Structural basis of RNA Polymerase III transcription initiation

Guillermo Abascal-Palacios[1], Ewan Phillip Ramsay[1], Fabienne Beuron[1], Edward Morris[1] & Alessandro Vannini[1*]

[1]The Institute of Cancer Research, London SW7 3RP, UK.

[*] Correspondence to A.V. (alessandro.vannini@icr.ac.uk)

SUMMARY

RNA Polymerase (Pol) III transcribes essential non-coding RNAs, including the entire pool of transfer RNAs (tRNAs), the 5S ribosomal RNA and the U6 spliceosomal RNA, and is often found deregulated in cancer cells. Initiation of Pol III gene transcription requires the activity of TFIIIB to form a transcriptionally active Pol III preinitiation complex (PIC). Here we present EM reconstructions of Pol III PICs at 3.4-4.0 Å and a reconstruction of unbound apo-Pol III at 3.1 Å. TFIIIB fully encircles the DNA and restructures Pol III. In particular, binding of the TFIIIB subunit Bdp1 rearranges Pol III specific subunits C37 and C34, promoting DNA opening. The unwound DNA directly contacts both sides of the Pol III cleft. Topologically, the Pol III PIC resembles the Pol II PIC, while the Pol I PIC is more divergent. The structures presented unravel the molecular mechanisms underlying the first steps of Pol III transcription, but also the general conserved mechanisms of gene transcription initiation.

INTRODUCTION

In the eukaryotic nucleus, RNA Polymerase (Pol) III catalyzes the DNA-dependent synthesis of short RNAs which are essential for cellular functions, such as transfer RNAs (tRNAs), the 5S ribosomal RNA (rRNA) and the U6 spliceosomal small nuclear RNA (snRNA). Pol III transcriptional output is a critical determinant of cellular and organismal growth[1] and misregulated Pol III activity plays a causative role in several diseases, including cancer[2-5].

Pol III is predominantly regulated at the level of transcription initiation. The assembly of a stable preinitiation complex (PIC), which is formed when Pol III is recruited to its target genes by a specific set of general transcription factors (GTFs), is a highly regulated process[6-8]. The TFIIIB complex is a multi-subunit GTF, which is ubiquitously required at Pol III genes *in vivo*[9-11] and is sufficient to form a transcriptionally-competent Pol III PIC *in vitro*[12-14]. TFIIIB consists of the TATA-box binding protein (TBP), the B-related factor 1 (Brf1), which is functionally related to the Pol II paralogue TFIIB[15], and B double prime (Bdp1)[16].

The 17-subunit Pol III is the most complex of the three eukaryotic polymerases. In addition to a 10-subunit enzymatic core and the 2-subunit peripheral stalk, which are modules conserved in both Pol I and Pol II, Pol III comprises five additional subunits, which are organized in subcomplexes. The C82/34/31 heterotrimeric subcomplex is essential for TFIIIB-dependent recruitment and for promoter opening[17], while the C53/37 heterodimeric subcomplex is involved in both initiation and termination of transcription[18,19].

Here we report the cryo-EM structures of an open Pol III PIC (OC-PIC) at 4.0 Å resolution, in which the DNA promoter has been spontaneously opened and the template strand engaged in the active site forming a full transcription bubble, a Pol III open complex (OC-POL3) at 3.4 Å resolution, which is similar to the OC-PIC but in which the upstream edge of the transcription bubble and TFIIIB are not resolved in the EM map, and finally the structure of unbound apo-Pol III (POL3) at 3.1 Å. Our results provide detailed insight into the mechanisms required for the initial steps of the Pol III transcription cycle, including DNA strand separation and template-strand engagement in the active site.

Visualization of the Pol III PIC

In order to obtain structural insights into Pol III transcription initiation, we purified by size-exclusion chromatography uncrosslinked Pol III PICs, assembled with a Brf1-TBP fusion

protein[20] and with a fully complementary DNA scaffold grafted from the U6 snRNA promoter (Methods, Fig. 1a).

Using a wild-type Pol III PIC, we obtained a reconstruction of the OC-PIC at an overall resolution of 4.0 Å (40,887 particles, 18.2% of the total) and a reconstruction of the OC-POL3 at an overall resolution of 3.7 Å (OC1-POL3, 101,484 particles, 47.4% of the total) (Fig. 1b, Extended Data Fig. 1a-b, Extended Data Fig. 2 a-c and Extended Data Fig. 3 a-b). For the Pol III PIC assembled in presence of Bdp1Δ (355-372), a Bdp1 mutant deficient in promoter opening[21], we obtained another OC-POL3 (OC2-POL3, 100,237 particles, 21.5% of the total) at an overall resolution of 3.4 Å and an unbound POL3 (178,779 particles, 38.3% of the total) at 3.1 Å (Extended Data Fig. 1c-e, Extended Data Fig. 2d-f and Extended Data Fig. 3 c-e). We further sub-classified the latter into cPOL3 and oPOL3 based on the closed or open conformation of the C82/34/31 subcomplex, at an overall resolution of 3.3 Å and 3.4 Å, respectively (Extended Data Figs. 1c-e and 3c-e). Because OC1- and OC2-POL3 are essentially identical and represent the same functional state, hereafter we will refer only to OC2-POL3, which was solved at higher resolution, as OC-POL3.

Cryo-EM maps were used to build and refine atomic models of the Pol III PICs (Extended Data Table 1). The structure of the OC-PIC reveals that the spontaneously formed transcription bubble is tightly stabilised by both TFIIIB and Pol III subunits (Fig. 1b). The Pol III clamp is in a closed state, in agreement with what was observed in Pol II PICs[22,23]. Upon formation of the OC-PIC, regions of the specific Pol III subunits C37, C34 and C31, which were mobile in the unbound and elongating structures of Pol III, are now visible in the cryo-EM maps (Fig. 1b, Extended Data Fig. 4). The structure of OC-POL3 suggests that once the DNA is melted and the template strand loaded in the active site, Pol III retains the downstream edge of the bubble in a correct orientation, even when the main interactions between Pol III and TFIIIB are disrupted (Extended Data Fig. 1b, c), in agreement with

previous studies[24]. This mechanism likely allows Pol III to maintain a correctly oriented bubble during promoter escape.

TFIIIB structure within the Pol III PIC

TFIIIB is centered around the TATA-box, which is recognized specifically by TBP (Fig. 2a). The architecture of the N-terminal part of Brf1 (residues 1-264) closely resembles that of TFIIB in the Pol II PIC[22,23]. The Brf1 Zn-ribbon domain is inserted through the active site cleft at the Pol III dock domain while the Brf1 C-terminal cyclin repeat is located between the Pol III wall and protrusion (Fig. 2a, Extended Data Fig. 4a,b). The linker region between the Brf1 Zn-ribbon and the N-terminal cyclin repeat, which in Pol II plays a role in stabilizing the template strand[22], is disordered in our structure. The Brf1 homology region I (HRI) includes a helical element, the "Brf1 helical pin", that directly contacts the DNA, TBP and the Brf1 C-terminal cyclin repeat (Fig. 2a and Extended Data Fig. 4a,b). The Brf1 helical pin sequence "PPSF/Y" is strictly conserved in eukaryotes and it is structurally homologous to the Brf2 "molecular pin" (amino acidic sequence "PPCML"), stabilizing the Brf1-TBP/DNA complex albeit without redox-sensing function[2] (Extended Data Fig. 5a, b). As previously described, the Brf1 homology region II (HRII) wraps around TBP and runs parallel to the promoter DNA[25] (Fig. 2a and Extended Data Fig. 4a,b). The Brf1 homology region III (HRIII), which forms a cryptic DNA binding domain[26], is not resolved in our structure.

Bdp1 adopts an intricate fold in the Pol III PIC (Fig. 2a and Extended Data Fig. 4a, c). The Bdp1 extended SANT domain and the Bdp1 linker bind to the major and the minor groove of the DNA, respectively, as observed in a human Bdp1-Brf2-TBP/DNA complex[27]. At its N-terminus, the Bdp1 linker extends towards the Pol III protrusion and forms a Bdp1 "tether" that interacts with the C37 "termination/initiation loop", which is disordered in cPOL3 and elongating Pol III[28] (Fig. 2b and Extended Data Fig. 4a, c, d). The Bdp1 tether and the C37 termination/initiation loop together form a composite "Bdp1/C37 platform" (Fig. 2b) that aids

the positioning and docking of the first and second winged helix domains (WHD1 and WHD2) of subunit C34 (Extended Data Fig. 4a, e), which are also stabilized only upon PIC formation (Fig. 2a, b), in agreement with previous photo-crosslinking studies[29,30]. The OC-PIC structure reveals that the inactive Bdp1 deletion mutants which fail to open the promoter[21] map to the Bdp1 tether and linker regions. Thus, the common transcriptional defect of these mutants is attributable to the disruption of the correct topology and architecture of the Bdp1/C37 platform. Supporting this evidence, a Pol III enzyme which completely lacks subunit C37 displays a similar transcription initiation defect[18] and mutations of charged residues in the C34 WHD2 impair binding to TFIIIB and promoter opening[17].

N-terminally to the Bdp1 tether, two α-helices form the "Bdp1 clip", a scissor-like element which binds the DNA in the major groove and contacts the Brf1 C-terminal cyclin repeat and the Bdp1 SANT domain (Fig. 2a, b and Extended Data Fig. 4a, c). At the C-terminus of the extended Bdp1 SANT domain, Bdp1 folds into a long α-helix, which we named the "Bdp1 stem" that participates, together with the Brf1 HRII and the Bdp1/C37 platform, in the rigid positioning of the C34 WHD2 over the cleft (Fig. 2a, b and Extended Data Fig. 4a, c, e).

TFIIIB forms an extensive number of contacts with the DNA and completely encloses the DNA promoter, explaining its unusually long half-life[27,31]. Upstream and downstream regions flanking the TATA box are bridged by TFIIIB, which acts as a molecular wrench pivoting around the C34 WHD2 and the Pol III protrusion.

DNA opening and template strand loading

In Pol III PICs, the formation of a transcription bubble occurs in a non-coordinated manner, with strand separation nucleating at the upstream edge of the transcription bubble and subsequently propagating downstream to the TSS[24]. In our structure of the OC-PIC, the upstream edge of the transcription bubble is at nucleobase -12, in close agreement with biochemical data[24] and with the structure of the human open Pol II PIC[22] (Fig. 1a, b). The

template and non-template strands are uncoupled by direct contacts with Pol III subunits (Fig. 3a, b). The unwound template strand is in close proximity to an exposed hydrophobic pocket on the clamp helices, which we named the "template strand pocket" (Fig. 3a). The template strand pocket is formed by Pol III residues W294, L298 and Y318, which are highly conserved, or substituted with chemically equivalent aminoacids, across the eukaryotic kingdom (Fig. 3a and Extended Data Fig. 5c). The nucleobase -9 establishes Van der Waals interactions with W294, which thus stabilize the unwound DNA (Fig. 3a). The template strand pocket is specific for Pol III since these residues are not conserved in Pol I and Pol II, except for L298 in Pol I (Extended Data Fig. 5c). Interestingly, part of the C160 “rudder" (Extended Data Fig. 4a), which is disordered in the OC-PIC, is stabilised in the structure of elongating Pol III[28], limiting the access of nucleic acids to the template strand pocket.

Concomitantly, the non-template strand is embraced between the C82 "cleft loop" and the Pol III C128 "lobe tip" (Fig. 3b and Extended Data Fig. 4a). Recognition and stabilization of the unwound non-template strand by Pol III lobe tip also rationalizes the role of this region during transcription termination. Deletion mutants of the lobe tip display a termination read-through phenotype[32], suggesting a direct role for this region in stabilizing the terminator signal, a stretch of five or more thymidine residues on the non-template strand. Sensing of the terminator signal is likely carried out together with the C37 termination/initiation loop, another hot-spot of termination read-through mutants[33], which lays in close proximity (Fig. 3b).

An extensive contact is formed between the DNA around positions -20 to -15 and a helix-turn-helix motif of Brf1, which is conserved in TFIIB[22] (Fig. 3a). Opposite to this site, the C34 WHD2 inserts a canonical recognition helix in the major groove of the DNA and the "wing" at the site where the DNA is melted, stabilizing the unwound structure (Fig. 3a, b).

148 The protrusion contacts the DNA backbone around register -19 and interacts with the C34
149 WHD2 recognition helix (Fig. 3a).
150 Once the DNA is unwound at the upstream edge of the transcription bubble, the template
151 strand is engaged at the polymerase active site while the bubble is extended downstream.
152 Comparison of the OC-PIC and POL3 structures reveals that, before PIC assembly, the path
153 for loading the template is obstructed by a loop (residues 1042-1061 of subunit C128) which
154 is part of the switch 3 region (Fig. 3c and Extended Data Fig. 4a). In the OC-PIC, the binding
155 of the Brf1 Zn-ribbon reconfigures the switch 3 loop in an open conformation, as observed in
156 elongating Pol II, which is capable of directly stabilizing the template strand (Fig. 3c).
157 Remarkably, this structural rearrangement can explain the transcriptional defects observed for
158 Brf1 N-ribbon deletion mutant *in vitro* and *in vivo*[21]. Specifically, this Brf1 mutant is
159 transcriptionally inactive due to its inability to extend the upstream transcription bubble but
160 can be rescued by pre-opening the nucleic acid scaffolds at the downstream edge of the
161 prospective bubble[21]. Thus, binding of the Brf1 Zn-ribbon induces a structural rearrangement
162 that clears the cleft, facilitating loading of the template strand at the active site and formation
163 of a full-size transcription bubble (Fig. 3c). The switch 3 loop adopts the same open
164 conformation in the structure of the elongating Pol III[28], suggesting that the structural
165 changes occurring in the cleft prepare the OC-PIC for the elongation step.
166 Stabilization of the downstream edge of the bubble, which is in proximity of the active site, is
167 reminiscent of the one observed in the Pol II PICs[22] (Fig. 3d). The fork-loop 2 (Extended
168 Data Fig. 4a) maintains the downstream transcription bubble by forming a wedge at the
169 branching point, a universal function which seems to be conserved from bacterial to
170 eukaryotic PICs[22]. The bridge helix, which in the unbound POL3 structure is kinked similarly
171 to the paused or inhibited bacterial polymerase[34,35], is completely folded (Extended Data Fig.
172 1 and 4a) and interacts with the template strand through residue Y884, which is conserved in

Pol I and Pol II (Fig. 3d and Extended Data Fig. 5d). The switch2 is disordered in the POL3 structures but folds into a helical element in the OC-PIC and interacts with the template strand (Fig. 3d and Extended Data Fig. 4a). The structural changes responsible for the stabilization of the downstream edge of the transcription bubble are also observed in the structure of elongating Pol III[28], reinforcing the idea that the OC-PIC is primed for elongation.

Sensing the open complex formation

Opening of the promoter DNA is mediated by a cascade of structural changes involving several regions of Pol III and TFIIIB, raising the question of how promoter opening is detected and how the structural rearrangements are co-ordinated.

Upon TFIIIB binding, stabilization of the C34 WHD2 over the cleft and DNA melting at the upstream edge of the transcription bubble are accompanied by a contraction of the clamp helices (Extended Data Fig. 4a, f), as compared to the POL3 or elongating Pol III structures[28] (Fig. 4, rearrangement I), in order to stabilize the unwound template strand. As a result, the C82/34/31 subcomplex, which lays on top of the clamp helices, is shifted towards the cleft (Fig. 4, rearrangement II) and undergoes a structural rearrangement that culminates in the stabilization of the C-terminal segment of subunit C34 (Fig. 4, rearrangement II and Extended Data Fig. 4a, e) and subunit C31 (Fig. 4, rearrangement III and Extended Data Fig. 4a, g), which are disordered in POL3 and the elongating Pol III[28] structures. Remarkably, the C31 "stalk bridge" folds into a α-helix that directly contacts the stalk, locking it at a defined angle from the C82/34/31 subcomplex (Fig. 4, rearrangement III). The Pol III stalk is tightly anchored to the Pol III core and directly contacts the clamp through a region which is stabilised in our OC-PIC (Fig. 4, rearrangement IV). The Pol III stalk is mobile and can relay structural changes to the clamp[28], a mechanism shared with the archaeal and Pol II

197 machineries[36,37]. Similarly, in these systems TFE and TFIIE directly bridge the stalk to the
198 clamp through a Zn-ribbon domain[22,23,38].
199 Taken together, our data suggest that formation of the open PIC is sensed through concerted
200 structural rearrangements that lock the Pol III clamp in a closed state, resulting in the
201 stabilisation of the downstream edge of the transcription bubble (Fig. 4, Supplemental Video
202 S1). We observe the same arrangement of the C82/34/31 subcomplex, the stalk and the clamp
203 in the OC-POL3 structure, suggesting that the locking mechanism of the clamp in a closed
204 conformation allows Pol III to tightly hold the downstream edge of the transcription bubble
205 (Extended Data Fig. 1b, c).

206 **Comparison of eukaryotic PICs**

207 The topology and architecture of the Pol III PIC closely resembles that of the Pol II PICs[22,23],
208 while the Pol I PIC is more dissimilar as it has evolved and specialized for the transcription
209 of a single gene (Extended Data Fig. 6a)[39-41]. The relative orientation of the polymerase with
210 respect to TBP is maintained due to the architectural and functional conservation of Brf1 and
211 TFIIB, which physically bridge the polymerase to TBP. The upstream and downstream edges
212 of the transcription bubble are placed in the same locations (Extended Data Fig. 6a). A
213 notable difference is observed in the upstream DNA pathway, as a result of the reduced TBP-
214 induced bending of the DNA caused by the binding of the Bdp1 clip domain (Extended Data
215 Fig. 6b).
216 The different architecture of the protrusion represents a remarkable difference among the 10-
217 subunit core of the three polymerases, which has implications in the assembly of the
218 eukaryotic PICs. While Pol I and Pol III utilize regions around the protrusion tip to directly
219 contact the promoter DNA, albeit Pol I does so in a more extensive manner, the Pol II
220 protrusion tip does not contact the DNA directly (Extended Data Fig. 6c). Pol III directly
221 contacts the promoter through the conserved residue K409 of subunit C128.

In the Pol II PIC, the TFIIF Tfg2 (RAP30 in humans) WHD binds and stabilizes the promoter DNA close to the upstream edge of the transcription bubble, in a manner which partially resembles C34 WHD2 (Fig. 5). However, while C34 WHD2 is directly involved in stabilizing the DNA at the site where strand separation occurs, the TFIIF Tfg2 WHD binds more upstream. From the opposite site, the TFIIE E-wing protrudes from the tip of the clamp helices towards the DNA. This structural element more strongly resembles the C82 cleft loop, which is however located more deeply into the cleft, stabilizing the non-template strand while the E-wing has been proposed to play an important role in strand separation[23] (Fig. 5). Taking also into account the functional conservation of TFIIE and C31 in physically bridging the stalk and the clamp, the C82/34/31 subcomplex can be regarded to as a TFIIF/TFIIE hybrid rather than simply a TFIIE-like subcomplex[15]. The strong structural homology between the dimerization domain of the TFIIF-like C53/37 subcomplex and TFIIF has been previously reported[28,42]. Comparison of the structures of the three eukaryotic PICs (Extended Data Fig. 6a) highlights the architectural and topological conservation between dissociable Pol II GTFs and stably associated Pol I and III specific subcomplexes, which acquired specific functional roles during evolution[15,43].

Conclusions

The structures of the OC-PIC and OC-POL3 provide insights into the molecular basis of specific promoter recognition and opening. Our results converge with previous biochemical data[21,24] on a model that requires a concerted allosteric mechanism, involving TFIIIB and Pol III-specific subunits, in order to form a transcriptionally-competent Pol III PIC (Extended Data Fig. 7, Fig. 4, Supplemental Video S1). The structural data unveil how Pol III opens the promoter DNA in the absence of ATP hydrolysis, using solely the binding energy generated by the large number of the newly established interactions formed during the assembly of an active PIC. Despite being structurally unrelated, the stabilization of unwound nucleobases by

aromatic residues, trapping of the non-template strand and movements of the clamp are all aspects shared with the bacterial transcription machinery, which also unwinds DNA in an ATP-independent manner[44,45]. Archaeal[46], Pol I[39-41] and certain Pol II promoters[23] can be also opened without ATP hydrolysis, suggesting that DNA opening occurs through similar mechanisms across the three kingdoms of life.

REFERENCES

1 Filer, D., Thompson, M. A., Takhaveev, V., Dobson, A. J., Kotronaki, I., Green, J. W. M., Heinemann, M., Tullet, J. M. A. and Alic, N. RNA polymerase III limits longevity downstream of TORC1. *Nature*, (2017), **doi:10.1038/nature25007**

2 Gouge, J., Satia, K., Guthertz, N., Widya, M., Thompson, A. J., Cousin, P., Dergai, O., Hernandez, N. and Vannini, A. Redox Signaling by the RNA Polymerase III TFIIB-Related Factor Brf2. *Cell* **163**, 1375-1387, (2015).

3 Goodarzi, H., Nguyen, H. C. B., Zhang, S., Dill, B. D., Molina, H. and Tavazoie, S. F. Modulated Expression of Specific tRNAs Drives Gene Expression and Cancer Progression. *Cell* **165**, 1416-1427, (2016).

4 Dauwerse, J. G., Dixon, J., Seland, S., Ruivenkamp, C. A., van Haeringen, A., Hoefsloot, L. H., Peters, D. J., Boers, A. C., Daumer-Haas, C., Maiwald, R., Zweier, C., Kerr, B., Cobo, A. M., Toral, J. F., Hoogeboom, A. J., Lohmann, D. R., Hehr, U., Dixon, M. J., Breuning, M. H. and Wieczorek, D. Mutations in genes encoding subunits of RNA polymerases I and III cause Treacher Collins syndrome. *Nat Genet* **43**, 20-22, (2011).

5 Thiffault, I., Wolf, N. I., Forget, D., Guerrero, K., Tran, L. T., Choquet, K., Lavallee-Adam, M., Poitras, C., Brais, B., Yoon, G., Sztriha, L., Webster, R. I., Timmann, D., van de Warrenburg, B. P., Seeger, J., Zimmermann, A., Mate, A., Goizet, C., Fung, E., van der Knaap, M. S., Fribourg, S., Vanderver, A., Simons, C., Taft, R. J., Yates, J. R., 3rd, Coulombe, B. and Bernard, G. Recessive mutations in POLR1C cause a leukodystrophy by impairing biogenesis of RNA polymerase III. *Nat Commun* **6**, 7623, (2015).

6 Lee, J., Moir, R. D. and Willis, I. M. Differential Phosphorylation of RNA Polymerase III and the Initiation Factor TFIIIB in Saccharomyces cerevisiae. *PLoS One* **10**, e0127225, (2015).

7 Fairley, J. A., Mitchell, L. E., Berg, T., Kenneth, N. S., von Schubert, C., Silljé, H. H. W., Medema, R. H., Nigg, E. A. and White, R. J. Direct Regulation of tRNA and 5S rRNA Gene Transcription by Polo-like Kinase 1. *Molecular Cell*, (2012).

8 Vannini, A., Ringel, R., Kusser, A. G., Berninghausen, O., Kassavetis, G. A. and Cramer, P. Molecular Basis of RNA Polymerase III Transcription Repression by Maf1. *Cell* **143**, 59-70, (2010).

9 Harismendy, O., Gendrel, C.-G., Soularue, P., Gidrol, X., Sentenac, A., Werner, M. and Lefebvre, O. Genome-wide location of yeast RNA polymerase III transcription machinery. *The EMBO Journal* **22**, 4738-4747, (2003).

10 Moqtaderi, Z., Wang, J., Raha, D., White, R. J., Snyder, M., Weng, Z. and Struhl, K. Genomic binding profiles of functionally distinct RNA polymerase III transcription complexes in human cells. *Nature Structural & Molecular Biology* **17**, 635-640, (2010).

11 Barski, A., Chepelev, I., Liko, D., Cuddapah, S., Fleming, A. B., Birch, J., Cui, K., White, R. J. and Zhao, K. Pol II and its associated epigenetic marks are present at Pol III–transcribed noncoding RNA genes. *Nature Publishing Group* **17**, 629-634, (2010).

12 Kassavetis, G. A., Letts, G. A. and Geiduschek, E. P. A minimal RNA polymerase III transcription system. *The EMBO Journal* **18**, 5042-5051, (1999).

13 Kassavetis, G. A., Braun, B. R., Nguyen, L. H. and Geiduschek, E. P. S. cerevisiae TFIIIB is the transcription initiation factor proper of RNA polymerase III, while TFIIIA and TFIIIC are assembly factors. *Cell* **60**, 235-245, (1990).

14 Dieci, G., Percudani, R., Giuliodori, S., Bottarelli, L. and Ottonello, S. TFIIIC-independent in vitro transcription of yeast tRNA genes. *Journal of Molecular Biology* **299**, 601-613, (2000).

15 Vannini, A. and Cramer, P. Conservation between the RNA Polymerase I, II, and III Transcription Initiation Machineries. *Molecular Cell* **45**, 439-446, (2012).

16 Ishiguro, A., Kassavetis, G. A. and Geiduschek, E. P. Essential roles of Bdp1, a subunit of RNA polymerase III initiation factor TFIIIB, in transcription and tRNA processing. *Molecular and cellular biology* **22**, 3264-3275, (2002).

17 Brun, I., Sentenac, A. and Werner, M. Dual role of the C34 subunit of RNA polymerase III in transcription initiation. *The EMBO Journal* **16**, 5730-5741, (1997).

18 Kassavetis, G. A., Prakash, P. and Shim, E. The C53/C37 subcomplex of RNA polymerase III lies near the active site and participates in promoter opening. *Journal of Biological Chemistry* **285**, 2695-2706, (2010).

19 Arimbasseri, A. G. and Maraia, R. J. Mechanism of Transcription Termination by RNA Polymerase III Utilizes a Non-template Strand Sequence-Specific Signal Element. *Mol Cell* **58**, 1124-1132, (2015).

20 Kassavetis, G. A., Soragni, E., Driscoll, R. and Geiduschek, E. P. Reconfiguring the connectivity of a multiprotein complex: fusions of yeast TATA-binding protein with Brf1, and the function of transcription factor IIIB. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 15406-15411, (2005).

21 Kassavetis, G. A., Letts, G. A. and Geiduschek, E. P. The RNA polymerase III transcription initiation factor TFIIIB participates in two steps of promoter opening. *EMBO J* **20**, 2823-2834, (2001).

22 He, Y., Yan, C., Fang, J., Inouye, C., Tjian, R., Ivanov, I. and Nogales, E. Near-atomic resolution visualization of human transcription promoter opening. *Nature* **533**, 359-365, (2016).

23 Plaschka, C., Hantsche, M., Dienemann, C., Burzinski, C., Plitzko, J. and Cramer, P. Transcription initiation complex structures elucidate DNA opening. *Nature* **533**, 353-358, (2016).

24 Kassavetis, G. A., Blanco, J. A., Johnson, T. E. and Geiduschek, E. P. Formation of open and elongating transcription complexes by RNA polymerase III. *J Mol Biol* **226**, 47-58, (1992).

25 Juo, Z. S., Kassavetis, G. A., Wang, J., Geiduschek, E. P. and Sigler, P. B. Crystal structure of a transcription factor IIIB core interface ternary complex. *Nature* **422**, 534-539, (2003).

26 Huet, J., Conesa, C., Carles, C. and Sentenac, A. A cryptic DNA binding domain at the COOH terminus of TFIIIB70 affects formation, stability, and function of preinitiation complexes. *The Journal of biological chemistry* **272**, 18341-18349, (1997).

27 Gouge, J., Guthertz, N., Kramm, K., Dergai, O., Abascal-Palacios, G., Satia, K., Cousin, P., Hernandez, N., Grohmann, D. and Vannini, A. Molecular mechanisms of Bdp1 in TFIIIB assembly and RNA polymerase III transcription initiation. *Nat Commun* **8**, 130, (2017).

28 Hoffmann, N. A., Jakobi, A. J., Moreno-Morcillo, M., Glatt, S., Kosinski, J., Hagen, W. J., Sachse, C. and Muller, C. W. Molecular structures of unbound and transcribing RNA polymerase III. *Nature* **528**, 231-236, (2015).

29 Hu, H. L., Wu, C. C., Lee, J. C. and Chen, H. T. A Region of Bdp1 Necessary for Transcription Initiation That Is Located within the RNA Polymerase III Active Site Cleft. *Mol Cell Biol* **35**, 2831-2840, (2015).

30 Wu, C. C., Lin, Y. C. and Chen, H. T. The TFIIF-Like Rpc37/53 Dimer Lies at the Center of a Protein Network To Connect TFIIIC, Bdp1, and the RNA Polymerase III Active Center. *Molecular and cellular biology* **31**, 2715-2728, (2011).

31 Cloutier, T. E., Librizzi, M. D., Mollah, A. K., Brenowitz, M. and Willis, I. M. Kinetic trapping of DNA by transcription factor IIIB. *Proc Natl Acad Sci U S A* **98**, 9581-9586, (2001).

32 Shaaban, S. A., Krupp, B. M. and Hall, B. D. Termination-altering mutations in the second-largest subunit of yeast RNA polymerase III. *Mol Cell Biol* **15**, 1467-1478, (1995).

33 Rijal, K. and Maraia, R. J. RNA polymerase III mutants in TFIIF -like C37 that cause terminator readthrough with no decrease in transcription output. *Nucleic acids research*, (2012).

34 Weixlbaumer, A., Leon, K., Landick, R. and Darst, S. A. Structural basis of transcriptional pausing in bacteria. *Cell* **152**, 431-441, (2013).

35 Tagami, S., Sekine, S., Kumarevel, T., Hino, N., Murayama, Y., Kamegamori, S., Yamamoto, M., Sakamoto, K. and Yokoyama, S. Crystal structure of bacterial RNA polymerase bound with a transcription inhibitor protein. *Nature* **468**, 978-982, (2010).

36 He, Y., Fang, J., Taatjes, D. J. and Nogales, E. Structural visualization of key steps in human transcription initiation. *Nature* **495**, 481-486, (2013).

37 Schulz, S., Gietl, A., Smollett, K., Tinnefeld, P., Werner, F. and Grohmann, D. TFE and Spt4/5 open and close the RNA polymerase clamp during the transcription cycle. *Proc Natl Acad Sci U S A* **113**, E1816-1825, (2016).

38 Jun, S. H., Hirata, A., Kanai, T., Santangelo, T. J., Imanaka, T. and Murakami, K. S. The X-ray crystal structure of the euryarchaeal RNA polymerase in an open-clamp configuration. *Nat Commun* **5**, 5132, (2014).

39 Engel, C., Gubbey, T., Neyer, S., Sainsbury, S., Oberthuer, C., Baejen, C., Bernecky, C. and Cramer, P. Structural Basis of RNA Polymerase I Transcription Initiation. *Cell* **169**, 120-131 e122, (2017).

40 Han, Y., Yan, C., Nguyen, T. H. D., Jackobel, A. J., Ivanov, I., Knutson, B. A. and He, Y. Structural mechanism of ATP-independent transcription initiation by RNA polymerase I. *Elife* **6**, (2017).

41 Sadian, Y., Tafur, L., Kosinski, J., Jakobi, A. J., Wetzel, R., Buczak, K., Hagen, W. J., Beck, M., Sachse, C. and Muller, C. W. Structural insights into transcription initiation by yeast RNA polymerase I. *EMBO J*, (2017).

42 Geiger, S. R., Lorenzen, K., Schreieck, A., Hanecker, P., Kostrewa, D., Heck, A. J. R. and Cramer, P. RNA polymerase I contains a TFIIF-related DNA-binding subcomplex. *Molecular Cell* **39**, 583-594, (2010).

43 Carter, R. and 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. *Molecular biology and evolution* **27**, 1035-1043, (2010).

44 Chakraborty, A., Wang, D., Ebright, Y. W., Korlann, Y., Kortkhonjia, E., Kim, T., Chowdhury, S., Wigneshweraraj, S., Irschik, H., Jansen, R., Nixon, B. T., Knight, J., Weiss, S. and Ebright, R. H. Opening and Closing of the Bacterial RNA Polymerase Clamp. *Science (New York, NY)* **337**, 591-595, (2012).

45 Zhang, Y., Feng, Y., Chatterjee, S., Tuske, S., Ho, M. X., Arnold, E. and Ebright, R. H. Structural basis of transcription initiation. *Science* **338**, 1076-1080, (2012).

46 Hausner, W. and Thomm, M. Events during initiation of archaeal transcription: open complex formation and DNA-protein interactions. *J Bacteriol* **183**, 3025-3031, (2001).

FIGURE LEGENDS

Figure 1. Cryo-EM structure of the *S. cerevisiae* RNA Polymerase III pre-initiation complex. a, DNA nucleotides modelled in the OC-PIC structure are depicted as solid circles and are numbered relative to TSS of the U6 gene. Mutation of nucleotide -24 to break the pseudo-symmetry of the TATA-box is coloured in red. The position of the active site magnesium ion is indicated as a magenta sphere. The TATA-box is highlighted with a grey box. Protein-DNA interactions are highlighted. **b,** Front and top view of yeast Pol III PIC structure (OC-PIC). Pol III core subunits are depicted in grey. Template and non-template strands of the DNA are shown in dark and light blue, respectively. Cryo-EM maps are shown as transparent surfaces. The colour code is used throughout.

Figure 2. Architecture of TFIIIB subunits and upstream DNA assembly. a, Position of TFIIIB transcription factor and Pol III interacting subunits. Brf1, TBP, Bdp1, c37 and C34 are depicted as ribbon models while Pol III protrusion and wall domains are depicted as molecular surfaces. **b,** Detailed view of the Bdp1/C37 platform. Bdp1, C37 and C34 subunits are depicted as ribbon models in the cryo-EM map.

Figure 3. Promoter opening and template strand loading. a, Detail of upstream promoter DNA opening. Side chains of residues involved in the template strand pocket are shown in stick representation. **b,** The non-template strand is threaded between the C82 cleft loop and C128 lobe tip. **c,** Detail of template DNA strand loading into the active site. The Brf1 Zinc-ribbon domain reconfigures Pol III switch loop 3 from a closed (cPOL3, wheat) to an 'open' state (OC-PIC, grey). DNA nucleotides are numbered relative to the TSS. **d,** Stabilisation of the DNA bubble downstream edge in the OC-PIC. Conserved Y884 residue is shown in stick representation.

Figure 4. Structural rearrangements upon open complex formation. Subunits involved in conformation changes are depicted as solid ribbon models. The core of Pol III is represented as a grey transparent surface. The numbering and close-up views refer to specific domain rearrangements occurring upon open PIC formation, which involve motions (dashed arrows), folding (dashed ellipsoids) and establishment of new contacts with neighbouring subunits (solid line with circles).

Figure 5. The C82/C34/C31 subcomplex is a TFIIF/TFIIE hybrid. The upstream region of the DNA bubble is stabilised by the winged-helix domains of C34 subunit in a similar manner to that of TFIIF subunits Tfg2 (wheat) and RAP30 (pink) in yeast and human, respectively. The ssDNA is stabilised by the C82 cleft loop, which resembles the yeast TFIIE E-wing element.

METHODS

Protein Expression and Purification. Endogenous Pol III, carrying a C-terminal Tap-tag on subunit AC40, was obtained through large-scale fermentation of a *S. cerevisiae* SC1613 strain (Euroscarf) in a BioFlo Pro 75l Fermentor (New Brunswick Scientific). After overnight growth, the yeast culture was harvested at OD_{600} of ~10 and the cells were flash-frozen in liquid nitrogen. The sample was then subjected to lysis using a 6870D Freezer/Mill Cryogenic Grinder (4 cycles, 2 min run, 2 min cool, 15 cps, SPEX Sample Prep). After cell lysis, the sample was resuspended in Buffer A (50 mM Tris-HCl pH 8, 20% glycerol, 250 mM $(NH_4)_2SO_4$, 1 mM EDTA, 10 mM $MgCl_2$, 10 µM $ZnCl_2$, 12 mM BME, 1 mM PMSF and 2 mM benzamidine) and cleared by double centrifugation at 38000 x g. for 40 min. The soluble fraction was then filtered and loaded into a gravity column containing ~125 ml of Heparin Sepharose 6 Fast Flow resin (GE Healthcare) equilibrated with buffer A. Next, the resin was washed with Buffer B (Buffer A + 0.5 mM EDTA, 1 mM $MgCl_2$, 10 µM $ZnCl_2$ and

1 mM BME in the absence of glycerol) and elution was performed in Buffer C (Buffer B + 1 M $(NH_4)_2SO_4$). Then, the sample was diluted to 500 mM $(NH_4)_2SO_4$ and incubated overnight at 4ºC with 10 ml IgG Sepharose 6 Fast Flow resin (GE Healthcare). After washing, the beads were incubated with Tobacco Etch Virus protease (TEV) for 6 h at 4ºC. The TAP-tag cleaved sample was collected, diluted to ~60 mM $(NH_4)_2SO_4$ and loaded into a MonoQ 5/50 GL column (GE Healthcare). Elution was performed through a linear gradient from 60 mM $(NH_4)_2SO_4$ to 1 M $(NH_4)_2SO_4$ in 40 mM Tris-HCl pH 8, 0.5 mM EDTA, 1 mM $MgCl_2$, 10 µM $ZnCl_2$, 10 mM DTT, 1 mM PMSF and 2 mM benzamidine. Two peaks corresponding to RNA Polymerase I and RNA Polymerase III eluted at ~340 mM $(NH_4)_2SO_4$ and ~470 mM $(NH_4)_2SO_4$, respectively. The fractions corresponding to Pol III were collected, supplemented with 10% glycerol and stored at -80ºC.

In our reconstitution, we used a Brf1-TBP fusion protein, which has been shown to functionally substitute for Brf1 and TBP *in vitro* and *in vivo*[20]. Purification of the Brf1-TBP fusion protein was performed as previously described[20].

The wild type Bdp1 protein or the transcriptionally inactive Bdp1Δ(355-372) mutant, which is unable to nucleate promoter DNA opening[21], were used for reconstitution of Pol III PICs. A pOPINJ plasmid containing *S. cerevisiae* Bdp1 genes was transformed into Rosetta™ (DE3) pLysS cells and grown in LSSB media at 37ºC to an OD_{600} of ~0.7. The culture was then cooled at 4ºC for 1 h and induction was performed O/N with 1 mM IPTG at 15ºC. The harvested cells were suspended in a buffer containing 10 mM Na_2HPO_4, 2 mM KH_2PO_4, 500 mM NaCl, 7 mM $MgCl_2$, 10 mM EDTA, 10% glycerol, 0.5% NP-40, 2.5 mM betaine, 10 mM BME, 0.5 mg/ml lysozyme, two protein inhibitor tablets (Roche) and a scoop of DNAse I, and incubated at 4ºC for 30 min. The cell suspension was subjected to sonication (12 cycles, 15 sec ON, 59 sec OFF, 60% amplitude) and the lysate was fractionated by centrifugation at 48000 x g. and 4ºC for 40 min. Next, the soluble fraction was filtered and

incubated for 2.5 h at 4ºC with ~5 ml Glutathione Sepharose 4 Fast Flow resin (GE Healthcare). Then, the beads were washed with 300 ml of lysis buffer and the protein was eluted in 20 ml of lysis buffer supplemented with 50 mM of reduced glutathione. The collected sample was diluted to 100 mM NaCl with buffer QA (20 mM Tris-HCl pH 8.8, 0.2 mM EDTA, 10% glycerol and 1 mM DTT) and loaded into a HiTrap Q HP 5 ml column (GE Healthcare) equilibrated with 5% of buffer QB (buffer QA supplemented with 2 M NaCl). After washing, the sample was eluted through a linear gradient from 5% to 100% of buffer QB in 30 CV. The elution was analysed through SDS-PAGE and fractions containing GST-ScBdp1 were pooled and incubated O/N with 3C protease at 4ºC. The cleaved sample was concentrated to 5 ml and loaded in a HiLoad 16/600 Superdex 200 pg column equilibrated with 20 mM Tris-HCl pH 7.5, 150 mM NaCl and 2 mM DTT. The fractions corresponding to pure ScBdp1 were pooled, concentrated to ~2.5 mg/ml, flash-frozen in liquid nitrogen and stored at -80ºC.

DNA oligonucleotides. The assembly of the Pol III PIC was performed using 70-nt complementary oligonucleotides based on the yeast U6 promoter (Template strand: CGAAGGGTTACTTCGCGAACACATAGTTGCGAAAAAAACATTTTTTTATAGTAGC CGAAAATAGTGGACG and Non-template Strand: CGTCCACTATTTTCGGCTACTATAAAAAAATGTTTTTTTCGCAACTATGTGTTCGC GAAGTAACCCTTCG; Integrated DNA technologies (IDT)). Because of the pseudo-symmetrical nature of the U6 snRNA TATA-box, which drives transcription in both directions in absence of TFIIIC[12], we mutated a single nucleotide of the TATA-box at position -24 from the transcriptional start site (TSS) in order to favour unidirectional positioning of TFIIIB and avoid heterogeneity, as described in the TBP-Brf1(437-506)/DNA crystal structure[25] (Fig. 1a). The oligonucleotides were suspended in annealing buffer (50 mM Tris-HCl pH 8, 1 mM EDTA and 5 mM $MgCl_2$) to a final concentration of ~1 µM,

mixed in equimolar concentration and heated at 95ºC for 5 min. In order to obtain dsDNA, the sample was cooled down to 10ºC at a rate of 1ºC/min.

Pol III PIC assembly. The formation of the Pol III PIC was carried out using 200 µg of Pol III. First, the Brf1-TBP fusion protein and the annealed DNA were mixed in a 5:2.5 molar ratio (relative to Pol III) and incubated at RT for 30 min. Next, 5-fold excess of Bdp1 was added to the mixture and subjected to a similar incubation process. Then, Pol III was added and the mixture was diluted 12x with buffer containing 50 mM Tris pH 8, 10% glycerol, 3 mM DTT and 10 mM $MgCl_2$, in order to reduce the salt concentration. The sample was concentrated to ~300 µl with a Vivaspin 6 50K MWCO spin concentrator and incubated at RT for 30 min. Finally, the sample was applied to a Superose 6 10/300 GL column (GE Healthcare) equilibrated with 40 mM Tris-HCl pH 8, 80 mM NaCl, 3 mM DTT and 10 mM $MgCl_2$ and collected in 100 µl fractions. Analysis of the elution peaks by SDS-PAGE and silver staining indicated the presence of the Pol III and the transcription factors in the first peak. The fractions corresponding to the front of the peak were collected, concentrated to ~0.1 mg/ml and used in the subsequent EM analysis. In absence of Bdp1, a Pol III PIC could not be assembled using the conditions specified above.

Cryo-EM data collection. Cryo-EM samples were prepared in Quantifoil R 1.2/1.3 copper grids for wild type Bdp1 (WT-PIC) sample or Quantifoil R2/2 Molybdenum grids for the Bdp1Δ(355-372) (Bdp1Δ-PIC) sample coated with a thin homemade carbon film. The grids were glow discharged for 20 sec at 15 mA using a PELCO EasyGlow glow discharger before sample addition. 2 µl of sample was added to the grids and incubated for 30 sec at 18ºC and 100% humidity. Then, the grid was blotted (drain time: 0.5 sec, blot force: 13, blot time: 4-5 sec) and plunge-frozen into liquid ethane using a VitroBot Mark IV (FEI) system.

Data was collected on a Titan Krios (FEI) transmission electron microscope at 300 keV using a Gatan Quantum energy filter and a K2 Summit direct detector. For the WT-PIC dataset,

519 4731 movies were collected at a 1.06 Å calibrated pixel size and a rate of 6 frames per
520 second. 24 frames were collected per movie within a defocus range from -1.6 µm to -3.4 µm
521 and at a dose rate of 6.5 $e^-/Å^2$/sec, which provided a total accumulated dose of 39 $e^-/Å^2$
522 (Extended Data Fig. 2a, Extended Data Table 1). For the Bdp1Δ-PIC dataset, a total of 6220
523 micrographs were collected in super-resolution mode (0.5269 Å pixel size) at a rate of 1.25
524 frames per second. 20 frames were collected for each movie within a similar defocus range
525 than the WT dataset and at a dose rate of 2.52 $e^-/Å^2$/sec, accumulating a total of ~40 $e^-/Å^2$
526 (Extended Data Fig. 2d, Extended Data Table 1).
527 **Cryo-EM and Image processing.** MotionCor2[47] was used to perform the frame alignment
528 and dose-weighting steps and CTFFIND 4.1.5[48] provided the estimation of the Contrast
529 Transfer Function (CTF) parameters. After CTF correction and movie alignment,
530 approximately 10,000 particles were manually picked, 2D classes calculated and used as
531 references for automatic picking. All the subsequent steps of particle picking, extraction,
532 classification, and post-processing of refined models were performed in Relion 2.0.2[49].
533 For the WT-PIC dataset, 836k particles were autopicked and selected for further 2D
534 classification steps (Extended Data Fig. 2b). After three steps of particle classification, we
535 obtained a dataset containing 214k particles that was subjected to 3D classification using a 60
536 Å-filtered map of the apo Pol III as a reference (emdb id: 3179, Extended Data Fig. 2c). This
537 process provided four major classes. Class 2 (91k part.) showed density corresponding to the
538 Pol III and the downstream open DNA but lacked density corresponding to TFIIIB or the
539 upstream DNA. Class 3 (34k part.) showed clear density for all the components and Class 4
540 (59k part.) showed density of Pol III, the DNA and TFIIIB but at a lower threshold. Then,
541 particles from Class 3 and Class 4 were joined and subjected to a second round of 3D
542 classification (Extended Data Fig. 2c). This process provided a single class (Class 1, 50k
543 part.) corresponding to a Pol III / open DNA complex and another class (Class 2, 43k part.)

corresponding to a PIC (Pol III / DNA / TFIIIB at a similar density threshold). The PIC particles were refined and further subjected to polishing, producing a final map at 4.1 Å-resolution according to the gold standard FSC cut-off criterion at 0.143 (Extended Data Fig. 2c). In order to improve the quality in specific regions of the map, we performed focused 3D classifications without alignment using masks around TFIIIB/C34-tWHD and C82/34/31/stalk. This hierarchical process provided a map at 4.0 Å (OC-PIC) according to the 0.143 cut-off criterion (Extended Data Fig. 2c and Extended Data Fig. 3a). The results were validated by *ab-initio* classification using Cryosparc[50]. Regarding the Pol III/DNA complexes, we joined the particles from Class 1 and Class 2 from the first round of 3D classification and Class 1 from the second round and performed a new classification process (Extended Data Fig. 2c). A predominant class (Class 1, 101k part.) was obtained, which after data processing provided a map of Pol III/DNA at 3.7 Å (OC1-POL3) (Extended Data Fig. 3b).

The Bdp1Δ-PIC dataset was subjected to an initial 2D classification similar to that described for the WT-PIC data (Extended Data Fig. 2e). The resulting 467k particles were 3D classified, which provided 3 classes (class 1, 3 & 4) corresponding to Pol III complexes with no TFIIIB visible (Extended Data Fig. 2f). These classes (298k part.) were selected, joined and classified using a focused mask around the Pol III cleft and the downstream DNA (Extended Data Fig. 2f). This process provided 3 classes corresponding to an unbound Pol III (Classes 2/3/4, 178k part.) and a class of Pol III/DNA (Class 5, 100k part.). The latter was subjected to 3D refinement and post-processing, which gave rise to a Pol III/DNA map (OC2-POL3) at 3.4 Å according to the 0.143 cut-off criterion (Extended Data Fig. 3c). For the unbound Pol III, we joined classes 2,3 and 4 and subjected them to 3D refinement and post-processing, which gave rise to an unbound POLIII map (POL3) at 3.1 Å. Then, we performed a focused classification using a mask around the C82/C34/C31 subcomplex

(Extended Data Fig. 2f). Two major classes (Class1, 62k part. & Class4, 54k part.), representing both open and close states of the clamp helices and C82/34/31, were identified. These classes were refined and post-processed to 3.4 Å (oPOL3) and 3.3 Å-resolution (cPOL3), respectively (Extended Data Fig. 3d, e). We noticed dissociation of Pol III from TFIIIB/DNA scaffolds only in the presence of the Bdp1Δ(355-372) mutant, suggesting a lower stability of the Pol III PIC, justifying the appearance of a large fraction of unbound apo Pol III in the imaged EM samples. Indeed, in absence of Bdp1, a Pol III PIC could not be assembled and purified by size-exclusion chromatography in the conditions used in this study. A substantial fraction of Pol III PICs assembled in the presence of Bdp1Δ(355-372) was still capable of efficiently melting the DNA (OC2-POL3, Extended Data Fig. 1c). This is likely due to the enhanced ability of the Brf1-TBP fusion protein to drive low-levels of accurate transcription even in the absence of Bdp1 in contrast to isolated TBP and Brf1[20]. The OC1- and OC2-POL3 represent a state in which Pol III holds on the downstream edge of the correctly formed transcription bubble, even in absence of TFIIIB, which is disordered or dissociated after formation of the open complex.

Model Building and Refinement. In an initial step, we fitted the model of Pol III elongation form (pdb id. 5FJ8) into our OC-PIC map, which confirmed the presence of all Pol III subunits and hinted at the existence of extra components such as TFIIIB or the C34 tandem winged-helix domain. Then, we fitted the models of Pol II open complexes (pdb ids. 5IYB and 5FYW) into the map, which strongly indicated the presence of yeast Brf1/TBP complex in an equivalent position to TFIIB/TBP. Using the crystal structure of yeast TBP/Brf1-Cter (pdb id. 1NGM) and homology models of Brf1 Zinc-Ribbon domain and cyclin-folds generated with I-TASSER[51], we obtained an initial model of the yeast TFIIIB. Then, we identified the position of Bpd1 SANT domain by fitting the human Brf2/TBP/Bdp1 crystallographic model (pdb id. 5N9G) into this region. A homology model of ScBdp1 SANT

domain was also generated with I-TASSER[51]. Density corresponding to the Bdp1 linker was observed in the major groove of the DNA and it was manually extended using COOT[52]. This approach left unassigned one major density close to the Pol III protrusion and three helical densities protruding from the SANT domain. Considering previous studies[8,17], we inferred that the first density corresponded to C34 WH1 and WH2, which were absent in the previous Pol III models. We fitted the crystallographic models of the mouse homolog C39 (pdb id: 2DK5 and 2DK8) and manually mutated and fitted in COOT[52]. Next, taking into consideration secondary structure predictions and crosslink data[29,30], we manually built the helical densities as part of Bdp1 N-terminal and C-terminal regions. Finally, in accordance with previous crosslinking results[29,30], the remaining small density close to C34 WHD1 and WHD2 was attributed to the Bdp1 tether region and to the "termination/initiation loop" of C37, which was flexible in previous structures. After a preliminary position for the different components had been assigned, we manually refined the models and built small unassigned regions using COOT[52]. The nucleic acids were initially placed using human Pol II open complex (pdb: 5IYB) as a guide and then manually fitted in the density using COOT[52]. Electron density supported the building of four nucleobases of the unwound template strand and three nucleobases of the unwound non-template strand, which are directly stabilized by Pol III subunits at the upstream edge of the transcription bubble, as well as seven nucleobases of the template strand, which are loaded into the active site, and two nucleobases of the non-template strand at the downstream edge of the bubble (Fig. 1a).

In the unbound Pol III (POL3, oPOL3, cPOL3), we observed two additional globular strong density peaks, as previously reported[28]. Thanks to the higher resolution of the reconstructions presented here, we confidently interpret these peaks as clusters of metal atoms, of unknown nature and function, stabilised in large mixed hydrophobic/hydrophilic pockets of Pol III

(Extended Data Fig. 8). Because we are not sure about the nature of these metal atoms, we did not include them in the deposited models.

Refinement of the models against the maps was performed using Refmac5[53] and the PHENIX Suite[54]. Figures were prepared with Chimera UCSF[55] and Pymol (Schrödinger, LLC). Local resolution was calculated with ResMap 1.1.4[56] as implemented in Relion 2.0.2[49].

Data availability. Cryo-EM maps of the OC-PIC, OC-POL3, POL3, cPOL3 and oPOL3 have been deposited in the Electron Microscopy Data Bank with accession codes EMD-3955 (OC-PIC), EMD-3956 (OC-POL3), EMD-3959 (POL3), EMD-3957 (oPOL3) and EMD-3958 (cPOL3). The coordinates of the corresponding atomic models have been deposited in the Protein Data Bank under accession code 6EU0 (OC-PIC), 6EU1 (OC-POL3), 6EU2 (oPOL3) and 6EU3 (cPOL3).

47 Zheng, S. Q., Palovcak, E., Armache, J. P., Verba, K. A., Cheng, Y. and Agard, D. A. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331-332, (2017).

48 Rohou, A. and Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol* **192**, 216-221, (2015).

49 Kimanius, D., Forsberg, B. O., Scheres, S. H. and Lindahl, E. Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *Elife* **5**, (2016).

50 Punjani, A., Rubinstein, J. L., Fleet, D. J. and Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296, (2017).

51 Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**, 40, (2008).

52 Emsley, P., Lohkamp, B., Scott, W. G. and Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501, (2010).

53 Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. and Vagin, A. A. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355-367, (2011).

54 Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N.

W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. and Zwart, P. H. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221, (2010).

55 Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. and Ferrin, T. E. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612, (2004).

56 Kucukelbir, A., Sigworth, F. J. and Tagare, H. D. Quantifying the local resolution of cryo-EM density maps. *Nat Methods* **11**, 63-65, (2014).

EXTENDED FIGURE LEGENDS

Extended Figure 1. Cryo-EM reconstructions and model fitting. a, Cryo-EM reconstruction of the OC-PIC *(left)*. Pol III core density is coloured in transparent grey and TFIIIB, heterodimer, heterotrimer, stalk and DNA are coloured as indicated in the boxes. Atomic models are represented as ribbon. Representative electron microscopy densities of different regions show the detail of the final reconstruction, where amino-acid side-chains are discernible as well as secondary structure features. Modelled DNA nucleotides are represented as indicated in Fig. 1. **b,** as in **a,** but for the OC1-POL3 reconstruction. **c,** as in **a,** but for the OC2-POL3 reconstruction. **d,** as in **a,** but for the oPOL3 reconstruction. **e,** as in **a,** but for the cPOL3 reconstruction.

Extended Figure 2. Cryo-EM data processing. a, d Representative raw micrographs of the Bdp1 wild-type and Bdp1 (Δ355-372) datasets, respectively. **b, e,** Fifteen representative reference-free 2D class averages of the Bdp1 wild-type and Bdp1 (Δ355-372) datasets, respectively. **c,** Three-dimensional classification of the Bdp1 wild-type dataset. The particles were subjected to a hierarchical process, which encompassed several rounds of classification using global masks or masks of specific regions of the complex, as indicated. The number of particles contributing to each class is indicated. The presence of density corresponding to TFIIIB and the downstream DNA, which guided the classification process, is represented

with red or green circles, respectively. **f,** as in **c,** but for the three-dimensional classification of the Bdp1 (Δ355-372) dataset.

Extended Figure 3. Resolution of cryo-EM reconstructions. a, Fourier shell correlation plot of the OC-PIC reconstruction with the estimated resolution at the gold-standard FSC (FSC=0143), is shown in the left panels. Lateral views of the orientation distribution sphere of the particles that contributed to the OC-PIC reconstruction are shown in the middle panels. The height of the surface bars indicates the relative number of particles in a given orientation. Resolution estimation represented by a lateral view and central slice of the OC-PIC cryo-EM map is shown in the right panel. The map is coloured according to the local resolution, as indicated in the scale bar. Local resolution was calculated with ResMap 1.1.4[56] as implemented in Relion 2.0.2[49]. **b,** as in **a,** but for the OC1-POL3 reconstruction. **c,** as in **a,** but for the OC2-POL3 reconstruction. **d,** as in **a,** but for the oPOL3 reconstruction. **e,** as in **a,** but for the cPOL3 reconstruction.

Extended Figure 4. Architecture of TFIIIB and Pol III subunits involved in PIC assembly. a, Domain architecture of TFIIIB and Pol III subunits. Protein regions are depicted based on their presence (solid colour boxes) or absence (empty boxes) in the OC-PIC structure. Regions built *de-novo* (for which previous structural information was not available) are highlighted with a black line (full atomic model) or with a dashed black line (backbone model). **b,** *S. cerevisiae* Brf1 domain architecture. Ribbon models and cryo-EM maps are coloured as in **a**. Regions absent in the density are indicated as a dashed line. C- and N-termini are indicated. **c,** Bdp1 domain architecture. Ribbon models and cryo-EM maps are coloured as in **a**. Regions absent in the density are indicated as a dashed line. C- and N-termini are indicated**. d,** C37 termination/initiation loop architecture. Ribbon models and cryo-EM maps are coloured as in **a**. **e,** C34 domain architecture. Ribbon models and cryo-EM

maps are coloured as in **a**. **f,** Clamp helices architecture. Ribbon models and cryo-EM maps are coloured as in **a**. **g,** C31 stalk bridge architecture. Ribbon models and cryo-EM maps are coloured as in **a**.

Extended Figure 5. Structural conservation. a, Structure alignment of *S. cerevisiae* Brf1 and *H. sapiens* Brf2 (Pdb id. 5N9G). *S. cerevisiae* TBP (pink) and Brf1 B-core cyclin repeats (green) are represented as molecular surfaces. Brf1 "helical pin" and Brf2 "molecular pin" are shown as green and wheat cylinders, respectively. **b,** Sequence alignment of *S. cerevisiae* Brf1 "helical pin" and *H. sapiens* Brf2 "molecular pin". Residues are coloured based on their percentage identity, with dark and light blue indicating high and low sequence identity, respectively. **c,** Multiple-sequence alignment of the clamp helices of RPA190 (Pol I), RPB1 (Pol II) and RPC160 (Pol III). Residues are coloured based on their percentage identity, with dark and light blue indicating high and low sequence identity, respectively. Residues participating in the template strand pocket (Trp294, Leu298 and Tyr318) are highlighted by a red box. **d,** Multiple-sequence alignment of the bridge helix of RPA190 (Pol I), RPB1 (Pol II) and RPC160 (Pol III). Color-coding as in **c**.

Extended Figure 6. General comparison of Pol I, Pol II and Pol III PICs. a, Front view of *S. cerevisiae* Pol I PIC (pdb id. 5OA1), Pol II OC (pdb id. 5FYW) and Pol III PIC (this work). Pol III subunits are coloured as described in Fig 1. Colour scheme of Pol I and Pol II is based on architectural similarities to the Pol III system. **b**, Upstream DNA path differences. The DNA pathway in the Pol III PIC (light and dark blue) is different than in yeast (wheat, pdb id. 5FYW) or human (pink, pdb id. 5IYB) Pol II PICs, likely due to the interaction with the Bdp1 clip domain (orange). **c**, Comparison of Pol I, Pol II and Pol III protrusion tip in the PICs. Pol I (pdb id. 5OA1) and Pol III (this work) contact the promoter DNA through residues of the protrusion tip. Pol I participates in an extensive network of interactions that

involve the binding of a α-helix to the major groove of the DNA, whereas Pol III binds to the non-template strand of the DNA (light blue) through the conserved residue K409. Pol II protrusion (pdb id. 5FYW) does not participate in direct contacts with the DNA.

Extended Figure 7. Mechanism of Pol III transcription initiation. a, In unbound Pol III the stalk region and the clamp are mobile and can adopt an open or closed conformation. **b,** Upon TFIIIB recruitment, C34 WHD1 and WHD2 are positioned over the cleft through the interaction with the Bdp1 tether and the C37 termination/initiation loop. **c,** The C34 WHD2 and the Brf1 N-terminal cyclin fold promote DNA melting, which occurs through stabilisation of the template strand in a template strand pocket of the clamp helices, and the non-template strand between the C82 cleft loop and the C128 tip lobe domain. Contraction of the clamp helices induces a conformational change in C82/34/31 subcomplex. The stalk and the clamp are now locked by the C31 stalk bridge. **d,** The template strand is loaded in the active site and the transcription bubble is fully expanded, as binding of the Brf1 Zn-ribbon domain clears the DNA loading pathway. The template strand is correctly engaged in the active site cleft, in a configuration primed for elongation. The transcription bubble around the active site is very stable and might even be maintained in circumstances where the main contacts with TFIIIB are disrupted. The clamp is locked in a closed conformation preventing the re-annealing of the transcription bubble. This might be particularly important during promoter escape, since short Pol III DNA/RNA hybrids are less tightly bound compared to Pol II[28]. The stalk and clamp are locked in a closed state preventing bubble reannealing. **e,** RNA synthesis starts, the clamp helices are released and the clamp and stalk are now unlocked. The clamp remains closed during elongation but can re-open during the following steps of the transcription cycle. The rudder is repositioned and occludes access to the template strand pocket, presumably to ensure Pol III processivity.

Extended Figure 8. High density peaks in unbound Pol III cryo-EM reconstructions. The globular density peaks (blue) observed in the cryo-EM maps of the unbound Pol III reconstructions are represented at three different threshold levels. The strong features of these regions are observed even at high threshold levels, suggesting the presence of metal clusters in hydrophilic/hydrophobic pockets of Pol III. Pol III core subunits are depicted in grey and the active site magnesium ion is represented as a magenta sphere.

ACKNOWLEDGMENTS

We thank N. Cronin at the Institute of Cancer Research for help with yeast fermentation and C. Richardson for the computing infrastructure. We thank C. Plaschka (MRC-LMB, Cambridge) for advice during data processing. We thank the staff at beamline M03 of the Diamond Light Source synchrotron (UK) for help with EM data collection (EM15629, EM16599 and EM166601). The authors acknowledge the support and the use of resources of iNEXT, in particular Carsten Sachse and Wim Hagen for EM data collection at EMBL Heidelberg (PID1956 and 2180). G. A.-P. is a recipient of a Marie Sklodowska-Curie Intra-European Fellowship (EU project 655238). E. M. is supported by Cancer Research UK (CR-UK C12209/A16749). A.V. is supported by a Biotechnology and Biological Sciences Research Council (BBSRC) New-Investigator Award (BB/K014390/1), a Cancer Research UK Programme Foundation (CR-UK C47547/A21536) and a Wellcome Trust Investigator Award (200818/Z/16/Z).

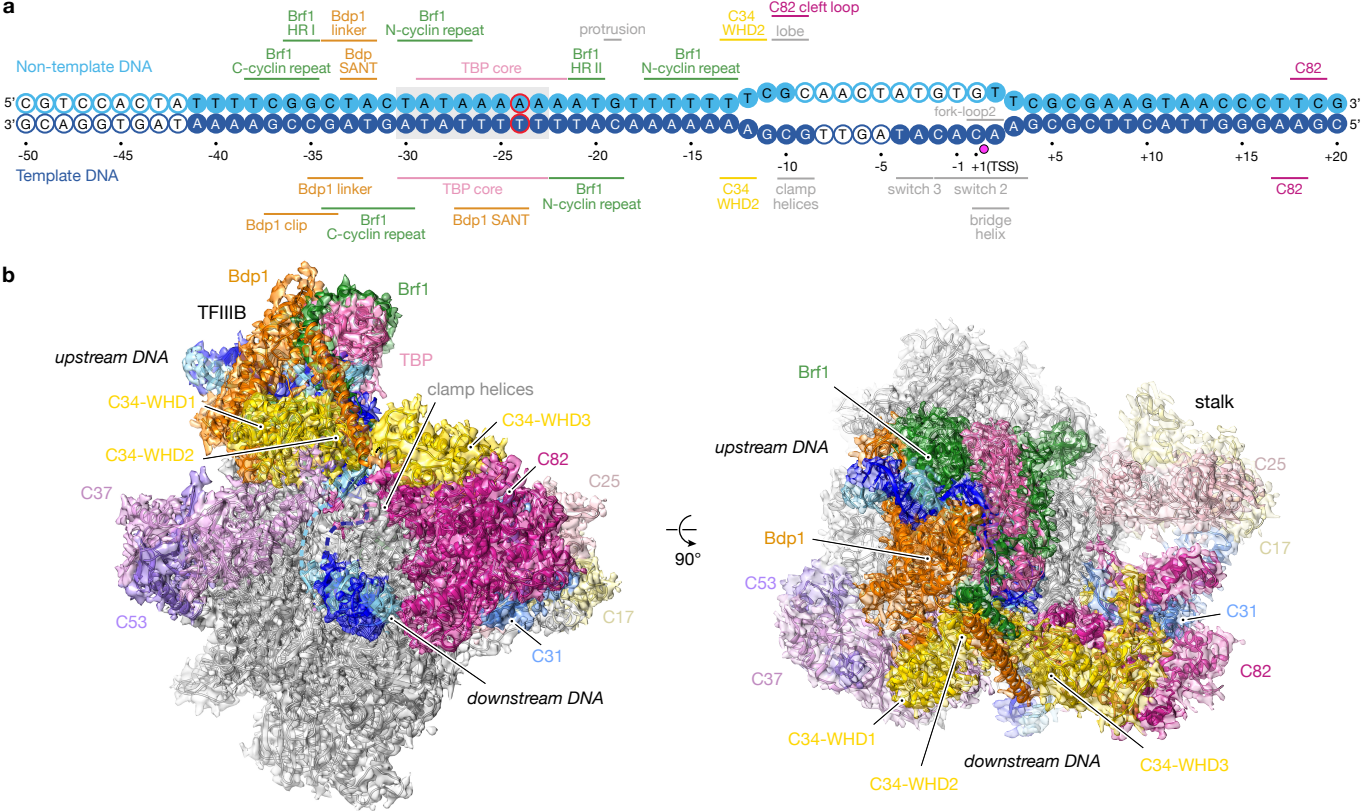
AUTHOR CONTRIBUTIONS

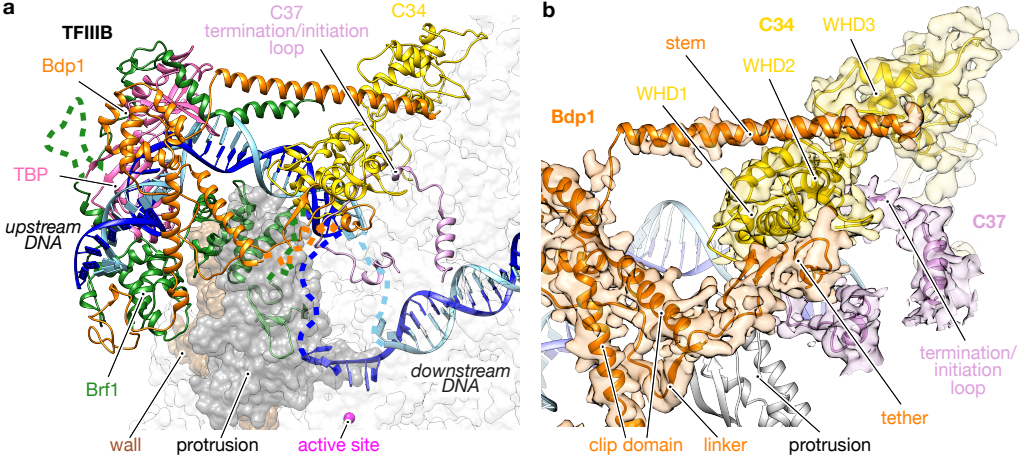
G. A-P carried out yeast fermentation, Pol III purification, Pol III-PICs reconstitution, EM specimen preparation, EM data collection and processing, model building and refinement. E. R. carried out EM data collection and processing of the Bdp1Δ(355-372) Pol III PIC and helped with EM data processing and analysis. F. B. carried out cryo-EM sample preparation,

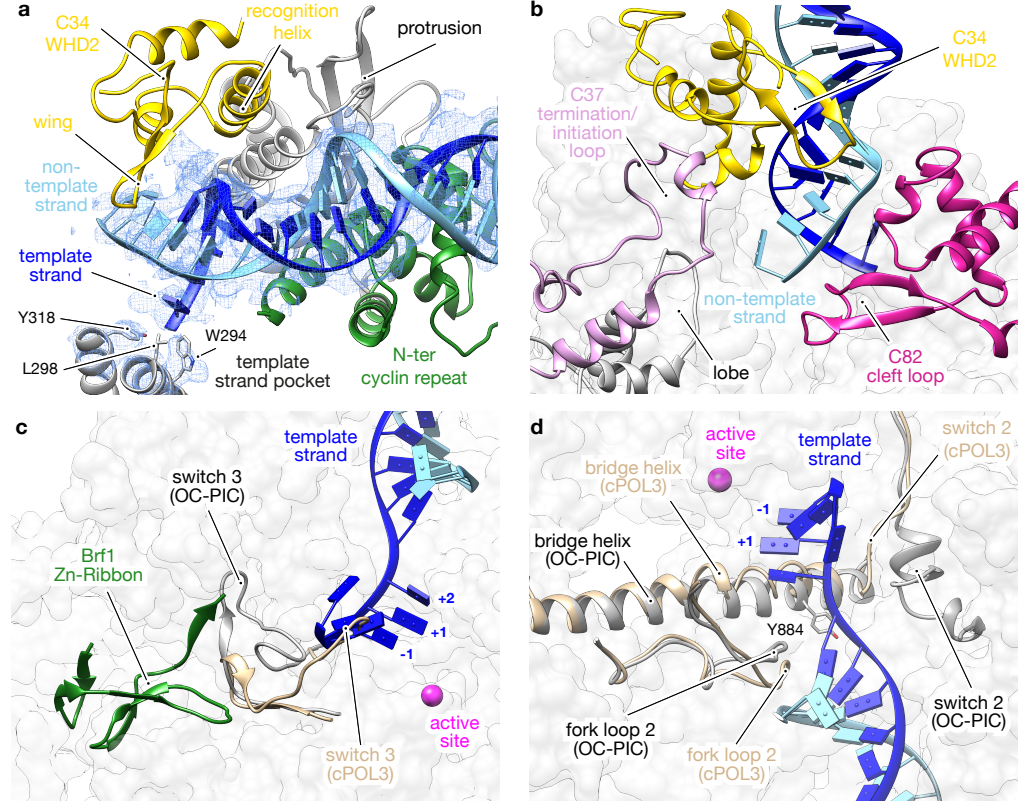
screening and sample collection. E. M. helped during initial EM characterization and data collection. A.V designed and supervised research, analyzed the structural data and prepared the manuscript with contributions from all authors.

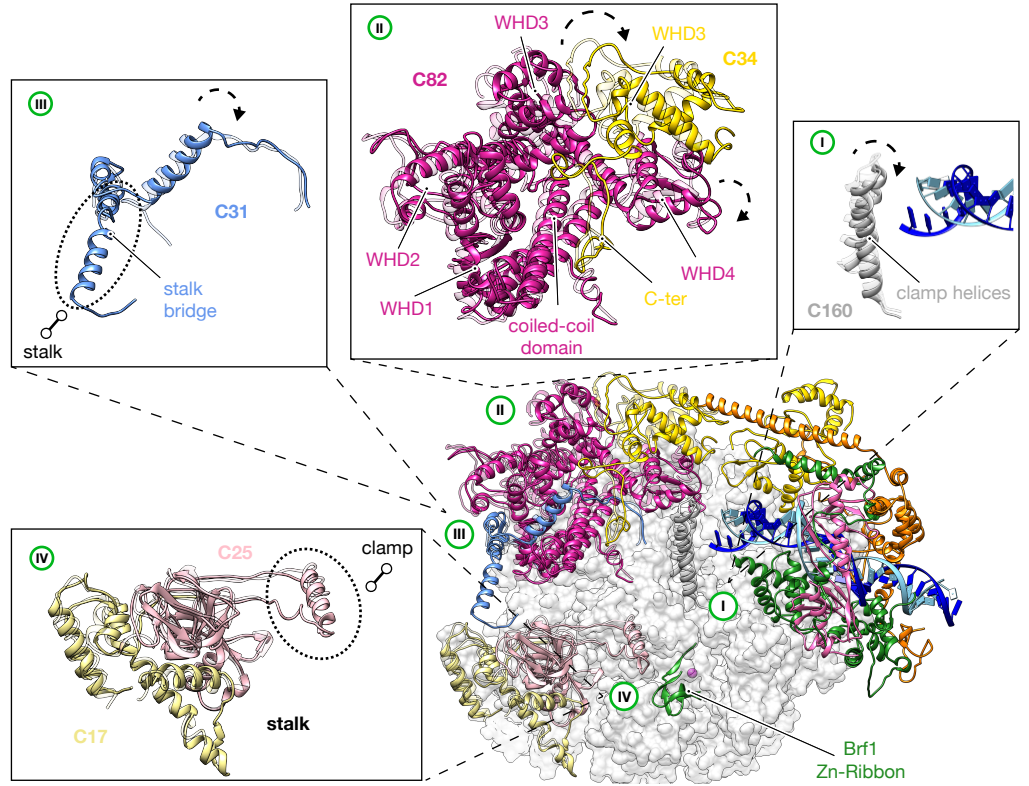
AUTHOR INFORMATION

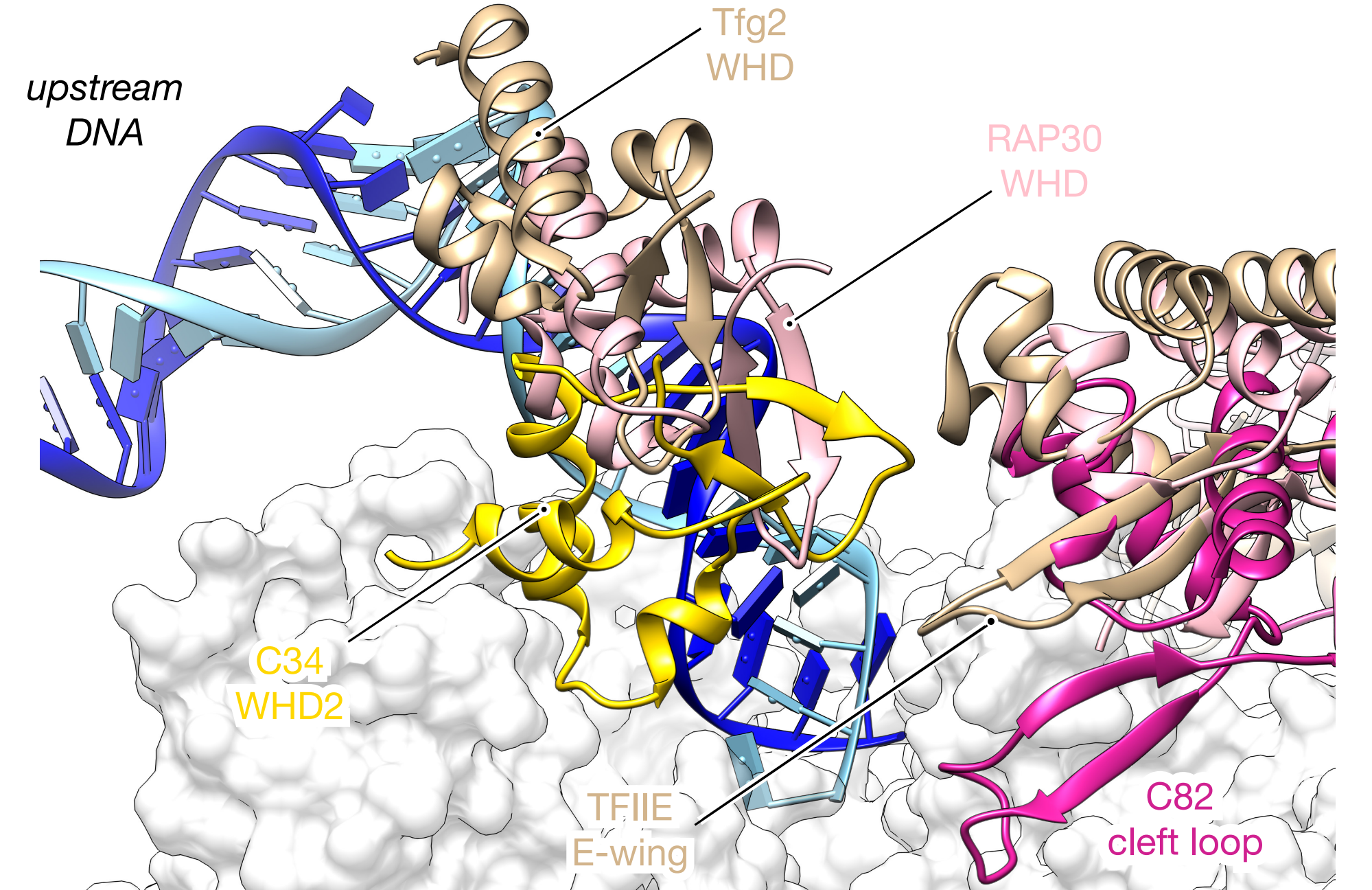
Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to A.V. (Alessandro.Vannini@icr.ac.uk).


a
Brf1 HR I
Bdp1 linker
Brf1 N-cyclin repeat
protrusion
C34 WHD2
C82 cleft loop
lobe
Brf1 C-cyclin repeat
Bdp SANT
TBP core
Brf1 HR II
Brf1 N-cyclin repeat
C82
Non-template DNA
fork-loop2
Template DNA
-50 -45 -40 -35 -30 -25 -20 -15 -10 -5 -1 +1(TSS) +5 +10 +15 +20
Bdp1 linker
TBP core
Brf1 N-cyclin repeat
C34 WHD2
clamp helices
switch 3
switch 2
C82
Bdp1 clip
Brf1 C-cyclin repeat
Bdp1 SANT
bridge helix
b
Bdp1
Brf1
TFIIIB
upstream DNA
TBP
clamp helices
C34-WHD1
C34-WHD3
C34-WHD2
C82
C25
C37
C53
C17
C31
downstream DNA
90°
Brf1
stalk
upstream DNA
C25
Bdp1
C17
C53
C31
C37
C82
C34-WHD1
downstream DNA
C34-WHD3
C34-WHD2


a
TFIIIB
C37
termination/initiation
loop
C34
Bdp1
TBP
upstream
DNA
Brf1
downstream
DNA
wall
protrusion
active site
b
C34
WHD3
stem
WHD2
WHD1
Bdp1
C37
termination/
initiation
loop
tether
clip domain
linker
protrusion


a
C34
WHD2
recognition
helix
protrusion
wing
non-
template
strand
template
strand
Y318
W294
L298
template
strand pocket
N-ter
cyclin repeat
b
C34
WHD2
C37
termination/
initiation
loop
non-template
strand
lobe
C82
cleft loop
c
template
strand
switch 3
(OC-PIC)
Brf1
Zn-Ribbon
+2
+1
-1
active
site
switch 3
(cPOL3)
d
active
site
template
strand
switch 2
(cPOL3)
bridge helix
(cPOL3)
-1
+1
bridge helix
(OC-PIC)
Y884
switch 2
(OC-PIC)
fork loop 2
(OC-PIC)
fork loop 2
(cPOL3)


II
WHD3
WHD3
C34
C82
WHD2
WHD1
WHD4
C-ter
coiled-coil
domain
III
C31
stalk
bridge
stalk
I
C160
clamp helices
II
III
I
IV
Brf1
Zn-Ribbon
IV
C25
clamp
C17
stalk


upstream
DNA
Tfg2
WHD
RAP30
WHD
C34
WHD2
TFIIE
E-wing
C82
cleft loop