Article

Structural basis of exoribonuclease-mediated mRNA transcription termination

Yuan Zeng[1,2], Hong-Wei Zhang[1,2], Xiao-Xian Wu[1] & Yu Zhang[1✉]

Efficient termination is required for robust gene transcription. Eukaryotic organisms use a conserved exoribonuclease-mediated mechanism to terminate the mRNA transcription by RNA polymerase II (Pol II)[1–5]. Here we report two cryogenic electron microscopy structures of *Saccharomyces cerevisiae* Pol II pre-termination transcription complexes bound to the 5′-to-3′ exoribonuclease Rat1 and its partner Rai1. Our structures show that Rat1 displaces the elongation factor Spt5 to dock at the Pol II stalk domain. Rat1 shields the RNA exit channel of Pol II, guides the nascent RNA towards its active centre and stacks three nucleotides at the 5′ terminus of the nascent RNA. The structures further show that Rat1 rotates towards Pol II as it shortens RNA. Our results provide the structural mechanism for the Rat1-mediated termination of mRNA transcription by Pol II in yeast and the exoribonuclease-mediated termination of mRNA transcription in other eukaryotes.

Termination is the last stage of gene transcription, during which RNA polymerases stop RNA extension, release RNA transcript and dissociate from the DNA template. In eukaryotes, Pol II uses distinct pathways to terminate transcription of protein-coding and non-coding genes[6–10]. The pathway for transcription termination of mRNAs is conserved in yeast, mammals and plants, requiring prior cleavage at the polyadenylation site (PAS) of the pre-mRNAs by the endonuclease subunit of the cleavage and polyadenylation complex[11–13] to release mRNA and subsequent cleavage of the Pol II-associated RNA by a 5′–3′ exoribonuclease[1–5] (Rat1 in yeast, XRN2 in humans and XRN3 in plants). The exoribonuclease is also involved in terminating transcription of premature mRNA at promoter-proximal PASs or decapped transcripts near transcription start sites[14–17].

Two prevailing models were proposed for the exoribonuclease-mediated termination of mRNA transcription by Pol II: the allosteric model and the torpedo model[7–9]. In the allosteric model, an exchange of elongation factors and termination factors causes conformational change of Pol II, leading to the dissociation of RNA and DNA from Pol II[18–21]. In the torpedo model, the exoribonuclease degrades the RNA in a 5′-to-3′ direction, catches up with the elongating Pol II, releases the RNA and dislodges Pol II from DNA[1–4]. Recently, it has become apparent that both mechanisms play an important role in the termination of mRNA transcription by exoribonucleases. This insight has led to the development of unified models, such as the unified allosteric–torpedo model and the sitting duck torpedo model[22–24]. Owing to the lack of structural information on Pol II transcription termination, how the 5′–3′ exoribonuclease Rat1 interacts with Pol II and the elongation factors to terminate mRNA transcription remains elusive.

Here we determined two cryogenic electron microscopy (cryo-EM) structures of *S. cerevisiae* Pol II pre-termination transcription complexes (Pol II Rat1–PTTC1 and Pol II Rat1–PTTC2) bound to the exoribonuclease Rat1–Rai1 heterodimer and the elongation factor Spt5 at 2.7 and 2.8 Å (Extended Data Table 1), respectively. Our structures provide insights for understanding the exoribonuclease-mediated termination of mRNA transcription by Pol II.
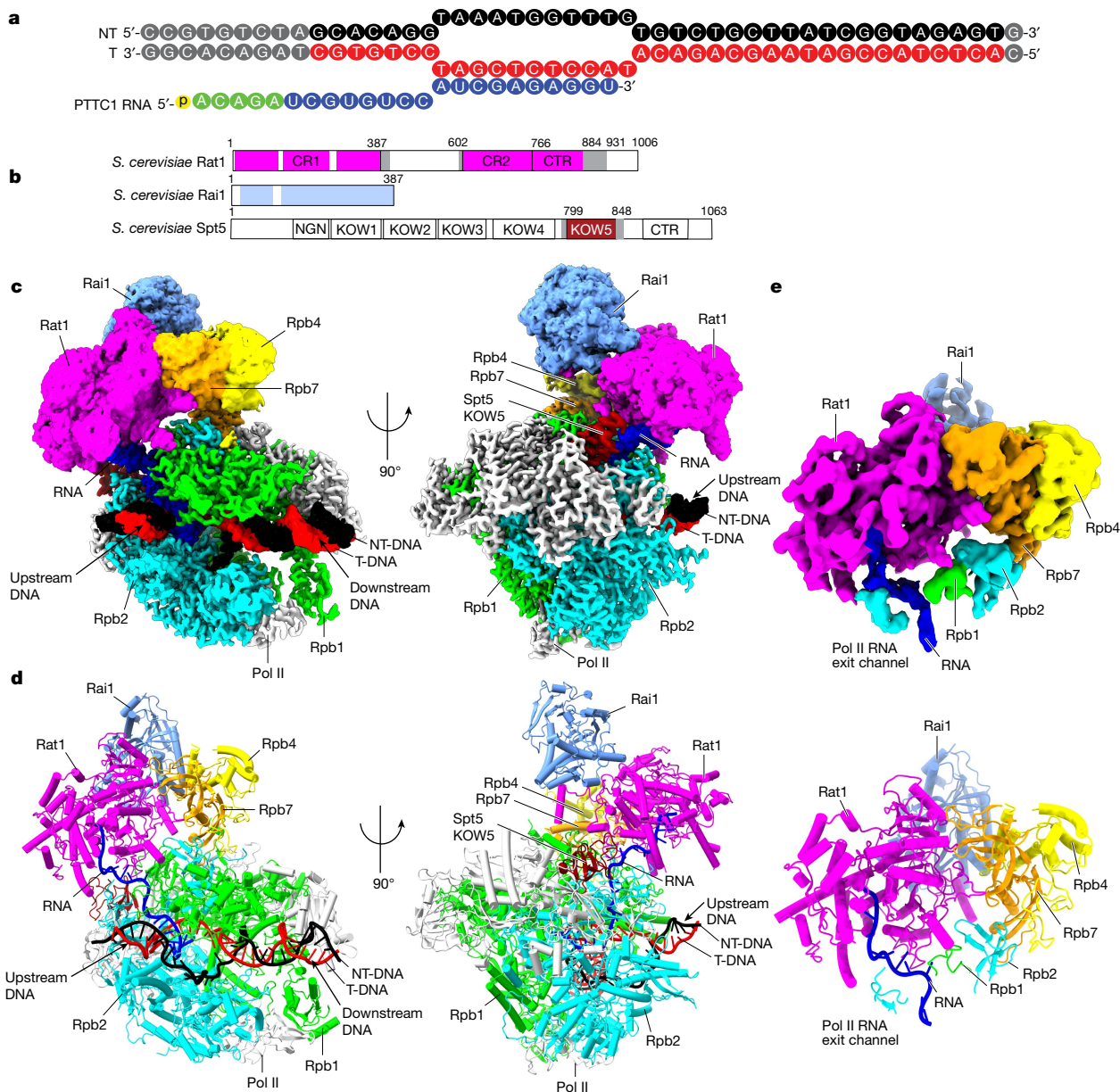
The overall structure of Pol II Rat1–PTTC1

To understand the structural basis for Rat1-mediated Pol II termination, we reconstituted a pre-termination complex comprising *S. cerevisiae* Pol II, the termination exoribonuclease–Rat1–Rai1 heterodimer, the elongation factors Spt4 and Spt5, and a nucleic acid scaffold that contains a 23-nucleotide (nt) RNA harbouring phosphorus-sulfur-bond modifications at the 5′ termini of the nucleotides (Fig. 1a,b, Methods and Extended Data Fig. 1). The cryo-EM structure of Pol II Rat1–PTTC1 was determined through a single-particle cryo-EM method at a nominal resolution of 2.7 Å (Extended Data Fig. 2).

The cryo-EM map (map 1a) calculated from Pol II Rat1–PTTC1 shows unambiguous signals near the RNA exit channel of Pol II that could be assigned to the Rat1–Rai1 heterodimer and the KOW5 domain of Spt5 (Fig. 1c,d and Extended Data Fig. 2). Local refinement focusing around the Rat1–Rai1 heterodimer resulted in a local map with a much-improved resolution of the Rat1–Rai1 heterodimer and the Pol II stalk domain, and continuous RNA signals from the Rat1 catalytic site to the RNA exit channel of Pol II (map 1b at 3.5 Å of Pol II Rat1–PTTC1; Fig. 1e and Extended Data Fig. 2e). We also obtained a cryo-EM map (map 1c) showing clear signals for all domains of Spt4 and Spt5 but not the Rat1–Rai1 heterodimer (Extended Data Fig. 2a,j), suggesting that Spt4 and Spt5 preclude interaction of the Rat1–Rai1 heterodimer with Pol II.

In the structure of Pol II Rat1–PTTC1, the Rat1–Rai1 heterodimer docks at the Pol II stalk domain and shields the RNA exit channel of Pol II; Rat1 and Pol II interact with the 5′ terminus and 3′ terminus of the nascent RNA, respectively; and only the KOW5 domain of Spt5 is associated with Pol II (Fig. 1c,d).

[1]Key Laboratory of Synthetic Biology, National Key Laboratory of Plant Design, CAS Center for Excellence in Molecular Plant Sciences, Shanghai Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai, China. [2]University of Chinese Academy of Sciences, Beijing, China. ✉e-mail: yzhang@cemps.ac.cn

Fig. 1 | The cryo-EM structure of Pol II Rat1–PTTC1. a, Sequences of the nucleic acid scaffold used in Pol II Rat1–PTTC1. Phosphorus-sulfur-bond modification of the RNA nucleotides is shown in green, and the mono-phosphorylation modification of the RNA 5′ terminus is shown in yellow. Nucleotides in grey are disordered in the structure. NT, non-template DNA; T, template DNA. **b**, Domain architecture of Rat1, Rai1 and Spt5. CR, conserved region; CTR, C-terminal region. The colour code presented here is used throughout all figures. All modelled regions are shown with colour. Rat1, magenta; Rai1, light blue; Spt5, dark red. **c**, The front and side views of the cryo-EM map (map 1a) of Pol II Rat1–PTTC1. Rpb1, green; Rpb2, cyan; Rpb4, yellow; Rpb7, orange; Pol II other subunits, light grey. **d**, The front and side views of the Pol II Rat1–PTTC1 structure. **e**, The local refined map (map 1b) and model of Rat1, Rai1, the Pol II subunits Rpb4 and Rpb7, and structure elements around the RNA exit channel of Pol II.

Rat1 displaces Spt5 to bind Pol II

Rai1 is a fungus-specific protein that stimulates Rat1 activity[25]. Rai1 is located at the opposite face of the active site of Rat1 and interacts with Rat1 through a 1,550 Å² otherwise exposed surface. Two concave surfaces of Rai1 embrace the carboxy-terminal structure elements of Rat1 (Extended Data Fig. 4). The smaller concave surface of Rai1 mainly interacts with three loops of Rat1 (residues 217–221, 313–318 and 853–865, respectively). This interaction has been observed in the crystal structure of the *Schizosaccharomyces pombe* Rat1–Rai1 complex[26] (Extended Data Fig. 4a–c). The larger concave surface of Rai1 interacts with another loop and a helix at the C terminus of Rat1 (residues 913–925; Extended Data Fig. 4a,d). This interaction has not been observed in the crystal structure of the *S. pombe* Rat1–Rai1 complex[26]. In summary, our structure shows that Rai1 forms a tight heterodimer with Rat1 through interaction with its C-terminal region.

Our structure of Pol II Rat1–PTTC1 reveals the interactions among the termination factor Rat1–Rai1 heterodimer, the elongation factor Spt5, and Pol II (Fig. 2). Both the Rat1–Rai1 heterodimer and the Spt5 KOW5 domain localize around the RNA exit channel of Pol II (Fig. 2a). The Rat1–Rai1 heterodimer binds on top of the RNA exit channel by making interactions with the stalk, clamp, dock and wall domains of Pol II (Fig. 2a,b). The Spt5 KOW5 domain fills in a cleft between the dock and wall domains of Pol II (Fig. 2a). The termination and elongation factors of Pol II together shield the RNA exit channel and direct the route of nascent RNA after it exits from Pol II (Fig. 2a,b).
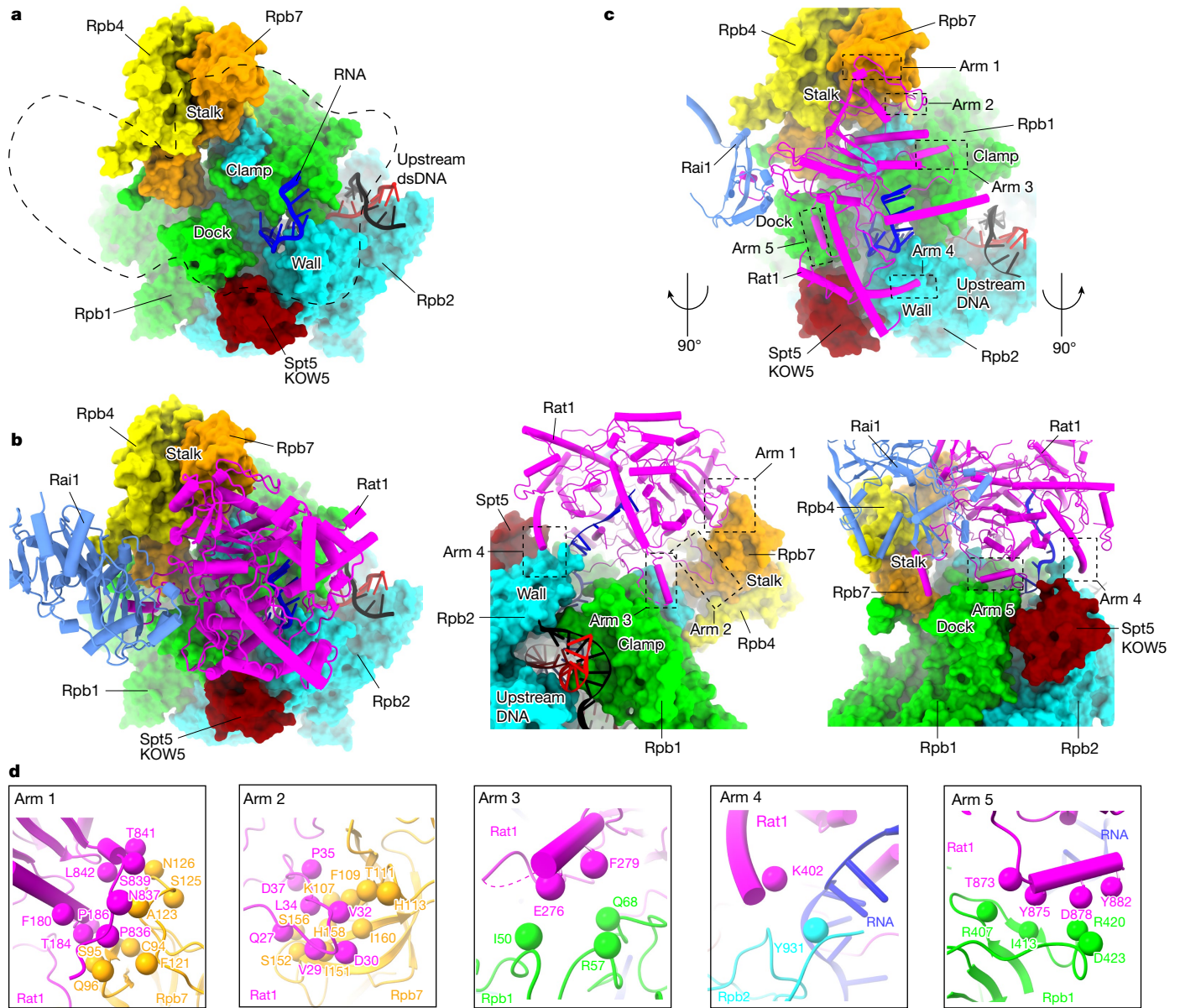
Fig. 2 | Protein interactions in Pol II Rat1–PTTC1. a, The RNA exit channel of Pol II. The dashed line represents the location of the Rat1–Rai1 complex. dsDNA, double-stranded DNA. **b**, The Rat1–Rai1 complex anchors on Pol II and covers the RNA exit channel of Pol II. **c**, Overview of the interface of Pol II and Rat1 in three view orientations. **d**, A detailed view of the interactions between Pol II and Rat1.

Rat1 of the Rat1–Rai1 heterodimer is mainly responsible for interactions with Pol II. Five arms of Rat1 make extensive contacts with the Rpb7 subunit of Pol II stalk, the dock and clamp domains of Rpb1, and the wall domain of Rpb2 (Fig. 2c,d and Extended Data Fig. 5), among which arm 1 and arm 2 of Rat1 form the largest interface with Pol II. Arm 1 and arm 2 anchor on the surface of the Rbp7 subunit of Pol II stalk. The interface residues probably include 15 residues of Rat1 (F180, T184, P186, P836, N837, S839, T841 and L842 in arm 1; Q27, V29, D30, V32, L34, P35 and D37 in arm 2) and 16 residues of Rpb7 (C94, S95, Q96, F121, A123, S125 and N126 interact with arm 1; K107, F109, T111, H113, I151, S152, S156, H158 and I160 interact with arm 2). The short helix of arm 3 contacts the clamp domain of Rpb1, and the interface residues probably include two residues of Rat1 (E276 and F279) and three residues of Rpb1 (I50, R57 and Q68). Arm 4 of Rat1 probably contacts residue Y931 of the Rpb2 wall domain through residue K402. The helix of arm 5 contacts the dock domain of Rbp1 through four residues of Rat1 (T873, Y875, D878 and Y882) and four residues of Rpb1 (R407, I413, R420 and D423).

Sequence alignment shows that about half of the residues on arms 1 and 2 of Rat1 are conserved in the functionally equivalent proteins in other eukaryotes (XRN3 in *Arabidopsis thaliana* and XRN2 in humans; Extended Data Fig. 6), consistent with the finding that arms 1 and 2 are the major dock site for the Rat1–Rai1 complex. The truncation of arms 1 and 2 impaired the interaction between Pol II and the Rat1–Rai1 complex (Extended Data Fig. 8c,d), a finding supporting our structure. In summary, our results show that the Rat1–Rai1 complex interacts extensively with Pol II, and the Pol II stalk domain is the main anchor site.

Our structure shows that Rat1 makes tight interactions with Pol II, raising the possibility that Rat1 can interact with Pol II independently of RNA. To test the hypothesis, we measured interactions between the recombinant *S. cerevisiae* Rat1–Rai1 complex and the *S. cerevisiae* Pol II core enzyme without DNA and RNA. The results of the size-exclusion chromatography show that the Rat1–Rai1 complex and Pol II were eluted in the same peak, and the result of the native polyacrylamide gel electrophoresis analysis indicates that the Rat1–Rai1 complex migrates with
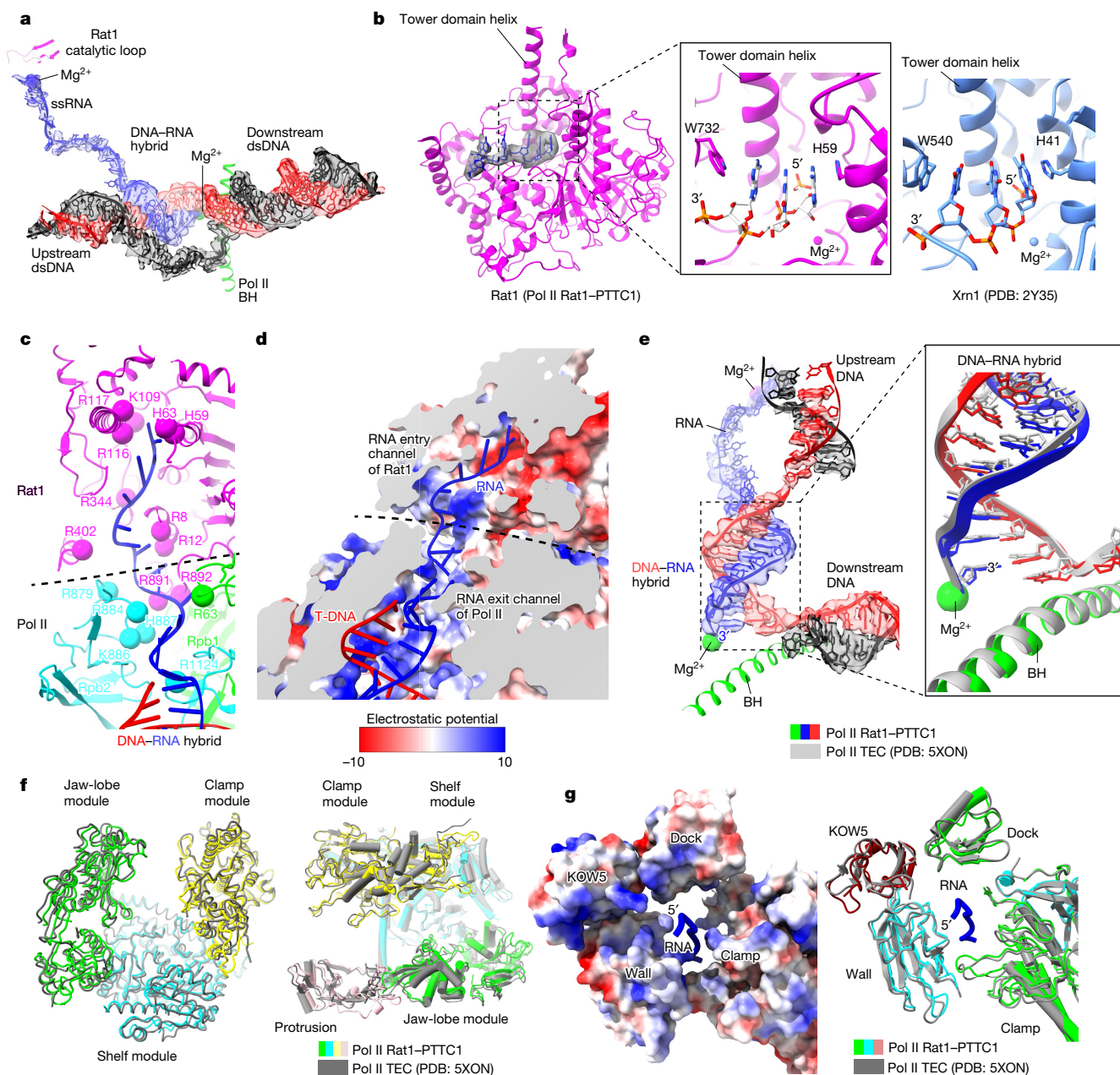
Fig. 3 | The RNA channel in Pol II Rat1–PTTC1. a, The cryo-EM map (map 1a) of the nucleic acid scaffold in the structure of Pol II Rat1–PTTC1. ssRNA, single-stranded RNA. **b**, Left: the Rat1 active site. RNA is stacked by W732 and H59 in Rat1. The cryo-EM density (map 1b) of RNA is coloured grey. Right: the Xrn1 active site. **c**, A detailed view of the interactions between the RNA and the polar residues in the RNA channel. **d**, The electrostatic potential surface of the RNA channel. **e**, Comparison of the DNA–RNA hybrid in the Pol II Rat1–PTTC1 structure (red and blue) with the post-translocated hybrid in Pol II TEC (Protein Data Bank (PDB): 5XON (ref. 32); grey). **f**, Conformational comparison between Pol II Rat1–PTTC1 and Pol II TEC in front and side views. **g**, The electrostatic potential surface (left) and a structural comparison (right) of the RNA exit channel between Pol II Rat1–PTTC1 and Pol II TEC.

Pol II (Extended Data Fig. 8a–c). The results support the hypothesis that Pol II can recruit the Rat1–Rai1 complex without prior interaction with RNA.

Structure superimposition between our structure of Pol II Rat1–PTTC1 and the Pol II transcription elongation complex (TEC) shows that Rat1 occupies the surface patches used for interaction with Spt6 and the KOW1–4 domains of Spt5 (Extended Data Fig. 7), suggesting that the interaction of Pol II termination factors is incompatible with the interaction of these Pol II elongation factors, consistent with the previous chromatin immunoprecipitation results showing that Spt6 dissociates after Pol II passes the PASs[18,27,28]. The cryo-EM map of Spt4-Spt5-bound

Pol II TEC (map 1c) shows unambiguous signals for Spt4 and Spt5 but not the Rat1–Rai1 complex (Extended Data Fig. 2j), indicating that Spt4 and Spt5 have to be displaced (or at least partially displaced) to allow interactions of the Rat1–Rai1 heterodimer with Pol II, explaining the reduced elongation rate of Pol II at the termination zone[24]. Our structure of Pol II Rat1–PTTC1 shows that the Spt5 KOW5 domain still attaches on the pre-termination complex (Fig. 1c,d), consistent with the retained chromatin immunoprecipitation signals of Spt5 after Pol II passes PASs[18,27,29], and supported by our pulldown result (Extended Data Fig. 8e, lane 8). Notably, the presence of Spt4 and Spt5 do not affect the amount of the Rat1–Rai1 complex bound to Pol II (Extended
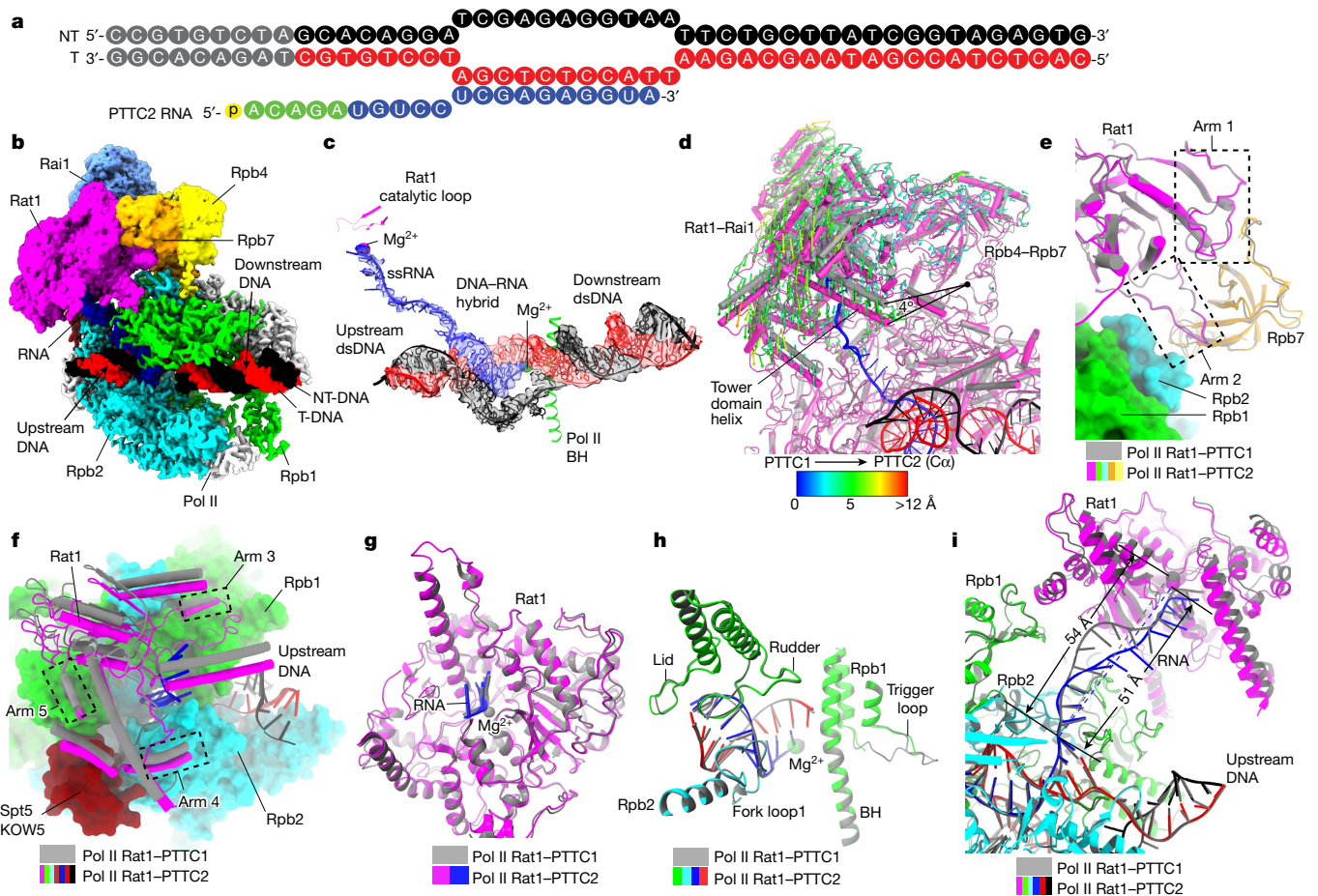
Fig. 4 | Rat1 rotates towards the RNA exit channel of Pol II as it shortens the nascent RNA. a, Sequences of the nucleic acid scaffold used in Pol II Rat1–PTTC2. Phosphorus-sulfur-bond modification of the RNA nucleotides is shown in green, and the mono-phosphorylation modification of the RNA 5′ terminus is shown in yellow. Nucleotides in grey are disordered in the structure. **b**, The cryo-EM map (map 2a) of Pol II Rat1–PTTC2. **c**, The cryo-EM map (map 2a) of the nucleic acid scaffold in the structure of Pol II Rat1–PTTC2. **d**, Rat1–Rai1–stalk in Pol II Rat1–PTTC2 rotates about 4° towards the RNA exit channel compared with that in Pol II Rat1–PTTC1. **e**, Superimposition of the two Pol II Rat1–PTTC structures suggests that the interaction of arms 1 and 2 with the stalk is retained during rotation. **f**, Arms 3, 4 and 5 slightly slide on the Pol II surface following the rotation of the Rat1–Rai1–stalk unit. **g**, Structure superimposition of the Rat1 active site of the two structures. **h**, Structure superimposition of the Pol II active site of the two structures. **i**, Structure superimposition of the semi-closed RNA channel of the two structures.

Data Fig. 8e, lanes 5 versus 8), which could be explained by the much stronger binding of the Rat1–Rai1 complex with Pol II than the Spt4-Spt5 complex (ΔKOW5-C-terminal region; Δ755–1063) (lanes 5 versus 7), an Spt4-Spt5 derivative containing the NGN and KOW1–4 domains of Spt5 that shows little interaction with Pol II under the same experimental conditions (lane 7). In summary, our results show that the Rat1–Rai1 complex competes with the elongation factors for binding to Pol II.

Rat1 guides RNA into its active site

The interactions among Pol II, Rat1 and Spt5 KOW5 connect the RNA exit channel of Pol II and the RNA entry channel of Rat1 (Fig. 3). In the semi-closed continuous channel, the nascent RNA exits from the RNA exit channel of Pol II and enters the active site of Rat1 (Fig. 3a), where three ribonucleotides of the RNA 5′ terminus are stacked between two aromatic residues of Rat1 (W732 and H59; Fig. 3b). A similar substrate interaction has also been observed in the crystal structure of a binary complex of Xrn1 and single-stranded DNA[30] (Fig. 3b). The 5′-phosphate group is stabilized by K109, Q113, R116 and R117 (Extended Data Fig. 9a). A triphosphate group would introduce severe steric clash in the active site, explaining the preference of RNA substrate with the mono-phosphate group at its 5′ terminus[31] (Extended Data Fig. 9c).

The phosphate backbone between the first and the second ribonucleotide is exposed to the catalytic loop of Rat1 (Fig. 3b and Extended Data Fig. 9a). The polar residues (residue R63 of the Pol II Rbp1 subunit, residues R879, R884, K886, H887 and R1124 of the Pol II Rpb2 subunit, and residues R8, R12, R344, R402, R891 and R892 of Rat1) along the channel probably define a charge-complementary path directing the nascent RNA from the Pol II exit channel to the Rat1 active site by making direct interactions with its phosphate backbones (Fig. 3c,d and Extended Data Fig. 9b).

Structural superimposition between our Pol II Rat1–PTTC1 structure and the previously reported Pol II TEC suggests similar conformations of Pol II, DNA and RNA (Fig. 3e–g). The DNA–RNA hybrid in the structure of Pol II Rat1–PTTC1 adopts a post-translocation state (Fig. 3e); the clamp, jaw-lobe, protrusion and shelf modules around the main cleft in the structure of Pol II Rat1–PTTC1 adopt a closed conformation as observed in Pol II TEC[32] (Fig. 3f). The interactions of the Rat1–Rai1 complex with Pol II do not induce a noticeable conformational change of the clamp, wall and dock domains around the RNA exit channel of Pol II (Fig. 3g). The structural comparison indicates that the exchange of termination and elongation factor does not induce a conformational change of Pol II, arguing against the allosteric model of Rat1-mediated transcription termination.
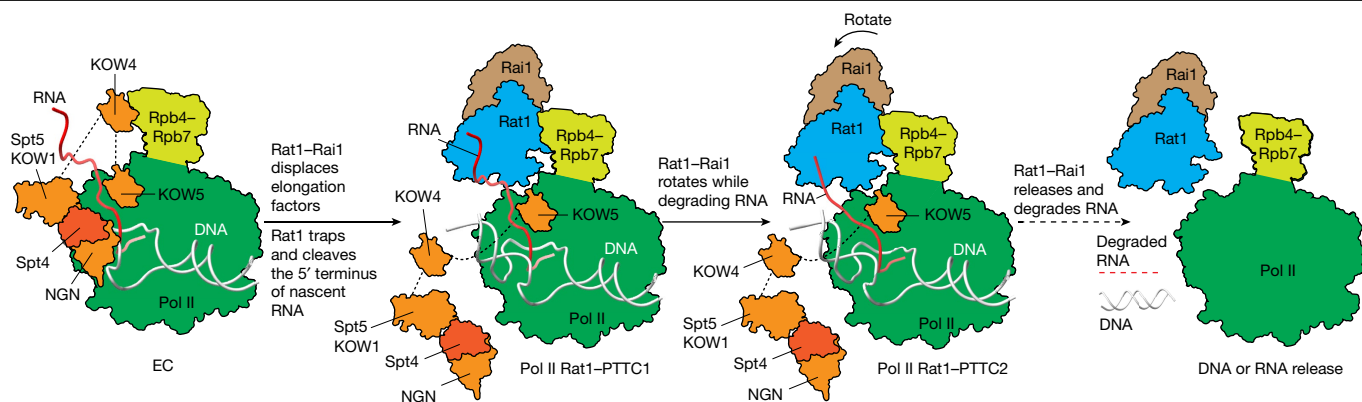
Fig. 5 | The proposed model of Rat1–Rai1-mediated Pol II termination. The Rat1–Rai1 complex engages with Pol II along with the dissociation of Pol II elongation factors and traps the 5′ terminus of the nascent RNA in its active site of Rat1 (Pol II Rat1–PTTC1). Subsequently, Rat1 cleaves the nascent RNA in a 5′-to-3′ direction and pulls the RNA towards its active centre. The cleavage stretches the RNA in the semi-closed RNA channel and causes rotation of Rat1 towards Pol II (Pol II Rat1–PTTC2). Rat1 continues cleaving the RNA in the 5′-to-3′ direction, pulls the RNA out from the Pol II active centre and digests the released RNA. EC, elongation complex.

Detailed analysis suggests that the RNA in the semi-closed channel of Pol II Rat1–PTTC1 adopts a relaxed conformation, in which the binding of ribonucleotides at the 5′ terminus of the nascent RNA in the Rat1 active site does not affect the binding of ribonucleotides at the 3′ terminus of the nascent RNA to the active site of Pol II, where the DNA–RNA hybrid adopts the post-translocated state (Fig. 3e). We predict that further cleavage of the nascent RNA in the 5′-to-3′ direction by Rat1 would pull ribonucleotides towards its active site, stretch the RNA in the semi-closed channel and apply mechanical force to the pre-termination complex of Pol II.

Rat1 rotates to Pol II during RNA cleaving

To investigate how Pol II pre-termination complex accommodates nascent RNA further shortened by Rat1, we reconstituted Pol II Rat1–PTTC2 with a nascent RNA of 20 nt in length, 3 nt shorter than that in Pol II Rat1–PTTC1 (Fig. 4a, Extended Data Fig. 1c–d and Extended Data Fig. 3). The cryo-EM map (map 2a) shows clear signals for Rat1, Rai1 and Spt5 KOW5 that surround the RNA exit channel of Pol II (Fig. 4b). To improve the resolution of Rat1–Rai1 and the nascent RNA, we obtained a local map focusing around the Rat1–Rai1 heterodimer at 3.4 Å (map 2b; Extended Data Fig. 3c). The overall and local cryo-EM maps (maps 2a and 2b) show clear signal for 10-nt RNA in the active-site cleft at the post-translocated state, a 7-nt RNA in the semi-closed extended RNA channel, and 3-nt RNA in the active site of Rat1 (Fig. 4c and Extended Data Fig. 3f). Structure superimposition between Pol II Rat1–PTTC2 and Pol II Rat1–PTTC1 shows that the Rat1–Rai1 complex together with Pol II stalk as a structure module rotates about 4° along an axis perpendicular to the plane of the transcription bubble (Fig. 4d). The rotation retains the interactions of arms 1 and 2 with Pol II stalk (Fig. 4e), but slides on the surface where arms 3, 4 and 5 make contact, consistent with the notion that the stalk is the primary dock point for Rat1 (Fig. 4f).

The rotation of the Rat1–Rai1–stalk module is attributed to the shortened RNA length. A detailed comparison shows that, in both structures, the three ribonucleotides at the 5′ terminus of nascent RNA are secured in the active site of Rat1 and make similar interactions with Rat1 (Fig. 4g) and that the ten nucleotides at the 3′ terminus of the nascent RNA are secured in the active site of Pol II adopting the post-translocated state (Fig. 4h). Therefore, to retain the above interactions with the shortened RNA, the Rat1–Rai1–stalk module rotates towards the RNA exit channel, which reduces the distance between the active site of Rat1 and the RNA exit channel of Pol II from 54 Å in Pol II Rat1–PTTC1 to 51 Å in Pol II Rat1–PTTC2 to accommodate the shortened linker region of the nascent RNA (Fig. 4i).

The structure of Pol II Rat1–PTTC2 shows that the linker RNA in the semi-closed channel is in a stretched state, and the Rat1–Rai1 complex is tightly engaged with Pol II (Fig. 4i). We predict that further cleavage and subsequent translocation of the nascent RNA in the 5′-to-3′ direction by Rat1 would interrupt interactions between ribonucleotides in the RNA exit channel and apply mechanical force to the DNA–RNA hybrid, leading to disruption of the DNA–RNA hybrid and RNA release.

Discussion

Our work supports the unified sitting duck torpedo model of the Rat1-mediated transcription termination[22,23]. We show that the Rat1–Rai1 complex is incompatible with the elongation factors for Pol II binding, including Spt5 and Spt6, and therefore, the recruitment of Rat1–Rai1 would cause dissociation of the elongation factors, explaining the reduced elongation rate of Pol II at the termination zone in the sitting duck torpedo model of Pol II termination[7–9,24,33,34]. However, the exchange of termination/elongation factor does not cause a conformational change of Pol II, arguing against the allosteric mechanism of Rat1-mediated transcription termination.

Furthermore, our structures show that Rat1 interacts extensively with Pol II by anchoring on the stalk domain and nascent RNA by trapping three nucleotides of the RNA 5′ terminus in its active site. The structures further show and that Rat1–Rai1–Rbp4–Rpb7 rotates as a structure unit towards the RNA exit channel as it shortens the length of nascent RNA. We predict that further cleavage of RNA in the 5′-to-3′ direction would generate mechanical force to pull the RNA out from the Pol II active centre and eventually release it from Pol II, highlighting the essential role of the Pol II interaction and exonuclease activity of the Rat1 torpedo in the sitting duck torpedo model[1,2].

Our work suggests that the Rat1-mediated transcription termination pathway probably involves three stages—the recruitment stage, the engagement stage and the termination stage. In the recruitment stage, the Rat1–Rai1 complex is first recruited to the Pol II transcription complex by Rtt103 recognizing the Pol II Ser2P C-terminal domain after Pol II passes through the PAS[1,12]. In the engagement stage, the Rat1–Rai1 complex engages with Pol II along with dissociation of Pol II elongation factors, and the 5′ terminus of the cleaved nascent RNA is trapped into the active site of Rat1[1,2,24,33,34]. Once the Rat1–Rai1 complex is engaged with Pol II, a race occurs between Rat1, which shortens the 5′ terminus of RNA, and Pol II, which extends the 3′ terminus of RNA[1,2,4,7,8]. In the termination stage, the cleavage and translocation of Rat1 pull RNA towards itself and release the RNA from Pol II. Our structures of Pol II Rat1–PTTC1 and Rat1–PTTC2 probably represent two intermediate

states at the termination stage of Pol II transcription termination before RNA release (Fig. 5). We propose that any factors impeding Pol II elongation would increase the probability of RNA release, consistent with previous in vivo and in vitro evidence showing that the RNA release is facilitated by the physical obstacle of a nucleosome on the path[35], the RNA sequence-induced Pol II pausing or backtracking[36,37] and NTP misincorporation[38].

By comparing Rat1-dependent termination of Pol II mRNA transcription in eukaryotes, Rho-dependent termination of bacterial transcription[39,40], FttA-dependent termination of archaeal transcription[41] and integrator-initiated termination of premature mRNA[42,43], we show that all of the RNA-associating termination motor proteins mentioned above anchor near the RNA exit channel, make physical interactions with RNA polymerase and apply mechanical force on the DNA–RNA hybrid to release RNA. The similarity suggests a general principle of transcription termination by multiple-subunit RNA polymerases. Despite the similarity, the termination pathways differ in the working mode of the motor proteins, the requirement of elongation factors and the fate of released RNA.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-024-07240-3.

1. Kim, M. et al. The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature* **432**, 517–522 (2004).
2. West, S., Gromak, N. & Proudfoot, N. J. Human 5′ -> 3′ exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. *Nature* **432**, 522–525 (2004).
3. Luo, W. & Bentley, D. A ribonucleolytic rat torpedoes RNA polymerase II. *Cell* **119**, 911–914 (2004).
4. Fong, N. et al. Effects of transcription elongation rate and Xrn2 exonuclease activity on RNA polymerase II termination suggest widespread kinetic competition. *Mol. Cell* **60**, 256–267 (2015).
5. Krzyszton, M. et al. Defective XRN3-mediated transcription termination in *Arabidopsis* affects the expression of protein-coding genes. *Plant J.* **93**, 1017–1031 (2018).
6. Proudfoot, N. J. Transcriptional termination in mammals: stopping the RNA polymerase II juggernaut. *Science* **352**, aad9926 (2016).
7. Girbig, M., Misiaszek, A. D. & Muller, C. W. Structural insights into nuclear transcription by eukaryotic DNA-dependent RNA polymerases. *Nat. Rev. Mol. Cell Biol.* **23**, 603–622 (2022).
8. Porrua, O. & Libri, D. Transcription termination and the control of the transcriptome: why, where and how to stop. *Nat. Rev. Mol. Cell Biol.* **16**, 190–202 (2015).
9. Eaton, J. D. & West, S. Termination of transcription by RNA polymerase II: BOOM! *Trends Genet.* **36**, 664–675 (2020).
10. Richard, P. & Manley, J. L. Transcription termination by nuclear RNA polymerases. *Genes Dev.* **23**, 1247–1269 (2009).
11. Kumar, A., Clerici, M., Muckenfuss, L. M., Passmore, L. A. & Jinek, M. Mechanistic insights into mRNA 3′-end processing. *Curr. Opin. Struc. Biol.* **59**, 143–150 (2019).
12. Lunde, B. M. et al. Cooperative interaction of transcription termination factors with the RNA polymerase II C-terminal domain. *Nat. Struct. Mol. Biol.* **17**, 1195–1201 (2010).
13. Carminati, M. et al. A direct interaction between CPF and RNA Pol II links RNA 3′ end processing to transcription. *Mol. Cell* **83**, 4461–4478 (2023).
14. Brannan, K. et al. mRNA decapping factors and the exonuclease Xrn2 function in widespread premature termination of RNA polymerase II transcription. *Mol. Cell* **46**, 311–324 (2012).
15. Nojima, T. et al. Mammalian NET-Seq reveals genome-wide nascent transcription coupled to RNA processing. *Cell* **161**, 526–540 (2015).
16. Jimeno-Gonzalez, S., Haaning, L. L., Malagon, F. & Jensen, T. H. The yeast 5′-3′ exonuclease Rat1p functions during transcription elongation by RNA polymerase II. *Mol. Cell* **37**, 580–587 (2010).
17. Cortazar, M. A. et al. Xrn2 substrate mapping identifies torpedo loading sites and extensive premature termination of RNA pol II transcription. *Genes Dev.* **36**, 1062–1078 (2022).
18. Kim, M., Ahn, S. H., Krogan, N. J., Greenblatt, J. F. & Buratowski, S. Transitions in RNA polymerase II elongation complexes at the 3′ ends of genes. *EMBO J.* **23**, 354–364 (2004).
19. Zhang, Z., Fu, J. & Gilmour, D. S. CTD-dependent dismantling of the RNA polymerase II elongation complex by the pre-mRNA 3′-end processing factor, Pcf11. *Genes Dev.* **19**, 1572–1580 (2005).
20. Zhang, H., Rigo, F. & Martinson, H. G. Poly(A) signal-dependent transcription termination occurs through a conformational change mechanism that does not require cleavage at the poly(A) site. *Mol. Cell* **59**, 437–448 (2015).
21. Logan, J., Falck-Pedersen, E., Darnell, J. E. Jr & Shenk, T. A poly(A) addition site and a downstream termination region are required for efficient cessation of transcription by RNA polymerase II in the mouse beta maj-globin gene. *Proc. Natl Acad. Sci. USA* **84**, 8306–8310 (1987).
22. Eaton, J. D., Francis, L., Davidson, L. & West, S. A unified allosteric/torpedo mechanism for transcriptional termination on human protein-coding genes. *Genes Dev.* **34**, 132–145 (2020).
23. Luo, W., Johnson, A. W. & Bentley, D. L. The role of Rat1 in coupling mRNA 3′-end processing to transcription termination: implications for a unified allosteric–torpedo model. *Genes Dev.* **20**, 954–965 (2006).
24. Cortazar, M. A. et al. Control of RNA Pol II speed by PNUTS-PP1 and Spt5 dephosphorylation facilitates termination by a "sitting duck torpedo" mechanism. *Mol. Cell* **76**, 896–908 e894 (2019).
25. Xue, Y. et al. *Saccharomyces cerevisiae* RAI1 (YGL246c) is homologous to human *DOM3Z* and encodes a protein that binds the nuclear exoribonuclease Rat1p. *Mol. Cell. Biol.* **20**, 4006–4015 (2000).
26. Xiang, S. et al. Structure and function of the 5′→3′ exoribonuclease Rat1 and its activating partner Rai1. *Nature* **458**, 784–788 (2009).
27. Mayer, A. et al. Uniform transitions of the general RNA polymerase II transcription complex. *Nat. Struct. Mol. Biol.* **17**, 1272–1278 (2010).
28. Narain, A. et al. Targeted protein degradation reveals a direct role of SPT6 in RNAPII elongation and termination. *Mol. Cell* **81**, 3110–3127 (2021).
29. Baejen, C. et al. Genome-wide analysis of RNA polymerase II termination at protein-coding genes. *Mol. Cell* **66**, 38–49 (2017).
30. Jinek, M., Coyle, S. M. & Doudna, J. A. Coupled 5′ nucleotide recognition and processivity in Xrn1-mediated mRNA decay. *Mol. Cell* **41**, 600–608 (2011).
31. Stevens, A. & Poole, T. L. 5′-exonuclease-2 of *Saccharomyces cerevisiae*. Purification and features of ribonuclease activity with comparison to 5′-exonuclease-1. *J. Biol. Chem.* **270**, 16063–16069 (1995).
32. Ehara, H. et al. Structure of the complete elongation complex of RNA polymerase II with basal factors. *Science* **357**, 921–924 (2017).
33. Parua, P. K. et al. A Cdk9–PP1 switch regulates the elongation–termination transition of RNA polymerase II. *Nature* **558**, 460–464 (2018).
34. Kecman, T. et al. Elongation/termination factor exchange mediated by PP1 phosphatase orchestrates transcription termination. *Cell Rep.* **25**, 259–269 (2018).
35. Grosso, A. R., de Almeida, S. F., Braga, J. & Carmo-Fonseca, M. Dynamic transitions in RNA polymerase II density profiles during transcription termination. *Genome Res.* **22**, 1447–1456 (2012).
36. Orozco, I. J., Kim, S. J. & Martinson, H. G. The poly(A) signal, without the assistance of any downstream element, directs RNA polymerase II to pause in vivo and then to release stochastically from the template. *J. Biol. Chem.* **277**, 42899–42911 (2002).
37. Gromak, N., West, S. & Proudfoot, N. J. Pause sites promote transcriptional termination of mammalian RNA polymerase II. *Mol. Cell. Biol.* **26**, 3986–3996 (2006).
38. James, K., Gamba, P., Cockell, S. J. & Zenkin, N. Misincorporation by RNA polymerase is a major source of transcription pausing in vivo. *Nucleic Acids Res.* **45**, 1105–1113 (2017).
39. Molodtsov, V., Wang, C., Firlar, E., Kaelber, J. T. & Ebright, R. H. Structural basis of Rho-dependent transcription termination. *Nature* **14**, 367–374 (2023).
40. Murayama, Y. et al. Structural basis of the transcription termination factor Rho engagement with transcribing RNA polymerase from *Thermus thermophilus*. *Sci. Adv.* **9**, eade7093 (2023).
41. Chengyuan, W. et al. Structural basis of archaeal FttA-dependent transcription termination. Preprint at *bioRxiv* https://doi.org/10.1101/2023.08.09.552649 (2023).
42. Fianu, I. et al. Structural basis of integrator-mediated transcription regulation. *Science* **374**, 883–887 (2021).
43. Zheng, H. et al. Structural basis of INTAC-regulated transcription. *Protein Cell* **14**, 698–702 (2023).

Article

Methods

Cloning and protein expression

DNA fragments encoding full-length *S. cerevisiae* Rat1, Rai1, Spt4, Spt5, Rpb4 and Rpb7 were amplified from *S. cerevisiae* genomic DNA using Phanta Max Super-fidelity DNA Polymerase (Vazyme Biotech). Rat1 with Rai1, Spt4 with Spt5, and Rpb4 with Rpb7 were cloned into the pETDuet1 vector, respectively, by a homologous recombination method (NovoRec Plus One Step PCR Cloning Kit; Novoprotein). The pETDuet1-Rat1(Δ1/2; Δ23–43/823–843)-Rai1 and pETDuet1-Spt4-Spt5(ΔKOW5-CTR; Δ755–1063) were constructed by site-directed mutagenesis (Transgene Biotech). The expression of the Rat1-Rai1 complex, Rat1(Δ1/2; Δ23–43/823–843)-Rai1 complex, Spt4-Spt5 complex, Spt4-Spt5(ΔKOW5-CTR; Δ755–1063) complex, and Rpb4-Rpb7 complex was induced by 0.4 mM isopropyl β-D-1-thiogalactopyanoside at 18 °C for 16 h in BL21 (DE3) codon plus cells. The cells were collected by centrifugation (3 min, 16 °C, 8,000 r.p.m.) and stored at −80 °C until protein purification.

A modified *S. cerevisiae* BJ2168 strain for expressing endogenous Pol II with an affinity tag at the C terminus of Rpb2 was generated as described previously[44]. Briefly, a DNA fragment encoding the last 15 residues of Rpb2 followed by an affinity tag (TEV-6×His-3×Flag) was inserted into the pESC-his vector through homologous recombination. The target DNA fragment amplified from the above vector was transformed into *S. cerevisiae* BJ2168. Positive colonies were screened by SD-URA medium, and protein expression was detected by western blot. The yeast cells were cultured in YPDA medium, collected at an optical density of 18 at 600 nm by centrifugation (3 min, 16 °C, 8,000 r.p.m.), and stored at −80 °C.

Protein purification

The cell pellets of wild-type and derivatives of the Rat1-Rai1 complex, Rpb4-Rpb7 and Spt4-Spt5 were suspended in lysis buffer (50 mM Tris-HCl pH 7.7, 0.2 M NaCl, 5% (v/v) glycerol, 5 mM β-mercaptoethanol), and lysed using an Avestin EmulsiFlex-C3 cell disrupter. The lysate was centrifuged at 16,000g for 50 min at 4 °C, and the supernatant was loaded onto a nickel nitrilotriacetic acid (Ni-NTA) affinity column (Smart-Lifesciences). The column was washed with lysis buffer supplemented with 20 mM imidazole and eluted with lysis buffer supplemented with 300 mM imidazole. For the Rat1–Rai1 complex, the protein sample was further purified by a Heparin column (HiTrap Heparin HP 5-ml column, Cytiva) using a salt gradient of buffer A (20 mM Tris-HCl pH 7.7, 0.1 M NaCl, 5% (v/v) glycerol, 1 mM dithiothreitol (DTT)) and buffer B (20 mM Tris-HCl pH 7.7, 1 M NaCl, 5% (v/v) glycerol, 1 mM DTT). The eluted fractions containing the target protein were diluted with buffer A, loaded onto a Q HP column (HiPrep Q HP 16/10, Cytiva), and eluted with a salt gradient of buffer A and buffer B. The eluted fractions containing target protein were collected and loaded onto a Superdex 200 10/300 GL column (Cytiva) in S200 buffer C (20 mM Tris-HCl pH 7.7, 0.1 M NaCl, 5% (v/v) glycerol, 1 mM DTT). Fractions containing the Rat1–Rai1 complex were collected, concentrated to 10 mg ml^{-1} and stored at −80 °C. For the Spt4-Spt5 and Rpb4-Rpb7 complexes, after Ni-NTA purification, the protein sample was purified by a Q HP column (as above) with a salt gradient of buffer A and buffer B. The fractions containing Spt4-Spt5 and Rpb4-Rpb7 were collected and loaded onto a Superdex 200 10/300 GL column (Cytiva) in S200 buffer C, and target proteins were concentrated at 3 mg ml^{-1} and stored at −80 °C.

For Pol II purification, cell pellets were suspended using lysis buffer (100 mM Tris-HCl pH 7.5, 0.25 M ammonium sulfate, 5% (v/v) glycerol, 1 mM EDTA, 2 mM DTT, 0.01 mM ZnCl$_2$, 0.1 mM phenylmethylsulfonyl fluoride and protease inhibitor cocktail (Biomake)) and lysed with an EmulsiFlex-C3 cell disrupter (Avestin). The lysate was centrifuged at 16,000g for 60 min at 4 °C; then the supernatant was precipitated with ammonium sulfate (30 g per 100 ml) and stirred at 4 °C overnight. The pellet was collected and dissolved with TEZ buffer (50 mM Tris-HCl

pH 7.5, 50 mM ammonium sulfate, 5% (v/v) glycerol, 2 mM DTT). Flag resin (Genscript) was subsequently added to the solution, followed by stirring at 4 °C for 3 h. Flag resin with target protein was collected, washed with lysis buffer and then eluted with 500 µg ml^{-1} Flag peptide (Genscript). The eluted protein was loaded onto a Mono Q column (Mono Q 10/100 GL, Cytiva) followed by a salt gradient of Mono Q buffer A (50 mM Tris-HCl pH 7.5, 50 mM NaCl, 5% glycerol, 2 mM DTT) and Mono Q buffer B (50 mM Tris pH 7.5, 1 M NaCl, 5% (v/v) glycerol, 2 mM DTT). Fractions containing target protein were collected and loaded onto a Superose 6 10/300 GL column (Cytiva) in S6 buffer C (20 mM Tris-HCl pH 7.7, 0.1 M NaCl, 5% (v/v) glycerol, 1 mM DTT). Fractions containing Pol II were collected, concentrated to 6 mg ml^{-1} and stored at −80 °C.

Gel-filtration assay

Complexes (500 µl) containing 1 µM Pol II and 4 µM Rat1-Rai1 were incubated for 30 min at room temperature in complex buffer B (20 mM HEPES pH 7.4, 0.1 M NaCl, 1 mM DTT). The sample was applied to a Superose 6 Increase 10/300 GL column (Cytiva) equilibrated with complex buffer B. Peak fractions were analysed by SDS–PAGE followed by Coomassie staining.

Gel-shift assay

Complexes (10 µl) contained 0.2 µM Pol II with a gradient (0.2–3.2 µM) of the wild-type Rat1–Rai1 complex or the Rat1(Δ1/2; Δ23–43/823–843)–Rai1 derivative was incubated for 60 min at room temperate in complex buffer A (20 mM Tris-HCl pH 7.7, 0.1 M KCl, 5% (v/v) glycerol, 3 mM MgCl$_2$, 1 mM DTT). The complexes were separated by 5% Tris-glycine gel in the Tris-glycine buffer (Sangon Biotech), followed by Coomassie staining.

Pol II release assay

The sequences of the nucleic acid scaffold used in the Pol II release assay are: 5′-CATGCGTACACGCGTTTGTACAGAAAAAAAAGAAA AATCCTGGTCGTTTCCTATA-3′ (template strand), 5′-TATAGGAAA CGACCAGGATTTTTCTTTTTTTTTCTGTACAAACGCGTGTACGCATGT AAC-biotin-3′ (non-template strand), 5′-p-rUrUrUrUrUrGrArCrCArGrGr ArUrU-3′ (RNA; RNA was mono-phosphorylated at the 5′ terminus). Template-strand DNA and RNA were mixed in a 1:1.2 molar ratio and annealed by 95 °C, 5 min followed by 2 °C-step cooling to 25 °C in annealing buffer (5 mM Tris-HCl pH 8.0, 200 mM NaCl). All following steps were carried out at ambient temperature unless otherwise indicated. The reactions were carried out in transcription buffer (TB; 20 mM Tris-HCl pH 8.0, 100 mM NaCl, 2.5% glycerol, 5 mM MgCl$_2$, 1 mM DTT). Pol II (50 nM) was mixed with template DNA–RNA scaffold in a 1:2 molar ratio and incubated for 20 min with gentle shaking. Next, a fourfold molar excess of non-template DNA was added and incubated for 20 min with gentle shaking to form ECs. Streptavidin beads (Smart-Lifesciences) pre-washed three times in TB with 100 µg ml^{-1} BSA were added into the reaction system and incubated for 30 min with gentle shaking. Then, the EC-bound beads were sequentially washed with TB + 100 µg ml^{-1} BSA, TB + 0.1% Triton X-100, TB + 0.5 M NaCl and TB to remove the unbound Pol II, DNA and RNA. The EC-bound beads were then resuspended in TB. A subset of NTP (100 µM UTP and 100 µM CTP) or a combination of the Rat1–Rai1 complex (1,200 nM) and a subset of NTP (100 µM UTP and 100 µM CTP) was subsequently incubated with the EC-bound beads for 1 h with gentle shaking at 30 °C. The 'total' (bead-bound and released Pol II) and 'supernatant' (released Pol II) fractions were analysed by 8% SDS–PAGE. Protein gel was transferred to PVDF membranes, and immunoblotted for Pol II with the anti-Flag antibody (YEASEN Biotechnology).

RNA cleavage assay

The sequence of the RNA used in the cleavage assay is: 5′-p-rArCrAr GrArUrCrGrU*rG*rUrCrCrArUrCrGrArGrArGrGrU-3′ (RNA was

mono-phosphorylated at the 5′ terminus; the asterisks represent phosphorus–sulfur-bond modifications to prevent Rat1 cleavage). The reaction mixture (20 µl) contained 1 µM RNA and 2 µM Rat1-Rai1 were incubated for 1 h at 30 °C in reaction buffer (20 mM Tris-HCl pH 8.0, 0.1 M KCl, 5% (v/v) glycerol, 3 mM MgCl₂, 1 mM DTT). The reaction was stopped by stopping buffer (1× TBE, 8 M urea) and boiling for 5 min at 95 °C. The sample was analysed with 20% urea gel in the TBE buffer (90 mM Tris-borate, pH 8.0, and 2 mM EDTA), and stained by SYBR Gold Nucleic Acid Gel Stain (Invitrogen).

Pulldown assay

The reaction mixtures (20 µl) containing Flag-tagged Pol II (0.8 µM) and 1.2 µM DNA–RNA scaffold (the same scaffold used in the Pol II Rat1–PTTC1 cryo-EM study) in binding buffer (20 mM Hepes pH 7.5, 100 mM NaCl, 5% glycerol, 2 mM DTT, 50 µg ml⁻¹ BSA) were incubated for 30 min. Subsequently, the Spt4-Spt5 complex (2.4 µM) or the Spt4-Spt5(ΔKOW5-CTR; Δ755–1063; 2.4 µM) complex were added and incubated for 30 min. The mixtures were then incubated with the Rat1–Rai1 complex (2.4 µM) for 30 min. All steps were carried out at ambient temperature unless otherwise indicated. Anti-Flag affinity resin (Genscript) was added to reaction mixtures and incubated at 30 °C for 1 h with shaking. After incubation, the anti-Flag affinity resin was washed four times with the wash buffer (20 mM Hepes pH 7.5, 100 mM NaCl, 0.05% Triton X-100, 2 mM DTT), and eluted with 500 µg ml⁻¹ Flag peptide. The eluted fractions were mixed with SDS loading buffer and subjected to electrophoresis at 150 V for 70 min on a 4–20% SDS–PAGE gel. Gels were analysed by Coomassie blue staining.

Cryo-EM sample preparation and image processing

The sequences of the nucleic acid scaffold used in Pol II Rat1–PTTC1 are: 5′-CACTCTACCGATAAGCAGACATACCTCTCGATCCTGTGC TAGACACGG-3′ (template strand), 5′-CCGTGTCTAGCACAGGTAA ATGGTTTGTGTCTGCTTATCGGTAGAGTG-3′ (non-template strand), 5′-p-rA*rC*rA*rG*rArUrCrGrUrGrUrCrCrArUrCrGrArGrArGrGrU-3′ (RNA; RNA was mono-phosphorylated at the 5′ terminus; the asterisks represent phosphorus–sulfur-bond modifications to prevent Rat1 cleavage).

The sequences of the nucleic acid scaffold used in Pol II Rat1–PTTC2 are: 5′-CACTCTACCGATAAGCAGAATTACCTCTCGATCCTG TGCTAGACACGC-3′ (template strand), 5′-GCGTGTCTAGCACAG GATCGAGAGGTAATTCTGCTTATCGGTAGAGTG-3′ (non-template strand), 5′-p-rA*rC*rA*rG*rArUrGrUrCrCrUrCrGrArGrArGrGrUrA-3′ (RNA; RNA was mono-phosphorylated at the 5′ terminus; the asterisks represent phosphorus–sulfur-bond modifications to prevent Rat1 cleavage).

Unless otherwise indicated, all of the following steps were carried out in EM buffer (20 mM Hepes pH 7.4, 100 mM KCl, 3 mM MgCl₂, 2 mM DTT). Pol II and the nucleic acid scaffold (Pol II TEC) were assembled at a molar ratio of 1:1.25, incubated at 4 °C for 1 h, and then purified by a Superose 6 10/300 GL column (Cytiva) equilibrated in EM buffer. Fractions containing Pol II TEC were collected, and then Rpb4-Rpb7 (a molar ratio of 5:1 with Pol II) was added and incubated at 4 °C for 30 min, Rat1-Rai1 (a molar ratio of 3:1 with Pol II) and Spt4-Spt5 (a molar ratio of 3:1 with Pol II) were added and incubated at 4 °C for 1 h. The complex was again purified by a Superose 6 10/300 GL column (Cytiva). Fractions containing the target proteins were collected, concentrated to 0.8 mg ml⁻¹, and then crosslinked with 0.15% glutaraldehyde on ice for 10 min. Subsequently, the crosslinking reactions were quenched with 100 mM Tris pH 8.0 and dialysed overnight in EM buffer. The sample was concentrated at 1.5 mg ml⁻¹.

The freshly purified sample was mixed with 0.05% Tween20 (final concentration) before grid preparation. A 3-µl sample was applied to glow-discharged UltraAuFoil R1.2/1.3 300 mesh holey Au grids (Quantifoil Micro Tools), and the grid was plunge-frozen in liquid ethane using a Vitrobot Mark IV (FEI) with 100% chamber humidity at 22 °C.

Cryo-EM imaging was carried out on an FEI Titan Krios equipped with a Gatan K3 Summit direct electron detector at the National Center for Protein Sciences Shanghai. Data were collected at a nominal magnification of 64,000 (1.1 Å per pixel) with a dose rate of 22.5 electrons per pixel per second on the sample.

For Pol II Rat1–PTTC1, 7,327 images were recorded using EPU software in counting mode (2.67 exposures in 40 subframes resulting in a total dose of 50 electrons per square ångström, defocus range of −1.0 to −1.8 µm). Frames of individual videos were aligned using MotionCor2 (ref. 45), and contrast transfer function estimations were carried out using CTFFIND[46]. About 1,000 particles were manually picked and subjected to two-dimensional (2D) classification in RELION 3.0 (ref. 47). The resulting distinct 2D classes served as templates for autopick. Protein aggregates and ice spots were removed, and the resulting particles were subjected to 2D classification in RELION 3.0 by specifying 50 classes. Poorly populated types were removed. We used a 60-Å low-pass-filtered map from our test data as the starting reference model for 3D classification ($n = 6$). Classes were combined, and 416,223 single particles were used for constructing the final cryo-EM map for Pol II Rat1–PTTC1. The steps of 3D auto-refinement, contrast transfer function refinement, Bayesian polishing and post-processing were finished in RELION 3.1. To improve the Rat1–Rai1 map signal, polished particles of Pol II Rat1–PTTC1 were imported to cryoSPARC v3.3.2 (ref. 48). Particles were then subjected to a non-uniform refinement, followed by particle subtraction using a mask that excludes Rat1-Rai1, Rp4-Rpb7 and structure elements around the Pol II RNA exit channel, and finally refined by local refinement. Gold-standard Fourier shell correlation indicated a mean map resolution of 2.7 Å for the Pol II Rat1–PTTC1 overall map (map 1a) and a resolution of 3.5 Å for the Rat1-Rai1-Rpb4-Rpb7 local refinement map (map 1b).

For Pol II Rat1–PTTC2, 5,313 images were recorded using EPU software in counting mode (2.67 exposures in 40 subframes resulting in a total dose of 50 electrons per square ångström, defocus range of −1.0 to −2.0 µm). Frames of individual videos were aligned using Motion-Cor2 (ref. 45). The following processes were primarily carried out in cryoSPARC v4.2.1 (ref. 48) unless specified otherwise. The Patch CTF module estimated each summed image's contrast transfer function. A total of 3,506,025 particles were picked from motion-corrected images using Template Picker (template from Pol II Rat1–PTTC1) and extracted with a box size of 100 pixels using particle extraction. The resulting particles were first sorted by 2D classification ($n = 50$) followed by ab initio reconstruction ($n = 4$). Particles from the good class were then classified by heterogeneous refinement ($n = 4$). Classes containing the Rat1-Rai1 or Spt4-Spt5 complex density were re-extracted using a box size of 400 pixels, followed by non-uniform refinement. For the Pol II Rat1–PTTC2 class, particles were further processed using RELION 3.1 Bayesian polishing. Polished particles were refined using non-uniform refinement, resulting in an overall map (map 2a) with a nominal resolution of 2.8 Å. The Rat1-Rai1-Rpb4-Rpb7 map signal was improved by particle subtraction (using Pol II core as a mask) and local refinement, resulting in a local map (map 2b) with a nominal resolution of 3.4 Å.

Model building and refinement

The cryo-EM maps were first fitted with the Pol II EC structure (PDB: 5VVS (ref. 49)), Spt5 structure (PDB: 7NKX (ref. 50)) and an Alpha-Fold2-predicted model[51] of Rat1-Rai1 using USCF Chimera[52], followed by manual adjustment in Coot[53] and real-space refinement in Phenix[54].

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Atomic coordinates have been deposited in the PDB under the accession codes 8JCH (Pol II Rat1–PTTC1) and 8K5P (Pol II Rat1–PTTC2).

Article

Corresponding cryo-EM density maps have been deposited in the Electron Microscopy Data Bank under the accession codes EMD-36162 (Pol II Rat1–PTTC1) and EMD-36908 (Pol II Rat1–PTTC2). Source data are provided with this paper.

44. Higo, T. et al. Development of a hexahistidine-3x FLAG-tandem affinity purification method for endogenous protein complexes in *Pichia pastoris*. *J. Struct. Funct. Genomics* **15**, 191–199 (2014).
45. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
46. Zhang, K. Gctf: real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
47. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**, e42166 (2018).
48. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
49. Xu, J. et al. Structural basis for the initiation of eukaryotic transcription-coupled DNA repair. *Nature* **551**, 653–657 (2017).
50. Farnung, L., Ochmann, M., Engeholm, M. & Cramer, P. Structural basis of nucleosome transcription mediated by Chd1 and FACT. *Nat. Struct. Mol. Biol.* **28**, 382–387 (2021).
51. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
52. Pettersen, E. F. et al. UCSF Chimera-a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
53. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
54. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
55. Ehara, H., Kujirai, T., Shirouzu, M., Kurumizaka, H. & Sekine, S. I. Structural basis of nucleosome disassembly and reassembly by RNAPII elongation complex with FACT. *Science* **377**, eabp9466 (2022).

Author contributions Y. Zeng collected the cryo-EM data, solved the cryo-EM structures and carried out biochemical experiments. H.-W.Z. purified yeast Pol II. X.-X.W. assisted in structure determination. Y. Zhang designed experiments, analysed data and wrote the manuscript.

Competing interests The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-024-07240-3.
Correspondence and requests for materials should be addressed to Yu Zhang.
Peer review information *Nature* thanks Seth Darst and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.
Reprints and permissions information is available at http://www.nature.com/reprints.

a

Pol II Rat1-PTTC1

Absorbance at 280nm (mAU)

b

Pol II Rat1-PTTC1

kDa / M

250 — Rpb1
*
150 — Rpb2
— Rat1/Spt5
100
75

50 — Rai1
37

25 — Rpb3
— Rpb4+5

20 — Rpb7
— Rpb6
— Rpb8

c

Pol II Rat1-PTTC2

Absorbance at 280nm (mAU)

**d**

Pol II Rat1-PTTC2

kDa / M

250 — Rpb1
150 — *
— Rpb2
100 — Rat1/Spt5
75

50 — Rai1
37

25 — Rpb3
— Rpb4+5
20 — Rpb7
— Rpb6
— Rpb8
15 — Rpb9+11
— Spt4
10 — Rpb10+12

e

NT–DNA 5'–TATAGGAAAC GACCAGGATT TTTCTTTTTTTTTCTGTACAAACGCGTGTACGCATGTAAC–Biotin–3'
T–DNA   3'–ATATCCTTTG CTGGTCCTAA AAAGAAAAAAAAGACATGTTTGCGCACATGCGTAC–5'
RNA   5'– p UUUUU GACCAGGAUU *[UTP]*[CTP]

Pol II GACCAGGATT
Stalled site

Rat1
Rai1

Streptavidin
Biotin

	BSA		Rat1/Rai1	
	T	S	T	S

anti-flag

f

RNA 5'– p ACAGAUCGU*G*UCCAUCGAGAGGU–3'
         1 nt   10 nt   23 nt
                    Stop site

Rat1/Rai1 – +
RNA   + +

RNA   23 nt
       ~15/14 nt

Extended Data Fig. 1 | The assembly of the Pol II Rat1-PTTC1 and Pol II Rat1-PTTC2. a-b, Elution peaks and SDS-PAGE results of Pol II Rat1-PTTC1 from a size-exclusion chromatography. **c-d,** Elution peaks and SDS-PAGE result of Pol II Rat1-PTTC2 from a size-exclusion chromatography. Asterisks indicate degradation products of Rpb1. **e,** The Pol II release result shows the Rat1-Rai1 complex releases Pol II from the elongation complex. Three biological repeats were done, see Supplementary Fig. 1. **f,** The RNA cleavage result shows the Rat1-Rai1 complex has exoribonuclease activity. Three biological repeats were done, see Supplementary Fig. 1.

Article
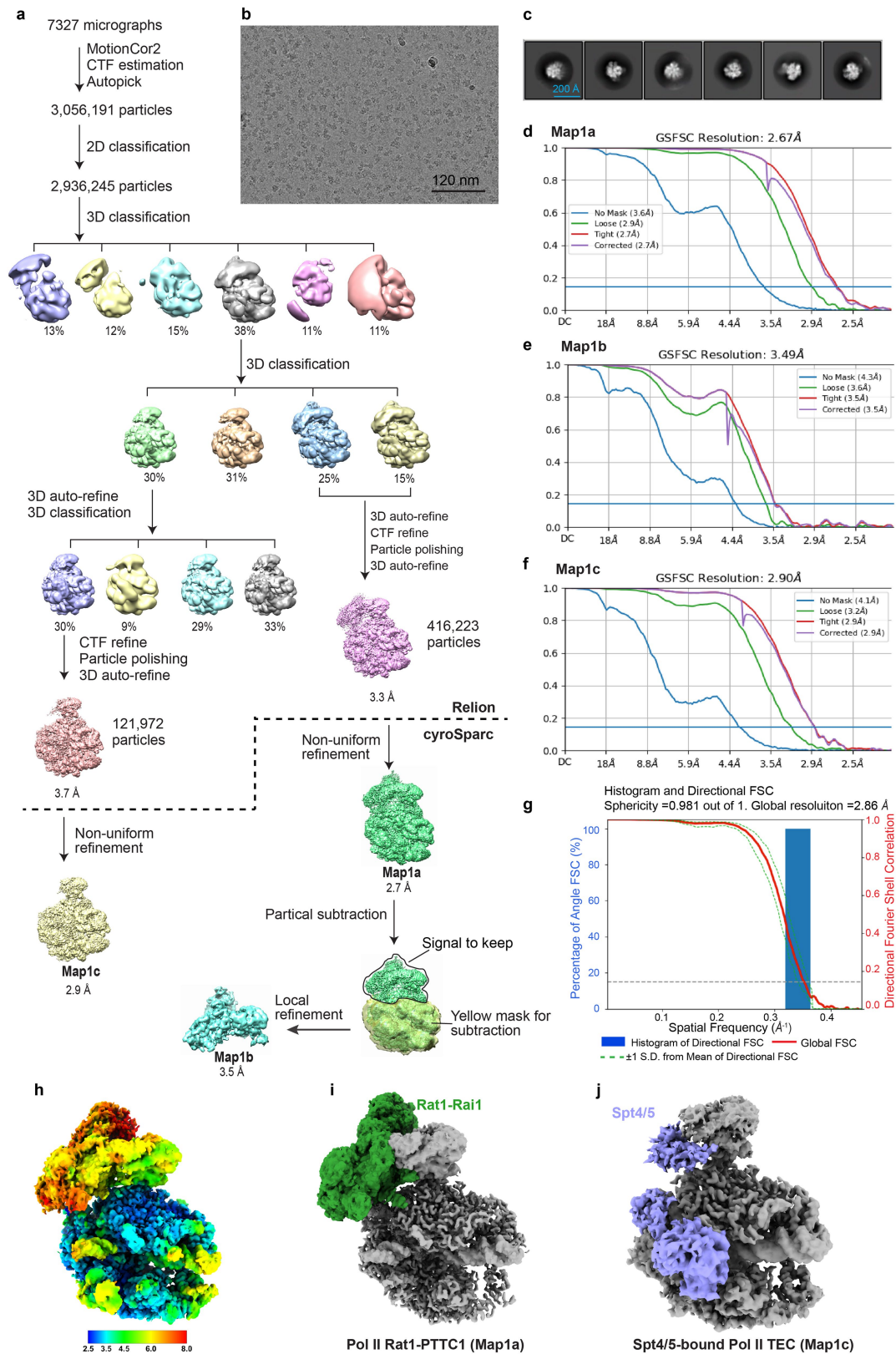
Extended Data Fig. 2 | Cryo-EM data processing of Pol II Rat1-PTTC1.
a, The flowchart of data processing for Pol II Rat1-PTTC1. **b**, Representative cryo-EM micrograph. **c**, Representative 2D class averages. **d-f**, Gold-standard FSC curves of map 1a, map 1b, and map 1c. **g**, The 3D FSC plot. The dotted line represents 0.143 cutoff of the global FSC curve. **h**, Nominal 2.7 Å-resolution cryo-EM reconstruction of map 2a filtered by local resolution. **i-j**, Map classified from Pol II Rat1-PTTC1. Map 1a represents the Pol II Rat1-PTTC1. Map 1c represents Pol II Spt4-Spt5 bound TEC. The Rat1-Rai1 map is colored green, and the Spt4-Spt5 map is colored blue.
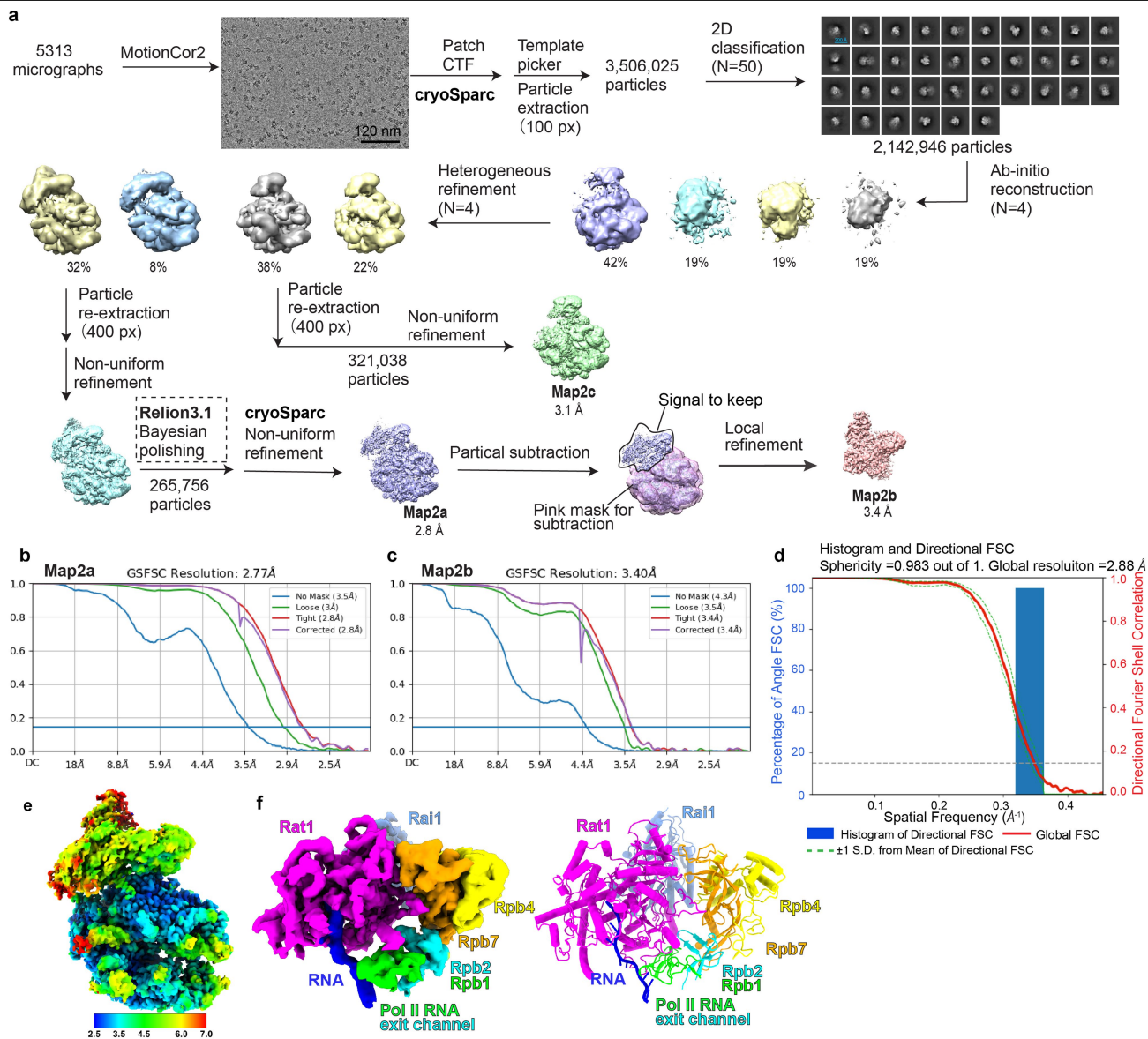
Extended Data Fig. 3 | Cryo-EM data processing of Pol II Rat1-PTTC2.
a, The flowchart of data processing for Pol II Rat1-PTTC2. **b-c**, Gold-standard FSC curves of map 2a, map 2b. **d**, The 3D FSC plot. The dotted line represents the 0.143 cutoff of the global FSC curve. **e**, Nominal 2.8 Å-resolution cryo-EM reconstruction of map 2a filtered by local resolution. **f**, Local refined map (map 2b) of Rat1, Rai1, Rpb4-Rpb7, and structure elements around the RNA exit channel of Pol II.
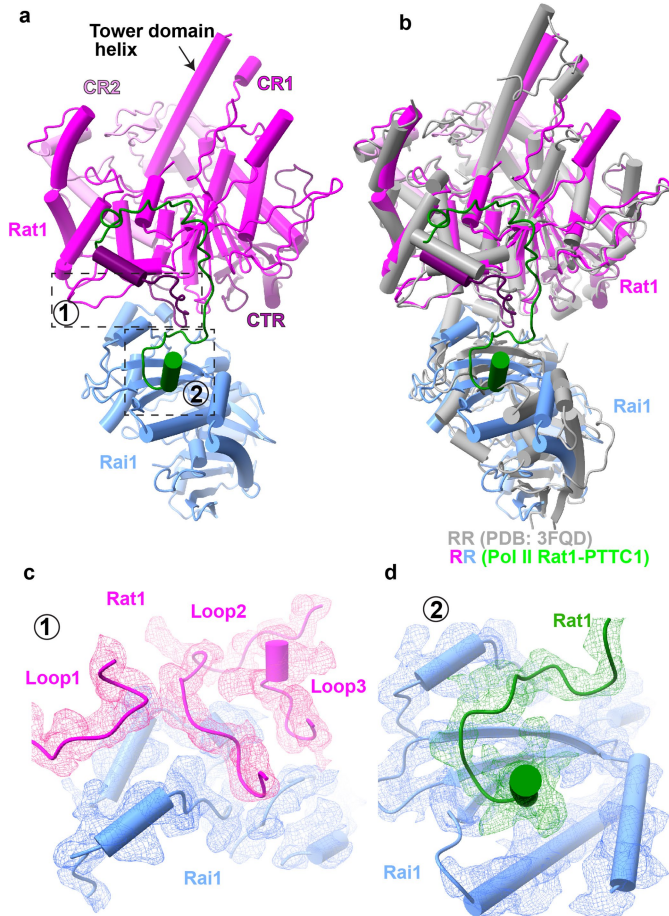
Extended Data Fig. 4 | The interface between Rat1 and Rai1 in Pol II Rat1-PTTC1. a, Overview of the interface between Rat1 and Rai1. The dashed boxes highlight the structure elements of Rat1 that contact the two concaves of Rai1. Green represents the Rat1 regions not observed in the crystal structure of *S. pombe* Rat1-Rai1 complex. **b**, Structural comparison of the Rat1-Rai1 complex in Pol II Rat1-PTTC1 and the crystal structure of *S. pombe* Rat1-Rai1 complex (PDB:3FQD[26]). **c-d**, The cryo-EM maps (map 1b) of interfaces 1 and 2.
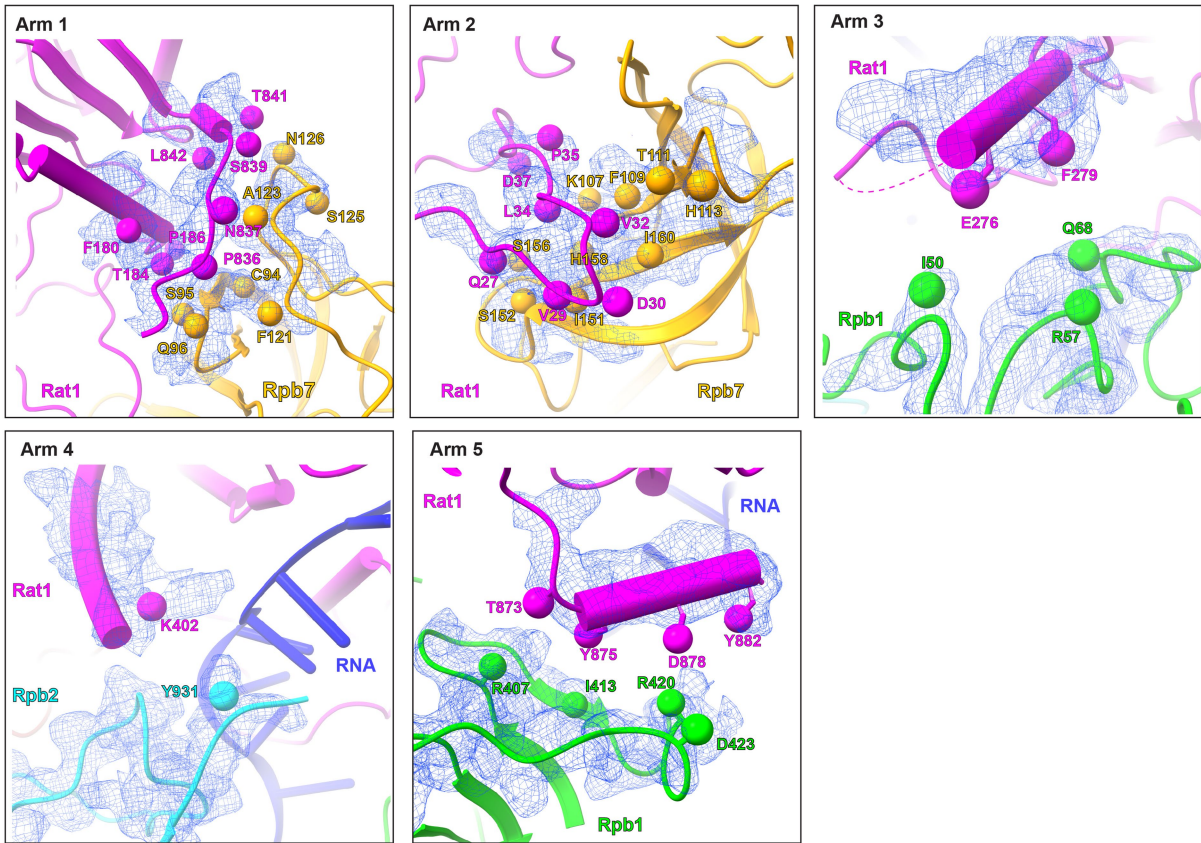
Extended Data Fig. 5 | The cryo-EM map of interfaces between Pol II and the Rat1-Rai1 complex in the structure of Pol II Rat1-PTTC1. The view orientation and cartoon color are the same as in Fig. 2d. The cryo-EM map is colored in blue.

Arm 1

Rat1_*S. cerevisiae*	176	YWTAFKLATDPGWKN	190	830	IQGGSLP......NLSTTLILK	845
Rat1_*S. pombe*	178	YYINKLNSDPCWRN	192	814	IDVTSADALQKYGSIDDNQSIS	835
XRN2_*H. sapiens*	176	YYIADRLNNDPGWKN	190	739	VPM.........LRDLTQNTVVS	752
XRN2_*M. musulus*	177	YYIADRLNNDPGWKN	190	739	VPM.........LRDLTQNTAVS	752
XRN2_*D. melanogaster*	178	YFVHDRQNNPAWKG	192	716	ISG.........LSDINDNITVT	729
XRN2_*C. elegans*	176	YYIHDRVTNDASWAN	191	765	VNH......DECQDLPTNCGIC	780
XRN3_*A. thaliana*	176	YYVHLRLNHDVGWKN	190	769	MEG.........MEDILTNQVIC	782

Arm 2

Rat1_*S. cerevisiae*	22	LEEQPQIV.DGVILPLDYSA	40
Rat1_*S. pombe*	22	IEAPTEKLPDGTEIEPDLSL	41
XRN2_*H. sapiens*	22	VEEKPKEC.NGVKIPVDASK	40
XRN2_*M. musulus*	22	VEEKPKEC.NGVKIPVDASK	40
XRN2_*D. melanogaster*	22	NENKQVDADTGRNIYEDPTL	41
XRN2_*C. elegans*	22	NEDRQRDQ.DGNRVPVDCTQ	40
XRN3_*A. thaliana*	22	IEEEPVEI.EGIKIPVDTSK	40

Arm 3

Rat1_*S. cerevisiae*	257	AQDNRKRNNLKDTINMTE..............EEKQFLQKQNSEQ	287
Rat1_*S. pombe*	259	FQQGSTKKTKEERL.........GIKR....LDDVSETNKVPVKK	290
XRN2_*H. sapiens*	256	KPNKPKPCGLCNQFGHEVKDCEGLPREKKGKHDELADSLPCAEG	299
XRN2_*M. musulus*	256	KPNKPKPCALCNQFGHEVKDCEGLPREKKGKHDELADSLPCAEG	299
XRN2_*D. melanogaster*	258	LPNKPRPCDICNGFGHEMDKCVGLGATAPTSANFKPDVPIGAEV	301
XRN2_*C. elegans*	257	VPNQPRACDLCGQYGHELKECRGAENETDLGDDY..CKPEQREK	298
XRN3_*A. thaliana*	257	TPGQQERCFLCGQMGHFASNCEGKPKKRAGESDEKGDGNDFVKK	300

Arm 4

Rat1_*S. cerevisiae*	395	RHIQEARKKEAFERRK	410
Rat1_*S. pombe*	398	LKQQEDRRNENYRRRQ	413
XRN2_*H. sapiens*	407	RKDDEDSFRRRQKEKR	422
XRN2_*M. musulus*	407	RKDDEDSFRRRQKEKR	422
XRN2_*D. melanogaster*	409	RQRREEQFKARDKARK	424
XRN2_*C. elegans*	406	RQQDEERFQENQRNKK	421
XRN3_*A. thaliana*	408	RAMQHQRQAERVKRDK	423

Arm 5

Rat1_*S. cerevisiae*	870	PVLTAYDLDSIMYKYNNQNYSRR	892
Rat1_*S. pombe*	860	RVLTPEDINQV...RAERSFSSR	879
XRN2_*H. sapiens*	779	AVLKPSDWEKS...SNGRQWKPQ	798
XRN2_*M. musulus*	779	TVLKPGDWEKS...SNGRQWKPQ	798
XRN2_*D. melanogaster*	756	QVLPN.......E...QSGNKHRPV	770
XRN2_*C. elegans*	807	KTLKPDDWNDR...RDG.RYQPQ	825
XRN3_*A. thaliana*	816	...PPALWHED...NGRRPMHNN	832

Extended Data Fig. 6 | Sequence alignment of Rat1 homologs. Sequence alignment shows that Rat1 residues in arms 1 and 2 contacting Pol II are generally conserved (red circles). Other Rat1 residues contacting Pol II are shown in blue circles.
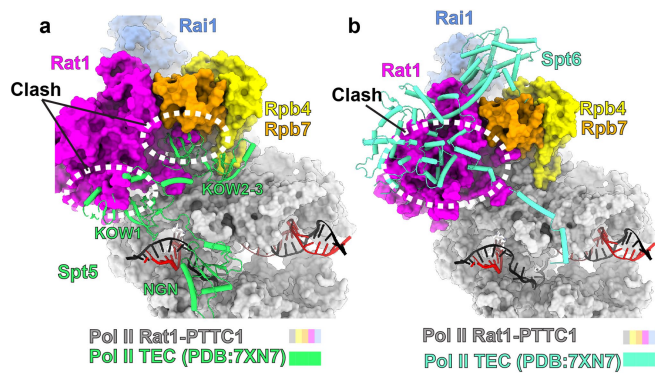
Extended Data Fig. 7 | Rat1 is incompatible with Pol II elongation factors.
a, Structural comparison between Pol II Rat1-PTTC1 and Pol II TEC (PDB:7XN7[55]) shows the steric clash of Rat1 and Spt5 KOW1/2/3 domains. **b**, Structural comparison between Pol II Rat1-PTTC1 and Pol II TEC (PDB:7XN7[55]) shows the steric clash of Rat1 and Spt6.
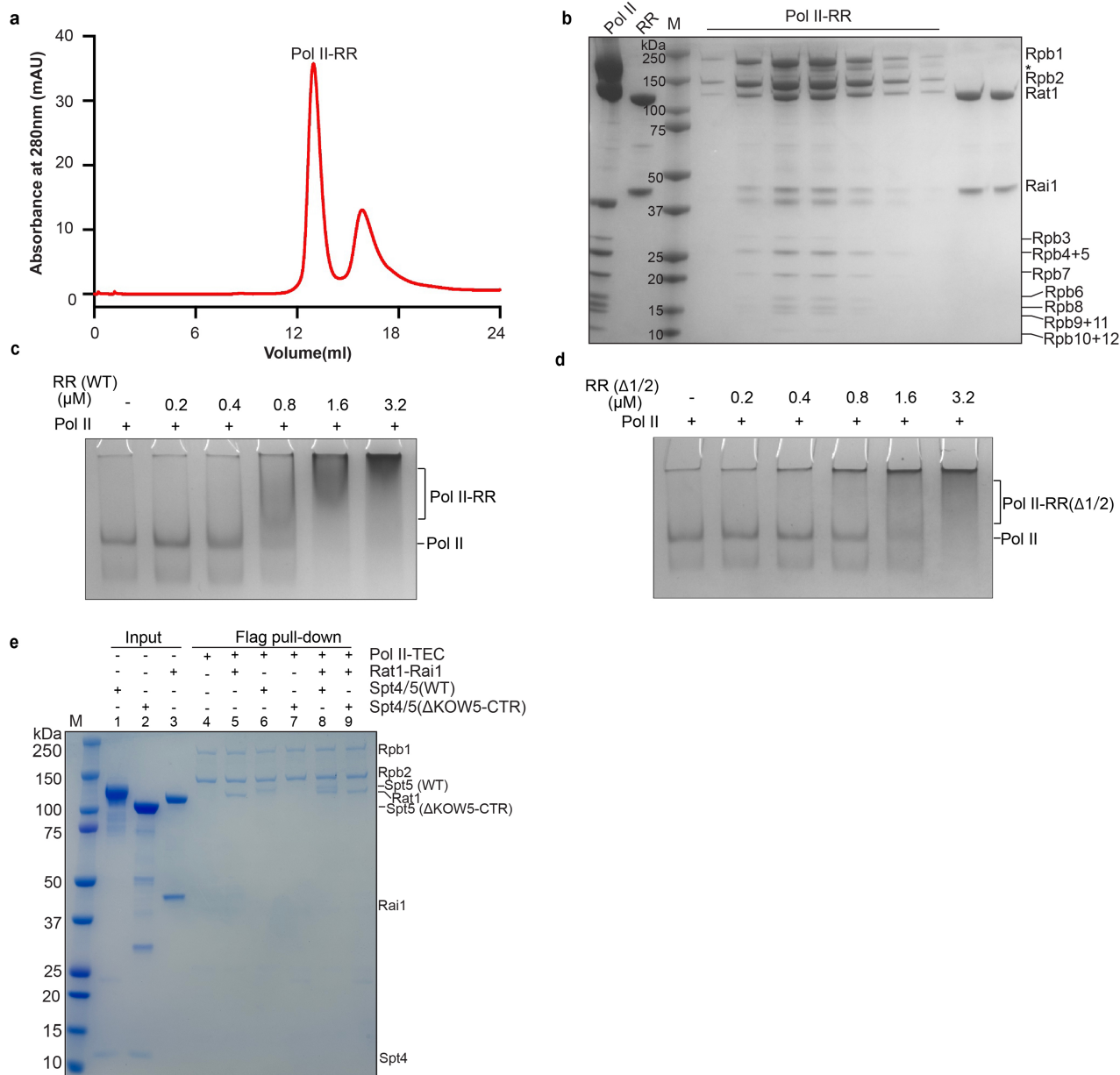
Extended Data Fig. 8 | Rat1 can be recruited to Pol II independent of RNA.
**a-b**, Elution peaks and SDS-PAGE result of Pol II-RR from a size-exclusion chromatography without a nucleic-acid scaffold. Two biological repeats were done, see source data and Supplementary Fig. 1. **c**, The native-page result shows that the Rat1-Rai1(RR) complex slows the migration of Pol II in a concentration-dependent way, suggesting that the Rat1-Rai1 complex forms a stable complex with Pol II. Three biological repeats were done, see Supplementary Fig. 1.
d, The native-page result shows that removing arms 1 and 2 weakens the interaction with Pol II. Three biological repeats were done, see Supplementary Fig. 1. **e**. The flag pull-down result shows Spt5 does not affect the interaction between Rat1-Rai1 and Pol II. Three biological repeats were done, see Supplementary Fig. 1.
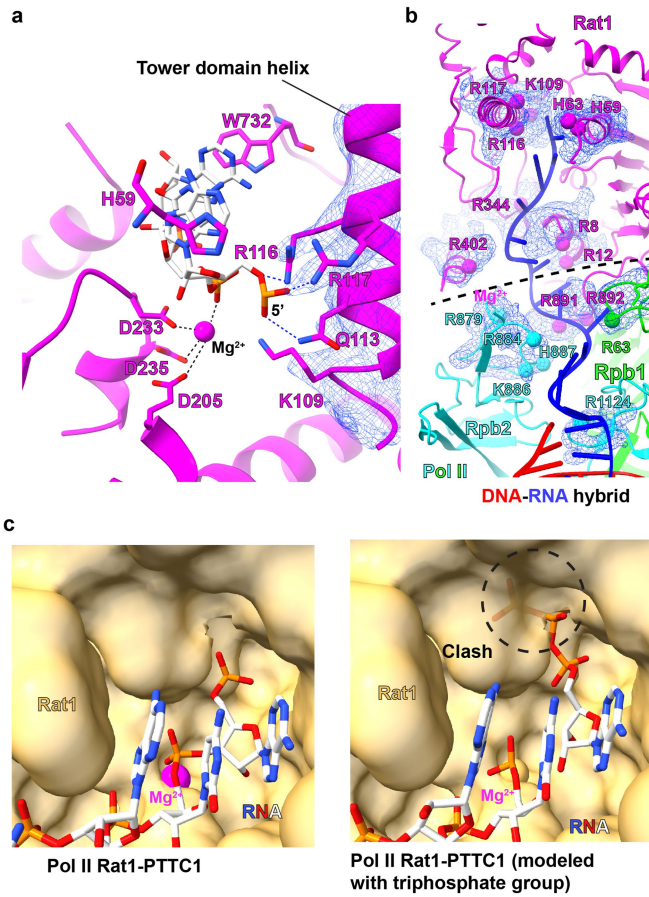
Extended Data Fig. 9 | The interaction of nascent RNA with Rat1 in Pol II Rat1-PTTC1. a, The interaction between the 5′ phosphate groups of the first two nucleotides with the active site of Rat1. Black dash, the coordination bonds. Blue dash, the H-bonds. **b,** The interaction of nascent RNA with Rat1 and Pol II. The cryo-EM map for the RNA-contacting residues is shown in blue. **C,** The modeled triphosphate group of the RNA 5′ terminus clashes with Rat1. Left, RNA with a monophosphate group in the structure of Pol II Rat1-PTTC1; right, RNA with a modeled triphosphate group in the structure of Pol II Rat1-PTTC1.

Article

Extended Data Table 1 | The statistics of cryo-EM structures in this study

	Pol II Rat1-PTTC1 (EMDB-36162) (PDB 8JCH)	Pol II Rat1-PTTC2 (EMDB-36908) (PDB 8K5P)
Data collection and processing		
Magnification	64,000	64,000
Voltage (kV)	300	300
Electron exposure (e–/Å^2)	50	50
Defocus range (μm)	-1.0 to -1.8	-1.0 to -2.0
Pixel size (Å)	1.1	1.1
Symmetry imposed	C1	C1
Initial particle images (no.)	3,056,191	3,506,025
Final particle images (no.)	416,223	265,756
Map resolution (Å)	2.7	2.8
FSC threshold	0.143	0.143
Map resolution range (Å)	2.5-8	2.5-7
Refinement		
Initial model used (PDB code)	5VVS	5VVS
Model resolution (Å)	6.4	6.4
FSC threshold	0.143	0.143
Model resolution range (Å)		
Map sharpening B factor (Å^2)	-79.9	-86.1
Model composition		
Non-hydrogen atoms	42,732	42,713
Protein residues	5,119	5,127
Ligands	8/3	8,3
B factors (Å^2)		
Protein	132.93	135.97
Ligand	169.08	168.81
R.m.s. deviations		
Bond lengths (Å)	0.002	0.002
Bond angles (°)	0.542	0.560
Validation		
MolProbity score	2.03	2.04
Clashscore	8.57	8.46
Poor rotamers (%)	2.29	2.24
Ramachandran plot		
Favored (%)	95.70	95.49
Allowed (%)	4.30	4.51
Disallowed (%)	0.00	0.00

Corresponding author(s): Yu Zhang

Last updated by author(s): Jan 23, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☒	☐	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	The EPU (version 2.12) was used to collect the cryo-EM micrographs.
Data analysis	MotionCor2, CTFFIND 4.1.8, RELION 3.0, RELION 3.1, cryoSPARC v3.3.2, and cryoSPARC v4.2.1 were used to process cryo-EM single particles. Phenix (version 1.20.1-4487) was used to perform structure refinement. Coot (version 0.9.6) was used to perform structure model building. UCSF Chimera (version 1.14) and chimeraX (version 1.5) was used for structure analysis. AlphaFold (version 2.0) was used for structure modeling. Prism (v.9.1.1) was used to analyse size exclusion chromatography result.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The cryo-EM map and coordinates were deposited in Protein Data Bank and Electron Microscopy Data Bank (Pol II Rat1-PTTC1: 8JCH, EMD-36162; Pol II Rat1-PTTC2: 8K5P and EMD-36108;). 5XON, 7XN7, and 3FQD were used for structure analysis.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The sample size has been described in the figure legends. No statistical method has been used to predetermine the sample size. Experiments were typically repeated in triplicates independently to ensure the robustness of conclusions.
Data exclusions	No data was excluded.
Replication	The replicate numbers for each experiments were described in the respective figure legends. Experiments were typically repeated in triplicates independently to ensure the robustness of conclusions. All attempts at replication were successful.
Randomization	Not applicable as no groups to be allocated.
Blinding	Not applicable as no groups to be allocated.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Mouse anti-flag monoclonal antibody (Yeasen Biotechnology, cat.no: 30503ES60). The dilution ratio used is 1:10000.
Validation	The mouse anti-flag antibody is commercially available and is validated by the supplier and literature (Dong W, Zhu Y, Chang H, et al. An SHR-SCR module specifies legume cortical cell fate to enable nodulation. Nature. 2021;589(7843):586-590. doi:10.1038/s41586-020-3016-z; Song Y, Guo Y, Li X, et al. RBM39 Alters Phosphorylation of c-Jun and Binds to Viral RNA to Promote PRRSV Proliferation. Front Immunol. 2021;12:664417. doi:10.3389/fimmu.2021.664417).