





Structure of the TFIIC subcomplex τ A provides insights into RNA polymerase III pre-initiation complex formation

Matthias K. Vorländer^{1,2}, Anna Jungblut ^{1,2}, Kai Karius³, Florence Baudin ¹, Helga Grötsch¹, Jan Kosinski ^{1,3} & Christoph W. Müller ¹✉

Transcription factor (TF) IIC is a conserved eukaryotic six-subunit protein complex with dual function. It serves as a general TF for most RNA polymerase (Pol) III genes by recruiting TFIIB, but it is also involved in chromatin organization and regulation of Pol II genes through interaction with CTCF and condensin II. Here, we report the structure of the *S. cerevisiae* TFIIC subcomplex τ A, which contains the most conserved subunits of TFIIC and is responsible for recruitment of TFIIB and transcription start site (TSS) selection at Pol III genes. We show that τ A binding to its promoter is auto-inhibited by a disordered acidic tail of subunit τ 95. We further provide a negative-stain reconstruction of τ A bound to the TFIIB subunits Brf1 and TBP. This shows that a ruler element in τ A achieves positioning of TFIIB upstream of the TSS, and suggests remodeling of the complex during assembly of TFIIB by TFIIC.

¹European Molecular Biology Laboratory (EMBL), Structural and Computational Biology Unit, Meyerhofstrasse 1, 69117 Heidelberg, Germany. ²Collaboration for joint PhD degree between EMBL and Heidelberg University, Faculty of Biosciences, 69120 Heidelberg, Germany. ³Centre for Structural Systems Biology (CSSB), DESY and EMBL Hamburg, 22607 Hamburg, Germany. ✉email: christoph.mueller@embl.de

RNA polymerase III (Pol III) transcribes a number of abundant, short, folded RNAs in the eukaryotic cell, including tRNAs, 5S rRNA, spliceosomal U6 RNA, and the signal recognition particle 7SL RNA. In yeast, transcription of all Pol III genes minimally requires two conserved transcription factors (TFs) to initiate transcription *in vivo*, namely TFIIC and TFIIB. *Saccharomyces cerevisiae* TFIIC consists of six subunits with a combined molecular weight of 520 kDa, and binds to promoter elements termed A-box and B-box that are downstream of the transcription start site (TSS). The A-box element is located 12–20 nucleotides (nt) downstream of the TSS, and the B-box element is located 30–60 nt downstream of the A-box. TFIIC then recruits TFIIB and places it 25–30 nt upstream of the TSS (reviewed in ref. ¹). TFIIB consists of subunits B-related factor 1 (Brf1), TATA-box-binding protein (TBP) and B double prime 1 (Bdp1), which together encircle the upstream promoter DNA in the pre-initiation complex and allosterically activate melting of the DNA double helix^{2,3}.

The six subunits of TFIIC are organized in two subcomplexes: τ A comprises subunit τ 131, τ 95, and τ 55, and τ B comprises subunits τ 138, τ 91, and τ 60. τ A binds the A-box with low affinity and is responsible for TFIIB recruitment^{4–6}. τ B binds the B-box with high affinity, and the quality of the B-box determines promoter strength⁶. Due to the variable distances between A- and B-box elements found among tRNA genes, τ A and τ B are thought to be connected by a flexible linker^{7–9}.

In addition to its function as a Pol III TF, TFIIC also plays important roles in shaping the three-dimensional (3D) organization of the genome, especially in higher eukaryotes. Human and yeast TFIIC is recruited to *extra TFIIC* (ETC) sites that vastly outnumber Pol III genes in humans^{10–13}. ETC sites are enriched at the boundaries of topologically associating chromatin domains, at Pol II promoters, and close to the architectural protein CTCF^{11,13}. TFIIC has been demonstrated to recruit cohesin¹⁴ and condensin II (ref. ¹⁵), thereby regulating the formation of chromatin loops. Binding of TFIIC to Alu-elements in the human genome has recently been demonstrated to control gene expression through chromatin looping and direct acetylation of histone tails¹⁶.

Despite its importance as a general Pol III TF and architectural chromatin-organizing factor, the structure and molecular mechanism of holo-TFIIC in Pol III recruitment have remained elusive, presumably due to its flexible nature. However, several crystal structures of TFIIC subcomplexes and domains are available, including the N-terminal TPR array of τ 131 (ref. ⁹), a histidine phosphatase domain (HPD) of τ 55 (ref. ¹⁷), a heterodimeric portion of τ 95– τ 55 (ref. ¹⁸), and a DNA-binding domain (DBD) of τ 95 (ref. ¹⁸). The structure of a τ 60– τ 91 heterodimer is also available¹⁹, as well as a winged-helix (WH) domain of τ 138 (ref. ⁹).

Here, we have used a divide-and-conquer approach to further our structural understanding of TFIIC. We report a 3.1 Å cryo-electron microscopy (cryo-EM) structure of τ A, which constitutes the most conserved part of TFIIC. We built an atomic model of τ A that allowed us to locate mutations that affect TFIIB binding and Pol III transcription. Our structural and biochemical studies identify an element of subunit τ 95 that functions in auto-inhibiting DNA binding by τ A. Lastly, we provide a negative stain map of τ A bound to the TFIIB subunits Brf1 and TBP—a complex that mimics a “probing” intermediate where the DNA is not yet engaged. We propose a model how a ruler element in τ A positions Brf1 and TBP upstream of the TSS, where TFIIB can assemble on a suitable DNA sequence by using a proofreading mechanism that contributes to TSS selection fidelity.

Results

Cryo-EM structure determination of τ A. We prepared recombinant τ A and TFIIC by co-expression of the respective subunits

in insect cells. τ A bound TFIIB (prepared using the Brf1–TBP fusion protein²⁰ and wild-type (wt) Bdp1) during size-exclusion chromatography (SEC) and stimulated faithful *in vitro* transcription of a TFIIC-dependent promoter, albeit at much lower levels than holo-TFIIC (Supplementary Fig. 1). This resembles the *in vitro* transcription properties of tRNA genes lacking a B-box⁶. We collected cryo-EM data of a sample containing τ A, TFIIB, and a DNA fragment comprising the A-box and the TFIIB-binding promoter elements of the His_tH(GUG)E2 gene. This allowed us to determine the structure of τ A at an overall resolution of 3.1 Å (Supplementary Fig. 2). However, TFIIB and DNA had dissociated from τ A during sample preparation and are not visible in the EM density.

Structure of τ A. The cryo-EM map allowed us to build an atomic model of τ A with excellent refinement statistics (Supplementary Table 1), but also revealed density for two molecules of the detergent CHAPSO, which was added before plunge-freezing to prevent adsorption to the air–water interface²¹, bound to τ A (Supplementary Fig. 3). Our structure includes the N-terminal TPR array of τ 131 (residues 131–573, the crystal structure of this construct has been described in ref. ⁹) and a previously not described C-terminal domain (residues 612–1025), that contains a helical domain (residues 612–732) and an additional TPR array with seven repeats (residues 733–1025; Fig. 1 and Supplementary Fig. 4). The C-terminal domain packs against the N-terminal TPRs 6–10, resulting in an overall conformation of τ 131 that resembles the letter “P” (Supplementary Fig. 4). τ 131 acts as a scaffold for the assembly of τ A, as described below.

The convex surface of the N-terminal τ 131 TPR array binds the τ 55 HPD and the triple β -barrel formed by the τ 55– τ 95 dimerization domains. Despite only being present in hemiascomycetes and being absent in other eukaryotes, the τ 55 HPD is well-ordered and an integral part of the τ A structure, rather than being flexibly attached to the τ A core. We note that the active site of the τ 55 HPD is solvent accessible in the context of τ A (Fig. 1 and Supplementary Fig. 4).

τ 95 is woven through the τ A structure in an intricate manner. Its N-terminus originates between the τ 55 HPD domain and the τ 55– τ 95 β -barrel. The β -barrel is followed by 75 amino acids (residues 161–236) that fold inside the superhelical groove of the C-terminal TPR of τ 131 into a disc-like domain, containing two α -helices (Supplementary Fig. 5). This region connects the τ 95 dimerization domain with the DBD. The τ 95 DBD (residues 263–509) is positioned between the τ 131 “ring” domain (residues 390–430, ref. ⁹) and the C-terminal TPR array of τ 131. Interestingly, a C-terminal *S. cerevisiae*-specific portion of τ 95, which we refer to as the acidic plug (residues 566–592) is bound to the predicted DNA-binding interface¹⁸ of the τ 95 DBD (Fig. 2a and Supplementary Fig. 4).

The C-terminus of τ 95 auto-inhibits DNA binding of τ A. The acidic plug contains a helix that is rich in negatively charged residues and is embedded in the positively charged DBD. Because the DBD of the *Schizosaccharomyces pombe* τ 95 homolog Sfc1 was shown to be auto-inhibited in DNA binding by a C-terminal acidic portion¹⁸, we wondered if the *S. cerevisiae* C-terminal region, which comprises the acidic plug followed by an acidic disordered “tail”, has a similar auto-inhibitory function in *S. cerevisiae* τ A.

To test if the acidic plug or the acidic tail inhibit DNA binding, we prepared two τ A variants with deletions in the C-terminus of τ 95 (Fig. 2b). The first mutant, τ 95 Δ^{plug} , lacks the acidic plug and the disordered acidic tail (stop codon introduced after residue 521). The second mutant, τ 95 Δ^{tail} , lacks only the

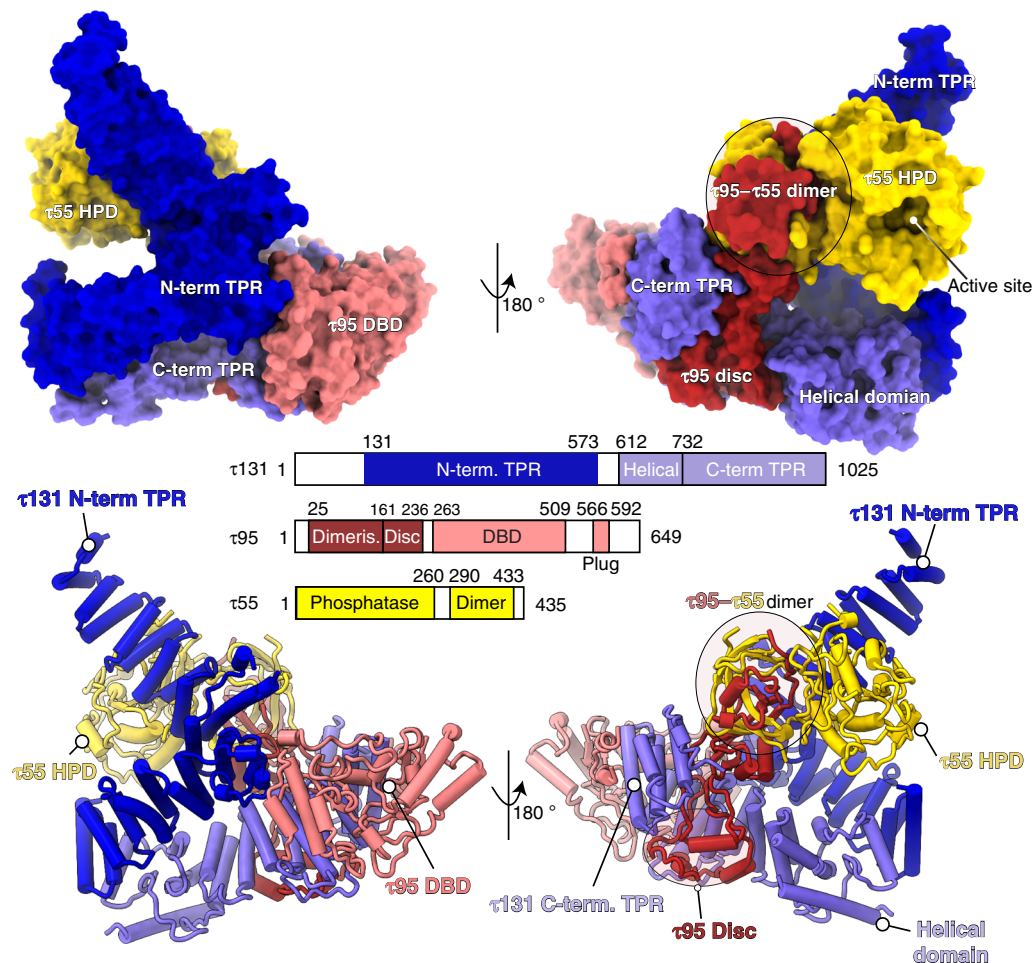


Fig. 1 Structure of τ A shown in two views, colored according to the bar diagram below. The top panel shows τ A as a surface rendering, and the bottom panel as a ribbon rendering with α -helices shown as cylinders.

disordered acidic tail (stop codon introduced after residue 592). Both variants bound A-box DNA in an electrophoretic mobility shift assay (EMSA) much stronger than full-length τ A, confirming the auto-inhibition of DNA binding in the context of *S. cerevisiae* τ A (Fig. 2c, left gel). However, the $\tau 95\Delta^{\text{tail}}$ mutant bound A-box DNA stronger than the $\tau 95\Delta^{\text{plug}}$ mutant. We obtained similar results using a filter-binding assay, but attempts to determine the affinity more precisely were hampered by the fact that we could not obtain τ A at sufficient concentrations to achieve complete saturation of binding.

The data are consistent with a model where the acidic plug is not displaced from the DBD during DNA binding, but the acidic tail transiently associates with the positively charged DBD, thereby competing with DNA and reducing the affinity of τ A to DNA. The relative reduction in DNA affinity of $\tau 95\Delta^{\text{plug}}$ compared to the $\tau 95\Delta^{\text{tail}}$ points to an architectural role of the plug in stabilizing τ A rather than a role in DNA binding. Note that wt τ A just begins to shift the probe at a protein concentration of $10\ \mu\text{M}$ and a DNA concentration of $1\ \mu\text{M}$, suggesting that the affinity to the tested A-box sequence falls roughly in the micromolar range.

We next introduced the $\tau 95\Delta^{\text{tail}}$ mutation in recombinant holo-TFIIC and tested DNA binding using EMSAs. Deletion of the acidic tail has only a mild effect on DNA binding by holo-TFIIC under our experimental conditions (Fig. 2c). This indicates that in holo-TFIIC, the effect of the tail mutation is buffered by the high-affinity interaction of τ B with the B-box⁶,

which presumably dominates initial formation of TFIIC–DNA complexes and thus leads to a high local DNA concentration, which facilitates engagement of the A-box.

The Δ^{tail} constructs were also tested in *in vitro* transcription assays. Consistent with the results from our EMSA assays, deletion of the tail has only a minor stimulatory effect on the transcriptional activity of holo-TFIIC in our experiments. Compared to wt τ A, $\tau\Delta^{\text{tail}}$ stimulated slightly higher transcription levels at increasing protein concentrations, but still has poor activity compared to holo-TFIIC (Fig. 2d). This indicates that τ B is required for full TFIIC function, potentially due to the contribution of the τ B subunit $\tau 60$ to TFIIB recruitment¹⁹.

We also introduced the $\tau 95\Delta^{\text{tail}}$ and $\tau 95\Delta^{\text{plug}}$ mutations in the endogenous TFIIC locus in yeast to test their functional importance *in vivo*. We observe a growth defect for both mutations at elevated temperature on rich media ($37\ ^\circ\text{C}$) and at optimal temperature on minimal media ($30\ ^\circ\text{C}$; Fig. 2e). Therefore, the C-terminus of $\tau 95$ is functionally important *in vivo*. We speculate that the acidic tail might increase the specificity of the A-box interaction by outcompeting suboptimal DNA sequences, whereas the acidic plug appears to stabilize τ A, consistent with its position at the interface of $\tau 95$ and $\tau 131$. Given that under exponential growth conditions, all tRNA genes are occupied by Pol III in yeast¹⁰, perturbations that affect TFIIC recruitment could easily lead to a reduction in the cellular tRNA pool, explaining the observed reduced growth rates of our $\tau 95$ mutant strains. This might not be captured in our *in vitro* transcription

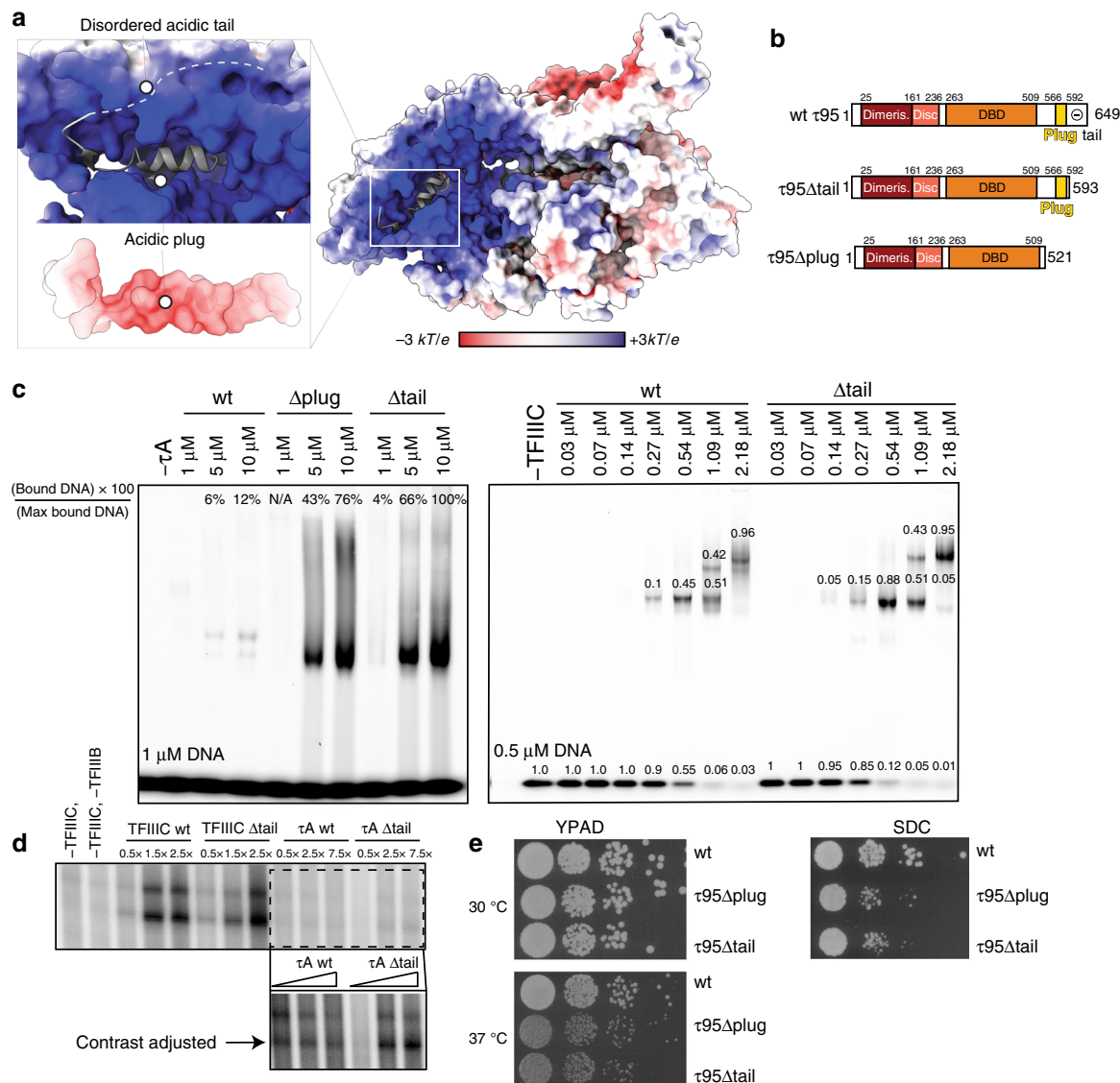


Fig. 2 The acidic C-terminal region of $\tau 95$ auto-inhibits DNA binding of τA . **a** Electrostatic potential mapped onto the surface of τA . The acidic plug is negatively charged and embedded in the positively charged DBD, followed by a disordered acidic tail (indicated by dashes). **b** Domain representation of $\tau 95$, $\tau 95\Delta tail$, and $\tau 95\Delta plug$. **c** EMSAs comparing the effect of deletions in the $\tau 95$ -C-terminus in the context of τA (left) and TFIIC (right). Left gel: labels (in %) at the top of lanes show the relative intensity of the bound DNA in that lane, normalized to the most intense band of bound DNA in the gel. Right gel: numbers above bands show the amount of DNA present in the band relative to the total DNA in the lane. **d** In vitro transcription assays comparing the effect of the $\Delta tail$ deletion on transcriptional activity of τA and TFIIC. The boxed region in the upper panel was cropped and the contrast was adjusted to better visualize differences between τA and $\tau A\Delta tail$. **e** Yeast viability assays of $\tau 95\Delta plug$ and $\tau 95\Delta tail$ strains compared to $\tau 95$ wt. Source data are provided as a Source data file.

assays due to the use of short DNA oligos as templates, and therefore lack of suboptimal, competing sequences.

$\tau 95$ cannot bind DNA in a canonical way. The $\tau 95$ DBD consists of a WH domain and a WH-interacting domain, as described for the *S. pombe* homolog Sfc1 (ref. 18). We searched the Protein Data Bank (PDB) for the closest structural matches of the $\tau 95$ dimerization domain and the $\tau 95$ DBD using the DALI server²², and found the Pol II TF TFIIF β ($\tau 95$ DBD to the TFIIF β WH domain with a Z-score of 5.4 and an r.m.s.d. of 1.8 Å, the $\tau 95$ dimerization domain to the TFIIF β dimerization domain with a Z-score of 3.1 and an r.m.s.d. of 4.4 Å). This mirrors the similarity of *S. pombe* Sfc1 and TFIIF β reported previously¹⁸. To test if τA might adopt a similar position in the

Pol III-PIC as TFIIF does in the Pol II-PIC, we superimposed τA onto the Pol II-PIC (PDB 5oqj²³), using only the $\tau 95$ DBD and the TFIIF β WH domain for alignment. While this illustrates the close structural similarity between the WH domains (Fig. 3a), it is clear that the $\tau 95$ DBD cannot bind to DNA as the TFIIF-WH domain does in the context of our τA structure. Both, the $\tau 95$ acidic plug and the $\tau 131$ C-terminal TPR array, severely clash with DNA in the complex modeled through superposition with TFIIF (Fig. 3b). TFIIF-like binding would thus require the $\tau 95$ DBD to dissociate from these elements, which together bury a large area (2265 and 2840 Å², respectively, calculated with COCOMAPS²⁴), arguing against this scenario. The superposition of the $\tau 95$ WH domain with a representative set of 56 structures of WH domains in complex

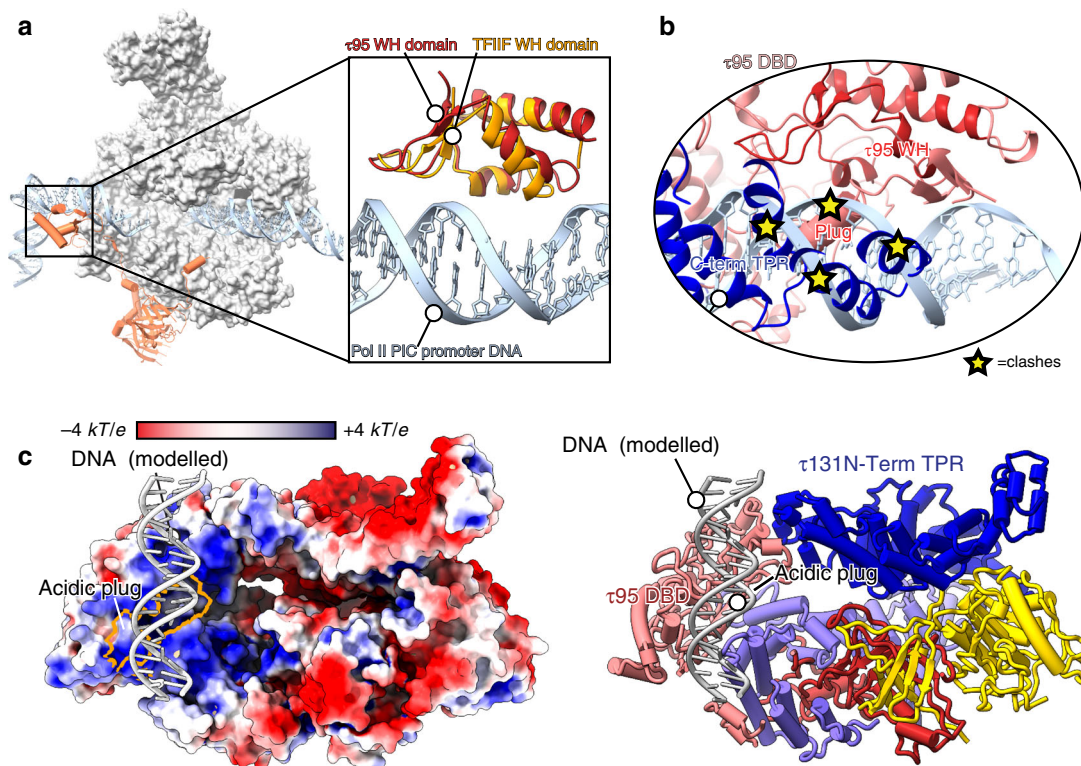


Fig. 3 The $\tau 95$ WH cannot bind promoter DNA in the same way as the TFIIF-WH domain. **a** Structure of the Pol II-PIC (PDB 5oqj), with TFIIF β shown in orange and $\tau 95$ shown in red. For clarity, other transcription factors of the Pol II-PIC have been omitted. The boxed region shows a superposition of the $\tau 95$ WH and the TFIIF-WH domains (PDB 5oqj). **b** In the context of τA , the $\tau 95$ WH domain cannot bind DNA like the TFIIF-WH domain due to steric clashes with the acidic plug and the $\tau 131$ C-term TPR. **c** Model of possible DNA binding to τA , with τA shown as a surface colored by electrostatic potential (negative (red) to positive (blue)). The acidic plug (outlined in orange) does not interfere with the modeled DNA (left). A ribbon model is shown on the right with the same color code as in Fig. 1.

with DNA (retrieved from the CATH database²⁵ and curated from structures of Pol I, Pol II, and Pol III in complex with their TFs) also did not allow identifying a binding mode that could accommodate an unperturbed double-stranded DNA without steric clashes. Therefore, τA likely employs a binding mode that is different to the frequently observed WH domain–DNA interaction, where the recognition helix protrudes into the major groove. For example, it is conceivable that τA binds its promoter element in a shallow positively charged groove instead. To illustrate this, we placed B-DNA over the most positively charged surface of τA (Fig. 3c), but due to the possibility of conformational changes in τA during DNA binding and the lack of structural data this remains speculative.

Recruitment of TFIIB by τA . The interaction between TFIIC and TFIIB has been extensively studied, leading to a sequential assembly model for TFIIB (reviewed in refs. 1,26). Assembly is initiated by recruitment of Brf1 to DNA-bound TFIIC through the N-terminal TPR array of $\tau 131$. Next, Brf1 recruits TBP, and, finally, Bdp1 enters the complex, driven by binding sites in the $\tau 131$ N-terminal TPR array and in Brf1 and TBP. Bdp1 renders the TFIIB–DNA complex extremely stable and resistant to high salt²⁷. The recruitment process involves a series of conformational changes^{28–31}, and the initial step of the reaction, binding of Brf1 to the N-terminus of $\tau 131$, has been studied in detail. In particular, each of the two arms of the $\tau 131$ N-terminal TPR array is capable of binding Brf1³⁰. In addition, gain-of-function mutations that increase Pol III transcription, loss-of-function mutations and mutations that rescue the loss-of-function mutations have been described in $\tau 131$ (refs. 29,32,33).

In order to understand how TSS selection is achieved in the Pol III system, structural information about the complex of TFIIB bound to τA /TFIIC is necessary. However, as mentioned earlier, TFIIB and DNA had dissociated from τA in our cryo-EM preparation. Interestingly, our cryo-EM map reveals a molecule of the detergent CHAPSO bound to the right arm of $\tau 131$, which is also predicted to bind TFIIB (Supplementary Fig. 3), and it is therefore possible that addition of CHAPSO competed with TFIIB for binding to $\tau 131$ and thereby contributed to dissociation of TFIIB from τA . However, sample prepared without addition of CHAPSO aggregated strongly on EM grids and was unsuitable for data collection. Attempts to stabilize the sample through chemical crosslinking resulted in disruption of the complex, presumably because the crosslinker modified lysine residues in the DNA-binding interface. We thus prepared a complex of τA and the Brf1–TBP fusion protein and crosslinked the sample using the GRAFIX method³⁴. Brf1–TBP bound τA with apparent 1:1 stoichiometry as assessed by co-elution on glycerol gradients in the absence of crosslinker (Supplementary Fig. 5). While attempts to obtain a high-resolution cryo-EM structure of the sample were unsuccessful, we obtained a negative stain reconstruction with a resolution of ~ 30 Å (Supplementary Fig. 6a). Note that our negative stain reconstruction is free of model bias, as we used the ab initio algorithm in CryoSPARC to obtain a 3D reference for particle alignment. We identified the most likely fit of the τA structure in this map, which led to the highest fitting scores with four different EM fitting metrics, and assigned the handedness of the map (see “Methods” section for details). Unfortunately, due to the low resolution, we could not unambiguously assign the orientation of Brf1 and TBP, although their overall location relative to τA can be extracted.

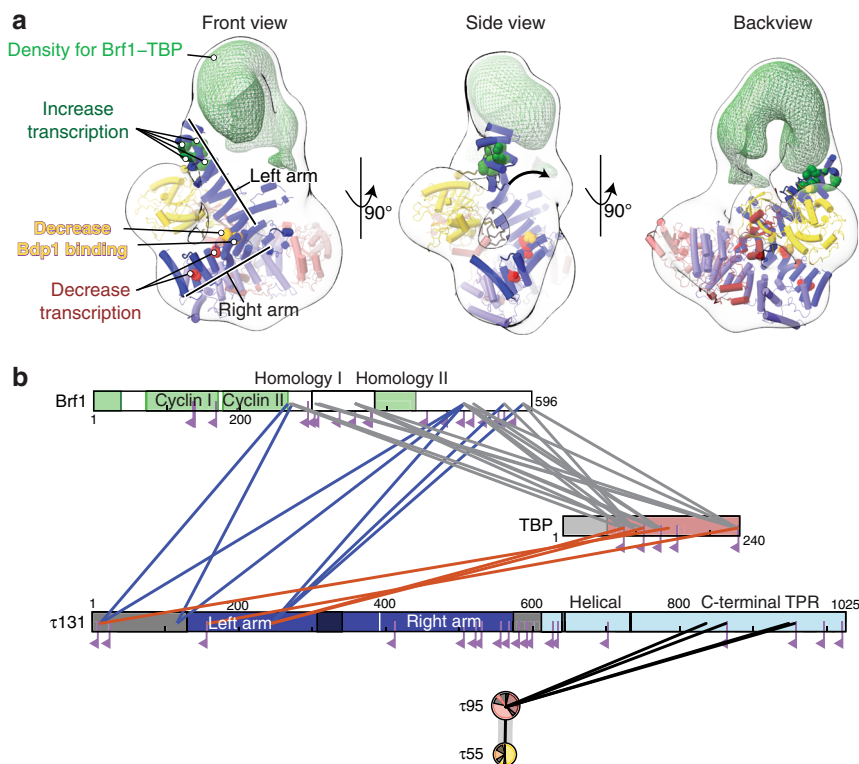


Fig. 4 Interactions between τ A and Brf1-TBP. **a** Negative stain reconstruction of τ A bound to Brf1-TBP. τ A was fitted as a rigid body into the density. τ A subunits are colored as in Fig. 1. Mutations in τ 131 that affect Pol III transcription are shown as spheres and colored green (increased transcription), yellow (decrease Bdp1 binding) or red (decrease transcription). The arrow in the middle panel (“side view”) indicates the movement of the τ 131 TPR array that is required to fit into the density. Additional density for Brf1-TBP is shown as a green mesh. **b** Schematic view of high-confidence inter-subunits crosslinks of the τ A-Brf1-TBP sample. Crosslinks between lysine residues with a score of >40 are displayed as solid lines. Mono-links are shown as flags. Domains of known structure in Brf1 are coloured green, parts of unknown structure/disorder are shown in white. τ 131 domains are coloured as in Fig. 1. τ 55 and τ 95 domain diagrams are collapsed into circles.

Relative to the fit of τ A, density for Brf1-TBP is located above the τ 131 N-terminal TPR array, in agreement with previous reports that the N-terminal TPR is the main interaction hub of τ 131 and Brf1 (refs. 26,29–31,35,36). We observe that the τ 131 N-terminal TPR array appears to be shifted toward this density (Fig. 4a, middle panel and Supplementary Movie 1). To assess if the observed shift is an artifact of negative staining, we also determined a negative stain structure of τ A in absence of Brf1-TBP. The negative stain envelope of the “ τ A-only” map fits better to our cryo-EM structure (Supplementary Fig. 5e), supporting the idea that the N-terminal TPR is shifted in the Brf1-TBP-bound state relative to the apo state. Furthermore, this portion of τ 131 also has the highest B-factors and lowest local resolution in the apo structure (Supplementary Fig. 4), and mutations that increase Pol III transcription map to this region^{9,29} (Fig. 4a). Presumably, these mutations shift the equilibrium between the conformations observed in the apo state and the Brf1-TBP-bound state toward the latter. Interestingly, the positive effect on Pol III transcription of one of these mutations (H190Y located in TPR2) has previously been interpreted as functioning by relieving an auto-inhibited state of τ 131 (ref. 31); this auto-inhibited state could correspond to the τ A conformation observed in the apo structure.

In contrast to τ A, fitting of Brf1-TBP into the negative stain density led to several alternative fits with similar scores, not allowing us to identify a unique unambiguous fit. Nevertheless, visual assessment of a putative, high-scoring fit indicates that the additional density on top of τ A can encompass the entire

Brf1-TBP complex (Supplementary Fig. 6c). Crosslinking-mass spectrometry (XL-MS) restraints generated from the same sample (Fig. 4b) also did not allow us to unambiguously determine the orientation of Brf1-TBP in the negative stain map, because of the low number of detected crosslinks between structured Brf1-TBP and τ A residues. Several factors might contribute to the difficulties in obtaining unambiguous fits. First, the available structures might change conformation upon complex formation, as seen for the τ 131 N-terminal TPR. As only structures of Brf1-TBP in DNA-bound form are available, the relative orientation of the domains might be different in absence of DNA. Second, the complex might be flexible to a degree (discussed below), resulting in averaging of different conformations in the map. Third, the negative stain map appears to enclose a larger volume than expected when considering the portions of τ A and Brf1-TBP of known structure. This could be due to ordering of flexible parts in the complex ($\sim 20\%$ of τ A and Brf1-TBP residues are not resolved in the available structures), which cannot be modeled based on low-resolution EM data.

Nonetheless, the overall location of the Brf1-TBP density relative to τ A is informative even without explicitly fitting the orientation of Brf1-TBP. Density for Brf1-TBP is located above the N-terminal TPR array of τ 131 and contacts mostly the first two TPRs (Fig. 4). Additional density extends toward the τ 95 DBD. However, no density is present close to the “right arm” of the τ 131 N-terminal TPR array, although this part (TPRs 6–10) has been shown to constitute a binding site for Brf1 with higher affinity than TPRs 1–5 (ref. 31). The implications of this finding are discussed below.

Discussion

Here, we report structural models for *S. cerevisiae* τ A alone, and for τ A bound to the TFIIB subunits Brf1 and TBP in absence of DNA, which provide insights into yeast-specific and general features of τ A. Both yeast and human TFIIC have enzymatic activity, although with clear differences: the purified *S. cerevisiae* τ 55 HPD domain has phosphatase activity and has been shown to be able to dephosphorylate peptides derived from the Pol II C-terminal domain that are phosphorylated on tyrosine 1 *in vitro*¹⁷, while several human TFIIC subunits possess histone acetyl transferase activity^{37,38}. However, because the τ 55– τ 95 heterodimer exists independently of TFIIC in yeast³⁹, it was unclear if TFIIC harbors enzymatic activity or not. Our structure shows that the τ 55 active site is accessible in τ A, arguing that TFIIC can directly modify substrates and the activity might be coupled to its function as a TF.

Our biochemical studies also confirm the existence of a disordered, auto-inhibitory C-terminal tail of subunit τ 95, which contributes to the low affinity of τ A to DNA. We have previously shown that a negatively charged tail at the C-terminus of τ 95 is conserved from yeast to humans¹⁸, and our *in vivo* experiments show deletion of this tail confers a temperature-sensitive growth defect in *S. cerevisiae*.

TFIIC serves as an assembly factor for TFIIB and our XL-MS data obtained from a sample containing τ A, and the TFIIB subunits Brf1 and TBP recapitulate published interactions between τ 131 and Brf1. We observe several crosslinks from the disordered N-terminus of τ 131 to Brf1 and to TBP, in agreement with the N-terminus being necessary for high-affinity interaction between τ 131 and Brf1 (refs. 9,30,40). We also observe crosslinks between τ 131 and the C-terminal part of Brf1, in agreement with the finding that a peptide corresponding to the Brf1 homology domain II inhibits assembly of TFIIC–TFIIB complex³⁵. However, at the chosen threshold (see “Methods” section), no crosslinks are found between the Brf1 cyclin domains or the homology domain II and τ 131, perhaps because lysine residues are buried in the interface and not accessible or not within crosslinking distance to lysine residues in τ 131. Unfortunately, this, together with the low resolution of the negative stain EM map, precludes us from modeling the exact position of these domains relative to τ A.

Nonetheless, the overall position of TFIIB relative to τ A, and the elements of τ A that are contacted by TFIIB in our structure can be extracted and are informative with regard to TFIIC’s primary function in Pol III transcription, which is to assemble TFIIB. While TFIIB is bound preferentially ~30 bp upstream of the TSS, previous work^{41,42} has demonstrated that TFIIB placement is codirected by TFIIC and by direct interaction of the TBP subunit with the upstream DNA sequence. TBP can select a suitable AT-rich sequence within a 20 bp (or 70 Å) window⁴², indicating some flexibility in the TFIIC–Brf1–TBP complex. We refer to the assembly in which Brf1 and TBP are bound to TFIIC, but not yet to the upstream sequence, as the “probing state”. It has been suggested that τ 131 is responsible for the flexibility in the probing state⁴².

Our negative stain reconstruction of τ A–Brf1–TBP lacks DNA, and therefore resembles the probing state, and thus an early assembly intermediate. This snapshot explains some of the biochemical observations concerning TFIIC and assembly of TFIIB made over the past decades; however, our interpretations remain to be validated by high-resolution structures.

First, we observe that the N-terminal TPR array of τ 131 might be shifted in the Brf1-bound state relative to the apo state (Fig. 4a). This movement might explain how mutations in the N-terminal TPR array can increase transcription by shifting the equilibrium toward the conformation observed in the Brf1-bound state, thereby overcoming a rate-limiting step in the assembly of TFIIB.

Second, we note that the negative stain density encloses a larger volume than expected when we only consider the regions of τ A and Brf1–TBP that are ordered in available structures. It is hence possible that the flexible N-terminus of τ 131, which is required for high-affinity binding to Brf1 (refs. 9,30,40), becomes structured in the complex with Brf1–TBP. In line with this, it has been previously demonstrated through circular dichroism spectroscopy that τ 131(residues 1–580), which encompasses the disordered N-terminus and the N-terminal TPR array, can form additional α -helices under certain conditions, and that binding to Brf1 to this construct also involves formation of new α -helices³⁰.

We would like to propose a model how additional binding sites in the τ 131 TPR array, which are not contacted by TFIIB in the probing state captured here, are used during later stages of TFIIB assembly (Fig. 5). To our surprise, in our negative stain structure, Brf1–TBP is only bound to the left arm (TPR 1–5) of the τ 131 TPR (Figs. 4 and 5a), but far away from the right arm (TPRs 6–10) of τ 131, although the right arm constitutes a second binding site for Brf1 that has higher affinity than the left arm³¹, and mutations that decrease Pol III transcription and Bdp1 binding also cluster to the right arm^{9,36,43} (Fig. 4a). We thus believe that binding of DNA to Brf1–TBP and addition of Bdp1 to the complex triggers a large conformational change (in agreement with evidence from DNA footprinting data^{44,45}) that brings the Brf1–cyclins and Bdp1 close to the right arm of τ 131 (Fig. 5b). This might be initiated by initial bending of the DNA by TBP and subsequently stabilized through addition of Bdp1 (Fig. 5c and d). This two-step mechanism of TFIIB assembly would open an opportunity to probing for an AT-rich upstream DNA element and allow for a proofreading mechanism, in which the lifetime of the initial TBP–DNA complex helps to select the correct sequence around which TFIIB assembles. Consistent with this model, recent single-molecule experiments showed that the lifetime of the bent state of human TBP–DNA complexes depends on the quality of the TATA box, with suboptimal sequences having a shorter lifetime⁴⁶. Hence, only a suitable AT-rich sequence could remain in the bent state for a sufficient amount of time to allow binding of Bdp1, whereas suboptimal sequences would revert to the unbent/disengaged state quickly. This mechanism could help to ensure the correct placement of TFIIB and faithful TSS selection.

Our structures also hint at the mechanism by which TFIIB is placed at a relatively constant position upstream of the A-box: the DBD auf τ 95 and the Brf1– τ 131 TPR array are located at opposite ends of τ A. Therefore, τ A might serve as a molecular ruler that positions TFIIB at a fixed distance upstream of the A-box, where it can scan for a suitable “bendable” region. While structurally unrelated, this resembles the Pol II factor TFIID, which also binds to promoter elements within the transcribed region and assembles TBP upstream of the TSS.

Lastly, we would like to discuss the role of TFIIC in transcription initiation. During the course of this project, substantial effort was invested in reconstituting a stable complex between TFIIC, TFIIB, and Pol III on different tRNA promoter sequences, yet without success. While the similar molecular weights of TFIIC and Pol III make it difficult to interpret co-elution of TFIIC and Pol III on size-exclusion columns or glycerol gradients, using recombinant τ A we can clearly show that τ A elutes separately from the TFIIB–Pol III–his promoter complex when purified over a glycerol gradient or size-exclusion column (Supplementary Fig. 7), despite τ A binding TFIIB when Pol III was omitted (Supplementary Fig. 1). This suggests that τ A, but also TFIIC, act as assembly factors for TFIIB, but are not bona fide components of the pre-initiation complex. Several findings are in line with this: when assembly of the Pol III–PIC on its promoter was monitored using chemical and enzymatic footprinting experiments, A-box protection

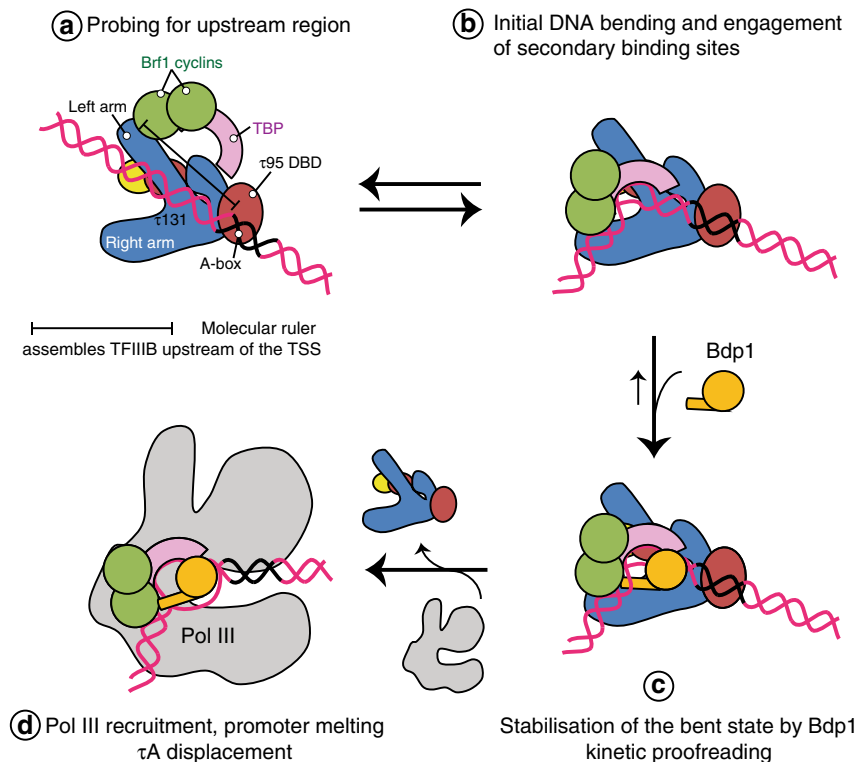


Fig. 5 Model of TFIIB assembly by τ A/TFIIIC. **a** Initially, Brf1 and TBP are recruited to the N-terminal TPR array of τ 131. The distance between the τ 95 DBD and the τ 131 TPR array serves as a molecular ruler that places TFIIB within a certain distance upstream of the TSS. **b** TBP binds and bends the upstream DNA sequence. The lifetime of this complex depends on the upstream DNA sequence engaged. **c** Bdp1 then enters the complex and stabilizes the bent state. **d** Recruitment of Pol III and promoter opening displaces τ A, freeing the transcribed region.

was lost when Pol III was added⁴⁷, indicating that Pol III dissociates τ A from its promoter element, although TFIIC might still be tethered to the gene through binding of τ B to the B-box. This agrees with previous in vitro transcription studies. Removal of TFIIC did not affect transcription in vitro, once TFIIB was assembled⁴⁸. Chromatin immunoprecipitation studies have established that Pol III occupancy and TFIIC occupancy are anticorrelated, with TFIIC occupancy peaking under repressive conditions, and Pol III occupancy peaking under conditions that favor high transcription rates^{11,12,49}. We speculate that displacement of τ A/TFIIIC from its promoter concomitant with Pol III recruitment is a mechanism to ensure efficient promoter escape of Pol III. Interestingly, TFIID has recently also been proposed to be displaced from its promoter during transcription initiation⁵⁰, showing further parallels between the two factors.

Methods

τ A expression and purification. Insect cell codon-optimized genes for τ A subunits were cloned into the pAceBac vector under a p10 promoter, and a TEV-cleavable his-tag was added to the N-terminus of τ 95, using overlap extension PCR cloning. Expression cassettes were combined into the pBig1a plasmid using the BigBac⁵¹ assembly method. Baculovirus was generated following standard procedures. Expression constructs for τ 95 Δ C mutants were generated via QuickChange mutagenesis PCRs.

τ A and TFIIC mutants were expressed in Hi5 insect cells. Cells were diluted 24 h before transfection at 5×10^5 cells/mL. Infection was performed at a cell density of 1 million cells/mL and protein was expressed for 72 h at 27 °C. Cells were harvested at $800 \times g$, resuspended in PBS buffer and pelleted again. Cells were lysed in a $3 \times$ volume of lysis buffer (500 mM NaCl, 10% glycerol, 50 mM Tris-HCl pH 7.5, 4 mM β -mercaptoethanol (bME), 25 mM imidazole, 2 mM $MgCl_2$, 1:500 DNase, 5 μ L benzonase, and SigmaFast EDTA-free inhibitor tablets) by sonication. The lysate was cleared by ultracentrifugation for 1 h at $30,000 \times g$ and filtered through 1.2 μ m glassfiber filters, loaded on a pre-equilibrated 5 mL GE Healthcare Nickel NTA FF column. The column was washed with 75 mL of lysis buffer and eluted with His-B buffer (100 mM NaCl, 50 mM Tris-HCl pH 7.5, 4 mM bME, and 100 mM imidazole). Peak fractions were pooled, diluted to a conductivity of

~ 20 mS and loaded onto a 5 mL Heparin FF column pre-equilibrated in HepA buffer (100 mM NaCl, 50 mM Tris-HCl pH 7.5, 5% glycerol, and 5 mM DDT). Protein elution was performed with a linear gradient to 100% HepB buffer (1 M NaCl, 50 mM Tris-HCl pH 7.5, 5% glycerol, and 5 mM DDT). Peak fractions were diluted to 100 mM NaCl and loaded onto a MonoQ 5/50 GL column and eluted with a linear gradient from 0 to 100% HepB over 100 mL. This separated τ A from excess τ 95- τ 55 dimer. The sample was concentrated and applied to a Superose 6 INCREASE 10/300 gel filtration column equilibrated in 150 mM NaCl, 20 mM HEPES pH 7.5, and 5 mM DTT.

Expression and purification of TFIIC. For holo-TFIIC, all six genes were cloned into a pBig2ab plasmid, with an N-terminal TEV-cleavable his-tag on τ 95 and a C-terminal TEV-cleavable TwinStrep tag on τ 138. Baculovirus was generated following standard procedures and TFIIC was expressed in Hi5 cells as described for τ A. TFIIC was purified over Ni-NTA as described for τ A, but eluted with 300 mM imidazole. The eluate was loaded on a 5 mL StrepTactin XP column (IBA Life Sciences), washed with 25 mL Strep-A buffer (150 mM NaCl, 50 mM Tris-HCl pH 7.5, and 5 mM DTT) and eluted with Strep-B buffer (Strep-A supplemented with 50 mM biotin). Peak fractions were concentrated and applied to a Superose 6 INCREASE 10/300 gel filtration column equilibrated in 150 mM NaCl, 20 mM HEPES pH 7.5, and 5 mM DTT.

Expression and purification of Brf1-TBP and Bdp1. The Brf1-TBP plasmid was transformed into BL21 Star (DE3) pRARE *E. coli* cells. Expression cultures were grown at 37 °C in TB medium to an optical density at 600 nm (OD) of ~ 1.0 and induced with 50 μ M IPTG overnight at 16 °C. Cells were pelleted for 5 min at 12,000 g and resuspended in 3 mL lysis buffer (1 M NaCl, 50 mM Tris-HCl pH 7.5, 2 mM bME), 20% glycerol, 10 μ g/mL DNase I, 1 x protease inhibitors (SigmaFast protease inhibitor cocktail EDTA free), 30 mM imidazole, and 50 mL His-A low salt (His-A but with 150 mM NaCl) and eluted with 50 mL His-B (200 mM NaCl, 2 mM bME, 5% glycerol, 300 mM imidazole). The eluate was loaded on a 5 mL HiTrap Heparin column (GE healthcare) pre-equilibrated in HepA buffer (like His-B but without imidazole). The column was washed with 6 column volumes (CV) containing 30% HepB (HepA but 1 M NaCl) and eluted with a linear gradient from 30% HepB to 70% HepB over 20 CV. Brf1-TBP eluted at ~ 600 mM NaCl. Peak fractions were

concentrated and applied to a HiLoad 16/600 Superdex 200 size-exclusion column equilibrated in 300 mM NaCl, 25 mM HEPES pH 7.5, 5 mM DTT and 5% glycerol. Purified Brf1–TBP was concentrated to ~6 mg/mL.

The Bdp1 plasmid30 was transformed in BL21 Star (DE3) pRARE *Escherichia coli* cells and grown in TB medium to an OD of ~1.0, cooled down, and induced with 100 μ M IPTG overnight at 18 °C. Cell harvesting, lysis, and Ni-NTA chromatography were performed as for Brf1–TBP. Eluted proteins were loaded on a 5 mL Heparin column pre-equilibrated in HepA. The column was washed with 6 CV of 20% HepB and eluted with a gradient from 20% HepB to 70% HepB over 30 CV. Bdp1 (in ~520 mM NaCl) was cleaved with TEV protease overnight at 4 °C and incubated with 3 mL Ni-NTA for 1 h. Bdp1 was recovered using 100 mM imidazole buffer and purified by SEC as for Brf1–TBP (but in a buffer containing 150 mM NaCl).

DNA oligonucleotides. Experiments were performed with different lengths of the His_{th}(GUG)E2 gene. For simplicity, only the sequence of the non-template strand is given here. All oligonucleotides were obtained from Sigma Aldrich. For all experiments, DNA was annealed by mixing equimolar amounts of the non-template and template strands in H₂O followed by heating to 95 °C for 10 min before cooling to 20 °C at a rate of 1.5 °C/min.

For transcription assay, we used the sequence of the His_{th}(GUG)E2 gene, including 44 base pairs upstream of the TSS (referred to as *His_{th} complete*, non-template 5'-GTATTACTCGAGCCCGTAATACAACAGTCTCCATTGAAAAGTCGCCATCTTAGTATAGTGGTATAGTACACATCGTTGTGGCCGATGAAACCCCTGGTTCGATTCTAGGAGATGGCATTTT-3', the +1 nt is printed in bold and A-box and B-box are underlined).

For TFIIB– τ A-binding experiments, we used *His_{th} upstream_Abox* oligonucleotide (non-template strand 5'-AGCCCGTAATACAACAGTCTCCATTGAAAAGTCGCCATCTTAGTATAGTGGTATAG-3').

For TFIIC–DNA-binding experiments, we used a *His_{th} Abox_Bbox* oligonucleotide encompassing the A- and B-box elements. For the EMSA assay, the non-template strand had a 6FAM fluorophore attached to the 3' end (non-template: 5'-CATCTTAGTATAGTGGTATAGTACACATCGTTGTGGCCGATGAAAACCTGGTTCGATTCTAGG6FAM-3'). For EMSAs with τ A, we used an oligonucleotide encompassing only the A-box with a 6FAM fluorophore label on the 3' end of the non-template strand (*His_{th} Abox*, non-template: 5'-CATCTTAGTATAGTGGTATAGT6FAM-3').

Electrophoretic mobility shift assay. τ A EMSA samples contained 1 μ M 6FAM labeled A-box DNA and 1 μ M, 5 μ M, or 10 μ M τ A construct in 100 mM NaCl, 20 mM HEPES pH 7.5, and 5 mM DTT.

TFIIC EMSA samples contained 0.5 μ M 6FAM labeled A-box–B-box DNA, and two-fold serial dilutions of TFIIC, ranging from 2.18 μ M to 0.03 μ M.

Samples were incubated for 1 h on ice, supplemented with 10% glycerol and loaded on a 3–8% Tris-acetate gel (Thermo Fisher Scientific) in the running buffer: 2.5 mM Tris base, 19 mM glycine, and 1 mM DTT. Gels were imaged on a Typhon FLA9500 phosphorimager.

τ A–TFIIB–Pol III reconstitution experiment. A total of 108 pmol of His-DNA were incubated with an equimolar amount of τ A for 15 min, followed by incubation with an equimolar amount of TFIIB for another 15 min. The sample was diluted fourfold with SEC buffer (100 mM KCl, 20 mM HEPES pH 7.5, 5 mM DTT, and 2 mM MgCl₂) and an equimolar amount of Pol III was added. The sample is incubated overnight before being applied to a Superose 6 INCREASE 3.2/300 column equilibrated in SEC buffer or loaded on a 15–45% glycerol gradient. After centrifugation, glycerol gradients were manually fractionated by removing 200 μ L fractions from the top of the gradient, and fractions were analyzed on 4–12% NuPAGE gels followed by silver staining.

EM sample preparation. τ A was incubated with equimolar amounts of double-stranded *His_{th} upstream_Abox* DNA, followed by the addition of the Brf1–TBP fusion protein and finally full-length Bdp1. The sample was incubated for 10 min before adding the next complex component, respectively. The complex was subsequently diluted to 1.3 mg/mL in dilution buffer (20 mM HEPES pH 7.5, 5 mM DTT, and 2 mM MgCl₂) and thereby adjusted to 75 mM NaCl. Directly before grid freezing 4 mM CHAPSO (3-([3-cholamidopropyl]dimethylammonio)-2-hydroxy-1-propanesulfonate in water) was added to the protein sample. A total of 2.5 μ L of protein sample were applied to Quantifoil Cu 2/1 which had been previously glow discharged with a Pelco EasyGlow instrument. Excess liquid was removed with a Vitrobot Mark IV chamber (Thermo Fischer Scientific) at 4 °C, and 100% humidity for 4 s and at a blot force of 2.

Cryo-EM and data processing. Cryo-EM data were collected on a Titan Krios microscope with a Gatan Quantum energy filter and a K2 Summit direct detector in counting mode. For τ A, data were collected at a magnification of 130,000 \times and a calibrated pixel size of 1.041 $\text{\AA}/\text{px}$. We collected 5824 movies with an accumulated dose of 48.7 $e^-/\text{\AA}^2$ over 36 frames, using a target defocus of –0.5 to –2 μ m. Movies were preprocessed on the fly using wARP 1.06 (ref. 52). The model parameters for motion correction and CTF estimation were set to 5 \times 5 \times 36 and

5 \times 5 \times 1, respectively. Particles were picked with BoxNet2_20180918 without retraining, using an expected diameter of 150 \AA . Particles were inverted, normalized and exported in a 300 pixel box. Particles were divided into three batches, and each batch classified using 3D classification in RELION 3.0 (ref. 53), using a 40 \AA low-pass filtered negative stain model of τ A as reference. The best class of each batch was retained, batches were joined and particles cleaned using 2D classification in RELION (setting the “Ignore CTFs until first peak” option to “yes”). Particles were refined, and then re-extracted from micrographs that were aligned using MotionCor2 and CTF corrected with gCTF to allow for Bayesian polishing inside RELION. Two rounds of CTF refinement (using per-particle defocus and beam tilt estimation) and Bayesian polishing were performed, yielding a map of 3.06 \AA resolution.

τ A model building and refinement. We initially placed the crystal structures of the τ 131 N-term TPR (PDB 5aem) and the τ 55 HPD (PDB 2yn0), as well as homology models of the τ 55– τ 95 β -barrel (modeled on PDB 4bjj) and the τ 95 dimerization domain (modeled on PDB 4bjj) into the density. Homology models were generated with Phyre2 (ref. 54). A partial model of the C-terminal TPR array of τ 131 was generated with the ARP/wARP⁵⁵ webserver. For this, we supplied the sequence of the C-terminal half of τ 131 (residues 573–1025) and a partial map obtained by subtracting the fitted densities and removing noise, using the UCSF chimera⁵⁶ volume subtraction and volume eraser tools. ARP/wARP generated a backbone model of a large portion of the C-terminal TPR and also contained a stretch of correctly assigned sequence. Starting from this partial model, we manually completed the model in COOT⁵⁷.

Due to the high quality of the map and relatively small variations in local resolution, we could unambiguously assign the correct sequence to all density except of two connected α -helices that pack against residues 984–933 in τ 131. We assign this density to subunit τ 95, as it is located between residues 236–263, which are missing in our structure and have included the region as a poly-alanine model in the PDB file. The model was refined using ISOLDE⁵⁸ and phenix real-space refine⁵⁹.

Negative stain of τ A–Brf1–TBP. For obtaining the negative stain map of τ A–Brf1–TBP, we used a Brf1–TBP construct lacking the first 70 amino acids (corresponding to the Zn-ribbon and linker) because unlike Brf1–TBP, Brf1–TBP Δ N did not form precipitates when diluted rapidly into low salt buffers. A total of 40 μ g τ A were mixed with equimolar amounts of Brf1–TBP Δ N, and the sample was diluted to 75 mM NaCl and incubated for 15 min at 20 °C. The sample was then applied to a 10–30% (v/v) glycerol gradient containing glutaraldehyde. The gradient was prepared by layering 2.2 mL of heavy buffer (10% glycerol (v/v), 100 mM NaCl, 20 mM HEPES pH 7.5, 5 mM DTT, and 0.1% glutaraldehyde) underneath 2.2 mL of light buffer (10% glycerol (v/v), 100 mM NaCl, 20 mM HEPES pH 7.5, and 5 mM DTT) and rotating the tube for 55 s at an angle of 86° and a speed of 20 r.p.m. The gradient was allowed to settle for 2 h at 4 °C before the sample was applied on top of the gradient and centrifuged for 16 h at 150,000 \times g in a SW55Ti rotor. The gradient was fractionated by removing 200 μ L fractions from the top of the gradient by pipetting. Crosslinking was quenched by adding 10 μ L 1 M Tris-HCl pH 7.5 to each fraction. Fractions were analyzed by SDS–PAGE and silver staining (when crosslinker was omitted) or InstantBlue staining (when crosslinking was performed). A total of 5 μ L of sample from the indicated fraction in Supplementary Fig. 5 were applied to a freshly glow-discharged EM grid, the sample was incubated for 1 min on the grid, and then washed twice with water and then incubated for 1 min with 2% uranyl acetate solution. Excess uranyl acetate was blotted away with filter paper and the grid was allowed to dry. The sample was imaged on a Tecnai Spirit microscope operated at 120 kV at a nominal magnification of 49,000 and a pixel size of 2.292 \AA . A total of 211 micrographs were collected with a target defocus of 1 μ m. Micrographs were CTF corrected using GCTF and templates for autopicking in RELION were generated by an initial round of autopicking, using the Laplacian-of-Gaussian autopicker in RELION followed by 2D classification. A total of 10,731 template-picked particles were extracted from CTF-corrected micrographs with a 200 px box and cleaned through 2D classification in CryoSPARC. An initial model was generated from 7398 particles using the CryoSPARC ab initio reconstruction program. The model was then refined in RELION to an estimated resolution of 30 \AA .

Crosslinking-mass spectrometry. A total of 40 μ g of τ A were incubated with an equimolar amount of the Brf1–TBP fusion protein at a protein concentration of 2.5 μ M in 75 mM NaCl, 20 mM HEPES pH 7.5, and 5 mM DTT for 30 min at 20 °C. The sample was crosslinked for 1 h at 20 °C by addition of 5 mM H12/D12 isotope-coded di-succinimidyl-suberate (Creative Molecules) and quenched by addition 50 mM of ammonium bicarbonate for 10 min. Crosslinked proteins were denatured using urea and Rapigest (Waters) at a final concentration of 4 M and 0.05% (w/v), respectively. Samples were reduced using 10 mM DTT (30 min at 37 °C), and cysteines were carbamidomethylated with 15 mM chloroacetamide for 30 min in the dark. Proteins were digested first using 1:100 (w/w) LysC (Wako Chemicals) for 4 h at 37 °C, and then the urea concentration was reduced to 1.5 M and digested was finalized with 1:50 (w/w) trypsin (Promega) overnight at 37 °C. Samples were acidified with 10% (v/v) TFA and desalted using OASIS® HLB

μ Elution Plate (Waters). Crosslinked peptides were enriched using SEC. In brief, desalted peptides were reconstituted with SEC buffer (30% (v/v) ACN in 0.1% (v/v) TFA) and fractionated using a Superdex Peptide PC 3.2/30 column (GE) on a 1200 Infinity HPLC system (Agilent) at a flow rate of 50 μ L/min. Fractions eluting between 50–70 μ L were evaporated to dryness and reconstituted in 30 μ L 4% (v/v) ACN in 1% (v/v) FA. Collected fractions were analyzed by liquid chromatography-coupled tandem mass spectrometry (MS/MS). The mass spectrometric analysis was conducted using an UltiMate™ 3000 RSLCnano system (Thermo Fisher Scientific) directly coupled to an Orbitrap Fusion Lumos (Thermo Fisher Scientific). Peptides were loaded onto the trapping cartridge (μ -Precolumn C18 PepMap 100, 5 μ m, 300 μ m i.d. \times 5 mm, 100 Å) for 5 min at 30 μ L/min (0.05% TFA in water). Peptides were eluted and separated on an analytical column (nanoEase MZ HSS T3 column, 100 Å, 1.8 μ m, 75 μ m \times 250 mm) with a constant flow of 0.3 μ L/min using solvent A (0.1% formic acid in LC-MS grade water, Fisher Chemicals) and solvent B (0.1% formic acid in LC-MS grade acetonitrile, Fisher Chemicals). Total analysis time was 60 min with a gradient containing an 8–25% solvent B elution step for 39 min (min 6–45), followed by an increase to 40% solvent B for 5 min, 85% B for 4 min and 6 min of a re-equilibration step to initial conditions (2% B). The LC system was online coupled to the mass spectrometer using a Nanospray-Flex ion source (Thermo Fisher Scientific) and a Pico-Tip Emitter 360 μ m OD \times 20 μ m ID; 10 μ m tip (New Objective). The MS was operated in positive mode and a spray voltage of 2.4 kV was applied for ionization with an ion transfer tube temperature of 275 °C. Full scan MS spectra were acquired in profile mode for a mass range of 375–1600 m/z at a resolution of 120,000 (RF Lens 30%, AGC target $2e^5$ ions, and maximum injection time of 250 ms). The instrument was operated in data-dependent mode for MS/MS acquisition. Peptide fragment spectra were acquired for charge states 3–7. Quadrupole isolation window was set to 0.8 m/z and peptides were fragmented via CID (35% NCE). Fragment mass spectra were recorded in the ion trap at normal scan rate for a maximum of $2e^4$ ions (AGC target) or 100 ms maximum injection time. The instrument acquired MS/MS spectra for up to ten scans between MS scans. Dynamic exclusion was set to 60 s.

The data analysis was performed using xQuest and xProphet and all crosslinks at the 5% false discovery rate (FDR) were exported (Supplementary Data 1)⁶⁰. To select an optimal crosslink score confidence threshold for analysis, the identified crosslinks were mapped to the τ A and Brf1–TBP structures using Xlink Analyzer⁶¹. At the ld score (linear discriminant, as calculated by xQuest) threshold of 40, 25 crosslinks could be mapped to the structures with four of them exceeding the distance of 35 Å (two within τ A and two within Brf1–TBP). Thus, only the crosslinks with the ld score at least 40 were selected for further analysis. We attribute the violated crosslinks to the flexibility of the complex in the absence of DNA and to the expected false positive identification rate at 5% FDR.

Yeast tagging and viability assay. To construct τ 95 Δ plug and τ 95 Δ tail yeast strains with a STOP codon after amino acid 521 or 592, respectively, C-terminal deletion cassettes with a ClonNat marker were amplified from pFA6a-NatNT2 (ref. 62) and transformed into yeast BSY17 (ref. 63). Colonies were selected on YPD plates containing 100 μ g/mL of ClonNat, and analyzed by colony PCR and subsequent sequencing. For the viability assay, serial dilutions of the deletions strains or the parental strain were spotted on YPD medium.

Fitting to the negative stain EM map. To fit the τ A and Brf1–TBP structures to the negative stain EM map and to determine the handedness of the map, we applied a procedure that explores a large number of possible fits and evaluates the fits, using four different EM fitting scores. First, we generated samples of alternative fits using the unbiased global fitting approach based on FitMap tool from UCSF Chimera software⁵⁶, which places fitted structure at random locations in the EM map, optimizes the fits locally, clusters them by similarity, and keeps best scoring solution belonging to the cluster. Prior to the fitting, the structures were converted into a simulated EM map at the resolution of 33.5 Å to approximate the resolution of the negative stain EM map. The “overlap” score of Chimera (sum of aligned density products)⁵⁶ was used for fitting, as out of all scores available in Chimera this score is more suitable for negative density EM maps, as confirmed by its high correlation with the envelope scores used below. Because for EM maps the correct mirror image is unknown without additional information (such as a known reference map or atomic structure that fits unambiguously to a large portion of the map), both mirror maps were used for fitting. The fitting runs were performed using 100,000 random initial placements of the structures in the map and the requirement of at least 30% of the model map to be covered by the density envelope defined at the threshold of 0.06 with high clustering thresholds and the number of optimization steps reduced from the default 2000 to 100, which resulted in alternative 20,000–24,000 fits per structure per mirror map. The fits were then scored and ranked with alternative EM fitting scores: overlap, cross-correlation, Chamfer distance, and the envelope score, with the latter two selected due to their ability of scoring matches between two surfaces (an intensity threshold of 0.06 was selected for defining the envelopes of the negative stain map and the fitted map simulated from the query structure). The envelope scores correspond to the interpretation of a negative stain EM map as an approximation of the surface of the protein complex. The scores were calculated using UCSF Chimera⁵⁶ and TEMPy library⁶⁴. The map shown in Fig. 4 (referred to as mirror map 1 in Supplementary Fig. 6) led to

higher scoring fits of τ A for all four score types and resulted in the same top fit regardless of the score (Fig. 4 and Supplementary Fig. 6a), indicating the mirror 1 as the correct image of the negative stain EM map. For neither of the mirror maps, however, the fits were statistically significant when evaluated as published previously⁶⁵. Thus, for each mirror map, we generated pairwise combinations of the τ A and Brf1–TBP fits, and assessed them with the above scores. For computational efficiency, to generate the pairwise combinations, the τ A and Brf1–TBP structures were fitted again, but now with low clustering thresholds resulting in lower number of representative fits (25–50 fits per structure per mirror map after clustering). The combinations of these fits were then generated using Integrative Modeling Platform⁶⁶. This led 1136 and 1200 combined fits, respectively, for mirror1 and mirror2, where a “combined fit” represents a candidate model, in which τ A and Brf1–TBP are fitted to the EM map. Higher scoring combined fits were obtained again for the mirror map 1, giving additional confidence that we identified the correct handedness. The fitting yielded one unique fit of τ A and several alternative fits of Brf1–TBP among the top scoring solutions (Supplementary Fig. 6b). The alternative fits of Brf1–TBP adopted the same location in the EM map relatively to the N-terminal arm of τ 131, but different orientations (with some orientations rotated 180° to the others). Thus, as the representative fit of τ A we selected the fit that (1) obtained the highest score out of all fits of τ A, and for all four scoring measures, (2) agrees with the interactions from the τ 131 arm to Brf1–TBP suggested previously (refs. 26,29–31,35,36) and the crosslinking data from this work (Supplementary Fig. 6b). However, due to relatively low difference in EM fitting scores and because only three crosslinks between Brf1–TBP and τ A could be mapped to the structurally resolved parts of the model (Supplementary Fig. 6b), the exact orientation of Brf1–TBP remained undefined, although its overall position in the map is defined by our confident identification of the τ A fit (Supplementary Fig. 6c).

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

Data availability

The cryo-EM map of the τ A complex has been deposited to the Electron Microscopy Data Bank (EMDB) under the accession code EMD-10817. The coordinates of the corresponding model have been deposited to the PDB under accession code 6YJ6. The negative stain map has been deposited at the EMDB under accession code EMD-10795. The negative stain map as well as the fitted τ A structure is also available as a chimera session (Supplementary Data 2). The MS proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE⁶⁷ partner repository with the dataset identifier PXD018232. Source data for Fig. 2 and Supplementary Figs. 1 and 7 are available with the paper online. Source data are provided with this paper.

Received: 14 April 2020; Accepted: 8 September 2020;

Published online: 30 September 2020

References

- Ramsay, E. P. & Vannini, A. Structural rearrangements of the RNA polymerase III machinery during tRNA transcription initiation. *Biochim. Biophys. Acta Gene Regul. Mech.* **1861**, 285–294 (2018).
- Vorländer, M. K., Khatter, H., Wetzler, R., Hagen, W. J. H. & Müller, C. W. Molecular mechanism of promoter opening by RNA polymerase III. *Nature* **553**, 295–300 (2018).
- Abascal-Palacios, G., Ramsay, E. P., Beuron, F., Morris, E. & Vannini, A. Structural basis of RNA polymerase III transcription initiation. *Nature* **553**, 301–306 (2018).
- Stillman, D. J. & Geiduschek, E. P. Differential binding of a *S. cerevisiae* RNA polymerase III transcription factor to two promoter segments of a tRNA gene. *EMBO J.* **3**, 847–853 (1984).
- Gerlach, V. L., Whitehall, S. K. & Geiduschek, E. P. TFIIB placement on a yeast U6 RNA gene in vivo is directed primarily by TFIIC rather than by sequence-specific DNA contacts. *Mol. Cell. Biol.* **15**, 1455–1466 (1995).
- Geiduschek, E. P. & Tocchini-Valentini, E. G. Transcription by RNA polymerase III. *Annu. Rev. Biochem.* **57**, 873–914 (1988).
- Schultz, P. et al. The two DNA-binding domains of yeast transcription factor as observed by scanning transmission electron microscopy. *EMBO J.* **8**, 3815–3824 (1989).
- Marzouki, N., Camier, S., Ruet, A., Moenne, A. & Sentenac, A. Selective proteolysis defines two DNA binding domains in yeast transcription factor τ . *Nature* **323**, 176–178 (1986).
- Male, G. et al. Architecture of TFIIC and its role in RNA polymerase III pre-initiation complex assembly. *Nat. Commun.* **6**, 7387 (2015).
- Moqtaderi, Z. & Struhl, K. Genome-wide occupancy profile of the RNA polymerase III machinery in *Saccharomyces cerevisiae* reveals loci with incomplete transcription complexes. *Mol. Cell. Biol.* **24**, 4118–4127 (2004).

11. Moqtaderi, Z. et al. Genomic binding profiles of functionally distinct RNA polymerase III transcription complexes in human cells. *Nat. Struct. Mol. Biol.* **17**, 635–640 (2010).
12. Roberts, D. N., Stewart, A. J., Huff, J. T. & Cairns, B. R. The RNA polymerase III transcriptome revealed by genome-wide localization and activity-occupancy relationships. *Proc. Natl Acad. Sci. USA* **100**, 14695–14700 (2003).
13. Oler, A. J. et al. Human RNA polymerase III transcriptomes and relationships to Pol II promoter chromatin and enhancer-binding factors. *Nat. Struct. Mol. Biol.* **17**, 620–628 (2010).
14. Büchel, G. et al. Association with aurora-A controls N-MYC-dependent promoter escape and pause release of RNA polymerase II during the cell cycle. *Cell Rep.* **21**, 3483–3497 (2017).
15. Yuen, K. C., Slaughter, B. D. & Gerton, J. L. Condensin II is anchored by TFIIC and H3K4me3 in the mammalian genome and supports the expression of active dense gene clusters. *Sci. Adv.* **3**, e1700191 (2017).
16. Ferrari, R. et al. TFIIC binding to Alu elements controls gene expression via chromatin looping and histone acetylation. *Mol. Cell* **77**, 475–487 (2020).
17. Taylor, N. M. I. et al. Structural and functional characterization of a phosphatase domain within yeast general transcription factor IIIC. *J. Biol. Chem.* **288**, 15110–15120 (2013).
18. Taylor, N. M. I., Baudin, F., Von Scheven, G. & Müller, C. W. RNA polymerase III-specific general transcription factor IIIC contains a heterodimer resembling TFIIF Rap30/Rap74. *Nucleic Acids Res.* **41**, 9183–9196 (2013).
19. Mylona, A. et al. Structure of the $\tau 60/\Delta\tau 91$ subcomplex of yeast transcription factor Iiic: insights into preinitiation complex assembly. *Mol. Cell* **24**, 221–232 (2006).
20. Kassavetis, G. A., Soragni, E., Driscoll, R. & 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. *Proc. Natl Acad. Sci. USA* **102**, 15406–15411 (2005).
21. Chen, J., Noble, A. J., Kang, J. Y. & Darst, S. A. Eliminating effects of particle adsorption to the air/water interface in single-particle cryo-electron microscopy: Bacterial RNA polymerase and CHAPSO. *J. Struct. Biol. X* **1**, 100005 (2019).
22. Holm, L. & Rosenström, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38**, W545–W549 (2010).
23. Schilbach, S. et al. Structures of transcription pre-initiation complex with TFIIB and Mediator. *Nature* **551**, 204–209 (2017).
24. Vangone, A., Spinelli, R., Scarano, V., Cavallo, L. & Oliva, R. COCOMAPS: A web application to analyze and visualize contacts at the interface of biomolecular complexes. *Bioinformatics* **27**, 2915–2916 (2011).
25. Dawson, N. L. et al. CATH: an expanded resource to predict protein function through structure and sequence. *Nucleic Acids Res.* **45**, 289–295 (2017).
26. Moir, R. D. & Willis, I. M. Tetratricopeptide repeats of TFC4 and a limiting step in the assembly of the initiation factor TFIIB. *Adv. Protein Chem.* **67**, 93–121 (2004).
27. Cloutier, T. E., Librizzi, M. D., Mollah, A. K., Brenowitz, M. & Willis, I. M. Kinetic trapping of DNA by transcription factor IIIB. *Proc. Natl Acad. Sci. USA* **98**, 9581–9586 (2001).
28. Geiduschek, E. P. & Kassavetis, G. A. The RNA polymerase III transcription apparatus. *J. Mol. Biol.* **310**, 1–26 (2001).
29. Moir, R. D., Sethy-Coraci, I., Puglia, K., Librizzi, M. D. & Willis, I. M. A tetratricopeptide repeat mutation in yeast transcription factor IIIC131 (TFIIC131) facilitates recruitment of TFIIB-related factor TFIIB70. *Mol. Cell Biol.* **17**, 7119–7125 (1997).
30. Moir, R. D., Puglia, K. V. & Willis, I. M. Interactions between the tetratricopeptide repeat-containing transcription factor TFIIC131 and its ligand, TFIIB70. Evidence for a conformational change in the complex. *J. Biol. Chem.* **275**, 26591–26598 (2000).
31. Moir, R. D., Puglia, K. V. & Willis, I. M. Autoinhibition of TFIIB70 binding by the tetratricopeptide repeat-containing subunit of TFIIC. *J. Biol. Chem.* **277**, 694–701 (2002).
32. Rameau, G., Puglia, K., Crowe, A., Sethy, I. & Willis, I. A. N. A mutation in the second largest subunit of TFIIC increases a rate-limiting step in transcription by RNA polymerase III. *Mol. Cell Biol.* **14**, 822–830 (1994).
33. Willis, I., Schmidt, P. & Söll, D. A selection for mutants of the RNA polymerase transcription apparatus: PCF1 stimulates transcription of tRNA and 5S RNA genes. *EMBO J.* **8**, 4281–4288 (1989).
34. Kastner, B. et al. GraFix: sample preparation for single-particle electron cryomicroscopy. *Nat. Methods* **5**, 53–55 (2008).
35. Liao, Y., Moir, R. D. & Willis, I. M. Interactions of Brf1 peptides with the tetratricopeptide repeat-containing subunit of TFIIC inhibit and promote preinitiation complex assembly. *Mol. Cell Biol.* **26**, 5946–5956 (2006).
36. Liao, Y., Willis, I. M. & Moir, R. D. The Brf1 and Bdp1 subunits of transcription factor TFIIB bind to overlapping sites in the tetratricopeptide repeats of Tfc4. *J. Biol. Chem.* **278**, 44467–44474 (2003).
37. Hsieh, Y. J., Kundu, T. K., Wang, Z., Kovelman, R. & Roeder, R. G. The TFIIC90 subunit of TFIIC interacts with multiple components of the RNA polymerase III machinery and contains a histone-specific acetyltransferase activity. *Mol. Cell Biol.* **19**, 7697–7704 (1999).
38. Kundu, T. K., Wang, Z. & Roeder, R. G. Human TFIIC relieves chromatin-mediated repression of RNA polymerase III transcription and contains an intrinsic histone acetyltransferase activity. *Mol. Cell Biol.* **19**, 1605–1615 (1999).
39. Manaud, N. et al. A chimeric subunit of yeast transcription factor IIIC forms a subcomplex with $\tau 95$. *Mol. Cell Biol.* **6**, 3191–3920 (1998).
40. Chaussivert, N., Conesa, C., Shaaban, S. & Sentenac, A. Complex interactions between yeast TFIIB and TFIIC. *J. Biol. Chem.* **270**, 15353–15358 (1995).
41. Kassavetis, G. A., Riggs, D. L., Negri, R., Nguyen, L. H. & Geiduschek, E. P. Transcription factor IIIB generates extended DNA interactions in RNA polymerase III transcription complexes on tRNA genes. *Mol. Cell Biol.* **9**, 2551–2566 (1989).
42. Joazeiro, C. A. P., Kassavetis, G. A. & Geiduschek, E. P. Alternative outcomes in assembly of promoter complexes: The roles of TBP and a flexible linker in placing TFIIB on tRNA genes. *Genes Dev.* **10**, 725–739 (1996).
43. Rozenfeld, S. & Thuriaux, P. Genetic interactions within TFIIC, the promoter-binding factor of yeast RNA polymerase III. *Mol. Genet. Genomics* **265**, 705–710 (2001).
44. Bartholomew, B., Kassavetis, G. A. & Geiduschek, E. P. Two components of *Saccharomyces cerevisiae* transcription factor IIIB (TFIIB) are stereospecifically located upstream of a tRNA gene and interact with the second-largest subunit of TFIIC. *Mol. Cell Biol.* **11**, 5181–5189 (1991).
45. Kassavetis, G. A. et al. The role of the TATA-binding protein in the assembly and function of the multisubunit yeast RNA polymerase III transcription factor, TFIIB. *Cell* **71**, 1055–1064 (1992).
46. Kramm, K. et al. DNA origami-based single-molecule force spectroscopy elucidates RNA Polymerase III pre-initiation complex stability. *Nat. Commun.* **11**, 2828 (2020).
47. Kassavetis, G. A., Blanco, J. A., Johnson, T. E. & Geiduschek, E. P. Formation of open and elongating transcription by rna polymerase III complexes. *J. Mol. Biol.* **226**, 47–58 (1992).
48. Kassavetis, G. A., Braun, B. R., Nguyen, L. H. & Peter Geiduschek, E. S. *cerevisiae* TFIIB is the transcription initiation factor proper of RNA polymerase III, while TFIIA and TFIIC are assembly factors. *Cell* **60**, 235–245 (1990).
49. Ciesla, M., Skowronek, E. & Boguta, M. Function of TFIIC, RNA polymerase III initiation factor, in activation and repression of tRNA gene transcription. *Nucleic Acids Res.* **46**, 9444–9455 (2018).
50. Patel, A. B. et al. Structure of human TFIID and mechanism of TBP loading onto promoter DNA. *Science* **362**, eaau8872 (2018).
51. Weissmann, F. et al. biGBac enables rapid gene assembly for the expression of large multisubunit protein complexes. *Proc. Natl Acad. Sci. USA* **19**, E2564–E2569 (2016).
52. Tegunov, D. & Cramer, P. Real-time cryo-electron microscopy data preprocessing with Warp. *Nat. Methods* **16**, 1146–1152 (2019).
53. Zivanov, J., Nakane, T. & Scheres, S. H. W. A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. *IUCr* **6**, 5–17 (2019).
54. Kelly, L. A., Mezulis, S., Yates, C., Wass, M. & Sternberg, M. The Phyre2 web portal for protein modelling, prediction, and analysis. *Nat. Protoc.* **10**, 845–858 (2015).
55. Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat. Protoc.* **3**, 1171–1179 (2008).
56. Pettersen, E. F. et al. UCSF Chimera - a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
57. Brown, A. et al. Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **71**, 136–153 (2015).
58. Croll, T. I. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 519–530 (2018).
59. Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **68**, 352–367 (2012).
60. Leitner, A., Walzthoeni, T. & Aebersold, R. Lysine-specific chemical cross-linking of protein complexes and identification of cross-linking sites using LC-MS/MS and the xQuest/xProphet software pipeline. *Nat. Protoc.* **9**, 120–137 (2014).
61. Kosinski, J. et al. Xlink analyzer: software for analysis and visualization of cross-linking data in the context of three-dimensional structures. *J. Struct. Biol.* **189**, 177–183 (2015).

62. 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).
63. Séraphin, B., Kretzner, L. & Rosbash, M. A U1 snRNA:pre-mRNA base pairing interaction is required early in yeast spliceosome assembly but does not uniquely define the 5' cleavage site. *EMBO J.* **7**, 2533–2538 (1988).
64. Farabella, I. et al. TEMPy: a Python library for assessment of three-dimensional electron microscopy density fits. *J. Appl. Crystallogr.* **48**, 1314–1323 (2015).
65. Dauden, M. I. et al. Architecture of the yeast Elongator complex. *EMBO Rep.* **18**, 264–279 (2017).
66. Russel, D. et al. Putting the pieces together: integrative modeling platform software for structure determination of macromolecular assemblies. *PLoS Biol.* **10**, e1001244 (2012).
67. Perez-Riverol, Y. et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* **47**, 442–450 (2019).

Acknowledgements

M.K.V. acknowledges support by the EMBL International PhD program. We thank H. Khatter for initial work on this project and W. Seifert-Davila for help with protein expression. We thank M. Rettel from the EMBL proteomics core facility for MS data acquisition and processing, and advice on XL-MS and deposition to the PRIDE database. We acknowledge W. Hagen and F. Weis for access to the EMBL Heidelberg cryo-EM platform, T. Hoffmann and J. Pecar for maintaining the high-performance computing environment at EMBL Heidelberg, and K. Remans for helpful discussions regarding insect cell expression construct design.

Author contributions

C.W.M. initiated and supervised the project. M.K.V. designed and interpreted experiments, cloned and purified proteins, collected cryo-EM data, built the model, and prepared figures. A.J. cloned, expressed and purified proteins, and prepared EM samples. F.B. performed the Pol III transcription assays. H.G. helped with cloning, protein expression, and performed yeast tagging and viability assays. K.K. and J.K. performed fitting into the negative stain density and XL-MS analysis. M.K.V. and C.W.M. wrote the manuscript with input from the other authors.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41467-020-18707-y>.

Correspondence and requests for materials should be addressed to C.W.M.

Peer review information *Nature Communications* thanks Feixia Chu, Christoph Engel, and Kenji Murakami for their contributions to the peer review of this report. Peer review reports are available.

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

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



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

© The Author(s) 2020