. Mills, R. Landick, P. Cramer, Architecture of a transcribing-translating expressome, *Science* **356**, 194-197 (2017).
12. G. Demo, A. Rasouly, N. Vasilyev, V. Svetlov, A. Loveland, R. Diaz-Avalos, N. Grigorieff, E. Nudler, A. Korostelev, Structure of RNA polymerase bound to ribosomal 30S subunit, *Elife* **6**, e28560 (2017).
13. J. Kang, R. Mooney, Y. Nedialkov, J. Saba, T. Mishanina, I. Artsimovitch, R. Landick, S. Darst, Structural basis for transcript elongation control by NusG family universal regulators, *Cell* **173**, 1650-1662 (2018).
14. X. Guo, G. Myasnikov, J. Chen, C. Crucifix, G. Papai, M. Takacs, P. Schultz, A. Weixlbaumer, Structural basis for NusA stabilized transcriptional pausing, *Mol. Cell* **69**, 816-827 (2018).
15. F. Krupp, N. Said, Y. Huang, B. Loll, J. Burger, T. Mielke, C. Spahn, M. Wahl, Structural basis for the action of an all-purpose transcription anti-termination factor, *Mol. Cell* **74**, 143-157 (2019).
16. J. Shi, X. Gao, T. Tian, Z. Yu, B. Gao, A. Wen, L. You, S. Chang, X. Zhang, Y. Zhang, Y. Feng, Structural basis of Q-dependent transcription antitermination., *Nature Commun.* **10**, 2925 (2019).
17. Z. Yin, J. Kaelber, R. Ebright, Structural basis of Q-dependent antitermination, *Proc. Natl. Acad. Sci. USA* **116**, 18384-18390 (2019).
18. S. A. Perdue, J. Roberts, Sigma(70)-dependent transcription pausing in *Escherichia coli*, *J. Mol. Biol.* **412**, 782-792 (2011).

19. S. Goldman, N. Nair, C. Wells, B. Nickels, A. Hochschild, The primary sigma factor in *Escherichia coli* can access the transcription elongation complex from solution *in vivo*, *Elife* **4**, e10514 (2015).

20. Y. Kang, V. Mishanina, J. Bellecourt, A. Mooney, A. Darst, R. Landick, RNA polymerase accommodates a pause RNA hairpin by global conformational rearrangements that prolong pausing, *Mol. Cell* **69**, 802-815 (2018).

21. J. Roberts, Mechanisms of bacterial transcription termination, *J. Mol. Biol.* **431**, 4030-4039 (2019).

22. I. Artsimovitch, V. Svetlov, K. Murakami, R. Landick, Co-overexpression of *Escherichia coli* RNA polymerase subunits allows isolation and analysis of mutant enzymes lacking lineage-specific sequence insertions, *J. Biol. Chem.* **278**, 12344-12355 (2003).

23. W. Lane, S. Darst, Molecular evolution of multisubunit RNA polymerases: sequence analysis., *J. Mol. Biol.* **395**, 671-685 (2010).

24. N. Opalka, J. Brown, W. Lane, K. Twist, R. Landick, F. Asturias, S. Darst, Complete structural model of *Escherichia coli* RNA polymerase from a hybrid approach., *PLoS Biol.* **8**, e1000483 (2010).

25. https://en.wikipedia.org/wiki/Buffers_and_chain_coupler

Figures:

Fig. 1. Structure determination: transcription-translation complexes (TTCs)

(A) Nucleic-acid scaffolds. Each scaffold comprises nontemplate- and template-strand oligodeoxyribonucleotides (black) and one of six oligoribonucleotides having spacer length n of 4, 5, 6, 7, 8, or 9 codons (red), corresponding to mRNA. Dashed black box labeled "TEC," portion of nucleic-acid scaffold that forms TEC upon addition of RNAP (10 nt nontemplate- and template-strand ssDNA segments forming "transcription bubble," 10 nt of mRNA engaged with template-strand DNA as RNA-DNA "hybrid," and 5 nt of mRNA, on diagonal, in RNAP RNA-exit channel); dashed black lines labeled "ribosome P-site," mRNA AUG codon intended to occupy ribosome active-center P site upon addition of ribosome and $tRNA^{fMet}$; "spacer," mRNA spacer between TEC and AUG codon in ribosome active-center P site.

(B) Cryo-EM structures of TTC-A (obtained with spacer lengths of 4-7 codons), TTC-B (obtained with spacer lengths of 7-9 codons), and TTC-C (obtained with spacer lengths of 7-9 codons). Structures shown are TTC-A1 (3.7 Å; n = 4; Table S1), TTC-B1 (7.0 Å; n = 9; Table S2), and TTC-C1 (10 Å; n = 9; Table S2). Images show EM density (gray surface) and fit (ribbons) for TEC and NusG-N (at top; direction of transcription, defined by downstream dsDNA, indicated by arrow) and for ribosome 30S and 50S subunits and P- and E-site tRNAs (at bottom). RNAP β', β, α^{I}, α^{II}, and ω subunits are in pink, cyan, light green, and dark green, and gray; 30S subunit, 50S subunit, P-site tRNA, E-site tRNA are in yellow, gray, green, and orange; DNA nontemplate strand, DNA template strand, and mRNA are in black, blue, and brick-red (brick-red dashed line where modeled). RNAP βSI2, NusG-N, and ribosomal protein S10 are in dark blue, red, and magenta.

Fig. 2. Cryo-EM structure of TTC-A

(A) Structure of TTC-A (TTC-A1; 3.7 Å; n = 4; Table S1). Two orthogonal views. Colors as in Fig. 1B.

(B) Accommodation of mRNA spacer lengths of 4, 5, 6 and 7 codons in TTC-A. EM density, blue mesh; mRNA, brick-red (disordered mRNA nucleotides indicated by dashed oval); template-strand DNA in

RNA-DNA hybrid, blue; RNAP active-center catalytic Mg^{2+}, purple sphere; tRNA in ribosome P site, green. Upper and lower black horizontal lines indicate edges of RNAP and ribosome.

(C) RNAP-ribosome interface in TTC-A (TTC-A1; n = 4; identical interface for n = 5, 6, or 7), showing interactions between RNAP β' zinc binding domain, (ZBD, pink; Zn^{2+} ion as black sphere), RNAP β flap, cyan, RNAP flap tip helix (β FTH; cyan dashed line), and RNAP α^I (green) interacting with ribosomal proteins S4 (forest green), S3 (orange), and S10 (magenta) and mRNA (brick red). Portions of RNAP β' and ribosome 30S not involved in interactions are shaded pink and yellow, respectively.

(D) RNAP-ribosome interactions involving RNAP β' ZBD and S4 (subpanel 1), RNAP β flap and S3 (subpanel 2; β FTH, dashed line;, β and S3 residues that interact with mRNA, cyan and orange spheres with red outlines; mRNA, brick-red), RNAP α^I and S3 (subpanel 3), and RNAP α^I and S10 (subpanel 4). Other colors as in (C).

(E) Absence of EM density for RNAP ω subunit. EM density, blue mesh; atomic models for RNAP β' and S2, pink ribbon and red-pink ribbon, respectively; location of missing EM density for ω, dashed oval; ω in TEC in absence of ribosome (PDB 6P19; *17*), white ribbon.

Fig. 3. Cryo-EM structure of TTC-B

(A) Structure of TTC-B (TTC-B1; 7.6 Å; n = 9; Table S2). Views and colors as in Fig. 2A.

(B) Subclasses of TTC-B. Superimposition of TTC-B subclasses B1-B6 (RNAP EM density colored by subclass; ribosome EM density in gray for subclass B1). View as in (A).

(C) RNAP-ribosome interactions and NusG bridging in TTC-B subclasses B1-B6. View as in Fig. 3A-left. Ribosomal proteins S2, S3, S5, and S10 in red-pink, orange, brown, and magenta. Dark blue ribbons, βSI2; dark blue dashed ovals, disordered βSI2 residues; brick-red dashed line, modeled mRNA; red dashed circle and red dashed line, modeled NusG-C and NusG-N/NusG-C linker; distances in Å, modeled distances between NusG-N residue 104 and NusG-C residue 125. Other colors as in (A).

(D) RNAP-ribosome interface in TTC-B (TTC-B1; n = 9), showing interaction between βSI2 (dark blue) and ribosomal proteins S2, S3 and S5 (red-pink, orange, and brown). Portions of RNAP β and ribosome 30S not involved in interactions are shaded cyan and yellow, respectively.

(E) Interaction between RNAP β pincer tip (βPT; cyan) and ribosomal protein S10 (magenta) and modeled NusG bridging, with NusG-N interacting with RNAP β' subunit (pink shading) and RNAP βPT, with NusG-C modeled interacting with S10 (red ribbon and dashed red circle), and modeled linker between NusG-N residue 104 and NusG-C residue 125 as dashed red line. 30S residues not involved in interactions are indicated with yellow shading.

Fig. 4. Cryo-EM structure of TTC-C

(A) Structure of TTC-C (TTC-C1; 10 Å; n = 9; Table S2). Views and colors as in Fig. 2A.

(B) Subclasses of TTC-C. Superimposition of TTC-C subclasses C1-C3 (RNAP EM density colored by subclass; ribosome EM density in gray for subclass C1). View as in (A).

(C) RNAP-ribosome interactions in TTC-C subclasses C1-C3. View as in (A), left. Ribosomal proteins S2, S3, S5, and S10 in red-pink, orange, brown, and magenta. Dark blue ribbons, βSI2; dark blue dashed oval, disordered βSI2 residues; brick-red dashed line, modeled mRNA. Other colors as in (A).

(D) RNAP-ribosome interface in TTC-C (TTC-C1; n = 9), showing interaction between βSI2 in dark blue) and ribosomal proteins S2, S3 and S5 (red-pink, orange, and brown). Portions of RNAP β subunit and ribosome 30S subunit not involved in interactions are shaded cyan and yellow, respectively.

Acknowledgements:

This work was supported by University of California discretionary funds to G.B., University of Michigan discretionary funds M.S., and National Institutes of Health (NIH) grant GM041376 to R.H.E. We thank the Rutgers University Cryo-EM Core facility, the University of Michigan Life Sciences Institute Cryo-EM Facility, the National Center for CryoEM Access and Training (supported by NIH grant GM129539, Simons Foundation grant SF349247, and New York state grants), and the Pacific Northwest Center for Cryo-EM (supported by NIH grant GM129547 and Department of Energy Environmental Molecular Sciences Laboratory) for microscope access; K. Kuznedelov and K. Severinov for plasmids; and L. Minakhin, B. Nickels, and J. Winkelman for discussion; and E. Eng and H. Wei for assistance.

Supplementary Materials:

Materials and Methods

Tables S1-S2

Figures S1-S8

Movies S1-S6

Supplemental References (*26-42*)

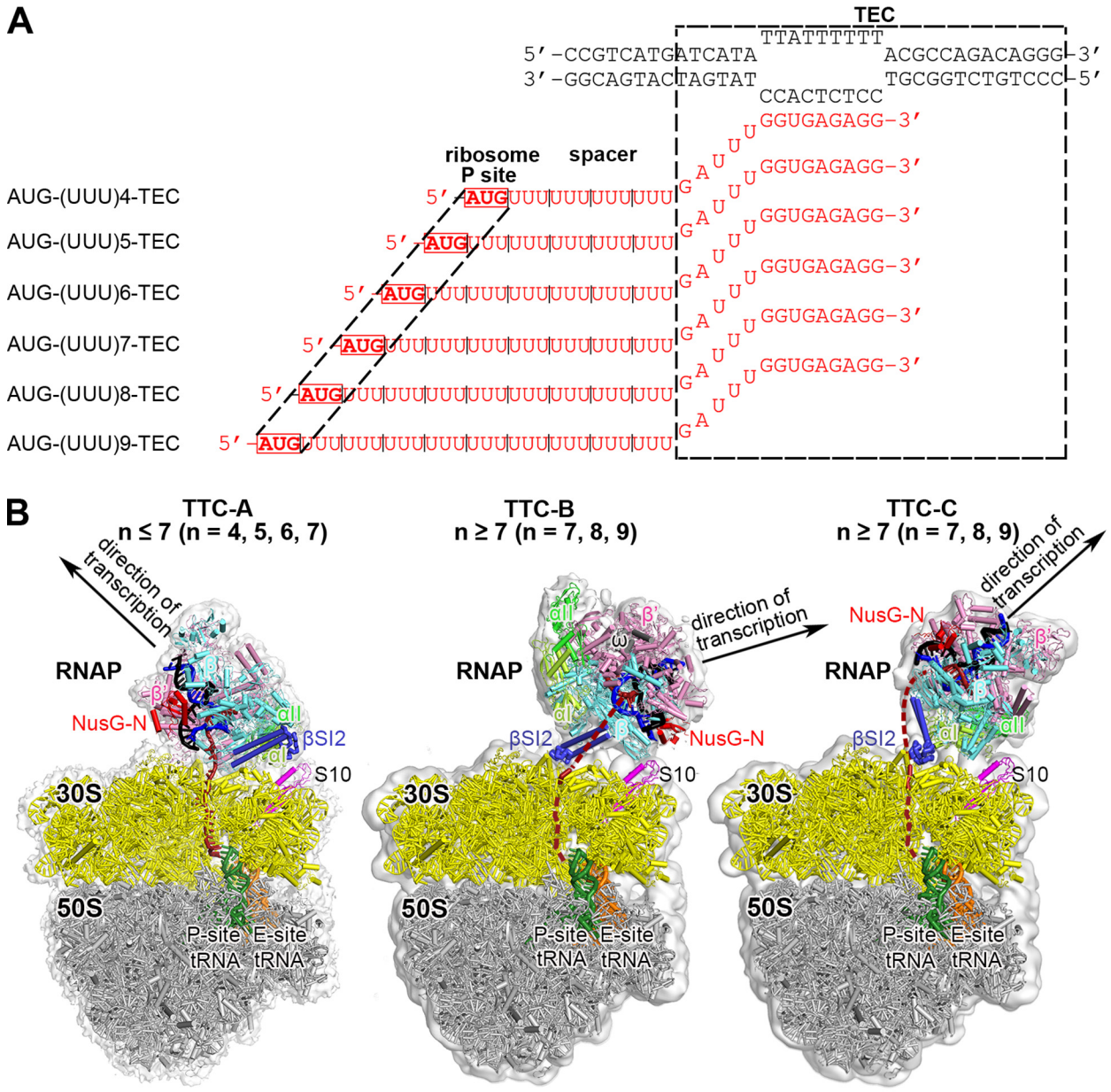
A
TEC
TTATTTTTT
5′-CCGTCATGATCATA
ACGCCAGACAGGG-3′
3′-GGCAGTACTAGTAT
TGCGGTCTGTCCC-5′
CCACTCTCC
GGUGAGAGG-3′
ribosome
P site
spacer
AUG-(UUU)4-TEC
AUG-(UUU)5-TEC
AUG-(UUU)6-TEC
AUG-(UUU)7-TEC
AUG-(UUU)8-TEC
AUG-(UUU)9-TEC
B
TTC-A
n ≤ 7 (n = 4, 5, 6, 7)
TTC-B
n ≥ 7 (n = 7, 8, 9)
TTC-C
n ≥ 7 (n = 7, 8, 9)
direction of transcription
RNAP
NusG-N
βSI2
αII
αI
S10
30S
50S
P-site tRNA
E-site tRNA

Figure 1

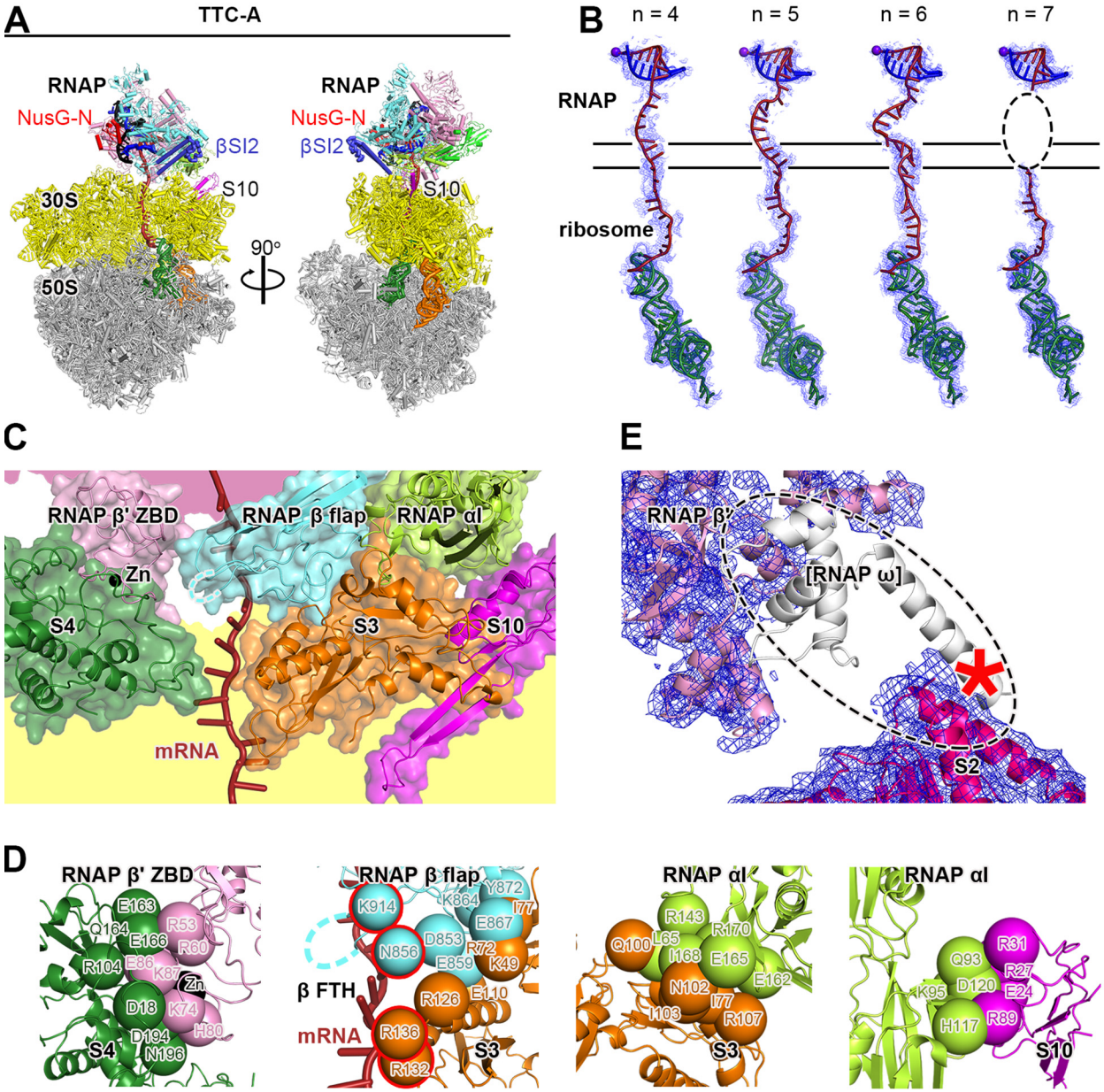
A
TTC-A
RNAP
NusG-N
βSI2
S10
30S
50S
90°
B
n = 4
n = 5
n = 6
n = 7
RNAP
ribosome
C
RNAP β' ZBD
RNAP β flap
RNAP αI
Zn
S4
S3
S10
mRNA
E
RNAP β'
[RNAP ω]
S2
D
RNAP β' ZBD
E163
Q164
E166
R53
R60
R104
E86
K87
Zn
D18
K74
H80
D194
N196
S4
RNAP β flap
K914
K864
Y872
I77
E867
N856
D853
R72
E859
K49
β FTH
R126
E110
mRNA
R136
R132
S3
RNAP αI
R143
R170
Q100
L65
I168
E165
E162
N102
I77
I103
R107
S3
RNAP αI
R31
Q93
R27
D120
E24
K95
H117
R89
S10

Figure 2

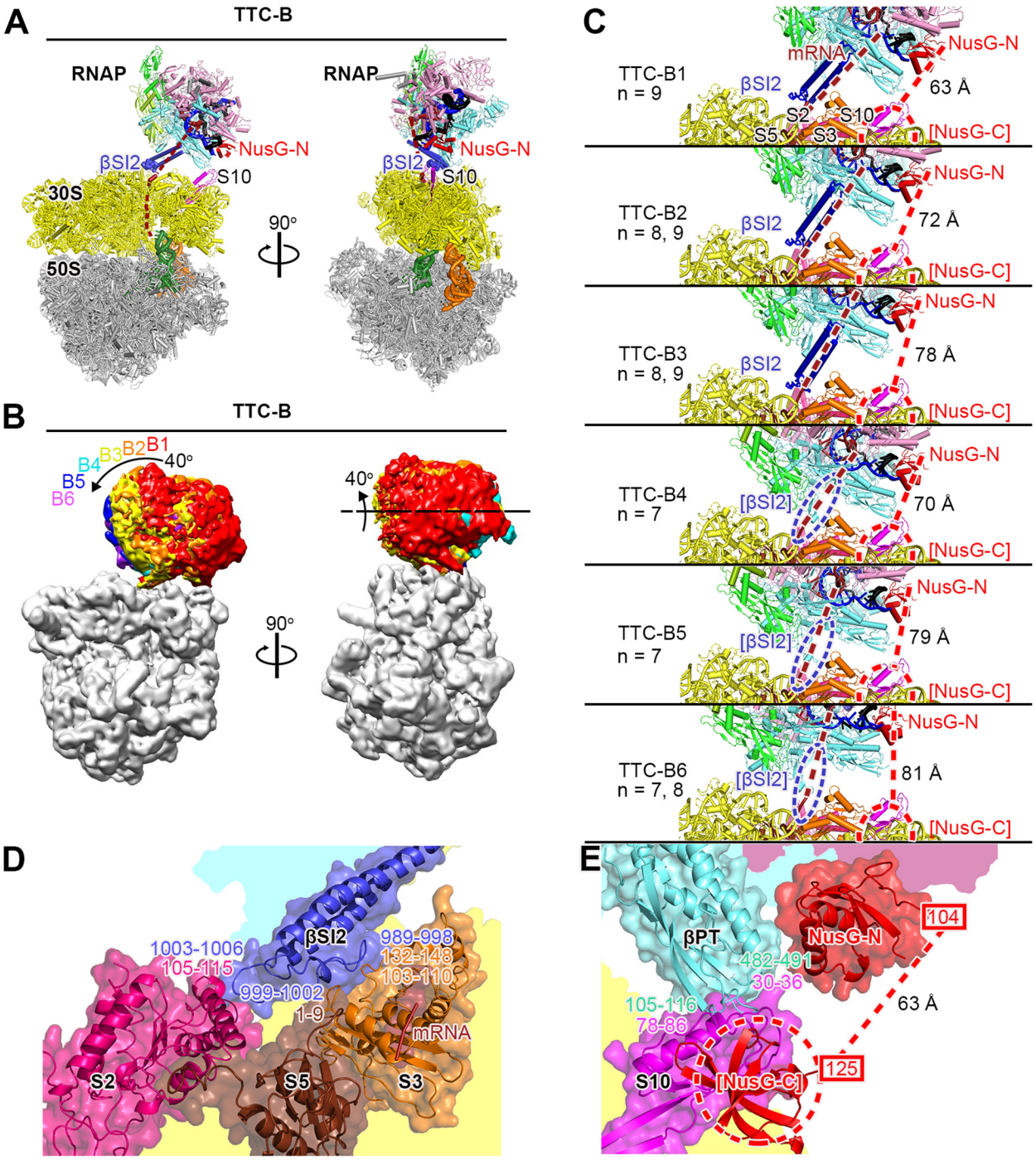
A
TTC-B
RNAP
RNAP
NusG-N
NusG-N
βSI2
βSI2
S10
S10
30S
50S
90°
B
TTC-B
B1
B2
B3
B4
B5
B6
40°
40°
90°
C
mRNA
NusG-N
TTC-B1
n = 9
βSI2
63 Å
S2
S10
S5
S3
[NusG-C]
NusG-N
TTC-B2
n = 8, 9
βSI2
72 Å
[NusG-C]
NusG-N
TTC-B3
n = 8, 9
βSI2
78 Å
[NusG-C]
NusG-N
TTC-B4
n = 7
[βSI2]
70 Å
[NusG-C]
NusG-N
TTC-B5
n = 7
[βSI2]
79 Å
[NusG-C]
NusG-N
TTC-B6
n = 7, 8
[βSI2]
81 Å
[NusG-C]
D
βSI2
1003-1006
105-115
989-998
132-148
103-110
999-1002
1-9
mRNA
S2
S5
S3
E
βPT
NusG-N
104
482-491
30-36
63 Å
105-116
78-86
S10
[NusG-C]
125

Figure 3

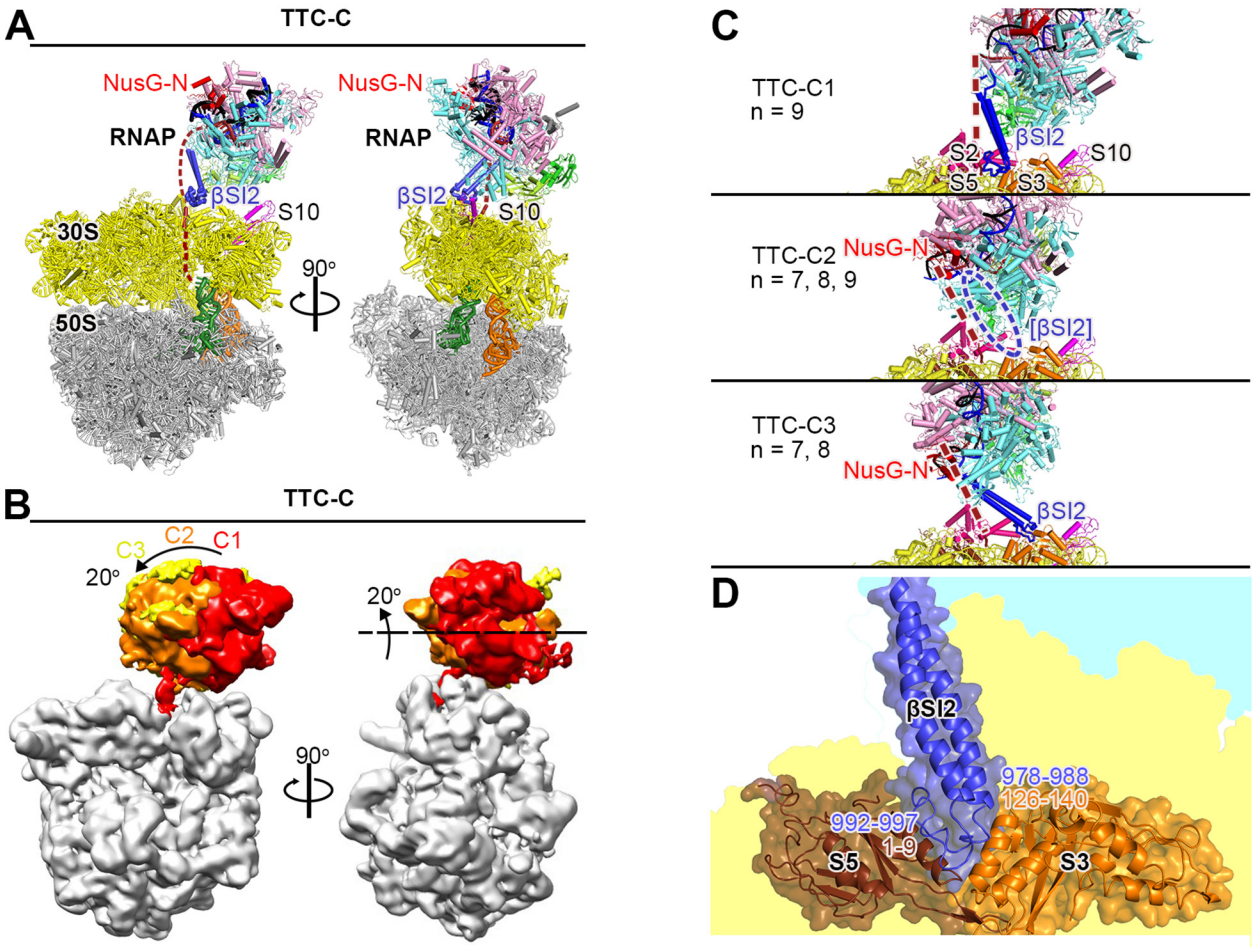
A
TTC-C
NusG-N
RNAP
βSI2
S10
30S
50S
90°
B
TTC-C
C3
C2
C1
20°
C
TTC-C1
n = 9
S2
S5
S3
TTC-C2
n = 7, 8, 9
[βSI2]
TTC-C3
n = 7, 8
D
978-988
126-140
992-997
1-9

Figure 4

Supplementary Materials:

Materials and Methods:

***E. coli* RNAP core enzyme**
E. coli RNAP core enzyme was prepared from *E. coli* strain BL21 Star (DE3) (ThermoFisher) transformed with plasmid pVS10 (*26*; encodes *E. coli* RNAP β' with C-terminal hexahistidine tag, β, α, and ω subunits), as described (*27*). The product (purity >95%) was stored in aliquots in RNAP storage buffer (10 mM Tris-HCl, pH 7.6, 100 mM NaCl, 0.1 mM EDTA, and 5 mM dithiothreitol) at -80ºC.

***E. coli* NusG**
E. coli NusG was prepared from *E. coli* strain BL21 Star (DE3) (ThermoFisher) transformed with plasmid pIA247 (*28?*; K. Kuznedelov and K. Severinov; encodes *E. coli* NusG with C-terminal hexahistidine tag), as described (*28*). The product (purity >95%) was stored in aliquots in RNAP storage buffer at -80ºC.

***E. coli* 70S ribosome**
E. coli 70S ribosomes were prepared from *E. coli* strain JE28 (*29*; grown in media containing 50 $^{g/ml}$ kanamycin to OD600 = 1.0 in a 50 L fermenter, harvested, flash frozen in liquid N_2, and stored at -80°C), as described (*30-31*).

Cells in 20 mM Tris-HCl, pH 7.6, 100 mM NH_4Cl, 10.5 mM $Mg(OAc)_2$, 0.5 mM EDTA, 4 mM 2-mercaptoethanol, and 1.5 mg/ml benzamidine HCl, and phenylmethylsulfonyl fluoride at 4°C were lysed [EmulsiFlex C3 (Avestin); 3 passages at 8,000 psi], and lysates were cleared by centrifugation (SS-34 rotor; 2 x 30 min at 16,000 rpm at 4°C). Ribosomes were pelleted by centrifugation through a 12 ml high-salt sucrose cushion [20 mM Tris-HCl, pH 7.5, 1.1 M sucrose, 500 mM NH_4Cl, 10.5 mM $Mg(OAc)_2$, 0.5 mM EDTA, and 4 mM 2-mercaptoethanol; Type 45 Ti rotor; 20 h at 43,000 rpm at 4°C; *32*]. Pelleted ribosomes were dissociated into ribosome 30S and 50S subunits by re-suspension in dissociation buffer [20 mM Tris-HCl, pH 7.6, 200 mM NH_4Cl, 1 mM $Mg(OAc)_2$, and 4 mM 2-mercaptoethanol], and 30S and 50S subunits were separated by loading an aliquot (250 A_{260} units) onto a 36 ml 10-30% sucrose gradient in dissociation buffer and centrifugation (SW28 rotor; 15 h at 20,000 rpm at 4°C). The gradient was fractionized, fractions containing 30S and 50S subunits were identified by UV absorbance and pooled, and pooled fractions containing 30S and 50S subunits were pelleted (Type 45 rotor; 21 h at 41,000 rpm at 4°C). Purified 30S and 50S subunits were reassociated by suspension in reassociation buffer (20 mM Tris-HCl, pH 7.5, 30 mM KCl, 20 mM $Mg(OAc)_2$, and 4 mM 2-mercaptoethanol), dilution to 200 A_{260} units/ml, and incubation 30 min at 40°C, and 70S ribosomes were separated from non-reassociated 30S and 50S subunits by loading a 1 ml aliquot onto a 36 ml 10-30% sucrose gradient in reassociation buffer and centrifugation (SW28 rotor; 15 h at 18,000 rpm at 4°C). The gradient was fractionized, fractions containing 70S ribosomes were identified by UV absorbance and pooled, and pooled fractions were pelleted (Type 45 rotor; 21 h at 41,000 rpm at 4°C). Pelleted 70S ribosomes were re-suspended in reassociation buffer, flash frozen in liquid N_2, and stored in aliquots at -80°C.

***E. coli* tRNA^{fMet}**
Escherichia coli tRNA^{fMet} was purchased (MP Biomedical), dissolved to 100 μM in 5 mM Tris-HCl, pH 7.5, and stored in aliquots at -80ºC.

Nucleic-acid scaffolds
Oligodeoxyribonucleotides and oligoribonucleotides (sequences in Fig. 1A) were purchased (Integrated DNA Technologies; PAGE-purified), dissolved to 1 mM in 5 mM Tris-HCl, pH 7.5, and stored in aliquots at-80ºC. Nucleic-acid scaffolds (sequences in Fig. 1A) were prepared by mixing 60 μM nontemplate-strand oligodeoxyribonucleotide, 60 μM template-strand oligodeoxyribonucleotide, and 60

μM oligoribonucleotide in 100 μl annealing buffer (5 mM Tris-HCl, pH 7.5), heating 10 min at 95ºC, and cooling slowly (3 h) to 22º, and nucleic-acid scaffolds were stored in aliquots at -80ºC.

Transcription-translation complexes (TTCs)
TTCs for structure determination in the presence of CHAPSO were prepared by mixing 10 μl 20 μM RNAP core enzyme (in RNAP storage buffer), 30 μl 0 or 30 μM NusG (in RNAP storage buffer), and 4 μl 60 μM nucleic-acid scaffold (in 5 mM Tris-HCl, pH 7.5), 10 μl 10x reassociation buffer, and 26 μl water; incubating 10 min at 22ºC; adding 10 μl 100 μM $tRNA^{fMet}$ (in 5 mM Tris-HCl, pH 7.5), 5 μl 40 μM 70S ribosome (in reassociation buffer), and 5 μl water; and further incubating 10 min at 22ºC. The resulting TTCs were transferred to a pre-chilled 0.5 ml Amicon Ultracel 30K concentrator (EMD Millipore), concentrated to 35 μl by centrifugation (15 min at 20,000xg at 4ºC), incubated 30 min on ice for 30 min, supplemented with 3.8 μl of ice-cold 80 mM CHAPSO, and immediately applied to grids (see "Cryo-EM structure determination: sample preparation").

TTCs for structure determination in the absence of CHAPSO were prepared by mixing 1 μl 20 μM RNAP core enzyme (in RNAP storage buffer), 3 μl 30 μM NusG (in RNAP storage buffer) and 0.4 μl 60 μM nucleic-acid scaffold (in 5 mM Tris-HCl, pH 7.5), 1 μl 10x reassociation buffer, and 2.6 μl water; incubating 10 min at 22ºC; adding 1 μl of 100 μM $tRNA^{fMet}$ (in 5 mM Tris-HCl, pH 7.5), 0.5 μl 40 μM 70S ribosomes (in reassociation buffer), and 0.5 μl water; and further incubated 10 min at 22ºC. The resulting TTCs were incubated 30 min on ice, mixed with 35 μl ice-cold reassociation buffer, and immediately applied to grids (see "Cryo-EM structure determination: sample preparation").

Cryo-EM structure determination: sample preparation
EM grids were prepared using a Vitrobot Mark IV autoplunger (FEI/ThermoFisher), with the environmental chamber at 22°C and 100% relative humidity. Samples (3 μl) were applied to 2/1 Quantifoil Cu 300 holey-carbon grids (Quantifoil); glow-discharged 60 s using a PELCO glow-discharge system (Ted Pella)]), grids were blotted with #595 filter paper (Ted Pella) for 7 s at 22°C, and grids were flash-frozen by plunging in a liquid ethane cooled with liquid N_2, and grids were stored in liquid N_2.

Cryo-EM structure determination: data collection and data reduction (TTC-A; n = 4, 5, 6, or 7; with CHAPSO)
Cryo-EM data for TTC-A (n = 5; with CHAPSO) were collected at the Rutgers University Cryo-EM core Facility, using a 200 kV Talos Arctica (FEI/ThermoFisher) electron microscope equipped with a GIF Quantum K2 direct electron detector (Gatan). Data were collected automatically in counting mode, using EPU (FEI/ThermoFisher), a nominal magnification of 130,000x, a calibrated pixel size of 1.038 Å/pixel, and a dose rate of 4.8 electrons/pixel/s. Movies were recorded at 200 ms/frame for 6 s (30 frames), resulting in a total radiation dose of 26.7 electrons/$Å^2$. Defocus range was varied between -1.25 μm and -2 μm. A total of 1,027 micrographs were recorded from one grid over two days. Micrographs were gain-normalized and defect-corrected.

Data were processed as summarized in Figs. S1A-D. Data processing was performed using a Tensor TS4 Linux GPU workstation with four GTX 1080 Ti graphic cards (NVIDIA). Dose weighting motion correction (3x3 tiles; b-factor = 150) were performed using Motioncor2 (*33*). Contrast-transfer-function (CTF) estimation was performed using CTFFIND-4.1 (*34*). Subsequent image processing was performed using Relion 3.0 (*35*). Automatic particle picking with Laplacian-of-Gaussian filtering yielded an initial set of 98,720 particles. Particles were extracted into 500x500 pixel boxes and subjected to rounds of reference-free 2D classification and removal of poorly populated classes, yielding a selected set of 27,378 particles. The selected set was 3D-classified with C1 symmetry, using a *de novo* 3D template created using 3D_initial_model under Relion 3.0. Of five classes obtained, four classes exhibited strong, well-defined density assignable to the TEC and having a spatial relationship to density for the ribosome consistent with a TTC. The 24,959 particles for these four classes were combined and 3D auto-refined

using a mask with a diameter of 450 Å. The resulting 3D auto-refined particles were further refined using a soft mask and solvent flattening and were post-processed, yielding a reconstruction at 3.7 Å overall resolution, as determined from gold-standard Fourier shell correlation (FSC; Fig. S1D; Table S1).

The initial atomic model for TTC-A (n = 5; with CHAPSO) was built by manual docking of RNAP β', β, a^{I},, a^{II} segments and NusG segments from a cryo-EM structure of an *E. coli* TEC bound to NusG (PDB 6C6U; *36*); DNA and RNA segments from a cryo-EM structure of an *E. coli* 21Q transcription antitermination "loaded" complex (PDB 6P19; *17*); ribosome 30S S2-S21 segments and 16S rRNA segments from a cryo-EM structure of an *E. coli* "expressome" (PDB 5MY1; *11*); ribosome 50S subunit from a cryo-EM structure of an *E. coli* 50S ribosomal subunit (PDB 6QDW); and P- and E-site tRNA segments from a crystal structure of a *Thermus thermophilus* 70S ribosome P- and E-site tRNA and mRNA (PDB 4V6G, *37*), using UCSF Chimera (*38*). For the RNAP β' N and C-termini (residues 1-15 and 1734-1407), the RNAP β flap-tip helix (residues 892-910), the RNAP α^{I} and α^{II} N-termini and C-terminal domain (residues 1-5 and 234-329), and the NusG-N/NusG-C linker and NusG-C (residues 117-182), density was absent, suggesting high segmental flexibility; these segments were not fitted. For the RNAP ω, density was absent, suggesting absence, low occupancy, or high segmental flexibility; RNAP ω was not fitted.

Refinement of the initial model was performed using real_space_refine under Phenix (*39*). The ribosome 30S and 50S subunits were rigid-body refined against the map, followed by real-space refinement with geometry, rotamer, Ramachandran-plot, Cβ, non-crystallographic-symmetry, secondary-structure, and reference-structure (initial model as reference) restraints, followed by global minimization and local-rotamer fitting. Secondary-structure annotation was inspected and edited using UCSF Chimera. RNAP and NusG were rigid-body refined against the map, and the RNAP β' Zn^{2+} binding domain (ZBD; residues 41-100), RNAP β flap (residues 835-891 and 911-937), RNAP α^{I} (residues 61-75 and 154-173), DNA, and mRNA segments were subjected to iterative cycles of model building and refinement in Coot (*40*). The final atomic model at was deposited in the Electron Microscopy Data Bank (EMDB) and the Protein Data Bank (PDB) with accession codes EMDB 21468 and PDB 6VYQ (Table S1).

The cryo-EM structure of TTC-A (n = 4; with CHAPSO) was determined in the same manner, but using data collected at the National Center for CryoEM Access and Training (NCCAT), using a 300 kV Krios Titan (FEI/ThermoFisher) electron microscope equipped with a Gatan K2 Summit direct electron detector (Gatan), Leginon (*41*), a nominal magnification of 105,000x, a calibrated pixel size of 1.096 Å/pixel, a dose rate of 6.33 electrons/pixel/s and 200 ms/frame for 10 s (50 frames; total radiation dose of 63.5 electrons/$Å^2$), and a defocus range of 1 µm to 3 µm, and recording a total of 6,724 micrographs were recorded from one grid over 1.5 days. The final map and atomic model had resolution of 3.8 Å (EMDB 21386; PDB 6VU3; Table 1; Fig. S1F).

Cryo-EM structures of TTC-A (n = 6; with CHAPSO) and TTC-A (n = 7; with CHAPSO) were determined in the same manner as for TTC-A (n = 5; with CHAPSO), yielding maps and atomic models with resolutions of 3.8 Å (EMDB 21469; PDB 6VYR) and 3.7 Å (EMDB 21470; PDB 6VYS), respectively (Table S1; Fig S1F).

The Cryo-EM structure of TTC-A (n = 5; with CHAPSO) in the absence of NusG was determined in the same manner as for TTC-A (n = 5; with CHAPSO), yielding a map and atomic model with resolution of 4.1 Å (EMDB 21494; PDB 6VZJ; Table S1; Fig. S2).

Cryo-EM structure determination: data collection and data reduction (TTC-A; n = 5 or 6; without CHAPSO)

Cryo-EM data for TTC-A (n = 5; without CHAPSO) were collected at the University of Michigan, Life Sciences Institute Cryo-EM Facility, using a 200 kV Glacios (FEI/ThermoFisher) electron microscope

equipped with a K2 summit direct electron detector (Gatan). Data were collected using a stage tilt of 40°, in order to overcome particle orientation preferences that precluded structure determination using data collected without stage tilt (Fig. S3C). Following insertion of the grid into the microscope, eucentric height was obtained, and the stage was tilted to 40°C by changing the stage α parameter. Data were collected automatically in counting mode, using Leginon (*41*), a nominal magnification of 45,000x, a calibrated pixel size of 0.98 Å/pixel, and a dose rate of 7 electrons/A^2/s. Movies were recorded at 200 ms/frame for 6 s (30 frames total), resulting in a total radiation dose of 42 electrons/$Å^2$. Defocus range was varied between -1 µm and -2 µm. A total of 1,212 micrographs were recorded from one grid over two days. Micrographs were gain-normalized and defect-corrected.

Data were processed as summarized in Figs. S3A-E. Dose weighting, motion correction (5x5 tiles; b-factor = 150) were performed using Motioncor2 (*33*). Per-particle CTF estimation was performed using goCTF (*42*), in order to estimate and account for the focus gradient in each tilted micrograph. Subsequent image processing was performed using Relion 3.0 (*35*). Automatic particle picking with Laplacian-of-Gaussian filtering yielded an initial set of 93,446 particles. Particles were binned 3x, extracted into 192x192 pixel boxes, and subjected to two rounds of reference-free 2D classification and removal of poorly populated classes, yielding a selected set of 41,868 particles. An *ab initio* model was generated using the SGD-based approach in Relion 3.0. 3D auto-refinement was performed against the selected set of particles, yielding a reconstruction with a global resolution that reached Nyquist frequency at 5.9 Å as determined from gold-standard Fourier shell correlation. Reference-free 2D classification and 3D auto-refinement also were performed with un-binned particles; this yielded a higher-resolution reconstruction, but did not yield clear improvement in subclass map quality. Therefore, subsequent steps were performed using subclass maps from 2D classification and 3D auto-refinement with binned data.

In the next step of data processing, particle subtraction was conducted to remove the signal from the ribosome 70S subunit, in order to characterize conformational states focused on the TEC. Because a mask accounting for the full range of conformational dynamics of the TEC was essential for the analysis, and because the segmented volume of the TEC was not reliable for generating such a mask (due to lower occupancy and higher conformational dynamics of the TEC as compared to the ribosome, and resulting weaker density for the TEC as compared to the ribosome in the map generated by 3D auto-refinement), a 3D volume was generated from signal-subtracted particles using the command relion_reconstruct, and this 3D volume was low-pass filtered to 60 Å resolution to generate a mask with an extended binary map in 9 pixels and a soft-edge in 3 pixels. The same 3D volume was low-pass filtered to 40 Å resolution and also used as an initial model for subsequent data processing. Focused 3D classifications without alignment was performed to sort different conformational states of the TEC relative to the ribosome. Of three subclasses obtained, two subclasses exhibited strong, well-defined density assignable to TEC: TTC-A1 (18%) and TTC-A2 (39%) (Fig. S3A). Non-signal-subtracted particles associated with each well-defined subclass were used to reconstruct a 3D volume using the command relion_reconstruct, and each volume was low-pass filtered to 30 Å resolution to generate an adaptive mask for subclass particles, with an extended binary map in 3 pixels and a soft-edge in 3 pixels. Local focused 3D refinement for each subclass then was performed with its adaptive mask using command relion_refine (-sigma_ang = 3). This process yielded reconstruction at 19 Å resolution for subclass A1 and at 14 Å resolution for subclass A2, as determined from gold-standard Fourier shell correlation (FSC; Fig. S3D; Table S2). Half maps were used to generate local-resolution maps using command relion_postprocess (-adhoc_bfac = 100) (Fig. S3E).

Initial atomic models for TTC-A1 and TTC-A2 (n = 5; without CHAPSO) were built by manual docking of RNAP β', β, α^I, α^{II} segments and NusG segments from a cryo-EM structure of an *E. coli* TEC bound to NusG (PDB 6C6U; *36*); DNA and RNA segments from a cryo-EM structure of an *E. coli* 21Q transcription antitermination "loaded" complex (PDB 6P19; *17*); ribosome 30S S2-S21 segments and 16S rRNA segments from a cryo-EM structure of an *E. coli* "expressome" (PDB 5MY1; *11*); ribosome 50S subunit from a cryo-EM structure of an *E. coli* 50S ribosomal subunit (PDB 6QDW); and P- and E-site

tRNA segments from a crystal structure of a *Thermus thermophilus* 70S ribosome P- and E-site tRNA and mRNA (PDB 4V6G, *37*), using UCSF Chimera (*38*). For the NusG-N/NusG-C linker and NusG-C (residues 117-182), density was absent, suggesting high segmental flexibility; the NusG-N/NusG-C linker and NusG-C segments were not fitted. RNAP SI2 (residues 936-1041) was subjected to iterative cycles of model building and refinement in Coot (*40*). The final atomic model for TTC-A2 (n = 5) has been deposited in the EMDB and PDB with accession codes EMDB 21471 and PDB 6VYT (Table S2).

Cryo-EM structures for TTC-A (n = 6; without CHAPSO) were determined in the same manner, yielding maps and atomic models with resolutions of 15 Å for TTC-A1 and 13 Å for TTC-A2 (Table S2).

Cryo-EM structure determination: data collection and data reduction (TTC-A, TTC-B, and TTC-C; n =7; without CHAPSO)

Cryo-EM data for TTC-A, TTC-B and TTC-C (n = 7; without CHAPSO) were collected at the University of Michigan, Life Sciences Institute Cryo-EM Facility. Data collection using a stage tilt of 40° was performed as in the preceding section. A total of 7,297 micrographs were recorded from one grid over five days. Micrographs were gain-normalized and defect-corrected.

Data were processed as summarized in Figs. S3A-E. Dose weighting, motion correction (5x5 tiles; b-factor = 150) were performed using Motioncor2 (*33*). Per-particle CTF estimation was performed using goCTF (*42*), in order to estimate and account for the focus gradient in each tilted micrograph. Subsequent image processing was performed using Relion 3.0. (*35*). Automatic particle picking with Laplacian-of-Gaussian filtering yielded an initial set of 765,225 particles. Particles were binned 3x, extracted into 192x192 pixel boxes, and subjected to two rounds of reference-free 2D classification and removal of poorly populated classes, yielding a selected set of 445,738 particles. An *ab initio* model was generated using the SGD-based approach in Relion 3.0. 3D auto-refinement was performed against the selected set of particles, yielding a reconstruction with a global resolution that reached Nyquist frequency at 5.9 Å as determined from gold-standard Fourier shell correlation. Reference-free 2D classification and 3D auto-refinement also were performed with un-binned particles; this yielded a higher-resolution reconstruction, but did not yield clear improvement in subclass map quality. Therefore, subsequent steps were performed using subclass maps from 2D classification and 3D auto-refinement with binned data.

In the next step of data processing, particle subtraction was conducted to remove the signal from the ribosome 70S subunit, in order to characterize conformational states focused on the TEC. Because a mask accounting for the full range of conformational dynamics of the TEC was essential for the analysis, and because the segmented volume of the TEC was not reliable for generating such a mask (due to lower occupancy and higher conformational dynamics of the TEC as compared to the ribosome, and resulting weaker density for the TEC as compared to the ribosome in the map generated by 3D auto-refinement), a 3D volume was generated from signal-subtracted particles using the command relion_reconstruct, and this 3D volume was low-pass filtered to 60 Å resolution to generate a mask with an extended binary map in 9 pixels and a soft-edge in 3 pixels. The same 3D volume was low-pass filtered to 40 Å resolution and also used as an initial model for subsequent data processing. Focused 3D classifications without alignment was performed to sort different conformational states of the TEC relative to the ribosome. Of seven subclasses obtained, six subclasses exhibited strong, well-defined density assignable to the TEC: TTC-A2 (5%), TTC-B4 (2.9%), TTC-B5 (2.8%), TTC-B6 (2.7%), TTC-C2 (3.2%), and TTC-C3 (3.3%). The three TTC-B subclasses were grouped together and subjected to an additional round of 3D classification into three subclasses without alignment. All well-defined subclasses of RNAP were selected (Figs S4A), and non-signal-subtracted particles associated with each subclass were used to reconstruct a 3D volume using the command relion_reconstruct, each volume was low-pass filtered to 30 Å resolution to generate an adaptive mask for subclass particles, with an extended binary map in 3 pixels and a soft-edge in 3 pixels. Local focused 3D refinement for each subclass was then performed with its adaptive mask using command relion_refine (-sigma_ang = 3). This process yielded reconstructions at 11 Å resolution for subclass A2; at 10 Å resolution for subclasses B4, B5, and B6; and at 9 Å resolution for

subclasses C2 and C3, as determined from gold-standard Fourier shell correlation (FSC; Figs S4D; Table S2). Half maps were used to generate local-resolution maps using command relion_postprocess (-adhoc_bfac = 100) (Fig. S4E).

Atomic models for TTC-A2, TTC-B4, TTC-B5, TTC-B6, TTC-C2, and TTC-C3 (n = 7; without CHAPSO) were built as described in the preceding section, yielding maps and atomic models with resolutions of 14 Å for TTC-A2, 9.9 Å for TTC-B4, (EMDB 21475; PDB 6VYX), 9.9 Å for TTC-B5 (EMDB 21476; PDB 6VYY), 9.9 Å for TTC-B6 (EMDB 21477; PDB 6VYZ), 8.9 Å for TTC-C2, and 8.9 Å for TTC-C3 (EMDB 21485; PDB 6VZ5) (Table S2).

Cryo-EM structure determination: data collection and data reduction (TTC-B, and TTC-C; n = 8 or 9; without CHAPSO)

Cryo-EM data for TTC-B and TTC-C (n = 9; without CHAPSO). Data collection using a stage tilt of 40° was performed as in the preceding two sections. A total of 11,989 micrographs were recorded from one grid over five days. Micrographs were gain-normalized and defect-corrected.

Data were processed as summarized in Figs. S5A-E. Dose weighting, motion correction (5x5 tiles; b-factor = 150) were performed using Motioncor2 (*33*). Per-particle CTF estimation was performed using goCTF (*42*), in order to estimate and account for the focus gradient in each tilted micrograph. Subsequent image processing was performed using Relion 3.0. (*35*). Automatic particle picking with Laplacian-of-Gaussian filtering yielded an initial set of 1,200,794 particles. Particles were binned 3x, extracted into 192x192 pixel boxes, and subjected to two rounds of reference-free 2D classification and removal of poorly populated classes, yielding a selected set of 569,067 articles. An *ab initio* model was generated using the SGD-based approach in Relion 3.0. 3D auto-refinement was performed against the selected set of particles, yielding a reconstruction with a global resolution that reached Nyquist frequency at 5.9 Å as determined from gold-standard Fourier shell correlation. Reference-free 2D classification and 3D auto-refinement also were performed with un-binned particles; this yielded a higher-resolution reconstruction, but did not yield clear improvement in subclass map quality. Therefore, subsequent steps were performed using subclass maps from 2D classification and 3D auto-refinement with binned data.

In the next step of data processing, particle subtraction was conducted to remove the signal from the ribosome 70S subunit, in order to characterize conformational states focused on the TEC. Because a mask accounting for the full range of conformational dynamics of the TEC was essential for the analysis, and because the segmented volume of the TEC was not reliable for generating such a mask (due to lower occupancy and higher conformational dynamics of the TEC as compared to the ribosome, and resulting weaker density for the TEC as compared to the ribosome in the map generated by 3D auto-refinement), a 3D volume was generated from signal-subtracted particles using the command relion_reconstruct, and this 3D volume was low-pass filtered to 60 Å resolution to generate a mask with an extended binary map in 9 pixels and a soft-edge in 3 pixels. The same 3D volume was low-pass filtered to 40 Å resolution and also used as an initial model for subsequent data processing. Focused 3D classifications without alignment was performed to sort different conformational states of the TEC relative to the ribosome. Of six subclasses obtained, five subclasses exhibited strong, well-defined density assignable to the TEC: TTC-B1 (2%), TTC-B2 (3%), TTC-B3 (6%), TTC-C1 (10%), and TTC-C2 (12%). The three TTC-B subclasses were grouped together and subjected to an additional round of 3D classification into three subclasses without alignment. All well-defined subclasses of RNAP were selected (Figs S5A), and non-signal-subtracted particles associated with each subclass were used to reconstruct a 3D volume using the command relion_reconstruct, each volume was low-pass filtered to 30 Å resolution to generate an adaptive mask for subclass particles, with an extended binary map in 3 pixels and a soft-edge in 3 pixels. Local focused 3D refinement for each subclass was then performed with its adaptive mask using command relion_refine (-sigma_ang = 3). This process yielded reconstructions at 7.6 Å resolution for subclass B1, at 7.0 Å resolution for subclasses B2 and B3, and at 10 Å resolution for subclasses C1 and C2, as determined from gold-standard Fourier shell correlation (FSC; Figs S5D; Table S2). Half maps

were used to generate local-resolution maps using command relion_postprocess (-adhoc_bfac = 100) (Fig. S5E).

Atomic models for TTC-B1, TTC-B2, TTC-B3, TTC-C1, and TTC-C2 (n = 9; without CHAPSO) were built as described in the preceding two sections. The final atomic models for TTC-B1, TTC-B2, TTC-B3, TTC-C1, and TTC-C2 (n = 9; without CHAPSO) have been deposited in the Electron Microscopy Data Bank (EMDB) and the PDB with accession codes EMD 21486 and PDB 6VZ7, EMD 21472 and PDB 6VYU, EMD 21474 and PDB6VYW, EMD-21482 and PDB 6VZ2, and EMD 21483 and PDB 6VZ3, respectively (Table S2).

Cryo-EM structures for TTC-B and TTC-C (n = 8; without CHAPSO) were determined in the same manner, yielding maps and atomic models with resolutions of 7.0 Å for TTC-B2, 7.0 Å for TTC-B3, 9.9 Å for TTC-B6, 10 Å for TTC-C2, 8.9 Å for TTC-C3, respectively (Table S2).

Supplementary References:

26. G. Belogurov, M. Vassylyeva, V. Svetlov, S. Klyuyev, N. Grishin, D. Vassylyev, I. Artsimovitch, Structural basis for converting a general transcription factor into an operon-specific virulence regulator. *Mol. Biosyst*. **26**, 117-129 (2007).

27. V. Molodtsov, E. Sineva, L. Zhang, X. Huang, M. Cashel, S. Ades, K. Murakami, Allosteric effector ppGpp potentiates the inhibition of transcript initiation by DksA. *Mol. Cell* **69**, 828 (2018).

28. I. Artsimovitch, R. Landick, Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. *Proc. Natl. Acad. Sci. USA* **97**, 7090-7095 (2000).

29. J. Ederth, C. Mandava, S. Dasgupta, S. Sanyal, A single-step method for purification of active His-tagged ribosomes from a genetically engineered *Escherichia coli. Nucl. Acids. Res*. **37**, (2009).

30. H. Fan, A. Conn, P. Williams, S. Diggs, J. Hahm, H. Gamper, Y. Hou, S. O'Leary, Y. Wang, G. Blaha, Transcription-translation coupling: direct interactions of RNA polymerase with ribosomes and ribosomal subunits*. Nucl. Acids Res*. **45**, 11043-11055 (2017).

31. G. Blaha, U. Stelzl, C. Spahn, R. Agrawal, J. Frank, K. Nierhaus, Preparation of functional ribosomal complexes and effect of buffer conditions on tRNA positions observed by cryoelectron microscopy. *Meths. Enzymol.* **317**, 292-309 (2000).

32. J. Robertson, H. Paulsen, W. Wintermeyer, Pre-steady-state kinetic studies on ribosomal translocation. *Meths. in Enzymol.* **164**, 581-597 (1988).

33. S. Zheng, E. Palovcak, J. Armache, K. Verba, Y. Cheng, D. Agard, MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature Meths*. **14**, 331-333 (2017).

34. A. Rohou, N. Grigorieff, CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol*. **192**, 216-221 (2015).

35. J. Zivanov, T. Nakane, B. Forsberg, D. Kimanius, W. Hagen, E. Lindahl, S. Scheres, New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**, (2018).

36. J. Kang, R. Mooney, Y. Nedialkov, J. Saba, T. Mishanina, I. Artsimovitch, R. Landick, S. Darst, Structural basis for transcript elongation control by NusG family universal regulators. *Cell* **173**, 1650-1662 (2018).

37. L. Jenner, N. Demeshkina, G. Yusupova, M. Yusupov, Structural aspects of messenger RNA reading frame maintenance by the ribosome. *Nature Structl. Mol. Biol.* **17**, 555-U548 (2010).

38. E. Pettersen, T. Goddard, C. Huang, G. Couch, D. Greenblatt, E. Meng, T. Ferrin, UCSF chimera - A visualization system for exploratory research and analysis. *J. Comp. Chem*. **25**, 1605-1612 (2004).

39. D. Liebschner, P. Afonine, M. Baker, G. Bunkoczi, V. Chen, T. Croll, B. Hintze, L. Hung, S. Jain, A. McCoy, N. Moriarty, R. Oeffner, B. Poon, M. Prisant, R. J. Read, J. Richardson, D. Richardson, M. Sammito, O. Sobolev, D. Stockwell, T. Terwilliger, A. Urzhumtsev, L. Videau, C. Williams, P. Adams, Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Cryst. D* **75**, 861-877 (2019).

40. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Cryst. D* **66**, 486 (2010).

41. C. Suloway, J. Pulokas, D. Fellmann, A. Cheng, F. Guerra, J. Quispe, S. Stagg, C. Potter, B. Carragher, Automated molecular microscopy: the new Leginon system. *J. Struct. Biol*. **151**, 41-60 (2005).

42. M. Su, goCTF: Geometrically optimized CTF determination for single-particle cryo-EM. *J. Struct. Biol.* **205**, 22-29 (2019).

Supplementary Figure Legends:

Fig. S1. Structure determination: TTC-A (TTC-A1; n = 4, 5, 6, or 7; with CHAPSO)
(A) Data processing scheme (Table S1).
(B) Representative electron micrograph and 2D class averages.
(C) Orientation distribution.
(D) Fourier-shell-correlation (FSC) plot.
(E) Cryo-EM density map colored by local resolution. View orientation as in Figs. 1B and 2A, left.
(F) Cryo-EM density maps for TTC-A1 obtained using nucleic-acid scaffolds with n = 4, 5, 6, and 7 (superimposition at left; individual cryo-EM maps and color scheme at right). View orientation as in Figs. 1A and 2A, left.
(G-I) Representative cryo-EM density (blue mesh) and fits (ribbons) for RNAP regions that interact with ribosome, for NusG-N, and for ribosomal proteins that interact with RNAP.
Data in panels A-E and G-I are for structure of TTC-A1 obtained with nucleic-acid scaffold having n = 4; similar data are obtained for structures of TTC-A1 obtained with nucleic-acid scaffolds having n = 5, 6, or 7 (panel F; Table S1).

Fig. S2. Structure determination: TTC-A in absence of NusG (n = 5; with CHAPSO)
(A) Data processing scheme (Table S1).
(B) Representative electron micrograph and 2D class averages.
(C) Orientation distribution.
(D) Fourier-shell-correlation (FSC) plot.
(E) Cryo-EM density map colored by local resolution. View orientation as in Figs. 1B and 2A, left.
(F) Cryo-EM density and fit for TTC-A1 (left subpanel; overall view at top, close up of NusG-N binding site below), and for TTC-A1 in absence of NusG (right; overall view at top, close up of NusG-N binding site below). Blue mesh, EM density; dashed black circle, absent EM density for NusG-N in absence of NusG. Other colors as in Figs. 1B and 2A.

Fig. S3. Structure determination: TTC-A (n = 5 or 6; without CHAPSO)
(A) Data processing scheme (Table S2). Two subclasses were obtained: A1 and A2.
(B) Representative electron micrograph and 2D class averages.
(C) Orientation distribution.
(D) Fourier-shell-correlation (FSC) plot.
(E) EM density maps colored by local resolution for subclasses A1 and A2. View orientation as in Figs. 1B and 2A, left.
(F) EM density and fit for subclasses A1 and A2.
(G) EM density (blue mesh) and fit (green, P-site tRNA; orange, E-site tRNA) for subclasses A1 and A2.
(H) Superimposition of EM density maps of subclasses A1 (red) and A2 (orange). View orientations as in (A).
(I) Absence of EM density for RNAP ω subunit in subclass A2. EM density, blue mesh; atomic models for RNAP β' and ribosomal protein S2, pink ribbon and red-pink ribbon, respectively; location of missing EM for ω, dashed oval; ω in TEC in absence of ribosome (PDB 6P19; *19*), white ribbon Analogous data for subclass A1 are in Fig 2E.
Data shown are for structures of TTC obtained with nucleic-acid scaffold having n = 5; similar data are obtained for structures with nucleic-acid scaffold having n = 6 (Table S2).

Fig. S4. Structure determination: TTC-A, TTC-B, and TTC-C (n = 7; without CHAPSO)
(A) Data processing scheme (Table S2). Six subclasses were obtained: TTC-A subclass A2; TTC-B subclasses B4, B5, and B6; and TCC-C subclasses C2 and C3.
(B) Representative electron micrograph and 2D class averages.
(C) Orientation distribution.
(D) Fourier-shell-correlation (FSC) plot.

(E) EM density maps colored by local resolution. View orientation as in Figs. 1B, 2A, 3A, and 4A.
(F) EM density and fit for subclasses A2, B4, B5, B6, C2, and C3.
(G) EM density (blue mesh) and fit (green, P-site tRNA; orange, E-site tRNA) for subclasses A2, B4, B5, B6, C2, and C3.

Fig. S5. Structure determination: TTC-B and TTC-C (n = 8 or 9; without CHAPSO)
(A) Data processing scheme (Table S2). Five subclasses were obtained: TTC-B subclasses B1, B2, and B3 and TTC-C subclasses C1 and C2.
(B) Representative electron micrograph and 2D class averages.
(C) Orientation distribution.
(D) Fourier-shell-correlation (FSC) plot.
(E) EM density maps colored by local resolution. View orientation as in Figs. 1B, 3A, and 4A.
(F) EM density and fit for subclasses B1, B2, B3, C1, and C2.
(G) EM density (blue mesh) and fit (green and orange for P-site tRNA and E-site tRNA) for subclasses B1, B2, B3, C1, and C2.
(H) EM density for βSI2 for subclasses B1, B2, B3, C1, and C2.
Data shown are for structures of TTC obtained with nucleic-acid scaffold having n = 9; similar data are obtained with nucleic-acid scaffold having n = 8 (Table S2).

Fig. S6. Compatibility of structures of TTC-A, TTC-B, and TTC-C with NusG bridging
Structures shown are TTC-A (left), TTC-B (center), and TTC-C (right). Black "X" indicates incompatibility with NusG bridging. NusG-N, red; ribosomal protein S10 (molecular target of NusG-C), magenta; modelled NusG-C, red with red dashed circle; measured length of shortest sterically allowed connection between NusG-N residue 104 and NusG-C residue 123, red dashed line and red distance in Å. The 22-residue linker between NusG-N residue 104 and NusG-C residue 125 can span 86 Å (22 residues x 3.9 Å/residue = 86 Å). Distances in TTC-A and TTC-C are incompatible with NusG bridging. Distance in TTC-B is compatible with NusG bridging.

Fig. S7. Compatibility of structures of TTC-A, TTC-B, and TTC-C with NusG bridging, binding of transcription-elongation factors, and formation of pause and termination RNA hairpins
Structures shown are TTC-A (left), TTC-B (center), and TTC-C (right). Black "X" indicates incompatibility.
(A) Compatibility with binding of transcription-elongation factor NusA (red; *14*). TTC-A is incompatible with binding of NusA in binding mode of *14* (steric clash of NusA NTD, S1, KH1, KH2, and AR1 with 30S subunit in TTC-A1; steric clash of NusA KH2 and AR1 with 30 S subunit in TTC-A2). TTC-B and TTC-C are compatible with binding of NusA.
(B) Compatibility with formation of bacteriophage λ N transcription antitermination complex (red; *15*). TTC-A is incompatible with formation of λ N transcription antitermination complex (steric clash of N, NusA NTD, NusA S1, NusA KH1, NusA KH2, NusB, NusE, and NusG-C with 30S subunit in TTC-A1; steric clash of N with 30S subunit in TTC-A2). TTC-B and TTC-C are compatible with formation of λ N transcription antitermination complex.
(C) Compatibility with formation of bacteriophage 21 Q transcription antitermination complex (red; *16-17*). TTC-A is incompatible with formation of 21 Q transcription antitermination complex (steric clash of Q body with 30S subunit). TTC-B and TTC-C are compatible with formation of 21 Q transcription antitermination complex.
(D) Compatibility with binding of σR2 (red; *16-10*). TTC-A is incompatible with binding of σR2 (steric clash of σR2 non-conserved region with 30S subunit). TTC-B and TTC-C are compatible with binding of σR2.
(E) Compatibility with formation of pause or termination hairpins (hairpin stem, red; hairpin loop, red dashed line; *14,20-21*). TTC-A1 is incompatible with formation of pause or termination hairpin (collision of hairpin loop of hairpin having stem ≥4 bp with ribosome in TTC-A1). TTC-A2, TTC-B, and TTC-C are compatible with formation of pause or termination hairpins.

Fig. S8. Lineage-specific sequence insertion βSI2

Sequence alignment of RNAP β subunits of Gram-negative bacteria and *Chlamydia* (top nine sequences; βSI2 present) and RNAP β subunits of Gram-positive bacteria, *Deinococcus*-*Thermus*-clade bacteria, and *Mycoplasma* (bottom seven sequences; βSI2 absent). For Gram-negative bacteria and *Chlamydia*, species are as follows: *Escherichia coli* (ECOLI), *Salmonella typhimurium* (SALTY), *Klebsiella pneumoniae* (KLEP7), *Enterobacter cloacae* (ENTCC), *Haemophilus influenzae* (HAEIN), *Neisseria gonorrhoeae* (NEIG1), *Acinetobacter baumannii* (ACIBC), *Pseudomonas aeruginosa* (PSEAE), *Chlamydia trachomatis* (CHLTR). For Gram-positive bacteria, *Deinococcus*-*Thermus*-clade bacteria, and *Mycoplasma*, species are as follows: *Staphylococcus aureus* (STAAU), *Enterococcus faecalis* (ENTFA), *Streptococcus pyogenes* (STRP1), *Mycobacterium tuberculosis* (MYCTU), *Deinococcus radiodurans* (DEIRA), *Thermus thermophilus* (THETH), and *Mycoplasma pneumoniae* (MYPN).

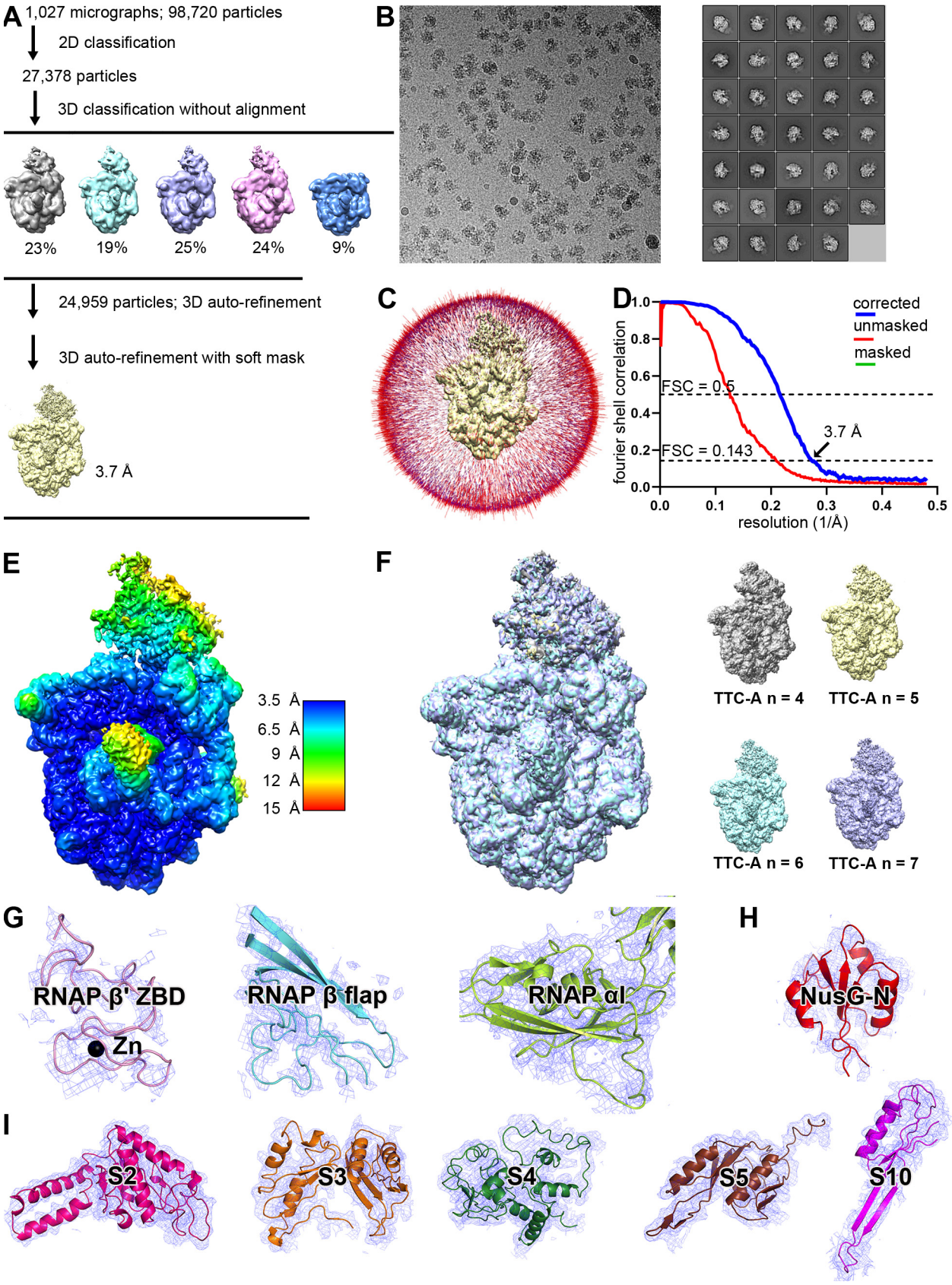
A
1,027 micrographs; 98,720 particles
2D classification
27,378 particles
3D classification without alignment
23%
19%
25%
24%
9%
24,959 particles; 3D auto-refinement
3D auto-refinement with soft mask
3.7 Å
B
C
D
corrected
unmasked
masked
FSC = 0.5
FSC = 0.143
3.7 Å
fourier shell correlation
resolution (1/Å)
E
3.5 Å
6.5 Å
9 Å
12 Å
15 Å
F
TTC-A n = 4
TTC-A n = 5
TTC-A n = 6
TTC-A n = 7
G
RNAP β' ZBD
Zn
RNAP β flap
RNAP αI
H
NusG-N
I
S2
S3
S4
S5
S10

Figure S1

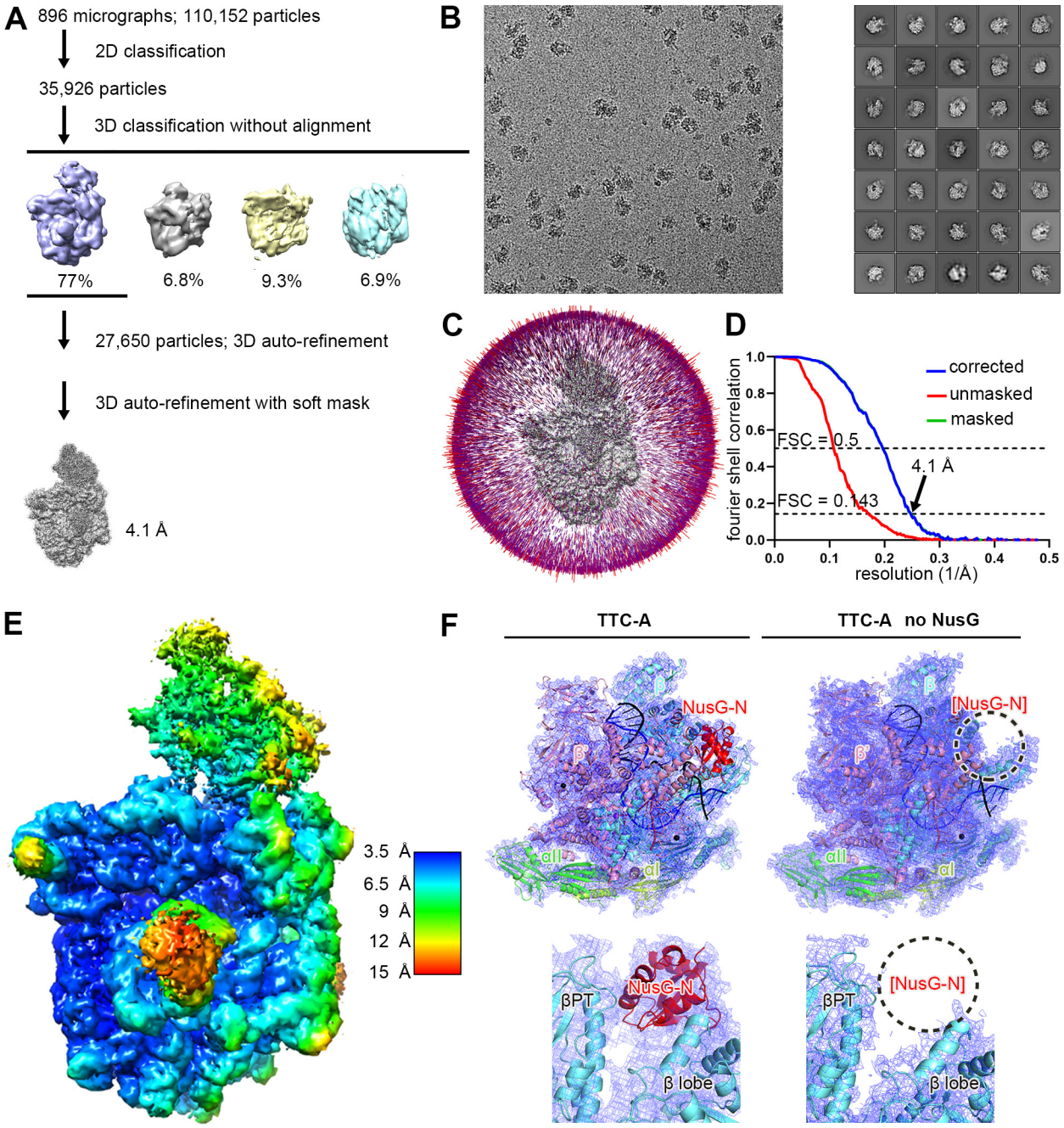
A
896 micrographs; 110,152 particles
2D classification
35,926 particles
3D classification without alignment
77%
6.8%
9.3%
6.9%
27,650 particles; 3D auto-refinement
3D auto-refinement with soft mask
4.1 Å
B
C
D
corrected
unmasked
masked
FSC = 0.5
FSC = 0.143
4.1 Å
fourier shell correlation
resolution (1/Å)
E
3.5 Å
6.5 Å
9 Å
12 Å
15 Å
F
TTC-A
TTC-A no NusG
NusG-N
[NusG-N]
βPT
β lobe

Figure S2

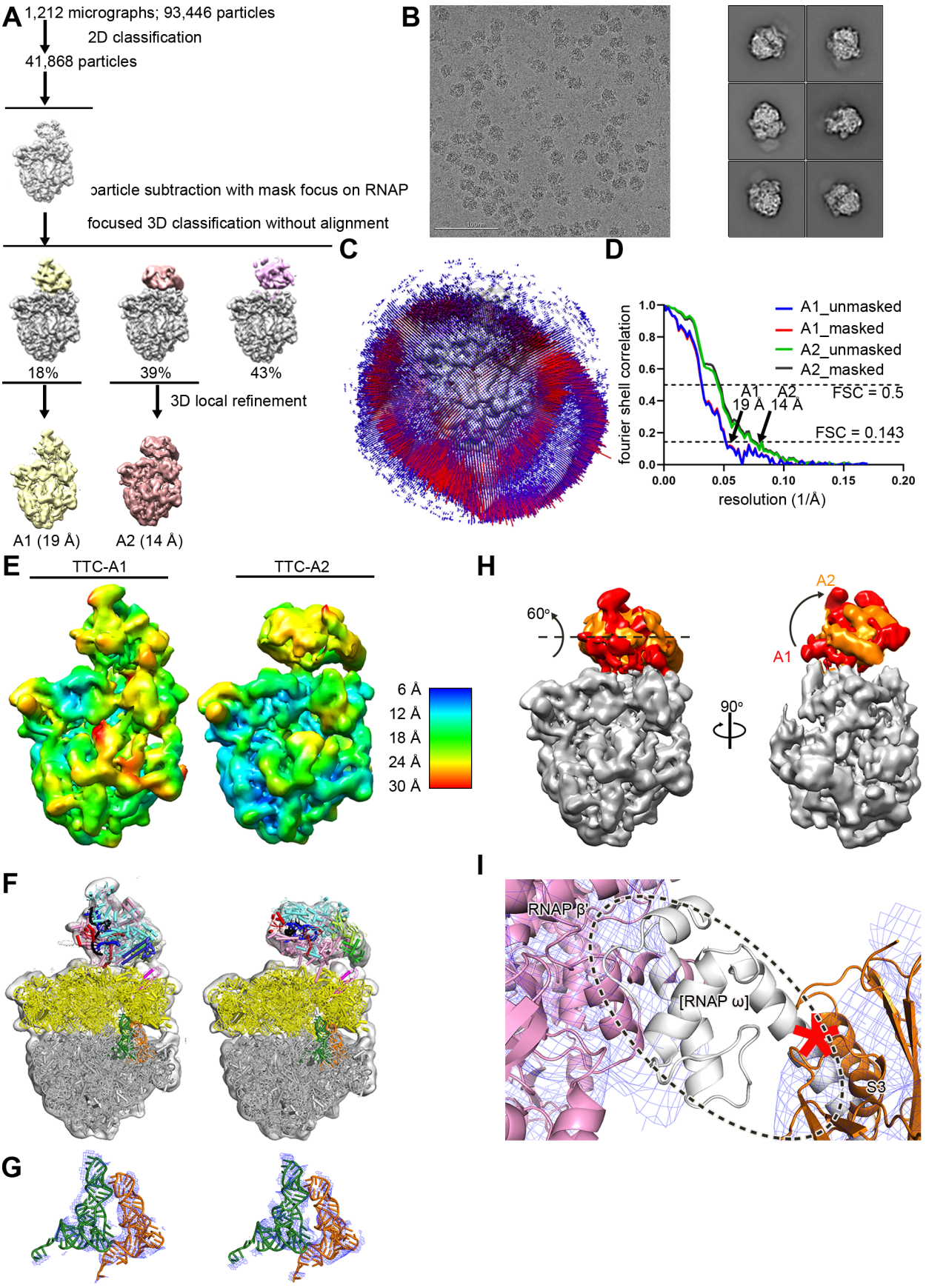
A
1,212 micrographs; 93,446 particles
2D classification
41,868 particles
particle subtraction with mask focus on RNAP
focused 3D classification without alignment
18%
39%
43%
3D local refinement
A1 (19 Å)
A2 (14 Å)
B
C
D
fourier shell correlation
resolution (1/Å)
A1_unmasked
A1_masked
A2_unmasked
A2_masked
A1
A2
19 Å
14 Å
FSC = 0.5
FSC = 0.143
E
TTC-A1
TTC-A2
6 Å
12 Å
18 Å
24 Å
30 Å
H
60°
90°
A2
A1
F
G
I
RNAP β'
[RNAP ω]
S3

Figure S3

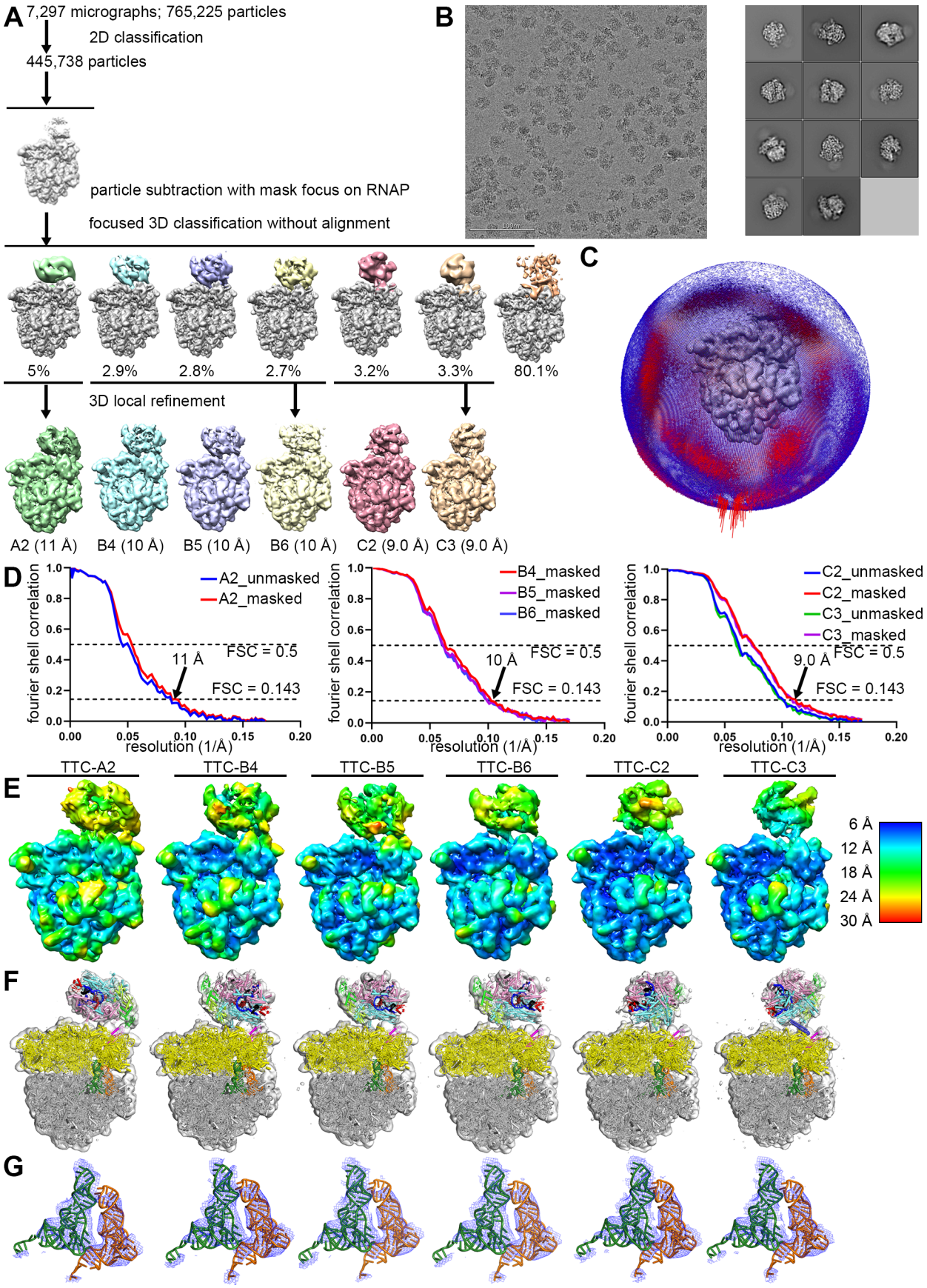
A
7,297 micrographs; 765,225 particles
2D classification
445,738 particles
particle subtraction with mask focus on RNAP
focused 3D classification without alignment
5%
2.9%
2.8%
2.7%
3.2%
3.3%
80.1%
3D local refinement
A2 (11 Å)
B4 (10 Å)
B5 (10 Å)
B6 (10 Å)
C2 (9.0 Å)
C3 (9.0 Å)
B
C
D
fourier shell correlation
resolution (1/Å)
A2_unmasked
A2_masked
11 Å
FSC = 0.5
FSC = 0.143
B4_masked
B5_masked
B6_masked
10 Å
C2_unmasked
C2_masked
C3_unmasked
C3_masked
9.0 Å
E
TTC-A2
TTC-B4
TTC-B5
TTC-B6
TTC-C2
TTC-C3
6 Å
12 Å
18 Å
24 Å
30 Å
F
G

Figure S4

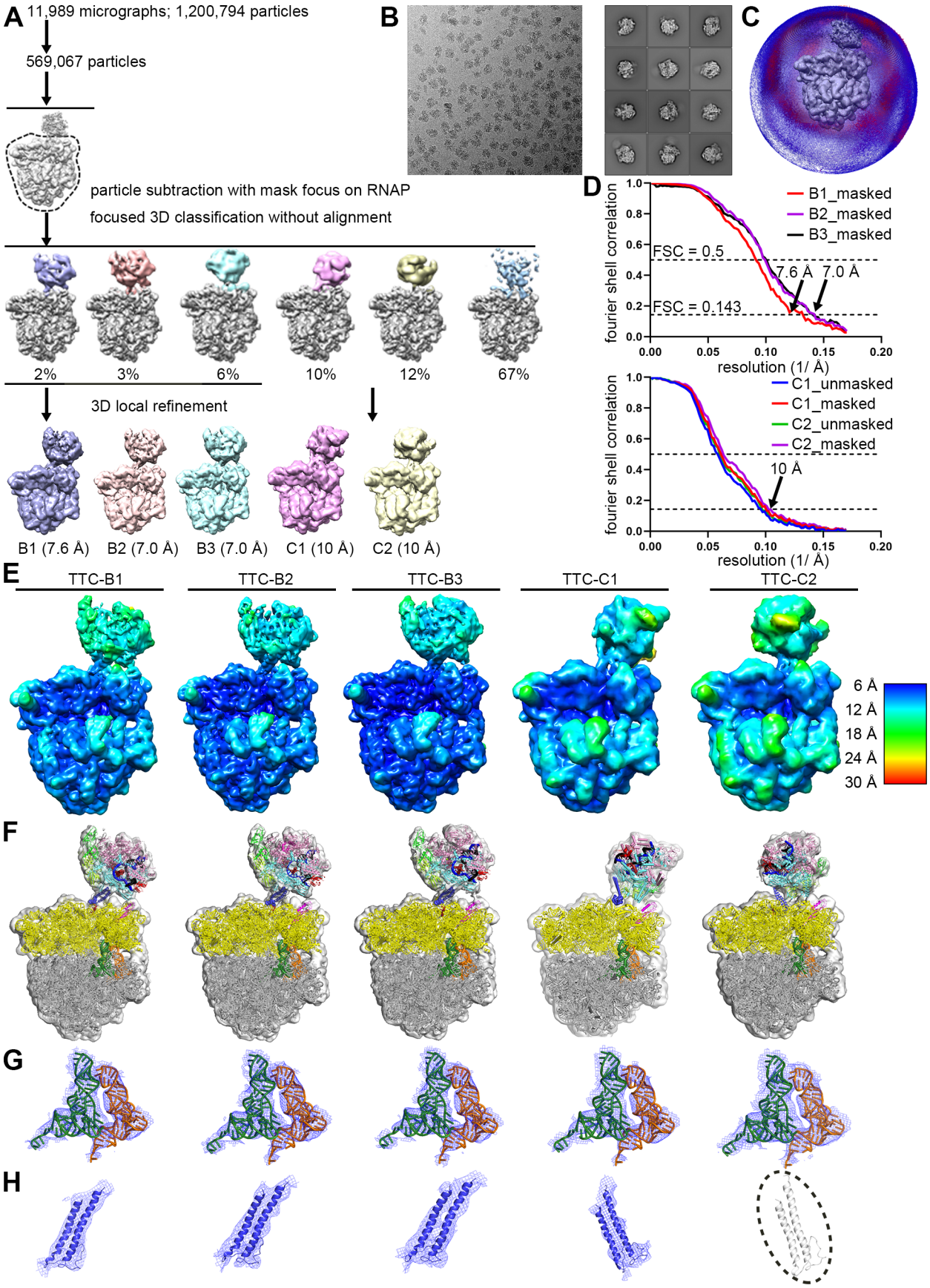
A
11,989 micrographs; 1,200,794 particles
569,067 particles
particle subtraction with mask focus on RNAP
focused 3D classification without alignment
2%
3%
6%
10%
12%
67%
3D local refinement
B1 (7.6 Å)
B2 (7.0 Å)
B3 (7.0 Å)
C1 (10 Å)
C2 (10 Å)
B
C
D
B1_masked
B2_masked
B3_masked
FSC = 0.5
FSC = 0.143
7.6 Å
7.0 Å
fourier shell correlation
resolution (1/ Å)
C1_unmasked
C1_masked
C2_unmasked
C2_masked
10 Å
E
TTC-B1
TTC-B2
TTC-B3
TTC-C1
TTC-C2
6 Å
12 Å
18 Å
24 Å
30 Å
F
G
H

Figure S5

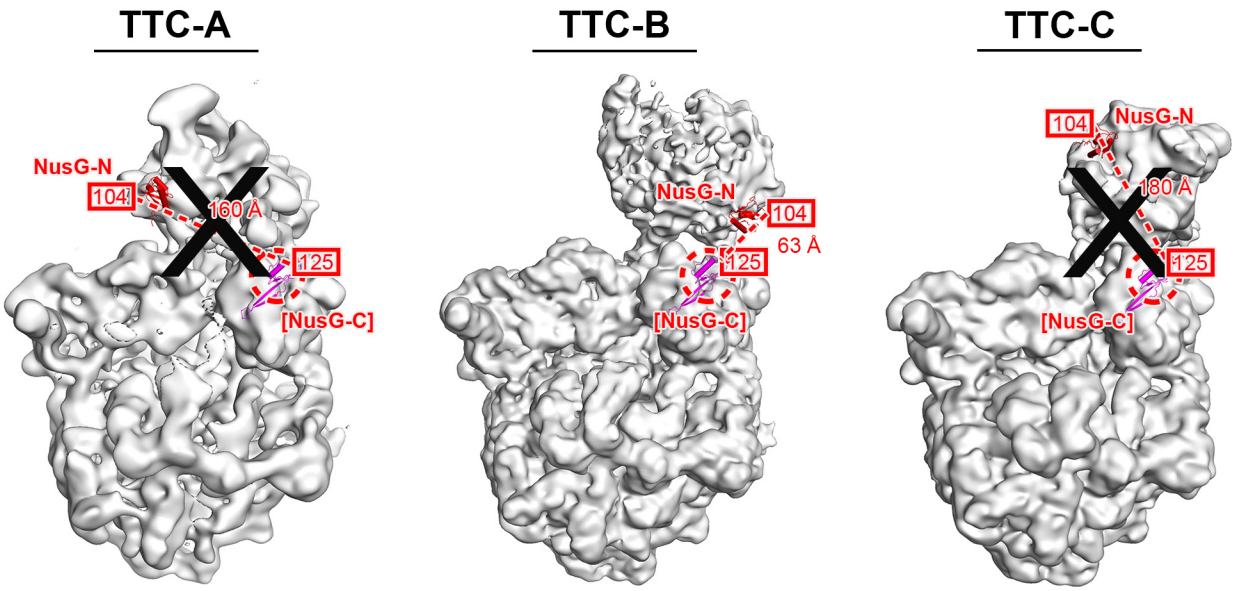
TTC-A
NusG-N
104
160 Å
125
[NusG-C]
TTC-B
NusG-N
104
63 Å
125
[NusG-C]
TTC-C
104
NusG-N
180 Å
125
[NusG-C]

Figure S6

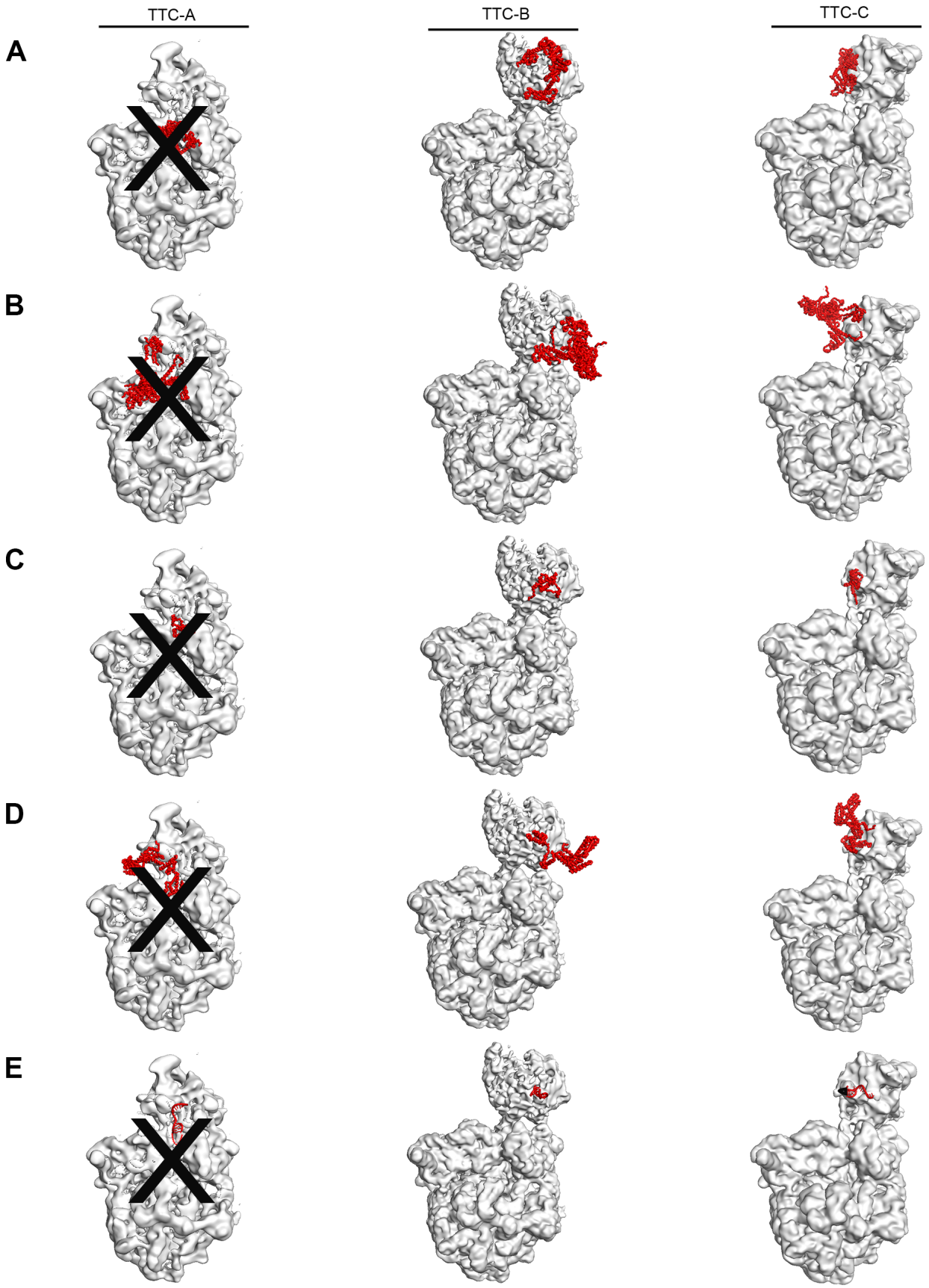

Figure S7

```
RPOB_ECOLI  GTVIDVQVFTR-DGVEKDKRALEIEEMQLKQAKKDLSEELQILEAGLFSRIRAVIVAGGVEA----------------------EKLDKLPRDRWLELGLTDEEKQNQLEQLAEQYDELKHEFEKKLEAKRRKITQGDDLAPGV
RPOB_SALTY  GTVIDVQVFTRDGVEKDKRALEIEEMQLKQAKKDLSEELQILEAGLFSRIRAVLVSSGVEAE----------------------KLDKLPRDRWLELGLTDEEKQNQLEQLAEQYDELKHEFEKKLEAKRRKITQGDDLAPGV
RPOB_KLEP7  GTVIDVQVFTRDGVEKDKRALEIEEMQLKQAKKDLSEELQILEAGLFSRIYAVLVSGGVEAE----------------------KLDKLPRDRWLELGLTDEEKQNQLEQLAEQYDELKHEFEKKLEAKRRKITQGDDLAPGV
RPOB_ENTCC  GTVIDVQVFTRDGVEKDKRALEIEEMQLKQAKKDLSEELQILEAGLFSRIYAVLVAGGVEAE----------------------KLDKLPRDRWLELGLTDEEKQNQLEQLAEQYDELKHEFEKKLEAKRRKITQGDDLAPGV
RPOB_HAEIN  GTVIDVQVFTR-DGVEKDKRALEIEEMQLREAKKDLTEELEILEAGLFARVRNLLISSGADA------------------AQLDKLD--RTKWLEQTIADEEKQNQLEQLAEQYEELRKEFEHKLEVKRKKIIKGD-DLAPGV
RPOB_NEIG1  GTVIDVQVFTREGIQRDKRAQSIILSELKRYRLDLNDQLRIFDNDAFDRIERMIVGQKANGGFMKL-TKG-------SEITTEYLAGLPSRHLWFDIRLTDEDLAKQLELIKLSLQQKREEADELYEIKKKKLTQGDELQPGV
RPOB_ACIBC  GTVIDVQVFTRDGLEKDDRALAIEKAQLDSYRKDLKEEYKIFEEAARERVIRLLKGQESNGGGSTK-R---------GDKLSEDLLSGLELVDLLEIQPADEAIAERLTQIQVFLKEKSAEIDEKFAEKKRKLATGDELTTGV
RPOB_PSEAE  GTVIDVQVFTR-DGVERDSRALSIEKMQLDQIRKDLNEEFRIVEGATFERLRAALVGAKAEGG-PALKKGT-----EITDDYLDGLE--RGQWFKLRMADDALNEQLEKAQAYISDRRQLLDDKFEDKKRKLQQGD-DLAPGV
RPOB_CHLTR  GVVMDVKVFSRKDRLSKSDDELVEEAVHLKDLQKEYKSQLAQLKVEHREKLGALLLNEKAPAAIIHRRSADILVQEGAIFDQETIELLERESLVDLLMAPCDMYDVLKDILSSYETAVQRLEVNYKTEAEHIKEGDADDLDHGV
RPOB_STAAU  GIVLDVKVFNREEGDDT-------------------------------------------------------------------------------------------------------------------------ISPGV
RPOB_ENTFA  GIVHDVKIFTREAGDE---------------------------------------------------------------------------------------------------------------------------ISPGV
RPOB_STRP1  GIVRDVKIFTRANGDE---------------------------------------------------------------------------------------------------------------------------LQSGV
RPOB_MYCTU  GKVIGIRVFSREDEDE-----------------------------------------------------------------------------------------------------------------------------IPAGV
RPOB_DEIRA  GIVVKTVRFRRGDEGV-----------------------------------------------------------------------------------------------------------------------------DLKPGV
RPOB_TTHER  GIVVRTVRLRR-GDPGV----------------------------------------------------------------------------------------------------------------------------ELKPGV
RPOB_MYCPN  GIVSCVKRFSIANGNE-----------------------------------------------------------------------------------------------------------------------------INDGV
```

Figure S8

Table S1. **Cryo-EM structures: TTC-A (n = 4, 5, 6, or 7; with CHAPSO)**

mRNA spacer	TTC class	TTC subclass	cryo-EM facility	number of particles	resolution	EMDB code	PDB code
4	TTC-A	TTC-A1	NCCAT	139,302	3.7 Å	21386	6VU3
5	TTC-A	TTC-A1	Rutgers	27,378	3.7 Å	21468	6VYQ
6	TTC-A	TTC-A1	Rutgers	24,582	3.8 Å	21469	6VYR
7	TTC-A	TTC-A1	Rutgers	29,704	3.7 Å	21470	6VYS
5	TTC-A, no NusG	TTC-A1	Rutgers	27,650	4.1 Å	21494	6VZJ

Table S2. **Cryo-EM structures: TTC-A, TTC-B, and TTC-C (n = 5, 6, 7, 8, or 9; without CHAPSO)**

mRNA spacer	TTC class	TTC subclass	cryo-EM facility	number of particles	resolution	EMDB code	PDB code
5	TTC-A	TTC-A1	U Michigan	2,898	19 Å		
5	TTC-A	TTC-A2	U Michigan	6,188	14 Å	21471	6VYT
6	TTC-A	TTC-A1	U Michigan	1,943	15 Å		
6	TTC-A	TTC-A2	U Michigan	9,631	13 Å		
7	TTC-A	TTC-A2	U Michigan	21,801	11 Å		
7	TTC-B	TTC-B4	U Michigan	6,233	9.9 Å	21475	6VYX
7	TTC-B	TTC-B5	U Michigan	6,084	9.9 Å	21476	6VYY
7	TTC-B	TTC-B6	U Michigan	5,913	9.9 Å	21477	6VYZ
7	TTC-C	TTC-C2	U Michigan	6,237	8.9 Å		
7	TTC-C	TTC-C3	U Michigan	6,440	8.9 Å	21485	6VZ5
8	TTC-B	TTC-B2	U Michigan	7,996	7.4 Å		
8	TTC-B	TTC-B3	U Michigan	4,898	9.9 Å		
8	TTC-B	TTC-B6	U Michigan	6,115	9.2 Å		
8	TTC-C	TTC-C2	U Michigan	5,622	9.7 Å		
8	TTC-C	TTC-C3	U Michigan	2,763	12 Å		
9	TTC-B	TTC-B1	U Michigan	5,979	7.6 Å	21486	6VZ7
9	TTC-B	TTC-B2	U Michigan	8,928	7.0 Å	21472	6VYU
9	TTC-C	TTC-B3	U Michigan	18,870	7.0 Å	21474	6VYW
9	TTC-C	TTC-C1	U Michigan	4,207	10 Å	21482	6VZ2
9	TTC-C	TTC-C2	U Michigan	4,960	10 Å	21483	6VZ3

Supplementary Movies:

Movie S1. Relationship between TTC-A subclasses A1 and A2
Movie shows the ~60° rotation that relates TTC-A subclasses A1 and A2. View and colors as in Fig. 1B (ribosome S50 subunit and tRNAs omitted for clarity; mRNA entering 30S subunit shown as in TTC-A1 for simplicity).

Movie S2. Relationship between TTC-A and TTC-B
Movie shows the ~90° rotation on one axis and ~180° rotation on an orthogonal axis that relate TTC-A1 to TTC-B1. View and colors as in Fig. 1B (ribosome S50 subunit and tRNAs omitted for clarity; mRNA entering 30S subunit shown as in TTC-A1 for simplicity).

Movie S3. Relationship between TTC-B subclasses B1, B2, B3, B4, B5, and B6
Movie shows the up-to-40° rotation that relates TTC-B subclasses B1, B2, B3, B4, B5, and B6. View and colors as in Fig. 1B (ribosome S50 subunit and tRNAs omitted for clarity; mRNA entering 30S subunit shown as in TTC-A1 for simplicity).

Movie S4. Relationship between TTC-A and TTC-C
Movie shows the ~45° rotation that relates TTC-A1 to TTC-C1. (ribosome S50 subunit and tRNAs omitted for clarity; mRNA entering 30S subunit shown as in TTC-A1 for simplicity).

Movie S5. Relationship between TTC-B and TTC-C
Movie shows the ~180° rotation that relates TTC-B1 to TTC-C1. (ribosome S50 subunit and tRNAs omitted for clarity; mRNA entering 30S subunit shown as in TTC-A1 for simplicity).

Movie S6. Relationship between TTC-C subclasses C1, C2, and C3
Movie shows the up-to-20° rotation that relates TTC-C subclasses C1, C2, and C3. (ribosome S50 subunit and tRNAs omitted for clarity; mRNA entering 30S subunit shown as in TTC-A1 for simplicity).