ARTICLE

doi:10.1038/nature12636

Crystal structure of the 14-subunit RNA polymerase I

Carlos Fernández-Tornero[1]*, María Moreno-Morcillo[2]*, Umar J. Rashid[2], Nicholas M. I. Taylor[1], Federico M. Ruiz[1], Tim Gruene[3], Pierre Legrand[4], Ulrich Steuerwald[2]† & Christoph W. Müller[2]

Protein biosynthesis depends on the availability of ribosomes, which in turn relies on ribosomal RNA production. In eukaryotes, this process is carried out by RNA polymerase I (Pol I), a 14-subunit enzyme, the activity of which is a major determinant of cell growth. Here we present the crystal structure of Pol I from *Saccharomyces cerevisiae* at 3.0 Å resolution. The Pol I structure shows a compact core with a wide DNA-binding cleft and a tightly anchored stalk. An extended loop mimics the DNA backbone in the cleft and may be involved in regulating Pol I transcription. Subunit A12.2 extends from the A190 jaw to the active site and inserts a transcription elongation factor TFIIS-like zinc ribbon into the nucleotide triphosphate entry pore, providing insight into the role of A12.2 in RNA cleavage and Pol I insensitivity to α-amanitin. The A49–A34.5 heterodimer embraces subunit A135 through extended arms, thereby contacting and potentially regulating subunit A12.2.

Ribosome biosynthesis is a central cellular process that in eukaryotes requires the coordinated action of three nuclear, DNA-dependent RNA polymerases. Pol I transcribes the ribosomal RNA (rRNA) precursor gene present in several copies in the nucleolus that, in yeast, is subsequently processed into 25S, 18S and 5.8S rRNAs. Pol II, synthesizing messenger RNA, and Pol III, mainly involved in transfer RNA synthesis, contribute to ribosome biogenesis by providing ribosomal protein mRNAs and 5S rRNA, respectively. As Pol I is the most active eukaryotic RNA polymerase contributing up to 60% of the total transcriptional activity[1], rRNA synthesis is a crucial point in cell regulation. Thus, alterations in cell proliferation are accompanied by changes in the synthesis rate of rRNA, and mis-regulation of Pol I has been associated with different types of cancer[2–4].

Despite conservation of the overall structure and function, differences between Pol I, Pol II and Pol III probably determine their specific activities[5]. The yeast Pol I enzyme has a total mass of 589 kilodalton (kDa) and consists of 14 subunits[6], whereas Pol II and Pol III contain 12 and 17 subunits, respectively[5]. All three RNA polymerases share a horseshoe-shaped core built of ten subunits. The Pol I core includes the two largest subunits, A190 and A135, forming the DNA binding cleft, plus five subunits present in all nuclear RNA polymerases (Rpb5, Rpb6, Rpb8, Rpb10 and Rpb12) and the AC40–AC19 heterodimer, shared with Pol III and homologous to Rpb3–Rpb11 in Pol II. The core is completed by A12.2, a subunit formed by two zinc ribbons that participates in RNA cleavage[7]. Whereas in Pol III subunit C11 performs a similar function, in Pol II the corresponding amino- and carboxy-terminal zinc ribbons are contributed by subunit Rpb9 and the elongation factor TFIIS, respectively[8]. Outside the core, the A43–A14 heterodimer, related to Rpb7–Rpb4 in Pol II and C25–C17 in Pol III, forms the stalk that provides a platform for initiation factors and interacts with newly synthesized RNA[7]. An additional peripheral heterodimer is formed by subunits A49 and A34.5, the dimerization module of which is related to similar modules in the Pol III C37–C53 heterodimer[9] and in Pol-II-specific TFIIF[10]. Additional homology has been detected between the tandem winged-helix (tWH) domain in A49 and the Pol III subunit C34 or Pol-II-specific TFIIE[9,10].

So far, atomic structures are only available for the yeast Pol II enzyme (reviewed in ref. 11). In contrast, only electron microscopy (EM) reconstructions at about 1-nm resolution have been reported for Pol I (ref. 7) and Pol III (ref. 12). The overall architecture of Pol I was initially revealed by EM studies[13,14], whereas crosslinking coupled to mass spectrometry suggested the approximate locations of A12.2 and A49–A34.5 (ref. 15). In addition, crystal structures of the A43–A14 stalk[7], the A49–A34.5 dimerization module and the A49 tWH domain[10] contributed to a better understanding of Pol I. To shed light on the molecular mechanism of Pol I transcription, we have determined the atomic structure of the complete yeast Pol I enzyme, providing unprecedented insight into its overall architecture and possible functional roles of its components.

Structure determination of yeast Pol I

The crystallographic analysis of yeast Pol I relied on a large-scale purification protocol, initially established for Pol III (refs 12, 16), that yields large amounts of pure, homogeneous and transcriptionally active Pol I. Three related crystal forms varying in cell dimensions were obtained in space group *C*2 and diffracted beyond 3.5 Å resolution (Extended Data Table 1). Initial molecular replacement phases using the Pol II core[17] as search model allowed the location of heavy atom positions in a Ta_6Br_{12} derivative that, after SIRAS phasing, yielded a map where the core and stalk could be built. Phasing to higher resolution using a Yb derivative, combined with the partial model, produced a map that allowed the completion of the model (Extended Data Fig. 1a–d). The correct chain tracing was confirmed by a selenomethionine derivative (Extended Data Fig. 1e, f and Extended Data Table 1). Refinement against the best diffracting crystal form, with a resolution of 3.0 Å according to $CC_{1/2} > 0.3$ criterion[18] (3.27 Å resolution according to the $I/\sigma I > 2.0$ criterion), yielded a model with good statistics and excellent geometry (Extended Data Table 1). The C-terminal A49 tWH

[1]Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, Ramiro de Maeztu 9, 28040 Madrid, Spain. [2]European Molecular Biology Laboratory, Structural and Computational Biology Unit, Meyerhofstrasse 1, 69117 Heidelberg, Germany. [3]Department of Structural Chemistry, Georg-August-University, Tammannstraße 4, 37077 Göttingen, Germany. [4]SOLEIL Synchrotron, L'Orme des Merisiers Saint Aubin, 91192 Gif-sur-Yvette, France. †Present address: Department of Cellular Biochemistry, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, D-37077 Göttingen, Germany.
*These authors contributed equally to this work.

domain and several other regions are disordered and could not be modelled.

Pol I has a wide cleft and a closed clamp

The crystal structure of Pol I reveals several striking new features compared to other RNA polymerases (Fig. 1a and Extended Data Fig. 2). Although the overall shape of the core is conserved, the Pol I DNA-binding cleft adopts the widest conformation ever reported. In comparison to other RNA polymerases the Pol I cleft is about 10 Å more open (Extended Data Fig. 3). Moreover, superposition of the Pol I structures derived from the three crystal forms allows dividing the enzyme into two modules of similar mass that independently move around five hinges to widen the cleft (Fig. 1b and Extended Data Fig. 3a–c). Interestingly, in all crystal forms the wide Pol I cleft is concomitant with a closed conformation of the clamp, an element that in Pol II has been described to rotate to open and close the DNA-binding cleft[17](Extended Data Fig. 3b, c). An open clamp has been observed in the Pol II core structure lacking the stalk subunits[17], whereas a closed clamp is present in the complete Pol II enzyme[19,20] and when bound to a transcription bubble[21,22]. In Pol I, the closed conformation of the clamp is probably due to the presence of unique helices in this element that tighten the interaction with subunit Rpb5 (Fig. 1b) and is consistent with the tight association of the Pol I core with the stalk subunits (see below). Accordingly, we speculate that the Pol I clamp always adopts a closed conformation, which might explain the high processivity of this enzyme able to transcribe the long rRNA precursor gene without dissociating from the DNA.

Concurrent with cleft widening, two critical structural elements in Pol I adopt unique conformations. First, the entrance gate to the RNA exit channel is partly occluded by the A190 lid loop (Fig. 1b), whereas the equivalent loop in Pol II and archaeal Pol establishes connections with the wall element above the RNA exit gate[17,23]. Second, the Pol I bridge helix unwinds at its middle region by a complete helical turn (A190 residues 1012–1016) thereby introducing an additional ~10° kink into the bridge helix compared to other RNA polymerases (Fig. 1b and Extended Data Fig. 3d, e). Partial unfolding of this helix has been proposed to occur during translocation in Pol II (ref. 17) but has only been observed in the bacterial Pol[24,25], where in the bacterial

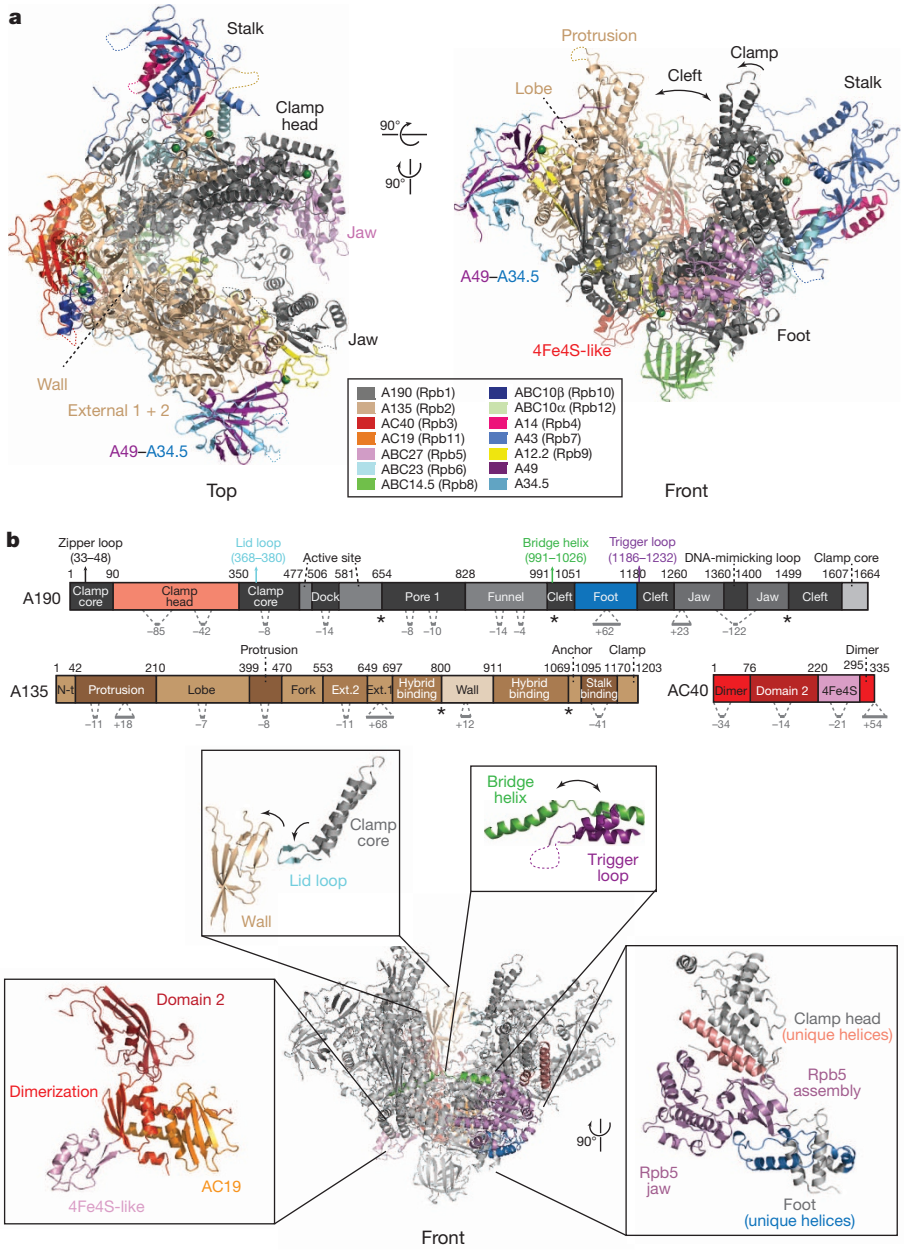


Figure 1 | Crystal structure of yeast RNA polymerase I. a, Two orthogonal views of the model with different domains and elements labelled. The colour code for the 14 subunits is given in the box with corresponding Pol II subunits in parentheses. Zn^{2+} ions are shown as green spheres. Arrows indicate the concomitant presence of a wide cleft and a closed clamp. **b**, Close-up views of specific regions in the Pol I core. Colouring of domains according to the bar diagrams. Insertions and deletions in Pol II with respect to Pol I are indicated below the bar diagrams. Black asterisks indicate 'hinges'.

holoenzyme two residues are flipped out of the bridge helix[25]. Interestingly, these two residues fall into the region that unwinds in Pol I, whereas mutating the corresponding residues in archaeal Pol in some cases increases its transcriptional activity[26], presumably because these mutations destabilize the bridge helix (Extended Data Fig. 3e). In accordance with previous systematic mutagenesis studies[26,27] and molecular dynamics simulations (reviewed in ref. 28), our results also suggest considerable flexibility of the bridge helix that in the case of Pol I even partially unwinds. Different from the bacterial Pol holoenzyme[25], the Pol I trigger loop immediately adjacent to the bridge helix is disordered as in the different Pol II structures[17,21,22] (Extended Data Fig. 3d).

To understand the functional implications of cleft widening better, we have modelled a Pol I elongation complex using the corresponding Pol II structure[22]. The resulting model (Extended Data Fig. 4a) shows that interactions of subunit A190 with downstream DNA and the beginning of the DNA–RNA duplex are maintained. However, most interactions of subunit A135 with nucleic acids in the vicinity of the active site, critical for NTP selection and translocation, and with the DNA–RNA duplex, can only be established if the Pol I modules move towards each other to close the cleft. This movement presumably involves complete folding of the bridge helix, opening of the RNA exit gate, and the approach of several A135 elements to anchor the transcription bubble (Extended Data Fig. 4b).

Subunits AC40 and AC19 form a heterodimer whose counterparts, Rpb3–Rpb11 in Pol II and archaeal Rpo3–Rpo11, have been described to constitute a platform for the assembly of the enzyme core[5]. AC40 comprises three domains (dimerization, domain 2 and 4Fe4S-like) preceded by a disordered N-terminal tail (Fig. 1b). Although the fold of the first two domains is conserved in other RNA polymerases, AC40 contains a 4Fe4S-like domain that is absent in Pol II but resembles the 4Fe4S cluster domain in Rpo3. However, in contrast to its archaeal homologue, AC40 lacks the four cysteine residues that coordinate the 4Fe4S cluster, further supporting a structural rather than a catalytic role of this domain[23].

An extended loop mimics DNA in the cleft

A notable feature in the Pol I structure is the presence of an extended loop inside the DNA-binding cleft (Fig. 2). We have assigned this loop, which is well ordered in two of the three crystal forms (Extended Data Fig. 5a), to a stretch of ~60 residues in the A190 jaw domain (residues 1340–1400). Although not conserved in Pol II and Pol III, this loop contains several acidic amino acid residues that are conserved within Pol I enzymes of different species (Extended Data Fig. 5b). Numerous hydrogen bonds and electrostatic interactions are formed between negatively charged residues of the loop and positively charged residues of the cleft. Of note, a patch of acidic residues in this loop (Asp 1385, Glu 1386, Asp 1388, Glu 1389) interacts with Arg 1015 in the unfolded region of the bridge helix. Ordering of the loop also coincides with the stabilization of the C-terminal Zn ribbon of subunit A12.2 (see below) that appears less flexible in the two crystal forms, suggesting that these unique Pol I regions stabilize each other. Although Pol II also contains an insertion and acidic loops in the Rpb1 jaw, they are too short to have a similar role.

The A190 jaw loop contains three short helices and extends from the inner face of the A135 lobe domain to the internal surface of the clamp domain on the opposite side of the cleft. As a consequence, it roughly overlaps with the expected position of the DNA backbone in the transcription bubble when the structure of elongating Pol II[21] is superimposed with Pol I (Fig. 2). We will therefore refer to it as the 'DNA-mimicking loop'. Because its position in the cleft is incompatible with elongation, ordering can only occur when the enzyme is not actively transcribing. The fact that this loop is partially disordered in one crystal form and that our Pol I preparation is transcriptionally active (Extended Data Fig. 5c) also suggests that it is a mobile element that only becomes ordered under certain conditions and presumably carries out a regulatory function. For example, it may prevent Pol I from nonspecifically binding nucleic acids or displace the DNA backbone during transcription termination. The DNA-mimicking loop is not essential. A yeast strain lacking most of the loop (residues 1361–1390) is viable but shows a subtle growth defect at 37 °C, but not at 30 °C, compared to the parental strain (see Methods and Extended Data Fig. 5d).

The stalk is tightly anchored to the core

The overall structure of the A43–A14 heterodimer bound to the enzyme core is similar to that determined for the isolated heterodimer[7]. In contrast to Pol II and archaeal Pol, Pol I forms a large interface between the core and the stalk, mainly involving the Pol-I-specific N-terminal tail of A43 (Fig. 3a). This tail forms a hydrophilic helix that provides a strong anchoring point for a Pol-I-specific insertion present at the C-terminal region of A135 (residues 1103–1075) termed stalk-binding domain. This new domain not only wraps around the A43 N-terminal helix in a double-hook fashion, but also contacts residues of the A43 core. In addition, and observed for the first time,

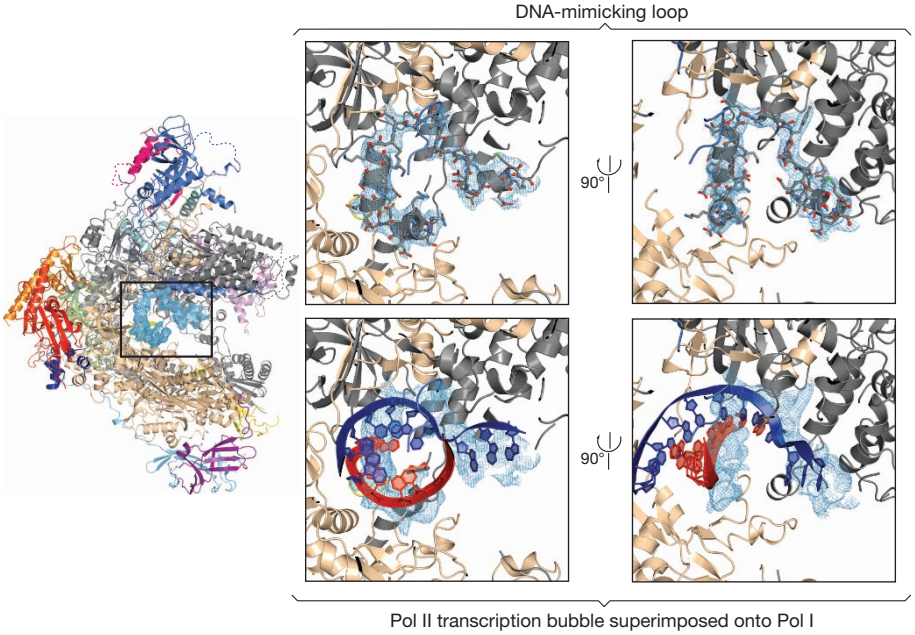


Figure 2 | An extended loop mimics the transcription bubble inside the DNA-binding cleft. Left panel: overall model of Pol I in crystal form C2-90. The σ_A-weighted electron density (contoured at 1σ) corresponds to the extended loop in the DNA-binding cleft. Right panels (top row): two orthogonal orientations of the DNA-mimicking loop in subunit A190 (residues 1361–1399). Right panels (bottom row): superposition of the Pol II elongation complex (Protein Data Bank accession 1I6H) onto the Pol I model. Only the DNA/RNA hybrid is shown with DNA and RNA depicted in blue and red, respectively.

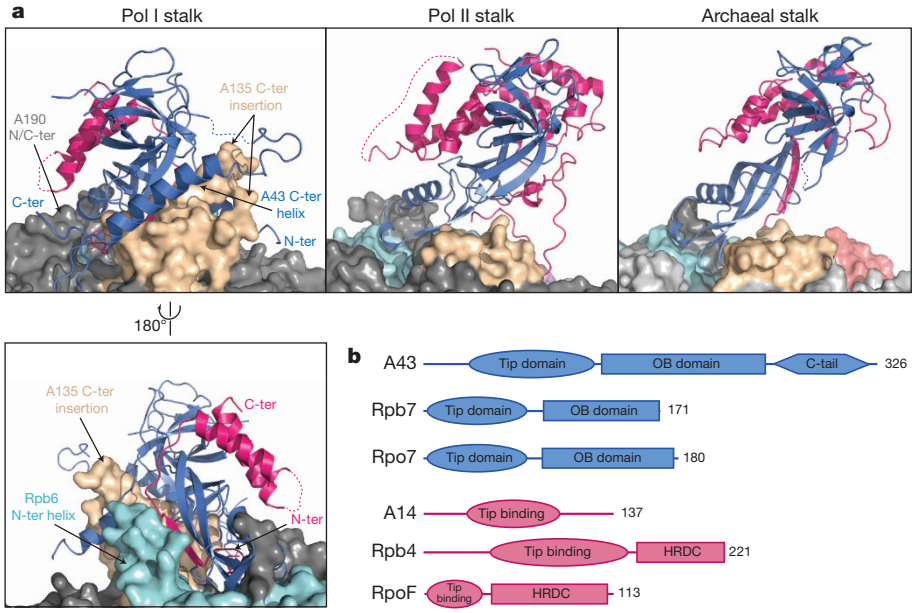


Figure 3 | Stalk anchoring in Pol I, Pol II and archaeal Pol. a, Ribbon representations of the Pol I, Pol II and archaeal Pol stalks show differences in their relative orientation with respect to the core (depicted as surface). Main additional connections between the Pol I stalk and the core are labelled. **b**, Schematic representation of the domains found in the stalks of all three polymerases, A43–A14 in Pol I, Rpb7–Rpb4 in Pol II and Rpo7–Rpo4 in archaeal Pol. C-ter, C terminal; HRDC, helicase RNaseD C-terminal domain; N-ter, N terminal; OB, oligonucleotide binding-fold domain.

the N-terminal region of Rpb6 (residues 55–71) adopts a helical conformation that provides additional contacts with A43 and A14. Finally, the A190 N and C termini also contribute to the formation of a well-defined pocket that tightly fixes the stalk. Indeed, in contrast to the dissociable Pol II stalk that can be readily removed from the core under *in vitro* and *in vivo* conditions[29,30], the Pol I stalk cannot dissociate even in the presence of moderate urea concentrations[31]. This also argues in favour of a permanently closed conformation of the Pol I clamp, as in Pol II the open clamp conformation has only been observed in the absence of the stalk[17]. The tight binding of the Pol I stalk produces a ~10° inclination towards the clamp when compared to that of Pol II, whereas the archaeal stalk bends in the opposite direction (Fig. 3). This unique bending may also be critical for the interaction with Pol-I-specific initiation factor Rrn3. The conditional rpa43 mutant (K63E, C118R and Q140R) disrupts the Pol I–Rrn3 interaction[32]. Whereas residues Lys 63 and Gln 140 are exposed to the solvent and may establish direct contacts with Rrn3 that are compatible with the proposed model of the Pol I–Rrn3 complex[33], residue Cys 118 points to the interior of the tip domain and its mutation rather affects the stalk stability.

In all three crystal forms, the stalk subunit A43 of a neighbouring molecule inserts between the clamp and protrusion domains, establishing contacts with both (Extended Data Fig. 6). The most extensive interaction involves the A43 C-terminal tail (residues 251–326), which folds into a long α-helix followed by a β-hairpin and runs between the A190 lid and zipper loops. These interactions establish a dimer interface around the crystallographic dyad and might contribute to widening the Pol I cleft. Pol I used for this study was monomeric in solution as shown by gel filtration and ultracentrifugation (data not shown). Dimerization of Pol I has been observed by cryo-EM[13] and in solution under certain conditions[34], although its functional relevance is still unclear.

A12.2 reaches the active site

Subunit A12.2 comprises two Zn-binding β-ribbon domains connected by a flexible linker. The N-terminal ribbon is located between the A190 jaw, the A135 lobe and the A49–A34.5 dimerization module, establishing contacts with all three (Fig. 4). Its position is equivalent to that of the Rpb9 N-terminal ribbon in Pol II as previously suggested[8], but it is slightly shifted and rotated, probably as a consequence of the presence of A49–A34.5. In marked contrast, the A12.2 C-terminal Zn ribbon is located inside the NTP entry pore, as shown in our electron density maps and validated by zinc anomalous difference maps. This position almost perfectly overlaps with that of the TFIIS C-terminal Zn ribbon in Pol II[35] and agrees with experiments showing that A12.2 is required for the intrinsic RNA cleavage activity of Pol I[7] and the approximate positioning of the C-terminal Zn ribbon by chemical crosslinking[15] (Fig. 4). Indeed, our structure shows that an acidic loop in this subunit lies next to the active site of the enzyme and is probably involved in RNA cleavage, as shown for the corresponding residues in TFIIS[36]. The two A12.2 Zn ribbons are connected by a linker (residues 45–80) that extends along the bottom face of the enzyme (Extended Data Fig. 7a). The N- and C-terminal portions of this linker follow equivalent paths to those of the Rpb9 and TFIIS linkers, respectively, whereas the middle region of A12.2 (residues 52–70) folds into a helix-containing 'mini-domain' that inserts between the A190 funnel and jaw domains. A Pol-I-specific insertion in subunit A190 (residues 863–887) interacts with this A12.2 region through a helix that would clash with the C-terminal domain of Rpb9 in Pol II.

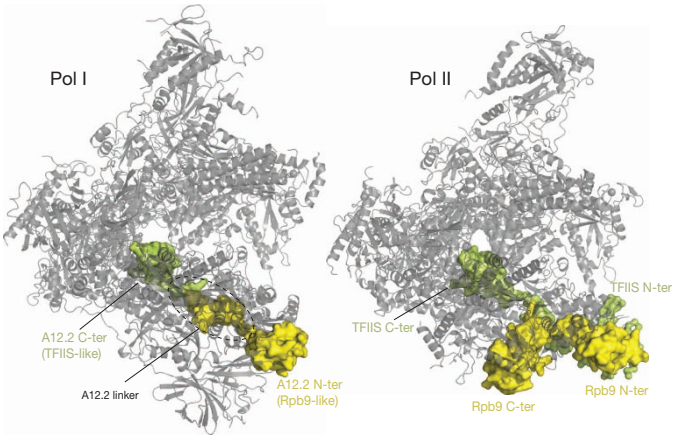


Figure 4 | Pol I subunit A12.2 extends from the jaw to the active site. On the left, A12.2 is depicted in surface representation with the N-terminal Zn ribbon (yellow) occupying the jaw region similar to the N-terminal domain of Rpb9 in Pol II, and its C-terminal part (green) residing in the active site similar to the C-terminal domain of TFIIS. On the right, the Pol II–TFIIS complex (Protein Data Bank accession 1Y1V) with Rpb9 (yellow) and TFIIS (green) depicted in surface representation.

The intrinsic RNA cleavage activity of Pol I has been described to be important for proofreading, termination and rRNA 3′ terminal processing during Pol I transcription[7,37]. Interestingly, a C-terminal Zn ribbon containing a TFIIS-like acidic loop is also present in subunit C11, the Pol III counterpart for A12.2, indicating a similar domain organization for C11 in Pol III. The position of the C-terminal Zn ribbon in Pol I may also explain the insensitivity of Pol I towards α-amanitin, a strong Pol II inhibitor[38]. Superposition of the Pol II/α-amanitin crystal structure with the Pol I structure shows that the inhibitor would need to displace the C-terminal Zn ribbon of A12.2, which is unlikely due to the extensive network of contacts that this domain establishes with the Pol I core (Extended Data Fig. 7b). The medium sensitivity of Pol III towards α-amanitin[39] indicates that the C-terminal Zn ribbon of C11 may be less tightly bound to the core Pol III enzyme or may occupy a slightly different position.

Extended arms fix A49–A34.5 on the lobe

The A49–A34.5 heterodimer comprises two folded regions: a dimerization module (A49 residues 6–102 and A34.5 residues 21–141) and the A49 tWH domain (residues 185–415). Whereas the A49 tWH domain seems to be mobile in all three crystal forms and thus cannot be modelled, the dimerization module forms a triple β-barrel—like its *Candida albicans* orthologue[10]—that docks on the lobe and external two domains of subunit A135, and also contacts the N-terminal Zn ribbon of subunit A12.2 in an unforeseen manner (Fig. 5). Its position, unambiguously confirmed by selenomethionine anomalous difference maps (Extended Data Fig. 1e), is roughly coherent with that observed in early cryo-EM studies[40] and by crosslinking experiments[15]. Moreover, on binding to the Pol I core, previously unobserved arms protrude from the C-terminal ends of both A49 and A34.5 dimerization domains. Notably, the A34.5 C-terminal tail has an essential role in heterodimer anchoring, as it extends by more than 70 Å along the outer face of A135 and reaches into a cavity at the junction of subunits A135, Rpb10 and AC40 (Fig. 5). Additionally, the first 11 residues in the A49 linker (residues 103–113) strongly interact with the A135 lobe domain and the A12.2 N-terminal Zn ribbon, further anchoring the heterodimer to the core (Fig. 5). These results are coherent with biochemical data showing that either the A49 linker or the A34.5 C-terminal tail is necessary for stable binding of the heterodimer to the core[10]. Moreover, interaction with A12.2 could explain the stimulatory role of A49–A34.5 in the intrinsic RNA cleavage of Pol I[10], as stabilization of the A12.2 N-terminal ribbon may also determine the position of its C-terminal ribbon in the cleft.

In Pol III, the C37–C53 heterodimer[9,10,12] seems to be the equivalent to the A49–A34.5 heterodimer, and the Pol I crystal structure can be perfectly placed into the cryo-EM reconstruction of Pol III[12] (Extended Data Fig. 8a). Nevertheless, as the domain organization is different in these subunits, the N- and C-terminal extensions of C37–C53 probably interact differently with the Pol III core. Our results can also be extended to the Pol II system, where an equivalent dimerization module is present in the TFIIF Rap74–Rap30 heterodimer (Extended Data Fig. 8b, c). Crosslinking[41,42] and cryo-EM studies[43] of Pol II in complex with TFIIF show a similar location for this module and the cryo-EM studies also locate the C-terminal winged helix (WH) domain of Rap30, the A34.5 counterpart, between the protrusion and lobe domains of Pol II. In Pol I, these domains are closer together and interact directly, thereby occluding the corresponding space, consistent with the absence of a WH domain in A34.5. Finally, the Pol I A49 tWH domain probably corresponds to the small subunit of the TFIIE heterodimer, which has also been observed in cryo-EM studies[43]. Whereas the A49 tWH domain is disordered in the Pol I crystal structure, we speculate that in a pre-initiation complex it may occupy an equivalent position over the transcription start site, in accordance to crosslinking data[15], thus participating in DNA melting as well as in the recruitment of the A43-interacting transcription factor Rrn3 (ref. 44).

The crystal structure of Pol I further corroborates the emerging view that the basic architecture of Pol I, Pol II and Pol III and their specific pre-initiation complexes is conserved[5]. The Pol I crystal structure presented here provides a structural framework to rationalize available genetic and biochemical data on Pol I transcription of ribosomal RNA genes and represents an important step towards a better understanding of eukaryotic transcription mechanisms. In addition, it will hopefully stimulate future research to target the Pol I transcription machinery by small molecule effectors as potential anticancer drugs.

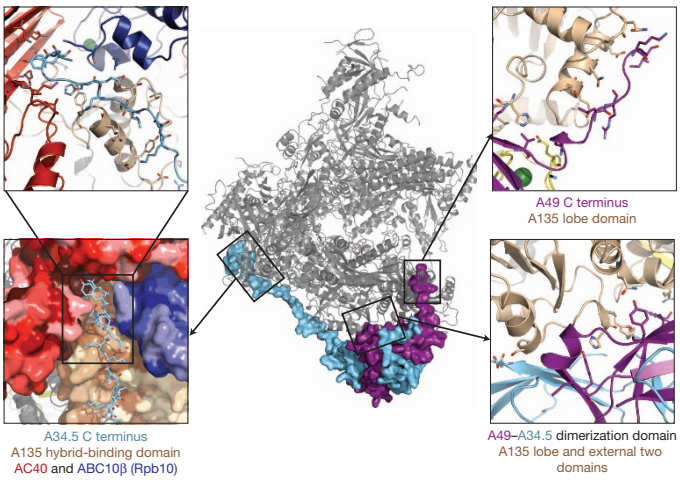


Figure 5 | Anchoring of the A49–A34.5 heterodimer onto the Pol I core. Central panel: A49 (purple) and A34.5 (cyan) are highlighted by surface representation. Left and right panels show residues of the A49–A34.5 heterodimer interacting with the core. Left panels: the C-terminal A34.5 hook inserts into a conserved cavity formed by subunits A135, AC40 and Rpb10. Subunits forming the cavity are coloured according to sequence conservation with darkest colour corresponding to the most conserved residues. Right panel, top: A49 linker residues interact with the lobe domain of A135. Right panel, bottom: A49–A34.5 dimerization domain contacts lobe and external two domains of A135.

METHODS SUMMARY

Endogenous *Saccharomyces cerevisiae* Pol I containing all 14 subunits was purified using several chromatographic steps. Crystallization using the sitting-drop method yielded crystals diffracting up to 3.0 Å resolution, according to recently defined criteria[18]. The Pol I crystal structure was solved by combining phase information from a molecular replacement solution using Pol II as search model with phases obtained from various heavy-atom derivatives, followed by solvent flattening and multi-crystal averaging. The resulting model was refined at 3.03 Å resolution with R_{work}/R_{free} = 19.7/22.9% and excellent geometry.

Online Content Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 27 May; accepted 4 September 2013.
Published online 23 October 2013.

1. Warner, J. R. The economics of ribosome biosynthesis in yeast. *Trends Biochem. Sci.* **24,** 437–440 (1999).
2. Drygin, D., Rice, W. G. & Grummt, I. The RNA polymerase I transcription machinery: an emerging target for the treatment of cancer. *Annu. Rev. Pharmacol. Toxicol.* **50,** 131–156 (2010).
3. Grummt, I. Life on a planet of its own: regulation of RNA polymerase I transcription in the nucleolus. *Genes Dev.* **17,** 1691–1702 (2003).
4. 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.* **64,** 29–49 (2007).
5. Vannini, A. & Cramer, P. Conservation between the RNA polymerase I, II, and III transcription initiation machineries. *Mol. Cell* **45,** 439–446 (2012).
6. Russell, J. & Zomerdijk, J. C. The RNA polymerase I transcription machinery. *Biochem. Soc. Symp.* **73,** 203–216 (2006).
7. Kuhn, C. D. *et al.* Functional architecture of RNA polymerase I. *Cell* **131,** 1260–1272 (2007).

8. Ruan, W., Lehmann, E., Thomm, M., Kostrewa, D. & Cramer, P. Evolution of two modes of intrinsic RNA polymerase transcript cleavage. *J. Biol. Chem.* **286,** 18701–18707 (2011).
9. Carter, R. & Drouin, G. The increase in the number of subunits in eukaryotic RNA polymerase III relative to RNA polymerase II is due to the permanent recruitment of general transcription factors. *Mol. Biol. Evol.* **27,** 1035–1043 (2010).
10. Geiger, S. R. *et al.* RNA polymerase I contains a TFIIF-related DNA-binding subcomplex. *Mol. Cell* **39,** 583–594 (2010).
11. Cramer, P. *et al.* Structure of eukaryotic RNA polymerases. *Ann. Rev. Biophys.* **37,** 337–352 (2008).
12. Fernández-Tornero, C. *et al.* Conformational flexibility of RNA polymerase III during transcriptional elongation. *EMBO J.* **29,** 3762–3772 (2010).
13. Bischler, N. *et al.* Localization of the yeast RNA polymerase I-specific subunits. *EMBO J.* **21,** 4136–4144 (2002).
14. Schultz, P., Celia, H., Riva, M., Sentenac, A. & Oudet, P. Three-dimensional model of yeast RNA polymerase I determined by electron microscopy of two-dimensional crystals. *EMBO J.* **12,** 2601–2607 (1993).
15. Jennebach, S., Herzog, F., Aebersold, R. & Cramer, P. Crosslinking-MS analysis reveals RNA polymerase I domain architecture and basis of rRNA cleavage. *Nucleic Acids Res.* **40,** 5591–5601 (2012).
16. Fernández-Tornero, C. *et al.* Insights into transcription initiation and termination from the electron microscopy structure of yeast RNA polymerase III. *Mol. Cell* **25,** 813–823 (2007).
17. Cramer, P., Bushnell, D. A. & Kornberg, R. D. Structural basis of transcription: RNA polymerase II at 2.8 Ångstrom resolution. *Science* **292,** 1863–1876 (2001).
18. Karplus, P. A. & Diederichs, K. Linking crystallographic model and data quality. *Science* **336,** 1030–1033 (2012).
19. Armache, K. J., Kettenberger, H. & Cramer, P. Architecture of initiation-competent 12-subunit RNA polymerase II. *Proc. Natl Acad. Sci. USA* **100,** 6964–6968 (2003).
20. Bushnell, D. A. & Kornberg, R. D. Complete, 12-subunit RNA polymerase II at 4.1-Å resolution: implications for the initiation of transcription. *Proc. Natl Acad. Sci. USA* **100,** 6969–6973 (2003).
21. 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).
22. Kettenberger, H., Armache, K. J. & Cramer, P. Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. *Mol. Cell* **16,** 955–965 (2004).
23. Hirata, A., Klein, B. J. & Murakami, K. S. The X-ray crystal structure of RNA polymerase from Archaea. *Nature* **451,** 851–854 (2008).
24. Zhang, G. *et al.* Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 Å resolution. *Cell* **98,** 811–824 (1999).
25. Vassylyev, D. G. *et al.* Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* **417,** 712–719 (2002).
26. Tan, L., Wiesler, S., Trzaska, D., Carney, H. C. & Weinzierl, R. O. Bridge helix and trigger loop perturbations generate superactive RNA polymerases. *J. Biol.* **7,** 40 (2008).
27. Jovanovic, M. *et al.* Activity map of the *Escherichia coli* RNA polymerase bridge helix. *J. Biol. Chem.* **286,** 14469–14479 (2011).
28. Weinzierl, R. O. The bridge helix of RNA polymerase acts as a central nanomechanical switchboard for coordinating catalysis and substrate movement. *Archaea* **2011,** 608385 (2011).
29. Edwards, A. M., Kane, C. M., Young, R. A. & Kornberg, R. D. Two dissociable subunits of yeast RNA polymerase II stimulate the initiation of transcription at a promoter *in vitro*. *J. Biol. Chem.* **266,** 71–75 (1991).
30. Mosley, A. L. *et al.* Quantitative proteomics demonstrates that the RNA polymerase II subunits Rpb4 and Rpb7 dissociate during transcription elongation. *Mol. Cell. Proteomics* **12,** 1230–1538 (2013).
31. 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).
32. 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).
33. Blattner, C. *et al.* Molecular basis of Rrn3-regulated RNA polymerase I initiation and cell growth. *Genes Dev.* **25,** 2093–2105 (2011).
34. Milkereit, P., Schultz, P. & Tschochner, H. Resolution of RNA polymerase I into dimers and monomers and their function in transcription. *Biol. Chem.* **378,** 1433–1443 (1997).
35. Kettenberger, H., Armache, K. J. & Cramer, P. Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. *Cell* **114,** 347–357 (2003).
36. Jeon, C., Yoon, H. & Agarwal, K. The transcription factor TFIIS zinc ribbon dipeptide Asp-Glu is critical for stimulation of elongation and RNA cleavage by RNA polymerase II. *Proc. Natl Acad. Sci. USA* **91,** 9106–9110 (1994).
37. Prescott, E. M. *et al.* Transcriptional termination by RNA polymerase I requires the small subunit Rpa12p. *Proc. Natl Acad. Sci. USA* **101,** 6068–6073 (2004).
38. Kedinger, C., Gniazdowski, M., Mandel, J. L. Jr, Gissinger, F. & Chambon, P. α-Amanitin: a specific inhibitor of one of two DNA-pendent RNA polymerase activities from calf thymus. *Biochem. Biophys. Res. Commun.* **38,** 165–171 (1970).
39. Weinmann, R. & Roeder, R. G. Role of DNA-dependent RNA polymerase 3 in the transcription of the tRNA and 5S RNA genes. *Proc. Natl Acad. Sci. USA* **71,** 1790–1794 (1974).
40. De Carlo, S., Carles, C., Riva, M. & Schultz, P. Cryo-negative staining reveals conformational flexibility within yeast RNA polymerase I. *J. Mol. Biol.* **329,** 891–902 (2003).
41. Chen, Z. A. *et al.* Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. *EMBO J.* **29,** 717–726 (2010).
42. Eichner, J., Chen, H. T., Warfield, L. & Hahn, S. Position of the general transcription factor TFIIF within the RNA polymerase II transcription preinitiation complex. *EMBO J.* **29,** 706–716 (2010).
43. He, Y., Fang, J., Taatjes, D. J. & Nogales, E. Structural visualization of key steps in human transcription initiation. *Nature* **495,** 481–486 (2013).
44. 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).

Acknowledgements We thank H. Grötsch for preparing the loop deletion yeast strain, and G. von Scheven and A. Scholz for technical assistance. We are also grateful to C. Vonrhein, G. Bricogne, S. Glatt and A. Romero for advice and discussions. We thank staff from the European synchrotrons SOLEIL, DESY, ESRF and SLS, where data were collected during different stages of the project. In particular, we thank A. Thompson for access and support at beamline Proxima 1 (Soleil) and T. Schneider and G. Bourenkov at beamline P14 (PETRA III). We also acknowledge support by the EMBL Heidelberg Protein Expression and Purification, Proteomics Core Facilities and Crystallization Platform, and the 'Fermentation et culture de microorganisms' (IFR88, CNRS). We are grateful to M. Bauzan, E. Poilpre and J. Scheurich for yeast fermentation. M.M.-M. and U.J.R. were supported by EMBO Long-Term fellowships, M.M.-M. by the Marie-Curie fellowship (FP7-PEOPLE-2011-IEF 301002), N.M.I.T. by a Fundación Futuro fellowship, F.M.R. by an ESF/CSIC funded JAE-DOC contract and T.G. by the Volkswagen Stiftung via the Niedersachsenprofessur of Prof. G. M. Sheldrick. This work was also partly funded by grant BFU2010-16336 of the Spanish Ministry of Science.

Author Contributions C.F.-T. and C.W.M. initiated the project. C.F.-T. and U.S. established the Pol I purification and obtained Pol I crystals. C.F.-T., M.M.-M. and U.J.R. further improved the Pol I crystals, and collected data and obtained experimental phase information with the help of N.M.I.T., T.G. and P.L. C.F.-T., M.M.-M., N.M.I.T., F.M.R., T.G. and P.L. carried out the crystallographic analysis and model refinement. C.F.-T., M.M.-M. and C.W.M. wrote the manuscript with input from the other authors.

Author Information Atomic coordinates and structure factors of the three related Pol I crystal forms have been deposited at the Protein Data Bank under accession numbers 4C3H, 4C3I and 4C3J. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to C.F.-T. (cftornero@cib.csic.es) or C.W.M. (cmueller@embl.de).

METHODS

Yeast Pol I expression and purification. The Pol I enzyme was isolated from *Saccharomyces cerevisiae* strain SC1613 (also called YPR110c, provided by Cellzome AG), modified to express endogenous AC40 fused with a C-terminal TAP-tag. Yeast cells were grown overnight in YPDA medium at 30 °C and 180 r.p.m. under controlled conditions and collected at an OD_{600} of 5–6. The cell paste was re-suspended in a buffer containing 250 mM Tris-HCl, pH 8, 40% glycerol, 250 mM $(NH_4)_2SO_4$, 1 mM EDTA, 10 mM $MgCl_2$, 10 μM $ZnCl_2$, 12 mM β-mercaptoethanol in the presence of a protease inhibitor cocktail (complete EDTA-free, Roche) and lysed with glass beads in a BeadBeater (BioSpec). After centrifugation at 14,000 r.p.m. for 1 h at 4 °C, the protein lysate was loaded on heparin-sepharose resin (GE Healthcare). The complex was eluted from the resin using high-salt buffer with 1 M $(NH_4)_2SO_4$ and further incubated with IgG Sepharose (GE-Healthcare) for 6 h. After washing, IgG beads were mixed with tobacco etch virus (TEV) protease and incubated overnight at 4 °C. TAP-tag cleaved Pol I was recovered and subsequently purified by ionic exchange on a Mono-Q column (GE-Healthcare). The pure Pol I enzyme was concentrated to 7 mg ml^{-1} in 15 mM Tris-HCl, pH 8.0, 150 mM NaCl and 10 mM dithiothreitol (DTT). Yeast cells containing selenomethionine-labelled Pol I were prepared from an initial glycerol stock pre-adapted to 80 μg ml^{-1} of selenomethionine. A synthetic complete medium, which included yeast nitrogen base lacking amino acids, D-glucose and an amino acid mix containing different concentrations of selenomethionine, was used for fermentation at 30 °C and 180 r.p.m. Purification was performed under the same conditions as for the native complex.

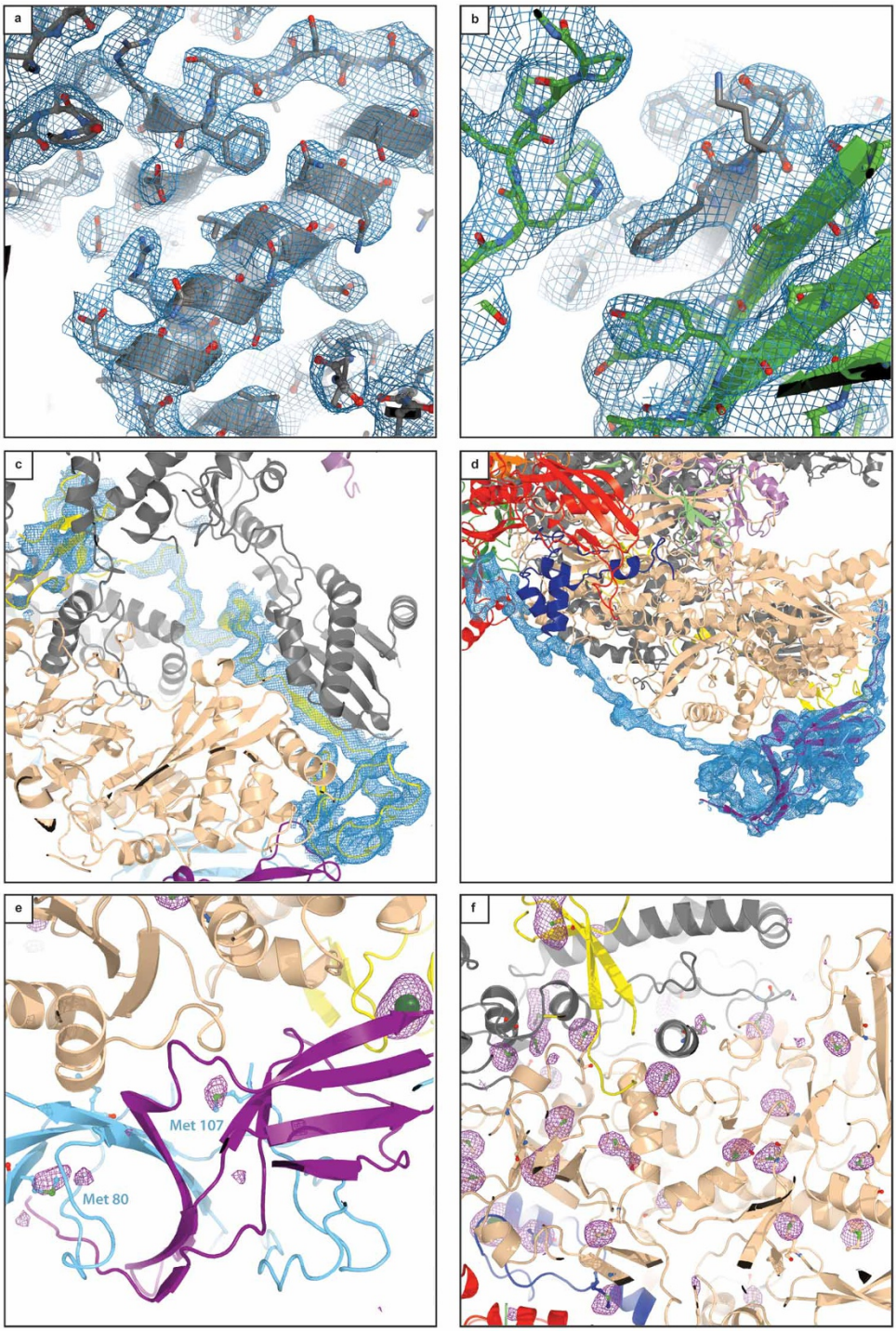
Yeast Pol I crystallization. Well-diffracting plate-shaped crystals grew during 4–7 days in two different conditions to a maximum size of 0.9 × 0.4 × 0.3 mm^3 at 18 °C using vapour diffusion, sitting-drop plates. Protein and crystallization solutions were mixed in a ratio 1:1. In the first condition (16–30% ethylenglycol (EG) and 100 mM 2-(*N*-morpholino)-ethane-sulphonic acid (MES) pH 6.3–6.9) cryo-protection was reached by stepwise increases of the concentration of EG up to 30% in the reservoir. The second condition included 1–12.5% 2-methyl-2,4-pentanediol (MPD) and 100 mM Tris-HCl, pH 6.3 to 6.9. In this case, crystals were directly soaked in a cryo-solution containing 22.5% MPD. Crystals were frozen in cryo-loops under a nitrogen stream at 100 K. For initial, low-resolution phasing, crystals grown in the presence of EG were soaked for 1 h in a solution containing 2 mM Ta_6Br_{12}. For optimized experimental phasing, an ytterbium derivative crystal was obtained by soaking in a 100 mM Yb-HPDO3A containing solution for 2 min. Selenomethionine-labelled Pol I crystals were grown in similar conditions.

Data collection and structure determination. Data were collected at synchrotrons Soleil, ESRF, SLS and DESY and processed with XDS[45]. All crystals belong to space group *C*2 but vary in cell dimensions (Extended Data Table 1). Crystal form C2-93 ($\beta = 93°$) was obtained in EG, whereas crystal forms C2-90 ($\beta = 90°$) and C2-100 ($\beta = 100°$) were grown in MPD. Anomalous data sets for tantalum bromide, selenomethionine and zinc were collected at the absorption edge of these elements using an inverse beam protocol. For the ytterbium derivative, the twofold symmetry axis of the crystal was aligned with the spindle axis to collect Bijvoet pairs on the same image and a multi-wavelength anomalous dispersion (MAD) data set was collected at four wavelengths. Initial phases were obtained in crystal form C2-93 through molecular replacement with Phaser[46] using the Pol II core structure, but did not allow further model building. After SIRAS phasing with SHARP[47] using a tantalum bromide derivative in this crystal form that diffracted to 6.5 Å and a native data set diffracting to 4.0 Å, followed by density modification and phase combination with the molecular replacement phases, an improved map was obtained. This allowed extension of the initial model with Coot[48] and subsequent refinement with Refmac[49]. Better phases were obtained with a MAD experiment in crystal form C2-90 performed on an ytterbium derivative diffracting to 4.0 Å, and the model could be further improved in this space group and refined using Refmac. Initial phases in crystal form C2-93 were improved by MIRAS using a 3.6 Å native data set combined with tantalum bromide, ytterbium and lead derivative data sets, and combined with model phases followed by density modification. Multi-crystal averaging using improved native data sets diffracting to 3.3 Å in crystal forms C2-93 and C2-90 yielded maps of excellent quality where most side chains were visible, and thus the structure could be completed and further refined using Phenix[50]. Finally, a data set was obtained in crystal form C2-100 and used for refinement with BUSTER[51] against data diffracting to 3.03 Å. Further validation was performed with a data set of a selenomethionine containing Pol I in crystal form C2-90, collected at the peak of the selenium absorption spectrum that allowed locating 90 out of 100 selenium atoms within a distance of less than 2.3 Å to corresponding methionine residues. Additionally, the position of the zinc atoms was also validated by inspection of anomalous difference maps. The final model in crystal form C2-100 has an R_{work}/R_{free} of 19.7/22.9% with excellent stereochemistry according to MolProbity[52] (Extended Data Table 1) and only 7 residues in disallowed regions of the Ramachandran plot. Further experimental details on the X-ray structure determination will be published elsewhere.

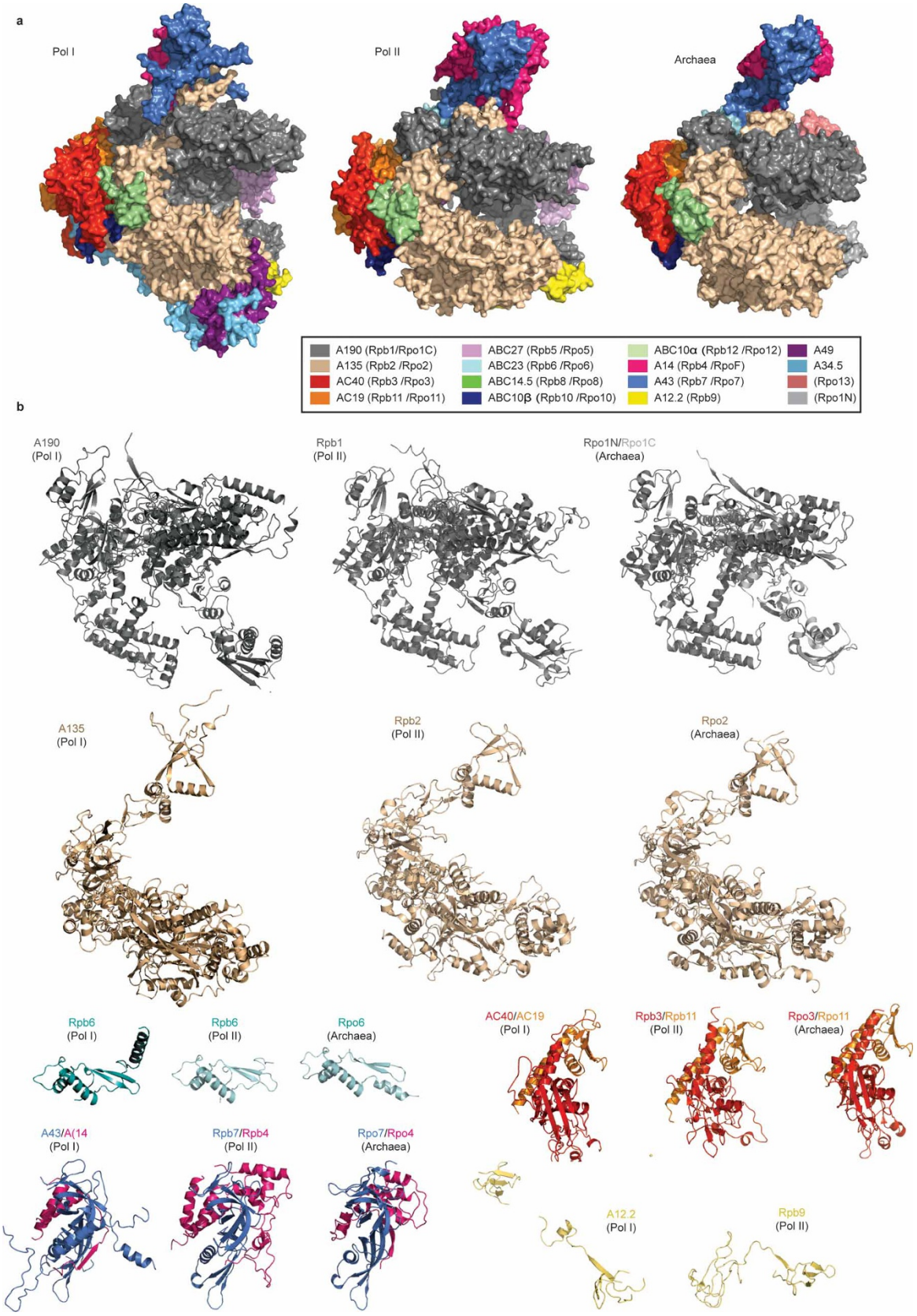
Pol I RNA extension assay. RNA (5′-AAGACCAGGC-3′) was first radioactively labelled with ^{32}P by T4 PNK and gel purified on denaturing 15% urea-PAGE. After annealing of 2 pmol of radioactive RNA with 2 pmol of single-stranded DNA (template DNA: 3′-CCGTCATGATCATTACTGGTCCGCATTCATGA ACTCGAACC-5′ or 3′-GCTCAGCCTG GTCCGCATGTGTCAGTC-5′) the DNA/RNA hybrid was incubated with 4 pmol of Pol I for 20 min at 20 °C in 20 mM Tris pH 7.5, 150 mM KCl, 10 mM DTT and 10 mM $MgCl_2$. RNA elongation was initiated by addition of 1 mM NTPs in 60 mM ammonium sulphate, 20 mM Tris pH 7.5, 8 mM $MgCl_2$, 10 μM zinc chloride, 10% glycerol, and 10 mM DTT. After incubation for 20 min at 28 °C, RNA extension was stopped by addition of loading buffer (8 M urea, TBE) and heating for 2 min at 95 °C. The resulting P^{32}-labelled RNA product was subsequently analysed on a denaturing polyacrylamide gel (17% PAGE, 8 M urea). The gel was fixed in 10% ethanol, 6% acetic acid, dried and autoradiographed with an X-Ray film (Biomax, MR-film, Kodak).

Construction of the yeast strain with DNA-mimicking loop deletion. To construct the yeast strain with a deletion of the DNA-mimicking loop in RPA190, a C-terminal fragment of *RPA190* was amplified from pNOY16[53] using the primers RPA190-PstI-s (5′-GTCGTGAAGCCGCTGCAG-3′) and RPA190-STOP-EcoRI-as (5′-ACAGAATTCTGTTGCTA GCCTAAGCCGCATTTGGAACCTTTG-3′) and cloned into PstI and EcoRI sites of pUC19, generating an NheI site after the *RPA190* STOP codon. Subsequently, the plasmid without the DNA-mimicking loop, pUC19-rpa90Δloop, was generated by restriction-free cloning using the primers RPA190woloop-s (5′-GCTACAGACAGATGTTGCAAATAGTTCTTC GGAAATTGAAACTATGAGAGAAGCTGAAAAGTCTTCT-3′) and RPA190 woloop-as (5′-AGAAGACTTTTCAGCTTCTCTCATAGTTTCAATTTCCGAA GAACTATTTGCAACATCTGTCTGTAGC-3′). The NatNT2 marker was amplified from pFA6a-NatNT2[54] using the primers NatN2-NheI-s (5′-AAGCTAGC CGTACGCTGCAGGTCGAC-3′) and S2 (5′-ATCGATGAATTCGAGCTCG-3′) and cloned into the NheI and EcoRI sites of pUCrpa90Δloop to generate pUCrpa 90Δloop-Nat. A deletion cassette was finally amplified from pUCrpa90Δloop-Nat using the primers loop-456-s (GGTCACGGTGCCGCTAATG) and LoopMarker as (5′-CCTTCTCCTTCAAATAAACTAATATTAAATCGTAATAATTATGG GACCTTTTGCCTGCTTGGCCAGTGAATTCGAGCTCG-3′), purified and transformed into yeast strain SC1613 by the lithium acetate method[55]. Colonies were selected on YPD plates containing 100 μg ml^{-1} of ClonNat and analysed by colony PCR and subsequent sequencing. Growth of the yeast strain was monitored in comparison to the parental strain SC1613. Serial dilutions were done on SDC and YPD plates and incubated at 25 °C, 30 °C and 37 °C.

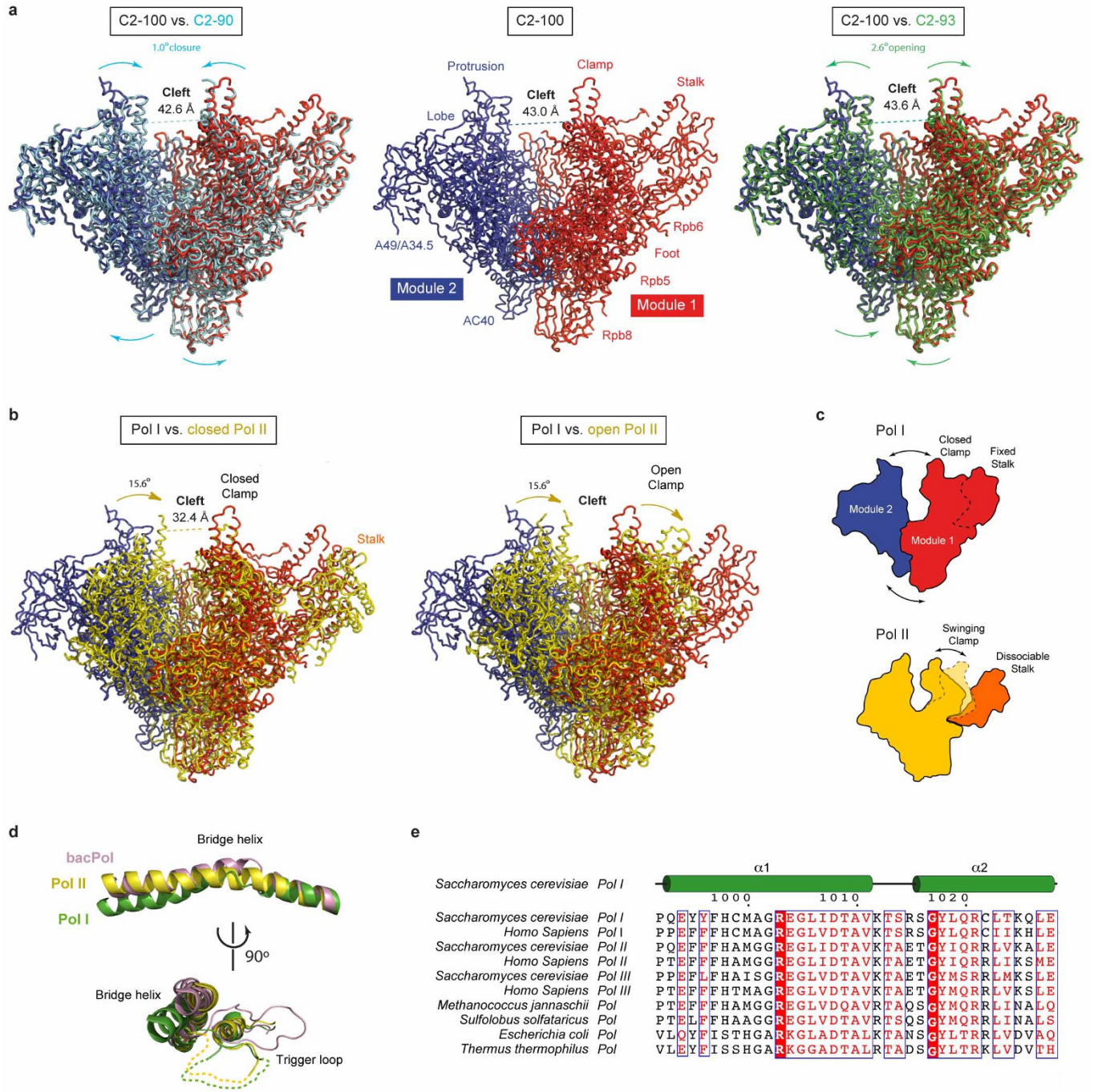
45. Kabsch, W. Xds. *Acta Crystallogr. D* **66,** 125–132 (2010).
46. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674 (2007).
47. Bricogne, G., Vonrhein, C., Flensburg, C., Schiltz, M. & Paciorek, W. Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. *Acta Crystallogr. D* **59,** 2023–2030 (2003).
48. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60,** 2126–2132 (2004).
49. Skubák, P., Murshudov, G. N. & Pannu, N. S. Direct incorporation of experimental phase information in model refinement. *Acta Crystallogr. D* **60,** 2196–2201 (2004).
50. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66,** 213–221 (2010).
51. Smart, O. S. *et al.* Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. *Acta Crystallogr. D* **68,** 368–380 (2012).
52. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66,** 12–21 (2010).
53. Wittekind, M. *et al.* Isolation and characterization of temperature-sensitive mutations in RPA190, the gene encoding the largest subunit of RNA polymerase I from *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **8,** 3997–4008 (1988).
54. Janke, C. *et al.* A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. *Yeast* **21,** 947–962 (2004).
55. Ito, H., Fukuda, Y., Murata, K. & Kimura, A. Transformation of intact yeast cells treated with alkali cations. *J. Bacteriol.* **153,** 163–168 (1983).



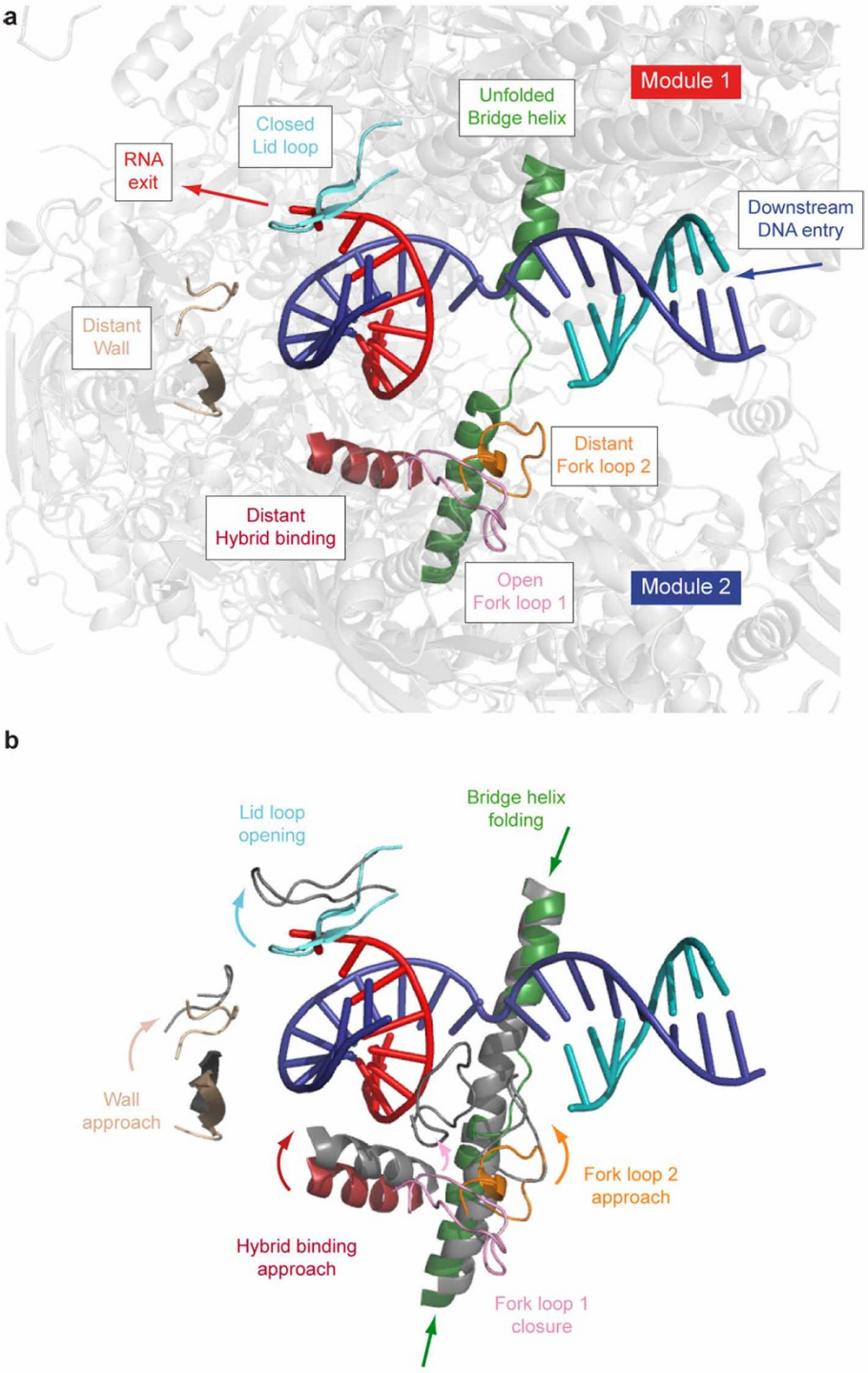
Extended Data Figure 1 | Electron densities of different regions of the Pol I structure. **a**, Helix in the cleft of subunit A190. Subunits are coloured according to the code given in Fig. 1a. In **a**–**d**, σ_A-weighted electron densities contoured at 1σ are depicted in blue. **b**, Interaction between subunit Rpb8 (green) and subunit A190 (grey). **c**, Linker between the two Zn ribbon domains in subunit A12.2. **d**, A49–A34.5 heterodimer and the anchoring onto the Pol I core by the A34.5 hook and the A49 linker region. **e**, Anomalous difference Fourier map (purple) calculated from partially selenomethionine-substituted Pol I contoured at 3σ (Extended Data Table 1). In the A49–A34.5 heterodimer two selenium peaks correspond to A34.5 Met 80 and Met 107. **f**, Anomalous difference Fourier map (purple) showing selenomethionine positions contoured at 3σ. In total, 90 out of 100 expected selenium positions were located within a distance of less than 2.3 Å to corresponding methionine residues.



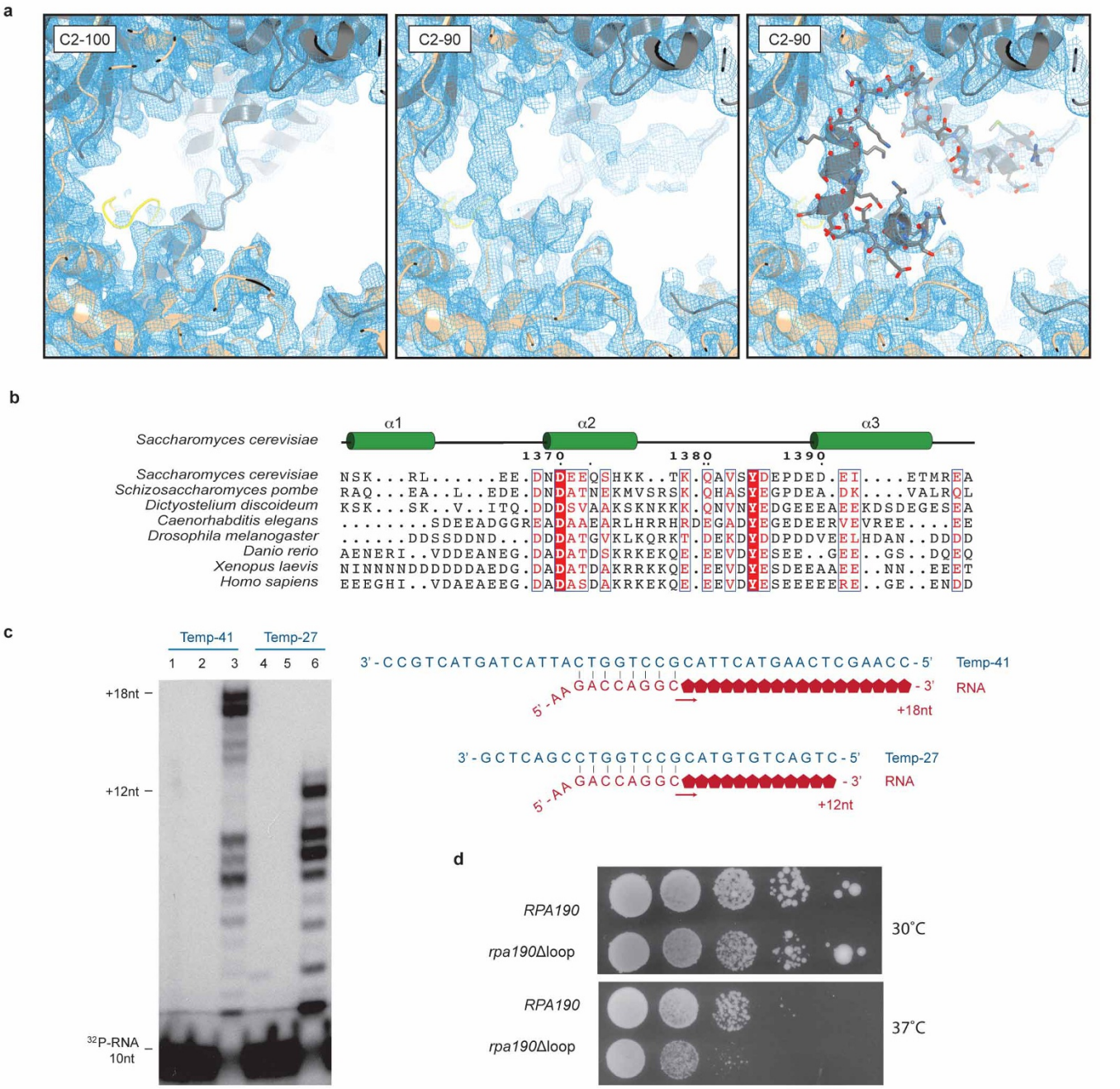
Extended Data Figure 2 | Comparison between Pol I, Pol II and archaeal Pol. **a**, Crystal structures of yeast Pol I and Pol II and the archaeal Pol are represented in the same orientation using the same colour code. Whereas the overall organization is conserved, additional subunits such as the A49–A34.5 heterodimer in Pol I and Rpo13 in Archaea are also present. The archaeal Pol lacks the orthologue of subunit A12.2 in Pol I (or Rpb9 in Pol II). The relative position of the stalk also varies between the three RNA polymerases. **b**, Crystal structures of the individual subunits varying between the three enzymes are depicted. The same colour code for corresponding or identical subunits in the three RNA polymerases is used.



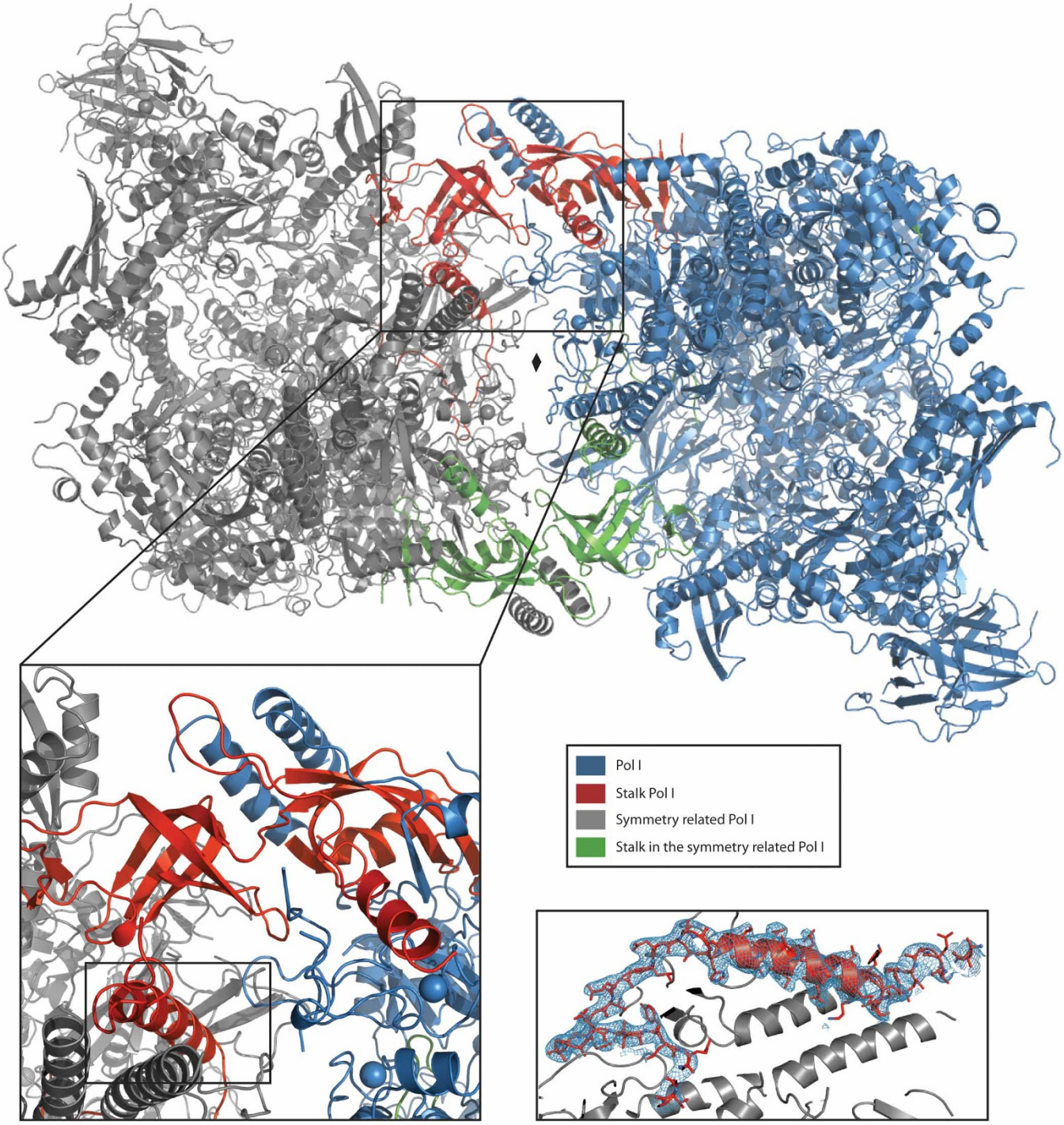
Extended Data Figure 3 | Opening of the cleft varies among different Pol I structures and between Pol I and Pol II. a, Middle panel: front view of the Pol I structure in crystal form C2-100 (see Extended Data Table 1). The complex is divided into two modules. Module 1 (red) is formed by the major part of subunit A190 (without the pore 1, funnel and jaw domains), the C terminus of A135, Rpb5, Rpb6, Rpb8 and the stalk subunits, whereas module 2 (blue) comprises the remaining A135 domains, the pore 1, funnel and jaw domains of A190, AC40–AC19, Rpb10, Rpb12, A12.2 and the A49–A34.5 heterodimer. These modules are held by three hinges in A190 (active site–pore 1 connection, bridge helix and jaw–cleft connection) and one hinge in A135 (hybrid binding–anchor connection) as indicated in Fig. 1b. Pol I structures obtained in crystal forms C2-90 (left panel) and C2-93 (right panel) were superimposed with the one obtained in crystal form C2-100 taking module 2 as reference. Differences between the different crystal forms in the cleft aperture and the tilting of the mobile modules are indicated. **b**, Pol II structure (Protein Data Bank accession 1WCM) is superimposed onto Pol I (C2-100) using module 1 as reference. In comparison, the Pol II cleft is closed by 10 Å and the modules rotate 15.6°. **c**, Schematic representation of Pol I and Pol II showing the mobility between Pol I modules, as well as the conformation of the stalk and the clamp. Colour coding is as in **b**, with the exception of the Pol II stalk, which is coloured deep orange. **d**, Conformation of the bridge helix of the bacterial *Thermus thermophilus* polymerase (bacPol, pink, Protein Data Bank accession 1IW7), Pol I (green) and Pol II (yellow, Protein Data Bank accession 1WCM). In addition, the trigger loop is shown for the *Thermus thermophilus* polymerase, where it is ordered, and as dotted lines for Pol I and Pol II, where it is disordered. **e**, Sequence alignment of the bridge helix of *Saccharomyces cerevisiae* and *Homo sapiens* Pol I, Pol II and Pol III, archaeal *Methanococcus jannaschii* and *Sulfolobus solfataricus*, and the bacterial *Escherichia coli* and *Thermus thermophilus* polymerases. The secondary structure of the Pol I bridge helix is shown above the alignment. In *Methanococcus jannaschii*, site-directed mutations Q823D and S824P in subunit A′ lead to increased transcriptional activity.



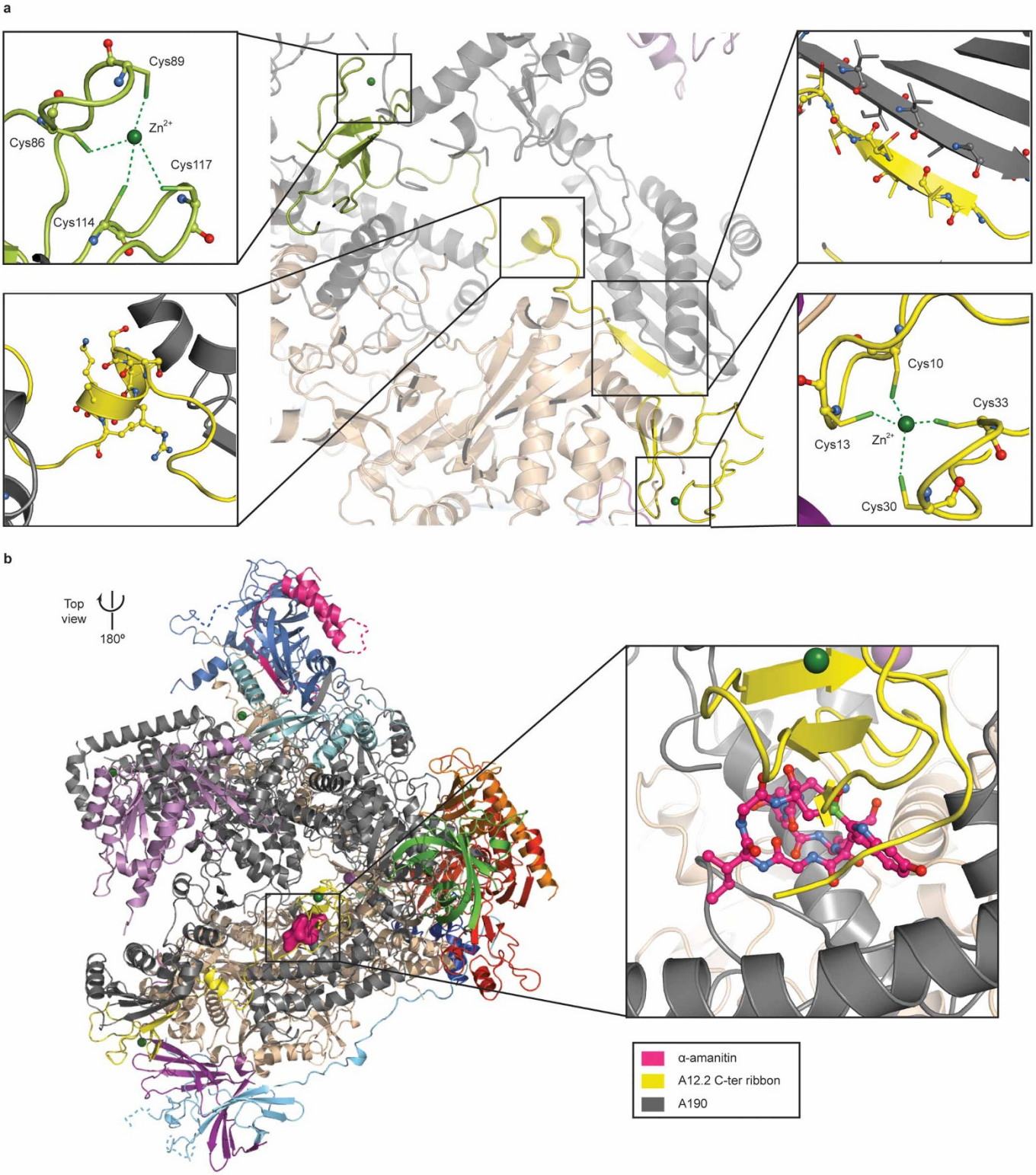
Extended Data Figure 4 | Elongation complex. a, Cartoon representation of a model of Pol I in complex with an elongation bubble, generated by superposition of the Pol II elongation complex crystal structure (Protein Data Bank accession 1Y1W) using the largest subunit as reference. Whereas Pol II is not shown, the coding and non-coding DNA strands are depicted in blue and cyan, respectively, and the RNA in red. The main Pol I elements putatively involved in nucleic acid interaction appear in different colours, whereas the rest of the Pol I structure is shown in light grey. **b**, Proposed rearrangements in elongating Pol I (coloured elements) in analogy with Pol II (grey elements). Closure of the cleft is expected to approach the wall (tan), the hybrid binding domain (dark red) and the fork loop 2 (orange) to the bubble, as well as to fold the bridge helix (green). Closure of the fork loop 1 (pink) and opening of the lid loop (cyan) would also be required.



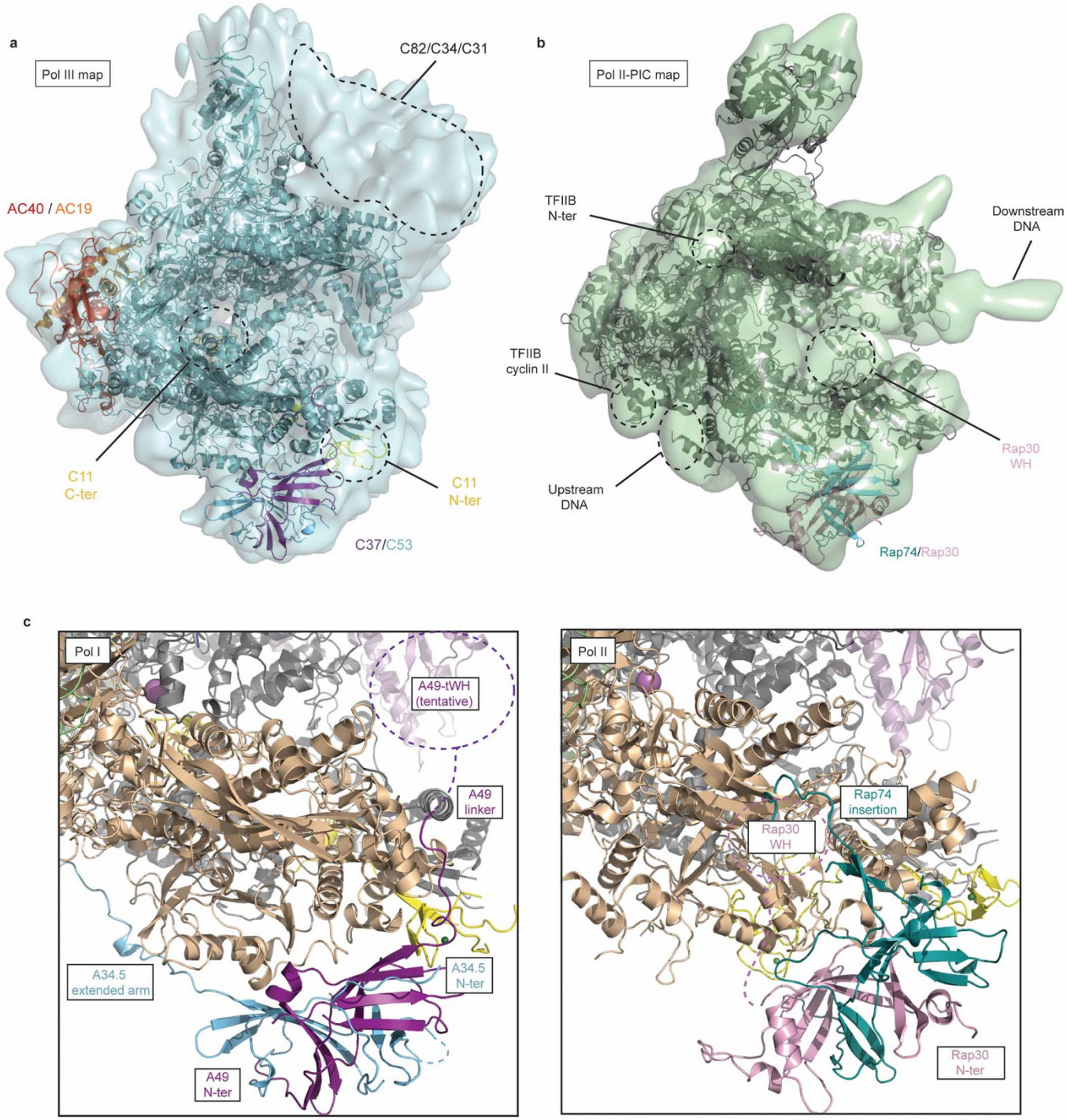
Extended Data Figure 5 | The DNA-mimicking loop of Pol I forms a mobile element. **a**, σ_A-weighted electron density contoured at 1σ of A190 jaw residues 1361–1399 (middle and right panel) in crystal form C2-90. Density is also present in crystal form C2-93 (data not shown), whereas it is absent in crystal form C2-100 at the same contour level (left panel). **b**, Sequence alignment of the Pol I DNA-mimicking loop across different species highlighting the conservation of this element. **c**, Purified Pol I shows elongation activity in an RNA extension assay. DNA templates (Temp-41 and Temp-27 of 41 and 27 nucleotides, respectively) and ^{32}P-labelled RNA sequences used for the assay are indicated. The autoradiogram shows the elongation of RNA by Pol I producing a run-off of 18 nucleotides (lane 3) or 12 nucleotides (lane 6) depending of the template used. Lanes 1 and 4: the DNA/RNA hybrids were incubated in the absence of Pol I. Lanes 2 and 5: Pol I–DNA–RNA complexes were incubated with a buffer without NTPs. **d**, Dot spots grown at the indicated temperatures of the parental *RPA190* strain and *rpa190*Δloop strain where the DNA-mimicking loop has been deleted. The *rpa190*Δloop strain shows a slight temperature-sensitive growth defect on SDC medium.



Extended Data Figure 6 | Pol I dimer in the crystal lattice. The A43 C-terminal tail establishes crystal contacts with a second molecule related by a crystallographic dyad. The A43 C-terminal helix is embedded between the clamp and the protrusion domain of a dyad related molecule. The σ_A-weighted electron density map (contoured at 1σ) shows clear density corresponding to residues A43 251–326. The two monomers are related by a crystallographic dyad, which is indicated by a dyad symbol.



Extended Data Figure 7 | Subunit A12.2 structure and its position in Pol I. **a**, Detailed views show the A12.2 Zn sites and the main contacts between its linker and the A190 subunit. The A12.2 linker extends the β-sheet of the A190 jaw. **b**, The overlap between subunit A12.2 and α-amanitin in the Pol I structure explains the insensitivity of Pol I for this fungal toxin. The Pol II–α-amanitin complex structure (Protein Data Bank accession 2VUM) was superimposed onto the Pol I crystal structure. In the left panel, the α-amanitin toxin is depicted in surface representation (pink). On the right, a detailed view of α-amanitin shows the overlap with the C-terminal Zn ribbon of A12.2.



Extended Data Figure 8 | Precise positioning of the A49–A34.5 heterodimer suggests similar positions for the related C37–C53 heterodimer in Pol III and the TFIIF heterodimer in Pol II. **a**, Pol I was fitted into the Pol III envelope (EM-1804)[12]. A49–A34.5 (corresponding to C37–C53 in Pol III), AC40–AC19 and A12.2 (corresponding to C11 in Pol III) are coloured as in Fig. 1. The approximate position of subcomplex C82–C34–C31 is also indicated. **b**, The proposed Pol II/TFIIF model was manually fitted into the Pol II/TFII-A-B-F/TBP/DNA EM density (EM-2305)[43]. **c**, Left panel: detailed view of the anchoring of the A49–A34.5 dimerization domain onto the Pol I core. Right panel: model for the TFIIF dimerization module bound to the Pol II core based on the crystal structures of the human Rap74–Rap30 complex (Protein Data Bank accession 1F3U) and Pol II (Protein Data Bank accession 1WCM).

Extended Data Table 1 | Data statistics of the Pol I structure determination

	C2-90 SeMet	C2-90	C2-93	C2-100
Data collection				
Space group	*C2*	*C2*	*C2*	*C2*
Cell dimensions				
a, *b*, *c* (Å)	403.5, 140.4, 142.1	401.0, 139.8, 140.9	425.2, 140.6, 139.7	400.5, 140.2, 122.9
β (°)	90.0	90.4	93.4	100.1
Beamline	SOLEIL (Proxima 1)	SOLEIL (Proxima 1)	SOLEIL (Proxima 1)	DESY PETRA III (P14)
Resolution (Å)	71-3.59 (3.69-3.59)*	80-3.35 (3.44-3.35)*	90-3.27 (3.35-3.27) *	84-3.03 (3.11-3.03)*
R_{merge}†	0.192 (3.711)	0.193 (2.163)	0.162 (5.838)	0.104 (2.322)
$I/\sigma I$	11.2 (0.9)	8.5 (0.9)	10.9 (0.5)	12.3 (0.8)
Completeness (%)	99.6 (95.2)	100 (100)	100 (99.9)	99.1 (99.2)
Redundancy	14.3 (9.8)	7.7 (7.9)	10.9 (11.2)	7.1 (7.1)
CC(1/2)‡	0.999 (0.375)	0.997 (0.429)	0.999 (0.317)	0.999 (0.336)
Refinement				
PDB ID		4C3J	4C3H	4C3I
Resolution (Å)	N/A§	50-3.35	50-3.27	84-3.03
No. reflections		105911	124088	128861
R_{work}/ R_{free}		21.3/22.0	22.1/22.6	19.7/22.9
No. atoms				
Protein		34545	34545	34221
Ligand/ion		7 (Zn)	7 (Zn)	31 (including 7 Zn)
Water		0	0	0
B-factors				
Protein		135.1	171.9	119.9
Ligand/ion		128.3	160.4	134.1
R.m.s deviations				
Bond lengths (Å)		0.008	0.008	0.008
Bond angles (°)		0.96	0.96	0.96

* Highest resolution shell is shown in parentheses.
† Data were collected from one crystal.
‡ According to ref. 18. For C2-100 the highest resolution is 3.15 Å and 3.27 Å resolution using the criteria $CC_{1/2} > 0.5$ and $I/\sigma I > 2.0$, respectively.
§ Data set C2-90 SeMet was only used for anomalous difference map, not for model building.