

RESEARCH ARTICLE SUMMARY

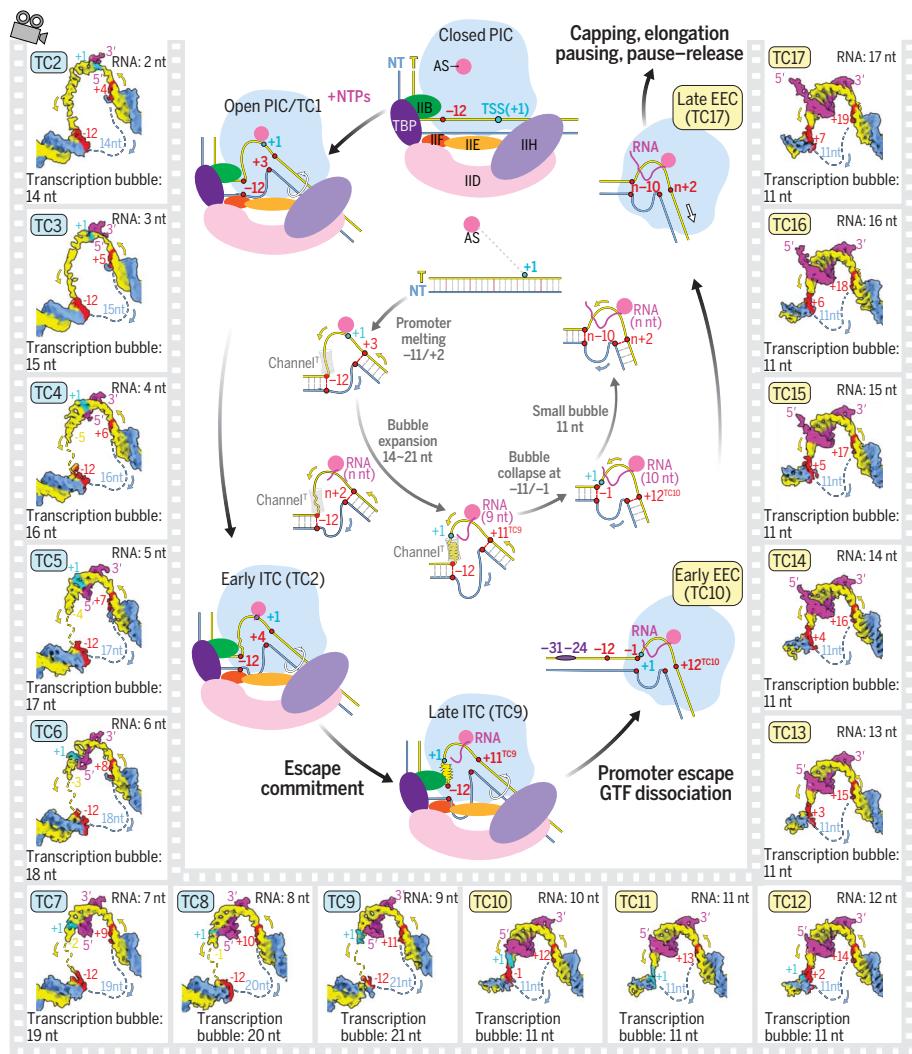
TRANSCRIPTION

Structural visualization of transcription initiation in action

Xizi Chen[†], Weida Liu[†], Qianmin Wang[†], Xinxin Wang[†], Yulei Ren, Xuechun Qu, Wanjun Li, Yanhui Xu*

INTRODUCTION: To achieve sophisticated transcription regulation, eukaryotic RNA polymerase II (Pol II) does not open the core promoter by itself and requires general transcription factors (GTFs) (~38 polypeptides in total) to

assemble a preinitiation complex (PIC). The PIC then sequentially transforms into an open-promoter PIC complex (OC), an initially transcribing complex (ITC), and an early elongation complex (EEC). Previous studies reported that



Schematic model of transcription initiation. Cryo-EM maps of nucleic acids are shown around the perimeter. Schematic models of representative complexes and nucleic acids are shown in the outer and inner rings of the central area, respectively. Functional states of Pol II and the transcription bubble are indicated. Directions of DNA translocation are indicated with arrows. IIB, IID, IIE, IIF, and IIH are transcription factors TFIIB, TFIID, TFIIE, TFIIF, and TFIH, respectively. AS, active site of Pol II; NT, nontemplate strand; T, template strand; TBP, TATA box binding protein.

notable functional and compositional changes occurred during these processes. However, the underlying molecular mechanisms remain largely elusive.

RATIONALE: We reconstituted the de novo transcribing complexes (TC2 to TC17) with Pol II halted on G-less promoters when nascent RNAs reach 2 to 17 nucleotides (nt) in length, respectively. We determined the cryo-electron microscopy (cryo-EM) structures and recapitulated the dynamic processes of transcription initiation by connecting structures of the PIC (Protein Data Bank ID 7EGB) and 16 TC complexes.

RESULTS: Structural analyses reveal marked complex reorganizations. Starting from the PIC, the GTF TFIH couples the energy of adenosine triphosphate (ATP) hydrolysis to translocate the downstream promoter to Pol II, resulting in an OC that contains a transcription bubble of ~13 nt with the transcription start site (TSS) positioned at the active site. The OC then transforms into the ITC (TC2 to TC9), during which the upstream promoter and GTFs remain bound to Pol II, and thus the transcription bubble expands from 14 to 21 nt within the active center. The melted template strand guides the synthesis of nascent RNA, which increases its association with Pol II as it grows from 2 to 9 nt, which is correlated with a transition from abortive transcription to escape commitment. The differences between TC9 and TC10 reveal a sharp ITC-EEC transition and suggest that nucleoside triphosphate (NTP)-driven RNA-DNA translocation and template-strand accumulation in a nearly sealed channel may balloon out the channel, resulting in GTF dissociation, bubble collapse, and promoter escape. The structures of TC10 to TC17 reveal EEC conformation, indicating that Pol II escapes the initiation region and proceeds on the promoter, followed by pausing, RNA capping, pause-release, and productive elongation.

CONCLUSION: GTFs help Pol II open the core promoter through multiple interactions, which also create an obstacle to the escape of Pol II from the promoter and therefore must eventually be disrupted. Our study provides structural visualization of transcription initiation in action for understanding why and how transcription machinery undergoes such notable changes. ■

The list of author affiliations is available in the full article online.

*Corresponding author. Email: xuh@fudan.edu.cn

[†]These authors contributed equally to this work.

Cite this article as X. Chen et al., *Science* 382, eadi5120 (2023). DOI: 10.1126/science.adl5120

S READ THE FULL ARTICLE AT
<https://doi.org/10.1126/science.adl5120>

RESEARCH ARTICLE

TRANSCRIPTION

Structural visualization of transcription initiation in action

Xizi Chen^{1,2†}, Weida Liu^{1†}, Qianmin Wang^{1†}, Xinxin Wang^{1†}, Yulei Ren¹, Xuechun Qu¹, Wanjun Li¹, Yanhui Xu^{1,2*}

Transcription initiation is a complex process, and its mechanism is incompletely understood. We determined the structures of de novo transcribing complexes TC2 to TC17 with RNA polymerase II halted on G-less promoters when nascent RNAs reach 2 to 17 nucleotides in length, respectively. Connecting these structures generated a movie and a working model. As initially synthesized RNA grows, general transcription factors (GTFs) remain bound to the promoter and the transcription bubble expands. Nucleoside triphosphate (NTP)-driven RNA-DNA translocation and template-strand accumulation in a nearly sealed channel may promote the transition from initially transcribing complexes (ITCs) (TC2 to TC9) to early elongation complexes (EECs) (TC10 to TC17). Our study shows dynamic processes of transcription initiation and reveals why ITCs require GTFs and bubble expansion for initial RNA synthesis, whereas EECs need GTF dissociation from the promoter and bubble collapse for promoter escape.

In eukaryotic cells, core promoters are made up of ~100 base pairs (bp), including elements that direct preinitiation complex (PIC) assembly and span the transcription start sites (TSSs) (1–3). Transcription initiation starts with the assembly of a PIC that consists of RNA polymerase II (Pol II) and general transcription factors (GTFs), including TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIH (4–6). TFIID recognizes the core promoter through TATA box binding protein (TBP) and TBP-associated factors (TAFs) and recruits Pol II and other GTFs to assemble a closed-promoter PIC complex (CC). Adenosine triphosphate (ATP) hydrolysis of TFIH results in the translocation of the downstream promoter to the Pol II active center. The PIC transforms into an open-promoter PIC complex (OC), which contains a transcription bubble with the template strand inserted into the Pol II active site to guide the synthesis of nascent RNA. Earlier structural studies revealed mechanisms of nucleoside triphosphate (NTP)-driven RNA growth and RNA-DNA translocation (7–13) as well as PIC assembly and promoter opening (14–24).

Extensive biochemical studies have shown marked functional and compositional changes

at the early stage of transcription initiation. Briefly, an initially transcribing complex (ITC) is formed when RNA synthesis starts, and abortive initiation can be observed when RNAs are 2 or 3 nucleotides (nt) in length (25–28). Escape commitment has been proposed to occur when RNA reaches 4 nt long (29–33). When RNA is 7 to 10 nt long, the transcription bubble expands up to 17 or 18 nt, followed by an abrupt bubble collapse (33–37). Pol II will eventually break its contact with GTFs, escape from the promoter, and form an early elongation complex (EEC) (38, 39). Although structures of reconstituted ITCs have been reported (40–46), these complexes were not started from a PIC but instead were assembled on artificial RNA-DNA hybrids that were formed by a mismatched transcription bubble in which the template strand paired with a synthetic complementary RNA. A recent study captured a structure of a yeast (*Saccharomyces cerevisiae*) backtracked ITC complex that started from a PIC (47). However, transcription initiation in yeast and mammals differs in various aspects. For example, yeast polymerase undergoes TSS scanning on the promoter for up to a hundred nucleotides (48, 49), which does not occur in mammals.

Reported structures to date provide fragmented glimpses of complexes that have varying compositions and different promoters. The present understanding of de novo transcription initiation in mammals, including GTF dissociation, bubble expansion and abrupt collapse, and the ITC-EEC transition, is limited by inconsistencies across various experimental contexts and the lack of continuously captured structures of transcribing complexes in action. Despite decades of biochemical and structural studies, the underlying molecular mechanisms of these processes remain largely elusive.

Results

In vitro reconstitution of de novo transcription initiation

To investigate de novo transcription initiation, we synthesized a series of modified super core promoters (SCPs) (Fig. 1A and data S1). Each promoter consists of an upstream TATA box, downstream TFIID-binding elements, and a G-less cassette between the putative TSS and a GG dinucleotide at varying positions on the nontemplate strand. The promoters are termed SCP^{nG}, with “n” denoting the G-stop position relative to the TSS, at which Pol II should stall if guanosine triphosphate (GTP) is not provided in the reaction. Mammalian Pol II and human GTFs were purified as previously described (15, 16) (fig. S1A). The reactions of de novo transcription initiation were respectively performed on SCP^{3G} to SCP^{17G} promoters in the presence of ATP, uridine triphosphate (UTP), [α -³²P]-UTP, cytidine triphosphate (CTP), and chain-terminating 3'-O-methylguanosine-5'-triphosphate (^{3'-OM}GTP). Note that reaction on SCP^{2G} was not performed because the generated AG dinucleotide would lack a radioactive U nucleotide and would not be able to be detected.

The reaction on each SCP^{nG} promoter resulted in a predominant G-stop RNA of the expected length (Fig. 1B; fig. S1, B to E; and supplementary text). Reaction on SCP^{3G} generated the highest level of G-stop RNAs among all the reactions, and reactions on SCP^{4G} to SCP^{6G} generated much more abundant G-stop RNAs compared with the reactions on SCP^{7G} to SCP^{17G}. This suggests that short (3 to 6 nt) “dead-end” G-stop products are unstable and dissociate more frequently from the halted Pol II, which allows for multiple cycles of transcription initiation. RNA products longer than 6 nt remain stably bound to Pol II for up to one round of initiation.

We observed two short RNA products of ~2 to 4 nt in length (Fig. 1B; fig. S1, B to E; and supplementary text). The RNA products at the lower position largely accumulated in the presence of TFIIS (fig. S1D), which suggests that these short RNAs were derived from backtracked complexes, in which nascent RNAs backtracked to the active site by several nucleotides followed by RNA cleavage. A time-course reaction on SCP^{17G} showed an accumulation of RNAs at a position above the cleaved RNAs (fig. S1D), which suggests that these RNAs are abortive products that have dissociated from the transcribing complexes. Reactions on SCP^{nG} and G-less-containing adenovirus major late (AdML)^{9G/10G} promoters under similar conditions generated similar heterogeneous RNA products that resulted from abortive transcription, cleaved RNAs from backtracked complexes, alternative TSS selection, and readthrough transcription (Fig. 1B and fig. S1, B to E). This result generally agrees with earlier *in vitro* transcription reactions on the AdML promoter (26–28, 33) and

¹Fudan University Shanghai Cancer Center, Institutes of Biomedical Sciences, New Cornerstone Science Laboratory, State Key Laboratory of Genetic Engineering, Department of Biochemistry and Biophysics, School of Life Sciences, Shanghai Key Laboratory of Radiation Oncology, and Shanghai Key Laboratory of Medical Epigenetics, Shanghai Medical College of Fudan University, Shanghai 200032, China. ²The International Co-laboratory of Medical Epigenetics and Metabolism, Ministry of Science and Technology, China, Department of Systems Biology for Medicine, School of Basic Medical Sciences, Shanghai Medical College of Fudan University, Shanghai 200032, China.

*Corresponding author. Email: xuyh@fudan.edu.cn

†These authors contributed equally to this work.

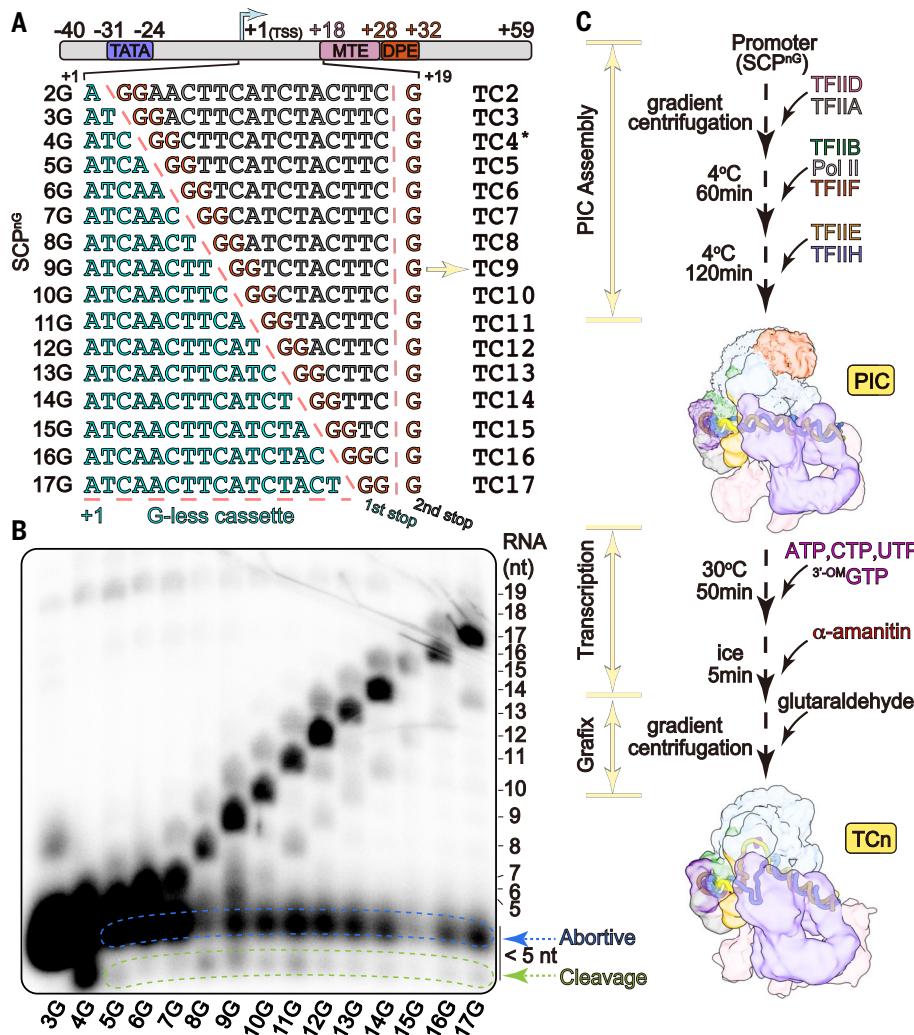


Fig. 1. In vitro reconstitution of de novo transcribing complexes. (A) Schematic diagram of SCP^{ng} promoters. The numbers above the promoter represent the positions relative to the TSS (+1). The GG dinucleotide at position "n" and the G nucleotide at +19 provide the first and second (if readthrough occurs) G-stop sites for Pol II. TCn represents the assembled transcribing complex on the indicated SCP^{ng} promoter. The asterisk indicates that TC4 on SCP^{ng} was used for the transcription assay and TC4 on SCP^{ng} mutant promoter was used for the structural study (fig. S1 and supplementary text). (B) In vitro transcription initiation assay. RNA products from one of the two independent experiments were visualized by autoradiography. The positions of abortive RNA and cleaved RNA from backtracked complexes are indicated. (C) Schematic diagram of the assembly of transcribing complexes. Transparent cryo-EM maps of the PIC and a representative TC show the changes in the promoter that is embedded within the complexes.

reported RNA heterogeneity in transcription initiation (50–52).

Structure determination

To visualize transcription initiation in action, we used $\text{SCP}^{2\text{G}}$ to $\text{SCP}^{17\text{G}}$ promoters to assemble de novo transcribing complexes termed TC2 to TC17, respectively (Fig. 1C). TC2 represents an ITC that forms the first RNA phosphodiester bond. TC17 was chosen as the last complex because a nascent RNA longer than 17 nt emerges from Pol II (53). TC4 was assembled on a $\text{SCP}^{4\text{G}}$ mutant promoter that generated a better-resolved structure (see supplementary text). For each complex assembly, we first as-

sembled a PIC on an SCP^{ng} promoter as previously described (15, 16). We then added ATP, CTP, UTP, and $^{3\prime}\text{-OM}$ GTP to start transcription initiation, which was followed by the addition of the mushroom toxin α -amanitin to inhibit the leaky activity of chain elongation. The reaction products were subjected to gradient fixation and cryo-electron microscopy (cryo-EM) structure determination (fig. S2 and data S2 to S4).

These transcribing complexes showed marked structural heterogeneity (fig. S2, data S2 to S4, and supplementary text), consistent with results from our in vitro assay (Fig. 1B and fig. S1, B to E) and earlier observations (30, 54–56). However, RNA heterogeneity in the reactions does not

necessarily correlate precisely with complex heterogeneity for structural studies because the latter represents a snapshot of the reaction at a specific point in time. For example, we found neither abortive complexes nor backtracked complexes in our cryo-EM data processing, likely owing to the following reason. Cryo-EM samples were prepared by freezing reaction products 50 min after the addition of NTP. DNA translocation and RNA synthesis were further inhibited by the addition of $^{3\prime}\text{-OM}$ GTP and α -amanitin. By contrast, abortive and backtracked complexes were transient and unlikely to accumulate in the reactions. Even though small quantities of these complexes might exist in the samples, minor differences in nucleic acids were not sufficient to distinguish them from TC complexes in cryo-EM three-dimensional (3D) classification.

To obtain structures at high resolution, we collected a large amount of data (580,409 images in total) with a 300-kV Titan Krios G3 microscope over ~110 days. Cryo-EM 3D classification and comparison defined particles that were used for the reconstruction of the TC2 to TC17 complexes (summarized in data S4). Cryo-EM maps of most of the complexes revealed Pol II at a resolution of 2.7 to 3.3 Å and well-ordered RNA-DNA hybrids in the active center (fig. S3 and data S2). Nucleic acids, Pol II, and TFIIF were modeled in all the complexes. Other GTFs were modeled by docking the structure templates into the cryo-EM maps of TC2 to TC9 but were not modeled in TC10 to TC17 because of high flexibility.

Overall structures of ITCs and EECs

Dynamic processes of transcription initiation can be recapitulated by connecting structures of the PIC (16) and the transcribing complexes from TC2 to TC17 (Fig. 2A, Movie 1, and movie S1). As detailed later in the text and in Figs. 3 to 6, TC2 to TC9 represent ITCs correlated with abortive transcription and escape commitment. TC10 to TC17 represent EECs correlated with promoter escape. The differences between TC9 and TC10 suggest a sharp transition from the ITC to the EEC that is correlated with bubble collapse and GTF dissociation.

Structures of TC2 to TC9 reveal a similar overall architecture of the ITC (Fig. 2, A to C; fig. S4; movies S2 to S4; and supplementary text). Similar to what occurs in the PIC (16), GTFs remain bound to Pol II and the TATA box is bent by TBP. The following promoter region at -23 to -12 (-23/-12) is stabilized by the TFIIB N-terminal core domain (B-core-N), the TFIIB linker (B-linker), the TFIIF winged-helix domain (WHT^{TFIIF}), and the Pol II protrusion. The upstream edge of the transcription bubble remains fixed at position -12 throughout TC2 to TC9. By contrast, the downstream edge of the transcription bubble is located at position n+2 in the TCn complex, which is

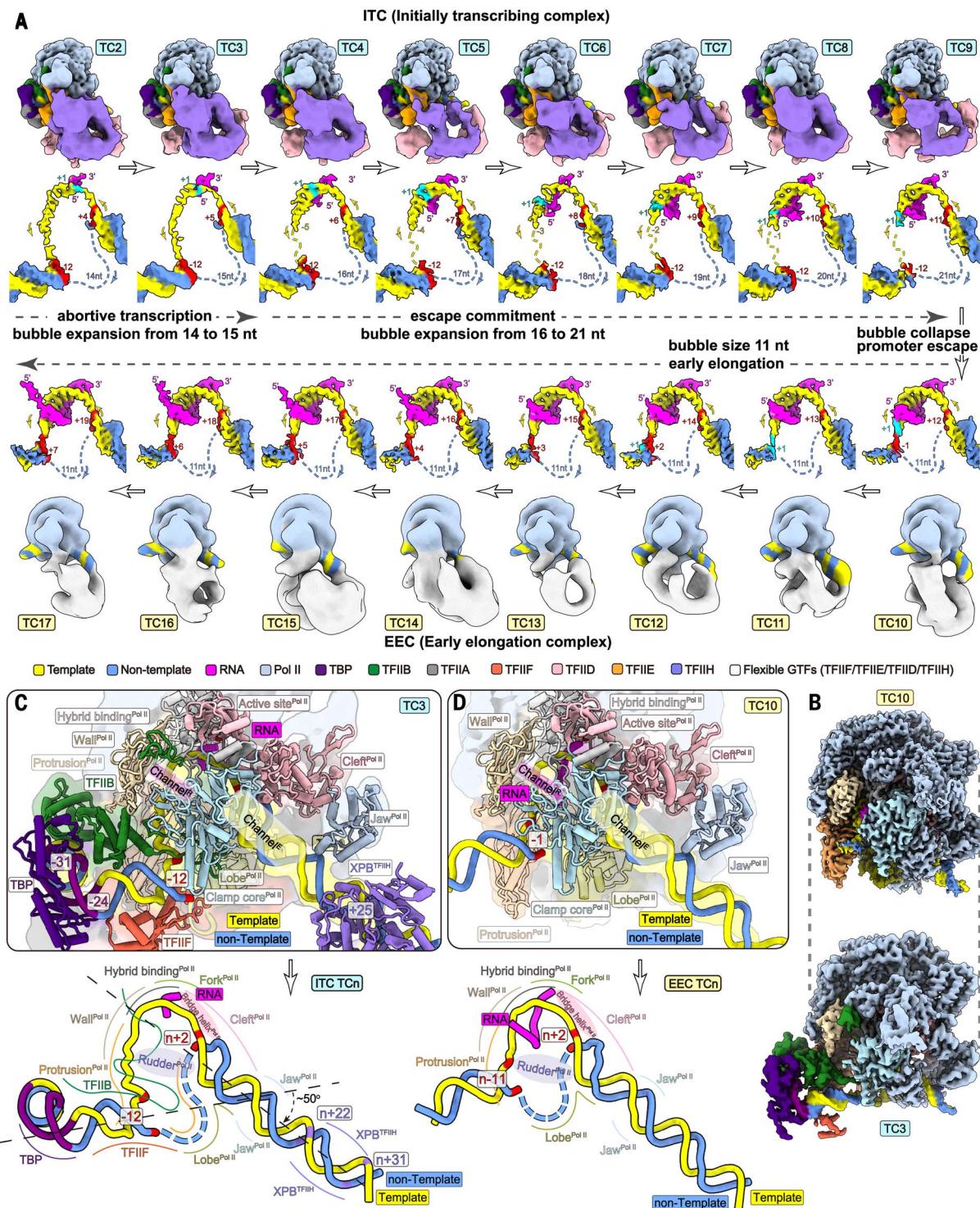


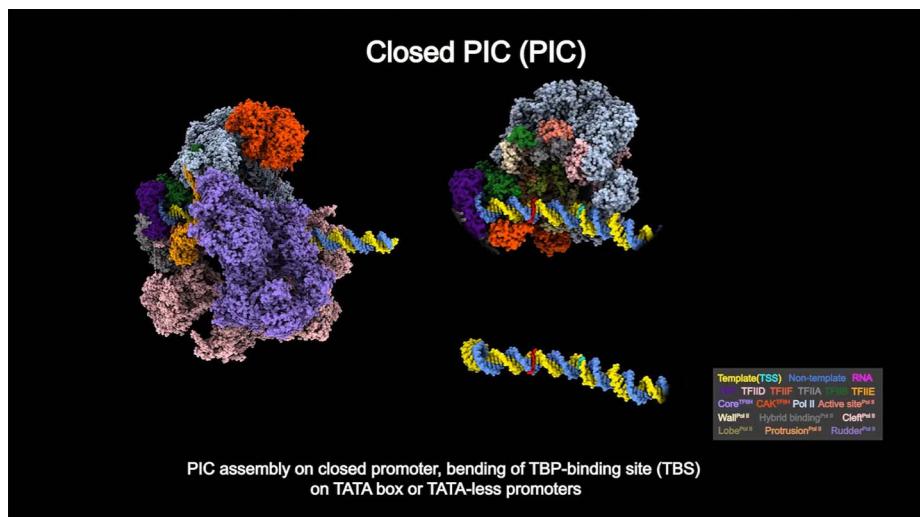
Fig. 2. Structures of the ITC and EEC. (A) Overall cryo-EM maps of TCn complexes and close-up views of nucleic acids. The color scheme is indicated at the bottom of the panel. The directions of strand translocation are indicated with arrows. The middle arrows indicate the functional states of Pol II and the transcription bubble. (B) Locally refined cryo-EM maps of Pol II, including Pol II, nucleic acids, and closely bound GTFs, in representative ITC and EEC complexes. (C and D) Close-up views of ITC (C) and EEC (D) structural models show protein-promoter interactions (top) and promoter conformations (bottom).

indicative of the continuous insertion of the downstream promoter into the active center. The promoter downstream of n+2 remains a straight double helix and is tilted by ~50° relative to the upstream duplex. The down-

stream promoter at n+2/n+11 inserts into the open DNA entry channel (channel^E). TFIIFH binds the downstream promoter at n+22/n+31, consistent with its role early in transcription initiation (25, 57–61), which is dependent on the

downstream promoter regions (62). Nascent RNA pairs with the template strand to form an RNA-DNA hybrid in the active center.

Cryo-EM maps of TC10 to TC17 show similar organization of the EEC (Fig. 2, A, B, and D;



Movie 1. A movie of transcription initiation. The movie was generated by connecting structures of the PIC (16) and TC2 to TC17 and shows the movements of Pol II relative to the upstream promoter. The image on the left shows the full complex. The images at the upper right show an omitted structure and a cross-sectional view of the complex. The image at the bottom right shows nucleic acids. The color scheme is also shown at the bottom right. Brief descriptions of each state are included at the bottom of the screen.

fig. S5; and movies S2 to S4). The RNA-DNA hybrid was well stabilized within the Pol II activity center, and a weak density of flanking RNA was observed in an open RNA exit channel (channel^R). The flanking upstream and downstream double-stranded promoter regions are flexible owing to the lack of stabilization by GTFs, which nearly dissociate from the promoter (Fig. 2B and fig. S5A). The TFIIIF dimerization module and TFIIIE remain bound to Pol II. The remaining GTFs, including TFIIID (with TBP), TFIIIA, TFIIIB, and TFIIIH, are either completely dissociated from or flexibly tethered to Pol II through the associations of TFIIIF, TFIIIE, and/or the exposed promoter elements. The flexibility of GTFs in the EEC may also explain seemingly inconsistent observations of GTFs on transcribing complexes, which vary under different experimental conditions (38, 39, 63–65). Thus, Pol II is no longer restrained by the upstream promoter-bound GTFs and is free to move forward and escape the promoter. Here, GTF dissociation refers to a dissociation from the promoter but not from Pol II.

The transition from the PIC to the early ITC

Compared to the PIC (16), the ITC shows the following structural differences (Fig. 3, Movie 1, fig. S6, and movie S1): The downstream promoter is tilted by ~30° relative to that of the PIC, leading to slight displacement and rotation of the TFIIIH core module and the TFIIIH-associated TFIIID away from Pol II. The cyclin-dependent kinase (CDK) activating kinase (CAK) module of TFIIIH was not observed, in agreement with the requirement that the TFIIIH core be in a proper position for loading CAK to Pol II (16). The Pol II active center is occupied by the melted promoter, and the ITC

exhibits a slight expansion in the direction of the DNA axis while becoming more compact in the vertical direction (as in Fig. 3B). TBP, the B-core-N, and the Pol II protrusion are slightly displaced toward the upstream direction, whereas the Pol II lobe moves toward the downstream direction. The clamp head and clamp helices move toward the lobe by up to 11 Å. The clamp-associated rudder and B-linker are more ordered and may contribute to stabilizing the melted promoter. These domains and motifs would generate steric clashes with a closed promoter, which suggests their roles in maintaining the promoter in an open conformation.

Early-stage transcription initiation is recapitulated by connecting structures from the PIC to TC3 (Figs. 3C and 6, Movie 1, and movie S1). Because Pol II by itself does not appear to efficiently open linear promoter DNA, the DNA translocase XPB of TFIIIH couples the energy of ATP hydrolysis to rotation and translocation of the downstream promoter to Pol II, resulting in promoter melting and displacement of the TSS by ~45 Å. The downstream promoter inserts into Pol II by translocation of ~16 bp, as measured by the displacement of the Pol II jaw-contacting point of the promoter, which is +5 in the PIC and +20 in TC3. This suggests an estimated insertion of 13 or 14 bp during the transition from the PIC to the OC. TFIIIH-generated torque for promoter melting would be counteracted by rotation and translocation of the upstream promoter if it were not restrained otherwise by Pol II-bound B-core-N, TBP, and WH^{TFIIIF}. Thus, extensive interactions among promoter, Pol II, and GTFs provide an integrated framework to open the promoter and maintain the open conformation. After OC formation, the melted template strand guides

Pol II to synthesize the first phosphodiester bond of nascent RNA, followed by cycles of nucleotide addition, RNA chain extension, and RNA-DNA hybrid translocation (66, 67).

TFIIH is required for efficient transcription initiation

TFIIH is required for transcription initiation (58, 61) but may be dispensable for promoter opening or transcription initiation under certain experimental contexts (24, 68–70). To test the effect of TFIIH in transcription initiation, we performed an *in vitro* transcription assay on representative templates (SCP^{4G}, SCP^{9G}, SCP^{10G}, and SCP^{17G}) in the presence and absence of TFIIH individually (fig. S1E). TFIIH supported the generation of G-stop RNAs, whereas the reactions in the absence of TFIIH showed largely decreased amounts of G-stop RNAs. Thus, TFIIH is essential for efficient *in vitro* transcription initiation. Transcription initiation did occur in the absence of TFIIH, albeit much less efficiently. The results may reconcile earlier studies that appeared to be controversial and suggest a context-dependent effect of TFIIH on transcription.

The absence of TFIIH led to only a slight decrease in the level of abortive RNAs, in sharp contrast to the largely decreased level of G-stop RNAs, which is indicative of an increase in the abortive/productive ratio (fig. S1E). Such an incremental abortive/productive ratio suggests that TFIIH may not only facilitate promoter opening but also prevent bubble reannealing (leading to RNA dissociation) in the early stage of transcription initiation. In the absence of TFIIH, nascent RNAs are more inclined to dissociate from Pol II to become abortive RNAs rather than productive RNAs. This observation agrees with the proposed function of TFIIH in pumping the downstream promoter to the active site, potentially aiding the formation and stabilization of the transcription bubble (25, 57–62). This phenomenon also accounts for the discrepancy between eukaryotic Pol II, which exhibits a much lower abortive/productive ratio, and bacterial RNA polymerase (RNAP), which has no TFIIH equivalent and is characterized by a high abortive/productive ratio (71, 72).

RNA-DNA hybrid stability is correlated with abortive transcription, escape commitment, and promoter escape

Throughout all the structures, Pol II adopts a posttranslocation state with 3'-OMGMP at the 3' end of the RNA paired with the first downstream cytosine at position +n of the template strand (Fig. 4, A and B; figs. S7 and S8; and supplementary text). At the nucleotide addition (A) site, a free 3'-OMGTP pairs with the second downstream cytosine at position n+1 of the template strand. The melted template strand exhibits a helical fold with phosphate

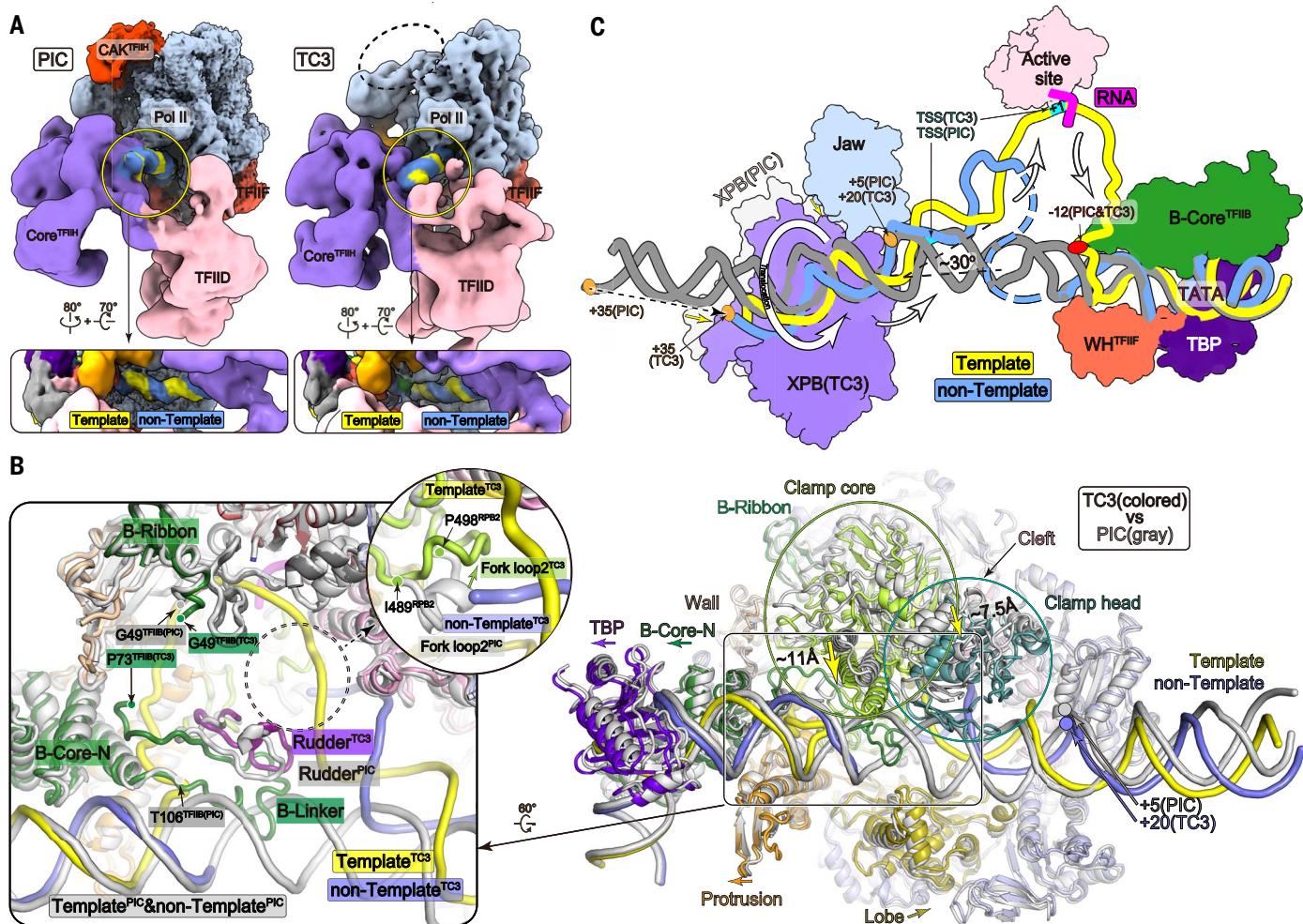


Fig. 3. The transition from the PIC to the early ITC. (A) Cryo-EM maps of the PIC (16) (top left) and a representative early ITC (TC3) (top right) shown with Pol II in a similar orientation. Close-up views of the promoters in the closed (bottom left) and open (bottom right) conformations are also shown. CAK is positioned on the Pol II surface in the PIC but is absent in TC3, as indicated by the white dashed oval. Although TC2 is almost identical to TC3, it has a less-stable RNA and thus was not used for comparison. (B) Superimposition of the PIC and TC3 structures with conformational differences indicated by

arrows, from PIC to TC3 (right). Close-up views of the differences are shown on the left. The circular close-up view shows fork loop 2 near the active site, which moves toward the bridge helix to avoid a clash with the melted nontemplate strand. (C) A model of the PIC-ITC transition, with molecular motions indicated with arrows. The closed promoter and XPB subunit in the PIC are colored gray. Molecules in the ITC are colored as indicated. Single-letter abbreviations for the amino acid residues are as follows: G, Gly; I, Ile; P, Pro; and T, Thr.

backbone packing against a positively charged path of the active center cleft. In TCn structures, the RNA-DNA hybrids within the active center adopt a conformation similar to that of the elongation complex (EC) (fig. S8, B and C). Conformational variations exist at the single-stranded template strand outside of the active center and the 5' flanking nucleotides of nascent RNA, which are flexible owing to the lack of stabilization.

RNA density is weak in TC2 and TC3 (Figs. 2A and 4B and movie S4), consistent with extensive abortive transcription on the SCP^{3G} promoter (Fig. 1B and fig. S1B). The short dead-end RNAs are prone to dissociate from the halted Pol II as a result of less stabilization by the template strand and Pol II. RNA dissociation may allow the unpaired template strand to retreat and guide a new cycle of RNA synthesis.

As the length of the RNA increases in complexes TC4 to TC9, cryo-EM maps of the RNA-DNA hybrid tend to be better resolved (Figs. 2A and 4B and movie S4), which is indicative of a metastable RNA-DNA hybrid correlated with escape commitment and a marked decrease in abortive initiation (29–33) (Fig. 1B and fig. S1B). Longer RNA makes more contact with the template strand and the Pol II active center cleft, which is rich in positively charged residues. We did not observe conformational differences in the Pol II active centers of TC2 to TC9. Thus, the transition from abortive transcription to escape commitment may result from an increase in RNA length rather than the proposed conformational changes of the transcribing complex (31).

TC10 to TC17 structures reveal well-ordered 10-bp RNA-DNA hybrids in the active center similar to those of EC structures (11, 73), consistent with a stable RNA binding and promoter escape (Fig. 2A, figs. S5 and S8, and movie S4). The nascent RNA protrudes out of the active center, winds through the open channel^R, and touches the Pol II dock (fig. S5C). Cryo-EM maps show the growth of RNA in TC10 to TC13 (fig. S5F). However, only 13 nt could be visualized in TC14 to TC17, and the 5' flanking nucleotides were flexible owing to the lack of stabilization.

The B-finger touches the 5' end of nascent RNA in the early ITC

Cryo-EM maps of the early ITC show a putative TFIIB finger (B-finger, residues 50 to 72,

also known as B-reader) in the Pol II active center (Fig. 4, B and C; and fig. S9), which has been biochemically detected (74, 75) and structurally observed (24, 40, 43–46). The B-finger

of TC3 contacts the template strand at position $-2/-3$ and touches the 5' end of nascent RNA, showing an insertion in the active center deeper than that of the previously reported

structures (40, 43–46). The B-finger in TC2 to TC6 is more separated from the active site and less stable as the RNA chain grows and is invisible in TC7 to TC9, which suggests a gradual

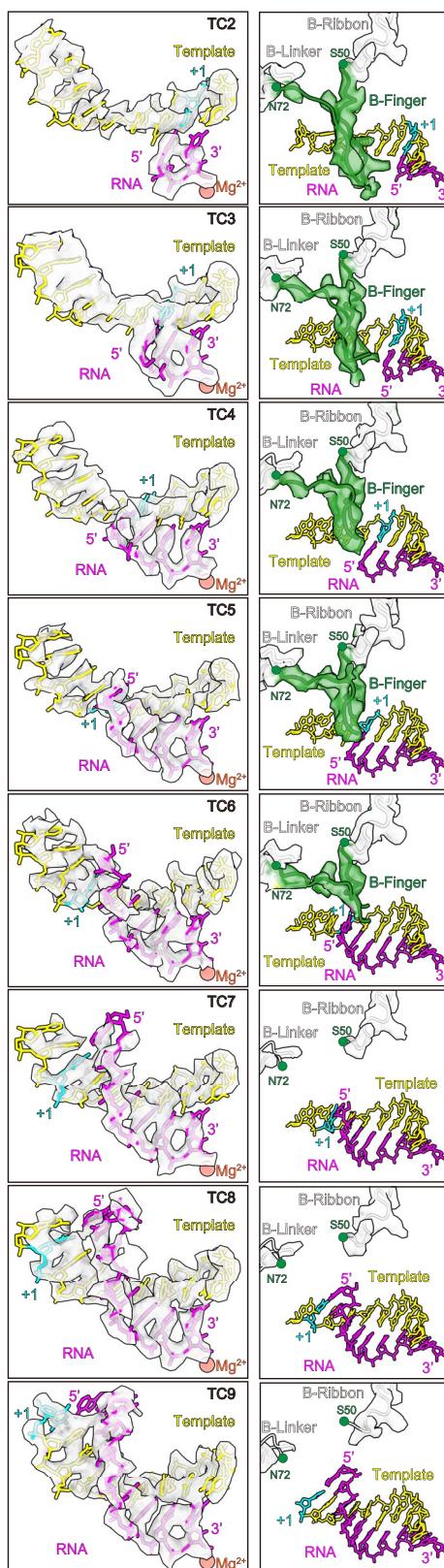
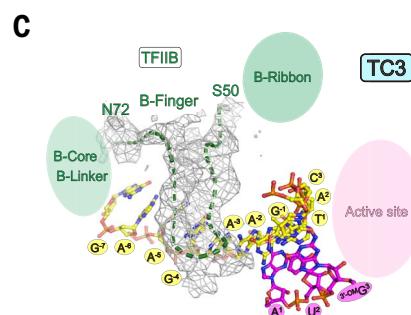
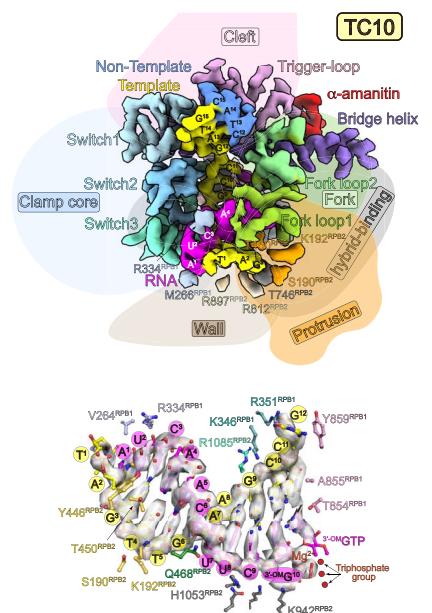
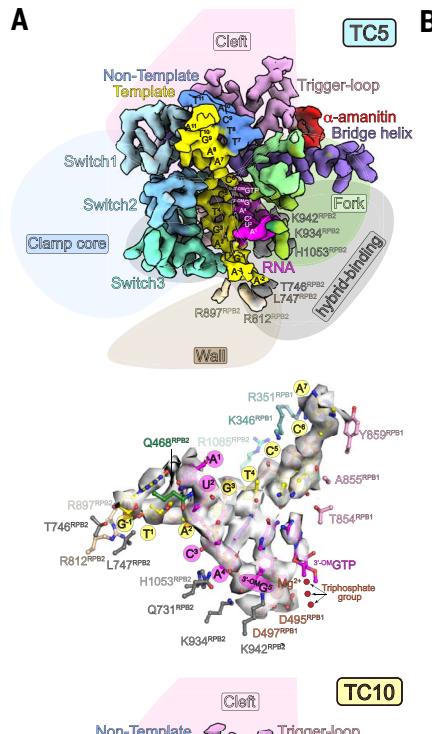


Fig. 4. RNA-DNA hybrid and the B-finger in the active center. (A) Cryo-EM maps of the active center of TC5 and TC10, which are representative ITC and EEC complexes, respectively. Because the active center is deeply buried, surrounding modules are simplified as background. Close-up views show RNA-DNA hybrid and Pol II residues that contribute to hybrid stabilization. Nucleic acids of the template strand and nascent RNA are numbered relative to the TSS. (B) Transparent cryo-EM maps of the RNA-DNA hybrid (left) and B-finger (right) in TC2 to TC9. Cryo-EM maps of the B-ribbon and B-linker are shown to mark the B-finger boundaries [residues Ser⁵⁰ (S50) and Asn⁷² (N72)] that remain similarly positioned throughout TC2 to TC9. The structural model of the B-finger was not built, and a putative trace is shown instead. Cryo-EM density of the B-finger is more upwardly positioned from TC2 to TC6 and is invisible in TC7 to TC9. (C) Cryo-EM map of the B-finger and structural model of the RNA-DNA hybrid in TC3. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; D, Asp; H, His; K, Lys; L, Leu; M, Met; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; and Y, Tyr.

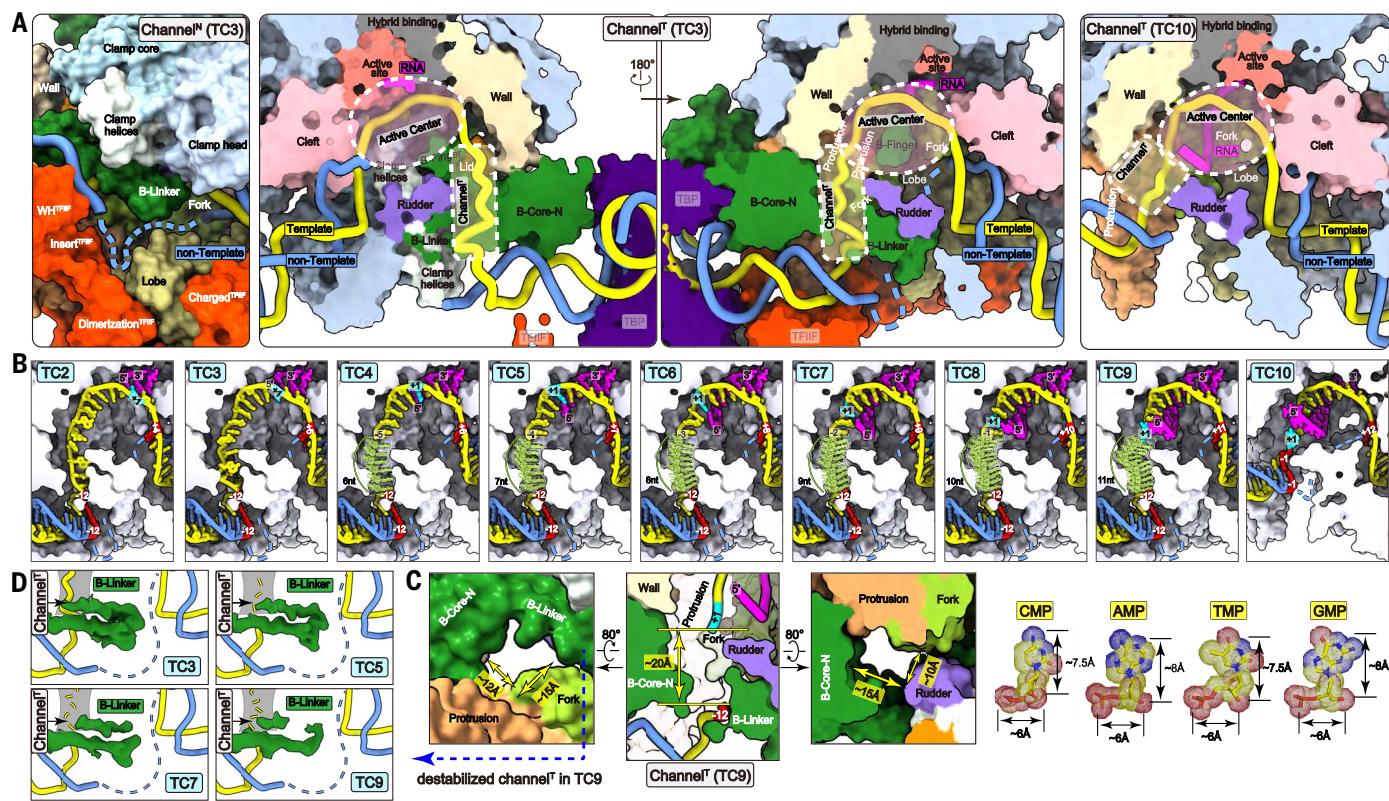


Fig. 5. Template strand accumulation in channel^T and the ITC-EEC transition.

(A) Cross-sectional views of the open channel^N of TC3 (left), the narrow channel^T of TC3 in two different views (middle), and the expended channel^T of TC10 (right). Proteins and nucleic acids are shown in surface and cartoon representations, respectively. The active center and channel^T are circled by white dashed lines. The invisible regions of the nontemplate strand are shown as dashed lines. (B) Cross-sectional views of channel^T in TC2 to TC10 for comparison. DNA and RNA are shown in stick representation. The channel^T is narrow in TC2 to TC9 and largely

expanded in TC10. Accumulation of the disordered template strand in channel^T of the ITC (TC4 to TC9) is diagrammed with green transparent nucleic acids. Each nontemplate strand is shown as a dashed line, and channel^N is covered in this view. Cryo-EM maps of nucleic acids are shown in fig. S10C. (C) Estimated dimensions of channel^T and nucleotides for comparison. Channel^T is shown in three different views. The blue dashed arrow indicates the position of the B-linker in (D). (D) Destabilization of the B-linker at the gate of channel^T in TC9. Cryo-EM maps of the B-linker of TC3, TC5, TC7, and TC9 are shown in comparable contour levels and are colored in green.

retreat out of the active center. As evidenced by its reported effects on transcription activity and TSS selection (including TSS scanning in yeast) (35, 52, 76–80), the placement of the B-finger in the early ITC suggests its supportive role in efficient initiation and promoting the stability of early ITCs.

Exit channels of the template and nontemplate strands

The melted template and nontemplate strands occupy otherwise empty channels, termed channel^T and channel^N, respectively (Fig. 5 and fig. S10). As exemplified by TC3 and TC17 structures, the weak but noticeable density of the non-template strand winds over an open channel^N formed by the Pol II lobe, clamp, and protrusion, which is generally similar to that of the Pol II EC (11, 73) and the Pol III pretermination complex (81). Consistent with the architecture of the open channel^N, the melted non-template strand is highly flexible and less restricted. This is a conserved feature in eukaryotic Pol II and differs from bacterial RNAP, in which

the σ factor restrains the non-template strand and regulates transcription initiation (82, 83).

The cross-sectional view of the ITC shows that the active center merges with the narrow channel^T formed by the B-core-N, the B-linker, and the Pol II active center cleft, hybrid binding, protrusion, rudder, lid, wall, and clamp helices (Fig. 5, A and B; and fig. S10C). The channel^T is nearly sealed, with two ends capped by the RNA-DNA hybrid and the upstream DNA duplex. The narrow channel^T may support the placement of the template strand in the active center that is necessary for RNA synthesis in the early ITC. Compared with the empty channel^T of the PIC, the DNA-occupied channel^T of the ITC is slightly enlarged by the displacement of Pol II protrusion and B-core-N (Fig. 3B), which suggests an adjustable size based on the molecules inside.

In the EEC, TFIIB is invisible and the channel^T is no longer sealed and serves as a wide and short extension of the active center (Fig. 5, A and B; and fig. S10C). The expended channel^T is similar to that of the Pol II EC (11, 73) (fig.

S5E). The transcription bubble is maintained in a fixed length of 11 nt, with the upstream (n-10) and downstream (n+2) edges stabilized by the Pol II rudder and bridge helix, respectively (Figs. 2, A and C, and 5; and fig. S4).

The transition from the ITC to the EEC is correlated with GTF dissociation, bubble collapse, and promoter escape

Comparison of the TC9 and TC10 structures with Pol II superimposed suggests considerable conformational changes in the ITC-EEC transition (Figs. 2 and 6, Movie 1, fig. S11, and movie S1). In TC10, the channel^T is largely expanded, and extensive contacts between GTFs and Pol II are lost. The template and non-template strands at -11/-1 collapse into a straight DNA duplex that merges with the upstream promoter. Pol II is more separated from the TATA box, which is likely unbent because Pol II and TFIIB are necessary for stable promoter bending (16). When the upstream promoter is superimposed as a fixed reference, Pol II undergoes considerable rotation and displacement by

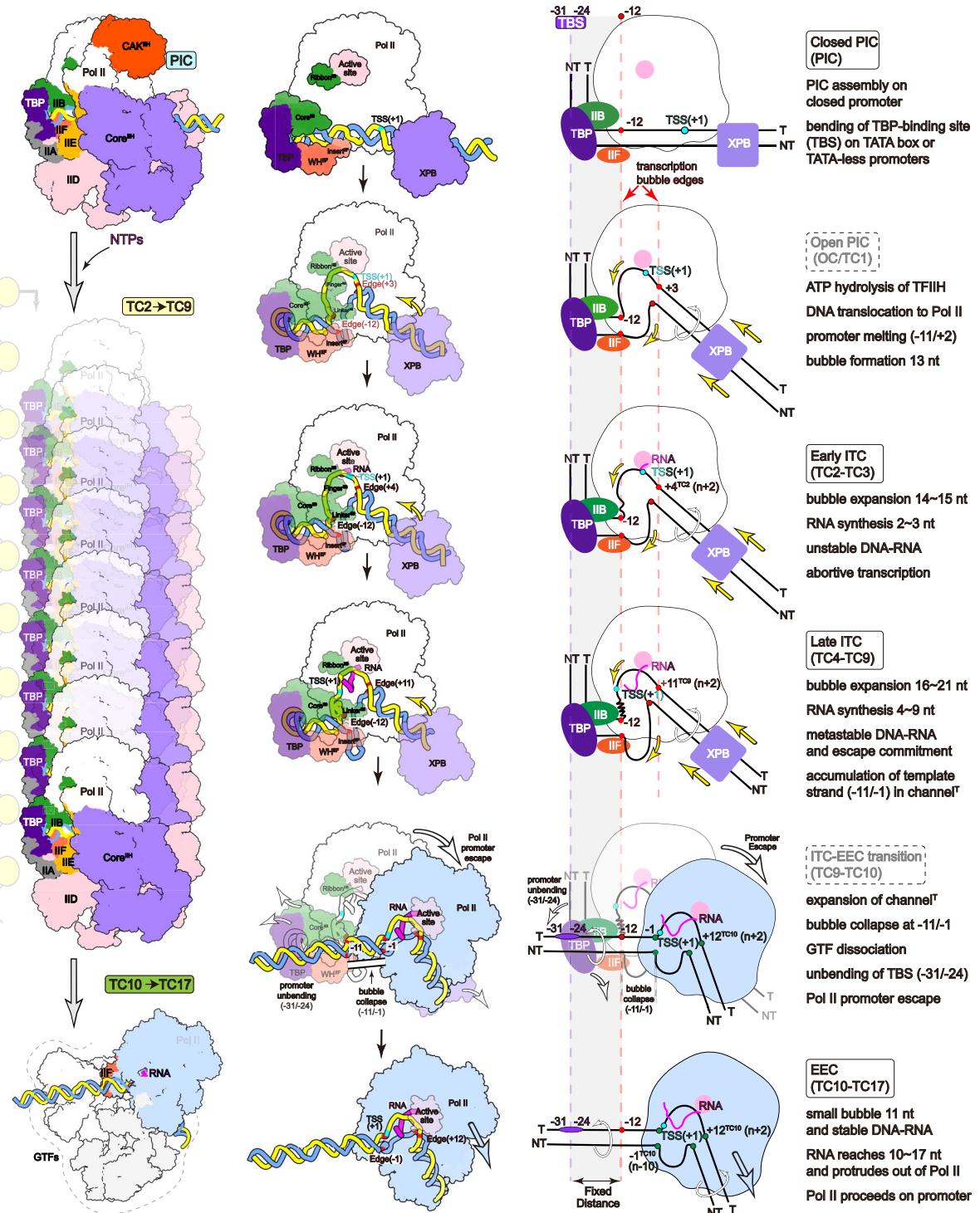


Fig. 6. A working model of transcription initiation. The transitions from the PIC (16) to the ITC and to the EEC are shown in cryo-EM maps of full complexes (left), simplified complexes (middle), and schematic models (right) with short descriptions. T and NT represent template and nontemplate strands, respectively. TBS represents the TBP-binding sites on TATA box and TATA-less promoters. The complexes from top to bottom are shown with the upstream promoter aligned to indicate a transcribing Pol II relative to the promoter. The open PIC and ITC-EEC transition schematics indicate proposed states by connecting the related structures.

~110 Å, which suggests that a process of promoter escape occurred on an active gene. As a result, Pol II proceeds on the promoter and the complex turns into an EEC.

The above result generally agrees with earlier biochemical studies that are not fully consistent

with each other. For example, by detecting changes in the transcription bubble, complex stability, or dissociation of GTFs (mainly TFIIIB), previous studies showed that the functional transition occurs when RNA reaches 7 (35), 8 (34), 9 (37), 10 (33), 7 or 9 (38), 12 or 13 (63),

or more than 9 or 10 nt (36, 57). This apparent discrepancy may result from the heterogeneity of transcription initiation, variations in promoters and experimental contexts, and limitations of the detection method (see supplementary text).

To test the ITC-EEC transition on another promoter, we assembled TC9 and TC10 on the AdML^{9G} and AdML^{10G} promoters, respectively, and determined cryo-EM structures (fig. S12 and data S2 to S4). The two structures are similar to their counterparts on the SCP^{9G} and SCP^{10G} promoters, respectively, which suggests a common point of conformational transition. Thus, comparing TC9 and TC10 structures in a similar experimental context may serve as a framework to unite previous studies and provide structural insights into the dynamic processes of the ITC-EEC transition, a rate-limiting step in transcription initiation (30, 32, 34).

GTfs remain associated with the early EEC, and TFIID dissociates from the late EEC

Cryo-EM maps of the EEC show the density of the Pol II-associated TFIIF and TFIIE (fig. S5B). As previously reported, TFIIE and TFIIF in the EEC are incompatible with DRB sensitivity-inducing factor (DSIF) in a paused elongation complex (PEC) (84) and an elongation factor super elongation complex (SEC) in SEC-Pol II complex (85). Thus, the retained TFIIE and TFIIF may function not only in supporting initiation but also in preventing the negative effects of pausing and elongation factors before the transition from the EEC to the PEC or the productive EC.

To investigate whether other GTfs dissociate from transcribing complexes, we assembled complexes of TC3 (early ITC), TC9 (late ITC), TC10 (early EEC), and TC17 (late EEC), followed by glycerol gradient ultracentrifugation (fig. S13A). Peak fractions (as in cryo-EM sample preparation) of each complex were subjected to Western blotting using antibodies against representative Pol II and GTF subunits, including RPB2 (Pol II subunit), TFIIB, TAF1 and TBP (two TFIID subunits), TFIIE β , TFIIF α , and XPB (TFIIFH subunit). All the GTfs maintained their association with both the early ITC and the late ITC in consistent stoichiometry. TFIID partially dissociated from the early EEC, whereas other GTfs remained in a stable association. In the fractions of the late EEC, TAF1 and TBP of TFIID were nearly undetectable, and there appeared to be a reduction in the levels of TFIIB, TFIIF, TFIIE, and TFIIFH.

The above results agree with our structural observation and provide complementary insights into GTfs in transcribing complexes (fig. S13B). TFIID seems to bind the exposed downstream promoter elements (DPEs) (positions 28 to 32) in TC10, which become buried within Pol II channel E and inaccessible to TFIID in TC17. TFIID in the apo form may bind TBP and inhibit its binding to the promoter (16, 86, 87). TFIIFH may tether with Pol II through the binding of TFIIE and TFIID. TFIIB may disengage from Pol II while maintaining partial connections with TBP, TFIIF, or the TFIIB-recognition element (BRE) of the upstream promoter. The

presence of GTfs within the EEC could potentially vary across different promoters owing to differences in GTF-binding elements. Their retention may be regulated by elongation factors, Mediator, CTD phosphorylation, or chromatin. Nevertheless, the GTfs that are associated with the EEC in peripheral positions might not contribute to the transcription reaction.

Template-strand accumulation in channel T and NTP-driven translocation may promote the ITC-EEC transition

The ITC-EEC transition may result from relative motion between nucleic acids and Pol II, a process of RNA-DNA translocation that is coupled to NTP loading, pyrophosphate release, and conformational changes of the bridge helix and trigger loop (7, 8, 13, 66, 67). We next investigated the correlation between RNA-DNA translocation and the ITC-EEC transition. Cryo-EM maps of TC2 and TC3 revealed the continuous density of template strands of 14 and 15 nt in length, respectively, spanning the upstream and downstream edges of the transcription bubble (Figs. 2A and 5B and fig. S10C). However, in TC4 to TC9, although the template strand becomes longer, only ~10 nt within the active center could be modeled, whereas 6 to 11 nt (in TC4 to TC9, respectively) outside of the active center are invisible. The disordered template strand does not appear to squeeze out of the nearly sealed channel T , and its flexibility suggests nonspecific placement of the unpaired nucleotides in channel T . In TC9, the Pol II active center is occupied by the 9-bp RNA-DNA duplex, and the 11-nt single-stranded template strand (and potentially with the B-finger) is likely accumulated in the channel T . The irregular channel T exhibits a volume of ~10 Å by 15 Å by 20 Å and may not accommodate single-stranded DNA longer than 11 nt (Fig. 5C).

The above analysis suggests a template stand-scrunching mechanism for the transition from the ITC to the EEC (Fig. 5B, Movie 1, fig. S10C, and movie S1). In the late ITC, RNA chain extension and RNA-DNA translocation lead to the accumulation of the template strand in the narrow channel T . The addition of one more nucleotide to the RNA chain and NTP-driven RNA-DNA translocation may provide a final push to balloon out the channel T . Because Pol II remains rigid, the B-linker and B-core-N dissociate from Pol II to release the accumulated template strand.

In support of this hypothesis, an earlier study showed a decrease in the complex stability on the AdML^{6G} to AdML^{9G} promoters (35). Cryo-EM density of the B-linker at the gate of channel T is continuous in TC3 and TC5 but becomes less ordered and disconnected in TC7 and TC9, which indicates the destabilization of channel T (Fig. 5D). Moreover, TFIIB dissociates from Pol II before RNA reaches 10 nt in length (38, 39). Previous studies in yeast Pol II have proposed that steric clashes between RNA and the B-

ribbon may lead to TFIIB dissociation (44, 46). This is not likely the case in mammals because RNA shorter than 12 nt does not reach the B-ribbon (fig. S5D).

More supporting evidence comes from a comparison of the mammalian ITC structures with the structure of the yeast ITC on a G-less promoter (G is at +26) (47) (fig. S6 and supplementary text). The yeast ITC structure represents a backtracked complex containing a 5-nt nascent RNA in the active center. The structure shows a partially open channel T , which may allow the template strand of more than a hundred nucleotides to thread out of the complex for TSS scanning. The differences in channel T (sealed versus open) and the template strand (which is melted within the catalytic pocket by up to 21 nt in the mammalian ITC versus threaded out of the catalytic pocket by more than a hundred nucleotides in yeast ITC) highlight the regulatory role of channel T in potentially controlling bubble size in mammals and suggest distinct underlying mechanisms of transcription initiation in mammals and yeast (48, 49).

Comparison with transcription initiation in bacteria

Our proposed mechanism of transcription initiation in mammals is generally similar to the well-characterized transcription initiation by bacterial RNAP, which forms CC, OC, ITC, and EC complexes and undergoes promoter opening, bubble expansion, bubble collapse, and promoter escape through a DNA-scrunching mechanism (88–91). The bacterial σ factor, which is functionally equivalent to eukaryotic GTfs, associates with RNAP and stabilizes the upstream promoter in the ITC. During promoter escape, the σ factor is forced to dissociate from RNAP by the scrunching DNA in the transcription bubble (92) and by the compressed σ -finger because of RNA growth (93–95). Distinct from what occurs in the mammalian system, the bacterial σ factor recognizes the single-stranded non-template DNA and contributes to promoter escape (82, 83) that is substantially variable dependent on the types of promoters (71, 72).

Discussion

Studies over the past three decades have established that there are functional transitions during eukaryotic transcription initiation. However, understanding this process at the molecular level has been a long-standing challenge because of the difficulty in obtaining de novo transcribing complexes, which are highly dynamic and heterogeneous (30, 54–56). Our work reconciles earlier studies and provides structural visualization of transcription initiation in action (Fig. 6, Movie 1, and movie S1). As detailed above, the transition from the PIC to the early ITC involves promoter melting, transcription bubble formation, nascent RNA synthesis, and abortive transcription. From the early ITC to the late ITC, the transcription bubble expands and RNA

is more stable within Pol II, which is correlated with escape commitment. From the late ITC to the EEC, the complex undergoes abrupt transitions, including channel^T expansion, GTF dissociation, bubble collapse, and promoter escape.

The proposed model may be applicable to Pol II-mediated transcription initiation on TATA-less promoters, which account for most core promoters (1). TATA-less promoters tend to exhibit lower transcriptional activity than TATA box promoters and often result in heterogeneous TSSs, limiting their usage in biochemical and structural studies. Our previous study showed that during PIC assembly, TFIID supports TBP to load and bend the upstream promoter region, termed the TBP-binding site (TBS), regardless of whether it contains a TATA box or not (16). PIC complexes assembled on the TATA box and TATA-less promoters adopt similar organizations, which suggests a general mechanism of transcription initiation. Although the selection of a TSS is fine-tuned by the surrounding sequences and the inherent heterogeneity (50–52), transcription seems to be started ~31 bp downstream of the TBS (35, 48), which is primarily determined by the site of PIC assembly, which is dependent on promoter sequence and chromatin organization (1, 14).

Our study also provides a structural basis for understanding why transcription initiation undergoes such notable changes. To avoid promiscuous transcription, Pol II by itself could neither find a target promoter nor open a promoter to initiate RNA synthesis. This is why GTFs are necessary for promoter recognition and transcription initiation. However, the stability that results from the multiple interactions among GTFs, Pol II, and the promoter also creates an obstacle for Pol II to escape the promoter and therefore must eventually be disrupted. The modular organization of the PIC and the ITC is well suited to achieve this purpose. In the early ITC, the active center and channel^T can accommodate the template strand and allow for bubble expansion and initial RNA synthesis. In the late ITC, the RNA-DNA hybrid is stable and long enough to support NTP-driven translocation and disruption of the GTF-Pol II-promoter contacts. In the EEC, the peripherally positioned GTFs would not restrain Pol II on the promoter. Therefore, coordinated changes of the GTFs and the transcription bubble support a functional transition from the ITC, which requires GTFs and promoter opening for initial RNA synthesis, to the EEC, which requires GTF dissociation and bubble collapse for promoter escape and stable RNA production.

Materials and methods

Cryo-EM sample preparation

Pol II was isolated from *Sus scrofa* thymus, and *Homo sapiens* TFIID, TFIIA, TFIIB, TFIIF, TFIIE, and TFIIH were prepared as previously described (15, 16). Promoter sequences (SCP^{NG}

and AdML^{9G/10G}) are listed in data S1. The promoter DNA was generated by annealing the template strand with equimolar amounts of single-stranded nontemplate DNA at a final concentration of 50 μM in water. The annealing reaction was carried out at 95°C for 10 min and gradually cooled down to 25°C within 2 hours.

To prepare each TC complex, the purified PIC components of 300 pmol of TFIID, 1500 pmol of TFIIA, and 450 pmol of promoter DNA (SCP^{NG} and AdML^{9G/10G}) were mixed at 4°C for 2 hours in a 150-μl mixture containing 25 mM HEPES-KOH pH 7.9, 100 mM KCl, 6 mM MgCl₂, 2 mM DTT, and 5% (v/v) glycerol. After incubation, the mixture was subjected to glycerol gradient ultracentrifugation. The gradient was generated from equal volumes of a 10% glycerol solution [10% (w/v) glycerol, 30 mM HEPES-KOH pH 7.9, 50 mM KCl, 6 mM MgCl₂, 2 mM DTT] and a 50% glycerol solution [50% (w/v) glycerol, 30 mM HEPES-KOH pH 7.9, 50 mM KCl, 6 mM MgCl₂, 2 mM DTT] with a BioComp Gradient Master 108 (BioComp Instruments). The ultrafiltration was performed using a swinging bucket rotor SW60Ti (Beckman) at 36,000 rpm for 14 hours at 4°C. Peak fractions containing promoter DNA-TFIID-TFIIA were pooled and concentrated to ~1.5 mg/ml. Subsequently, the 100 pmol of preassembled promoter DNA-TFIID-TFIIA was mixed with 500 pmol of TFIIB, 110 pmol of Pol II, and 160 pmol of TFIIF at 4°C for 1 hour in a 150-μl mixture containing 25 mM HEPES-KOH pH 7.9, 75 mM KCl, 6 mM MgCl₂, 2 mM DTT, and 5% (v/v) glycerol. Next, 300 pmol of TFIIE and 200 pmol of TFIIH were added to the mixture, and it was further incubated for 2 hours to assemble the PIC complex. After the incubation, 15 μl of 20× NTP mix [containing 40 mM ATP, 40 mM CTP, 40 mM UTP, 2 mM 3'-O-methylguanosine-5'-triphosphate (APExBIO)] was added to the 285 μl of PIC mixture to initiate the transcription reaction. The reactions were carried out for 50 min at 30°C and then quenched by adding 0.5 μl of 0.4 mM α-amanitin, followed by GraFix (96). For Grafix, the glycerol gradient was prepared using a low buffer containing 15% (v/v) glycerol, 30 mM HEPES-KOH pH 7.9, 100 mM KCl, 2 mM MgCl₂, 10 μM ZnCl₂, and 2 mM DTT and a high buffer containing 35% (v/v) glycerol, 30 mM HEPES-KOH pH 7.9, 100 mM KCl, 2 mM MgCl₂, 10 μM ZnCl₂, 2 mM DTT, and 0.01% glutaraldehyde. The samples were centrifuged at 38,000 rpm at 4°C for 14 hours using an SW41 Ti rotor (Beckman Coulter). Subsequently, fractions containing cross-linked complexes were quenched with 50 mM Tris pH 7.4 (25°C). The peak fractions were assessed by negative-stain EM. Fractions of interest were pooled, concentrated, and subjected to buffer exchange into a buffer containing 30 mM HEPES-KOH pH 7.9, 100 mM KCl, 2 mM MgCl₂, 10 μM ZnCl₂, 2 mM DTT, and 0.8% (v/v) glycerol using a 100-kDa cut-off centrifugation filter unit (Amicon Ultra). For

each assembly step, the samples were monitored by SDS-PAGE and visualized by Coomassie blue staining. TC3, TC9, TC10, and TC17 were assembled and subjected to glycerol gradient ultracentrifugation without cross-linking, and peak fractions (as in the cryo-EM sample) of each complex were subjected to Western blotting using the indicated antibodies (anti-TAF1 antibody: A303-505A, Bethyl; anti-RPB2 antibody: A5928, Abclonal; anti-XPB antibody: 10580-1-AP, Proteintech; anti-TFIIFα antibody: 10093-2-AP, Proteintech; anti-TBP antibody: A2192, Abclonal; anti-TFIIEβ antibody: A305-054A, Bethyl; and anti-TFIIB antibody: A1708, Abclonal).

For cryo-EM grid preparation, 3 μl of the sample at a concentration of ~0.2 μM was applied to freshly glow-discharged Quantifoil R1.2/1.3 Au holey carbon grids. After incubation of 5 s at a temperature of 4°C and a humidity of 100%, the grids were blotted for 0.5 to 1.0 s in Vitrobot Mark IV (Thermo Fisher Scientific), plunge-frozen in liquid ethane, and stored at liquid nitrogen temperature. The grids were prepared in the H₂/O₂ mixture for 15 s using a Gatan 950 Solarus plasma cleaning system with a power of 5 W. The ø 55/20-mm blotting paper is made by TED PELLA and used for plunge freezing.

In vitro transcription initiation assay

The transcription reaction was conducted by mixing 0.4 pmol of promoter DNA, 0.5 pmol of TFIID, 1.5 pmol of TFIIA, 1.5 pmol of TFIB, 1 pmol of RNA Pol II, 1.5 pmol of TFIIF, 1.5 pmol of TFIIE, and 0.8 pmol of TFIIH in an 18.8-μl reaction mixture containing 25 mM HEPES-KOH pH 7.9, 50 mM KCl, 5 mM MgCl₂, 1 mM DTT, 10% (v/v) glycerol, 200 μg/ml bovine serum albumin (BSA), and 8 units of RiboLock RNase Inhibitor (Thermo Fisher Scientific). The mixture was incubated at 25°C for 30 min. After the incubation, reactions were initiated by the addition of 1 μl of 20× NTP mix containing 1.2 mM ATP, 0.2 mM CTP, 0.02 mM UTP, 2.4 mM 3'-O-methylguanosine-5'-triphosphate (APExBIO), and 0.2 μl of 10 mCi/ml [α -³²P]-UTP (Perkin Elmer). The transcription reactions were carried out for 50 min at 30°C and stopped by heating the reaction mixture to 68°C for 3 min. In the time-course assay, the reactions were stopped at time intervals of 5, 10, 30, and 50 min. Subsequently, after cooling on ice, the reaction mixtures were treated with 1 μl of calf intestinal alkaline phosphatase (CIAP) (20 U/ml, Thermo Fisher Scientific) at 37°C for 30 min. The RNA products were analyzed by 23% urea polyacrylamide gels, and the autoradiograph was obtained by storage-phosphor scanning (Typhoon; Cytiva).

Cryo-EM data collection and image processing

Detailed parameters of data collection and image processing are shown in data S2 and S3 and summarized in data S4. A brief description is provided below, and fig. S2 shows the

data collection and image processing of TC5 and TC10, two representative complexes.

The cryo-EM grids were loaded onto a Titan Krios transmission electron microscope (Thermo Fisher Scientific) operated at 300 kV for data collection. The cryo-EM images were automatically recorded by a post-GIF Gatan K3 Summit direct electron detector in the super-resolution counting mode using Serial-EM with a nominal magnification of 64,000 \times in the EFTEM mode, which yielded a super-resolution pixel size of 0.667 Å on the image plane, and with defocus values ranging from -1.5 to -2.5 μm. Each micrograph stack was dose-fractionated to 40 frames with an exposure rate of 21 e⁻ pixel⁻¹ s⁻¹, a total electron exposure of ~50 e⁻ Å⁻², and a total exposure time of 4.2 s.

For TC5 and TC10 complexes, 49,009/58,688 ("/" separates parameters in TC5 and TC10) micrographs were collected for further processing. Drift and beam-induced motion correction were applied on the super-resolution movie stacks using MotionCor2 (97) and binned twofold to a calibrated pixel size of 1.334 Å pixel⁻¹. The defocus values were estimated by Getf (98) from summed images without dose weighting. Other procedures of cryo-EM data processing were performed within RELION v3.1 (99, 100) and cryoSPARC v3 (101) using the dose-weighted micrographs. Particles of 28,786,688/33,235,621 were picked by reference autopicking in RELION using high-resolution Pol II map [Electron Microscopy Data Bank (EMDB) ID 3218] as a 3D reference, 0.05 picking threshold, and 180 minimum interparticle distance, with other parameters set to default, and were subjected to reference-free 2D classification using a mask diameter of 420 Å, iterations of 30, Tau2Fudge factor of 2, classes of 200, e-step of -1 Å, and ignore CTFs until first peak (Yes) (other parameters were set to default). Particles of 19,288,625/21,463,872 were selected from good 2D classes for 3D classification in RELION.

All the 3D classification and 3D autorefine in RELION used an initial low-pass filter of 40 Å, Tau2Fudge factor of 4, iterations of 160, classes of 9, and e-step of -1 Å (other parameters were set to default). Particles of 12,743,286/12,467,838 were selected from 3D classes using a mask diameter of 420 Å. The yielded particles were applied for subtraction with a Pol II mask. The subtracted particles were applied for 3D classification using a mask diameter of 280 Å to get the ITC and EEC conformations, respectively. The particles of the ITC and EEC were used for refinement in cryoSPARC Local Refinement, yielding a reconstruction at 2.98 Å/2.82 Å resolution for Pol II (including the bound nucleic acids). All the refinement of Pol II has been subjected to CTF refinement, postprocessing, and Bayesian polishing in RELION. The subtracted particles of Pol II in ITC and EEC conformation were reverted to original particles,

followed by 3D classification to obtain overall maps of ITCs and EECs. The overall maps were refined with a box size of 200 pixels and a down-sampled pixel size of 2.668 Å.

All reported resolutions are based on the gold-standard Fourier shell correlation (FSC) = 0.143 criterion. The gold-standard FSC curves were corrected for the effects of a soft mask with high-resolution noise substitution. We also used the 3D FSC server to evaluate our maps. All the visualization and evaluation of the 3D volume map were performed using UCSF Chimera (102), and the local resolution variations were calculated using cryoSPARC.

Model building and structure refinement

The overall cryo-EM maps and locally refined maps of all samples were used for model building. The structures of the mammalian PIC complex [Protein Data Bank (PDB) ID 7EGB] (15) and mammalian elongation complex (PDB ID 5FLM) (11) were used as initial structural templates. For ITC complexes, structural modules of Pol II and GTFs were individually docked into the cryo-EM maps by rigid-body fitting using UCSF Chimera. For EEC complexes, Pol II and TFIIF were docked into the cryo-EM maps, whereas other GTFs were not built because the density did not support confident model docking. The structural models were manually adjusted in COOT (103).

Overfitting of the model was monitored by refining the model in one of the two half maps from the gold-standard refinement approach and testing the refined model against the other map. The structural models (except the flexibly docked TFIIE, TFIIF, and TFIID) were evaluated using Phenix real-space refine and MolProbity (104) using resolution at 6.5 Å for the ITC and 3.5 Å for the EEC and constraints of rotamer, Ramachandran, and secondary structure. Map and model representations in the figures and movies were prepared by PyMOL (105), UCSF Chimera, or UCSF ChimeraX (102).

REFERENCES AND NOTES

- V. Haberle, A. Stark, Eukaryotic core promoters and the functional basis of transcription initiation. *Nat. Rev. Mol. Cell Biol.* **19**, 621–637 (2018). doi: 10.1038/s41580-018-0028-8; pmid: 29946135
- D. S. Luse, Promoter clearance by RNA polymerase II. *Biochim. Biophys. Acta* **1829**, 63–68 (2013). doi: 10.1016/j.bbaram.2012.08.010; pmid: 22982364
- A. Saunders, L. J. Core, J. T. Lis, Breaking barriers to transcription elongation. *Nat. Rev. Mol. Cell Biol.* **7**, 557–567 (2006). doi: 10.1038/nrm1981; pmid: 16936696
- R. G. Roeder, The role of general initiation factors in transcription by RNA polymerase II. *Trends Biochem. Sci.* **21**, 327–335 (1996). doi: 10.1016/0968-0004(96)10050-5; pmid: 8870495
- L. Zawel, D. Reinberg, Initiation of transcription by RNA polymerase II: A multi-step process. *Prog. Nucleic Acid Res. Mol. Biol.* **44**, 67–108 (1993). doi: 10.1016/S0079-6603(08)60217-2; pmid: 8434126
- M. C. Thomas, C. M. Chiang, The general transcription machinery and general cofactors. *Crit. Rev. Biochem. Mol. Biol.* **41**, 105–178 (2006). doi: 10.1080/10409230600648736; pmid: 16858867
- A. L. Gnatt, P. Cramer, J. Fu, D. A. Bushnell, R. D. Kornberg, Structural basis of transcription: An RNA polymerase II elongation complex at 3.3 Å resolution. *Science* **292**, 1876–1882 (2001). doi: 10.1126/science.1059495; pmid: 11313499
- D. Wang, D. A. Bushnell, K. D. Westover, C. D. Kaplan, R. D. Kornberg, Structural basis of transcription: Role of the trigger loop in substrate specificity and catalysis. *Cell* **127**, 941–954 (2006). doi: 10.1016/j.cell.2006.11.023; pmid: 17129781
- K. D. Westover, D. A. Bushnell, R. D. Kornberg, Structural basis of transcription: Nucleotide selection by rotation in the RNA polymerase II active center. *Cell* **119**, 481–489 (2004). doi: 10.1016/j.cell.2004.10.016; pmid: 15537538
- K. D. Westover, D. A. Bushnell, R. D. Kornberg, Structural basis of transcription: Separation of RNA from DNA by RNA polymerase II. *Science* **303**, 1014–1016 (2004). doi: 10.1126/science.1090839; pmid: 14963331
- C. Bernnecky, F. Herzog, W. Baumeister, J. M. Plitzko, P. Cramer, Structure of transcribing mammalian RNA polymerase II. *Nature* **529**, 551–554 (2016). doi: 10.1038/nature16482; pmid: 26789250
- A. C. Cheung, P. Cramer, Structural basis of RNA polymerase II backtracking, arrest and reactivation. *Nature* **471**, 249–253 (2011). doi: 10.1038/nature09785; pmid: 21346759
- P. Cramer, D. A. Bushnell, R. D. Kornberg, Structural basis of transcription: RNA polymerase II at 2.8 Ångstrom resolution. *Science* **292**, 1863–1876 (2001). doi: 10.1126/science.1059493; pmid: 11313498
- X. Chen et al., Structures of +1 nucleosome-bound PIC-Mediator complex. *Science* **378**, 62–68 (2022). doi: 10.1126/science.abn8131; pmid: 36201575
- X. Chen et al., Structures of the human Mediator and Mediator-bound preinitiation complex. *Science* **372**, eabg0635 (2021). doi: 10.1126/science.abg0635; pmid: 33958484
- X. Chen et al., Structural insights into preinitiation complex assembly on core promoters. *Science* **372**, eaba8490 (2021). doi: 10.1126/science.aba8490; pmid: 33795473
- S. Schilbach, S. Albara, C. Dienemann, F. Grubbe, P. Cramer, Structure of RNA polymerase II pre-initiation complex at 2.9 Å defines initial DNA opening. *Cell* **184**, 4064–4072.e28 (2021). doi: 10.1016/j.cell.2021.05.012; pmid: 34133942
- S. Rengachari, S. Schilbach, S. Albara, C. Dienemann, P. Cramer, Structure of the human Mediator-RNA polymerase II pre-initiation complex. *Nature* **594**, 129–133 (2021). doi: 10.1038/s41586-021-03555-7; pmid: 33902108
- S. Albara, S. Schilbach, P. Cramer, Structures of mammalian RNA polymerase II pre-initiation complexes. *Nature* **594**, 124–128 (2021). doi: 10.1038/s41586-021-03554-8; pmid: 33902107
- R. Abdella et al., Structure of the human Mediator-bound transcription preinitiation complex. *Science* **372**, 52–56 (2021). doi: 10.1126/science.abg3074; pmid: 33707221
- S. Schilbach et al., Structures of transcription pre-initiation complex with TFIIF and Mediator. *Nature* **551**, 204–209 (2017). doi: 10.1038/nature24282; pmid: 29088706
- Y. He, J. Fang, D. J. Taatjes, E. Nogales, Structural visualization of key steps in human transcription initiation. *Nature* **495**, 481–486 (2013). doi: 10.1038/nature11991; pmid: 23446344
- P. J. Robinson et al., Structure of a complete Mediator-RNA polymerase II pre-initiation complex. *Cell* **166**, 1411–1422.e16 (2016). doi: 10.1016/j.cell.2016.08.050; pmid: 27610567
- S. Rengachari et al., Structural basis of SNAPc-dependent snRNA transcription initiation by RNA polymerase II. *Nat. Struct. Mol. Biol.* **29**, 1159–1169 (2022). doi: 10.1038/s41994-022-00857-w; pmid: 36424526
- F. C. Holstege, P. C. van der Vliet, H. T. Timmers, Opening of an RNA polymerase II promoter occurs in two distinct steps and requires the basal transcription factors IIE and IIH. *EMBO J.* **15**, 1666–1677 (1996). doi: 10.1002/j.1460-2075.1996.tb00512.x; pmid: 8612591
- H. Cai, D. S. Luse, Transcription initiation by RNA polymerase II in vitro. Properties of preinitiation, initiation, and elongation complexes. *J. Biol. Chem.* **262**, 298–304 (1987). doi: 10.1016/S0021-9258(19)75926-2; pmid: 2432061
- D. S. Luse, G. A. Jacob, Abortive initiation by RNA polymerase II in vitro at the adenovirus 2 major late promoter. *J. Biol. Chem.* **262**, 14990–14997 (1987). doi: 10.1016/S0021-9258(18)48127-6; pmid: 3667620
- D. S. Luse, T. Kochel, E. D. Kuempel, J. A. Coppola, H. Cai, Transcription initiation by RNA polymerase II in vitro. At least two nucleotides must be added to form a stable ternary complex. *J. Biol. Chem.* **262**, 289–297 (1987). doi: 10.1016/S0021-9258(19)75925-0; pmid: 2432060
- H. A. Ferguson, J. F. Kugel, J. A. Goodrich, Kinetic and mechanistic analysis of the RNA polymerase II transcription

- reaction at the human interleukin-2 promoter. *J. Mol. Biol.* **314**, 993–1006 (2001). doi: [10.1006/jmbi.2000.5215](https://doi.org/10.1006/jmbi.2000.5215); pmid: [11743717](https://pubmed.ncbi.nlm.nih.gov/11743717/)
30. J. F. Kugel, J. A. Goodrich, Promoter escape limits the rate of RNA polymerase II transcription and is enhanced by TFIIE, TFIIB, and ATP on negatively supercoiled DNA. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 9232–9237 (1998). doi: [10.1073/pnas.95.16.9232](https://doi.org/10.1073/pnas.95.16.9232); pmid: [9689063](https://pubmed.ncbi.nlm.nih.gov/9689063/)
31. J. F. Kugel, J. A. Goodrich, Translocation after synthesis of a four-nucleotide RNA commits RNA polymerase II to promoter escape. *Mol. Cell. Biol.* **22**, 762–773 (2002). doi: [10.1128/MCB.22.3.762-773.2002](https://doi.org/10.1128/MCB.22.3.762-773.2002); pmid: [11784853](https://pubmed.ncbi.nlm.nih.gov/11784853/)
32. J. F. Kugel, J. A. Goodrich, A kinetic model for the early steps of RNA synthesis by human RNA polymerase II. *J. Biol. Chem.* **275**, 40483–40491 (2000). doi: [10.1074/jbc.M006401200](https://doi.org/10.1074/jbc.M006401200); pmid: [10982810](https://pubmed.ncbi.nlm.nih.gov/10982810/)
33. F. C. Holstege, U. Fiedler, H. T. Timmers, Three transitions in the RNA polymerase II transcription complex during initiation. *EMBO J.* **16**, 7468–7480 (1997). doi: [10.1093/emboj/16.24.7468](https://doi.org/10.1093/emboj/16.24.7468); pmid: [9405375](https://pubmed.ncbi.nlm.nih.gov/9405375/)
34. A. R. Hieb, S. Baran, J. A. Goodrich, J. F. Kugel, An 8 nt RNA triggers a rate-limiting shift of RNA polymerase II complexes into elongation. *EMBO J.* **25**, 3100–3109 (2006). doi: [10.1038/sj.emboj.7601197](https://doi.org/10.1038/sj.emboj.7601197); pmid: [16778763](https://pubmed.ncbi.nlm.nih.gov/16778763/)
35. M. Pal, A. S. Ponticelli, D. S. Luse, The role of the transcription bubble and TFIIB in promoter clearance by RNA polymerase II. *Mol. Cell* **19**, 101–110 (2005). doi: [10.1016/j.molcel.2005.05.024](https://doi.org/10.1016/j.molcel.2005.05.024); pmid: [15989968](https://pubmed.ncbi.nlm.nih.gov/15989968/)
36. J. A. Coppola, D. S. Luse, Purification and characterization of ternary complexes containing accurately initiated RNA polymerase II and less than 20 nucleotides of RNA. *J. Mol. Biol.* **178**, 415–437 (1984). doi: [10.1016/0022-2836\(84\)90151-7](https://doi.org/10.1016/0022-2836(84)90151-7); pmid: [6492155](https://pubmed.ncbi.nlm.nih.gov/6492155/)
37. M. Pal, D. S. Luse, The initiation-elongation transition: Lateral mobility of RNA in RNA polymerase II complexes is greatly reduced at +8/+9 and absent by +23. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 5700–5705 (2003). doi: [10.1073/pnas.1037057100](https://doi.org/10.1073/pnas.1037057100); pmid: [12719526](https://pubmed.ncbi.nlm.nih.gov/12719526/)
38. E. Ly, A. E. Powell, J. A. Goodrich, J. F. Kugel, Release of human TFIIB from actively transcribing complexes is triggered upon synthesis of 7- and 9-nt RNAs. *J. Mol. Biol.* **432**, 4049–4060 (2020). doi: [10.1016/j.jmb.2020.05.005](https://doi.org/10.1016/j.jmb.2020.05.005); pmid: [32417370](https://pubmed.ncbi.nlm.nih.gov/32417370/)
39. L. Zawel, K. P. Kumar, D. Reinberg, Recycling of the general transcription factors during RNA polymerase II transcription. *Genes Dev.* **9**, 1479–1490 (1995). doi: [10.1101/gad.9.12.1479](https://doi.org/10.1101/gad.9.12.1479); pmid: [7601352](https://pubmed.ncbi.nlm.nih.gov/7601352/)
40. X. Liu, D. A. Bushnell, D. Wang, G. Calero, R. D. Kornberg, Structure of an RNA polymerase II-TFIIB complex and the transcription initiation mechanism. *Science* **327**, 206–209 (2010). doi: [10.1126/science.1182015](https://doi.org/10.1126/science.1182015); pmid: [19965383](https://pubmed.ncbi.nlm.nih.gov/19965383/)
41. X. Liu, D. A. Bushnell, D. A. Silva, X. Huang, R. D. Kornberg, Initiation complex structure and promoter proofreading. *Science* **333**, 633–637 (2011). doi: [10.1126/science.1206629](https://doi.org/10.1126/science.1206629); pmid: [21798951](https://pubmed.ncbi.nlm.nih.gov/21798951/)
42. A. C. Cheung, S. Sainsbury, P. Cramer, Structural basis of initial RNA polymerase II transcription. *EMBO J.* **30**, 4755–4763 (2011). doi: [10.1038/embj.2011.396](https://doi.org/10.1038/embj.2011.396); pmid: [22056778](https://pubmed.ncbi.nlm.nih.gov/22056778/)
43. Y. He et al., Near-atomic resolution visualization of human transcription promoter opening. *Nature* **533**, 359–365 (2016). doi: [10.1038/nature17970](https://doi.org/10.1038/nature17970); pmid: [27193682](https://pubmed.ncbi.nlm.nih.gov/27193682/)
44. S. Sainsbury, J. Niesser, P. Cramer, Structure and function of the initially transcribing RNA polymerase II-TFIIB complex. *Nature* **493**, 437–440 (2013). doi: [10.1038/nature11715](https://doi.org/10.1038/nature11715); pmid: [23151482](https://pubmed.ncbi.nlm.nih.gov/23151482/)
45. D. A. Bushnell, K. D. Westover, R. E. Davis, R. D. Kornberg, Structural basis of transcription: An RNA polymerase II-TFIIB cocrystal at 4.5 angstroms. *Science* **303**, 983–988 (2004). doi: [10.1126/science.1090838](https://doi.org/10.1126/science.1090838); pmid: [14963322](https://pubmed.ncbi.nlm.nih.gov/14963322/)
46. D. Kostrewa et al., RNA polymerase II-TFIIB structure and mechanism of transcription initiation. *Nature* **462**, 323–330 (2009). doi: [10.1038/nature08548](https://doi.org/10.1038/nature08548); pmid: [19820686](https://pubmed.ncbi.nlm.nih.gov/19820686/)
47. C. Yang et al., Structural visualization of de novo transcription initiation by *Saccharomyces cerevisiae* RNA polymerase II. *Mol. Cell* **82**, 660–676.e9 (2022). doi: [10.1016/j.molcel.2021.10.020](https://doi.org/10.1016/j.molcel.2021.10.020); pmid: [35051353](https://pubmed.ncbi.nlm.nih.gov/35051353/)
48. C. Giardina, J. T. Lis, DNA melting on yeast RNA polymerase II promoters. *Science* **261**, 759–762 (1993). doi: [10.1126/science.8342041](https://doi.org/10.1126/science.8342041); pmid: [8342041](https://pubmed.ncbi.nlm.nih.gov/8342041/)
49. F. M. Fazal, C. A. Meng, K. Murakami, R. D. Kornberg, S. M. Block, Real-time observation of the initiation of RNA polymerase II transcription. *Nature* **525**, 274–277 (2015). doi: [10.1038/nature14882](https://doi.org/10.1038/nature14882); pmid: [26331540](https://pubmed.ncbi.nlm.nih.gov/26331540/)
50. L. Vo ngoc, C. J. Cassidy, C. Y. Huang, S. H. Duttke, J. T. Kadonaga, The human initiator is a distinct and abundant element that is precisely positioned in focused core promoters. *Genes Dev.* **31**, 6–11 (2017). doi: [10.1101/gad.293837.116](https://doi.org/10.1101/gad.293837.116); pmid: [28108472](https://pubmed.ncbi.nlm.nih.gov/28108472/)
51. M. Pal, D. S. Luse, Strong natural pausing by RNA polymerase II within 10 bases of transcription start may result in repeated slippage and reextension of the nascent RNA. *Mol. Cell. Biol.* **22**, 30–40 (2002). doi: [10.1128/MCB.22.1.30-40.2002](https://doi.org/10.1128/MCB.22.1.30-40.2002); pmid: [11739720](https://pubmed.ncbi.nlm.nih.gov/11739720/)
52. J. A. Fairley, R. Evans, N. A. Hawkes, S. G. Roberts, Core promoter-dependent TFIIB conformation and a role for TFIIB conformation in transcription start site selection. *Mol. Cell. Biol.* **22**, 6697–6705 (2002). doi: [10.1128/MCB.22.19.6697-6705.2002](https://doi.org/10.1128/MCB.22.19.6697-6705.2002); pmid: [12215527](https://pubmed.ncbi.nlm.nih.gov/12215527/)
53. W. Gu, M. Wind, D. Reines, Increased accommodation of nascent RNA in a product site on RNA polymerase II during arrest. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6935–6940 (1996). doi: [10.1073/pnas.93.14.6935](https://doi.org/10.1073/pnas.93.14.6935); pmid: [8692922](https://pubmed.ncbi.nlm.nih.gov/8692922/)
54. R. T. Kamakaka, C. M. Tyree, J. T. Kadonaga, Accurate and efficient RNA polymerase II transcription with a soluble nuclear fraction derived from *Drosophila* embryos. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 1024–1028 (1991). doi: [10.1073/pnas.88.3.1024](https://doi.org/10.1073/pnas.88.3.1024); pmid: [199453](https://pubmed.ncbi.nlm.nih.gov/199453/)
55. J. T. Kadonaga, Assembly and disassembly of the *Drosophila* RNA polymerase II complex during transcription. *J. Biol. Chem.* **265**, 2624–2631 (1990). doi: [10.1016/S0021-9258\(19\)39847-3](https://doi.org/10.1016/S0021-9258(19)39847-3); pmid: [1689291](https://pubmed.ncbi.nlm.nih.gov/1689291/)
56. A. E. Horn, J. F. Kugel, J. A. Goodrich, Single molecule microscopy reveals mechanistic insight into RNA polymerase II preinitiation complex assembly and transcriptional activity. *Nucleic Acids Res.* **44**, 7132–7143 (2016). doi: [10.1093/nar/gkw321](https://doi.org/10.1093/nar/gkw321); pmid: [2712574](https://pubmed.ncbi.nlm.nih.gov/2712574/)
57. A. Dvir, R. C. Conaway, J. W. Conaway, Promoter escape by RNA polymerase II. A role for an ATP cofactor in suppression of arrest by polymerase at promoter-proximal sites. *J. Biol. Chem.* **271**, 23352–23356 (1996). doi: [10.1074/jbc.271.38.23352](https://doi.org/10.1074/jbc.271.38.23352); pmid: [8798537](https://pubmed.ncbi.nlm.nih.gov/8798537/)
58. T. K. Kim, R. H. Ebright, D. Reinberg, Mechanism of ATP-dependent promoter melting by transcription factor IIH. *Science* **288**, 1418–1421 (2000). doi: [10.1126/science.288.5470.1418](https://doi.org/10.1126/science.288.5470.1418); pmid: [10827951](https://pubmed.ncbi.nlm.nih.gov/10827951/)
59. M. Douziech et al., Mechanism of promoter melting by the xeroderma pigmentosum complementation group B helicase of transcription factor IIH revealed by protein-DNA photo-cross-linking. *Mol. Cell. Biol.* **20**, 8168–8177 (2000). doi: [10.1128/MCB.20.21.8168-8177.2000](https://doi.org/10.1128/MCB.20.21.8168-8177.2000); pmid: [11027286](https://pubmed.ncbi.nlm.nih.gov/11027286/)
60. A. Dvir, R. C. Conaway, J. W. Conaway, A role for TFIIB in controlling the activity of early RNA polymerase II elongation complexes. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 9006–9010 (1997). doi: [10.1073/pnas.94.17.9006](https://doi.org/10.1073/pnas.94.17.9006); pmid: [9256425](https://pubmed.ncbi.nlm.nih.gov/9256425/)
61. J. A. Goodrich, R. Tjian, Transcription factors IIE and IIH and ATP hydrolysis direct promoter clearance by RNA polymerase II. *Cell* **77**, 145–156 (1994). doi: [10.1016/0092-8674\(94\)90242-9](https://doi.org/10.1016/0092-8674(94)90242-9); pmid: [8156590](https://pubmed.ncbi.nlm.nih.gov/8156590/)
62. L. Spangler, X. Wang, J. W. Conaway, R. C. Conaway, A. Dvir, TFIIB action in transcription initiation and promoter escape requires distinct regions of downstream promoter DNA. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 5544–5549 (2001). doi: [10.1073/pnas.011004498](https://doi.org/10.1073/pnas.011004498); pmid: [11331764](https://pubmed.ncbi.nlm.nih.gov/11331764/)
63. P. Čabárt, A. Újvári, M. Pal, D. S. Luse, Transcription factor TFIIF is not required for initiation by RNA polymerase II, but it is essential to stabilize transcription factor TFIIB in early elongation complexes. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 15786–15791 (2011). doi: [10.1073/pnas.1104591108](https://doi.org/10.1073/pnas.1104591108); pmid: [21896726](https://pubmed.ncbi.nlm.nih.gov/21896726/)
64. R. Fujiwara, N. Damodaran, J. E. Wilusz, K. Murakami, The capping enzyme facilitates promoter escape and assembly of a follow-on preinitiation complex for reinitiation. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 22573–22582 (2019). doi: [10.1073/pnas.1905449116](https://doi.org/10.1073/pnas.1905449116); pmid: [31591205](https://pubmed.ncbi.nlm.nih.gov/31591205/)
65. N. Yudkovsky, J. A. Ranish, S. Hahn, A transcription reinitiation intermediate that is stabilized by activator. *Nature* **408**, 225–229 (2000). doi: [10.1038/35041603](https://doi.org/10.1038/35041603); pmid: [11089979](https://pubmed.ncbi.nlm.nih.gov/11089979/)
66. F. W. Martinez-Rucobo, P. Cramer, Structural basis of transcription elongation. *Biochim. Biophys. Acta* **1829**, 9–19 (2013). doi: [10.1016/j.bbagen.2012.09.002](https://doi.org/10.1016/j.bbagen.2012.09.002); pmid: [22982352](https://pubmed.ncbi.nlm.nih.gov/22982352/)
67. M. Kireeva, M. Kashlev, Z. F. Burton, Translocation by multi-subunit RNA polymerases. *Biochim. Biophys. Acta* **1799**, 389–401 (2010). doi: [10.1016/j.bbagen.2010.01.007](https://doi.org/10.1016/j.bbagen.2010.01.007); pmid: [2097318](https://pubmed.ncbi.nlm.nih.gov/2097318/)
68. S. Alekseev et al., Transcription without XPB establishes a unified helicase-independent mechanism of promoter opening in eukaryotic gene expression. *Mol. Cell* **65**, 504–514.e4 (2017). doi: [10.1016/j.molcel.2017.01.012](https://doi.org/10.1016/j.molcel.2017.01.012); pmid: [28157507](https://pubmed.ncbi.nlm.nih.gov/28157507/)
69. C. Plaschka et al., Transcription initiation complex structures elucidate DNA opening. *Nature* **533**, 353–358 (2016). doi: [10.1038/nature17990](https://doi.org/10.1038/nature17990); pmid: [27193681](https://pubmed.ncbi.nlm.nih.gov/27193681/)
70. J. F. Santana, G. S. Collins, M. Parida, D. S. Luse, D. H. Price, Differential dependencies of human RNA polymerase II promoters on TBP, TAFI, TFIIB and XPB. *Nucleic Acids Res.* **50**, 9127–9148 (2022). doi: [10.1093/nar/gkac678](https://doi.org/10.1093/nar/gkac678); pmid: [3547745](https://pubmed.ncbi.nlm.nih.gov/3547745/)
71. L. M. Hsu, N. V. Vo, C. M. Kane, M. J. Chamberlin, In vitro studies of transcript initiation by *Escherichia coli* RNA polymerase. 1. RNA chain initiation, abortive initiation, and promoter escape at three bacteriophage promoters. *Biochemistry* **42**, 3777–3786 (2003). doi: [10.1021/bi026954e](https://doi.org/10.1021/bi026954e); pmid: [12667069](https://pubmed.ncbi.nlm.nih.gov/12667069/)
72. K. L. Henderson et al., Mechanism of transcription initiation and promoter escape by *E. coli* RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E3032–E3040 (2017). doi: [10.1073/pnas.1618675114](https://doi.org/10.1073/pnas.1618675114); pmid: [28348246](https://pubmed.ncbi.nlm.nih.gov/28348246/)
73. X. Liu, L. Farnung, C. Wigge, P. Cramer, Cryo-EM structure of a mammalian RNA polymerase II elongation complex inhibited by α -amanitin. *J. Biol. Chem.* **293**, 7189–7194 (2018). doi: [10.1074/jbc.RA118.002545](https://doi.org/10.1074/jbc.RA118.002545); pmid: [29550768](https://pubmed.ncbi.nlm.nih.gov/29550768/)
74. M. J. Bick, S. Malik, A. Mustaev, S. A. Darst, TFIIB is only ~9 Å away from the 5'-end of a trimeric RNA primer in a functional RNA polymerase II preinitiation complex. *PLOS ONE* **10**, e019007 (2015). doi: [10.1371/journal.pone.019007](https://doi.org/10.1371/journal.pone.019007); pmid: [25774659](https://pubmed.ncbi.nlm.nih.gov/25774659/)
75. H. T. Chen, S. Hahn, Mapping the location of TFIIB within the RNA polymerase II transcription preinitiation complex: A model for the structure of the PIC. *Cell* **119**, 169–180 (2004). doi: [10.1016/j.cell.2004.09.028](https://doi.org/10.1016/j.cell.2004.09.028); pmid: [15479635](https://pubmed.ncbi.nlm.nih.gov/15479635/)
76. B. S. Chen, M. Hampsey, Functional interaction between TFIIB and the Rpb2 subunit of RNA polymerase II: Implications for the mechanism of transcription initiation. *Mol. Cell. Biol.* **24**, 3983–3991 (2004). doi: [10.1128/MCB.24.9.3983-3991.2004](https://doi.org/10.1128/MCB.24.9.3983-3991.2004); pmid: [15082791](https://pubmed.ncbi.nlm.nih.gov/15082791/)
77. E. J. Cho, S. Buratowski, Evidence that transcription factor IIB is required for a post-assembly step in transcription initiation. *J. Biol. Chem.* **274**, 25807–25813 (1999). doi: [10.1074/jbc.274.36.25807](https://doi.org/10.1074/jbc.274.36.25807); pmid: [10464320](https://pubmed.ncbi.nlm.nih.gov/10464320/)
78. T. S. Pardee, C. S. Bangur, A. S. Ponticelli, The N-terminal region of yeast TFIIB contains two adjacent functional domains involved in stable RNA polymerase II binding and transcription start site selection. *J. Biol. Chem.* **273**, 17859–17864 (1998). doi: [10.1074/jbc.273.28.17859](https://doi.org/10.1074/jbc.273.28.17859); pmid: [9651390](https://pubmed.ncbi.nlm.nih.gov/9651390/)
79. N. A. Hawkes, S. G. Roberts, The role of human TFIIB in transcription start site selection in vitro and in vivo. *J. Biol. Chem.* **274**, 14337–14343 (1999). doi: [10.1074/jbc.274.20.14337](https://doi.org/10.1074/jbc.274.20.14337); pmid: [10318856](https://pubmed.ncbi.nlm.nih.gov/10318856/)
80. K. Tran, J. D. Gralla, Control of the timing of promoter escape and RNA catalysis by the transcription factor IIB fingertip. *J. Biol. Chem.* **283**, 15665–15671 (2008). doi: [10.1074/jbc.M801439200](https://doi.org/10.1074/jbc.M801439200); pmid: [18411280](https://pubmed.ncbi.nlm.nih.gov/18411280/)
81. H. Hou et al., Structural insights into RNA polymerase III-mediated transcription termination through trapping poly-deoxythymidine. *Nat. Commun.* **12**, 6135 (2021). doi: [10.1038/s41467-021-26402-9](https://doi.org/10.1038/s41467-021-26402-9); pmid: [34675218](https://pubmed.ncbi.nlm.nih.gov/34675218/)
82. A. Feklistov, S. A. Darst, Structural basis for promoter-10 element recognition by the bacterial RNA polymerase σ subunit. *Cell* **147**, 1257–1269 (2011). doi: [10.1016/j.cell.2011.10.041](https://doi.org/10.1016/j.cell.2011.10.041); pmid: [22136875](https://pubmed.ncbi.nlm.nih.gov/22136875/)
83. Y. Zhang et al., Structural basis of transcription initiation. *Science* **338**, 1076–1080 (2012). doi: [10.1126/science.1227786](https://doi.org/10.1126/science.1227786); pmid: [23086998](https://pubmed.ncbi.nlm.nih.gov/23086998/)
84. S. M. Vos, L. Farnung, H. Urlaub, P. Cramer, Structure of paused transcription complex Pol II-DSIF-NELF. *Nature* **560**, 601–606 (2018). doi: [10.1038/s41586-018-0442-2](https://doi.org/10.1038/s41586-018-0442-2); pmid: [30135580](https://pubmed.ncbi.nlm.nih.gov/30135580/)
85. Y. Chen et al., Allosteric transcription stimulation by RNA polymerase II super elongation complex. *Mol. Cell* **81**, 3386–3399.e10 (2021). doi: [10.1016/j.molcel.2021.06.019](https://doi.org/10.1016/j.molcel.2021.06.019); pmid: [34265249](https://pubmed.ncbi.nlm.nih.gov/34265249/)

86. T. Kokubo, S. Yamashita, M. Horikoshi, R. G. Roeder, Y. Nakatani, Interaction between the N-terminal domain of the 230-kDa subunit and the TATA box-binding subunit of TFIID negatively regulates TATA-box binding. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 3520–3524 (1994). doi: [10.1073/pnas.91.9.3520](https://doi.org/10.1073/pnas.91.9.3520); pmid: [8170939](https://pubmed.ncbi.nlm.nih.gov/8170939/)
87. A. B. Patel *et al.*, Structure of human TFIID and mechanism of TBP loading onto promoter DNA. *Science* **362**, eaau8872 (2018). doi: [10.1126/science.aau8872](https://doi.org/10.1126/science.aau8872); pmid: [30442764](https://pubmed.ncbi.nlm.nih.gov/30442764/)
88. J. Chen, H. Boyaci, E. A. Campbell, Diverse and unified mechanisms of transcription initiation in bacteria. *Nat. Rev. Microbiol.* **19**, 95–109 (2021). doi: [10.1038/s41579-020-00450-2](https://doi.org/10.1038/s41579-020-00450-2); pmid: [33122819](https://pubmed.ncbi.nlm.nih.gov/33122819/)
89. Y. Liu *et al.*, Structural and mechanistic basis of reiterative transcription initiation. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2115746119 (2022). doi: [10.1073/pnas.2115746119](https://doi.org/10.1073/pnas.2115746119); pmid: [35082149](https://pubmed.ncbi.nlm.nih.gov/35082149/)
90. A. Revyakin, C. Liu, R. H. Ebright, T. R. Strick, Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. *Science* **314**, 1139–1143 (2006). doi: [10.1126/science.1131398](https://doi.org/10.1126/science.1131398); pmid: [17110577](https://pubmed.ncbi.nlm.nih.gov/17110577/)
91. A. N. Kapanidis *et al.*, Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science* **314**, 1144–1147 (2006). doi: [10.1126/science.1131399](https://doi.org/10.1126/science.1131399); pmid: [17110578](https://pubmed.ncbi.nlm.nih.gov/17110578/)
92. J. T. Winkelman *et al.*, Crosslink mapping at amino acid-base resolution reveals the path of scrunched DNA in initial transcribing complexes. *Mol. Cell* **59**, 768–780 (2015). doi: [10.1016/j.molcel.2015.06.037](https://doi.org/10.1016/j.molcel.2015.06.037); pmid: [26257284](https://pubmed.ncbi.nlm.nih.gov/26257284/)
93. L. Li, V. Molodtsov, W. Lin, R. H. Ebright, Y. Zhang, RNA extension drives a stepwise displacement of an initiation-factor structural module in initial transcription. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 5801–5809 (2020). doi: [10.1073/pnas.1920747117](https://doi.org/10.1073/pnas.1920747117); pmid: [32127479](https://pubmed.ncbi.nlm.nih.gov/32127479/)
94. S. Samanta, C. T. Martin, Insights into the mechanism of initial transcription in *Escherichia coli* RNA polymerase. *J. Biol. Chem.* **288**, 31993–32003 (2013). doi: [10.1074/jbc.M113.497669](https://doi.org/10.1074/jbc.M113.497669); pmid: [24047893](https://pubmed.ncbi.nlm.nih.gov/24047893/)
95. D. Pupov, I. Kuzin, I. Bass, A. Kulbachinskiy, Distinct functions of the RNA polymerase σ subunit region 3.2 in RNA priming and promoter escape. *Nucleic Acids Res.* **42**, 4494–4504 (2014). doi: [10.1093/nar/gkt1384](https://doi.org/10.1093/nar/gkt1384); pmid: [24452800](https://pubmed.ncbi.nlm.nih.gov/24452800/)
96. B. Kastner *et al.*, GraFix: Sample preparation for single-particle electron cryomicroscopy. *Nat. Methods* **5**, 53–55 (2008). doi: [10.1038/nmeth1139](https://doi.org/10.1038/nmeth1139); pmid: [18157137](https://pubmed.ncbi.nlm.nih.gov/18157137/)
97. S. Q. Zheng *et al.*, MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017). doi: [10.1038/nmeth.4193](https://doi.org/10.1038/nmeth.4193); pmid: [28250466](https://pubmed.ncbi.nlm.nih.gov/28250466/)
98. K. Zhang, Getf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016). doi: [10.1016/j.jsb.2015.11.003](https://doi.org/10.1016/j.jsb.2015.11.003); pmid: [26592709](https://pubmed.ncbi.nlm.nih.gov/26592709/)
99. S. H. Scheres, RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012). doi: [10.1016/j.jsb.2012.09.006](https://doi.org/10.1016/j.jsb.2012.09.006); pmid: [23000701](https://pubmed.ncbi.nlm.nih.gov/23000701/)
100. D. Kimanius, B. O. Forsberg, S. H. Scheres, E. Lindahl, Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *eLife* **5**, e18722 (2016). doi: [10.7554/eLife.18722](https://doi.org/10.7554/eLife.18722); pmid: [27845625](https://pubmed.ncbi.nlm.nih.gov/27845625/)
101. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017). doi: [10.1038/nmeth.4169](https://doi.org/10.1038/nmeth.4169); pmid: [28165473](https://pubmed.ncbi.nlm.nih.gov/28165473/)
102. E. F. Petersen *et al.*, UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004). doi: [10.1002/jcc.20084](https://doi.org/10.1002/jcc.20084); pmid: [15264254](https://pubmed.ncbi.nlm.nih.gov/15264254/)
103. P. Emsley, K. Cowtan, Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004). doi: [10.1107/S0907444904019158](https://doi.org/10.1107/S0907444904019158); pmid: [15572765](https://pubmed.ncbi.nlm.nih.gov/15572765/)
104. V. B. Chen *et al.*, MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010). doi: [10.1107/S0907444909042073](https://doi.org/10.1107/S0907444909042073); pmid: [20057044](https://pubmed.ncbi.nlm.nih.gov/20057044/)
105. W. L. DeLano, The PyMOL Molecular Graphics System (2002); <http://www.pymol.org>.

ACKNOWLEDGMENTS

We thank Y. Zhang and H. Zhang for their help with the transcription initiation assay and the Center of Cryo-Electron Microscopy at Fudan University for their support with cryo-EM data collection.

Funding: This work was supported by grants from the National Key R&D Program of China (2021YFA1300100 and 2022YFA1303200), the National Natural Science Foundation of China (32271242, 32030055, 31830107, and 31821002), the Basic Research Special Zone Plan of Shanghai (22TQ020), the XPLORER Prize (Y.X.), the New Cornerstone Investigator Program (Y.X.), the Shanghai Oriental Scholar Distinguished Professor Project (X.C.), and the Young Elite Scientists Sponsorship Program by CAST (X.C.). **Author contributions:** X.C. and X.W. prepared the samples with the help of X.Q. and W.L.; We.L. and Q.W. performed the cryo-EM analyses and model building with the help of Y.R.; X.C. and Y.X. designed the experiments, analyzed the data, and wrote the manuscript; and Y.X. supervised the project. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The cryo-EM maps and coordinates have been deposited to the Electron Microscopy Data Bank (EMDB) and Protein Data Bank (PDB). The transcribing complexes include TC2 (EMD-37395, PDB ID 8WAK), TC3 (EMD-37396, PDB ID 8WAL), TC4 (EMD-37398, PDB ID 8WAN), TC5 (EMD-37399, PDB ID 8WAO), TC6 (EMD-37400, PDB ID 8WAP), TC7 (EMD-37401, PDB ID 8WAQ), TC8 (EMD-37402, PDB ID 8WAR), TC9 (EMD-37403, PDB ID 8WAS), TC10 (EMD-37404, PDB ID 8WAT), TC11 (EMD-37405, PDB ID 8WAU), TC12 (EMD-37406, PDB ID 8WAV), TC13 (EMD-37407, PDB ID 8WAW), TC14 (EMD-37408, PDB ID 8WAX), TC15 (EMD-37409, PDB ID 8WAY), TC16 (EMD-37410, PDB ID 8WAZ), TC17 (EMD-37411, PDB ID 8WBO), TC9 on AdML^{3g} (EMD-37412), and TC10 on AdML^{10g} (EMD-37413). **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. <https://www.science.org/about/science-licenses-journal-article-reuse>

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adl5120

Supplementary Text

Figs. S1 to S13

References ([106](#), [107](#))

MDAR Reproducibility Checklist

Movies S1 to S4

Data S1 to S4

Submitted 1 May 2023; accepted 11 November 2023

10.1126/science.adl5120