

Cryo-EM uncovers a sequential mechanism for RNA polymerase I pausing and stalling at abasic DNA lesions

Received: 18 August 2024

Accepted: 26 May 2025

Published online: 06 June 2025



Alicia Santos-Aledo ^{1,6}, Adrián Plaza-Pegueroles^{1,6}, Marta Sanz-Murillo ^{1,6}, Federico M. Ruiz ¹, Peini Hou ², Jun Xu ², David Gil-Carton ^{3,4,5}, Dong Wang ²  & Carlos Fernández-Tornero ¹ 

During synthesis of the ribosomal RNA precursor, RNA polymerase I (Pol I) monitors DNA integrity but its response to DNA damage remains poorly studied. Abasic sites are among the most prevalent DNA lesions in eukaryotic cells, and their detection is critical for cell survival. We report cryo-EM structures of Pol I in different stages of stalling at abasic sites, supported by in vitro transcription studies. Slow nucleotide addition opposite abasic sites occurs through base sandwiching between the RNA 3'-end and the Pol I bridge helix. Templating abasic sites can also cause Pol I cleft opening, which enables the A12 subunit to access the active center. Nucleotide addition opposite the lesion induces a translocation intermediate where DNA bases tilt to form hydrogen bonds with the new RNA base. These findings reveal unique mechanisms of Pol I stalling at abasic sites, differing from arrest by bulky lesions or abasic site handling by RNA polymerase II.

Genome integrity is critical for cell homeostasis, an equilibrium that is challenged by DNA damage. Among the >50,000 DNA lesions per day produced in a human cell, the most abundant are abasic (Ap) sites. Around 10,000 apurinic sites are formed daily in a cell under physiological conditions, in contrast to only 500 apyrimidinic sites¹. Ap sites arise either from spontaneous depurination or from defects in base excision repair (BER) of oxidated or alkylated bases². Cells mainly repair Ap sites through additional activation of the BER pathway, while nucleotide excision repair (NER) operates as a parallel mechanism. Repair of these DNA lesions protects both from mutations in the genetic information due to loss of the base code and from genome instability due to interference with replication or transcription.

Ap sites block replication by interfering with the progression of replicative DNA polymerases (DNAPs). To overcome the obstruction, replicative DNAPs can be replaced by translesion DNAPs, which have developed different strategies to handle Ap sites. Most DNAPs follow the so-called 'A-rule' or 'purine rule' because of their preference for

adenine incorporation, or guanine in a lesser degree, opposite Ap sites³, essentially due to the higher base stacking ability of purines^{4,5}. Alternative translesion DNA synthesis mechanisms include cytosine addition opposite a templating arginine that occupies the gap left by the missing base, as well as generation of a -1 frameshift due to Ap site extrusion from the active center^{6–8}.

Beside replication, Ap sites in the DNA template strand also induce stalling of DNA-dependent RNA polymerases (RNAP). Three types of RNAPs transcribe the nuclear genome in eukaryotic cells. RNA polymerase I (Pol I) produces the ribosomal RNA (rRNA) precursor that after maturation leads to the major RNA components of the ribosome. RNA polymerase II (Pol II) transcribes genes encoding for messenger RNA precursors and some non-coding RNAs such as small nuclear RNAs (snRNAs) and microRNAs. RNA polymerase III (Pol III) synthesizes short, untranslated RNAs such as transfer RNA, the 5S rRNA and U6 snRNA. In addition, RNAP IV and V are plant-specific enzymes involved in gene silencing through RNA-directed DNA methylation.

¹Centro de Investigaciones Biológicas Margarita Salas (CIB), CSIC, Madrid, Spain. ²Skaggs School of Pharmacy and Pharmaceutical Sciences, UCSD, La Jolla, CA, USA. ³Basque Resource for Electron Microscopy, Leioa, Spain. ⁴Instituto Biofisika (UPV/EHU, CSIC), University of the Basque Country, Leioa, Spain.

⁵Ikerbasque, Basque Foundation for Science, Bilbao, Spain. ⁶These authors contributed equally: Alicia Santos-Aledo, Adrián Plaza-Pegueroles, Marta Sanz-Murillo. ✉ e-mail: dow003@health.ucsd.edu; cftornero@cib.csic.es

Previous studies of Ap site management by Pol II show initial enzyme stalling at the lesion, followed by slow bypass obeying the A-rule⁹.

The nucleotide addition cycle has been best characterized in Pol II. A DNA nucleotide in the templating position ($i + 1$) faces an empty active center^{10,11}, where three aspartates coordinate a Mg^{2+} ion (Mg A). Arrival of a nucleoside triphosphate (NTP) opposite the templating nucleotide, in the so-called addition site (A-site), promotes closure of the trigger loop to sense base pairing between the incoming and templating nucleotides¹². At this stage, the NTP phosphates, which coordinate a second Mg^{2+} ion (Mg B), are positioned so that RNA nucleotide addition to the 3'-end of the nascent RNA is catalyzed, with release of pyrophosphate and Mg B^{12,13}. This provokes the backward movement of the trigger loop, which leads to enzyme translocation by one nucleotide due to conformational rearrangements of both the bridge helix and trigger loop¹⁴. As a result, the newly-formed base pair occupies position $i - 1$ in the DNA/RNA hybrid, while the next DNA nucleotide locates at $i + 1$ to start a new addition cycle. Apart from the A-site for nucleotide addition, Pol II harbors an entry site (E-site) where NTPs bind in an inverted position relative to the A-site¹⁵. Interestingly, the A- and E-sites in Pol II overlap partially and, thus, cannot be occupied simultaneously. While only the cognate NTP can occupy the A-site, any NTP can enter the E-site, suggesting an NTP-probing role for this site. Closure of the trigger loop during nucleotide addition prevents an NTP in the A-site from returning to the E-site due to steric hindrance.

RNAPs are known to pause or reduce the elongation rate in response to different signals, including misincorporations in the growing RNA that alter the DNA/RNA hybrid¹⁶. Sensing of a distorted hybrid induces reverse translocation of the enzyme, i.e., backtracking, which inactivates nucleotide addition because the RNA 3'-end is moved away from the active center¹⁷. Reactivation of the enzyme from a backtracked state requires RNA cleavage, which in Pol II is stimulated by binding of the C-terminal zinc ribbon (ZnR) from transcription factor IIS (TFIIS) to the enzyme funnel^{13,17}.

In growing eukaryotic cells, most transcriptional activity is devoted to rRNA synthesis and, consequently, Pol I function is a major regulator of cell homeostasis¹⁸. Among the 14 subunits constituting Pol I in *Saccharomyces cerevisiae* (13 in most species) the two largest, A190 and A135, form the DNA-binding cleft and harbor the active center^{19,20}. Assembly of these subunits is stabilized by the AC40/AC19 heterodimer, which is shared with Pol III, while peripheral subunits Rpb5, Rpb6, Rpb8, Rpb10, and Rpb12 are shared among the three nuclear RNAPs. Subunits A43 and A14 form the stalk that interacts with Rrn3^{21–23}, an activating factor that binds Pol I prior to promoter recruitment^{24,25}. Subunit A12 is composed by two ZnR that share homology with Pol II subunit Rpb9 and TFIIS²⁶. While the N-terminal ZnR (A12-Nt) stimulates elongation^{27,28}, the C-terminal ZnR (A12-Ct) is involved in RNA cleavage and proofreading^{29,30}. The A49/A34.5 heterodimer, containing domains that present homology with TFIIF and TFIIE in Pol II, stimulates backtracking and RNA cleavage^{30,31}. This heterodimer can reversibly dissociate from Pol I to generate Pol I*³², thus uncovering a binding site for A12-Ct that is distant from the active center³³. Accordingly, yeast Pol I* was proposed to coexist with Pol I in vivo and play a physiological role^{33,34}. Moreover, the mammalian counterpart of the A49/A34.5 heterodimer, termed PAF53/PAF49, was shown to dissociate from Pol I in starved cells³⁵. Furthermore, a yeast strain with an A12 deletion remains viable and Pol I purified from this strain lacks subunits A12, A49, and A34.5^{26,36}.

Previous work on Pol I stalling by DNA lesions is restricted to cyclobutane pyrimidine dimers (CPD), a common bulky DNA lesion induced by UV light³⁷. These studies showed that Pol I and Pol II handle bulky lesions differently, with Pol I stalling firmly before the CPD can access the active center while Pol II is able to slowly incorporate nucleotides opposite the lesion. To evaluate if this differential behavior is specific for bulky lesions or may also operate for other lesions, we studied the impact of Ap sites on Pol I activity.

We present biochemical and structural data showing that Pol I from *S. cerevisiae* handles Ap sites in a unique manner. Contrary to CPD lesions, Ap sites can template preferential addition of adenine opposite the lesion, following the A-rule. Nevertheless, in contrast to Pol II, Pol I stalls after nucleotide incorporation opposite the Ap site with minimal bypass. Moreover, our electron cryo-microscopy (cryo-EM) structures, at resolutions ranging from 2.8 to 3.5 Å, capture cleft opening and A12-Ct access in the enzyme funnel next to the active center, provide the structural basis for adenine addition opposite the Ap site, present atomic evidence for subsequent stalling, and uncover the presence of an E-site in Pol I. Altogether, our results shed light on the mechanism of Ap site recognition and handling by Pol I.

Results

Pol I stalls at abasic sites and cleaves slowly-inserted purines

To explore the effect of Ap sites on Pol I elongation, we performed in vitro transcription tests using a mismatched transcription bubble that contains an Ap site analog (Fig. 1a) at the templating $i + 1$ position ('a = Ap' in Fig. 1). The scaffold includes a radiolabeled 12mer RNA forming a 9 base pair DNA/RNA hybrid with the template strand, one nucleotide before the Ap site (Fig. 1b). Two different sequences for the base pair at $i - 1$ were assayed to study the influence of purine or pyrimidine at the RNA 3'-end. Control experiments with an undamaged nucleotide at $i + 1$ ('a = dT' in Fig. 1) show that Pol I and Pol II elongate the RNA molecule in a similar fashion, but Pol I is more prone to dissociate at intermediate pauses and exhibits residual RNA cleavage activity (Fig. 1c). In contrast, in the presence of an Ap site, Pol I slowly incorporates one nucleotide opposite the Ap site (Fig. 1c, 13mer band) but further RNA extension is disfavored. While nucleotide incorporation opposite the Ap site is also slow in the case of Pol II, this enzyme elongates the RNA molecule more efficiently than Pol I, in agreement with previous reports^{9,38}. Interestingly, the presence of purine at the RNA 3'-end is more favorable for nucleotide addition opposite the Ap site (Fig. 1c), suggesting that stacking with the incoming NTP plays a prominent role.

We then investigated the effect of Ap sites on nucleotide selectivity and found that Pol I follows the A-rule for nucleotide addition opposite the Ap site (Fig. 1d). Nevertheless, guanine can also be incorporated though at a slower rate, while pyrimidine incorporation is strongly disfavored. Our measured dNTP selectivity of Pol I opposite an Ap site is 1A:0.5G:0.1T:0.1U, further stressing the importance of base stacking for nucleotide addition opposite Ap sites. Besides, we detected RNA cleavage upon prolonged incubation with individual NTPs, with the exception of GTP when the nucleotide at the $i - 1$ position of the template strand in the bubble is cytosine (Fig. 1b and d, left panel). This can be explained by the fact that, after nucleotide addition, a dinucleotide is cleaved that leaves cytosine at $i + 1$, which can serve as template for effective GTP addition (but no other individual NTP). All these results were also confirmed using a simplified nucleotide scaffold lacking the upstream region of the bubble and the unpaired nucleotides in the non-template strand (Supplementary Fig. 1). Overall, we show that Pol I slowly inserts a purine opposite Ap sites, with preference for adenine, while the Pol I intrinsic RNA cleavage activity reduces lesion bypass.

Structure of Pol I paused at an abasic site

To uncover the structural basis of Pol I stalling at Ap sites, we prepared Pol I elongation complexes using a mismatched transcription bubble containing an Ap site at the DNA templating position, i.e. $i + 1$ (Fig. 2a). A non-hydrolysable ATP analog (AMPCPP) was included in the mixture to mimic the insertion step of nucleotide addition opposite the Ap site. We applied cryo-EM to obtain a consensus map at 2.8 Å resolution (Map-I), with regions attaining 2.6 Å resolution at the Pol I core including the active center (Supplementary Figs. 2–5). The structure corresponds to an Ap site-induced paused elongation complex (Ap-

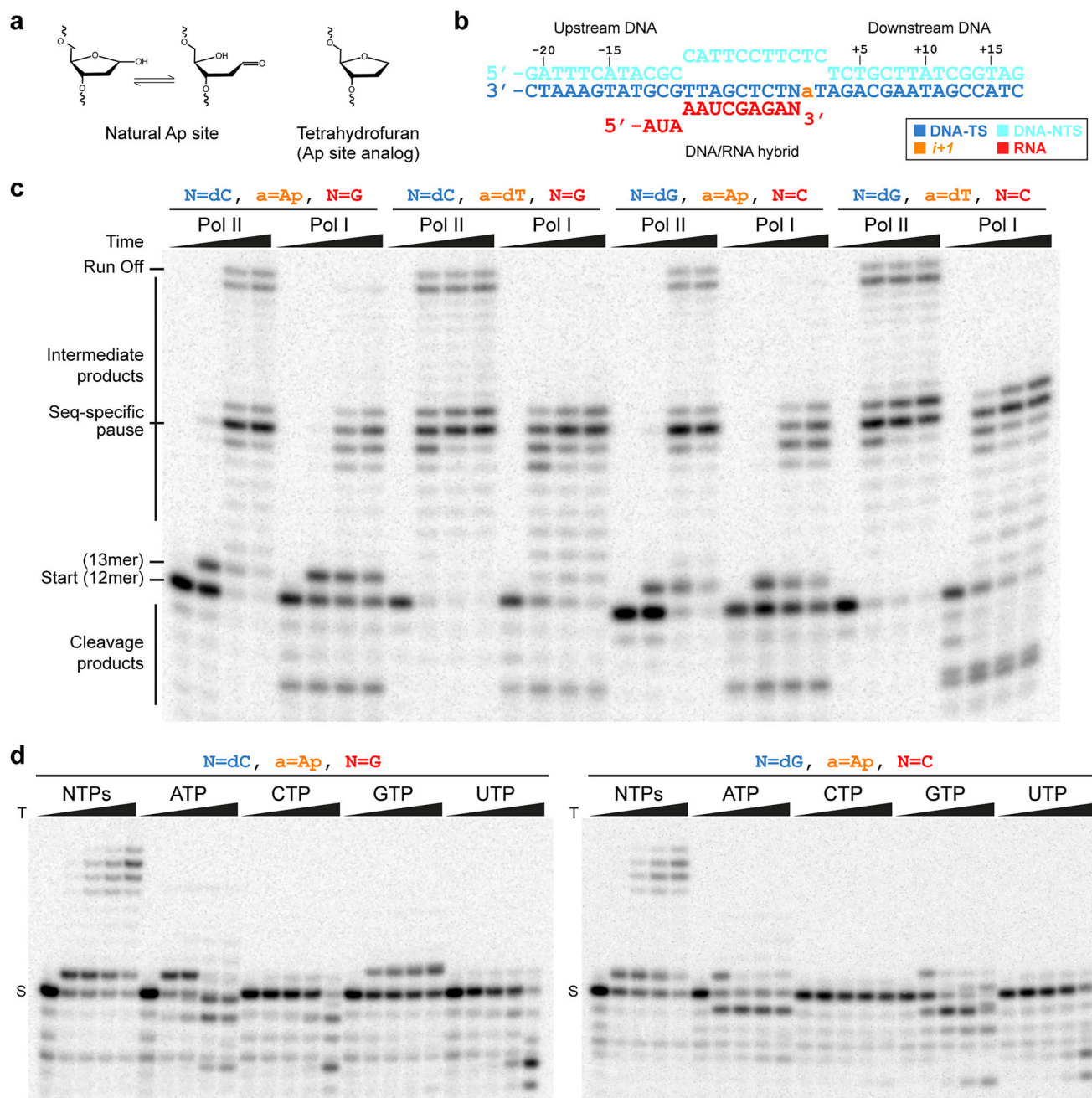


Fig. 1 | Ap site has distinct effects on Pol I and Pol II elongation. a The natural Ap site exists in equilibrium between the closed sugar hemiacetal and the open aldehyde. The synthetic analog tetrahydrofuran (hereafter termed Ap for simplicity) was used in this study as a mimic of the natural Ap site. **b** Full mismatched bubble scaffold used for in vitro transcription assays. **c** In vitro RNA extension assay on four

different mismatched bubbles. Each lane corresponds to incubation times of 0, 1, 15, and 60 min. **d** RNA extension assay with the same scaffold as in (a) but using individual NTPs. 'T' and 'S' state for 'Time' and 'Start (12mer)'. These experiments were repeated twice. Source data are provided as a Source Data file.

PEC), where Pol I adopts a post-translocated configuration with the Ap site at the *i*+1 position (Fig. 2b). To shed light on conformational changes of Pol I paused at Ap sites, we conducted 3D classification of the particles (Supplementary Fig. 2). One of the classes produced Map1a, which is most similar to the consensus both overall and around the active center, where resolution ranges between 3.6 and 3.8 Å (Supplementary Figs. 3 and 4). The resulting structure presents a closed DNA-binding cleft (Closed Ap-PEC) with an aperture of 28 Å (Fig. 2c). Like in the consensus structure, the Ap site occupies the *i*+1 position and exhibits certain flexibility, as judged from weaker density around the lesion (Supplementary Fig. 5). In spite of divergent sequences, comparison with the reported structure of post-translocated Pol I EC

with undamaged DNA³⁹ provides hints into the overall DNA/RNA hybrid configuration (Fig. 2d). Apart from minor differences at the Ap site, template DNA nucleotides at positions *i*-1 and *i*-2 are tilted and shifted towards the bridge helix, likely due to base absence at the Ap site. Superposition with the structure of Pol I stalled at a CPD lesion shows that the template DNA strand lies backwards in the CPD lesion structure, which affects the configuration of the DNA/RNA hybrid (Fig. 2e). Importantly, while the Ap site surpasses the bridge helix barrier and occupies the templating position, the CPD lesion lies behind the bridge helix. This explains that RNA extension by Pol I stops before the CPD lesion³⁷, while Ap sites induce enzyme stalling after slow nucleotide addition opposite the lesion (Fig. 1).

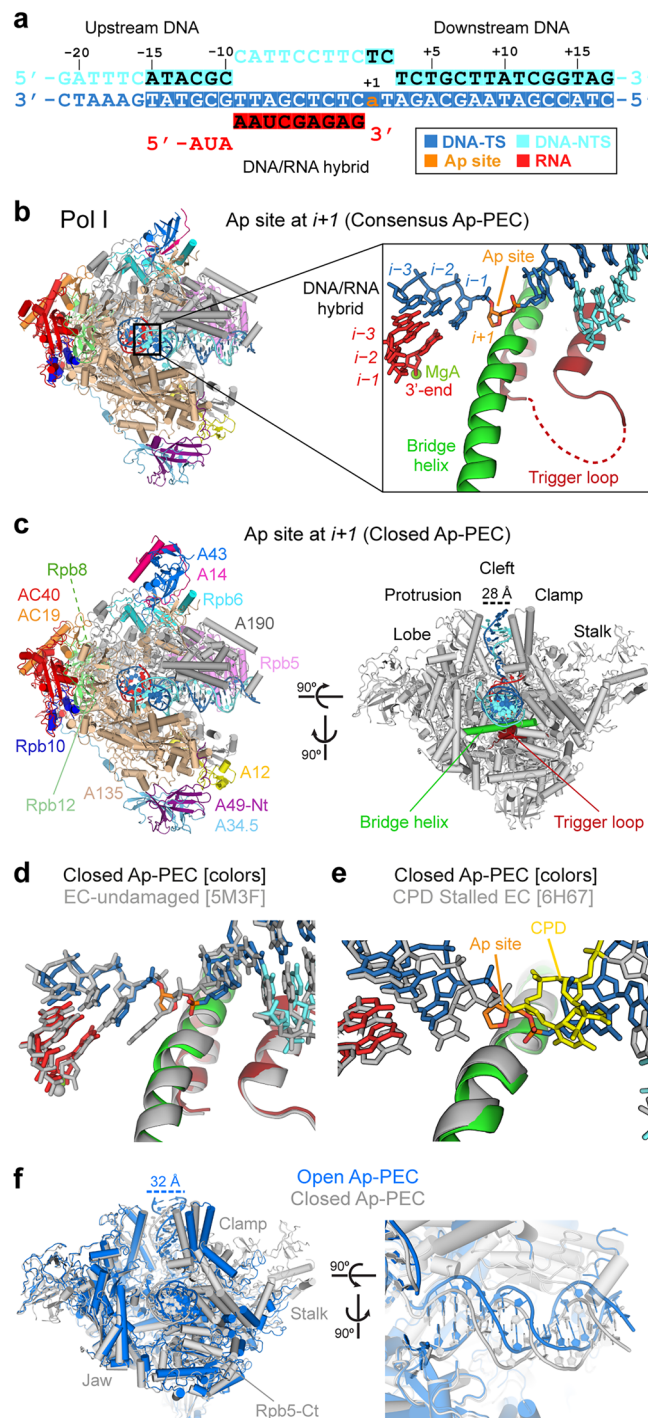


Fig. 2 | Structure of Pol I paused by an Ap site at *i* + 1. **a** Diagram of the nucleic acid scaffold with an Ap site at *i* + 1. Filled squares denote nucleotides visible in the cryo-EM map that were modelled. **b** Consensus Ap-PEC structure with subunits colored as in (c) and close-up view of the active center. **c** Two views of the Closed Ap-PEC structure indicating the different subunits and structural elements in the enzyme. **d** Superposition between Closed Ap-PEC (colors) and Pol I with undamaged DNA in a post-translocated state (PDB code 5M3F, grey). **e** Superposition of Closed Ap-PEC (colors) with Pol I stalled at a CPD lesion (PDB code 6H67, grey and CPD lesion in yellow). **f** Comparison between Closed (grey) and Open (blue) Ap-PEC structures in two different views.

A second class yielded Map-Ib, which is similar to the consensus around the active center but presents an open cleft that involves partial unfolding of the central region in the bridge helix (Supplementary Figs. 2–5). Conformational changes in the cleft occur through

swiveling along the downstream DNA axis (Supplementary Movie 1), consistent with previous observations^{20,40}. In the Open Ap-PEC structure, the cleft aperture attains 32 Å and results in displacement of downstream DNA together with the clamp (Fig. 2f). While Maps-Ia and Ib exhibit the same overall resolution of 3.5 Å, several domains are disordered in the Open Ap-PEC, including the external half of the clamp and part of the jaw in A190, the C-terminal region of Rpb5 and most of the stalk. Moreover, densities for both upstream and downstream DNA are poor in Map-Ib, indicating that nucleic acids in the Open Ap-PEC are more mobile (Supplementary Movie 1). Therefore, we interpret this open configuration as a Pol I paused state that may be prone to dissociation from DNA.

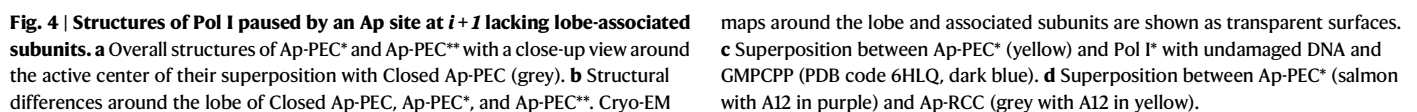
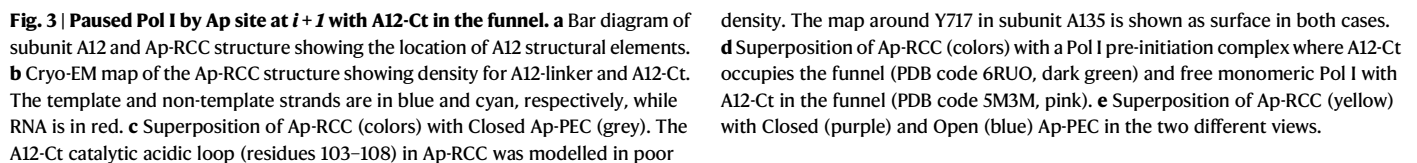
The A12 C-terminal Zn ribbon inserts in the funnel to catalyze RNA cleavage

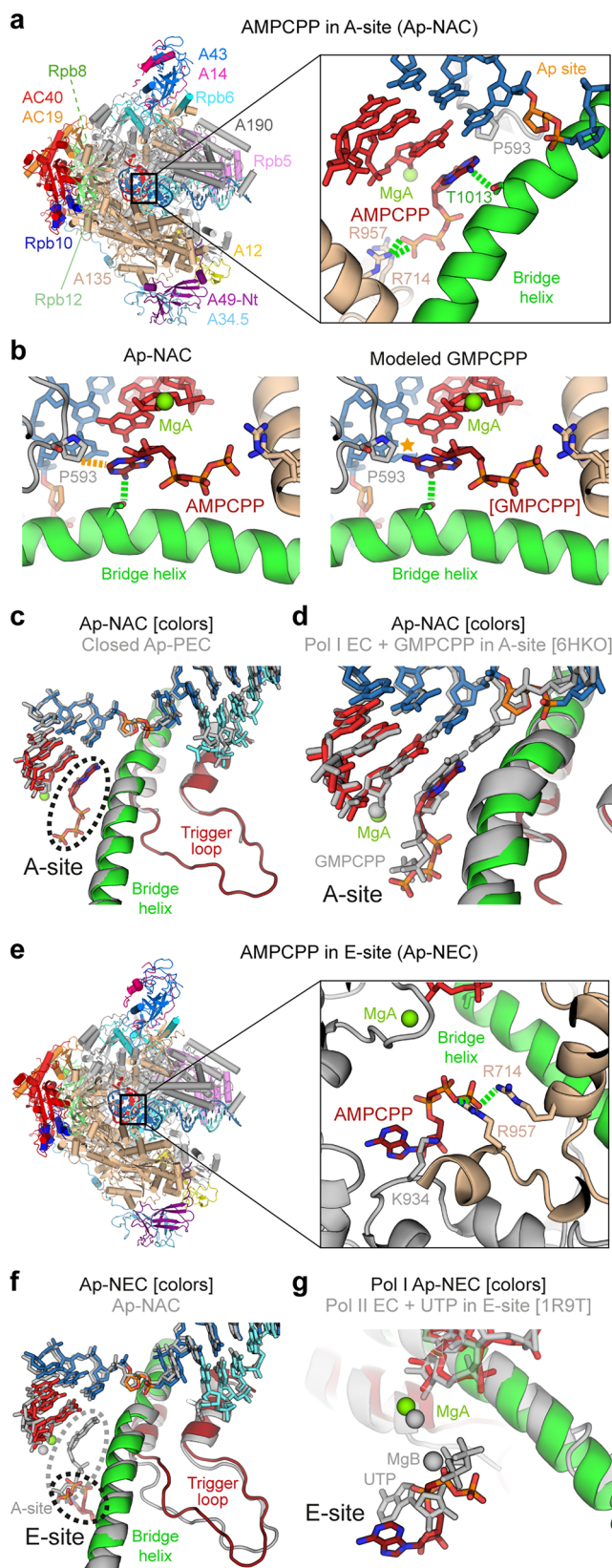
A third class showed additional density for A12-Ct bound to the enzyme funnel. Focused 3D classification using a mask covering the funnel yielded a locally improved map with an overall resolution of 3.3 Å, hereafter Map-II (Supplementary Figs. 2–5). The complete A12 subunit including both ZnR (residues 1–43 and 80–125) and the connecting linker could be modelled (Fig. 3a–c), in spite of poor density for the catalytic loop harboring acidic residues D105 and E106. Insertion of A12-Ct in the funnel has been observed in Pol I devoid of nucleic acids^{19,20,39} or in a pre-initiation complex lacking RNA⁴¹ (Fig. 3d). In contrast, our structure shows clear density for the full DNA/RNA hybrid and, thus, corresponds to an Ap site-induced RNA cleavage complex (Ap-RCC). Accommodation of A12-Ct within the funnel involves partial opening of the DNA-binding cleft by 30 Å (Fig. 3e), associated to unfolding of the central region in the bridge helix (Supplementary Fig. 5). A similar configuration has been reported for Pol II complexed to TFIIIS¹³ and for Pol III in complex with the integrase of the Ty1 retrotransposon⁴². Additionally, residue Y717 in A135 moves away from the A12 acidic loop to avoid clashes (Fig. 3c). The equivalent residue in Pol II, Y769 in subunit Rpb2, was proposed to limit the extent of backtracking¹³. The overall configuration of the hybrid is similar to that in the Closed Ap-PEC, although the Ap site and downstream DNA accompany the clamp as the cleft opens (Fig. 3e). This suggests that an Ap site at the templating *i* + 1 position induces an RNA cleavage-prone configuration in Pol I, consistent with the presence of residual RNA cleavage products in our RNA extension assays (Fig. 1). The Ap-RCC structure may also represent a post-cleavage configuration in which A12-Ct has not abandoned the funnel.

Structures of Pol I lacking lobe-associated subunits

Two additional classes lacked density for subunits binding next to the A135 lobe. Map-III, showing a resolution of 3.2 Å, corresponds to Pol I* paused at an Ap site as it lacks density for the A49/A34.5 heterodimer (Supplementary Figs. 2–5). The nucleic acid configuration in the resulting Ap-PEC* structure is virtually identical to that observed in the Closed Ap-PEC (Fig. 4a). In the absence of A49/A34.5, A12-Ct binds external domains 1 and 2 (ED1 and ED2), a region in the A135 lobe that is inaccessible in the presence of A49/A34.5 (Fig. 4b). This A12-Ct location is equivalent to that observed for the Pol I* EC with undamaged DNA³³, which is overall similar to Ap-PEC* except that downstream DNA lies backwards in the latter (Fig. 4c). Comparison with the Ap-RCC structure shows that A12-Ct undergoes a movement of ~60 Å and a rotation of ~90°, using the central mini-helix (residues 59–65) in the A12-linker as a hinge (Fig. 4d).

A final class arising from overall 3D classification produced Map-IV, which presents an overall resolution of 3.5 Å and lacks subunits A49, A34.5, and A12 (Supplementary Figs. 2–5). Unexpectedly, the resulting Ap-PEC** structure not only lacks density for these three subunits, but also for the entire A190 jaw domain (residues 1241–1540) and helices α6–α7 in the A135 lobe (residues 281–324) (Fig. 4a, b). This indicates that the N-terminal ZnR in A12 (A12-Nt) plays a stabilizing role of the





A190 jaw and part of the lobe, which are flexible when A12-Nt is absent. As a result, about two thirds of downstream DNA become exposed to the solvent, which may allow access of DNA-binding proteins into the Pol I cleft (Fig. 4b). The conformation of the enzyme is similar to that observed in the Closed Ap-PEC or Ap-PEC*, although the cleft is slightly more open in Ap-PEC*. This suggests that the jaw is able to move

Fig. 5 | Structures of Pol I paused by an Ap site at $i+1$ in the presence of AMPCPP. **a** Overall Ap-NAC structure and close-up view around AMPCPP in the A-site. H-bonds are shown as green dotted lines. **b** Comparison between Ap-NAC and the same structure where GMPCPP has been modelled. H-bonds and hydrophobic contacts are shown as green and orange dotted lines, while an orange star denotes a clash. **c** Superposition of Ap-NAC (colors) with Closed Ap-PEC (grey). **d** Superposition between Ap-NAC (colors) and Pol I with undamaged DNA in a pre-insertion state in the presence of GMPCPP in the A-site (PDB code 6HKO, grey). **e** Overall Ap-NEC structure and close-up view around AMPCPP in the E-site. The close-up view is rotated -120° along the bridge helix with respect to that in (a). **f** Superposition of Ap-NEC (colors) with Ap-NAC (grey). **g** Superposition of Ap-NEC (colors) with Pol II with ATP in the E-site (PDB code 1R9T, grey). View is as in (e).

independently using the cleft-jaw link as a hinge. Moreover, the nucleic acid scaffold also adopts an equivalent configuration, suggesting that this Pol I form may be able to transcribe *in vivo*, consistent with cell viability of a $\Delta A12$ strain where the Pol I enzyme lacks the three lobe-associated subunits^{26,34,36,43}.

Purine addition opposite the abasic site occurs via base stacking

Focused 3D classification with a spherical mask around the active center identified two classes with density for AMPCPP. The refined Map-Va and Map-Vb attained resolutions of 3.0 and 3.3 Å, respectively (Supplementary Figs. 2–4). The former exhibits clear density next to the RNA 3'-end corresponding to the adenine and phosphate moieties of AMPCPP, while the ribose appears flexible (Supplementary Fig. 5). In the resulting structure, the adenine moiety of AMPCPP is sandwiched between the base at the RNA 3'-end and residue T1013 in the bridge helix (Fig. 5a). P593 in subunit A190, a residue that is strictly conserved among nuclear RNAP (Supplementary Fig. 6), lies at a distance suitable for establishing Van der Waals interactions with adenine, likely explaining the observed preference versus guanine (Fig. 1d). Consistently, modeling of a guanine as incoming base would clash with this proline residue (Fig. 5b). R714 and R957 in subunit A135, both strictly conserved among nuclear RNAP (Supplementary Fig. 7), establish hydrogen bonds (H-bonds) with the γ -phosphate. The active center configuration is equivalent to the closed Ap-PEC, apart from minor changes around the Ap site (Fig. 5c). Comparison with Pol I EC with undamaged DNA containing GMPCPP³³ shows that AMPCPP in our structure occupies the canonical NTP addition site (A-site) and, thus, represents a nucleotide addition complex at Ap sites (Ap-NAC) (Fig. 5d). Overall, this configuration of the active center is likely derived from lack of templating base at the Ap site and provides structural evidence for preferential adenine addition (Fig. 1d).

Map-Vb exhibits density for AMPCPP in an alternative location beneath the active center (Supplementary Fig. 5). In the resulting structure, the adenine moiety of AMPCPP lies next to K934 in A190, while the γ -phosphate lies at H-bond distance from R714 and R957 in A135 (Fig. 5e–g). These three residues are strictly conserved among nuclear RNAP (Supplementary Fig. 6 and 7). We assign this location to the NTP entry site (E-site) previously described in a Pol II EC structure in complex with a non-cognate NTP¹⁵ (Fig. 5g). As in Pol II, the β - and γ -phosphates binding sites overlap in the A- and E-sites, implying that they cannot be occupied simultaneously (Fig. 5f). However, in spite of similarities between the two enzymes, the nucleotide in the Pol I E-site appears to fit more snugly, which may be due to absence of MgB. In our structure with AMPCPP in the E-site (Ap-NEC), the bridge helix is partly unfolded next to the Ap site, while the trigger loop is retracted and can be fully modelled (Fig. 5f). Additionally, DNA around the Ap site appears flexible due to weak density in this region of the map (Supplementary Fig. 5).

Structure of Pol I stalled after addition opposite an abasic site

Finally, we studied the cryo-EM structure of Pol I stalled after addition of adenine opposite the Ap site, using a nucleic acid scaffold where the

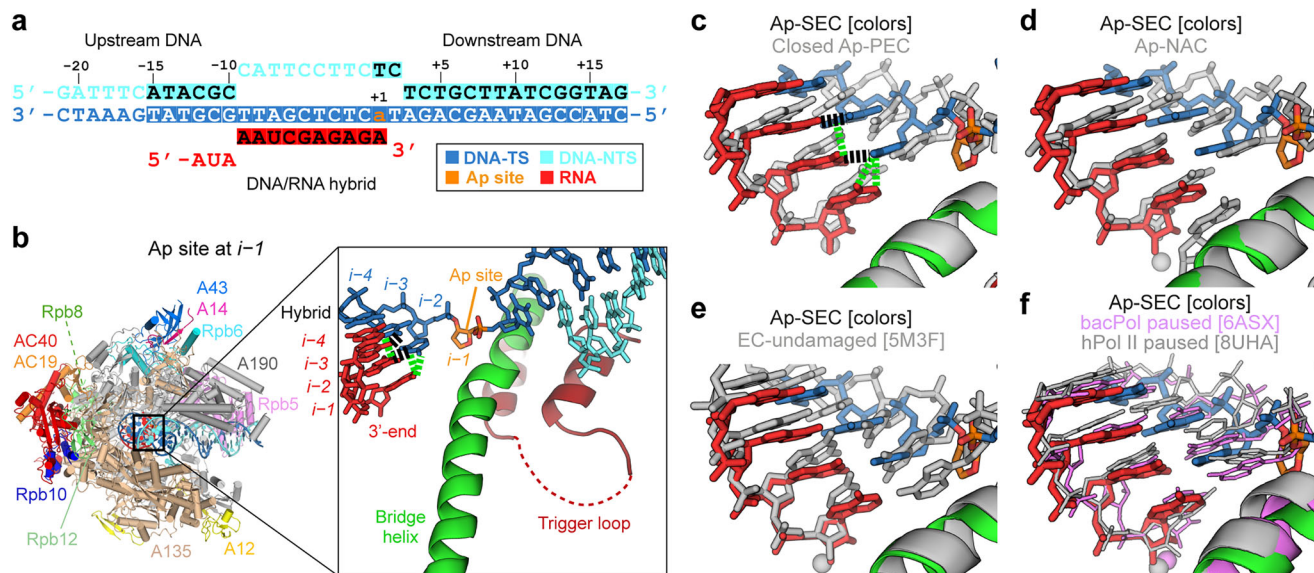


Fig. 6 | Structure of Pol I stalled by an Ap site at *i*-1. **a** Schematic diagram of the nucleic acid scaffold with an Ap site at *i*-1. Filled squares denote nucleotides visible in the cryo-EM map that were modelled. **b** Overall view of Pol I stalled at an Ap site (Ap-SEC) indicating the different subunits and close-up view of the active center. Canonical Watson-Crick H-bonds and other H-bonds between bases are shown as

black and green dotted lines, respectively. **c-f** Close-up views comparing the position of nucleotides in the DNA/RNA hybrid between Ap-SEC (colors) and Closed Ap-PEC (grey, (c)), Ap-NAC (grey, (d)), Pol I EC with undamaged DNA (PDB code 5M3F, grey, (e)), paused bacterial RNA polymerase (PDB code 6ASX, pink; (f)), or paused human Pol II (PDB code 8UHA, grey; (f)).

Ap:adenine unpaired duo is placed at the *i* - 1 position (Fig. 6a). Focused 3D classification with a mask covering the DNA-binding cleft identified a population where the Ap:adenine unpaired duo is clearly defined (Supplementary Figs. 2-5). To our surprise, the refined 2.8 Å Map-VI corresponds to Pol I*, i.e., lacking A49/A34.5 and with A12-Ct bound on the lobe. In the derived structure, nucleic acids adopt a configuration that is compatible with an intermediate of translocation (Fig. 6b), corresponding to an Ap site-induced stalled elongation complex (Ap-SEC). Notably, the DNA base at position *i* - 2 is tilted by ~30° so that it forms H-bonds with RNA bases at positions *i* - 2 and *i* - 1, the latter of which would otherwise be unpaired due to lack of base in the Ap site (Fig. 6b, c). Equivalently, the DNA base at position *i* - 3 is tilted so that it forms H-bonds with RNA bases at positions *i* - 3 and *i* - 2. The position of the Ap site is clearly defined, slightly advanced with respect to the Closed Ap-PEC and Ap-NAC (Fig. 6c, d) but far from the post-translocated state (Fig. 6e). Downstream DNA also presents an intermediate state of translocation, where the *i* + 1 nucleotide lies behind the bridge helix and cannot template nucleotide addition (Fig. 6b). RNA nucleotides are closer to the post-translocated state but lie 1.6 Å behind the canonical position, thus partly occluding the A-site. While the central region of the bridge helix is fully folded, it lies closer to the trigger loop relative to the post-translocated Closed Ap-PEC structure, leading to slight opening of the cleft. Overall, this provokes a less tight interaction of the template DNA with Pol I and may be related to lack of subunits A49 and A34.5. In conclusion, Ap-SEC presents an intermediate translocation state that is disfavored for subsequent RNA elongation. This configuration is different from the bacterial RNA polymerase swiveled state⁴⁴ or the mammalian Pol II tilted state^{45,46}. In particular, the central region of the bridge helix lies ~3 Å closer to the trigger loop, while DNA bases in *i* - 2 and *i* - 3 occupy the space of *i* - 1 and *i* - 2 bases in other paused states (Fig. 6f). This suggests that Pol I stalls at Ap sites through a unique mechanism.

Discussion

DNA lesions on the template strand challenge transcription elongation by slowing or blocking the advance of RNAP. In this work, we provide mechanistic insights into Pol I stalling at Ap sites, which are the most common DNA lesions in living cells, by combining biochemical analysis

with cryo-EM structures of several conformational states in the process. Our results allow us to propose a pathway for Ap site-induced pausing and stalling that consists of several steps (Fig. 7). Contrary to CPD lesions, which induce firm Pol I stalling before nucleotide addition opposite the lesion³⁷, Ap sites allow for untemplated purine addition opposite the lesion at slow rate, followed by Pol I stalling with minimal bypass after prolonged incubation. Importantly, as shown for CPD lesions, the intrinsic RNA cleavage activity residing in A12-Ct likely reduces Ap site bypass by Pol I.

Pol I swiveling at templating Ap sites

Our structures of Pol I with an Ap site at *i* + 1 show that nucleic acids in the cleft adopt a post-translocated configuration where the lesion occupies the templating position. Nevertheless, the absence of a templating base induces changes in the hybrid and the bridge helix, which likely make this post-translocated state suboptimal for addition. Consistently, two conformations of the cleft aperture are observed, open and closed (Movie S1), suggesting that the lack of a templating base is sensed by Pol I. In the open state, the downstream and upstream regions of DNA are loosely bound to the enzyme, which may lead to Pol I fall off from DNA at templating Ap sites. In agreement, the Pol I elongation complex presents low stability as compared to other RNAP⁴⁷.

Alternatively, cleft opening may be necessary to allow access of A12-Ct into the funnel. This is supported by our structure of Pol I with A12-Ct in the funnel, which exhibits partial cleft opening, as previously observed for an initiation open complex in the absence of RNA⁴¹. Our structure indicates that lack of a templating base may induce a configuration that is prone to RNA cleavage, although our biochemical assays reveal that RNA cleavage mainly occurs after nucleotide addition opposite the Ap site. Importantly, the location of A12-Ct in the funnel leaves a channel for NTP entrance into the active center (Fig. 3a). Moreover, flexibility of the A12-Ct acidic loop in our structure likely allows NTP binding to the E- or the A-site for nucleotide addition.

Purine addition occurs through base sandwiching

The structures of Pol I paused at Ap sites complexed to AMPCPP uncover the E- and A-sites for nucleotide incorporation opposite

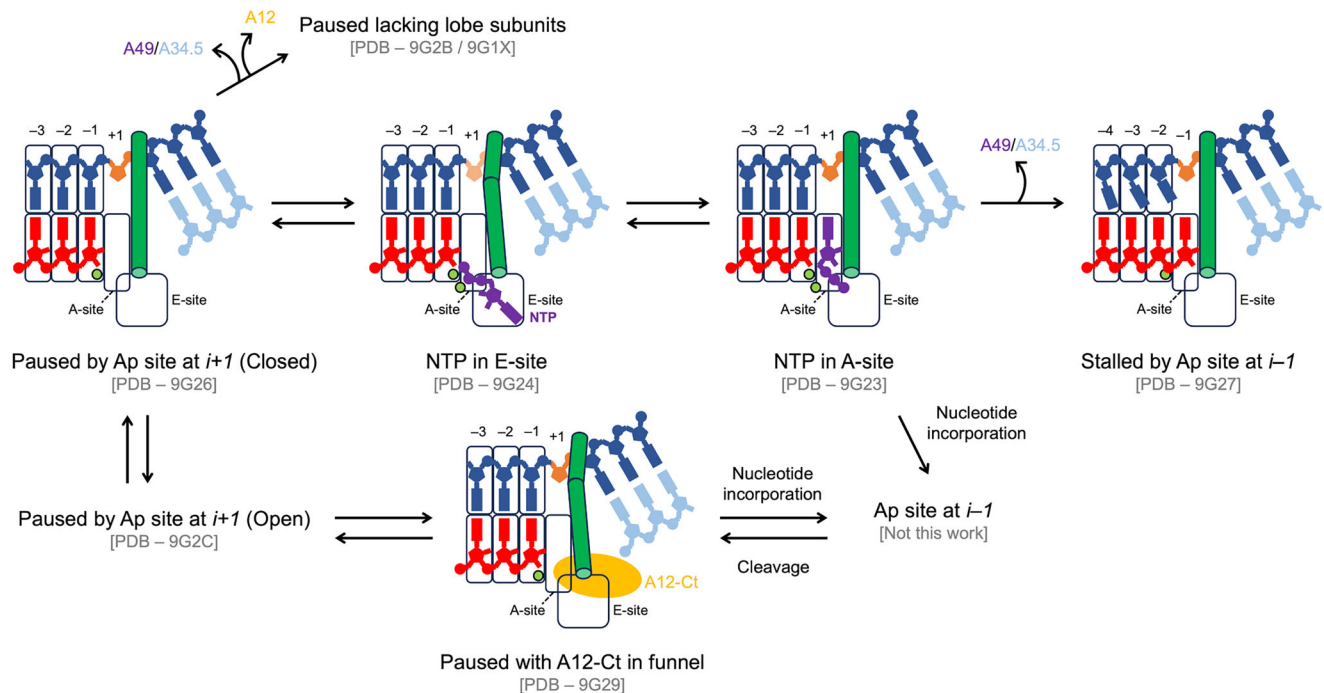


Fig. 7 | Mechanistic model of Pol I pausing and stalling at Ap sites. The DNA template and non-template strands are in blue and cyan, RNA is in red, the Ap site is in orange, the incoming nucleotide is in purple, A12 is in yellow, the bridge helix is in dark green and magnesium ions are in light green. Pol I is initially paused as the Ap site occupies the templating position ($i+1$), which leads to cleft swiveling. This may allow access of A12-Ct into the funnel for subsequent RNA cleavage or lead to Pol I dissociation from DNA in the open cleft state. Pol I stalled at Ap site allows NTP

entry into the E-site (entry), which enables access into the A-site (addition). In the A-site, purines are stabilized by sandwiching between the RNA 3'-end and the bridge helix, with preference for ATP. Phosphodiester bond formation likely leads to an altered hybrid configuration that induces RNA cleavage to minimize lesion bypass. Alternatively A49/A34.5 are lost, which hampers RNA cleavage but induces an intermediate of translocation that compromises nucleotide addition, thus stalling Pol I.

abasic lesions. The NTP in the E-site binds mainly through its terminal diphosphate moiety to two arginine residues in subunit A135 that are conserved among nuclear RNAP. Interestingly, the NTP base forms stabilizing interactions with a lysine in A190 that is conserved in all RNAP. A pyrimidine is likely to establish equivalent weak interactions with this Pol I region. A second Mg ion, as seen in Pol II^{9,15}, is likely to further stabilize the interaction on the phosphate moiety and favor rotation into the A-site. The existence of an E-site in Pol I may be related to transcriptional speed, as shown for Pol II⁴⁸. This may enhance NTP residence in the funnel, consistent with faster elongation rates of Pol I with respect to Pol II^{47,49}.

In the absence of a DNA templating base, NTP in the A-site is able to occupy the canonical position, thus enabling addition opposite the Ap site, as shown in our RNA extension assays. This is possible due to NTP base sandwiching by the bridge helix and the base at the 3'-end of the RNA, which explains the preference for purines as their stacking capacity is higher than for pyrimidines. The minor preference of adenine versus guanine addition is likely due a conserved proline next to the active center. Our biochemical results also show that the presence of purine at the RNA 3'-end is more favorable than the presence of pyrimidine, further underscoring the importance of base stacking for nucleotide addition opposite Ap sites.

Stalling after purine addition is incompatible with RNA extension

According to our results, NTP addition opposite Ap sites leads to Pol I translocation stalling, which affects both the template DNA strand and the RNA. On the RNA side, incomplete translocation leads to partial occlusion of the A-site. On the DNA side, the $i+1$ nucleotide lies behind the bridge helix, hampering its templating role. This is due to stabilization of the unpaired base at the RNA 3'-end by the DNA base in the hybrid that is adjacent to the Ap site. Importantly,

the intermediate translocation state observed in our structure differs from described translocation intermediates^{14,44–46}. In this unique paused state, which we term retracted pause, nucleotide addition is compromised, consistent with our RNA extension assays (Fig. 1). Moreover, these assays suggest that the retracted pause induces Pol I backtracking and RNA cleavage, leading to persistent stalling after nucleotide addition opposite Ap sites, with minimal bypass upon long incubation.

Role of lobe-associated subunits A12 and A49/34.5

Lobe-associated subunits A49, A34.5, and A12 provide distinctive functions to Pol I⁵⁰ in spite of being dispensable for yeast growth^{34,43,51}. We find that the A49/A34.5 heterodimer can dissociate from Pol I when an Ap site in the template DNA strand is positioned at the $i+1$ or $i-1$ site. This heterodimer was proposed to reversibly bind the Pol I core with a putative in vivo role in initiation and elongation³³. Our data support this hypothesis but further suggest a role in DNA lesion detection or, more generally, in transcriptional pausing. While we show that backtracking and RNA cleavage are induced at Ap sites, Pol I stalls after NTP incorporation opposite the lesion due to a non-productive translocation intermediate. Certain Pol I stalled configurations may induce dissociation of A49/A34.5, while this heterodimer would travel with Pol I during standard rRNA elongation. This is consistent with ChIP data showing that the heterodimer associates with the entire pre-rRNA gene^{52,53}.

A49/A34.5 dissociation has two major effects on the Pol I EC. On one hand, the A12-Ct binding site at Pol I external domains becomes exposed, leading to A12-Ct fixation on the lobe. This is consistent with reduced RNA cleavage activity relative to the complete enzyme³⁰. On the other hand, the binding site for the C-terminal domain of A49 next to upstream DNA becomes available. In Pol II, processivity factor Spt4/5 has been shown to bind in the same region⁵⁴ and Spt5 was shown to

bind Pol I⁵⁵. These two effects could contribute to increased lesion bypass. Interestingly, as opposed to Pol I stalled at CPD lesions³⁷, no density for A49-Ct is observed in particles containing A49/A34.5. This suggests that bypass of Ap sites may be favored compared to CPD lesions.

Unexpectedly, we found particles that lack all three lobe-associated subunits when an Ap site occupies the templating position. This structure shows a highly-exposed downstream DNA due to flexibility of the A190 jaw domain and a pair of helices in the lobe. This may enable access to factors that bind downstream DNA for repair. Nevertheless, the structure is compatible with nucleotide addition, consistent with studies showing that the presence of A12 destabilizes the Pol I elongation complex⁵⁶. Moreover, biochemical analysis of Pol I lacking the three lobe-associated subunits showed that this form of the enzyme is deficient in NTP selection and induces errors in the transcript³⁰. Therefore, loss of lobe-associated subunits may favor Ap site bypass.

Different stalling mechanisms in Pol I and Pol II

Our biochemical and structural results suggest that Pol I stalling at Ap sites involves two major steps (Fig. 7). In the first step, Pol I is able to slowly add a nucleotide opposite the lesion, with preference for adenine, which is equivalent to Pol II⁹. In the second step, Pol I stalls firmly due to impaired nucleotide addition and fast RNA cleavage, leading to minimal bypass. In contrast, Pol II presents a greater tendency to bypass Ap sites after nucleotide addition opposite the lesion. On one hand, this is due to the fact that Pol II is able to fully translocate after nucleotide addition opposite the lesion, while translocation is unfavored in Pol I (Fig. 6). On the other hand, Pol I-intrinsic RNA cleavage activity acts as failsafe mechanism to minimize bypass, whereas RNA cleavage in Pol II is negligible in the absence of TFIIIS. Because Pol I is less prone to bypass than Pol II, we speculate that putative mutations on rRNA derived from Ap site bypass could be more deleterious for the cell than those generated from bypass by Pol II, as previously suggested for CPD lesions³⁷.

While the main pathway to repair Ap sites is BER, transcription-coupled nucleotide excision repair has been reported to operate for Ap site removal from the template DNA strand in yeast⁵⁷. Moreover, NER was shown to repair Ap sites in genes that are highly transcribed by Pol II^{38,58,59}. Given rDNA is transcribed at high rates, it is plausible that a similar mechanism operates for Pol I stalled at Ap sites, as NER was shown to function in this transcription system^{60,61}.

Methods

Protein purification

Yeast Pol II was purified following described protocols⁶². Purification of yeast Pol I was performed following described procedures⁶³ with minor modifications. A *S. cerevisiae* strain harboring a tandem affinity-purification tag at the C-terminus of subunit A190 was grown in a fermenter containing 50 L of YPDA for 20 h at 30 °C to an OD₆₀₀ of 5.5, harvested by centrifugation and stored at -80 °C until use. Subsequent steps were performed at 4 °C. Five hundred grams of cells were resuspended in a buffer containing 250 mM Tris-HCl pH 7.4, 20% glycerol, 250 mM ammonium sulfate, 1 mM EDTA, 10 mM MgCl₂, 10 μM ZnCl₂, 10 mM β-mercaptoethanol, supplemented with protease-inhibitors (1 mM phenylmethanesulfonyl fluoride, 2 μg/ml leupeptine, 4 mM benzamidine, 1.4 μg/ml pepstatine A) and DNase (Roche), and lysed with glass beads using a BeadBeater (Biospec). After centrifugation, the supernatant was incubated with 4 ml of IgG Sepharose (GE Healthcare) for 6 h, washed with a buffer containing 50 mM Tris-HCl pH 7.4, 5% glycerol, 200 mM NaCl, 1 mM MgCl₂, 10 μM ZnCl₂ and 5 mM dithiothreitol (DTT) and incubated overnight with tobacco etch virus (TEV) protease. The TEV eluate was further purified using a Mono Q column (GE Healthcare) and a gradient from 0.2 to 1 M NaCl in buffer 20 mM Tris-HCl pH 7.4, 1 mM MgCl₂, 10 μM ZnCl₂ and 5 mM DTT. Pol

I-containing fractions were pooled and concentrated to 5 mg/ml, frozen in liquid nitrogen and stored at -80 °C for further use.

In vitro transcription assays with mismatched bubble scaffolds

Non-template DNA (5'-GATTCATACGCCATTCCTTCTCTGCTTATC GGTAG-3'), undamaged template DNA (5'-CTACCGATAAGCA-GATTCTCTCGATTGCGTATGAAATC-3' or 5'-CTACCGATAAGCAGATT GTCTCGATTGCGTATGAAATC-3') or template DNA with lesion (5'-CTACCGATAAGCAGATaCTCTCGATTGCGTATGAAATC-3' or 5'-CTACCGATAAGCAGATaGTCTCGATTGCGTATGAAATC-3') where 'a' represents the abasic site analog tetrahydrofuran, and 12mer RNA (5'-AUAAAUCGAGAG-3' or 5'-AUAAAUCGAGAC-3') oligonucleotides were purchased from IDT. Pol I and Pol II elongation complexes with full bubble scaffolds were assembled by stepwise method as previously described⁶². The molar ratio of RNA:TS:NTS is 1:1.5:2. First the 5'-³²P labeled RNA was annealed to TS in assembly buffer (20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5 mM DTT) by heating to 95 °C and slowly cooling down to room temperature. Then Pol I or Pol II was incubated with RNA-TS hybrid on ice for 5 min and followed by incubating at room temperature for 5 min, then 10 min at 35 °C. NTS was added into polymerase/RNA/TS mixture and incubated at 35 °C for 5 min, followed by slowly cooling down to room temperature for 40 min to form elongation complexes. The elongation complex was incubated at 30 °C for 10 min, then transcription reaction was initiated by adding 2 mM NTP mix or 2 mM individual NTP diluted with elongation buffer (150 mM NaCl, 20 mM Tris-HCl, pH 7.5, 5 mM DTT, 10 mM MgCl₂). Final concentrations of each component in the reaction system are 30 nM scaffold and 200 nM polymerase. All the reactions were conducted at 30 °C and quenched by adding quench-loading buffer (90% formamide, 50 mM EDTA, 0.05% xylene cyanol and 0.05% bromophenol blue) at each desired time points (1 min, 5 min, 15 min, 60 min or 1 min, 20 min, 60 min). Samples were boiled for 10 min at 95 °C and product RNAs were analyzed by 12% (wt/vol) denaturing PAGE (6 M urea). The gels were visualized by phosphorimaging and analyzed using Image Laboratory software (BioRad).

In vitro transcription assays with minimal scaffold

Non-template DNA (5'-TCTGCTTATCGGTA-3'), template DNA (5'-CTACCGATAAGCAGATaCTCTCGATG-3') where 'a' represents the abasic site analog tetrahydrofuran, and 8mer RNA (5'-AUCGAGAG-3') or 9mer RNA (5'-AUCGAGAGA-3') were purchased from IDT. The transcription elongation assay was performed based on previous reported methods with slight modifications⁹. Briefly, an aliquot of 5'-³²P labeled RNA was annealed with a 1.5-fold amount of template DNA strand (TS) and 2.0-fold amount of non-template DNA strand (NTS) to form the RNA/DNA mini-scaffold in a buffer composed of 20 mM Tris-HCl (pH 7.5), 150 mM KCl, 5 mM DTT. An aliquot of annealed scaffold was then incubated with a five excess amount of purified Pol I or Pol II on ice for 5 min, followed by incubation at room temperature (23 °C) for 15 min to ensure the formation of Pol I or Pol II elongation complex. The transcription was chased by adding an equal volume of solution containing 150 mM NaCl, 20 mM Tris-HCl (pH 7.5), 5 mM DTT, 10 mM MgCl₂, and 2 mM NTP. Final reaction concentrations after mixing were 30 nM scaffold, 150 nM Pol I or Pol II, 10 mM MgCl₂, and 1 mM NTP. Reactions were quenched at various times (0, 0.25, 1, 4, 15, 60 min or 0, 1, 4, 15, 60 min) by addition of one volume of 0.5 M EDTA (pH 8.0). The RNA transcript was analyzed by 12% (wt/vol) denaturing urea/PAGE. The gel was visualized by phosphorimaging and analyzed using ImageLab software (BioRad).

Preparation of Pol I EC stalled by an abasic site for structural studies

Non-template DNA (5'-GATTCATACGCCATTCCTTCTCTGCTTATC GGTAG-3'), template DNA (5'-CTAAAGTATGCGTTAGCTCTCaTAG ACGAATAGCCATC-3'), where 'a' represents the abasic site analog

tetrahydrofuran) and RNA-1 (5'-AUGAUCGAGAG-3') or RNA-2 (5'-AUGAUCGAGAGA-3') from IDT were resuspended in a buffer containing 10 mM Hepes pH 8.0, 150 mM NaCl. The DNA strands were mixed in equimolar amounts, heated to 95 °C and slow-cooled to 4 °C. After, either RNA-1 or RNA-2 was added in equimolar amounts, the mixture was heated to 45 °C and slow-cooled to 4 °C. The scaffold was then incubated in equimolar amounts with Pol I overnight at 4 °C in a buffer containing 10 mM Tris pH 7.4, 150 mM NaCl and 5 mM DTT. For the scaffold containing RNA-1, the buffer was supplemented with 5 mM MgCl₂ and 1 mM AMPCPP. For the scaffold containing RNA-2, 1 mM EDTA was added.

Cryo-EM sample preparation and data collection

For the sample containing RNA-1, we collected data from two different cryo-EM samples due to strong preferential orientation. For the first dataset, the final Pol I concentration was 0.07 mg/ml and 8 mM CHAPSO was added prior to grid preparation. Three microliters of sample were applied to freshly glow-discharged Quantifoil R2/1, 400 mesh grids coated with a second layer of homemade thin continuous carbon. After 1 min incubation, grids were blotted for 2 s and blot force of −5 on a Vitrobot Mark IV (FEI) at room temperature and 100% humidity, then plunged into liquid ethane. For the second dataset, the final Pol I concentration in the stalled EC sample was 0.24 mg/ml and no further chemicals were added. Four microliters of sample were applied to freshly glow-discharged Quantifoil R2/2, 300 mesh grids. After 30 s incubation, grids were blotted for 3 s and offset −3 mm on a Vitrobot Mark II (FEI) at room temperature and 100% humidity, then plunged into liquid ethane. In both cases, vitrified grids were stored in liquid nitrogen. Both datasets were collected on the Titan Krios I electron microscope (FEI) at the Diamond Light Source operated at 300 kV, using a K2 summit direct electron detector (Gatan) and the EPU automated single-particle acquisition software (FEI). In both cases, data were acquired with a final pixel size of 1.047 Å at the specimen level, using defocus values between −1 and −3.5 μm in 0.3 μm intervals, and the camera was operated in dose-fractionation counting mode. For grids coated with thin continuous carbon, 4425 movies were collected at tilt angle 0°, with a total dose of 42.5 e[−]/Å² accumulated over 8 s and fractionated across 40 frames. For holey grids lacking continuous carbon, 6424 movies were collected at tilt angles 0°, 20°, and 30° with a total dose of 40.1 e[−]/Å² accumulated over 8 s and fractionated across 40 frames.

For the sample containing RNA-2, the final Pol I concentration was 0.4 mg/ml. Four microliters of sample were applied to freshly glow-discharged UltrAufoil RL2/L3, 300 mesh grids. After 30 s incubation, grids were blotted for 2 s and blot force of −5 on a Vitrobot Mark IV (FEI) at 10 °C and 100% humidity, then plunged into liquid ethane and stored in liquid nitrogen. Data were collected on the Titan Krios (FEI) at the BREM Facility operated at 300 kV, using a K3 direct electron detector (Gatan) and the EPU automated single-particle acquisition software (FEI). Data were acquired with a final pixel size of 0.646 Å at the specimen level, using defocus values between −1 and −2.5 μm in 0.3 μm intervals, and the camera was operated in dose-fractionation counting mode. A total of 23488 movies were collected at tilt angle 30° with a total dose of 47.2 e[−]/Å² accumulated over 1.7 s and fractionated across 40 frames.

Cryo-EM data processing

For the sample containing RNA-1, a total of 10849 movies from both datasets were aligned with MotionCorr2⁶⁴ and global CTF parameters were estimated using CtfFind⁶⁵. 2137 of those were discarded after manual inspection of average micrographs and their corresponding power spectra. For the first dataset, around 1000 particles were manually-picked, reference-free 2D classified and the resulting averages used as templates for autopicking in Relion 3.0⁶⁶. For the

second dataset, particle locations were obtained with crYOLO⁶⁷, then used to estimate per-particle CTF defocus values with Gctf⁶⁸. Relion 3.0 was used for further processing. Particles from both datasets were separately extracted using a box of 288 pixels and subjected to several rounds of 2D classification. 199450 and 635120 particles from the first and second datasets, respectively, belonging to classes that produce averages showing secondary structure features were joined and used to generate an initial model. All particles were jointly refined using the initial model as a reference, then polished and their CTF parameters refined. A global 3D classification into 6 classes was performed. Two classes lacking two or three lobe-associated subunits, plus 2 classes with different cleft opening were refined. A fifth class showed density in the funnel and was used for focused 3D classification using a mask around the funnel. The resulting class showing strong density in the funnel was subsequently refined. To identify AMPCPP binding sites, we performed focused 3D classification with a mask around the active center using CryoSPARC⁶⁹. Two classes showing density accounting for AMPCPP at distinct locations in the vicinity of the active center were refined. In all cases, final post-processing was performed using automatic masking and B-factor sharpening.

For the sample containing RNA-2, a total of 23488 movies were imported into CryoSPARC, where subsequent image analysis was performed. After movie alignment and CTF estimation, 4234 movies were discarded. Particles were selected using the blob picking utility, then extracted using a box of 500 pixels and subjected to several rounds of 2D classification. 1282765 particles belonging to classes with averages showing secondary structure features were used to generate an ab initio model. These particles were subjected to focused 3D classification with a mask around the DNA-binding cleft. One of the resulting classes showing density for the RNA 3'-end was refined using non-uniform refinement.

Model building and refinement

The available structure of undamaged Pol I EC containing GMPCPP in the A-site (PDB code: 6HKO) was fitted into the 3D maps using UCSF Chimera⁷⁰ and used as starting point for model building with Coot⁷¹ and subsequent real-space refinement as implemented in Phenix⁷². Refinement statistics are summarized in Supplementary Table 1. Figures were prepared using PyMOL (Schrödinger Inc.) and ChimeraX⁷³. Sequence alignments are performed with Clustalw⁷⁴ and represented with ESPrpt3⁷⁵.

Reporting summary

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

Data availability

The cryo-EM structures generated in this study have been deposited in the Protein Data Bank under accession codes 9GIV, 9G26, 9G2C, 9G29, 9G2B, 9GIX, 9G23, 9G24, 9G27. Their corresponding electronic density maps have been deposited in the Electron Microscopy Data Bank under accession codes EMD-50955, EMD-50965, EMD-50972, EMD-50970, EMD-50971, EMD-50956, EMD-50962, EMD-50963, EMD-50966. These structures and maps are detailed in Supplementary Table 1. The additional atomic structures used in this study are available in the Protein Data Bank under accession codes 5M3F, 6H67, 6RUO, 5M3M, 6HLQ, 6HKO, 1R9T, 6ASX, 8UHA. Source data are provided with this paper.

References

- Lindahl, T. Instability and decay of the primary structure of DNA. *Nature* **362**, 709–715 (1993).
- Loeb, L. A. & Preston, B. D. Mutagenesis by apurinic/aprimidinic sites. *Annu. Rev. Genet.* **20**, 201–230 (1986).

3. Strauss, B. S. The 'A rule' of mutagen specificity: a consequence of DNA polymerase bypass of non-instructional lesions?. *BioEssays* **13**, 79–84 (1991).
4. Zahn, K. E., Belrhali, H., Wallace, S. S. & Doublé, S. Caught bending the A-rule: crystal structures of translesion DNA synthesis with a non-natural nucleotide. *Biochemistry* **46**, 10551–10561 (2007).
5. Obeid, S. et al. Replication through an abasic DNA lesion: structural basis for adenine selectivity. *EMBO J.* **29**, 1738–1747 (2010).
6. Obeid, S., Welte, W., Diederichs, K. & Marx, A. Amino acid templating mechanisms in selection of nucleotides opposite abasic sites by a family a DNA polymerase*. *J. Biol. Chem.* **287**, 14099–14108 (2012).
7. Beard, W. A., Shock, D. D., Batra, V. K., Pedersen, L. C. & Wilson, S. H. DNA polymerase β substrate specificity. *J. Biol. Chem.* **284**, 31680–31689 (2009).
8. Nair, D. T., Johnson, R. E., Prakash, L., Prakash, S. & Aggarwal, A. K. DNA synthesis across an abasic lesion by human DNA polymerase ϵ . *Structure* **17**, 530–537 (2009).
9. Wang, W., Walmacq, C., Chong, J., Kashlev, M. & Wang, D. Structural basis of transcriptional stalling and bypass of abasic DNA lesion by RNA polymerase II. *Proc. Natl. Acad. Sci. USA* **115**, E2538–E2545 (2018).
10. Gnatt, A. L., Cramer, P., Fu, J., Bushnell, D. A. & Kornberg, R. D. Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science* **292**, 1876–1882 (2001).
11. Kettenberger, H., Armache, K.-J. & Cramer, P. Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIIS. *Mol. Cell* **16**, 955–965 (2004).
12. Wang, D., Bushnell, D. A., Westover, K. D., Kaplan, C. D. & Kornberg, R. D. Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell* **127**, 941–954 (2006).
13. Cheung, A. C. M. & Cramer, P. Structural basis of RNA polymerase II backtracking, arrest and reactivation. *Nature* **471**, 249–253 (2011).
14. Brueckner, F. & Cramer, P. Structural basis of transcription inhibition by α -amanitin and implications for RNA polymerase II translocation. *Nat. Struct. Mol. Biol.* **15**, 811–818 (2008).
15. Westover, K. D., Bushnell, D. A. & Kornberg, R. D. Structural basis of transcription: nucleotide selection by rotation in the RNA polymerase II active center. *Cell* **119**, 481–489 (2004).
16. Landick, R. The regulatory roles and mechanism of transcriptional pausing. *Biochem. Soc. Trans.* **34**, 1062–1066 (2006).
17. Wang, D. et al. Structural basis of transcription: backtracked RNA polymerase II at 3.4 Å resolution. *Science* **324**, 1203–1206 (2009).
18. Warner, J. R. The economics of ribosome biosynthesis in yeast. *Trends Biochem. Sci.* **24**, 437–440 (1999).
19. Engel, C., Sainsbury, S., Cheung, A. C., Kostrewa, D. & Cramer, P. RNA polymerase I structure and transcription regulation. *Nature* **502**, 650–655 (2013).
20. Fernández-Tornero, C. et al. Crystal structure of the 14-subunit RNA polymerase I. *Nature* **502**, 644–649 (2013).
21. Engel, C., Plitzko, J. & Cramer, P. RNA polymerase I–Rrn3 complex at 4.8 Å resolution. *Nat. Commun.* **7**, 12129 (2016).
22. Pils, M. et al. Structure of the initiation-competent RNA polymerase I and its implication for transcription. *Nat. Commun.* **7**, 12126 (2016).
23. Correia, E. et al. The dynamic assembly of distinct RNA polymerase I complexes modulates rDNA transcription. *eLife* **6**, e20832 (2017).
24. Yamamoto, R. T., Nogi, Y., Dodd, J. A. & Nomura, M. RRN3 gene of *Saccharomyces cerevisiae* encodes an essential RNA polymerase I transcription factor which interacts with the polymerase independently of DNA template. *EMBO J.* **15**, 3964–3973 (1996).
25. Keener, J., Josaitis, C. A., Dodd, J. A. & Nomura, M. Reconstitution of yeast RNA polymerase I transcription in vitro from purified components. *J. Biol. Chem.* **273**, 33795–33802 (1998).
26. Van Mullem, V., Landrieux, E., Vandenhaute, J. & Thuriaux, P. Rpa12p, a conserved RNA polymerase I subunit with two functional domains. *Mol. Microbiol.* **43**, 1105–1113 (2002).
27. Appling, F. D., Schneider, D. A. & Lucius, A. L. Multisubunit RNA polymerase cleavage factors modulate the kinetics and energetics of nucleotide incorporation: an RNA polymerase I case study. *Biochemistry* **56**, 5654–5662 (2017).
28. Scull, C. E., Lucius, A. L. & Schneider, D. A. The N-terminal domain of the A12.2 subunit stimulates RNA polymerase I transcription elongation. *Biophys. J.* **120**, 1883–1893 (2021).
29. Kuhn, C.-D. et al. Functional architecture of RNA polymerase I. *Cell* **131**, 1260–1272 (2007).
30. Schwank, K. et al. RNA polymerase I (Pol I) lobe-binding subunit Rpa12.2 promotes RNA cleavage and proofreading. *J. Biol. Chem.* **298**, 101862 (2022).
31. Geiger, S. R. et al. RNA polymerase I contains a TFIIF-related DNA-binding subcomplex. *Mol. Cell* **39**, 583–594 (2010).
32. Huet, J., Buhler, J. M., Sentenac, A. & Fromageot, P. Dissociation of two polypeptide chains from yeast RNA polymerase A. *Proc. Natl. Acad. Sci. USA* **72**, 3034–3038 (1975).
33. Tafur, L. et al. The cryo-EM structure of a 12-subunit variant of RNA polymerase I reveals dissociation of the A49–A34.5 heterodimer and rearrangement of subunit A12.2. *eLife* **8**, e43204 (2019).
34. Gadai, O. et al. A34.5, a nonessential component of yeast RNA polymerase I, cooperates with subunit A14 and DNA topoisomerase I to produce a functional rRNA synthesis machine. *Mol. Cell. Biol.* **17**, 1787–1795 (1997).
35. Penrod, Y., Rothblum, K., Cavanaugh, A. & Rothblum, L. I. Regulation of the association of the PAF53/PAF49 heterodimer with RNA polymerase I. *Gene* **556**, 61–67 (2015).
36. Merkl, P. E. et al. RNA polymerase I (Pol I) passage through nucleosomes depends on Pol I subunits binding its lobe structure. *J. Biol. Chem.* **295**, 4782–4795 (2020).
37. Sanz-Murillo, M. et al. Structural basis of RNA polymerase I stalling at UV light-induced DNA damage. *Proc. Natl. Acad. Sci. USA* **115**, 8972–8977 (2018).
38. Tornaletti, S., Maeda, L. S. & Hanawalt, P. C. Transcription arrest at an abasic site in the transcribed strand of template DNA. *Chem. Res. Toxicol.* **19**, 1215–1220 (2006).
39. Neyer, S. et al. Structure of RNA polymerase I transcribing ribosomal DNA genes. *Nature* **540**, 607–610 (2016).
40. Fernández-Tornero, C. RNA polymerase I activation and hibernation: unique mechanisms for unique genes. *Transcription* **9**, 248–254 (2018).
41. Sadian, Y. et al. Molecular insight into RNA polymerase I promoter recognition and promoter melting. *Nat. Commun.* **10**, 5543 (2019).
42. Nguyen, P. Q. et al. Structural basis of Ty1 integrase tethering to RNA polymerase III for targeted retrotransposon integration. *Nat. Commun.* **14**, 1729 (2023).
43. Mccusker, J. H., Yamagishi, M., Kolb, J. M. & Nomura, M. Suppressor analysis of temperature-sensitive RNA polymerase I mutations in *Saccharomyces cerevisiae*: suppression of mutations in a zinc-binding motif by transposed mutant genes. *Mol. Cell. Biol.* **11**, 746–753 (1991).
44. Kang, J. Y. et al. RNA polymerase accommodates a pause RNA hairpin by global conformational rearrangements that prolong pausing. *Mol. Cell* **69**, 802–815.e5 (2018).
45. Su, B. G. & Vos, S. M. Distinct negative elongation factor conformations regulate RNA polymerase II promoter-proximal pausing. *Mol. Cell* **84**, 1243–1256.e5 (2024).
46. Vos, S. M., Farnung, L., Urlaub, H. & Cramer, P. Structure of paused transcription complex Pol II–DSIF–NELF. *Nature* **560**, 601–606 (2018).
47. Jacobs, R. Q., Carter, Z. I., Lucius, A. L. & Schneider, D. A. Uncovering the mechanisms of transcription elongation by eukaryotic RNA polymerases I, II, and III. *iScience* **25**, 105306 (2022).

48. Batada, N. N., Westover, K. D., Bushnell, D. A., Levitt, M. & Kornberg, R. D. Diffusion of nucleoside triphosphates and role of the entry site to the RNA polymerase II active center. *Proc. Natl. Acad. Sci.* **101**, 17361–17364 (2004).
49. Lisica, A. et al. Mechanisms of backtrack recovery by RNA polymerases I and II. *Proc. Natl. Acad. Sci. USA* **113**, 2946–2951 (2016).
50. Schwank, K. et al. Features of yeast RNA polymerase I with special consideration of the lobe binding subunits. *Biol. Chem.* **404**, 979–1002 (2023).
51. Liljelund, P., Mariotte, S., Buhler, J. M. & Sentenac, A. Characterization and mutagenesis of the gene encoding the A49 subunit of RNA polymerase A in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci.* **89**, 9302–9305 (1992).
52. Beckouet, F. et al. Two RNA polymerase I subunits control the binding and release of Rrn3 during transcription. *Mol. Cell. Biol.* **28**, 1596–1605 (2008).
53. Rossi, M. J. et al. A high-resolution protein architecture of the budding yeast genome. *Nature* **592**, 309–314 (2021).
54. Farnung, L., Ochmann, M., Engholm, M. & Cramer, P. Structural basis of nucleosome transcription mediated by Chd1 and FACT. *Nat. Struct. Mol. Biol.* **28**, 382–387 (2021).
55. Viktorovskaya, O. V., Appling, F. D. & Schneider, D. A. Yeast transcription elongation factor Spt5 associates with RNA polymerase I and RNA polymerase II directly. *J. Biol. Chem.* **286**, 18825–18833 (2011).
56. Appling, F. D., Scull, C. E., Lucius, A. L. & Schneider, D. A. The A12.2 subunit is an intrinsic destabilizer of the RNA polymerase I elongation complex. *Biophys. J.* **114**, 2507–2515 (2018).
57. Kim, N. & Jinks-Robertson, S. Abasic sites in the transcribed strand of yeast DNA are removed by transcription-coupled nucleotide excision repair. *Mol. Cell. Biol.* **30**, 3206–3215 (2010).
58. Swanson, R. L., Morey, N. J., Doetsch, P. W. & Jinks-Robertson, S. Overlapping specificities of base excision repair, nucleotide excision repair, recombination, and translesion synthesis pathways for DNA base damage in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **19**, 2929–2935 (1999).
59. Torres-Ramos, C. A., Johnson, R. E., Prakash, L. & Prakash, S. Evidence for the involvement of nucleotide excision repair in the removal of abasic sites in yeast. *Mol. Cell. Biol.* **20**, 3522–3528 (2000).
60. Hanawalt, P. C. & Spivak, G. Transcription-coupled DNA repair: two decades of progress and surprises. *Nat. Rev. Mol. Cell Biol.* **9**, 958–970 (2008).
61. Daniel, L. et al. Mechanistic insights in transcription-coupled nucleotide excision repair of ribosomal DNA. *Proc. Natl. Acad. Sci. USA* **115**, E6770–E6779 (2018).
62. Xu, J. et al. Structural basis for the initiation of eukaryotic transcription-coupled DNA repair. *Nature* **551**, 653–657 (2017).
63. Moreno-Morcillo, M. et al. Solving the RNA polymerase I structural puzzle. *Acta Crystallogr. D. Biol. Crystallogr.* **70**, 2570–2582 (2014).
64. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
65. Rohou, A. & Grigorieff, N. CTFFIND4: fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
66. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).
67. Wagner, T. et al. SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. *Commun. Biol.* **2**, 1–13 (2019).
68. Zhang, K. Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
69. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
70. Pettersen, E. F. et al. UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
71. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 486–501 (2010).
72. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. Sect. Struct. Biol.* **75**, 861–877 (2019).
73. Meng, E. C. et al. UCSF ChimeraX: tools for structure building and analysis. *Protein Sci.* **32**, e4792 (2023).
74. Sievers, F. & Higgins, D. G. Clustal Omega for making accurate alignments of many protein sequences. *Protein Sci. Publ. Protein Soc.* **27**, 135–145 (2018).
75. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320–W324 (2014).

Acknowledgements

We thank J. Varela and C. Quevedo at the Protein Chemistry Facility (CIB-CSIC) and R. Núñez at the Electron Microscopy Facility (CIB-CSIC) for technical assistance. We acknowledge the Diamond Light Source for access and support of the cryo-EM facilities at the UK National Electron Bio-Imaging Centre (eBIC). We also thank the Basque Resource for Electron Microscopy (BREM) located at Instituto Biofisika (UPV/EHU, CSIC), supported by the Department of Education and the Innovation Fund of the Basque Government, Fundación Biofísica Bizkaia and MCIN with funding from European Union NextGenerationEU (PRTR-C17.I1). This work was supported by grant PID2020-116722GB-I00 funded by MICIU/AEI/10.13039/501100011033 (to C.F.-T.); grant PRE2021-098362 (to A.S.-A.) and grant PRE2018-087012 (to A.P.-P.) funded by MICIU/AEI/10.13039/501100011033, 'ESF Investing in your future' and ESF+; intramural funding from CSIC (2020AEP152 and PIE-202120E047-Conexiones-Life to C.F.-T.); and grant R01GM102362 funded by NIGMS (to D.W.). A.S.-A. acknowledges support from the Madrid City Council for a living stipend at Residencia de Estudiantes.

Author contributions

A.S.-A., A.P.-P. and M.S.-M. prepared Pol I complexes with DNA, collected cryo-EM data with D.G.-C. and performed cryo-EM analysis and structure determination with F.M.R. M.S.-M., P.H. and J.X. performed in vitro transcription experiments. D.W. and C.F.-T. conceived and supervised the work. C.F.-T. wrote the paper with input from all authors.

Competing interests

The authors declare no competing interest.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-60536-4>.

Correspondence and requests for materials should be addressed to Dong Wang or Carlos Fernández-Tornero.

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

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

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

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

© The Author(s) 2025