



# Structural basis of RNA polymerase I pre-initiation complex formation and promoter melting

Michael Pilsl <sup>1</sup> & Christoph Engel <sup>1</sup>✉

Transcription of the ribosomal RNA precursor by RNA polymerase (Pol) I is a prerequisite for the biosynthesis of ribosomes in eukaryotes. Compared to Pols II and III, the mechanisms underlying promoter recognition, initiation complex formation and DNA melting by Pol I substantially diverge. Here, we report the high-resolution cryo-EM reconstruction of a Pol I early initiation intermediate assembled on a double-stranded promoter scaffold that prevents the establishment of downstream DNA contacts. Our analyses demonstrate how efficient promoter-backbone interaction is achieved by combined re-arrangements of flexible regions in the 'core factor' subunits Rrn7 and Rrn11. Furthermore, structure-function analysis illustrates how destabilization of the melted DNA region correlates with contraction of the polymerase cleft upon transcription activation, thereby combining promoter recruitment with DNA-melting. This suggests that molecular mechanisms and structural features of Pol I initiation have co-evolved to support the efficient melting, initial transcription and promoter clearance required for high-level rRNA synthesis.

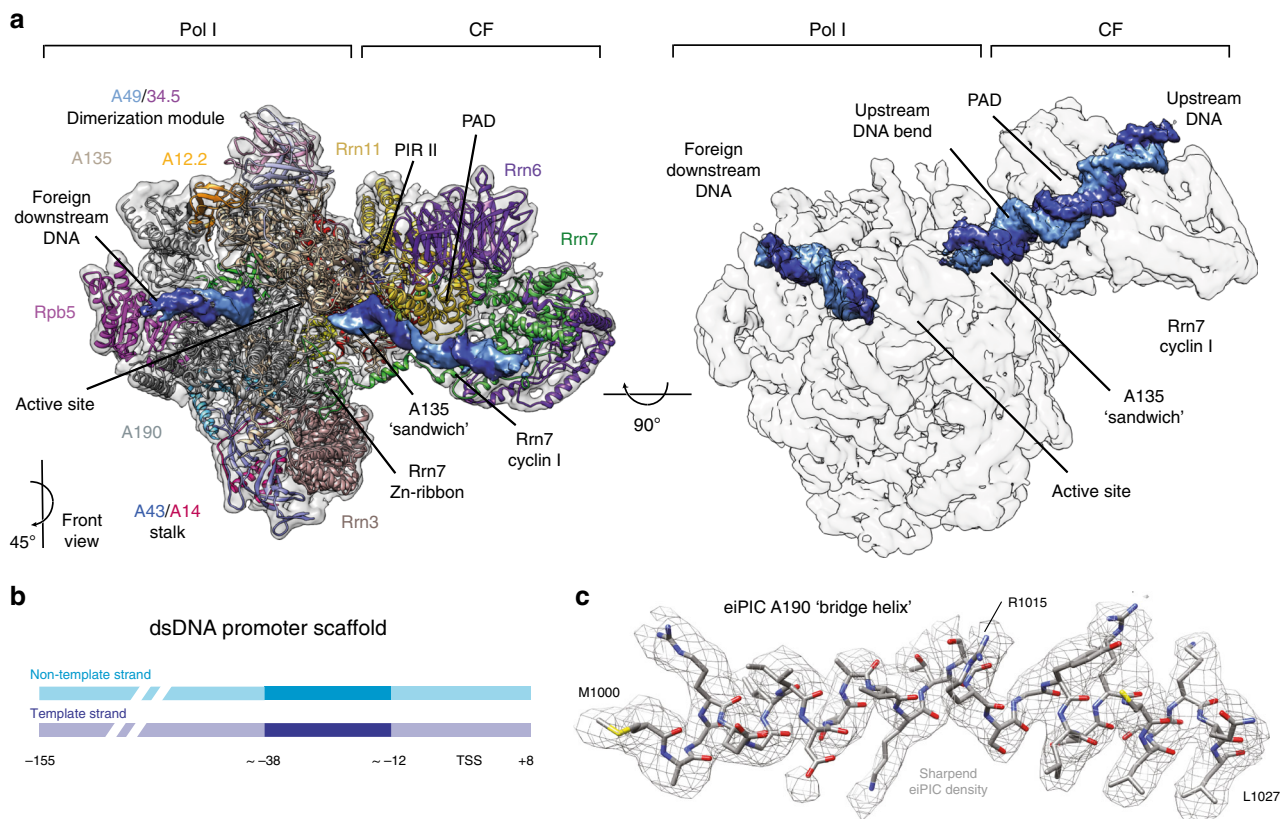
<sup>1</sup>Regensburg Center for Biochemistry, University of Regensburg, Regensburg, Germany. ✉email: [christoph.engel@ur.de](mailto:christoph.engel@ur.de)

The transcription of the ribosomal RNA (rRNA) precursor by RNA polymerase (Pol) I is a prerequisite for ribosome biosynthesis in all known eukaryotes<sup>1</sup>. As such, Pol I transcription is tightly regulated, mostly at the level of pre-initiation complex (PIC) formation<sup>2–6</sup>. Whereas Pol II and Pol III use related initiation mechanisms, the processes underlying Pol I promoter recognition, PIC formation and DNA melting substantially diverge<sup>7–9</sup>. In bakers' yeast *Saccharomyces cerevisiae*, a basal initiation system required for Pol I activity consists of the promoter DNA core element (CE), specific initiation factor Rrn3 and heterotrimeric core factor (CF)<sup>10</sup>. CF binds a CE stretch between ~15 and 38 base pairs (bps) upstream of the transcription start site (TSS)<sup>11</sup> and recruits Rrn3-stabilized Pol I that is primed for initiation<sup>12–14</sup>. DNA melting occurs at a position slightly upstream of the TSS between the Pol I 'clamp core' and 'protrusion' domains<sup>15–17</sup>. No additional factors are required to commence initial transcription and promoter escape. In a complete system, however, upstream activating factor (UAF) recognizes an upstream element (UE) and cooperates with the TATA-binding protein (TBP) to stabilize CF association with the promoter, increasing Pol I initiation rates by up to 40-fold in vitro<sup>18–21</sup>. Furthermore, the factor Net1 may reside at Pol I promoters and has been described to enhance initiation in vivo and in vitro<sup>22,23</sup>.

During transcription initiation, Pols are recruited to their promoters by a set of general transcription factors, forming a 'closed complex' (CC). After melting of both DNA strands, an 'open complex' (OC) is established, transitioning into an 'initially

transcribing complex' (ITC) with the beginning of RNA chain synthesis. In ITCs, a stable DNA/RNA hybrid is formed and the polymerase has initiated movement into the gene before establishment of a processive elongation complex ('EC'; for a review of initiation phases compare refs. 24,25). Previous structural analyses of Pol I initiation complexes by us and others relied on an artificially stabilized, mismatched bubble scaffold assembled with an initially transcribed RNA sequence and a double-stranded DNA (dsDNA) sequence extending to up to 24 bps downstream of the TSS<sup>15–17</sup>. This experimental approach originates from the analysis of Pol II elongation complexes (ECs), preventing heterogenic sample conformations and making use of the tight DNA/RNA hybrid association with the polymerase<sup>26,27</sup>. In the case of the Pol I PIC a similar experimental strategy results in the visualization of late initiation intermediates. Consequently, an inconsistent occupancy of Rrn3 and divergent localization of the tandem-winged helix (twh) domain of Pol I subunit A49 and the C-terminal domain of subunit A12.2 have been observed<sup>15–17</sup>, leaving room for speculation with regard to the functional roles and temporal classification of the analyzed conformations during initiation<sup>7–9</sup>.

Therefore, we aimed at analyzing Pol I initiation mechanisms at an early initiation stage, allowing the visualization of promoter recognition, Pol I recruitment and DNA melting in a scenario as close to the native situation as possible. For this purpose, we assembled a complete initiation complex on double-stranded (ds) promoter DNA and performed single-particle cryo-EM analysis. The dsDNA scaffold was truncated on its downstream edge at



**Fig. 1** Cryo-EM reconstruction of a Pol I early intermediate PIC. **a** Overview of the Pol I eiPIC cryo-EM reconstruction at 3.5 Å resolution (unsharpened; transparent gray envelope) overlaid with the PDB model (colored ribbon) and DNA (space filling). The right panel shows transparent density (gray) for protein components and solid density for the DNA path (template strand in blue and non-template in light blue). PAD promoter-associated domain (of Rrn11); PIR polymerase interacting region (of CF). **b** Schematic representation of promoter dsDNA used for PIC assembly, densities observed in the eiPIC reconstruction are highlighted in blue and light blue for template strand and non-template strand, respectively. **c** Atomic model of the bridge helix in subunit A190 overlaid with sharpened eiPIC density (gray mesh) indicates residue orientations.

position +8 relative to the TSS, thus preventing a contact with the clamp core and jaw domains of the polymerase. Three-dimensional particle reconstruction, cryo-EM density refinement and structural modeling allow the placement of basal PIC components and a comparative PIC analysis of the three eukaryotic Pols. Furthermore, structure-guided analysis indicates how Pol-I-specific ribosomal DNA (rDNA) promoter melting may be achieved.

## Results

**Complex formation and cryo-EM analysis.** To study promoter recognition and DNA melting, we formed a complete Pol I initiation complex *in vitro*. UAF was assembled on a dsDNA promoter scaffold ranging from position –155 to +8 relative to the TSS together with TBP, CF, and a fragment of the protein Net1<sup>22,28,29</sup> (Methods). Endogenously purified Pol I<sup>13,30,31</sup> was pre-incubated with recombinant Rrn3<sup>32</sup> to reconstitute a complete early PIC that was stable throughout size exclusion chromatography (Supplementary Fig. 1a; Methods). Accordingly, Pol I could be recruited to a UAF/TPB/Net1/CF-bound promoter scaffold lacking sequence stretches required for forming extended downstream contacts with the jaw- and clamp-head domains of the polymerase. Single-particle cryo-EM data was collected on a Titan Krios equipped with Gatan K2 summit direct electron detector basically as described<sup>12,13</sup>. Following pre-processing, two-dimensional (2D)- and three-dimensional (3D)-classification in RELION<sup>33</sup>, a total of 122,099 particles were selected from 4,088 micrograph movies (Methods; Supplementary Fig. 1). A final cryo-EM reconstruction exhibits an overall resolution of 3.5 Å and shows a Pol I early intermediate PIC (eiPIC; Fig. 1 and Supplementary Fig. 1). The cryo-EM density clearly reveals secondary structure features for the entire particle and side chain orientations in most regions (Fig. 1c and Supplementary Fig. 2a–f). Despite protein–protein crosslinking, TBP, UAF, and Net1-CTR remain flexible, although apparently stabilizing CF similar to the co-activator ‘mediator’ in context of a Pol II PIC<sup>34</sup>.

### An early intermediate PIC exhibits a well-defined architecture.

Initial assignment located template and non-template DNA strands, Pol I, CF subunits, and Rrn3, followed by manual model building and real-space refinement, resulting in a model of high quality (Methods, Table 1). Upstream DNA is well-ordered between CF-interacting regions and entry into the Pol I active center cleft. Following the canonical DNA-path further downstream, however, no density is visible around the active center itself, but ≥12 well-defined base-pairs can be placed on the downstream edge between bridge helix and the clamp-head/jaw domains, even though our scaffold should not extend this far. Most likely the conserved<sup>35</sup> and highly charged region is bound by foreign DNA or the far upstream end of our scaffold. A similar effect was observed for patches of the nucleosome, after transcription by Pol II ‘peeled’ off supercoiled DNA<sup>36</sup>. Well in line, *in vitro* initiation assays previously showed a strong preference for Pol I to initiate from dsDNA ends of synthetic sequences<sup>15</sup>.

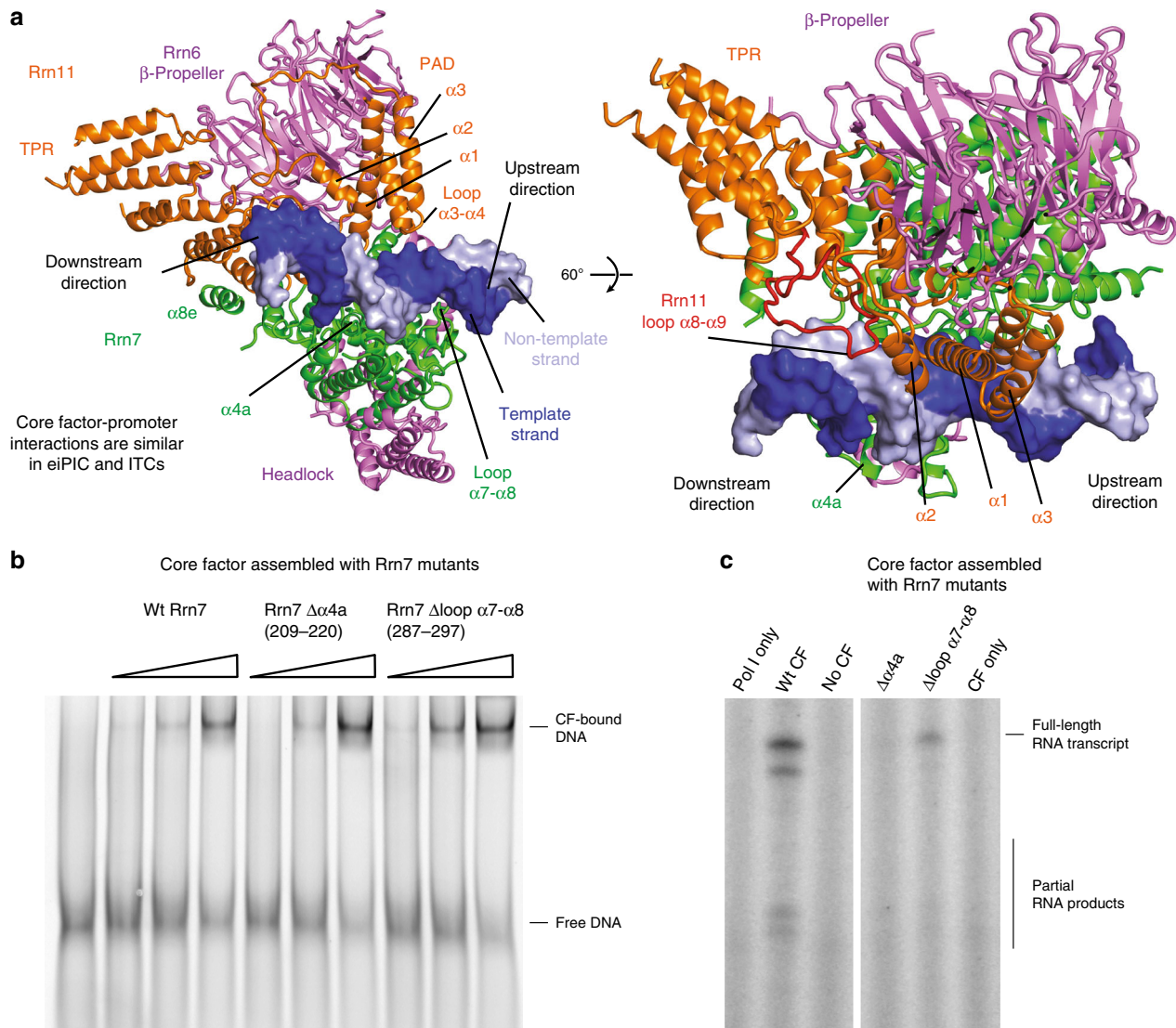
Initiation factor Rrn3 is tightly bound to Pol I ‘stalk’ and ‘dock’ subdomains<sup>12,13</sup> in all analyzed particles, agreeing with chromatin immunoprecipitation (ChIP) and biochemical studies in yeast<sup>2,32,37</sup> and mouse<sup>38,39</sup> cells. CF is associated with the Pol I core via its polymerase interacting regions (PIR) similar to ITC conformations<sup>15–17</sup>. Excellent quality of the cryo-EM density allowed us, to rebuild the CF subunits Rrn6, Rrn7 and Rrn11, consolidating divergent assignments in the crystal structure<sup>15</sup> (PDB 5O7X) and an ITC EM-based model<sup>17</sup> (PDB 5W66). In contrast to inactive Pol I<sup>30,31,40</sup>, the ‘expander’ and ‘connector’ subdomains are flexible and the central bridge helix is refolded in

**Table 1 Cryo-EM data collection, refinement, and validation statistics.**

	eiPIC EMDB-10544 PDB 6TPS	CF in eiPIC EMDB-10663
Data collection and processing		
Magnification	105,000	105,000
Voltage (kV)	300	300
Electron exposure (e <sup>-</sup> /Å <sup>2</sup> )	56	56
Defocus range (μm)	–1.5 to –3.1	–1.5 to –3.1
Pixel size (Å)	1.09 (0.545 superres)	1.09 (0.545 superres)
Symmetry imposed	C1	C1
Initial particle images (no.)	311,557	311,557
Final particle images (no.)	122,099	122,099
Map resolution (Å)	3.54	3.91
FSC threshold	0.143	0.143
Map resolution range (Å)	3.3 to 9.9	3.4 to 16.7
Refinement		
Initial model used (PDB code)	6TPS	
Model resolution (Å)	3.5	
FSC threshold	0.143	
Map sharpening B factor (Å <sup>2</sup> )	–75	
Model composition		
Non-hydrogen atoms	50,070	
Protein residues	6,109	
Ligands	8 (Zn and Mg)	
B factors (Å <sup>2</sup> )		
Protein	65.6	
Ligand	102.9	
R.m.s. deviations		
Bond lengths (Å)	0.009	
Bond angles (°)	0.985	
Validation		
MolProbity score	1.85	
Clashscore	5.96	
Poor rotamers (%)	0.59	
Ramachandran plot		
Favored (%)	91.10	
Allowed (%)	8.75	
Disallowed (%)	0.15	

the eiPIC (Fig. 1c) as expected from EC structures<sup>41,42</sup>. The C-terminal domain of subunit A12.2 shows only residual density in funnel domain of subunit A190 (Supplementary Fig. 2b), but is not localized on the A135 lobe as observed in a 12-subunit EC<sup>43</sup>. Our eiPIC reconstruction shows strong density for the A49/A34.5 dimerization and A34.5 C-terminal tail domains (Supplementary Fig. 2e), indicating that the heterodimer is constitutively attached. The twh and linker domains of subunit A49 are detached in the eiPIC, agreeing with a proposed role in promoter escape<sup>17</sup>.

**Core factor embraces the promoter DNA.** The eiPIC density allows the construction of a CF model, which we found to resemble the overall ITC conformation. To define the structural changes that take place upon promoter recruitment, we compared the architecture of CF in free (PDB 5O7X) and promoter-engaged eiPIC conformation (Supplementary Fig. 3). This shows that CF module I and II retract from each other by up to 12 Å upon



**Fig. 2 Core Factor—promoter interactions in eiPIC. a** Model of promoter-bound CF in the eiPIC. The same regions of Rrn7 and Rrn11 contribute to promoter phosphate backbone interactions compared to ITC reconstructions. **b** Electrophoretic mobility shift assay (EMSA) shows that wild-type CF interacts with double-stranded promoter DNA (0.25 pmol, 0.5 pmol, and 1 pmol CF added). Mutation of Rrn7 ( $\Delta\alpha 4a$  and  $\Delta$ loop  $\alpha 7$ - $\alpha 8$ ) does not impair promoter-DNA association. **c** In contrast to DNA binding, initiation efficiency of CF assembled with Rrn7 mutants  $\Delta\alpha 4a$  and  $\Delta$ loop  $\alpha 7$ - $\alpha 8$  is impaired (promoter-dependent in vitro transcription assay from a minimal scaffold).

binding of the CE promoter sequence. This retraction leads to the exposure of positively charged residues that are now free to engage the phosphate backbone (Supplementary Fig. 3a–c). These DNA-binding regions lie within the Rrn11 promoter-associated domain ('PAD') and the cyclin domains of Rrn7. The same regions engage the DNA in ITCs<sup>15–17</sup> and have been described in detail in late ITCs devoid of Rrn3<sup>17</sup>. Remarkably, the Rrn7 residues involved in DNA-binding are not conserved within TFIIB or Brl1, which share a similarity in their overall fold<sup>44–46</sup> and would clash with TBP<sup>15</sup> in canonical TFIIB-TBP<sup>47</sup> or Brl1-TBP<sup>48,49</sup> complex.

Comparison of free and promoter-engaged CF also shows that the Rrn7-specific helix  $\alpha 4a$  in the N-terminal cyclin domain shifts and is inserted into the minor groove of the CE promoter DNA, while loop  $\alpha 7$ - $\alpha 8$  in cyclin II becomes well-structured and contacts the major groove further upstream upon eiPIC formation (Fig. 2a). Thereby, the distal upstream DNA-path is modified towards the C-terminal domain of Rrn7 and the  $\beta$ -propeller-domain of Rrn6. Thus, promoter binding by Rrn7-

specific regions on one face and by the TFIIB-unrelated CF subunit Rrn11 on the opposite face tightly squeeze the DNA. This may explain why the basal Pol I initiation system does not require TBP association opposite of the Rrn7 cyclins.

To address the importance of these residues, we constructed CF mutants with deletions in helix  $\alpha 4a$  and in loop  $\alpha 7$ - $\alpha 8$ . Both can still associate with promoter DNA (Fig. 2b), but show defects in basal initiation in vitro (Fig. 2c). Engagement of these regions may therefore be important to induce a specific DNA conformation required for Pol I recruitment or promoter melting.

#### The Pol I 'sandwich' region is important for PIC formation.

We have previously described a Pol-I-specific proximal upstream promoter-binding region consisting of loop  $\alpha 11a$ - $\alpha 12$  (residues 452–456) and the loop  $\beta 28$ - $\beta 28$  (residues 815–818) in the protrusion and wall domains of Pol I subunit A135, respectively<sup>15</sup>. In the eiPIC, a positively charged loop (892–895, wall domain of subunit A135) re-orientates towards the promoter DNA, contributing

additional phosphate-backbone interactions (Supplementary Fig. 4) similar to other ITC/PIC structures<sup>16,17</sup>. These promoter interactions are all specific to Pol I, because the residues are not conserved in Pol II<sup>50</sup> and III<sup>51</sup>. Furthermore, DNA is occluded from the corresponding region in Pol II and III PICs by the N-terminal cyclin domains of TFIIB<sup>34,52,53</sup> and Brf1/Brf2<sup>48,49,54,55</sup>, respectively. Fittingly, this Pol I region was previously named ‘sandwich’<sup>17</sup>.

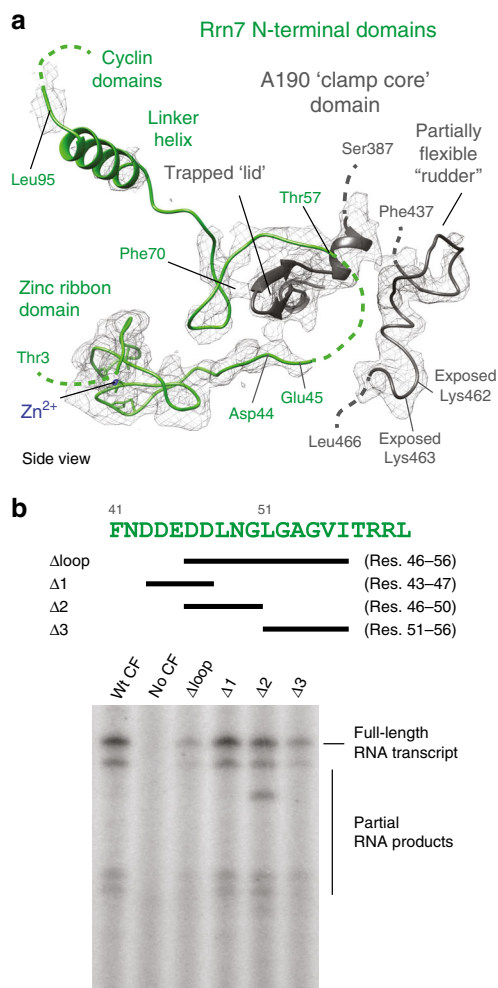
In the eiPIC, the sandwich region tightly holds the promoter in place between the wall and protrusion domains at the bottom of the cleft. sandwich elements contact both DNA strands, therefore rendering it specific for an un-melted duplex. Density for the DNA directly downstream of the sandwich is not observed, indicating a higher degree of flexibility. Consequently, the recruitment of the Pol-I-Rrn3 complex seems to mainly rely on (1) contacts between the promoter and the sandwich and (2) protein–protein contacts between CF and the Pol-I-Rrn3 complex. In contrast, further promoter contacts with the Pol I cleft or downstream elements and/or A49 appear not to be required for recruitment.

**TFIIB-related elements in Rrn7 adopt divergent positions.** The TFIIB-related ‘reader’ and ‘linker’ elements within Rrn7<sup>44,46</sup> are mostly ordered in the active center cleft of the eiPIC, with the exception of the residues 46–56 (B-reader homologous<sup>56</sup>). The protein backbone extends from the N-terminal zinc ribbon into the Pol I cleft, apparently trapping the well-ordered ‘lid’ sub-domain of Pol I subunit A190 before forming two anti-parallel strands and exiting the Pol I upstream face on the side of the shelf module (Fig. 3a). The path of Rrn7 differs from a Pol I ITC<sup>17</sup> and from TFIIB in complex with Pol II<sup>57</sup> (Supplementary Fig. 5). During Pol II initiation, the TFIIB-reader-loop contacts the ‘rudder’ and the ‘fork loop I’ domains, while the TFIIB-linker binds the top of the rudder and forms a helix that interacts with the clamp core domain<sup>57</sup>. In the eiPIC, rudder and fork loop I apparently interact neither with each other nor with the TFIIB-reader-homologous regions of Rrn7. Instead, rudder and fork loop I are oriented towards the bridge helix and an Rrn7 helix that is similar to the TFIIB linker connects to CF module II.

In addition to a divergent path of Rrn7 compared to TFIIB, the residues contacting the template strand in a Pol I ITC<sup>17</sup> and Pol II ITC<sup>57</sup> are mostly flexible in the eiPIC, but not in Pol II CCs<sup>52</sup> or in a Pol II-TFIIB<sup>57</sup> complex. Furthermore, TFIIB reader-loop arginine residue 78, which is important for TSS selection by Pol II<sup>58</sup>, does not exist in Rrn7. This adds to overall sequence<sup>44,46</sup> and architecture differences<sup>15</sup> between Rrn7 and TFIIB.

To clarify the importance of Rrn7 loop residues disordered in the eiPIC, we mutated the entire loop or smaller stretches and analyzed CF initiation activity in a basal assay (Fig. 3b). The loop-deletion Rrn7 mutant shows strongly reduced initiation efficiency, which can mainly be attributed to the residues 51–56, but not to residues 43–50. The Rrn7 version with loop-deletion still assembles well with Rrn6 and Rrn11 and is able to form a basal PIC in vitro (Supplementary Fig. 3d, e). Thus, the Rrn7-reader-loop is likely important for promoter melting.

**Pol I is primed for initiation at the eiPIC stage.** Modeling of the active center based on our eiPIC density indicates, that aspartate 629 in subunit A190 (Asp483 in Pol II subunit Rpb1) has apparently changed its orientation with respect to the dimeric crystal structures<sup>30,31</sup> (Fig. 4). Assuming its active orientation in the eiPIC, Asp629 now allows coordination of the catalytic magnesium ion (‘metal A’), together with Asp627 and Asp631 for which we observe a clear cryo-EM density peak (Fig. 4b). In addition, the hybrid-binding domain of subunit A135 re-arranges to form a one-turn helix in the eiPIC. This helix also resembles

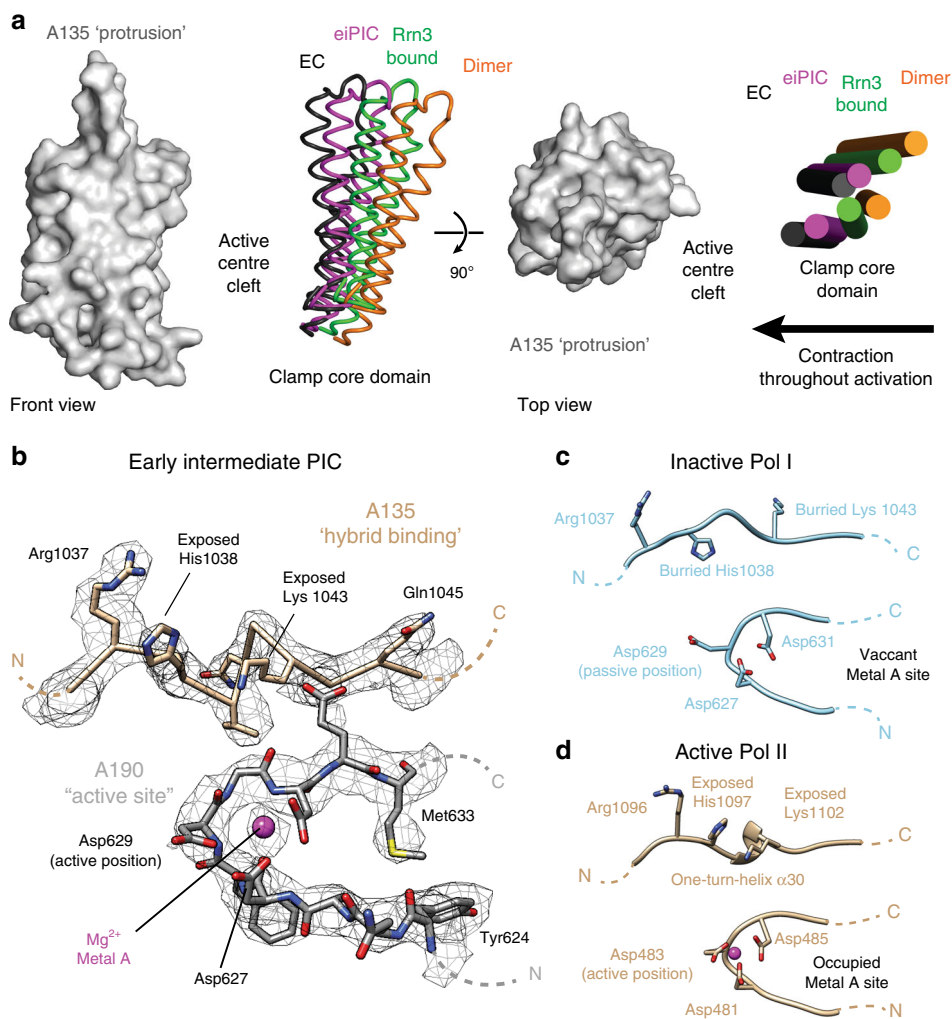


**Fig. 3 The N-terminal region of Rrn7 is partially ordered within the eiPIC.**

**a** Ribbon model of TFIIB-homologous regions in the N-terminus of Rrn7 (green) overlaid with sharpened eiPIC density (gray mesh). The lid domain of Pol I subunit A190 (dark gray ribbon) is trapped between well-ordered regions of Rrn7. Residues 46 to 56 of Rrn7 are partially flexible, hinting at a function during promoter melting. **b** Amino acid sequence of flexible Rrn7 region is shown in green with schematic representation of deletion mutants indicated by black bars. Basal in vitro initiation assay shows the effect of Rrn7 mutations within this loop: Deletion of the entire loop and its C-terminal part (51 to 56) show reduced initiation activity.

the active Pol I, II, and III EC conformations and its formation exposes histidine 1038 to the bottom of the cleft, which is now free to contact the hybrid upon initial transcription as observed in ITCs. Furthermore, the previously buried lysines 462 and 463 in subunit A190 become exposed in the eiPIC (Supplementary Fig. 2f), now resembling the active Pol-II-fold<sup>59</sup> and contacting the first visible downstream DNA base pair. This may contribute to a high affinity for foreign DNA and to the Pol I preference for initiation from ends of dsDNA. With the described structural changes upon eiPIC formation, Pol I enters a conformation that is primed for initial transcription via a conserved mechanism<sup>60</sup> in the presence of NTPs.

We also observe, that the Pol I cleft continues to contract downstream of the sandwich region, adapting an intermediate conformation between the Rrn3-bound and ITC/actively elongating states (Fig. 4a). This adds an additional intermediate to the set of Pol I structures<sup>9</sup>, but is in line with the suggestion, that cleft modulation is a major regulatory mechanism of Pol I



**Fig. 4 Pol I is primed for initiation in the eiPIC. a** Cleft contraction throughout activation stages. Pol I structural models were overlaid via their A135 subunits (protrusion subdomain in gray, space filling). Cleft contraction is indicated by colored clamp core helices (subunit A190). Monomeric Pol I and ITC stages are similar to Rrn3-bound- and EC-conformations, respectively (not shown for clarity). PDB models displayed: 4C2M (orange), 5G5L (green), eiPIC (magenta) and 5M3F (black). **b** Atomic model of the active center and hybrid-binding domains within Pol I subunits A190 and A135, respectively. Overlaid with sharpened eiPIC density (gray mesh). The metal A site is occupied and a one-turn-helix  $\alpha$ 30 is formed in A135, exposing positively charged residues. **c** Inactive Pol I (PDB 4C2M) region for comparison to **b**. **d** Active Pol II region for comparison to **b** and **c**.

transcription<sup>14,30,61</sup>. At the stage of DNA-melting during the transition from CC to OC states, dsDNA cannot be accommodated between clamp core and protrusion domains any longer<sup>15</sup>. Hence, simultaneous promoter loading and cleft contraction allosterically destabilize the upstream duplex at the position of the clamp core and may foster spontaneous melting at this position. Notably, the initially melted region shows the highest conservation among rDNA promoters identified thus far<sup>62</sup>. Thus, the eiPIC apparently represents a trapped CC-OC transition intermediate conformation, which is important for spontaneous DNA-melting to take place during promoter association of the polymerase.

## Discussion

Within this work, we describe an early intermediate initiation complex. The structure enables the independent discussion of promoter recruitment and DNA-melting in a sequential manner. Apparently the polymerase is recruited to its dsDNA promoter but cannot complete the melting process due to a lack of fixated downstream DNA. We described the eiPIC reconstruction in the context of PIC formation and continue to update our model of

Pol I recruitment and DNA-melting in light of these findings. Our interpretation is well in line with the idea that targeting of the initiation machinery to the rDNA promoter depends mostly on UAF, and TBP serves to position CF downstream of the UE, while interacting with the promoter using a divergent interface<sup>63</sup>. Recruitment of the Pol-I-Rrn3 complex then relies on a specific DNA architecture<sup>64</sup>, namely a bendability that allows interactions of the Rrn11 TPR domain with the Pol I protrusion<sup>15</sup> and binding of a promoter element to the Pol I sandwich region (Supplementary Fig. 4). Since our assembly originally comprised UAF and TBP, and only a single reconstruction was obtained from 39% of all recorded particles, it is likely that we capture a physiologically relevant conformation, while factors were artificially positioned by DNA/RNA hybrid scaffolds simulating initial transcription in previous analyses<sup>15–17</sup>, even though RNA was lost in one case<sup>16</sup>.

Within the eiPIC structure, re-arrangements between CF module I and II enable Rrn7 and Rrn11 to bind promoter DNA, mainly by phosphate backbone interactions of basic loops. This explains the (low) sequence specificity of DNA-binding by CF and thus the overall similar eiPIC architecture compared to ITCs and late PIC reconstructions. Likely, Rrn7-specific

DNA-interacting loops contribute to DNA-conformational modulation (compare Fig. 2). We further confirm cleft contraction between the protrusion and clamp core domains and exposure of basic residues bottom of the cleft during DNA-melting by Pol I in the eiPIC.

While our findings do not oppose the idea of an upstream ratcheting mechanism to open Pol I promoter DNA, we also see no evidence to support such a mechanism deduced from shifts in CF-positions observed in ITC reconstructions<sup>17</sup>.

Instead, we propose a simplified melting-mechanism based on steric DNA-distortion and electrostatic single-strand trapping which, in this combination, is only possible in Pol I, but not in Pol II and III. Firstly, Pol I recruitment relies on DNA-duplex binding to the sandwiching region and DNA positioning within the expanded cleft of the Pol-I-Rrn3 complex (Fig. 4a and Supplementary Fig. 4). Sequence specificity is determined by proximal upstream bendability<sup>15,65</sup> and distal upstream recognition by UAF, which is linked to the PIC via CF and TBP. Divergent TFIIB reader-loop elements within Rrn7 are placed in the Pol I cleft, may play a role in duplex-destabilization and bind the melted template strand similar to observations in ITCs<sup>17</sup>. In addition, allosteric duplex-destabilization resulting from a cleft contraction between the clamp and protrusion domains observed in the eiPIC likely contributes to melting (Fig. 4a). This contraction primes Pol I for initial transcription by re-ordering previously inactivated regions (Figs. 1 and 4 and Supplementary Fig. 2). Exposed basic residues can then contribute to stabilization of the initially melted template strand and ultimately the DNA/RNA hybrid at the bottom of the cleft. Furthermore, the non-template strand may be bound by the A49 linker (as observed in ref. 17), thereby preventing collapse of the early bubble similar to the  $\sigma$ -factor in bacterial Pol<sup>66,67</sup>. Only after initial transcription, the growing RNA chain can interact with Rrn7 and would finally clash with reader/linker elements, freeing the exit channel and expelling Rrn7 from the polymerase. This is probably concerted with the association of the flexible A49 twh domain at the back of the clamp core domain, leading to dissociation of CF and Rrn3 and preventing re-association, thereby fostering promoter escape.

In Pol II and Pol III initiation complexes<sup>48,49,53,68</sup>, TFIIB/Brf1 cyclin domains occlude the sandwiching region and reader/linker domains diverge from Rrn7, preventing a similar mechanism. Arguing for a model of combined adaptations, a number of CF-mutations impaired in vitro initiation rates, but only large deletions completely abolished functionality<sup>15,45</sup>. Furthermore, a 12-subunit Pol I lacking A49/A34.5 is still able to initiate from its native promoter (although the lack of A49 linker-positioning strongly impaired the process)<sup>13,37</sup>, TBP is not necessary for basal transcription<sup>11,21</sup> and single A49 mutations have only minor effects on Pol I function<sup>69</sup>. Thus, the overall functionality of the system is robust and highly adaptive to conditional variations. However, full initiation rates required for physiological growth depend on the combined action of all Pol-I-specific elements that have accumulated throughout evolutionary adaptation and are basically conserved throughout eukaryotic organisms<sup>10,70,71</sup>. These adaptations increase initial transcription to such efficiency, that formation of a stable closed complex under physiological conditions appears unlikely. While such a state may be transiently established, the instant cleft contraction and Rrn7-dependent duplex-destabilization by the combined action of Pol I and CF elements directly lead to melting and prime the polymerase for initial transcription and hybrid stabilization.

During the final stages of revision of this work, a related study was published<sup>72</sup>. Sadian et al. provide an excellent description of CF-promoter contacts in detail and investigate the role of an acidic loop in Rrn3, based on higher resolution reconstructions. Compared to our results, interpretation relies on a minor subset

of 0.7% or 0.5% of particles from two datasets indicating a transient nature of CCs. In our UAF/TBP-containing samples, however, 39% of initial particles contribute to the final reconstruction and divergent CF-positions are not observed. This may be due to a lack of available particles in our datasets, or due to stabilization of a more 'native' CF-orientation in the presence of UAF/TBP. Detailed structure-function analysis of UAF- and TBP-contributions are now instrumental to understanding the process of Pol I initiation in its entirety.

## Methods

**Protein expression and purification.** Partially purified endogenous *Saccharomyces cerevisiae* Pol I is a by-product of Pol III purification via a TAP-tag on subunit AC40<sup>49</sup>. The Pol-I-containing MonoQ fractions were a gift from A. Vannini and G. Abascal-Palacios. Fractions were pooled, diluted fivefold in buffer A (20 mM HEPES/KOH pH 7.8, 10% glycerol, 1 mM MgCl<sub>2</sub>, 10  $\mu$ M ZnCl<sub>2</sub>, 5 mM dithiothreitol (DTT)) and loaded onto a MonoS HR 5/5 column (GE Healthcare). Separation was performed with a gradient from 10–37.5% buffer B (buffer A with 2 M KAc) with a 2 CV plateau at 17.5% B. Pol I eluted at 470 mM KAc<sup>73</sup>, peak fractions were pooled, flash frozen in liquid nitrogen and stored at –80 °C.

Rrn3<sup>32</sup> was expressed in BL21(DE3) pRIL (Agilent) cells, by autoinduction in TB medium (1.2% tryptone; 2.4% yeast extract; 0.5% glycerol); 1/10 volume of a sterile solution containing 0.17 M KH<sub>2</sub>PO<sub>4</sub> and 0.72 M K<sub>2</sub>HPO<sub>4</sub> and 1/50 volume of a sterile solution containing 25% glycerol; 10% lactose and 1% glucose were added. A culture was grown at 37 °C to an OD<sub>600</sub> of 0.6, after cooling the culture on ice, incubation was continued at 16 °C overnight. Cells were harvested (6000 g; 10 min), resuspended in lysis buffer (50 mM HEPES at pH 7.8, 200 mM NaCl, 3 mM DTT, 10% glycerol). A 3 ml Ni-NTA column (Qiagen) was equilibrated with lysis buffer, the supernatant loaded, and the column was washed with lysis buffer containing 25 mM imidazole. Elution was carried out in lysis buffer containing 150 mM imidazole. Next, Rrn3 was further purified by anion exchange chromatography (Mono Q 5/50, GE Healthcare). The column was equilibrated in MonoQ buffer 1 (50 mM HEPES at pH 7.8, 5 mM DTT, 10% glycerol), and proteins were eluted with a linear gradient of 20 column volumes from 100 mM to 1 M NaCl. After concentration (Amicon, 35 kDa cutoff), the sample was applied to a Superdex 200 increase 10/300 size exclusion column (GE Healthcare) equilibrated with buffer Rrn3-SEC (20 mM HEPES at pH 7.8, 300 mM NaCl, 5 mM DTT).

CF subunits<sup>15</sup> were co-expressed in *E. coli* BL21-CodonPlus(DE3)-RIL cells (Agilent) from two plasmids. A 4 l culture was grown in LB medium at 37 °C until OD<sub>600</sub> reached 0.5–0.7. Cultures were cooled on ice for 20 min and expression was induced with 0.1 mM IPTG. Cells were grown at 18 °C overnight. Cells were harvested by centrifugation, washed with phosphate-buffered saline (PBS) at 4 °C, flash frozen in liquid nitrogen and stored at –80 °C. One pellet was suspended in buffer CF-A (20 mM imidazole, 350 mM NaCl, 10 mM MgCl<sub>2</sub>, 10% (v/v) glycerol, 20 mM HEPES pH 7.8, 1 mM DTT, 1x protease inhibitor). Cells were lysed by sonication using a Branson Digital Sonifier, the lysate was cleared by centrifugation and the supernatant was filtered with a 0.22  $\mu$ m filter (Millipore) to remove cell debris. Cell lysate was then applied to a Ni-NTA column (5 ml, GE Healthcare) and bound CF washed with 5 CV of buffer CF-B (25 mM imidazole, 200 mM NaCl, 10 mM MgCl<sub>2</sub>, 10% (v/v) glycerol, 20 mM HEPES pH 7.8, 1 mM DTT) at 4 °C. The column was transferred to room temperature, washed with 2.5 CV of buffer CF-C (50 mM imidazole, 200 mM NaCl, 10 mM MgCl<sub>2</sub>, 10% (v/v) glycerol, 20 mM HEPES pH 7.8, 1 mM DTT, 5 mM ATP, 2 mg/ml denatured protein), incubated for 10 min, and washed again with 2.5 CV buffer CF-C. The column was transferred to 4 °C and washed with 5 CV buffer CF-D (50 mM imidazole, 200 mM NaCl, 10 mM MgCl<sub>2</sub>, 10% (v/v) glycerol, 20 mM HEPES pH 7.8, 1 mM DTT). Elution was performed with 5 CV of buffer CF-E (350 mM imidazole, 200 mM NaCl, 10 mM MgCl<sub>2</sub>, 10% (v/v) glycerol, 20 mM HEPES pH 7.8, 1 mM DTT). Protein was then loaded on a 5 ml heparin column (GE Healthcare) in buffer CF-F (200 mM NaCl, 1 mM MgCl<sub>2</sub>, 10% (v/v) glycerol, 20 mM HEPES pH 7.8, 1 mM DTT) and eluted with a gradient ranging from 0.2 to 2.0 M NaCl, including a plateau at 550 mM NaCl of 2 CVs. CF-containing fractions were concentrated using a 100 kDa cutoff centrifugal filter (Millipore). Size exclusion chromatography was carried out with a Superose 6 increase 10/300 column (GE Healthcare) in buffer CF-G (200 mM NaCl, 1 mM MgCl<sub>2</sub>, 5% (v/v) glycerol, 10 mM HEPES pH 7.8, 10  $\mu$ M ZnCl<sub>2</sub>, 1 mM DTT). CF-containing fractions were concentrated using a 100 kDa cutoff centrifugal filter (Millipore) and directly used or flash frozen in liquid nitrogen for storage at –80 °C.

*S. cerevisiae* TBP was cloned into vector pET28b via NheI/Not I restriction sites (compare Supplementary Table 1). Recombinant His<sub>6</sub>-TBP protein was expressed in BL21(DE3) pRIL (Agilent) cells, by autoinduction in TB medium (1.2% tryptone; 2.4% yeast extract; 0.5% glycerol); 1/10 volume of a sterile solution containing 0.17 M KH<sub>2</sub>PO<sub>4</sub> and 0.72 M K<sub>2</sub>HPO<sub>4</sub> and 1/50 volume of a sterile solution containing 25% glycerol; 10% lactose and 1% glucose were added. A culture was grown at 37 °C to an OD<sub>600</sub> of 0.6, after cooling the culture on ice, incubation was continued at 16 °C overnight. Cells were harvested (6000 g; 10 min), resuspended in lysis buffer (50 mM HEPES/KOH; 10% glycerol; 10 mM MgAc<sub>2</sub>; 200 mM KCl; 10 mM imidazole; 5 mM  $\beta$ -mercaptoethanol; 1 mM

phenylmethylsulphonyl fluoride (PMSF); 2 mM benzamidine), and lysed by sonication (Branson Sonifier 250 macrotip, cooling in icewater). The cell extract was cleared twice (40,000 g for 40 min at 4 °C) and incubated with 1 ml equilibrated NiNTA Agarose (Qiagen) at 4 °C for 2 h on a rotating wheel. The resin was transferred to a polypropylene column (Bio-Rad), washed with wash buffer 1 (20 mM HEPES/KOH; 10% glycerol; 5 mM MgAc<sub>2</sub>; 1 M KCl; 20 mM imidazole; 5 mM β-mercaptoethanol), wash buffer 2 (as wash buffer 1 but with 0.2 M KCl) and eluted with elution buffer (20 mM HEPES/KOH; 10% glycerol; 5 mM MgAc<sub>2</sub>; 0.2 M KCl; 200 mM imidazole; and 5 mM β-mercaptoethanol). The sample was diluted with buffer C (20 mM HEPES/KOH pH 7.8, 10% glycerol, 1 mM MgCl<sub>2</sub>, 5 mM DTT) to ≈100 mM KCl and loaded onto a MonoS 5/50 GL column (GE Healthcare) and eluted with a linear gradient from 10–100% buffer D (buffer C + 1 M KCl). TBP containing fractions were pooled, concentrated and loaded onto a Superdex 75 10/300 Increase column (GE Healthcare) equilibrated in buffer E (20 mM HEPES/KOH pH 7.8, 10% glycerol, 200 mM KCl, 1 mM MgCl<sub>2</sub>, 5 mM DTT) peak fractions were pooled, flash frozen in liquid nitrogen and stored at –80 °C.

Net1-CTR-TAP was expressed in baculovirus infected SF21 cells as recently published<sup>22,74–76</sup>. Specifically, 50 × 10<sup>6</sup> cells were resuspended in buffer F (20 mM HEPES/KOH pH 7.8, 10% glycerol, 200 mM KCl, 1 mM MgCl<sub>2</sub>, 0.2% NP40, 1 mM DTT), lysed by sonication (Branson Sonifier 250 macrotip, cooling in icewater) and cleared by centrifugation (40,000 × g for 40 min at 4 °C). The supernatant was incubated four hours with 1 ml IgG Sepharose 6 Fast Flow (GE Healthcare), washed with buffer F and incubated with TEV protease for two hours at 16 °C. Eluate fractions were collected, flash frozen in liquid nitrogen and stored at –80 °C. The MultiBac system was also used to generate a baculovirus co-expressing Rrn5-HA, Rrn9-Flag, Uaf30-His7, Rrn10, Histones H3 and 4 in SF21 cells (Supplementary Table 1). In all, 1 × 10<sup>9</sup> cells were resuspended in lysis buffer (50 mM HEPES/KOH pH 7.8, 10% glycerol, 400 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 10 mM MgCl<sub>2</sub>, 20 mM imidazole, 1 mM DTT) and lysed by sonication (Branson Sonifier 250, cooling in icewater). The cleared lysate (2 × 70,000 g, 45 min, 4 °C) was incubated with 2 ml Ni-NTA Agarose (Qiagen) for 2 h, beads were washed with buffer G (20 mM HEPES/KOH pH 7.8, 10% glycerol, 1 M KCl, 5 mM MgCl<sub>2</sub>, 20 mM imidazole, 1 mM DTT), buffer H (20 mM HEPES/KOH pH 7.8, 10% glycerol, 400 mM KCl, 2 mM MgCl<sub>2</sub>, 50 mM imidazole, 1 mM DTT) and eluted with buffer I (20 mM HEPES/KOH pH 7.8, 10% glycerol, 400 mM KCl, 2 mM MgCl<sub>2</sub>, 300 mM imidazole, 1 mM DTT). The eluate was diluted with buffer C, loaded onto a MonoS HR 5/5 column and eluted with a gradient to buffer D (see above). Eluate fractions were collected, concentrated, flash frozen in liquid nitrogen and stored at –80 °C. Raw SDS-PAGE gels are shown in Supplementary Fig. 6.

**Promoter-dependent in vitro transcription.** Promoter-dependent in vitro transcriptions were performed following our previously published protocols<sup>13,15</sup> on core promoter scaffolds from position –38 to +24 relative to the TSS (Supplementary Table 1). Specifically, Promoter-dependent in vitro transcription reactions were performed as follows: A total of 50 ng template dsDNA template were used for each transcription reaction (25 μl reaction volume). CF was added to a final concentration of 20 nM pre-incubated Pol-I-Rrn3 complex was added to a final concentration of 4 nM. 20 mM HEPES/KOH pH 7.8 and 2 M KAc were added to adjust volume and salt concentration to the final reaction conditions of 150 mM KAc in 25 μl. Transcription was started by the addition of 5 μl 5x transcription buffer (100 mM HEPES/KOH pH 7.8, 50 mM MgCl<sub>2</sub>, 25 mM EGTA, 0.25 mM EDTA, 3 mM DTT, 1 mM ATP, 1 mM UTP, 1 mM CTP, 0.05 mM GTP complemented with 0.3 μl [α-<sup>32</sup>P]GTP (10 mCi/ml; Hartmann Analytic). The samples were incubated at 24 °C for 30 min. Next, 200 μl Proteinase K buffer (0.5 mg/ml Proteinase K in 0.3 M NaCl, 10 mM Tris/HCl pH 7.5, 5 mM EDTA and 0.6% SDS) was added to stop transcription. The samples were incubated at 56 °C for 15 min. Ethanol (700 μl) was added to allow precipitation of nucleic acids (30 min at –80 °C). The samples were centrifuged for 10 min at 12,000 × g, the supernatant was removed and the precipitate was washed with 150 μl 70% ethanol. After centrifugation, the supernatant was removed and the pellets were dried at 95 °C for 1 min. RNA in the pellet was dissolved in 12 μl 80% formamide, 0.1 M TRIS-Borate-EDTA (TBE), 0.02% bromophenol blue and 0.02% xylene cyanol. Samples were heated for 2 min under vigorous shaking at 95 °C and briefly centrifuged. After separation on a 20% polyacrylamide gel containing 8 M urea and 1x TBE. Radiolabelled transcripts are visualized using a PhosphorImager (GE Healthcare). Raw gels are shown in Supplementary Fig. 6.

**Electrophoretic mobility shift assays (EMSA).** For EMSA experiments, a fluorescently labeled promoter fragment (–83 to +26 relative to the TSS) was annealed from oligonucleotides labeled with fluorescent dyes (NTS position –3 Atto647N and TS position –5 Cy3) as described below. 0.2 pmol dsDNA was incubated for 30 min without or with increasing amounts (0.25, 0.5, 1 pmol) of recombinant mutant CF in incubation buffer containing 20 mM HEPES/KOH pH 7.8, 10% glycerol, 200 mM KCl, 1 mM MgCl<sub>2</sub>, 0.1 mg/ml BSA, 1 mM DTT. The reaction was loaded on a pre-run 6% native acrylamide gel in 0.5x TBE buffer and imaged on a Typhoon FLA 9000 (GE Healthcare) imaging system. Raw gels are shown in Supplementary Fig. 6.

**Pol I PIC assembly.** The Pol I PIC was assembled on complementary rDNA promoter oligonucleotides AGCTTAAATTGAAGTTTTCTCGGCGAGAAATACGT AGTTAAGGCAGACGACAGAGAGGGCAAAGAAAATAAAGTAAAGATT TAGTTTGTAAATGGGAGGGGGGTTTAGTCATGGAGTACAAGTGTGAGGA AAAGTAGTTGGGAGGTACTTCATGCGAAA (NTS), TTTCGCATGAAGTACC TCCCAACTACTTTCTCCACACTTGACTCCATGACTAAACCCCCCTCCC ATTACAAACTAAAATCTTACTTTTATTTCTTCTGCTCTCTGCTGCTCT GCCTTAACTACGTATTTCTCGCCGAGAAAACTTCAATTTAAGCT (TS) (Integrated DNA Technologies). Oligonucleotides were dissolved in TE buffer (10 mM Tris pH 8, 0.5 mM EDTA), mixed in equimolar amounts to a final concentration of 10 μM each, heated to 95 °C and slowly cooled down to 10 °C with a cooling rate of 1 °C/min.

In all, 0.11 nmol promoter DNA was incubated with equimolar amounts of UAF, and threefold molecular excess of TBP and Net1-CTR. After 20 min incubation at 28 °C, 0.17 nmol CF was added and incubated for additional 20 min. 0.095 nmol Pol I, pre-incubated overnight with fivefold molar excess of Rrn3 on ice, was added and the sample was diluted with buffer G (20 mM HEPES/KOH pH 7.8, 2 mM MgCl<sub>2</sub>, 5 mM DTT) to final assembly conditions (20 mM HEPES/KOH pH 7.8, 50 mM KAc, 50 mM KCl, 2 mM MgCl<sub>2</sub>, 10 μM ZnCl<sub>2</sub>, 5 mM DTT; buffer H), incubated further 30 min and concentrated to 50 μl. The sample was crosslinked with 1 mM (bis(sulfosuccinimidyl) suberate) (BS3) for 30 min at 28 °C. The crosslinking reaction was quenched with 100 mM NH<sub>4</sub>HCO<sub>3</sub> final concentration for 15 min at 28 °C. The sample was loaded onto a Superose 6 PC 3.2/30 column (GE Healthcare) equilibrated with buffer H (20 mM HEPES/KOH pH 7.8, 50 mM KAc, 50 mM KCl, 2 mM MgCl<sub>2</sub>, 10 μM ZnCl<sub>2</sub>, 5 mM DTT) and collected in 60 μl fractions. A raw sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) gel of non-crosslinked sample is shown in Supplementary Fig. 6.

**Cryo-EM sample preparation and data acquisition.** Grids were glow discharged in Argon/Oxygen plasma 90/10 (Fischione) for one minute. Four microliters of sample was applied to a grid (Quantifoil R 2/1 + 2 nm carbon, Quantifoil), incubated for 30 s, blotted 4 s with blot force ‘8’, at 100% humidity and 4 °C in a Vitrobot Mark IV (FEI) and plunged in liquid ethane.

Images were collected on a Cs-corrected Titan Krios microscope (FEI), operated at 300 kV using the multi-shot feature of the serialEM software<sup>77</sup> for automated data collection. Movie frames were acquired on a 4k × 4k Gatan K2 summit direct electron detector in super-resolution mode at a nominal magnification of 105,000, which yielded a pixel size of 0.545 Å. Forty movie frames were recorded at a dose of 1.4 electrons per Å<sup>2</sup> per frame corresponding to a total dose of 56 e/Å<sup>2</sup>.

**Image processing.** Movie frames were aligned, dose-weighted, binned by a factor of 2 and averaged using MotionCorr<sup>78</sup>. Contrast Transfer Function (CTF) parameters were estimated with the Gctf<sup>79</sup> program. The RELION 3-beta suite<sup>33</sup> was used for the whole-image processing workflow unless stated otherwise. The dataset was divided into four subsets with ~1000 images each. In a first step the reference-free auto-picking procedure based on a Laplacian-of-Gaussian (LoG) filter was used to identify ~100,000 starting coordinates (per subset), which were used to extract particles with threefold binning in a 140 pixel box and the particles were grouped by reference-free 2D classification. Classes with contamination and damaged particles were discarded and the remaining particles were aligned on a reference generated from the PDB entry 5G5L low-pass filtered to 40 Å. Three-dimensional (3D) classes containing only Pol I and Rrn3, or damaged particles were discarded. The remaining 227,718 particles from the four subsets were merged, re-extracted without binning and refined against an initial model generated in RELION. CTF Refinement and Bayesian polishing was performed and the polished particles were refined and 2D and 3D classification without alignment were performed to remove misaligned particles and the remaining 168,532 particles were subjected to a second round of CTF refinement. A 3D classification without sampling and a CF-only mask revealed one class with partial CF occupancy and another with damaged particles that were both discarded. Refinement of the remaining 122,099 particles resulted in an early intermediate PIC reconstruction. For details, compare Supplementary Fig. 1. During post-processing in RELION, a B-factor of –75 Å<sup>2</sup> was determined and applied for map sharpening, resulting in an overall resolution of 3.5 Å. Focused refinements with a Pol-I-Rrn3 mask (3.5 Å after post-processing) or a CF-DNA mask (3.9 Å after post-processing) were additionally carried out to assist subdomain conformation determination and aid CF chain tracing, respectively. Directional FSC were calculated as described<sup>80</sup>.

**Model building.** At a resolution of 3.5 Å, we derive an atomic model of an early intermediate PIC. We first placed Pol I domains as described for PDB 5G5L<sup>12</sup> originating from the crystal structure (PDB 4C2M<sup>30</sup>), an Rrn3 monomer (PDB 3TJ1<sup>32</sup>), a CF monomer (PDB 5O7X<sup>15</sup>) and an ITC DNA (PDB 5W66<sup>17</sup>) in the unsharpened ePIC map generated with RELION 3 (beta version)<sup>33</sup>. Using COOT<sup>81</sup>, we adjusted protein backbone traces consulting focused maps of CF or the Pol-I-Rrn3 complex and finally build side chain residues where appropriate. DNA-sequences were mutated to poly-A (–T, –G, –C). For the structure-based modeling of the TFIIB-related domains in the N-terminal region of Rrn7, the strong density for aromatic residue Phe70 was used as a marker. The final model



was refined using the real-space refinement tool of the Phenix suite<sup>82</sup> and evaluated using MolProbity<sup>83</sup>. Figures were prepared with UCSF Chimera<sup>84</sup> or PyMOL (pymol.org).

It should be noted that promoter-binding regions within CF are highly flexible and thus poorly ordered in DNA-free CF crystals<sup>15</sup>. While we refrained from building most of these regions in the crystal structure, the putatively assigned residue numbers within helix  $\alpha 2$  of CF subunit Rrn11 were now adjusted in the eiPIC, similar to a de novo built model based on a cryo-EM reconstruction of an ITC<sup>17</sup>.

An additional cryo-EM density stretch between the Rrn7 ribbon and the Pol I wall domain may potentially be attributed to a flexible loop in Rrn3 (249–323) or to a part of the Rrn6 C-terminal domain. Whereas the latter assignment would agree with a previously published crosslinking/mass spectrometry analysis<sup>16</sup> and direct Rrn6–Rrn3 interaction studies<sup>85</sup>, it remains as speculative at this point.

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

## Data availability

The cryo-EM density of the Pol I eiPIC has been deposited in the Electron Microscopy Data Bank under accession code EMD-10544 and coordinates of the eiPIC model have been deposited with the Protein Data Bank under accession code 6TPS. Focused refinement density of CF has been deposited under EMD-10663. Other data are available from the corresponding author upon reasonable request.

Received: 13 May 2019; Accepted: 15 February 2020;

Published online: 05 March 2020

## References

- Klinge, S. & Woolford, J. L. Ribosome assembly coming into focus. *Nat. Rev. Mol. Cell Biol.* **20**, 116–131 (2019).
- Milkereit, P. & Tschochner, H. A specialized form of RNA polymerase I, essential for initiation and growth-dependent regulation of rRNA synthesis, is disrupted during transcription. *EMBO J.* **17**, 3692–3703 (1998).
- Bodem, J. et al. TIF-IA, the factor mediating growth-dependent control of ribosomal RNA synthesis, is the mammalian homolog of yeast Rrn3p. *EMBO Rep.* **1**, 171–175 (2000).
- Mayer, C., Zhao, J., Yuan, X. & Grummt, I. mTOR-dependent activation of the transcription factor TIF-IA links rRNA synthesis to nutrient availability. *Genes Dev.* **18**, 423–434 (2004).
- Fath, S., Kobor, M. S., Philippi, A., Greenblatt, J. & Tschochner, H. Dephosphorylation of RNA polymerase I by Fcp1p is required for efficient rRNA synthesis. *J. Biol. Chem.* **279**, 25251–25259 (2004).
- Gerber, J. et al. Site specific phosphorylation of yeast RNA polymerase I. *Nucleic Acids Res.* **36**, 793–802 (2008).
- Jochem, L., Ramsay, E. P. & Vannini, A. RNA polymerase I, bending the rules? *EMBO J.* **36**, 2664–2666 (2017).
- Khatler, H., Vorländer, M. K. & Müller, C. W. RNA polymerase I and III: similar yet unique. *Curr. Opin. Struct. Biol.* **47**, 88–94 (2017).
- Engel, C., Neyer, S. & Cramer, P. Distinct Mechanisms of transcription initiation by RNA polymerases I and II. *Annu. Rev. Biophys.* **47**, 425–446 (2018).
- Moss, T., Langlois, F., Gagnon-Kugler, T. & Stefanovsky, V. A housekeeper with power of attorney: the rRNA genes in ribosome biogenesis. *Cell. Mol. Life Sci.: CMLS* **64**, 29–49 (2007).
- Keener, J., Josaitis, C. A., Dodd, J. A. & Nomura, M. Reconstitution of yeast RNA polymerase I transcription in vitro from purified components. TATA-binding protein is not required for basal transcription. *J. Biol. Chem.* **273**, 33795–33802 (1998).
- Engel, C., Plitzko, J. & Cramer, P. RNA polymerase I–Rrn3 complex at 4.8 Å resolution. *Nat. Commun.* **7**, 12129 (2016).
- Pils, M. et al. Structure of the initiation-competent RNA polymerase I and its implication for transcription. *Nat. Commun.* **7**, 12126 (2016).
- Torreira, E. et al. The dynamic assembly of distinct RNA polymerase I complexes modulates rDNA transcription. *eLife* **6**; <https://doi.org/10.7554/eLife.20832> (2017).
- Engel, C. et al. Structural basis of RNA polymerase I transcription initiation. *Cell* **169**, 120–131.e22 (2017).
- Sadian, Y. et al. Structural insights into transcription initiation by yeast RNA polymerase I. *EMBO J.* **36**, 2698–2709 (2017).
- Han, Y. et al. Structural mechanism of ATP-independent transcription initiation by RNA polymerase I. *eLife* **6**; <https://doi.org/10.7554/eLife.27414> (2017).
- Steffan, J. S., Keys, D. A., Dodd, J. A. & Nomura, M. The role of TBP in rDNA transcription by RNA polymerase I in *Saccharomyces cerevisiae*: TBP is required for upstream activation factor-dependent recruitment of core factor. *Genes Dev.* **10**, 2551–2563 (1996).
- Steffan, J. S., Keys, D. A., Vu, L. & Nomura, M. Interaction of TATA-binding protein with upstream activation factor is required for activated transcription of ribosomal DNA by RNA polymerase I in *Saccharomyces cerevisiae* in vivo. *Mol. Cell Biol.* **18**, 3752–3761 (1998).
- Keys, D. A. et al. Multiprotein transcription factor UAF interacts with the upstream element of the yeast RNA polymerase I promoter and forms a stable preinitiation complex. *Genes Dev.* **10**, 887–903 (1996).
- Bedwell, G. J., Appling, F. D., Anderson, S. J. & Schneider, D. A. Efficient transcription by RNA polymerase I using recombinant core factor. *Gene* **492**, 94–99; <https://doi.org/10.1016/j.gene.2011.10.049> (2012).
- Hannig, K. et al. The C-terminal region of Net1 is an activator of RNA polymerase I transcription with conserved features from yeast to human. *PLoS Genet.* **15**, e1008006 (2019).
- Shou, W. et al. Net1 stimulates RNA polymerase I transcription and regulates nucleolar structure independently of controlling mitotic exit. *Mol. Cell* **8**, 45–55 (2001).
- Sainsbury, S., Bernecky, C. & Cramer, P. Structural basis of transcription initiation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* **16**, 129–143 (2015).
- Vannini, A. & Cramer, P. Conservation between the RNA polymerase I, II, and III transcription initiation machineries. *Mol. Cell* **45**, 439–446 (2012).
- 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 (N. Y., N. Y.)* **292**, 1876–1882 (2001).
- 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).
- Huang, J. & Moazed, D. Association of the RENT complex with nontranscribed and coding regions of rDNA and a regional requirement for the replication fork block protein Fob1 in rDNA silencing. *Genes Dev.* **17**, 2162–2176; <https://doi.org/10.1101/gad.1108403> (2003).
- Straight, A. F. et al. Net1, a Sir2-associated nucleolar protein required for rDNA silencing and nucleolar integrity. *Cell* **97**, 245–256 (1999).
- Engel, C., Sainsbury, S., Cheung, A. C., Kostrewa, D. & Cramer, P. RNA polymerase I structure and transcription regulation. *Nature* **502**, 650–655 (2013).
- Fernández-Tornero, C. et al. Crystal structure of the 14-subunit RNA polymerase I. *Nature* **502**, 644–649 (2013).
- Blattner, C. et al. Molecular basis of Rrn3-regulated RNA polymerase I initiation and cell growth. *Genes Dev.* **25**, 2093–2105 (2011).
- Zivanov, J. et al. RELION-3: new tools for automated high-resolution cryo-EM structure determination. *bioRxiv* <https://doi.org/10.1101/421123> (2018).
- Plaschka, C. et al. Transcription initiation complex structures elucidate DNA opening. *Nature* **533**, 353–358 (2016).
- Bernecky, C., Herzog, F., Baumeister, W., Plitzko, J. M. & Cramer, P. Structure of transcribing mammalian RNA polymerase II. *Nature* **529**, 551–554 (2016).
- Kujirai, T. et al. Structural basis of the nucleosome transition during RNA polymerase II passage. *Science (N. Y., N. Y.)* **362**, 595–598 (2018).
- 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).
- Herdman, C. et al. A unique enhancer boundary complex on the mouse ribosomal RNA genes persists after loss of Rrn3 or UBF and the inactivation of RNA polymerase I transcription. *PLoS Genet.* **13**, e1006899 (2017).
- Schnapp, A., Schnapp, G., Erny, B. & Grummt, I. Function of the growth-regulated transcription initiation factor TIF-IA in initiation complex formation at the murine ribosomal gene promoter. *Mol. Cell Biol.* **13**, 6723–6732 (1993).
- Kostrewa, D., Kuhn, C. -D., Engel, C. & Cramer, P. An alternative RNA polymerase I structure reveals a dimer hinge. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **71**, 1850–1855 (2015).
- Neyer, S. et al. Structure of RNA polymerase I transcribing ribosomal DNA genes. *Nature* <https://doi.org/10.1038/nature20561> (2016).
- Tafur, L. et al. Molecular structures of transcribing RNA polymerase I. *Mol. Cell* **64**, 1135–1143 (2016).
- 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**; <https://doi.org/10.7554/eLife.43204> (2019).
- Knutson, B. A. & Hahn, S. Yeast Rrn7 and human TAF1B are TFIIB-related RNA polymerase I general transcription factors. *Science (N. Y., N. Y.)* **333**, 1637–1640 (2011).
- Knutson, B. A., Luo, J., Ranish, J. & Hahn, S. Architecture of the *Saccharomyces cerevisiae* RNA polymerase I core factor complex. *Nat. Struct. Mol. Biol.* **21**, 810–816 (2014).

46. Naidu, S., Friedrich, J. K., Russell, J. & Zomerdijk, J. C. B. M. TAF1B is a TFIIB-like component of the basal transcription machinery for RNA polymerase I. *Science (N. Y., N. Y.)* **333**, 1640–1642 (2011).
47. Nikolov, D. B. et al. Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature* **377**, 119–128 (1995).
48. Vorländer, M. K., Khatter, H., Wetzel, R., Hagen, W. J. H. & Müller, C. W. Molecular mechanism of promoter opening by RNA polymerase III. *Nature* **553**, 295–300 (2018).
49. Abascal-Palacios, G., Ramsay, E. P., Beuron, F., Morris, E. & Vannini, A. Structural basis of RNA polymerase III transcription initiation. *Nature* **553**, 301–306 (2018).
50. Cramer, P., Bushnell, D. A. & Kornberg, R. D. Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science (N. Y., N. Y.)* **292**, 1863–1876 (2001).
51. Hoffmann, N. A. et al. Molecular structures of unbound and transcribing RNA polymerase III. *Nature* **528**, 231–236 (2015).
52. Dienemann, C., Schwab, B., Schilbach, S. & Cramer, P. Promoter distortion and opening in the RNA polymerase II cleft. *Mol. Cell* **73**, 97–106.e4 (2019).
53. He, Y. et al. Near-atomic resolution visualization of human transcription promoter opening. *Nature* **533**, 359–365 (2016).
54. Gouge, J. et al. Redox signaling by the RNA polymerase III TFIIB-related factor Brf2. *Cell* **163**, 1375–1387 (2015).
55. Han, Y., Yan, C., Fishbain, S., Ivanov, I. & He, Y. Structural visualization of RNA polymerase III transcription machineries. *Cell Discov.* **4**, 40 (2018).
56. Kostrewa, D. et al. *Crystal Structure of RNA Polymerase II in Complex with TFIIB* (RCSB PDB, 2009).
57. Sainsbury, S., Niesser, J. & Cramer, P. Structure and function of the initially transcribing RNA polymerase II-TFIIB complex. *Nature* **493**, 437–440 (2013).
58. Chen, B. -S. & Hampsey, M. 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).
59. Armache, K. -J., Mitterweger, S., Meinhart, A. & Cramer, P. Structures of complete RNA polymerase II and its subcomplex, Rpb4/7. *J. Biol. Chem.* **280**, 7131–7134 (2005).
60. Cheung, A. C. M. & Cramer, P. A movie of RNA polymerase II transcription. *Cell* **149**, 1431–1437 (2012).
61. Fernández-Tornero, C. RNA polymerase I activation and hibernation: unique mechanisms for unique genes. *Transcription* **9**, 248–254 (2018).
62. Ganley, A. R. D., Hayashi, K., Horiuchi, T. & Kobayashi, T. Identifying gene-independent noncoding functional elements in the yeast ribosomal DNA by phylogenetic footprinting. *Proc. Natl Acad. Sci. USA* **102**, 11787–11792 (2005).
63. Bric, A., Radebaugh, C. A. & Paule, M. R. Photocross-linking of the RNA polymerase I preinitiation and immediate postinitiation complexes: implications for promoter recruitment. *J. Biol. Chem.* **279**, 31259–31267 (2004).
64. Kownin, P., Bateman, E. & Paule, M. R. Eukaryotic RNA polymerase I promoter binding is directed by protein contacts with transcription initiation factor and is DNA sequence-independent. *Cell* **50**, 693–699 (1987).
65. Jackobel, A. J., Zeberl, B. J., Glover, D. M., Fakhouri, A. M. & Knutson, B. A. DNA binding preferences of *S. cerevisiae* RNA polymerase I Core Factor reveal a preference for the GC-minor groove and a conserved binding mechanism. *Biochim. Biophys. Acta Gene Regul. Mech.* **1862**, 194408 (2019).
66. Boyaci, H., Chen, J., Jansen, R., Darst, S. A. & Campbell, E. A. Structures of an RNA polymerase promoter melting intermediate elucidate DNA unwinding. *Nature* **565**, 382–385 (2019).
67. Feklistov, A. et al. RNA polymerase motions during promoter melting. *Science (N. Y., N. Y.)* **356**, 863–866 (2017).
68. Schilbach, S. et al. Structures of transcription pre-initiation complex with TFIIB and Mediator. *Nature* **551**, 204–209 (2017).
69. Geiger, S. R. et al. RNA polymerase I contains a TFIIB-related DNA-binding subcomplex. *Mol. Cell* **39**, 583–594 (2010).
70. Russell, J. & Zomerdijk, J. C. B. M. The RNA polymerase I transcription machinery. *Biochem. Soc. Symp.* **73**, 203–216 (2006).
71. Moorefield, B., Greene, E. A. & Reeder, R. H. RNA polymerase I transcription factor Rrn3 is functionally conserved between yeast and human. *Proc. Natl Acad. Sci. USA* **97**, 4724–4729 (2000).
72. Sadian, Y. et al. Molecular insight into RNA polymerase I promoter recognition and promoter melting. *Nat. Commun.* **10**, 5543 (2019).
73. Engel, C. Purification of crystallization-grade RNA polymerase I from *S. cerevisiae*. *Methods Mol. Biol. (Clifton, N. J.)* **1455**, 85–97 (2016).
74. Smith, M. L., Cui, W., Jackobel, A. J., Walker-Kopp, N. & Knutson, B. A. Reconstitution of RNA polymerase I upstream activating factor and the roles of histones H3 and H4 in complex assembly. *J. Mol. Biol.* **430**, 641–654 (2018).
75. Berger, I., Fitzgerald, D. J. & Richmond, T. J. Baculovirus expression system for heterologous multiprotein complexes. *Nat. Biotechnol.* **22**, 1583–1587 (2004).
76. Fitzgerald, D. J. et al. Protein complex expression by using multigene baculoviral vectors. *Nat. Methods* **3**, 1021–1032 (2006).
77. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).
78. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
79. Zhang, K. Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
80. Tan, Y. Z. et al. Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat. Methods* **14**, 793–796 (2017).
81. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **66**, 486–501 (2010).
82. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **66**, 213–221 (2010).
83. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **66**, 12–21 (2010).
84. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
85. Peyroche, G. et al. The recruitment of RNA polymerase I on rDNA is mediated by the interaction of the A43 subunit with Rrn3. *EMBO J.* **19**, 5473–5482 (2000).

## Acknowledgements

We thank Herbert Tschochner, Joachim Griesenbeck, and Philipp Milkereit for advice and critical discussion, Patrick Schultz and Corinne Crucifix for help and equipment access, and Lena Heuschneider, Gerhard Lehmann, Norbert Eichner and Gisela Pöll for technical assistance. Thanks to Guillermo Abascal-Palacios and Alessandro Vannini for sharing their by-product polymerase I. Data were collected at the IGBMC cryo-EM facility (Strasbourg, France). M.P. and C.E. were supported by Deutsche Forschungsgemeinschaft SFB 960 and the Emmy-Noether Programm (DFG grant no. EN 1204/1-1 to C.E.).

## Author contributions

M.P. carried out experiments and data analysis. C.E. designed and supervised research. M.P. and C.E. carried out model building and prepared the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information


**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41467-020-15052-y>.

**Correspondence** and requests for materials should be addressed to C.E.

**Peer review information** *Nature Communications* thanks Dong Wang and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permission 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 4.0 International License, which permits use, sharing, adaptation, 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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020