

Structural basis for λ N-dependent processive transcription antitermination

Nelly Said¹, Ferdinand Krupp^{2†}, Ekaterina Anedchenko^{1†‡}, Karine F. Santos^{1‡}, Olexandr Dybkov³, Yong-Heng Huang¹, Chung-Tien Lee^{4,5}, Bernhard Loll¹, Elmar Behrmann^{2‡}, Jörg Bürger^{2,6}, Thorsten Mielke⁵, Justus Loerke², Henning Urlaub^{4,5}, Christian M. T. Spahn², Gert Weber^{1‡} and Markus C. Wahl^{1,7*}

λ N-mediated processive antitermination constitutes a paradigmatic transcription regulatory event, during which phage protein λ N, host factors NusA, NusB, NusE and NusG, and an RNA *nut* site render elongating RNA polymerase termination-resistant. The structural basis of the process has so far remained elusive. Here we describe a crystal structure of a λ N-NusA-NusB-NusE-*nut* site complex and an electron cryo-microscopic structure of a complete transcription antitermination complex, comprising RNA polymerase, DNA, *nut* site RNA, all Nus factors and λ N, validated by crosslinking/mass spectrometry. Due to intrinsic disorder, λ N can act as a multiprotein/RNA interaction hub, which, together with *nut* site RNA, arranges NusA, NusB and NusE into a triangular complex. This complex docks via the NusA N-terminal domain and the λ N C-terminus next to the RNA exit channel on RNA polymerase. Based on the structures, comparative crosslinking analyses and structure-guided mutagenesis, we hypothesize that λ N mounts a multipronged strategy to reprogram the transcriptional machinery, which may include (1) the λ N C terminus clamping the RNA exit channel, thus stabilizing the DNA:RNA hybrid; (2) repositioning of NusA and RNAP elements, thus redirecting nascent RNA and sequestering the upstream branch of a terminator hairpin; and (3) hindering RNA engagement of termination factor ρ and/or obstructing ρ translocation on the transcript.

In all free-living organisms, transcription is carried out by multi-subunit RNA polymerases (RNAPs). Bacteria harbour a single RNAP that receives regulatory inputs from the substrate DNA, nascent RNA and protein transcription factors¹. The *Escherichia coli* RNAP core enzyme has a $\alpha_2\beta\beta'$ subunit composition and associates with a σ factor to form a holoenzyme that can initiate transcription at promoters. After promoter escape, σ is replaced by elongation factors NusA and NusG (ref. 2), which together with core RNAP form a stable transcription elongation complex (TEC). During elongation, *E. coli* NusA enhances RNAP pausing at specific sites³, while NusG increases the RNA chain elongation rate⁴. Transcription can be terminated via an intrinsic mechanism, elicited by a stable stem-loop structure followed by a stretch of uridine residues in the nascent RNA, which leads to conformational changes in RNAP, destabilization of the DNA:RNA hybrid and release of the transcript⁵. Alternatively, the ρ factor can terminate transcription by engaging nascent RNA via ρ -utilization (*rut*) sequences, translocating on the transcript in the 5'-to-3' direction and, upon encountering RNAP, extracting the transcript^{4,6}. NusA facilitates intrinsic termination⁵ and can support or counteract ρ -dependent termination depending on the context⁶. NusG can directly contact ρ via its C-terminal domain (CTD) and supports ρ -dependent termination⁴.

During lytic growth, lambdoid phages switch from immediate-early to delayed-early gene expression by processive transcription antitermination⁵. Upon transcription of N-utilization (*nut*) sequences on the λ genome, phage protein λ N, together with host N-utilization substances (Nus) A, B, E (equivalent to ribosomal (r) protein S10) and G, assemble an RNA-protein complex (RNP) on RNAP, which enables the enzyme to read through intrinsic and ρ -dependent terminators⁵. A similar mechanism, with several r-proteins replacing λ N, is at work during transcription of ribosomal RNA (*rrn*) genes⁷. The overall organization of the λ N-based transcription antitermination complex (TAC) is unknown. Here, we used X-ray crystallography, single-particle electron cryo-microscopy (cryo-EM), chemical crosslinking/mass spectrometry (CX-MS) and structure-guided mutational analyses to elucidate the functional organization of an intact λ N-TAC. Our results suggest that λ N binds and reorganizes RNAP elements, and repositions NusA on RNAP, thereby redirecting nascent RNA. Through these activities, it apparently mounts a multipronged approach to prevent termination, which might involve clamping of the RNAP RNA exit channel, sequestering the upstream branch of a terminator hairpin, and sterically interfering with ρ action.

¹Laboratory of Structural Biochemistry, Freie Universität Berlin, Takustraße 6, D-14195 Berlin, Germany. ²Medizinische Physik und Biophysik, Charité – Universitätsmedizin Berlin, Charitéplatz 1, D-10117 Berlin, Germany. ³Department of Cellular Biochemistry, Max Planck Institut für biophysikalische Chemie, Am Fassberg 11, D-37077 Göttingen, Germany. ⁴Bioanalytical Mass Spectrometry, Max Planck Institut für biophysikalische Chemie, Am Fassberg 11, D-37077 Göttingen, Germany. ⁵Institut für Klinische Chemie, Bioanalytik, Universitätsmedizin Göttingen, Robert-Koch-Straße 40, D-35075 Göttingen, Germany. ⁶Microscopy and Cryo-Electron Microscopy Group, Max-Planck-Institut für Molekulare Genetik, Ihnestr. 63-73, D-14195 Berlin, Germany. ⁷Macromolecular Crystallography, Helmholtz-Zentrum Berlin für Materialien und Energie, Albert-Einstein-Straße 15, D-12489 Berlin, Germany. [†]These authors contributed equally to this work. [‡]Present addresses: Molecular Cell Biology, Humboldt-Universität zu Berlin, Philippstr. 13, D-10099 Berlin, Germany (E.A.). moloX GmbH, Takustraße 6, D-14195 Berlin, Germany (K.F.S.). Max Planck Research Group Structural Dynamics of Proteins, Center of Advanced European Studies and Research (caesar), Ludwig-Erhard-Allee 2, D-53175 Bonn, Germany (E.B.). Molekulare Strukturbiologie, Ernst-Moritz-Arndt-Universität Greifswald, Felix-Hausdorff-Str. 4, D-17487 Greifswald, Germany (G.W.). *e-mail: mwahl@zedat.fu-berlin.de

Results

Crystal structure of a λ N–NusA–NusB–NusE–*nut* RNP. We recombinantly produced λ N and all *E. coli* Nus factors and generated a 36-nucleotide (nt) *nut* RNA containing a linear *boxA* element and a *boxB* hairpin, by *in vitro* transcription (Fig. 1a). Although a complex containing λ N, NusA or NusA^{ΔAR2} (a variant lacking 70 C-terminal residues that comprise the second of two C-terminal acidic repeats (ARs)), NusB, NusE and *nut* RNA could be reconstituted, but NusG failed to stably integrate (Fig. 2a,b). We determined the crystal structure of the λ N–NusA^{ΔAR2}–NusB–NusE–*nut* RNP at 4.0 Å resolution. Residues lacking electron density (85–107 of λ N and the first five residues of *nut* RNA) guided the design of a slightly further truncated complex, which yielded a crystal structure at 3.35 Å resolution (Fig. 1b and Supplementary Table 1).

The λ N^{1–84}–NusA^{ΔAR2}–NusB–NusE–*nut* RNP resembles a triskelion (Fig. 1b). NusA^{ΔAR2} comprises an N-terminal domain (NTD) followed by S1, hnRNP K-homology (KH) 1, KH2 and AR1 domains, and forms an arch-like structure along two arms of the triskelion, with KH1 positioned at the centre. NusA NTD and S1 as well as KH2 and AR1 are linked via long helices, while S1, KH1 and KH2 are connected by short peptides and large interfaces, as in the isolated protein⁸. The third arm of the triskelion is dominated by the NusB–NusE heterodimer, which resembles the structure of the isolated subcomplex^{9,10}. *Nut* RNA adopts an elongated conformation with *boxA* extending across NusB–NusE (Fig. 1c), the *boxA*–*boxB* spacer running along NusA KH1 and the *boxB* hairpin neighbouring NusA KH2 (Fig. 1d), in agreement with the previous implication of the KH domains in RNA binding¹¹. Although NusA S1 is a bona fide RNA-binding domain, it does not contact *nut* RNA.

λ N is intrinsically disordered in isolation¹². In the complex, its N-terminal 20 residues form a kinked helix, one face of which binds along the major groove and the loop of *boxB*, similar to the isolated subcomplex^{13,14}, while the other flank is facing NusA KH2 (Fig. 1b, region I). The 5-nt *boxB* loop adopts a GNRA (G, guanine nucleotide; N, any nucleotide; R, purine; A, adenine nucleotide) tetraloop-like conformation with Gua25 extruded. Gua25 is stacked between R8 of λ N and I318 of NusA, presenting it for hydrogen bonding to the backbone of the neighbouring GXXG motif of NusA KH2 (Fig. 1e) and thus explaining why mutations in Gua25 abolish NusA binding to a λ N–*nut* complex¹⁵. The following five residues of λ N further encircle NusA KH2, expanding the domain's β -sheet (Fig. 1b, region II). The irregularly structured λ N^{21–46} element extends across the NusA KH1 and KH2 surfaces that face away from *nut* RNA (Fig. 1b, region III). λ N^{47–52} folds as a short helix that rests in a cavity formed by the NusA NTD–S1 connector helix, NusA S1, KH1 and the globular portion of NusE ('NusA–NusE cavity'; Fig. 1b, region IV). The following irregular linker leads to a long, amphipathic helix (residues 57–79), that forms a three-helix bundle with the N-terminal helix of NusA and the NTD–S1 connector helix (Fig. 1b, region V). λ N^{34–47} has previously been shown to bind NusA AR1^{12,16,17} but is remote from AR1 in the present complex, consistent with the λ N^{34–47}–AR1 interaction being dispensable for antitermination^{16,18} and possibly only formed transiently during complex assembly (Supplementary Discussion). CX-MS of a λ N–NusA^{ΔAR2}–NusB–NusE–*nut* RNP indicated a very similar overall structure of the complex in solution as in the crystal (Fig. 1b, left and Supplementary Table 2). Crosslinks at variance with the structure are explained by the presence of a minor complex based on the alternative λ N^{34–47}–NusA AR1 interaction in solution^{12,16,17}.

Structure-guided interaction studies. Based on the structure, we probed the stability of individual interfaces and the importance of contact-mediating residues for RNP formation by analytical

size-exclusion chromatography (Fig. 2). In the absence of *nut* RNA, only binary λ N–NusA and NusB–NusE interactions prevailed (Fig. 2c), showing that the RNA is required to interconnect the two protein complexes (Fig. 1b, right). *Nut* RNA bound stably to λ N, NusB–NusE or NusA^{ΔAR2} but not to full-length NusA (Fig. 2d–g), consistent with AR2 auto-inhibiting NusA–RNA interactions by folding back onto the S1–KH1–KH2 region¹⁹. NusA joined a λ N–*nut* RNP (Fig. 2h), at least in part via NusA– λ N contacts, as NusA did not bind a NusB–NusE–*nut* RNP (Fig. 2i). NusB–NusE did not stably associate with NusA^{ΔAR2}–*nut* RNP (Fig. 2f) or λ N–*nut* RNP (Fig. 2j). Besides the *boxA*-binding site, a NusB–NusE–RNA structure revealed an unexpected binding site for *boxB*-like elements¹⁰ (Fig. 1f), which is not maintained in the presence of NusA^{ΔAR2} and λ N (Fig. 1c). Thus, NusA^{ΔAR2} or λ N sequester *boxB*, weakening the NusB–NusE–*nut* RNP and rendering NusB–NusE entry dependent on both λ N and NusA. NusE alone assembled with NusA and λ N on *nut* RNA (Fig. 2k), while NusB in the absence of NusE failed to do so (Fig. 2l), indicating that NusE is required for stable integration of NusB. The S1–KH1–KH2 region of NusA was not sufficient for formation of an intact RNP (Fig. 2m), presumably due to the lack of the NusA NTD–S1 connector helix and consequent destabilization of the NusA–NusE cavity. Residues 73–107 of λ N, which contact NusA NTD, were dispensable for RNP formation (Fig. 2n). By contrast, while still binding NusA and *nut* RNA, a λ N fragment truncated after residue 47, which lacks the region binding to the NusA–NusE cavity, was defective in higher-order RNP formation (Fig. 2o). Alanine exchanges of NusA E212, E218 and R255, which mediate binary NusA–NusE contacts, did not destabilize the complex. By contrast, an alanine exchange of R258, which is involved in a NusA–NusE–*boxA*–*boxB* spacer triple interaction (Fig. 1h), abrogated complex formation (Fig. 2p). These results expand previous interaction studies^{15,16,20} showing that while complex stability is considerably buffered with respect to the loss of individual binary contacts, it critically depends on residues (such as NusA R258) and regions (such as λ N^{47–52} or the NusA–NusE cavity) that mediate interactions between multiple other subunits.

Structural mechanisms underlying the effects of subunit variants. Our structure rationalizes λ N-related defects associated with numerous previously studied variants of complex components (Fig. 3 and Table 1). A7D, V8A, V8E, A11D, V12D and L31E exchanges in NusA show defects in the propagation of lambdoid phages and/or in λ N-mediated antitermination¹⁸. Residues A7, V8, A11 and V12 of NusA form the surface of the N-terminal helix of NusA that is involved in binding the C-terminal helix of λ N (Fig. 3b), while L31 is part of the hydrophobic core and thus crucial for the fold stability of the NusA NTD. NusA^{R199A} abrogates λ N-mediated antitermination and reduces NusA interaction with a λ N–*nut* RNP^{19,21}, in agreement with R199 stabilizing the S1–KH1 arrangement that bridges between *nut* RNA and λ N regions (Fig. 3c). Thus, in contrast to previous expectations²¹, NusA^{R199A} does not abrogate direct *nut* RNA contacts. NusA^{L183R} (encoded by the *nusA1* allele) blocks^{22,23} while NusA^{R104H/E212K} (*nusA10*) reduces λ N-dependent antitermination based on the E212K exchange²⁴. L183 in S1 interacts with the upper part of the NTD–S1 connector helix (Fig. 3d) and E212 in NusA KH1 forms part of the NusA–NusE interface (Fig. 3e). Thus, both residues help to shape the NusA–NusE cavity that accommodates λ N^{47–52}. In NusA944, four residues (153–156) in a loop of the S1 domain are replaced by nine residues found at this position in *Salmonella typhimurium* NusA²⁵. In line with NusA944 and *S. typhimurium* NusA still binding a λ N–*nut* RNP but failing to support antitermination^{21,25}, our structure reveals that the insertion will alter NusA– λ N contacts (Fig. 3f), but again does not

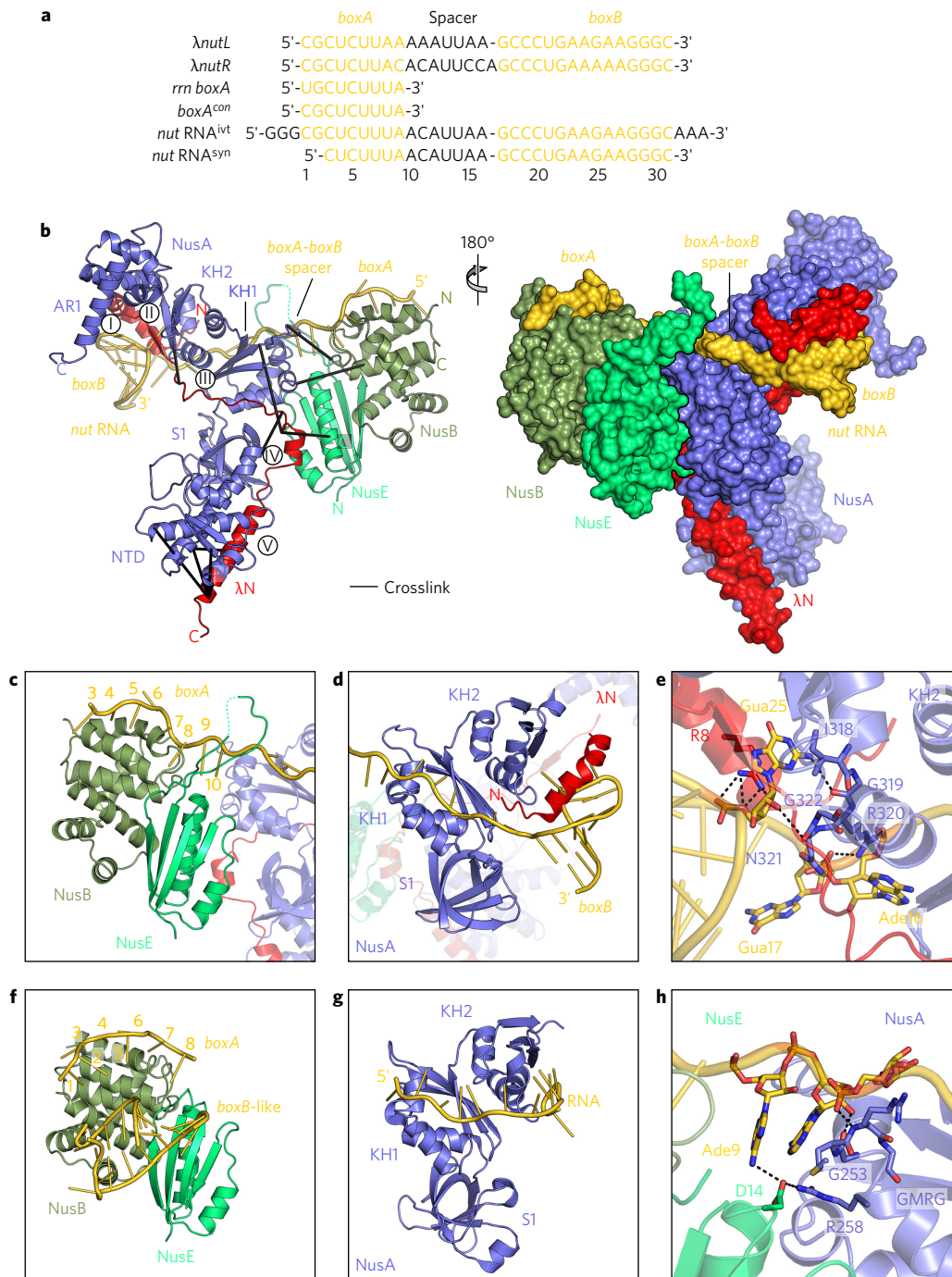


Figure 1 | Structure of a λ N¹⁻⁸⁴-NusA^{ΔAR2}-NusB-NusE-*nut* RNP. **a**, λ , ribosomal (r) RNA (rrn) and consensus (con) *nut* RNA sequences and RNAs used for crystallization (ivt, *in vitro* transcribed; syn, chemically synthesized). **b**, Diametric cartoon and surface representations of the λ N¹⁻⁸⁴-NusA^{ΔAR2}-NusB-NusE-*nut* RNP. N, N termini; C, C termini. Black lines in cartoon view, intermolecular crosslinks observed in the λ N¹⁻⁸⁴-NusA^{ΔAR2}-NusB-NusE-*nut* RNP in solution. Roman numerals indicate λ N-NusA interaction regions. Black line in surface view, border between the NusB-NusE-*boxA* and λ N¹⁻⁸⁴-NusA^{ΔAR2}-*boxB* half-RNPs. **c**, NusB-NusE-*boxA* unit of the λ N¹⁻⁸⁴-NusA^{ΔAR2}-NusB-NusE-*nut* RNP. Nucleotide numbering according to **a**. Orientation as in **b**, right. **d**, λ N¹⁻²⁰-NusA^{51-KH1-KH2}-*boxB*-spacer unit of the λ N¹⁻⁸⁴-NusA^{ΔAR2}-NusB-NusE-*nut* RNP. Rotated 30° about the horizontal axis (top to front) compared to **b**, right. **e**, Interaction of λ N and NusA with the extruded nucleobase Gua25, sandwiched between λ N R8, NusA I318 and the following GXXG of NusA KH2. R320 of the GXXG motif additionally binds the backbone of Ade16-Gua17. Orientation as in **b**, left. **f**, Comparison of the structure of an isolated NusB-NusE-RNA complex (PDB ID: 3R2C). **g**, Comparison of the structure of an isolated NusA^{51-KH1-KH2}-RNA complex (PDB ID: 2ATW). **h**, Close-up view of the triple interaction between NusA KH1, NusE and the *boxA*-*boxB* spacer. Rotated 10° clockwise about the vertical axis compared to **b**, right. GMRG, glycine-methionine-arginine-glycine motif of the NusA KH1 domain.

directly affect NusA-*nut* RNA interactions. NusA^{G253D} fails to bind a λ N-*nut* RNP or to support λ N-dependent antitermination¹¹, as explained by G253 being part of the KH1 GXXG motif that binds

the *boxA*-*boxB* spacer and R258 following this motif engaging in interactions with NusE (Fig. 1h). Consistent with NusA^{G319D} similarly abrogating NusA interaction with a λ N-*nut* RNP and

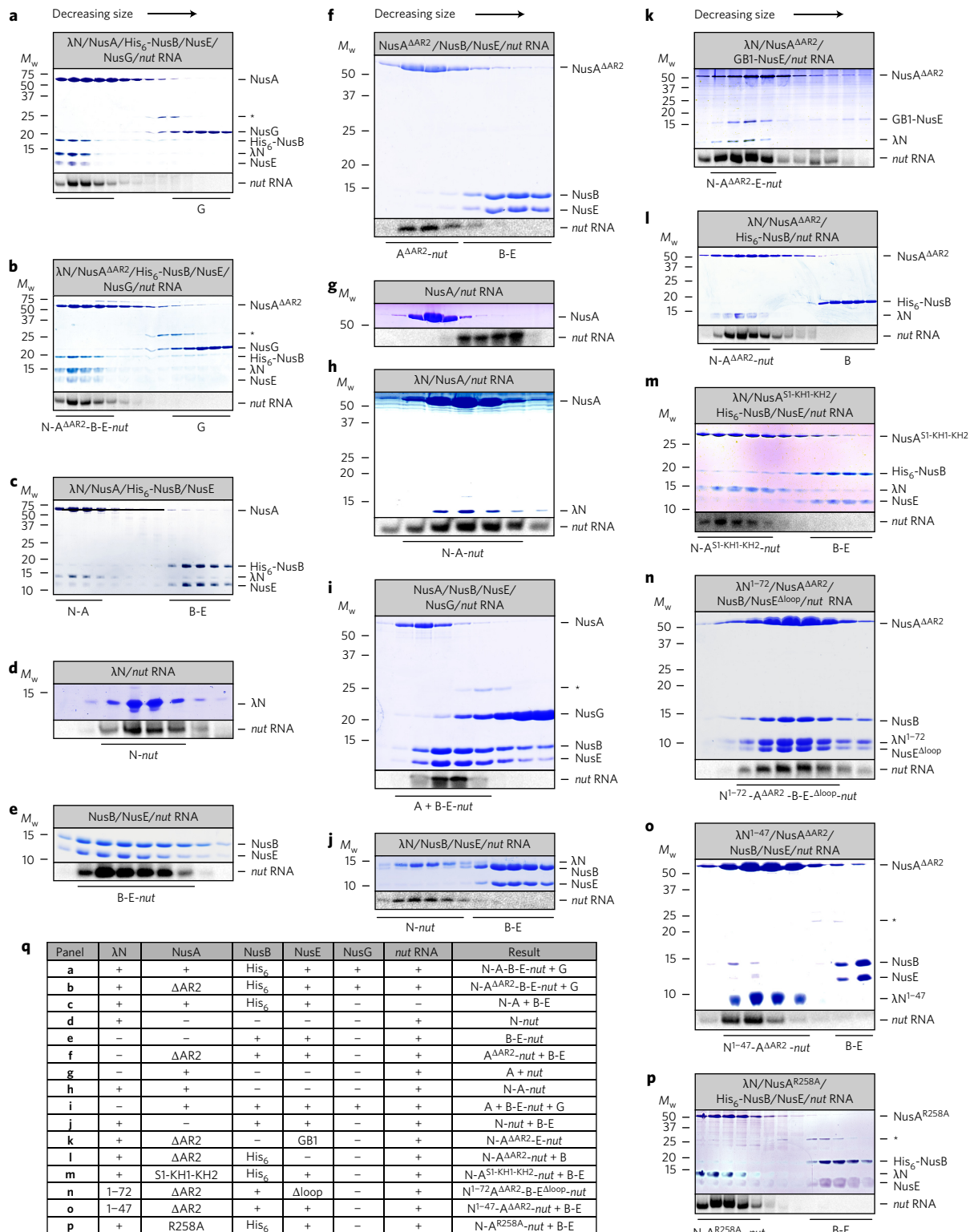


Figure 2 | Interaction studies. **a–p**, Coomassie-stained SDS-PAGE (proteins, top) and ethidium bromide-stained urea PAGE (RNA, bottom) analyses of analytical size-exclusion chromatography runs, monitoring interactions among λN, Nus factors and *nut* RNA. Mixtures that were loaded for each run are indicated in the boxes above the gels. Molecular weights are shown on the left (M_w). Bands are identified on the right, and complexes formed are indicated below the gels. N, λN; A, NusA; B, NusB; E, NusE; G, NusG; *nut*, *nut* RNA; NusE^{Δloop}/E^{Δloop}, NusE variant with the long protruding loop deleted, which has no known functional consequences for λN-dependent antitermination⁹. *unidentified contaminants. **q**, Overview of the components mixed in the experiments shown in **a–p** and the results obtained. Gels are representative examples of at least two repetitions.

exhibiting a mild defect in supporting λN action¹¹, G319 belongs to the KH2 GXXG motif involved in binding the *boxB* loop (Fig. 1e). NusE^{A86D} (*nusE71*) blocks λN-dependent antitermination²⁶,

reconciled by A86 being located in the NusE–λN interface, where an aspartate would be forced into unfavourable interactions with the neighbouring L52 of λN (Fig. 3g). NusB^{D118N} (*nusB101*)

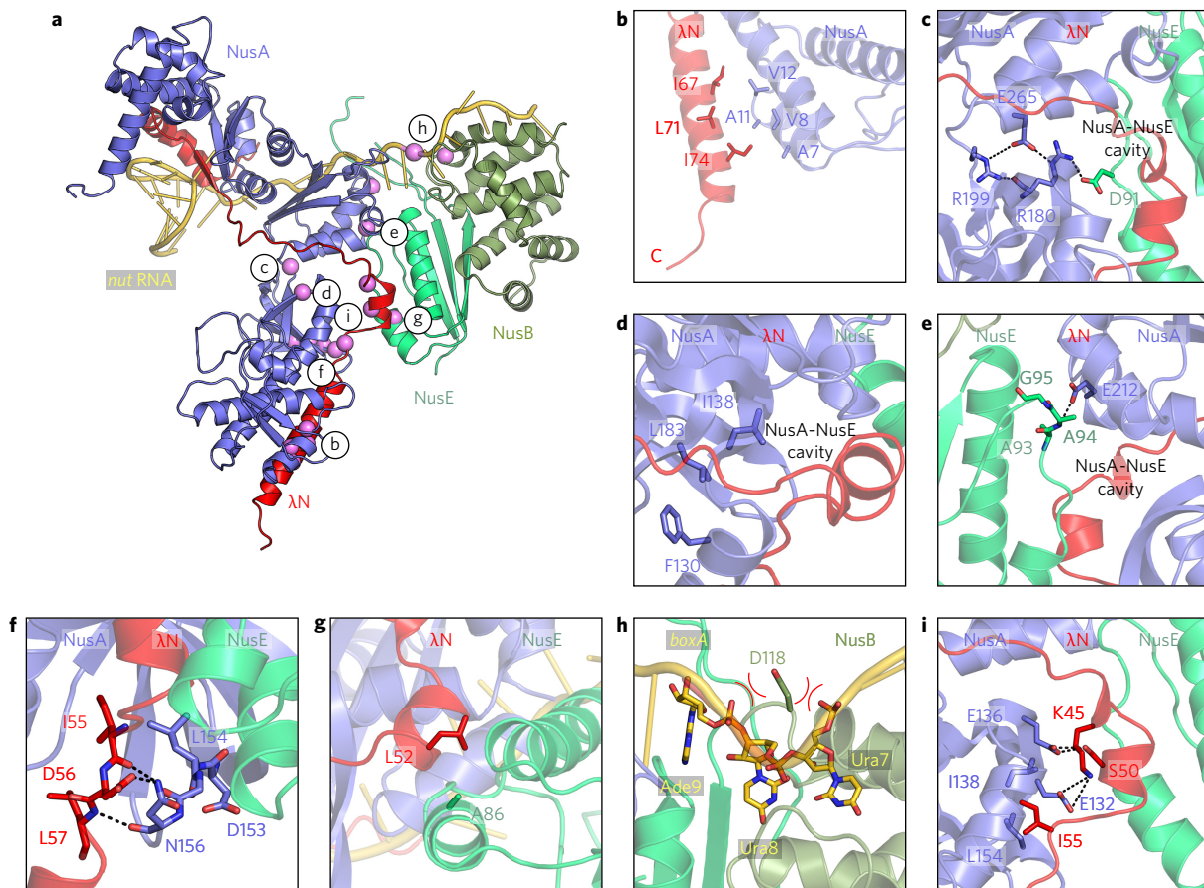


Figure 3 | Mapping of subunit variants. **a**, Overview of previously investigated variants of subunits, with affected residues mapped to the λN^{1-84} -NusA^{ΔAR2}-NusB-NusE-nut RNP structure. Magenta spheres, Cα or phosphorus atoms of affected protein or RNA residues, respectively. Letters refer to close-up views in the following panels. **b**, Close-up view of the region affected by the A7D, V8A, V8E, A11D and V12D exchanges in NusA. Rotated 90° anticlockwise about the vertical axis relative to **a**. **c**, Close-up view of the region affected by the NusA^{R199A} variant. Rotated 20° anticlockwise about the vertical axis relative to **a**. **d**, Close-up view of the region affected by the nusA1 (L183R) allele. Rotated 60° about the horizontal axis (top to front) relative to **a**. **e**, Close-up view of the region affected by the nusA10 (E212K and R104H) allele. Rotated 150° anticlockwise about the vertical axis relative to **a**. **f**, Close-up view of the region affected by an insertion via the nusA944 allele (NusA residues 153-156). Rotated 60° clockwise about the vertical axis relative to **a**. **g**, Close-up view of region affected by the nusE71 (A86D) allele. NusB cut away. Rotated 30° clockwise about the vertical axis and 60° about the horizontal axis (top to back) relative to **a**. **h**, Close-up view of the region affected by the nusB101 (D118N) allele. Red double arcs, repulsive electrostatic effects. Rotated 20° clockwise about the viewing axis and 40° about the horizontal axis (top to front) relative to **a**. **i**, Close-up view of the region affected by the pun alleles, encoding K45R, S50R, I55M exchanges in λN . The orientation is the same as in **a**.

exhibits enhanced affinity to *boxA*⁹ and suppresses *nusA1* and *nusE71* defects²⁷. These findings can be rationalized by D118 neighbouring the negatively charged sugar-phosphate backbone (Fig. 3h), which will decrease the NusB-*boxA* affinity compared to an asparagine at this position. K45R, S50R and I55M in λN (*pun134*, *pun150* and *pun165*, respectively) suppress *nusA1* defects^{28,29}. λN^{K45R} probably enhances contacts to NusA E132, λN^{S50R} presumably introduces additional salt bridges to NusA E132 and/or E136 and λN^{I55M} might sandwich more stably between the NTD-S1 connector helix and the S1 domain (Fig. 3i). Finally, consistent with the identity of the nucleotide at *boxA* position 9 modulating λN -mediated antitermination^{30,31}, Ade9 in our structures mediates the only nucleobase contact of *nusA* RNA to NusE (Fig. 1h). Taken together, the effects of these variants in light of the present structure emphasize the importance of a precise relative orientation of the NusA S1-KH1-KH2 domains, around which *nusA* RNA, λN and NusE are organized and point to the NusA-NusE- λN interaction region as a crucial node in the complex.

Structure of a λN -based TAC. To further investigate the structural basis of λN -mediated antitermination, we assembled a complete

TAC comprising RNAP, double-stranded DNA bearing a transcription bubble, *nusA* RNA that can hybridize with the template strand in the transcription bubble, all Nus factors and λN (Fig. 4a,b). We determined a cryo-EM map of the λN -TAC at 9.8 Å resolution, in which all subunits could be located (Fig. 4c,d and Supplementary Fig. 1). CX-MS analysis of the TAC supported the structural assignment (Fig. 4e and Supplementary Table 3). In the TAC, the λN -NusA-NusB-NusE-*nusA* RNA towers above the RNA exit channel formed by the β flap domain (including the flexible β flap tip helix; FTH), C-terminal residues of β, the β' dock domain and the β' zinc-finger (ZnF; Fig. 4d-f). The NusG NTD is located next to the RNAP primary channel, contacting the β' clamp helices and spanning one wall of the upstream DNA-binding site, in agreement with the location of the NusG NTD in a recent RNAP-NusG co-crystal structure³² and of the homologous Spt5 NTD on eukaryotic RNA polymerase II (ref. 33); however, NusG and Spt5 NTDs have been modelled in different orientations in these previous studies, which cannot be distinguished at the present resolution of the λN -TAC. The NusG C-terminal domain (CTD) is detached from the site seen in an RNAP-NusG crystal structure³². Instead, it is bound to NusE, as

Table 1 | List of subunit variants and their effects.

Factor	Mutation	Allele	Effect	Figure	Ref.
NusA	A7D	-	Defective phage propagation/antitermination	3b	18
	V8A	-	Defective phage propagation/antitermination	3b	
	V8E	-	Defective phage propagation/antitermination	3b	
	A11D	-	Defective phage propagation/antitermination	3b	
	V12D	-	Defective phage propagation/antitermination	3b	
	L31E	-	Defective phage propagation/antitermination	-	
	R199A	-	Abrogation of λ N antitermination, reduced binding to λ N- <i>nut</i> RNP	3c	19,21
	L183R	<i>nusA1</i>	Block of λ N antitermination	3d	22,23
	R104H/E212K	<i>nusA10</i>	Reduction of AN antitermination	3e	24
	944	<i>nusA944</i>	Binding to λ N- <i>nut</i> RNP, failure to support antitermination	3f	21,25
	G253D	-	No binding to λ N- <i>nut</i> RNP, failure to support antitermination	1h	11
	G319D	-	No binding to λ N- <i>nut</i> RNP, mild defect in supporting λ N action	1e	
NusE	A86D	<i>nusE71</i>	Block of λ N but not of <i>rm</i> antitermination	3g	26
NusB	D118N	<i>nusB101</i>	Enhanced affinity to <i>boxA</i> , suppression of <i>nusA1</i> and <i>nusE71</i> defects	3h	9,27
	K45R	<i>pun134</i>	Suppression of <i>nusA1</i> defects	3i	28,29
λ N	S50R	<i>pun150</i>	Suppression of <i>nusA1</i> defects	3i	
	I55M	<i>pun165</i>	Suppression of <i>nusA1</i> defects	3i	
<i>nut</i>	<i>boxA Ade9Uri</i>	-	Abrogation of λ N antitermination	1h	30,31

previously characterized in isolation³⁴, and also contacts NusA S1–KH1, stabilizing the λ N-binding NusA–NusE cavity (Fig. 4d). In this arrangement, additional contacts of NusG^{S163F} (*nusG4*) to NusA could ensue and counteract *nusA1*- and *nusE71*-mediated destabilization of the NusA–NusE– λ N interaction network³⁵.

The above analyses suggest strategies by which λ N modulates termination. First, as indicated by λ N crosslinks to β flap residues 844, 937 and 1041, to β flap tip residues 890 and 914, to β residues 496 and 503 (preceding the β pincer), to β residue 1242 (neighbouring the β' rudder) and to β' dock residues 371, 378, 395 and 399 (Supplementary Table 3), the λ N C terminus meanders around the rim of the RNA exit channel (Fig. 4f) and clamps RNAP elements that bind the DNA:RNA hybrid and are crucial for the stability of elongating complexes³⁶. RNAP clamping by the λ N C terminus would reinforce hybrid stability and hinder hairpin-induced conformational changes in RNAP, explaining how the protein could interfere with termination-related pausing, RNA slippage and hybrid melting^{37,38}. In the same manner, λ N could counteract ρ -mediated destabilization of the hybrid. In agreement with the path of the λ N C terminus, mutations in the β flap, β pincer, β' lid and β' rudder affect λ N-dependent antitermination^{39–42} and hydroxyl radicals generated from a modified C terminus of the phage H-19B N protein target the RNAP active site cavity⁴³. Furthermore, the same contact areas are also used by σ 70 region 3.2 to access the interior of RNAP⁴⁴, explaining binding competition between σ 70 and λ N¹².

Second, in a TEC lacking λ N, NusA NTD and S1 domains are thought to reside next to and guide nascent RNA along the β' dock domain^{45,46} (Fig. 5a). Conversely, in the TAC, the NusA NTD bound to the C-terminal λ N helix is moved towards the centre of the β flap domain, placing the NusA S1 domain next to the β' ZnF (Figs 4e and 5b). The β flap tip maintains interactions with the NusA NTD and is thus displaced from the β' ZnF and β' dock domain, consistent with the suggestion that λ N leads to altered NusA NTD– β flap interactions and a rearrangement of the β flap and β' dock domains³⁹. This rearrangement could allow nascent RNA 15–33 nt upstream of the 3'-end to run between the λ N C terminus, the β flap tip and the β' ZnF towards the NusA S1 domain (Figs 4a and 5b). In striking agreement with this organization, crosslinking analyses have shown that λ N directly contacts this RNA region and reinforces its interactions with the RNAP upstream RNA-binding site (UBS, composed of the β' ZnF, C-terminal residues of β and α CTD) and NusA S1 (ref. 37). The upstream branch of a terminator hairpin at the time of termination could lie in the RNA region guided along NusA S1 and thus remote

from the complementary downstream branch just emerging from the exit channel. Thus, our analysis provides a structural explanation for the suggested sequestration of the upstream terminator branch by λ N, RNAP UBS and NusA S1 to prevent termination³⁷ (Fig. 5b).

Structure-guided functional analyses. To test the suggested λ N-mediated repositioning of NusA, we analysed a TEC lacking λ N by CX-MS (Supplementary Table 4). In agreement with our model, crosslinks of the NusA NTD and S1 domain to the β flap, β' ZnF, β' dock and UBS regions were strongly altered in the absence of λ N. Only in the absence of λ N did the NusA S1 domain crosslink to the β' dock domain and the C terminus of β (Fig. 4e), fully consistent with its presumed positioning next to these elements in the unmodified TEC. We also determined a crystal structure of a large NusA fragment (NusA^{100–426}) in isolation (Supplementary Table 1), revealing a flexible relative orientation between the NTD and the remainder of NusA (Fig. 5b, inset). Thus, λ N repositions NusA on RNAP, possibly capitalizing on an intrinsic flexibility of NusA.

To further test the suggested antitermination mechanisms, we conducted *in vitro* transcription assays with rationally engineered components (Fig. 6a,b). According to our model, λ N^{1–84} is expected to still maintain contacts to the core of the β flap domain, reposition NusA NTD and thus support sequestration of the upstream terminator branch, but lacks the ability to suppress pausing and hybrid melting by RNAP clamping. The fragment retained 83% of the antitermination efficiency of full-length λ N (Fig. 6c, lanes 2 and 3), in good agreement with the estimated ~10% of the antitermination activity of λ N in the presence of NusA relying on pause suppression³⁷. λ N^{1–72} lacks all RNAP-contacting residues of λ N. Although the fragment still assembled with Nus factors and *nut* RNA on RNAP, the resulting complex lacked antitermination activity (Fig. 6c, lanes 2 and 4). Removal of the N-terminal helix of NusA (NusA^{17–426}) or further deletions up to the NTD–S1 connector helix (NusA^{41–426} and NusA^{136–426}) should leave RNAP clamping by the λ N C terminus unaffected but partly interfere with sequestration of the upstream terminator branch by hampering rearrangements of the RNAP β flap tip and proper positioning of NusA S1. Consistently, these variants exhibited ~85% of the antitermination efficiency of full-length NusA (Fig. 6d).

Discussion

Here, we have elucidated the structure of a λ N^{1–84}–NusA ^{Δ AR2}–NusB–NusE–*nut* RNP and the global architecture of a complete λ N-TAC. Interaction studies, mapping of subunit variants,

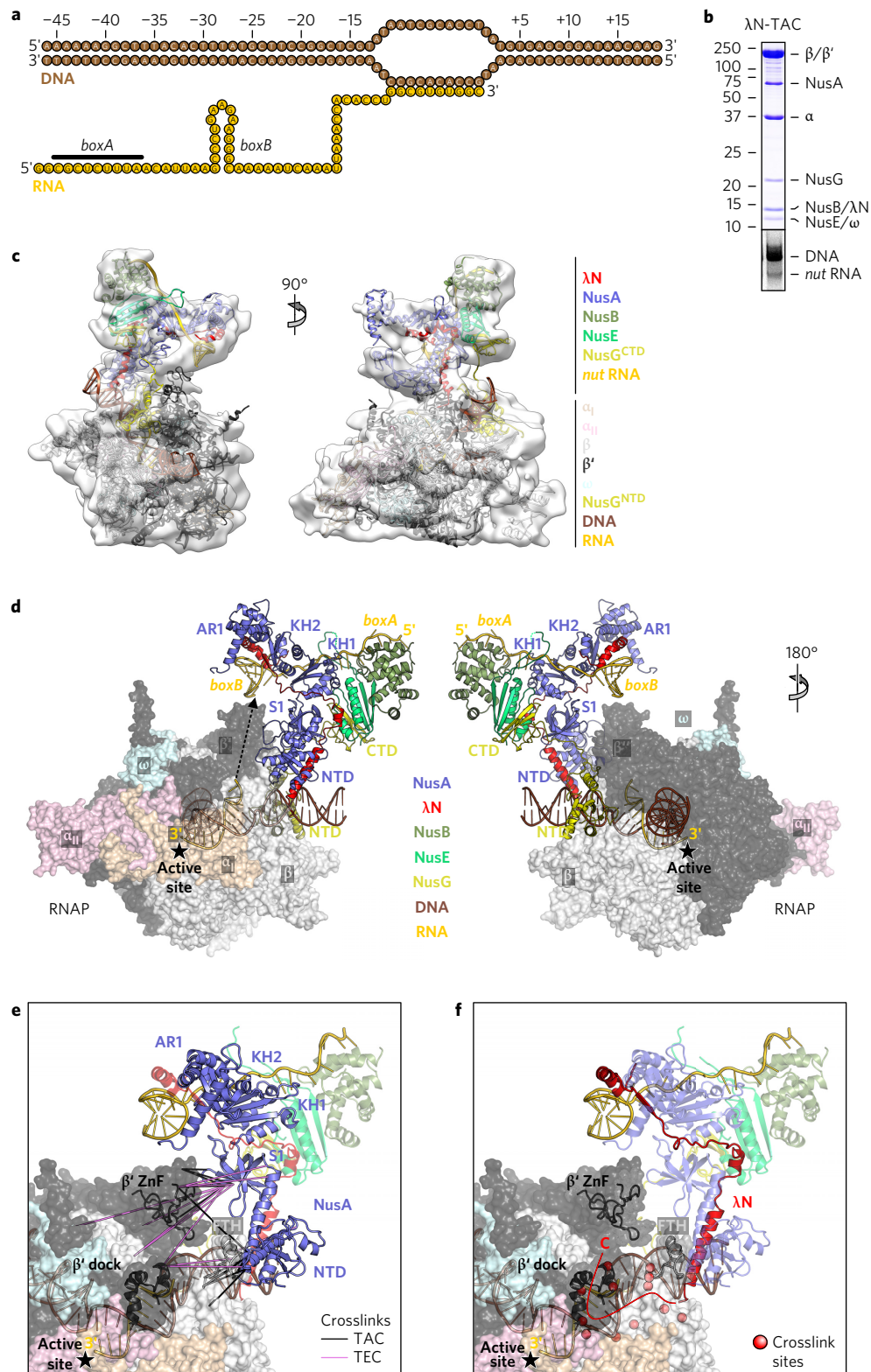


Figure 4 | Structure of a λ N-based transcription antitermination complex. **a**, Scheme of nucleic acids used for transcription antitermination complex (TAC) assembly. **b**, Analysis of the assembled TAC. Top: Coomassie-stained SDS-PAGE (proteins). Bottom: Ethidium bromide-stained urea PAGE (nucleic acids). Molecular weights are shown on the left (M_w). Gels shown are representative examples of three repetitions. **c**, Fitting of the TAC components to the cryo-EM map. At the present resolution, the positions of the NusA AR2 domain and of the α -CTDs could not be modelled reliably and have thus been omitted from the model. **d**, Structural model of the λ N-based TAC. Star, active site. Dashed arrow, connection between RNA emerging from the exit channel and boxB. Orientations of the λ N-NusA-NusB-NusE-nut RNP as in Fig. 1b, left. **e**, Crosslinks between NusA and RNAP subunits in the TAC (black lines) and a TEC lacking λ N (violet lines). Elements of the RNA exit channel are shown as cartoons. Rotated 30° anticlockwise about the vertical axis and 30° about the horizontal axis (top to front) relative to **d**, left. **f**, Crosslinks between λ N C-terminal residues absent in the crystal structure and RNAP, indicating the path of the λ N C terminus around the RNA exit channel (red line). Red spheres, crosslink sites on RNAP. Orientation as in **e**.

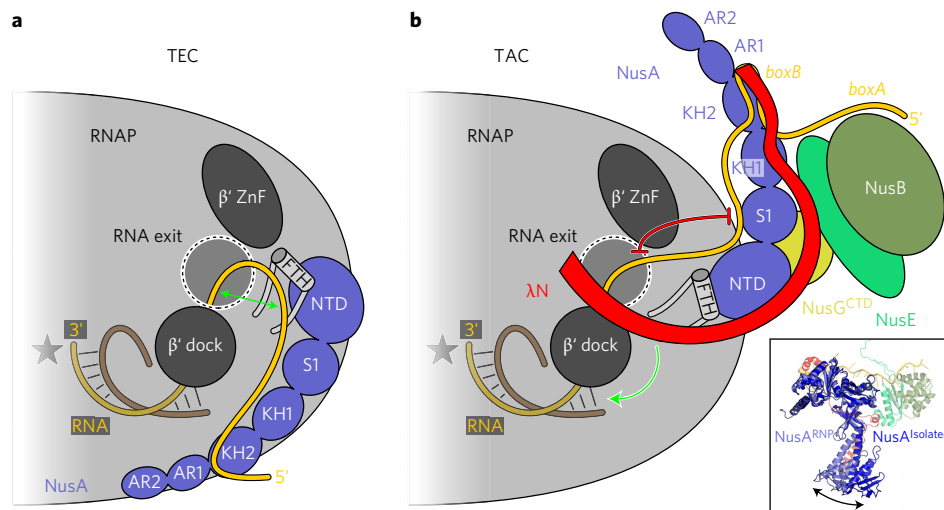


Figure 5 | Mechanism of λ N-mediated processive antitermination. a, Model for NusA activity in a TEC lacking λ N. Nascent RNA is threaded between the RNAP β' dock domain and NusA. The NusA NTD can present the upstream branch of a terminator hairpin for base-pairing with the downstream branch emerging from the RNA exit channel (green double arrow), thus supporting intrinsic termination. **b**, Organization of the λ N-based TAC. Repositioning of NusA and associated factors by λ N relocates the NusA S1 domain and displaces the β flap tip, thus opening a gate for the nascent RNA to approach NusA S1. The latter is not blocked by *nut* RNA and can thus sequester the upstream branch of a terminator hairpin during emergence of the downstream branch from the RNA exit channel (red symbol). Additionally, the λ N C terminus clamps RNAP elements around the RNA exit channel, thus indirectly stabilizing the DNA:RNA hybrid (green arrow) and preventing RNA slippage/hybrid melting. Inset: Structural comparison of NusA in the λ N¹⁻⁸⁴-NusA^{AR2}-NusB-NusE-*nut* RNP and isolated NusA¹⁰⁰⁻⁴²⁶ with an NTD modelled according to the orientation of the NTD-S1 connector helix (dark blue) after superposition according to the S1-KH1-KH2 regions. Double arrow in inset, flexible positioning of the NTD with respect to the remainder of the NusA. Rotated 45° about the horizontal axis (top to front) relative to Fig. 1b, left.

CX-MS and cryo-EM analyses are in full agreement with the crystal structure of the isolated RNP, indicating that the latter was not influenced by truncations of λ N and NusA. The structures, together with comparative CX-MS analyses of a λ N-TAC and a TEC lacking λ N, suggest molecular mechanisms underlying λ N-mediated antitermination that are consistent with our structure-guided mutational analyses. We propose that λ N, in conjunction with Nus factors and *nut* RNA, launches a multitiered strategy to suppress termination.

Basal antitermination activity of λ N and support by Nus factors and *nut* RNA. As shown by our crosslinking data, the λ N C terminus binds elements of the RNAP RNA exit channel, probably stabilizing their relative organization and counteracting hybrid melting due to terminator hairpin formation or due to the action of termination factor ρ . Additionally, the λ N C terminus may reposition RNAP elements and thereby redirect nascent RNA, possibly creating a steric barrier for terminator hairpin formation. Based on our structural results, λ N could also exert these functions in the absence of Nus factors and *nut* RNA, in agreement with λ N alone exhibiting basal antitermination activity^{37,47}. Our results also suggest how λ N-based antitermination is rendered processive by the Nus factors and *nut* RNA⁵. Nus factors and *nut* RNA almost certainly stabilize and possibly orient λ N on RNAP, as indicated by additional NusA-RNAP interactions and by the numerous mutual contacts within the λ N-NusA-NusB-NusE-*nut* RNP.

RNAP binding by λ N and its cooperation with Nus factors and *nut* RNA hinges on λ N being intrinsically unstructured, as a consequence of which essentially all of its side chains are surface-exposed and the protein is conformationally malleable. Despite its relatively small size, λ N can thus engage via numerous, densely arrayed, short epitopes in interactions with *boxB*, multiple domains of NusA, NusE, various parts of RNAP and nascent RNA at the same time. Additionally, intrinsic disorder may allow overlapping regions of λ N to adapt to different interaction platforms, such as residues

34–52 to NusA AR1 or the NusA-NusE cavity, during TAC assembly (Supplementary Discussion).

Reprogramming of NusA into an antitermination factor. A low-resolution EM analysis of a RNAP-NusA complex³, mapping of the NusA NTD on RNAP by NMR (ref. 48) and biochemical data^{45,46} indicate that in a TEC lacking λ N, the NusA NTD resides next to the β flap tip and β' ZnF, with the S1 and KH domains extending towards and beyond the β' dock domain. In this situation, nascent RNA emerging from the RNA exit channel is suggested to run between the RNAP β' dock domain on one side and the NusA NTD, S1 and KH domains on the other^{45,46}. NusA may thus display the upstream branch of a terminator hairpin in close vicinity to the downstream branch emerging from the exit channel (Fig. 5a), consistent with the suggested direct stabilization by NusA of RNA duplexes formed in the exit channel⁴⁹. The NusA NTD contributes most of this termination-promoting activity of NusA, while the remaining domains seem to predominantly enhance NusA association with the TEC (ref. 46). Clearly, a full understanding of how NusA functions in a TEC will require elucidation of a high-resolution TEC structure and structure-guided mutational analyses.

Our cryo-EM structure of a λ N-TAC and our comparative CX-MS analyses indicate that λ N repositions NusA and the β flap tip on RNAP, possibly allowing nascent RNA to thread along the NusA NTD and S1 domain towards the *boxB* element of the *nut* site, the latter being suspended across the NusA KH domains and the NusB-NusE heterodimer remote from the RNA exit (Fig. 5b). The upstream branch of a terminator hairpin could then be bound on a surface of the λ N-NusA-NusB-NusE-*nut* RNP, remote from the downstream branch when it just emerges from the exit channel (Fig. 5b). Our structural results suggest that the NusA S1 domain may sequester the upper terminator branch, but it could be supported by other elements such as the NusA NTD. Additionally or alternatively, repositioned RNAP and/or NusA elements around the redirected nascent RNA may sterically

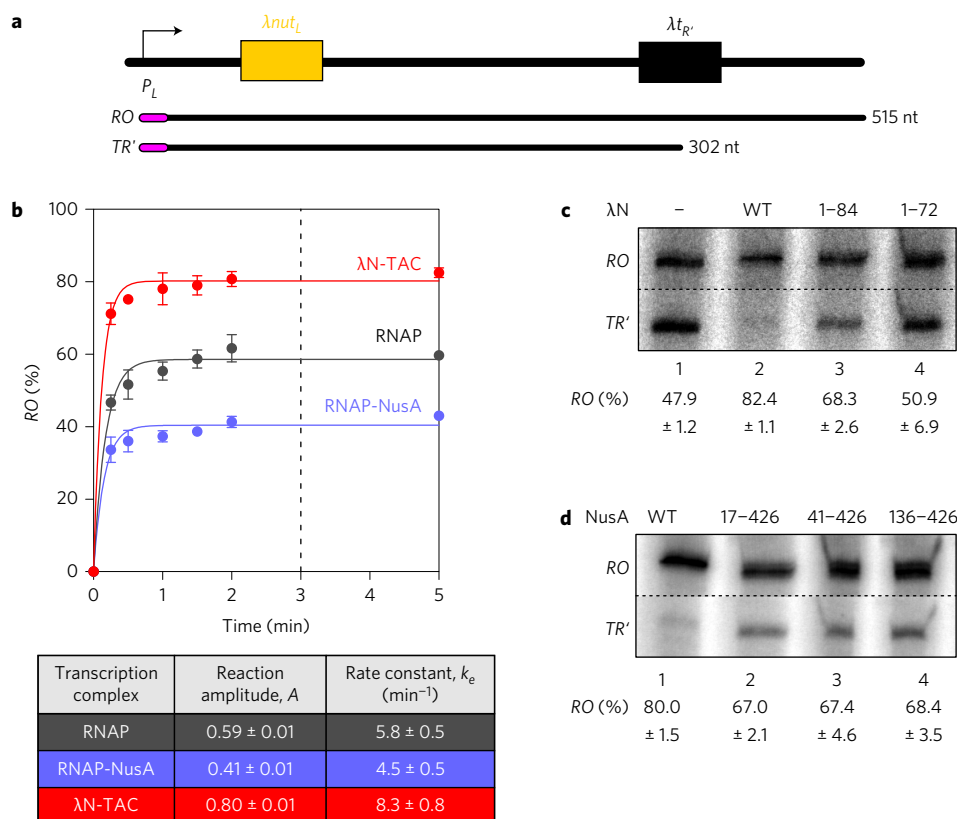


Figure 6 | In vitro functional analysis of antitermination activity. **a**, Scheme of the DNA template used for transcription assays and the observed products. RO, runoff product; TR', product terminated at $\lambda t_{R'}$; magenta regions, radioactively labelled initial transcripts. **b**, Top: Time courses of transcription by the indicated machineries, showing that product levels do not significantly change after 3 min (dashed line). Data represent means \pm s.e.m. of three independent experiments. Bottom: Data were fit to a first-order reaction (fraction RO = $A[1 - \exp(-k_e t)]$, where A is the amplitude of the reaction, k_e is the apparent first-order rate constant of transcription elongation, and t is time). **c,d**, Transcription assays monitoring antitermination efficiency at 3 min time points by TACs bearing the indicated λ N (**c**) or NusA (**d**) variants. Quantified data represent means \pm s.e.m. of three independent experiments. Lanes in **c** and **d** are from the same gel, but duplicate lanes and sections between RO and TR' have been removed (dashed lines, cut/paste sites). Gels shown are representative examples of three repetitions.

prevent terminator hairpin formation. In the future, these hypotheses have to be further scrutinized, for example by devising a higher-resolution structure of a λ N-TAC that would allow precisely targeted mutational studies.

Although *nut* RNA has previously been proposed to merely facilitate recruitment of λ N (ref. 37), our findings suggest that it also aids λ N in redirecting NusA functions. In the λ N-TAC, the N-terminal 20 residues of λ N are sandwiched between NusA KH2 and bound *boxB*, and the following λ N residues reorganize the NusA KH2 GXXG motif and occupy an RNA-binding platform neighbouring the KH2 β -sheet (Fig. 1d,e), thereby preventing continuous RNA binding along an extended KH1–KH2 surface as seen in an isolated NusA–RNA complex⁵⁰ (Fig. 1g). Thus, occupation of RNA binding sites on NusA KH1 and KH2 by λ N and *nut* RNA, further supported by NusB–NusE binding *boxA*, will prevent interaction of the NusA domains with alternative nascent RNA sequences and might modulate NusA function in certain scenarios. For example, NusA alone decreases termination by a terminator hairpin bearing an elongated loop, which was previously explained by postulating two NusA molecules per RNAP (ref. 37). However, our analyses suggest that in the absence of λ N, NusA KH1 and KH2 would not be occupied by *nut* RNA and be available for sequestering the upstream branch of a long-loop terminator distal to the exit channel.

Possible mechanisms for inhibition of ρ -dependent termination. Our results also suggest how the Nus factors and *nut* RNP may

support λ N in opposing termination factor ρ . ρ action via a *rut* site upstream of a *nut* site could be prevented if the λ N–NusA–NusB–NusE–*nut* RNP posed a stable roadblock for ρ translocating towards RNAP. A *rut* site either up- or downstream but in close proximity to a *nut* site could also be sequestered via surfaces of the λ N–NusA–NusB–NusE–*nut* RNP. In addition, the bulky λ N–NusA–NusB–NusE–*nut* RNP may sterically hinder ρ engagement of a *rut* site close to a *nut* site. Furthermore, consistent with the observation that ρ associates with RNAP early during transcription², ρ might be recruited directly by the NusG CTD to RNAP (ref. 4) before being handed off to an emerging *rut* site. RNA downstream of the *nut* site is expected to loop out on the concave side of NusA in the λ N-TAC (Fig. 2d). A downstream *rut* site would thus be guided away from NusG-bound ρ by the intervening λ N–NusA–NusB–NusE–*nut* RNP (Fig. 2d), disfavoured ρ engagement of the transcript. Structures of TECs and TACs stalled in the presence of ρ will be required to eventually elucidate the mode of ρ action and how it is prevented during processive antitermination.

Methods

Protein production. DNA fragments encoding NusA and NusG were generated by PCR from *E. coli* chromosomal DNA and cloned into pETM-11 vectors (European Molecular Biology Laboratories) under the control of a T7 promoter via *Bsa*I and *Xho*I restriction sites to produce proteins with a TEV-cleavable N-terminal His₆-tag. A DNA fragment encoding λ N was produced by assembly PCR and cloned into the pGEX-6P-1 vector using *Bam*HI and *Xho*I restriction sites to produce a protein with

a PreScission-cleavable N-terminal GST-tag. NusB–NusE (ref. 9) and RNAP (ref. 51) were produced based on previously described plasmids. Mutations were introduced by site-directed mutagenesis using the QuikChange protocol (Stratagene). All constructs were verified by DNA sequencing (SeqLab).

Plasmids containing the genes of interest were transformed into *E. coli* Rosetta2 (DE3) cells. The cells were grown in auto-inducing medium⁵² to an optical density at 600 nm (OD_{600}) of 0.3 at 37 °C and incubated for an additional 16 h at 18 °C. Cells were collected at 4 °C by centrifugation, resuspended, lysed by sonication using a Sonoplus ultrasonic homogenizer HD3100 (Bandelin), and centrifuged at 4 °C and 21,500 r.p.m. for 1 h. NusA, NusB, NusE, NusG, RNAP and $\sigma 70$ proteins and variants were purified as described previously^{9,34,53,54}. GST– λ N in 50 mM Tris–HCl, pH 7.6, 500 mM NaCl, 10 mM EDTA, 10% (vol/vol) glycerol, 10 mM dithiothreitol (DTT) was captured on glutathione sepharose beads (GE Healthcare), washed, eluted with 20 mM reduced glutathione (Sigma–Aldrich) supplemented with PreScission protease, and dialysed against 50 mM Tris–HCl, pH 7.6, 150 mM NaCl, 0.1 mM EDTA, 1 mM DTT. The protein solution was diluted to 50 mM NaCl and applied to a MonoS cation exchange column (GE Healthcare) in 20 mM Tris–HCl, pH 7.6, 50 mM NaCl, 0.1 mM EDTA, 5% (vol/vol) glycerol, 1 mM DTT and eluted with a gradient to 1 M NaCl. Pooled fractions were concentrated and further separated via a Superdex 75 size-exclusion column (GE Healthcare) in 10 mM Tris–HCl, pH 7.6, 200 mM NaCl, 1 mM DTT. Peak fractions were pooled, concentrated, and stored at –80 °C.

RNA production. The shortened *nut* site RNA used to obtain optimized crystals was chemically synthesized (Dharmacon). All other RNAs were produced by T7 RNAP-based *in vitro* transcription and purified via a non-denaturing method using strong anion-exchange chromatography and size-exclusion chromatography^{55,56}. To this end, DNA templates containing a T7 promoter and downstream region of interest were produced by assembly PCR and cloned into the pUC18 vector (Thermo Scientific) via *Xba*I and *Kpn*I restriction sites. For *in vitro* transcription, a PCR product was generated using reverse oligos that harboured two terminal 2'-O-methylated nucleotides to reduce 3'-end heterogeneity^{57,58}. Transcription reactions were performed in 120 mM HEPES–NaOH, pH 7.5, 3 mM rNTPs (ATP, GTP, UTP, CTP), 40 mM DTT, 4.8 mM spermidine, 0.125 mg ml^{–1} acetylated BSA, 16 mM MgCl₂, 0.04 μ l μ l^{–1} reaction PPase, 10 ng μ l^{–1} PCR product, 0.15 μ g μ l^{–1} T7 RNAP at 37 °C overnight. The reactions were treated with DNaseI, loaded on a MonoQ anion-exchange column (GE Healthcare) in 50 mM HEPES–NaOH, pH 6.9, 0.2 mM EDTA and eluted in a gradient to 2 M NaCl. Pooled fractions were further purified via Superdex 75 size-exclusion chromatography in 10 mM HEPES–NaOH, pH 7.5, 50 mM NaCl. Peak fractions were pooled, concentrated and stored at –80 °C.

Interaction assays. Interactions were studied using analytical size-exclusion chromatography. Stock solutions of proteins and/or nucleic acids were combined to generate equimolar mixtures (20 μ l final concentration for each component) in 10 mM HEPES–NaOH, pH 7.5, 50 mM NaCl, 1 mM DTT and incubated for 15 min at room temperature. The mixtures (50 μ l) were loaded on a Superdex 200 Increase 3.2/300 column (GE Healthcare) and chromatographed at a flow rate of 50 μ l min^{–1} in 10 mM HEPES–NaOH, pH 7.5, 50 mM NaCl, 1 mM DTT at 4 °C. Fractions (50 μ l) were collected and analysed by SDS–PAGE (17.5% separating gel) and urea PAGE (15%, 8 M urea) to reveal protein and nucleic acid contents, respectively.

Transcription assays. A linear DNA fragment, containing the λP_L promoter, λnut_L , λt_R , intrinsic terminator and additional vector-derived sequences to distinguish between products of termination and read through, was generated by PCR based on pKC30 (ref. 59). In control reactions we used a DNA fragment in which the λnut_L element was deleted. *In vitro* transcription was assayed in single-round format as described in ref. 60. Briefly, 100 nM *E. coli* core RNAP, 100 nM $\sigma 70$ factor and 10 nM template DNA were mixed in a 12.5 μ l reaction in transcription buffer (20 mM Tris–acetate, pH 7.9, 100 mM potassium acetate, 5 mM magnesium acetate, 5% (vol/vol) glycerol, 1 mM DTT) with 100 μ M ApU, 2 μ M ATP, 2 μ M GTP, 2 μ M CTP, 2 μ Ci [α -³²P] ATP and incubated for 10 min at 32 °C. The initial incubation allowed the formation of a halted ternary RNAP– $\sigma 70$ –DNA complex containing a 15 nucleotide initial transcript, as UTP was omitted from the reaction. RNA chain elongation was started by addition of 12.5 μ l of a preheated (32 °C) mixture containing 200 nM NusA, 1 μ M NusG, 1 μ M NusB–NusE, 500 nM λ N, 4 mM ATP, 100 μ M GTP, 100 μ M CTP and 100 μ M UTP in transcription buffer and incubated for defined times. Transcription was stopped by the addition of 5 μ l proteinase K (5 mg ml^{–1}) and incubated for 5 min at 50 °C followed by phenol/chloroform/isoamyl alcohol extraction and ethanol precipitation (3 vol. 1:30 3 M sodium acetate: ethanol, 0.5 μ l glycoblue). The RNA was washed three times with 70% ethanol and dried. It was then dissolved in 1xtris/borate/EDTA buffer that also contained 0.5 mg ml^{–1} urea, 2 mg ml^{–1} xylene cyanol, 2 mg ml^{–1} bromophenol blue, 200 μ l ml^{–1} 0.5 M EDTA, pH 8.0, incubated for 2 min at 96 °C, and loaded on a preheated 6 M urea, 4% polyacrylamide gel. RNA bands were visualized using a Storm PhosphorImager and quantified using Image-Quant software (GE Healthcare). The antitermination efficiency was determined as a ratio of the read through to the sum of the terminated and read-through products.

Crystallographic procedures. NusA^{ΔAR2}, NusB, NusE, λ N and *in vitro* transcribed *nut* RNA (λ N complex) or NusA^{ΔAR2}, NusB, NusE, λ N^{1–84} and synthetic *nut* RNA (λ N^{1–84} complex) were mixed in equimolar ratios in 10 mM HEPES–NaOH, pH 7.5, 50 mM NaCl, 1 mM DTT and passed over a Superdex 200 (GE Healthcare) size-exclusion column. Fractions containing the target complex were pooled and concentrated to 25 mg ml^{–1}. Crystallization was performed using the sitting-drop vapour diffusion technique at 22 °C in 48-well plates. Crystals of the λ N complex were obtained by mixing 1 μ l complex solution with 0.2 μ l of 0.1 M potassium tetracyanoplatinate and 0.8 μ l of reservoir solution (0.1 M MES–NaOH, pH 6.3, 14% (wt/vol) PEG 12000). Crystals of the λ N^{1–84} complex were obtained by mixing equal volumes of complex solution and reservoir solution (0.1 M HEPES–NaOH, pH 7.5, 40% (vol/vol) ethylene glycol, 5% (wt/vol) PEG 3000). Crystals in their mother liquor (λ N^{1–84} complex) or after transfer into mother liquor containing 20% (vol/vol) ethylene glycol (λ N complex) were flash-cooled in liquid nitrogen.

Diffraction data were collected at 100 K on beamline P14 of the Petra III synchrotron (Hamburg, Germany) and on beamline 14.1 of the BESSY II storage ring (Berlin, Germany). The data were processed with XDS⁶¹ and the structure of the λ N complex was solved by molecular replacement using the structure coordinates of NusA domains (*Thermotoga maritima* S1–KH1–KH2; PDB ID 1L2F (ref. 62); *E. coli* NTD, PDB ID 2KWP; *E. coli* AR1, PDB ID 1U9L (ref. 17)), *E. coli* NusB–NusE (PDB ID 3B3D (ref. 9)) and λ N^{1–22}–boxB (PDB ID 1QFQ (ref. 14)) as search models with Phaser⁶³ and completed by model building in Coot⁶⁴. The structure of the λ N^{1–84} complex was solved by molecular replacement using the structure coordinates of the λ N complex and by alternating rounds of model building in Coot and automated refinement using phenix.refine⁶⁵ and CNS⁶⁶.

TAC assembly. For assembly of a TAC, we followed a strategy as used previously for the structural investigation of bacterial RNAP-based^{67,68} and eukaryotic RNA polymerase II-based TECs^{69–71}. We generated a 66 nt RNA by *in vitro* transcription, whose 3'-terminal 10 nt could pair to the template DNA strand within a 14 nt non-complementary region of a 65 nt DNA duplex. The 5'-terminal 33 nt of this RNA contained the same *nut* site as used in crystallization. TAC was assembled by pre-incubating RNAP with the DNA duplex and mixing the RNAP–DNA complex with equimolar amounts of NusA, NusG, NusB, NusE, λ N and RNA. The mixture was incubated for 15 min at room temperature and separated by Superdex 200 size-exclusion chromatography in 10 mM HEPES–NaOH, pH 7.5, 50 mM NaCl, 1 mM DTT. Fractions containing the target complex were pooled and used without further concentration.

EM procedures. For the determination of an initial map, the sample was applied to a freshly glow-discharged holey carbon grid (Quantifoil) covered with an additional thin carbon support film and negatively stained using 3.5 μ l of 2% uranyl acetate solution. Grids were imaged on a Tecnai Spirit electron microscope operated at 120 kV equipped with a 2k × 2k Eagle charge-coupled device (CCD) camera (FEI). Micrographs with a pixel size of 5.19 Å px^{–1} were acquired fully automatically using Leginon⁷². Contrast transfer function (CTF) parameters were estimated using CTFFind 3 (ref. 73). A total of 25,543 particle images were selected manually from 128 micrographs using EMAN2 (ref. 74), centred, and grouped into homogeneous classes using the ISAC algorithm⁷⁵. This resulted in 240 stable classes, accounting for 17,825 particle images. Suitable class averages exhibiting distinctive structural features and representing possible projections of the complex were selected and used for reconstruction of an initial map via the VIPER algorithm^{76,77}. The map was used to determine the initial parameters by template matching for all particle images using SPARX (ref. 78).

For cryo-EM data acquisition, the sample was applied to freshly glow-discharged R3/3 holey carbon grids (Quantifoil) covered with a 2 nm carbon support film and plunge-frozen into liquid ethane using a Vitrobot plunger (FEI) with the blotting chamber set to 4 °C and 100% humidity. Grids were frozen using a blotting time of 2–4 s and stored in liquid nitrogen until use. For validation of the initial map, a small cryo-EM data set was acquired on a Tecnai Spirit electron microscope operated at 120 kV equipped with a 2k × 2k Eagle CCD camera (FEI). Micrographs were recorded with a pixel size of 5.19 Å px^{–1} using Leginon and processed as described before. A total of 46,327 particle images were selected with EMAN2 and used for template matching using SPARX against the previously determined initial map.

For the final cryo-EM reconstruction, micrographs were acquired on a Tecnai G2 Polara electron microscope (FEI) operated at 300 kV, equipped with a K2 direct electron detector (Gatan) operated in super-resolution mode and using the Leginon system. This resulted in 1,721 micrographs acquired within a defocus range of 1–8 μ m under low-dose conditions with a pixel size of 0.64 Å px^{–1}. Defocus estimation was performed using the CTFFIND4 package⁷⁹. Dose-fractionated image stacks were aligned using MotionCorr (ref. 80). Using EMAN2, 102,374 particle images were picked manually. All subsequent image-processing steps were performed with Relion 1.4 (refs 81, 82) unless otherwise noted. Fourfold binned particle images were subjected to two-dimensional classification and further selected by image quality (using a figure-of-merit criterion of better than 8 Å in CTFFIND4), resulting in a total of 79,946 suitable particle images. To find the initial parameters, three-dimensional refinement was performed with all twofold binned particles against the previously determined initial map. Three-dimensional classification

without alignment was then performed to sort the particles into four classes, with the classes containing 18, 28, 24 and 30% of the particles images, respectively. Three-dimensional refinement was performed for each class independently, resulting in three low-quality maps and one medium-quality class with an overall gold-standard resolution of 9.8 Å according to the FSC_{0.143} criterion (Supplementary Fig. 1e,g). Power spectrum readjustment was applied to the final map in SPIDER⁸³ after calculating a spectral weighting function from the raw map and a map simulated from the atomic model and map was low-pass-filtered to the global FSC estimate. The local resolution of the raw map was calculated using ResMap (ref. 84; Supplementary Fig. 1h,i).

The low overall resolution can be explained by the strong orientational preference of the particle images (Supplementary Fig. 1j). Although results from ResMap indicate a large portion of voxels with a resolution better than 9 Å (Supplementary Fig. 1h), the map itself does not show the details expected at this resolution. We therefore chose the FSC_{0.143} criterion as a more reliable representation of the map resolution. The strong orientational preference might have an impact on the ResMap calculations.

Modelling the TAC. A crystal structure of *E. coli* RNAP (PDB ID: 4YG2; ref. 44) was manually docked into the cryo-EM map and the fit was optimized by rigid body refinement in Coot. The RNAP model was locally adjusted by transferring the FTH coordinates from the structure of the *E. coli* RNAP in complex with RapA (PDB ID: 4S20 (ref. 85)) after superpositioning of the RNAP cores. Nucleic acids were transferred from structures of *Thermus thermophilus* RNAP (PDB ID: 2O5I (ref. 67); DNA:RNA hybrid and RNA in the RNA exit channel) and bovine Pol II (PDB ID: 5FLM (ref. 86); upstream dsDNA) by superpositioning of RNAP core subunits. The λ N¹⁻⁸⁴-NusA^{ΔAR2}-NusB-NusE-nut RNP was docked into a distinctly shaped, unoccupied region of the cryo-EM map and adjusted by rigid body refinement. The structures of NusG NTD and NusG CTD (PDB ID: 2K06 and 2JVJ (ref. 87)) were added according to the structure of archaeal Spt4/5 bound to the β' clamp helices of Pol II (PDB ID: 3QQC (ref. 33)) and an *E. coli* NusE-NusG-CTD structure (PDB ID: 2KVQ (ref. 34)), respectively.

Chemical crosslinking and crosslink identification by MS. For CX-MS we used the homobifunctional, primary amino group-reactive BS3 (bis(sulfosuccinimidyl) suberate) crosslinker, which can bridge N termini or lysine side chains when their Ca atoms are <35 Å apart, or the hetero-bifunctional 'zero-length' EDC (N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride) and DMTMM (4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride) crosslinkers, which mediate the direct formation of amide bonds between carboxylate (Asp or Glu side chains) and primary amine (Lys side chains) groups.

A total of 25 pmol of λ N-NusA^{ΔAR2}-NusB-NusE-nut RNP was crosslinked with 150 μM BS3 or 30 mM DMTMM in 10 mM HEPES, pH 7.5, 50 mM NaCl and 1 mM DTT and incubated for 30 min at room temperature. The reaction was stopped by 25 mM Tris-HCl. Crosslinking efficiency was analysed by SDS-PAGE. The crosslinked λ N-NusA^{ΔAR2}-NusB-NusE-nut RNP was excised from the gel and digested with trypsin⁸⁸. The digested proteins were dissolved in 20 μl of 5% (vol/vol) acetonitrile and 0.1% (vol/vol) formic acid and subjected to liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis using a Q Exactive (Thermo Scientific) mass spectrometer.

For TAC and TEC, 100 pmol of purified recombinant complex was crosslinked with 150 μM BS3 for 30 min in 10 mM HEPES-NaOH, pH 7.5, 50 mM NaCl or with 4 mM EDC, 1 mM N-hydroxysuccinimide for 60 min in 10 mM MES-NaOH, pH 6.5, 50 mM NaCl at 25 °C. The samples were acetone-precipitated and analysed essentially as described before⁸⁹, with the following modifications. Precipitated material was dissolved in 4 M urea/50 mM ammonium bicarbonate, reduced with DTT, alkylated with iodoacetamide, diluted to 1 M urea and digested with trypsin (1:20 (wt/wt)). Peptides were reversed-phase extracted using Sep-Pak Vac tC18 1cc cartridges (Waters) and fractionated by size-exclusion on a Superdex Peptide PC3.2/30 column (GE Healthcare). Fractions (50 μl) corresponding to elution volumes of 1.2–1.8 ml were analysed on a Q Exactive HF (Thermo Scientific) mass spectrometer. Protein-protein crosslinks were identified via the pLink1.22 search engine (<http://pfind.ict.ac.cn/software/pLink>) and filtered at false discovery rate 1% according to the recommendations of the developers⁹⁰. A crosslink score was evaluated as the negative value of the common logarithm of the original pLink Score (score = -log₁₀[pLink score]). The crosslinks observed with >4 spectra for BS3 and >7 spectra for EDC and DMTMM are listed in Supplementary Tables 2–4.

Data availability. Coordinates and structure factors have been deposited in the RCSB Protein Data Bank (www.pdb.org) under accession codes 5LM7 (NusA^{ΔAR2}-NusB-NusE- λ N¹⁻⁸⁴-nut RNA complex) and 5LM9 (NusA¹⁰⁰⁻⁴²⁶). Cryo-EM data have been deposited in the RCSB Protein Data Bank under accession code 5MS0 and in the Electron Microscopy Data Bank (www.emdatabank.org) under accession code EMD-3561.

Received 15 December 2016; accepted 24 March 2017;
published 28 April 2017

References

- Zhang, J. & Landick, R. A two-way street: regulatory interplay between RNA polymerase and nascent RNA structure. *Trends Biochem. Sci.* **41**, 293–310 (2016).
- Mooney, R. A. *et al.* Regulator trafficking on bacterial transcription units *in vivo*. *Mol. Cell* **33**, 97–108 (2009).
- Yang, X. & Lewis, P. J. The interaction between RNA polymerase and the elongation factor NusA. *RNA Biol.* **7**, 272–275 (2010).
- Tomar, S. K. & Artsimovitch, I. NusG-Spt5 proteins—universal tools for transcription modification and communication. *Chem. Rev.* **113**, 8604–8619 (2013).
- Nudler, E. & Gottesman, M. E. Transcription termination and anti-termination in *E. coli*. *Genes Cells* **7**, 755–768 (2002).
- Ciampi, M. S. Rho-dependent terminators and transcription termination. *Microbiology* **152**, 2515–2528 (2006).
- Torres, M., Condon, C., Balada, J. M., Squires, C. & Squires, C. L. Ribosomal protein S4 is a transcription factor with properties remarkably similar to nusA, a protein involved in both non-ribosomal and ribosomal RNA antitermination. *EMBO J.* **20**, 3811–3820 (2001).
- Worbs, M., Bourenkov, G. P., Bartunik, H. D., Huber, R. & Wahl, M. C. An extended RNA binding surface through arrayed S1 and KH domains in transcription factor NusA. *Mol. Cell* **7**, 1177–1189 (2001).
- Luo, X. *et al.* Structural and functional analysis of the *E. coli* NusB-S10 transcription antitermination complex. *Mol. Cell* **32**, 791–802 (2008).
- Stagno, J. R. *et al.* Structural basis for RNA recognition by NusB and NusE in the initiation of transcription antitermination. *Nucleic Acids Res.* **39**, 7803–7815 (2011).
- Zhou, Y., Mah, T. F., Greenblatt, J. & Friedman, D. I. Evidence that the KH RNA-binding domains influence the action of the *E. coli* NusA protein. *J. Mol. Biol.* **318**, 1175–1188 (2002).
- Mogridge, J. *et al.* Independent ligand-induced folding of the RNA-binding domain and two functionally distinct antitermination regions in the phage lambda N protein. *Mol. Cell* **1**, 265–275 (1998).
- Legault, P., Li, J., Mogridge, J., Kay, L. E. & Greenblatt, J. NMR structure of the bacteriophage lambda N peptide/boxB RNA complex: recognition of a GNRA fold by an arginine-rich motif. *Cell* **93**, 289–299 (1998).
- Scharpf, M. *et al.* Antitermination in bacteriophage lambda. The structure of the N36 peptide-boxB RNA complex. *Eur. J. Biochem.* **267**, 2397–2408 (2000).
- Mogridge, J., Mah, T. F. & Greenblatt, J. Involvement of boxA nucleotides in the formation of a stable ribonucleoprotein complex containing the bacteriophage lambda N protein. *J. Biol. Chem.* **273**, 4143–4148 (1998).
- Mah, T. F., Li, J., Davidson, A. R. & Greenblatt, J. Functional importance of regions in *Escherichia coli* elongation factor NusA that interact with RNA polymerase, the bacteriophage lambda N protein and RNA. *Mol. Microbiol.* **34**, 523–537 (1999).
- Bonin, I. *et al.* Structural basis for the interaction of *Escherichia coli* NusA with protein N of phage lambda. *Proc. Natl. Acad. Sci. USA* **101**, 13762–13767 (2004).
- Mishra, S., Mohan, S., Godavarthi, S. & Sen, R. The interaction surface of a bacterial transcription elongation factor required for complex formation with an antiterminator during transcription antitermination. *J. Biol. Chem.* **288**, 28089–28103 (2013).
- Mah, T. F., Kuznedelov, K., Mushegian, A., Severinov, K. & Greenblatt, J. The alpha subunit of *E. coli* RNA polymerase activates RNA binding by NusA. *Genes Dev.* **14**, 2664–2675 (2000).
- Mogridge, J., Mah, T. F. & Greenblatt, J. A protein-RNA interaction network facilitates the template-independent cooperative assembly on RNA polymerase of a stable antitermination complex containing the lambda N protein. *Genes Dev.* **9**, 2831–2845 (1995).
- Zhou, Y. *et al.* Interactions of an Arg-rich region of transcription elongation protein NusA with NUT RNA: implications for the order of assembly of the lambda N antitermination complex *in vivo*. *J. Mol. Biol.* **310**, 33–49 (2001).
- Friedman, D. I. & Baron, L. S. Genetic characterization of a bacterial locus involved in the activity of the N function of phage lambda. *Virology* **58**, 141–148 (1974).
- Ward, D. F. & Gottesman, M. E. The nus mutations affect transcription termination in *Escherichia coli*. *Nature* **292**, 212–215 (1981).
- Craven, M. G. & Friedman, D. I. Analysis of the *Escherichia coli* nusA10(Cs) allele: relating nucleotide changes to phenotypes. *J. Bacteriol.* **173**, 1485–1491 (1991).
- Craven, M. G. *et al.* *Escherichia coli*-*Salmonella typhimurium* hybrid nusA genes: identification of a short motif required for action of the lambda N transcription antitermination protein. *J. Bacteriol.* **176**, 1394–1404 (1994).
- Friedman, D. I. *et al.* Interactions of bacteriophage and host macromolecules in the growth of bacteriophage lambda. *Microbiol. Rev.* **48**, 299–325 (1984).
- Ward, D. F., DeLong, A. & Gottesman, M. E. *Escherichia coli* nusB mutations that suppress nusA1 exhibit lambda N specificity. *J. Mol. Biol.* **168**, 73–85 (1983).

28. Friedman, D. I., Schauer, A. T., Baumann, M. R., Baron, L. S. & Adhya, S. L. Evidence that ribosomal protein S10 participates in control of transcription termination. *Proc. Natl Acad. Sci. USA* **78**, 1115–1118 (1981).
29. Sparkowski, J. & Das, A. Simultaneous gain and loss of functions caused by a single amino acid substitution in the beta subunit of *Escherichia coli* RNA polymerase: suppression of *nusA* and rho mutations and conditional lethality. *Genetics* **130**, 411–428 (1992).
30. Friedman, D. I. & Olson, E. R. Evidence that a nucleotide sequence, 'boxA,' is involved in the action of the NusA protein. *Cell* **34**, 143–149 (1983).
31. Friedman, D. I., Olson, E. R., Johnson, L. L., Alessi, D. & Craven, M. G. Transcription-dependent competition for a host factor: the function and optimal sequence of the phage lambda *boxA* transcription antitermination signal. *Genes Dev.* **4**, 2210–2222 (1990).
32. Liu, B. & Steitz, T. A. Structural insights into NusG regulating transcription elongation. *Nucleic Acids Res.* **45**, 968–974 (2017).
33. Martinez-Rucobo, F. W., Sainsbury, S., Cheung, A. C. M. & Cramer, P. Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity. *EMBO J.* **30**, 1302–1310 (2011).
34. Burmann, B. M. *et al.* A NusE:NusG complex links transcription and translation. *Science* **328**, 501–504 (2010).
35. Sullivan, S. L., Ward, D. F. & Gottesman, M. E. Effect of *Escherichia coli* *nusG* function on lambda N-mediated transcription antitermination. *J. Bacteriol.* **174**, 1339–1344 (1992).
36. Kuznedelov, K., Korzhova, N., Mustaev, A. & Severinov, K. Structure-based analysis of RNA polymerase function: the largest subunit's rudder contributes critically to elongation complex stability and is not involved in the maintenance of RNA–DNA hybrid length. *EMBO J.* **21**, 1369–1378 (2002).
37. Gusarov, I. & Nudler, E. Control of intrinsic transcription termination by N and NusA: the basic mechanisms. *Cell* **107**, 437–449 (2001).
38. Parks, A. R. *et al.* Bacteriophage lambda N protein inhibits transcription slippage by *Escherichia coli* RNA polymerase. *Nucleic Acids Res.* **42**, 5823–5829 (2014).
39. Mishra, S. & Sen, R. N protein from lambdaoid phages transforms NusA into an antiterminator by modulating NusA–RNA polymerase flap domain interactions. *Nucleic Acids Res.* **43**, 5744–5758 (2015).
40. Cheeran, A. *et al.* *Escherichia coli* RNA polymerase mutations located near the upstream edge of an RNA:DNA hybrid and the beginning of the RNA-exit channel are defective for transcription antitermination by the N protein from lambdaoid phage H-19B. *J. Mol. Biol.* **352**, 28–43 (2005).
41. Georgopoulos, C. P. Bacterial mutants in which gene N function of bacteriophage lambda is blocked have an altered RNA polymerase. *Proc. Natl Acad. Sci. USA* **68**, 2977–2981 (1971).
42. Jin, D. J. *et al.* Effects of rifampicin resistant *rpoB* mutations on antitermination and interaction with *nusA* in *Escherichia coli*. *J. Mol. Biol.* **204**, 247–261 (1988).
43. Cheeran, A., Koll, N. R. & Sen, R. The site of action of the antiterminator protein N from the lambdaoid phage H-19B. *J. Biol. Chem.* **282**, 30997–31007 (2007).
44. Murakami, K. S. X-ray crystal structure of *Escherichia coli* RNA polymerase sigma70 holoenzyme. *J. Biol. Chem.* **288**, 9126–9134 (2013).
45. Toulkhonov, I., Artsimovitch, I. & Landick, R. Allosteric control of RNA polymerase by a site that contacts nascent RNA hairpins. *Science* **292**, 730–733 (2001).
46. Ha, K. S., Toulkhonov, I., Vassilyev, D. G. & Landick, R. The NusA N-terminal domain is necessary and sufficient for enhancement of transcriptional pausing via interaction with the RNA exit channel of RNA polymerase. *J. Mol. Biol.* **401**, 708–725 (2010).
47. Rees, W. A., Weitzel, S. E., Yager, T. D., Das, A. & von Hippel, P. H. Bacteriophage lambda N protein alone can induce transcription antitermination *in vitro*. *Proc. Natl Acad. Sci. USA* **93**, 342–346 (1996).
48. Drogemuller, J. *et al.* Determination of RNA polymerase binding surfaces of transcription factors by NMR spectroscopy. *Sci. Rep.* **5**, 16428 (2015).
49. Kolb, K. E., Hein, P. P. & Landick, R. Antisense oligonucleotide-stimulated transcriptional pausing reveals RNA exit channel specificity of RNA polymerase and mechanistic contributions of NusA and rfaH. *J. Biol. Chem.* **289**, 1151–1163 (2014).
50. Beuth, B., Pennell, S., Arnvig, K. B., Martin, S. R. & Taylor, I. A. Structure of a *Mycobacterium tuberculosis* NusA–RNA complex. *EMBO J.* **24**, 3576–3587 (2005).
51. Artsimovitch, I., Svetlov, V., Murakami, K. S. & Landick, R. 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).
52. Studier, F. W. Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.* **41**, 207–234 (2005).
53. Kyzer, S., Ha, K. S., Landick, R. & Palangat, M. Direct versus limited-step reconstitution reveals key features of an RNA hairpin-stabilized paused transcription complex. *J. Biol. Chem.* **282**, 19020–19028 (2007).
54. Belogurov, G. A. *et al.* Structural basis for converting a general transcription factor into an operon-specific virulence regulator. *Mol. Cell* **26**, 117–129 (2007).
55. Easton, L. E., Shibata, Y. & Lukavsky, P. J. Rapid, nondenaturing RNA purification using weak anion-exchange fast performance liquid chromatography. *RNA* **16**, 647–653 (2010).
56. McKenna, S. A. *et al.* Purification and characterization of transcribed RNAs using gel filtration chromatography. *Nat. Protoc.* **2**, 3270–3277 (2007).
57. Kao, C., Zheng, M. & Rudisser, S. A simple and efficient method to reduce nontemplated nucleotide addition at the 3' terminus of RNAs transcribed by T7 RNA polymerase. *RNA* **5**, 1268–1272 (1999).
58. Sherlin, L. D. *et al.* Chemical and enzymatic synthesis of tRNAs for high-throughput crystallization. *RNA* **7**, 1671–1678 (2001).
59. Rao, R. N. Construction and properties of plasmid pKC30, a pBR322 derivative containing the pL-N region of phage lambda. *Gene* **31**, 247–250 (1984).
60. Artsimovitch, I. & Henkin, T. M. *In vitro* approaches to analysis of transcription termination. *Methods* **47**, 37–43 (2009).
61. Kabsch, W. XDS. *Acta Crystallogr. D* **66**, 125–132 (2010).
62. Shin, D. H. *et al.* Crystal structure of NusA from *Thermotoga maritima* and functional implication of the N-terminal domain. *Biochemistry* **42**, 13429–13437 (2003).
63. McCoy, A. J. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. D* **63**, 32–41 (2007).
64. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
65. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D* **68**, 352–367 (2012).
66. Schroder, G. F., Levitt, M. & Brunger, A. T. Deformable elastic network refinement for low-resolution macromolecular crystallography. *Acta Crystallogr. D* **70**, 2241–2255 (2014).
67. Vassilyev, D. G., Vassilyeva, M. N., Perederina, A., Tahirov, T. H. & Artsimovitch, I. Structural basis for transcription elongation by bacterial RNA polymerase. *Nature* **448**, 157–162 (2007).
68. Vassilyev, D. G. *et al.* Structural basis for substrate loading in bacterial RNA polymerase. *Nature* **448**, 163–168 (2007).
69. Westover, K. D., Bushnell, D. A. & Kornberg, R. D. Structural basis of transcription: separation of RNA from DNA by RNA polymerase II. *Science* **303**, 1014–1016 (2004).
70. Westover, K. D., Bushnell, D. A. & Kornberg, R. D. Structural basis of transcription: nucleotide selection by rotation in the RNA polymerase II active center. *Cell* **119**, 481–489 (2004).
71. Kettenberger, H., Armache, K. J. & Cramer, P. Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIIS. *Mol. Cell* **16**, 955–965 (2004).
72. Suloway, C. *et al.* Automated molecular microscopy: the new legimon system. *J. Struct. Biol.* **151**, 41–60 (2005).
73. Mindell, J. A. & Grigorieff, N. Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.* **142**, 334–347 (2003).
74. Tang, G. *et al.* EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38–46 (2007).
75. Yang, Z., Fang, J., Chittuluru, J., Asturias, F. J. & Penczek, P. A. Iterative stable alignment and clustering of 2D transmission electron microscope images. *Structure* **20**, 237–247 (2012).
76. Penczek, P. A. & Asturias, F. J. *Ab initio* cryo-EM structure determination as a validation problem. *Proc. IEEE Int. Conf. Image Processing (ICIP)* **2014**, 2090–2094 (2014).
77. Cheng, Y., Grigorieff, N., Penczek, P. A. & Walz, T. A primer to single-particle cryo-electron microscopy. *Cell* **161**, 438–449 (2015).
78. Hohn, M. *et al.* SPARX, a new environment for cryo-EM image processing. *J. Struct. Biol.* **157**, 47–55 (2007).
79. Rohou, A. & Grigorieff, N. CTFFIND4: fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
80. Li, X. *et al.* Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat. Methods* **10**, 584–590 (2013).
81. Scheres, S. H. Semi-automated selection of cryo-EM particles in RELION-1.3. *J. Struct. Biol.* **189**, 114–122 (2015).
82. Bharat, T. A., Russo, C. J., Lowe, J., Passmore, L. A. & Scheres, S. H. Advances in single-particle electron cryomicroscopy structure determination applied to subtomogram averaging. *Structure* **23**, 1743–1753 (2015).
83. Frank, J., Shimkin, B. & Dowse, H. Spider—a modular software system for electron image-processing. *Ultramicroscopy* **6**, 343–357 (1981).
84. Kucukelbir, A., Sigworth, F. J. & Tagare, H. D. Quantifying the local resolution of cryo-EM density maps. *Nat. Methods* **11**, 63–65 (2014).
85. Liu, B., Zuo, Y. & Steitz, T. A. Structural basis for transcription reactivation by RapA. *Proc. Natl Acad. Sci. USA* **112**, 2006–2010 (2015).
86. Bernecky, C., Herzog, F., Baumeister, W., Plitzko, J. M. & Cramer, P. Structure of transcribing mammalian RNA polymerase II. *Nature* **529**, 551–554 (2016).
87. Mooney, R. A., Schweimer, K., Rosch, P., Gottesman, M. & Landick, R. Two structurally independent domains of *E. coli* NusG create regulatory plasticity via distinct interactions with RNA polymerase and regulators. *J. Mol. Biol.* **391**, 341–358 (2009).

88. Schmidt, C. & Urlaub, H. iTRAQ-labeling of in-gel digested proteins for relative quantification. *Methods Mol. Biol.* **564**, 207–226 (2009).
89. Leitner, A., Walzthoeni, T. & Aebersold, R. Lysine-specific chemical cross-linking of protein complexes and identification of cross-linking sites using LC-MS/MS and the xQuest/xProphet software pipeline. *Nat. Protoc.* **9**, 120–137 (2014).
90. Yang, B. *et al.* Identification of cross-linked peptides from complex samples. *Nat. Methods* **9**, 904–906 (2012).

Acknowledgements

The authors thank C. Alings (Freie Universität Berlin) for help with crystallization, I. Artsimovitch (Ohio State University) for plasmids pVS-10 (T7P- α - β -His6- ω) and pIA1127 (T7P-His6-TEV- σ 70), used for RNAP production, and M. Gottesman (New York University) for plasmid pKC30, used in transcription assays. The authors acknowledge access to beamlines of the BESSY II storage ring (Berlin, Germany) via the Joint Berlin MX-Laboratory sponsored by the Helmholtz Zentrum Berlin für Materialien und Energie, the Freie Universität Berlin, the Humboldt-Universität zu Berlin, the Max-Delbrück-Centrum and the Leibniz-Institut für Molekulare Pharmakologie and to beamline 14–1 at the Petra III storage ring (EMBL, Hamburg, Germany). The authors acknowledge access to the resources of the North-German Supercomputing Alliance (HLRN). This work was supported by the Deutsche Forschungsgemeinschaft (SFB 740 to T.M., C.M.S. and M.C.W. and grant WA 1126/5-1 to M.C.W.). K.F.S. was supported by a Dahlem International Network PostDoc grant from Freie Universität Berlin. E.B. holds a Freigeist Fellowship

from the Volkswagen Foundation and acknowledges continuing support from the Caesar Foundation.

Author contributions

N.S., F.K., E.A., K.F.S., O.D., Y.-H.H., C.-T.L., B.L., E.B., J.B., T.M., J.L. and G.W. performed the experiments. All authors contributed to analysis of the data and interpretation of the results. H.U., C.M.T.S. and M.C.W. supervised the studies. N.S. and M.C.W. wrote the manuscript.

Additional information

Supplementary information is available for this paper.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to M.C.W.

How to cite this article: Said, N. *et al.* Structural basis for λ N-dependent processive transcription antitermination. *Nat. Microbiol.* **2**, 17062 (2017).

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

Competing interests

The authors declare no competing financial interests.