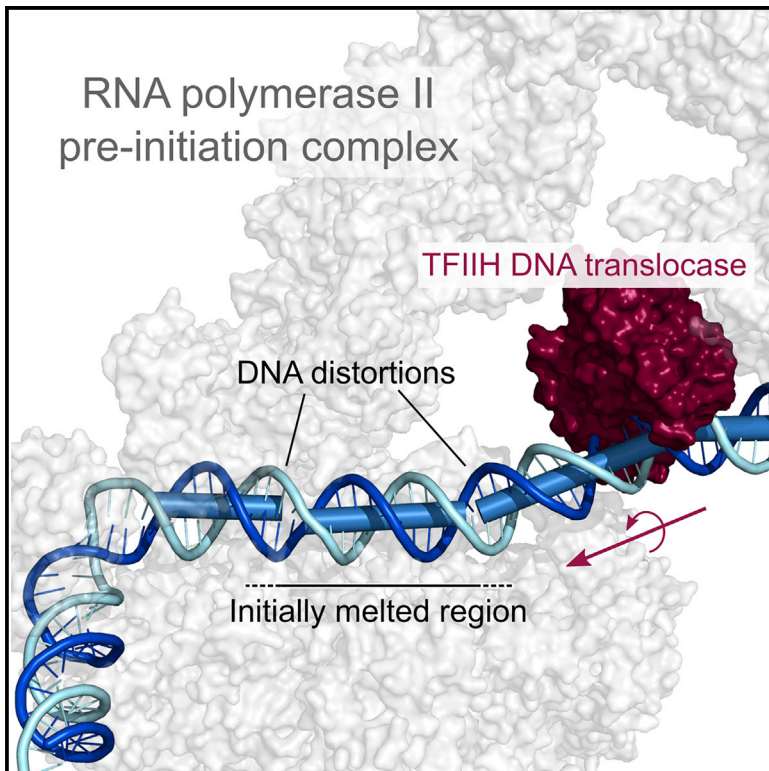


# Molecular Cell

## Promoter Distortion and Opening in the RNA Polymerase II Cleft

### Graphical Abstract



### Authors

Christian Dienemann, Björn Schwalb, Sandra Schilbach, Patrick Cramer

### Correspondence

patrick.cramer@mpibpc.mpg.de

### In Brief

Dienemann et al. show by cryo-EM structures and *in vivo* experiments how the pre-initiation complex primes promoter DNA of a gene for melting. The meltability of the promoter DNA modulates the dependence on the DNA translocase of TFIID. They identify general features of DNA opening by multi-subunit RNAPs.

### Highlights

- Promoter meltability defines requirement of TFIID for initiation
- DNA distortions are induced by clamp closure prior to DNA melting
- The initially melted region is structurally pre-defined by DNA distortions
- Clamp closure and DNA distortion are general features of multi-subunit RNAPs

# Promoter Distortion and Opening in the RNA Polymerase II Cleft

Christian Dienemann,<sup>1</sup> Björn Schwalb,<sup>1</sup> Sandra Schilbach,<sup>1</sup> and Patrick Cramer<sup>1,2,\*</sup>

<sup>1</sup>Max Planck Institute for Biophysical Chemistry, Department of Molecular Biology, Am Fassberg 11, 37077 Göttingen, Germany

<sup>2</sup>Lead Contact

\*Correspondence: [patrick.cramer@mpibpc.mpg.de](mailto:patrick.cramer@mpibpc.mpg.de)

<https://doi.org/10.1016/j.molcel.2018.10.014>

## SUMMARY

Transcription initiation requires opening of promoter DNA in the RNA polymerase II (Pol II) pre-initiation complex (PIC), but it remains unclear how this is achieved. Here we report the cryo-electron microscopic (cryo-EM) structure of a yeast PIC that contains underwound, distorted promoter DNA in the closed Pol II cleft. The DNA duplex axis is offset at the upstream edge of the initially melted DNA region (IMR) where DNA opening begins. Unstable IMRs are found in a subset of yeast promoters that we show can still initiate transcription after depletion of the transcription factor (TF) IIH (TFIIH) translocase Ssl2 (XPB in human) from the nucleus *in vivo*. PIC-induced DNA distortions may thus prime the IMR for melting and may explain how unstable IMRs that are predicted in promoters of Pol I and Pol III can open spontaneously. These results suggest that DNA distortion in the polymerase cleft is a general mechanism that contributes to promoter opening.

## INTRODUCTION

Transcription of a eukaryotic protein-coding gene starts with assembly of the RNA polymerase II (Pol II) transcription initiation machinery at the promoter (Grünberg and Hahn, 2013; Nogales et al., 2017; Sainsbury et al., 2015). The basal transcription factors (TFs) TBP (as part of TFIID), TFIIA, and TFIIB first bind to promoter DNA upstream of the transcription start site (TSS) and recruit the Pol II-TFIIF complex. TFIIE binding then leads to the core pre-initiation complex (cPIC) that associates with TFIIH to form the pre-initiation complex (PIC). Opening of the initially melted DNA region (IMR) occurs ~20–30 base pairs (bp) downstream of the TATA box (Holstege et al., 1997) and converts the closed complex (CC) to an open complex (OC). RNA synthesis then leads to the initially transcribing complex (ITC) and enables Pol II to escape from the promoter as an elongation complex (EC) (Martinez-Rucobo and Cramer, 2013).

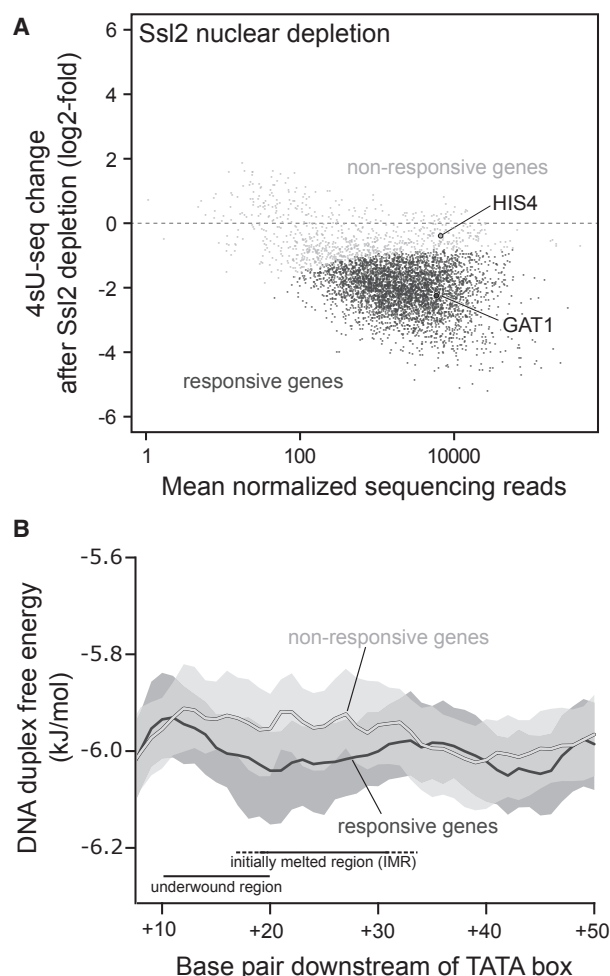
Human and yeast cPIC structures were determined in the form of the CC, OC, and ITC (He et al., 2013, 2016; Murakami et al.,

2015; Plaschka et al., 2015, 2016). The OC and ITC structures are very similar in the human and yeast systems with respect to the position of DNA and TFs and show a closed active center cleft with a closed clamp domain. By contrast, CC structures exhibit differences in the human and yeast systems. In human CCs, promoter DNA has been observed in an open cleft (He et al., 2013, 2016), whereas a yeast CC structure showed DNA located above the closed cleft (Plaschka et al., 2016). It was suggested that cleft opening and DNA loading into the cleft are stabilized by the formation of additional contacts between TFIIE and the clamp (He et al., 2016; Plaschka et al., 2016).

Promoter DNA opening involves TFIIE and TFIIH. TFIIH contains the ATP-dependent DNA translocase subunit Ssl2/XPB (yeast/human). Ssl2/XPB is located ~60 bp downstream of the TATA box and acts as a molecular wrench to rotate and translocate downstream promoter DNA into the Pol II cleft (Fishburn et al., 2015; He et al., 2016; Schilbach et al., 2017). TFIIE and TFIIH are generally recruited to Pol II promoters *in vivo* (Rhee and Pugh, 2012), but TFIIE and the ATPase activity of TFIIH are not required for transcription initiation either at certain promoters (Goodrich and Tjian, 1994; Pan and Greenblatt, 1994; Plaschka et al., 2016) or on supercoiled DNA (Holstege et al., 1995; Parvin and Sharp, 1993; Tyree et al., 1993). Indeed, TFIIH-independent DNA opening was observed for the yeast CC *in vitro* (Plaschka et al., 2016), and transcription can occur *in vivo* upon depletion of XPB in human cells (Alekseev et al., 2017). The mechanism for such ATP-independent promoter opening, however, remains unclear.

Despite these recent advances in our understanding of promoter opening in the Pol II system, several questions remain. First, what are the molecular changes during the CC-to-OC transition that lead to promoter opening in a defined region and formation of the IMR? Second, do TFIIH-dependent and -independent DNA opening use the same molecular pathway? Third, how do the mechanisms of DNA opening differ between Pol II and other eukaryotic transcription systems that do not require ATP for DNA opening?

Here we address these questions using a combination of functional and structural approaches. We report a previously unobserved structure of a Pol II PIC containing a distorted, closed promoter DNA, which may represent an intermediate of initiation. Although the heterogeneity of Pol II promoters indicates that several aspects of the mechanism of transcription initiation may differ between genes, we suggest that coupled DNA distortion and polymerase cleft closure could be a general feature of the DNA opening mechanism.



**Figure 1. Ssl2 Dependence for Transcription Varies In Vivo**

(A) Transcription of a subset of ~18% of protein-coding genes are not significantly affected after nuclear depletion of the TFIIF translocase subunit Ssl2. Points mark the log2-fold change of newly synthesized RNA upon Ssl2 nuclear depletion versus the normalized mean read count across replicates (Anders and Huber, 2010). Genes that respond significantly to Ssl2 depletion are shown in dark gray, non-responding genes without significantly altered transcription in light gray.

(B) DNA duplex free energy of responding and non-responding genes from (A) differs in the IMR of their promoters. The DNA duplex free energy (SantaLucia, 1998) of the promoter region is shown for responsive (dark gray) and non-responsive (light gray) TATA-containing genes. Sequences were aligned at the TATA box, and base coordinates are given relative to the beginning of the TATA box. Confidence intervals are shown as areas around the traces.

## RESULTS

### TFIIF Translocase Dependence for Transcription Varies In Vivo

Our previous results obtained with the yeast system *in vitro* (Plaschka et al., 2016) suggested that the TFIIF translocase subunit Ssl2 (human XPB) is not strictly required to open promoter DNA and to initiate transcription, but this was not tested *in vivo*. To investigate the genome-wide dependence of transcription

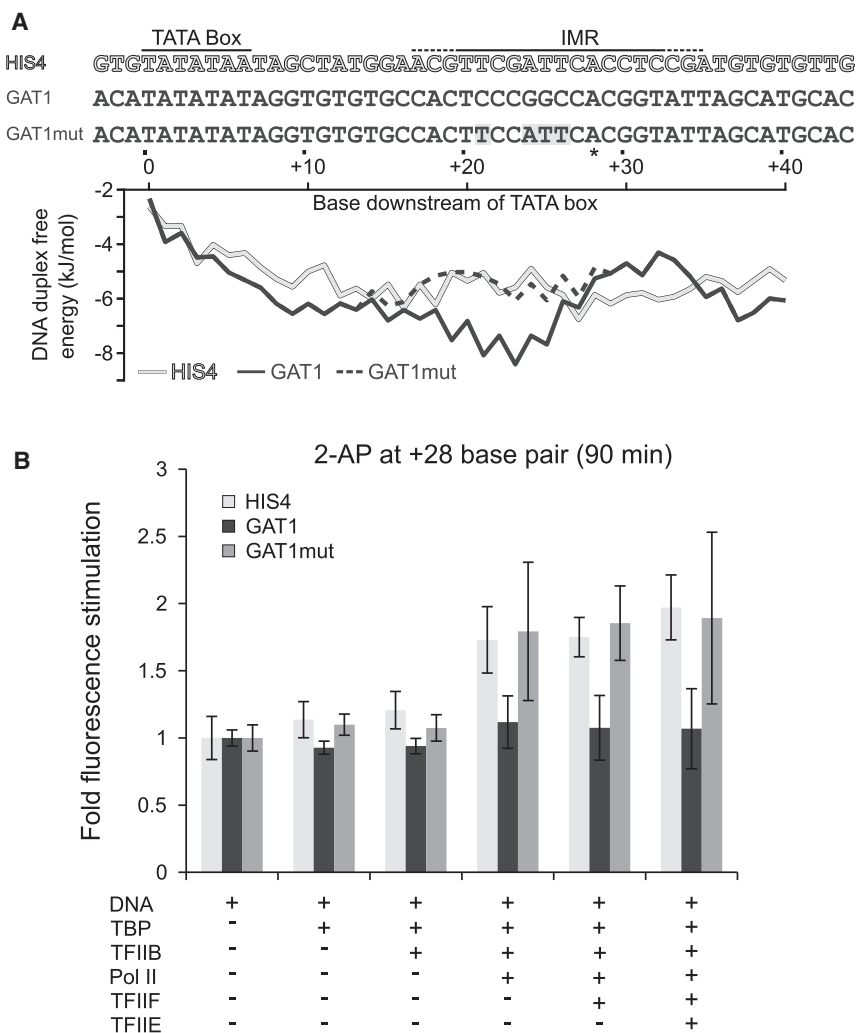
initiation on Ssl2, we transiently depleted Ssl2 from the yeast nucleus using the anchor-away technique (Haruki et al., 2008). We verified by western blotting that nuclear Ssl2 protein levels are strongly reduced after depletion (Figure S1A). Although *in vitro* reconstitution shows that the core of TFIIF is stable in the absence of Ssl2 (Figure S1B), we cannot exclude that other TFIIF subunits are depleted together with Ssl2. We further showed by chromatin immunoprecipitation that Ssl2 occupancy at genes is almost entirely lost under depletion conditions (Figure S1C). We then monitored newly synthesized RNA by 4-thiouracil sequencing (4sU-seq) (Schulz et al., 2013) and analyzed spike-in normalized sequencing reads for differential expression (STAR Methods).

We obtained groups of responsive and non-responsive genes (Figure 1A). The majority of genes responded to the nuclear depletion of Ssl2, confirming the importance of Ssl2 for transcription *in vivo*. However, RNA synthesis from 18% of genes was not significantly altered upon Ssl2 depletion, indicating that a fraction of yeast promoters may be opened in the absence of Ssl2. This fraction was enriched for genes involved in ribosome biogenesis, which clearly depend less on Ssl2 (Figure S1D). This group of genes is essential for cell growth and transcribes at high levels when nutrients are available. Responsive and non-responsive genes were indistinguishable with respect to the presence of a TATA box element in their promoters, suggesting that DNA opening is independent of the TATA box. These results show that the dependence of transcription on the TFIIF translocase varies between genes and indicate that spontaneous DNA opening may occur for a fraction of yeast Pol II promoters *in vivo*.

### Translocase Dependence Correlates with Promoter Stability

The observation that a fraction of genes can be transcribed after depletion of the translocase Ssl2 *in vivo* suggests that the meltability of the corresponding promoters is higher. To find possible determinants of promoter meltability, we analyzed IMRs in promoters of two groups of genes that were either or were not responding to Ssl2 depletion. Neither of the two gene groups showed a bias in base pair composition (Figure S1E). However, when calculating the free energy of the DNA duplex by the nearest-neighbor method (SantaLucia, 1998), we found that genes that were not responding to Ssl2 depletion contained less-stable DNA duplexes between positions +15 and +30 downstream of the TATA box (Figure 1B). Although the average difference was not very pronounced, the difference can be as large as 2.5 kJ/mol when individual promoters are compared. These results indicate that lower DNA duplex stability within the IMR leads to a higher promoter meltability and can compromise or circumvent a Ssl2 dependence for transcription *in vivo*.

Genes that did not respond to Ssl2 depletion included *HIS4* (Figure 1A), suggesting that the *HIS4* promoter can melt easily. This explained why we obtained OCs in our previous structural studies of PICs with *HIS4* promoter DNA, although closed DNA was used and TFIIF and ATP were absent in these experiments (Plaschka et al., 2016). In search for a promoter that would melt less easily, we picked from our *in vivo* data the *GAT1* promoter. *GAT1* belongs to the genes that strongly respond to Ssl2



**Figure 2. Promoter Opening Depends on the Initially Melted Region**

(A) Promoter DNA scaffolds used in DNA opening assays. The sequence of the non-template strand is shown aligned at the TATA box. The DNA-duplex free energy (SantaLucia, 1998) of each scaffold is shown below. The position of the 2-aminopurine (2-AP) label is marked with an asterisk. Bases of the GAT1 promoter that are mutated in GAT1mut are highlighted.

(B) The HIS4 promoter but not the GAT1 promoter shows OC formation *in vitro* under conditions of our assay. Melting of the GAT1 promoter is achieved after its sequence was changed in the IMR to resemble the HIS4 promoter (GAT1mut). Bar plots show the fluorescence increase normalized to the DNA only reaction. Data are shown as mean  $\pm$  SEM.

depletion (Figure 1A), and its IMR is predicted to be much more stable than the IMR in the HIS4 promoter (Figure 2A).

### Promoter Stability Defines DNA Meltability

We then tested the meltability of the HIS4 and GAT1 promoters in a reconstituted, TFIH-free *in vitro* system (Figure 2). To monitor DNA opening, we incorporated 2-aminopurine (2-AP) in the IMR at the base pair located +28 positions downstream of TATA and measured the increase in fluorescence upon single-strand formation (Kashkina et al., 2007). In the presence of only DNA, TBP, and TFIIB, the fluorescence signal did not change after 90 min of incubation, providing a negative control (Figure 2B). When Pol II was added, DNA opening was detected for the HIS4 promoter, but not for the GAT1 promoter (Figure 2B), consistent with our *in vivo* data. Promoter DNA melting could be observed across the entire IMR, and the signal corresponds to ~40% OC formation when compared with a control DNA that contained a mismatched region (Figures S2A and S2B). The GAT1 promoter showed DNA opening only after 22 hr of incubation (Figure S2C). Opening required TBP and TFIIB, but not TFIIE and TFIIF, resembling the archaeal transcription system,

which only requires counterparts of TBP and TFIIB for initiation (Spitalny and Thomm, 2003). The long incubation times required to observe DNA opening are likely due to our use of a minimal *in vitro* system and incubation at room temperature. Although in these assays we cannot exclude a contribution of protein binding to the measured signals, the results reveal a difference in meltability between IMR sequences.

To confirm that the observed differences in promoter meltability stem from differences in the sequence of their IMRs, we mutated the IMR in the GAT1 promoter. We changed positions +21, +24, +25, and +26 downstream of TATA, resulting in a mutant GAT1 promoter, GAT1mut (Figure 2A). Indeed, GAT1mut DNA opened

within 90 min, like the HIS4 promoter (Figure 2B). In agreement with this, the DNA duplex free energy of the GAT1mut promoter IMR is very similar to that of the HIS4 promoter (Figure 2A). Thus, the sequence and the DNA duplex stability of the IMR influences OC formation *in vitro*. This is in agreement with our finding that an unstable IMR correlates with Ssl2 dependence *in vivo*. Thus, the stability of the IMR defines promoter meltability in this system.

### Cryo-EM Reveals Distinct Closed Complex Structures

Based on these results, we used GAT1 promoter DNA to form a more stable yeast CC *in vitro*. We prepared core PICs with GAT1 promoter DNA, TBP, TFIIB, TFIIE, TFIIF, and Pol II for cryo-EM analysis and collected ~3,800 micrograph images on a Titan Krios with a Gatan K2 direct electron detector in EFTEM mode (Figure S3A; Table 1). The initially picked set of particles was cleaned by 2D classification (Figure S3B). The remaining 385,000 particle images were classified in 3D, applying a mask around the IMR of promoter DNA. This led to 24% CC and 76% OC particles (Figure S3C). Thus, spontaneous DNA opening still occurred during sample preparation but was reduced



**Table 1. Data Collection and Processing Statistics**

	cCC1	cCC <sup>dist</sup>	CC <sup>dist</sup>
<b>Data Collection</b>			
Acc. Voltage (kV)	300	300	300 <sup>a</sup>
Detector	K2 Summit (GIF)	K2 Summit (GIF)	K2 Summit (GIF) <sup>a</sup>
Electron dose (e <sup>-</sup> /Å <sup>2</sup> )	37	37	42 <sup>a</sup>
Defocus range (μm)	−0.8 to −3.0	−0.8 to −3.0	−0.5 to −5.0 <sup>a</sup>
Magnification (x)	130 000	130 000	105 000 <sup>a</sup>
Pixel size (Å/px)	1.07	1.07	1.37 <sup>a</sup>
Particles	39 000	38 000	60 000 <sup>a</sup>
Symmetry imposed	C1	C1	C1 <sup>a</sup>
<b>Reconstruction</b>			
Map overall resolution	5.1	4.8	6.7
Map sharpening B-factor (Å <sup>2</sup> )	−100	−100	−100
FSC threshold	0.143	0.143	0.143
<b>Model Composition</b>			
Non-hydrogen atoms	40,627	43,398	63,013
Protein residues	4,953	5,362	8,338
DNA bases	112	112	150
Ligand atoms	10	11	18
<b>Model Validation</b>			
MolProbity score	1.94	1.95	1.96
Clash score	9.12	9.22	9.18
Rotamer outliers (%)	0.1	0.1	0.3
Cβ-deviations	0	0	0
Ramachandran Favored (%)	92.9	92.8	92.4
Ramachandran Allowed (%)	8.0	6.2	6.4
Ramachandran Disallowed (%)	0.9	1.0	1.2

<sup>a</sup>Values from Schilbach et al., 2017

when compared with our previous study due to the use of the *GAT1* DNA that is more difficult to melt than the previously used *HIS4* DNA (Plaschka et al., 2016).

As a result, we obtained enough yeast CC particles for structure determination at higher resolution than previously possible (Plaschka et al., 2016). Classification of particles revealed two different structures of the core CC (cCC) (Figures S3C and S3D). Reconstructions were obtained at nominal resolutions of 5.1 Å and 4.8 Å with a well-resolved core and lower resolution at the periphery of the complex (Figure S3D). Structures were built based on our previous PIC structures (Plaschka et al., 2016; Schilbach et al., 2017) and refined in real space to yield good stereochemistry (STAR Methods).

The two cCC structures differed in the position of closed promoter DNA (Figure 3; Video S1). The first structure (cCC1) contained canonical B-DNA along the upper Pol II cleft (Figure 3A). This structure resembles the previously reported yeast cCC (Plaschka et al., 2016) and lacks TFIIE and the Tfg2 WH domain

(Figure S3D), which also showed weak density in the previous structure. In the second structure, we observed a region of promoter DNA that deviated substantially from B-DNA (Figures 3B and S3E), and we therefore called it distorted cCC (cCC<sup>dist</sup>). This structure revealed all components of the cPIC including TFIIE and the Tfg2 WH domain (Figure 3B). In both structures, the clamp is closed and the TFIIB B-linker is ordered (Figure S4). Thus, a single CC preparation gave rise to two different CC structures and the OC, indicating that the CC is dynamic and can be converted to the OC, and suggesting that the CC structures may correspond to intermediates on the path from the CC to the OC.

### Distorted DNA in the Closed Pol II Cleft

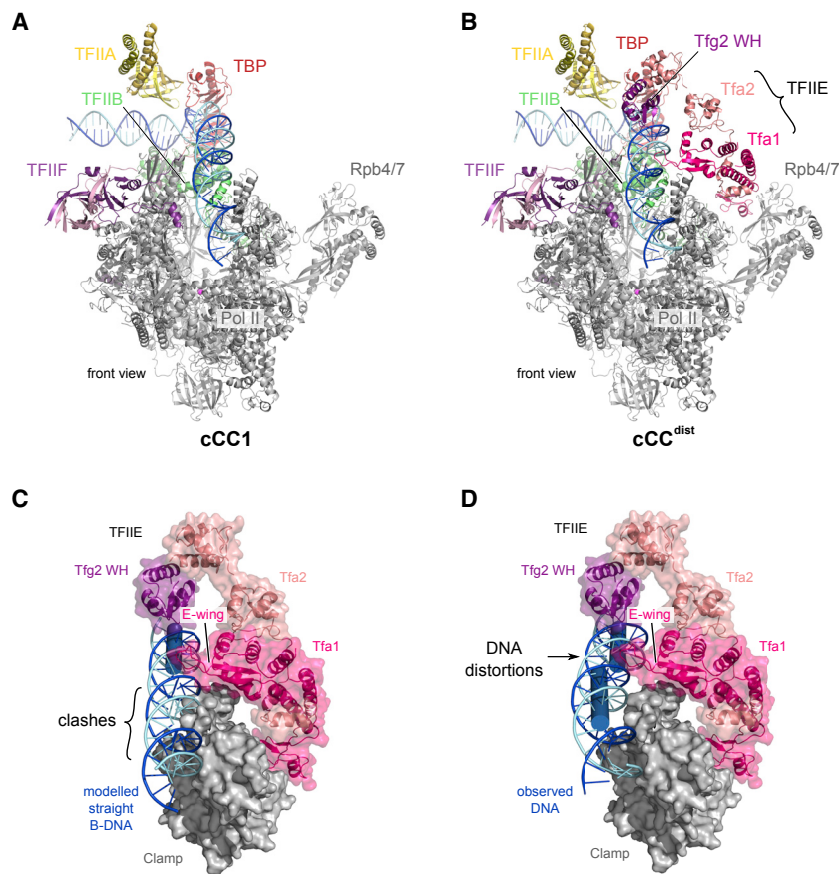
The cCC<sup>dist</sup> structure represents a previously unobserved conformation that may correspond to an initiation intermediate. In this structure, promoter DNA is loaded ~7 Å deeper into the cleft compared to cCC1 (Figure 4; Table S1). We observe an offset of the DNA helical axis by ~5 Å around base pair +20 downstream of the TATA box (Figure 4B), which corresponds to the position of the upstream edge of the IMR (Giardina and Lis, 1993). The DNA distortion is required for the clamp and cleft to be closed as observed here. Modeling shows that straight B-form DNA results in a clash with the clamp when the cleft is closed (Figure 3C). Consistent with this, a human CC structure with straight, undistorted DNA showed an open clamp and cleft (He et al., 2013, 2016) (Table S1).

The DNA distortion also includes an underwinding of base pairs between positions +10 and +20 downstream of TATA (Figure 4C; Video S1). DNA underwinding in this region is released in an OC (Figures 4B and 4C), suggesting that DNA is strained in cCC<sup>dist</sup> and that release of this strain facilitates DNA opening and the transition to the OC. The distorted, underwound DNA region interacts with the Tfg2 WH domain and the TFIIE extended winged helix (eWH) domain (Figure 3D), which may stabilize distorted DNA in the cleft. A hairpin protruding from the TFIIE eWH domain, the E-wing, binds above the distorted DNA ~23 bp downstream of the TATA box.

### Complete PIC Structure with Distorted DNA

To investigate whether the observed DNA distortions also occur in the presence of TFIH, we further solved the structure of the CC in the presence of TFIH. We re-classified particles from a large dataset of yeast PICs comprising TFIH and *HIS4* promoter DNA (Schilbach et al., 2017). Although this dataset contains mainly OCs, we could obtain 142,000 PIC particles with closed DNA (Figure S5A). A complete PIC structure with closed DNA could be refined from a subset of 60,000 particles to a nominal resolution of 6.7 Å (Figure 5 and Figure S5B; Table 1). In this structure, the distorted DNA region is observed at a local resolution of ~5 Å (Figure S5C), and this was sufficient to define its conformation.

The structure revealed the same DNA distortion that we observed in cCC<sup>dist</sup>, and we therefore refer to it as CC<sup>dist</sup> (Figure 5). Despite the different promoter sequence, the DNA is again underwound in the region spanning from +10 to +20 downstream of TATA, and we again observe an offset in the helical axis around +20 downstream of TATA. Because these DNA distortions were not altered in the presence of TFIH,



**Figure 3. Cryo-EM Structures of Core Closed Complexes**

(A) Structure of Pol II core closed complex cCC1. Proteins are shown in cartoon representation and transcription factors are colored according to their subunits. Pol II is shown in gray, the non-template strand in light cyan and the template strand in blue. (B) Structure of Pol II core closed complex (cCC) cCC<sup>dist</sup>. Coloring as in (A).

(C) The closed clamp in cCC<sup>dist</sup> would clash with undistorted, straight B-DNA. The Pol II clamp and TFIIE are shown as transparent surface with cartoon inside. Other protein components were omitted for clarity. Straight B-DNA was extended from the yeast cOC structure (PDB: 5fyw). The DNA helical axis of the OC DNA used as start for B-DNA modeling is shown as blue cylinder.

(D) The Tfg2 WH and TFIIE lock the DNA at the site of DNA distortion. The Pol II clamp, TFIIE, and Tfg2 WH are shown in transparent surface representation with cartoon inside. Other proteins have been omitted for clarity. DNA helical axis is indicated as blue cylinders.

they are apparently induced and stabilized by components of the cPIC, in particular the Pol II clamp, TFIIE, and TFIIF (Figure 3D). Closed promoter DNA has been observed in a similar location within the yeast PIC before (Murakami et al., 2015), but at the available resolution the DNA distortion was not detected at that time.

In CC<sup>dist</sup> we observe an additional  $\sim 20^\circ$  bend of the DNA helical axis around position +33 downstream of TATA (Figure 5C) that is not present in the cCC<sup>dist</sup> and thus apparently induced by the presence of TFIIF. The region where this additional bend occurs corresponds to the nearest downstream site where transcription initiation may occur, because 30–35 nucleotides of DNA are required to reach from the upstream end of TATA to the active site of Pol II in an OC (Kostrewa et al., 2009). The DNA bend at +33 from TATA and the DNA axis offset at +20 from TATA delimit a DNA region of  $\sim 13$  bp that corresponds to the IMR in the OC (Figure 5C). It is known that the DNA duplex is weakened by DNA underwinding (Kannan et al., 2006) and DNA bending (Ramstein and Lavery, 1988). We thus conclude that the IMR is structurally pre-defined by two flanking DNA distortions in the complete PIC CC<sup>dist</sup> intermediate.

#### Promoters of Pol I and Pol III Contain Unstable IMRs

The lack of Ssl2 dependence at a subset of Pol II promoters resembles the situation during transcription initiation by Pol I

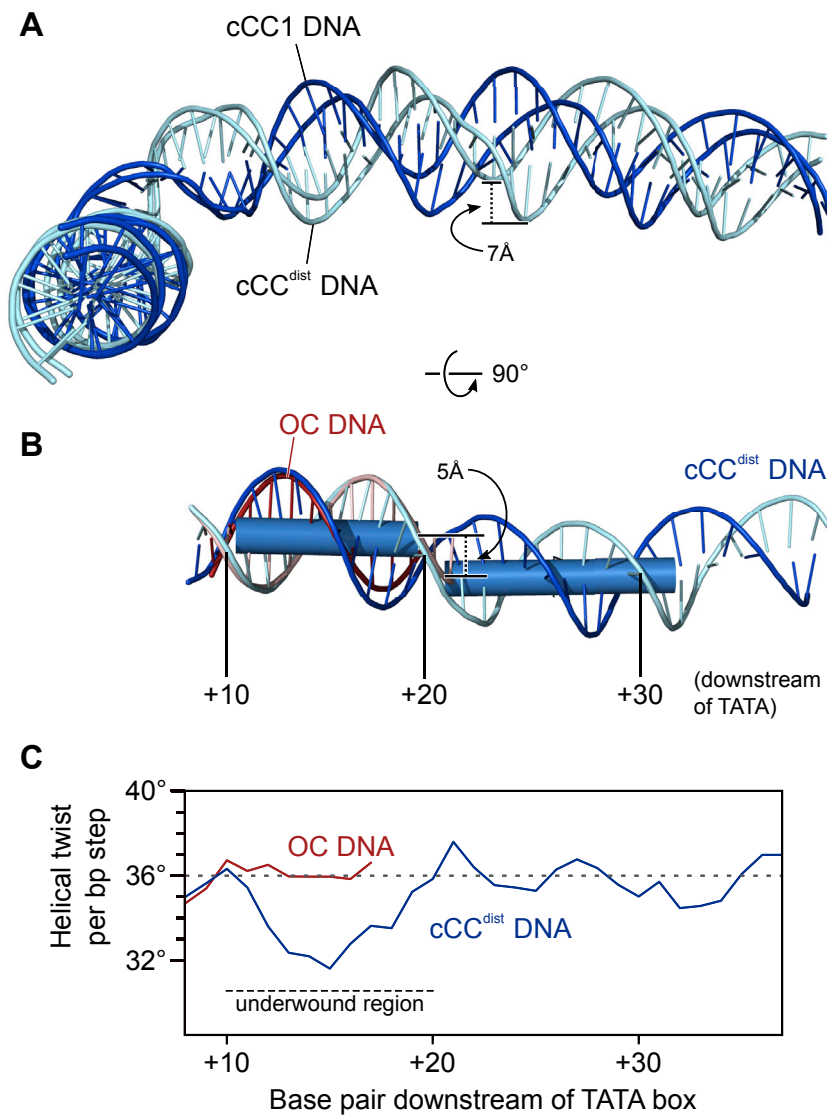
and Pol III, which do not require a translocase. Whereas Pol I uses unique initiation factors, the Pol III initiation machinery is closely related to that of Pol II (Vannini and Cramer, 2012). To investigate whether Pol III genes may also show enhanced meltability in their IMR, we calculated the average DNA duplex free energy (SantaLucia, 1998) of all yeast tRNA genes aligned at their TSS. Indeed, the IMR upstream of the TSS shows a weaker DNA duplex than the region downstream of the TSS (Figure 6). This indicates that Pol III promoters have a less stable IMR, explaining why they are prone to opening easily.

Pol I promoters were also suggested to be prone to DNA melting in their IMR (Engel et al., 2017). We therefore calculated the average DNA duplex free energy of Pol I promoters from eight different species (Moss et al., 2007) aligned at their TSS. Stability of these promoters was significantly lower in the region upstream and around the TSS (Figure 6). Thus, IMRs in Pol I promoters also show lower stability and thus greater meltability, similar to Pol III promoters and Pol II promoters that do not respond to Ssl2 depletion. These findings suggest that the strategy for DNA opening by Pol II shows similarities to the DNA opening mechanisms employed by the Pol I and Pol III machineries.

## DISCUSSION

### Closed Promoter Complex with Distorted DNA May Represent an Initiation Intermediate

Here we describe the structures of Pol II transcription initiation complexes CC1 and CC<sup>dist</sup> that both contain closed promoter DNA but in different conformations (Video S1). Whereas CC1 resembles known structures of the yeast CC (Plaschka et al.,



**Figure 4. Distortion of Promoter DNA**

(A) Promoter DNA is loaded 7 Å deeper into the Pol II cleft in the cCC<sup>dist</sup> structure compared with cCC1. The template and non-template strands are in blue and cyan, respectively.

(B) Distortion of promoter DNA in cCC<sup>dist</sup>. The helical DNA axis was calculated by w3DNA (Zheng et al., 2009) and is indicated by blue cylinders. OC DNA (PDB: 5fyw) is shown in red. OC and cCC<sup>dist</sup> were superimposed with their TATA boxes. The sequence register relative to the upstream end of the TATA box is indicated for DNA in cCC<sup>dist</sup>.

(C) The DNA region around the upstream edge of the IMR is underwound in cCC<sup>dist</sup>. The DNA helical twist per base pair step as calculated with w3DNA (Zheng et al., 2009) is shown for OC (red) and cCC<sup>dist</sup> (cyan/blue) structures. The average helical twist for canonical B-DNA (36°) is indicated with a dotted gray line. Values were calculated as average on a running window of three base pairs.

may now occur. We show that the degree of dependence on TFIIH for DNA opening varies for different promoters, but since some Ssl2 remains in the depleted nuclei, we cannot entirely exclude the possibility that low levels of Ssl2 still support transcription *in vivo*. When the IMR is stable, which is usually the case for Pol II promoters, TFIIH is required and uses its ATP-dependent translocase activity in subunit Ssl2 (human XPB) to further underwind and open DNA, leading to the OC. When the IMR is unstable, DNA distortion and underwinding alone are sufficient to promote DNA opening, explaining why some genes may be transcribed without TFIIH translocase activity.

In both cases, the DNA template strand gets loaded into the cleft and DNA interacts more extensively with Pol II, and this stabilizes the OC and prevents re-annealing of the DNA strands.

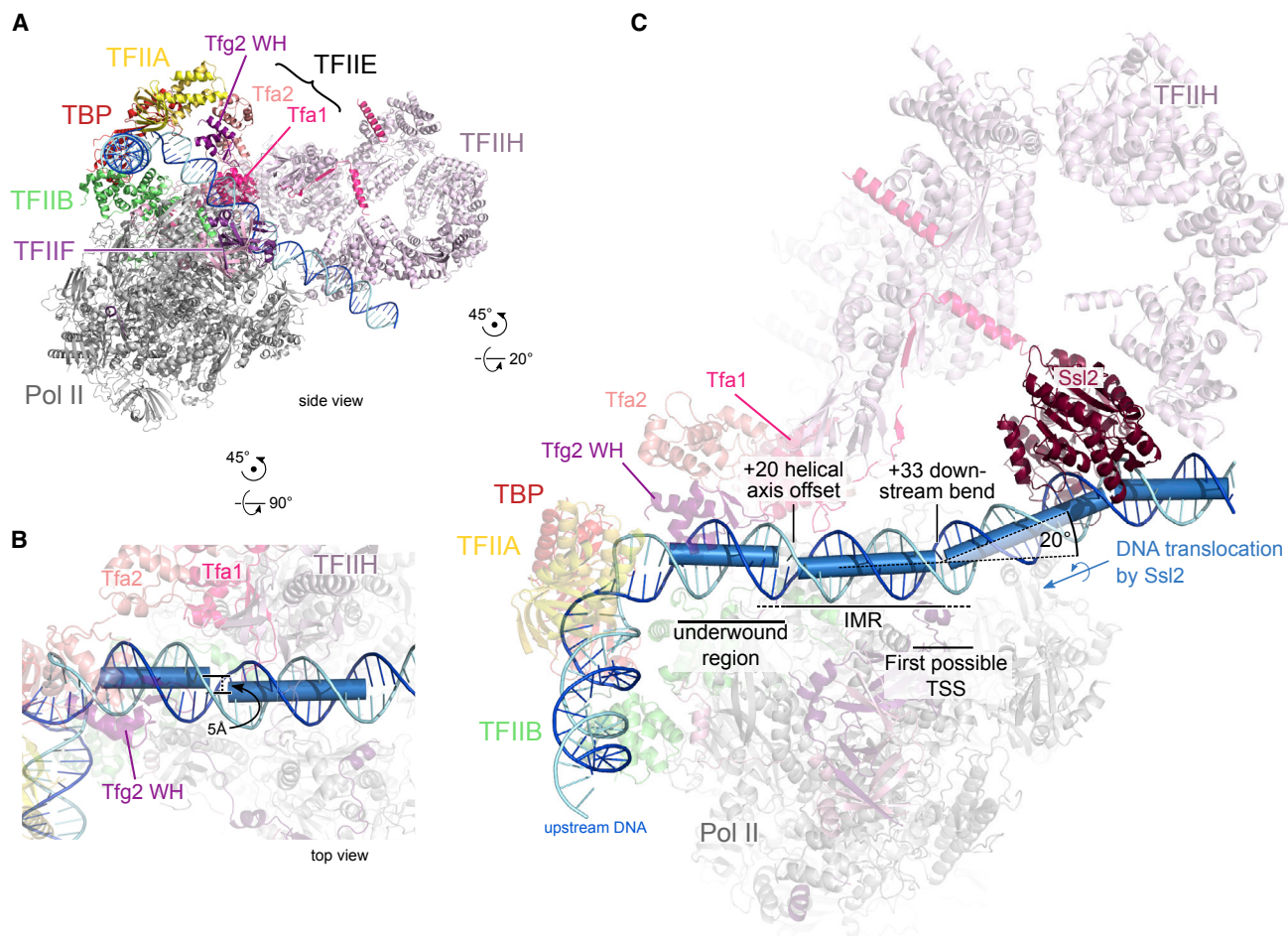
### DNA Distortion and Initial Opening

Our results suggest that promoter opening may spontaneously nucleate at the upstream edge of the IMR in the underwound DNA region in cCC<sup>dist</sup>. Supporting this, TFIIH activity is required only for opening of the downstream part of the IMR (Holstege et al., 1996). The cCC<sup>dist</sup> structure also indicates why DNA opening propagates downstream, but not upstream, from the initial point of opening. Interactions of the upstream DNA duplex with the Tfg2 WH and TFIIIE eWH domains are formed in cCC<sup>dist</sup> and may prevent propagation of the initial DNA bubble upstream toward the TATA box. However, the strain introduced by TFIIH translocase activity can be released when the bubble is extended downstream toward the TSS.

2016), CC<sup>dist</sup> represents a new conformation and was observed both in the absence and in the presence of TFIIH. CC<sup>dist</sup> contains distorted promoter DNA in a closed Pol II cleft. This structure likely represents an intermediate of transcription initiation that has not been observed before and that reflects the state of the PIC just before DNA opening and conversion to the OC. In contrast, the previously reported human CC structures (He et al., 2013, 2016) show an open clamp and contain canonical B-form DNA, indicating that they represent a CC state that is readily interchangeable with CC1, and therefore it may be referred to as CC2.

Together with the obtained functional data, these structures suggest an extended model for promoter opening and transcription initiation (Figure 7). In the PIC, promoter DNA is first positioned above the Pol II cleft (CC1). Cleft opening then allows for DNA swinging into the cleft (CC2). Clamp closure is coupled to DNA distortion and underwinding in the cleft (CC<sup>dist</sup>). In CC<sup>dist</sup>, the IMR is pre-defined, and DNA opening





**Figure 5. PIC Structure with Closed, Distorted DNA**

(A) Structure of Pol II PIC including TFIH in the CC<sup>dist</sup> state. The structure is shown in side view and colored according to Figure 3A with TFIH in pale pink. (B) DNA helical axis offset. Close-up view of CC<sup>dist</sup> promoter DNA with the DNA helical axis indicated as blue cylinders. Proteins around the DNA are shown in transparent cartoon. (C) The presence of TFIH induces a 20° bend at the downstream edge of the IMR.

The energetic cost for distorting DNA in CC<sup>dist</sup> is apparently compensated by a gain of binding energy; otherwise, we could not have trapped CC<sup>dist</sup> *in vitro*. Conversion of CC1 to CC<sup>dist</sup> may yield additional binding energy because the TFIIE eWH domain and the Tfg2 WH domain form contacts above the cleft. In addition, the B-linker element in TFIIB likely stabilizes CC<sup>dist</sup>, because we observe good density for it (Figure S4). The B-linker may help to keep the clamp in a closed state. The B-linker is also known to be involved in promoter opening in the homologous archaeal system, and its mutation induces a growth phenotype in yeast (Kostrewa et al., 2009). In contrast, CC2 (He et al., 2013, 2016) shows a mobile B-linker and undistorted DNA that would clash with a modeled B-linker. Thus, a transition from CC2 to CC<sup>dist</sup> involves clamp closure, DNA distortion, and B-linker ordering.

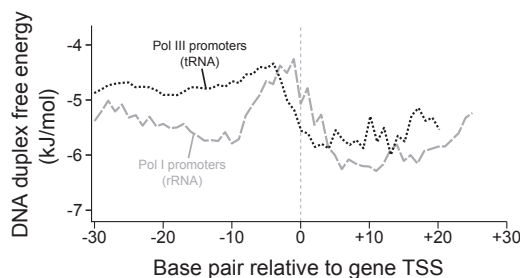
### General Features of Transcription Initiation

Our results also provide insights into the evolution of the three eukaryotic transcription machineries and help explain why the

two other nuclear RNA polymerases, Pol I and Pol III, require neither energy from ATP hydrolysis nor a TFIH-like factor for DNA opening. Recent studies on Pol III transcription initiation complexes revealed a very similar path of promoter DNA and equivalent structural elements for most of the basal Pol II TFs (Abascal-Palacios et al., 2018; Vorländer et al., 2018). Two CCs were resolved with promoter DNA along the closed cleft of Pol III (Vorländer et al., 2018) that correspond to the Pol II cCC1. It was proposed that opening of the Pol III clamp allows promoter DNA to enter the cleft and that clamp closure leads to DNA opening (Vorländer et al., 2018), also consistent with the Pol II DNA opening mechanism.

For Pol I, structural studies of initiation complexes showed that promoter DNA is positioned further inside the cleft compared to the Pol II and Pol III systems (Engel et al., 2017; Han et al., 2017; Sadian et al., 2017). Promoter DNA was observed running into the cleft directly over the Pol I wall, rather than being held well above the wall as in the Pol II and Pol III systems. A Pol I CC





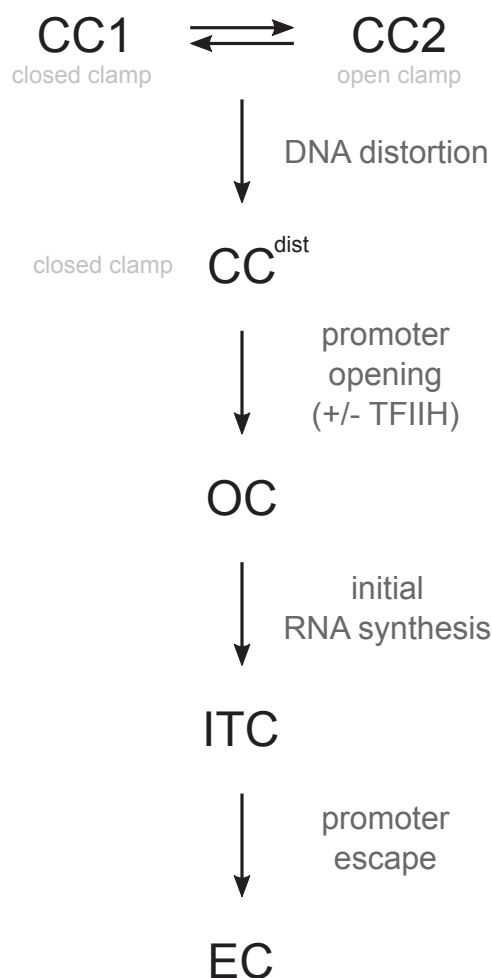
**Figure 6. IMRs of Pol I and Pol III Promoters Are Unstable**

DNA duplex free energy calculated based on nearest-neighbor thermodynamics (SantaLucia, 1998) for Pol I (dashed) and Pol III (dotted) promoters. Sequences were aligned at the gene TSS. Free energy was calculated for the first base pair of the central dinucleotide of an eight base pairs running window for more stable estimates.

was not resolved, but modeling suggested that DNA loading into the cleft and subsequent promoter opening require clamp opening and closing, respectively (Engel et al., 2017), strongly suggesting a similar overall mechanism to what we describe here for Pol II. Thus, despite the differences between the three eukaryotic nuclear transcription machineries, features of all three systems include DNA loading, distortion and opening in the cleft, and the necessity of opening and closing of the clamp.

Whereas Pol I and Pol III promoters apparently open easily, most Pol II promoters have evolved to contain stable IMRs. This may have rendered them TFIIF dependent and more regulatable during evolution. In addition, DNA opening may be facilitated in the Pol I and Pol III systems because DNA is more extensively bound by the PICs of Pol I and Pol III and is loaded more deeply into the cleft when compared to Pol II (Table S1). This may allow for tighter DNA binding and more severe DNA distortions during the conversion to a  $CC^{dist}$  intermediate. As a result, the efficiency of the transition from  $CC^{dist}$  to OC would be higher and DNA opening would be easier in the Pol I and Pol III systems. A lower stability of  $CC^{dist}$  may explain why this structure could thus far not be trapped for Pol I and Pol III.

The archaeal initiation system is highly homologous to a minimal Pol II system, and changes of the clamp state have been observed during promoter opening (Schulz et al., 2016). The initiation mechanism differs for bacterial RNA polymerase. In bacterial CC structures, promoter DNA resides further above the cleft and is contacted only by the accessory sigma factors (Glyde et al., 2017; Murakami et al., 2002). Promoter opening is initiated by the factor sigma-70 in the upper region of the active center cleft by trapping of bases that are flipped-out (Feklistov and Darst, 2011). Structural studies of bacterial RNAP with sigma-54 revealed that DNA opening and cleft loading are coupled and involve DNA distortion (Glyde et al., 2018). DNA bubble extension and DNA loading into the bacterial RNAP active center cleft are known to require clamp opening and closing (Chakraborty et al., 2012; Feklistov and Darst, 2011). Based on these conceptual similarities with our observations in the eukaryotic system, we suggest that the coupling of clamp closure to DNA distortion in the cleft is a universal feature of



**Figure 7. General Model of Transcription Initiation**

Proposed interconversion of initiation complex intermediates during DNA opening and the initiation-elongation transition. Compare text for details.

transcription initiation that facilitates DNA opening by multi-subunit RNA polymerases, although promoter-specific variations likely occur.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Ssl2 nuclear depletion and 4sU-sequencing
  - Nuclei isolation and western blotting
  - Bioinformatics data analysis
  - DNA opening assay
  - Preparation of cPICs
  - Cryo-EM and image processing of cCC1 and cCC<sup>dist</sup>

- Data processing and CC<sup>dist</sup> reconstruction
- Model building and refinement
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY

## SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures, one table, one video, and a wwPDB validation report and can be found with this article online at <https://doi.org/10.1016/j.molcel.2018.10.014>.

## ACKNOWLEDGMENTS

We thank C. Bernecky, D. Tegunov, S. Neyer, K. Maier, P. Rus, C. Burzinski, C. Wigge, and other members of the Cramer laboratory for help. P.C. was supported by ERC Advanced Grant TRANSREGULON (grant agreement No 693023) of the European Research Council, the Deutsche Forschungsgemeinschaft (SFB 860), and the Volkswagen Foundation.

## AUTHOR CONTRIBUTIONS

C.D. carried out all experiments and data analysis except for the following. B.S. carried out bioinformatics analysis of Ssl2 anchor-away 4sU-sequencing data. B.S. and C.D. interpreted Ssl2 anchor-away data. S.S. carried out initial processing of the CC<sup>dist</sup> data set. P.C. designed and supervised research. C.D. and P.C. wrote the manuscript.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 24, 2018

Revised: July 4, 2018

Accepted: October 9, 2018

Published: November 21, 2018

## REFERENCES

- Abascal-Palacios, G., Ramsay, E.P., Beuron, F., Morris, E., and Vannini, A. (2018). Structural basis of RNA polymerase III transcription initiation. *Nature* 553, 301–306.
- Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* 66, 213–221.
- Alekseev, S., Nagy, Z., Sandoz, J., Weiss, A., Egly, J.M., Le May, N., and Coin, F. (2017). Transcription without XPB establishes a unified helicase-independent mechanism of promoter opening in eukaryotic gene expression. *Mol. Cell* 65, 504–514.
- Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biol.* 11, R106.
- Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166–169.
- Baejen, C., Andreani, J., Torkler, P., Battaglia, S., Schwalb, B., Lidschreiber, M., Maier, K.C., Boltendahl, A., Rus, P., Esslinger, S., et al. (2017). Genome-wide analysis of RNA polymerase II termination at protein-coding genes. *Mol. Cell* 66, 38–49.
- Chakraborty, A., Wang, D., Ebright, Y.W., Korlann, Y., Kortkhonja, E., Kim, T., Chowdhury, S., Wigneshweraraj, S., Irschik, H., Jansen, R., et al. (2012). Opening and closing of the bacterial RNA polymerase clamp. *Science* 337, 591–595.
- Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21.
- Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* 60, 2126–2132.
- Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66, 486–501, <https://doi.org/10.1107/S0907444910007493>.
- Engel, C., Gubbey, T., Neyer, S., Sainsbury, S., Oberthuer, C., Baejen, C., Bernecky, C., and Cramer, P. (2017). Structural basis of RNA polymerase I transcription initiation. *Cell* 169, 120–131.
- Feklistov, A., and Darst, S.A. (2011). Structural basis for promoter-10 element recognition by the bacterial RNA polymerase  $\sigma$  subunit. *Cell* 147, 1257–1269.
- Fishburn, J., Tomko, E., Galburt, E., and Hahn, S. (2015). Double-stranded DNA translocase activity of transcription factor TFIH and the mechanism of RNA polymerase II open complex formation. *Proc. Natl. Acad. Sci. USA* 112, 3961–3966.
- Giardina, C., and Lis, J.T. (1993). DNA melting on yeast RNA polymerase II promoters. *Science* 261, 759–762.
- Glyde, R., Ye, F., Darbari, V.C., Zhang, N., Buck, M., and Zhang, X. (2017). Structures of RNA polymerase closed and intermediate complexes reveal mechanisms of DNA opening and transcription initiation. *Mol. Cell* 67, 106–116.
- Glyde, R., Ye, F., Jovanovic, M., Kotta-Loizou, I., Buck, M., and Zhang, X. (2018). Structures of bacterial RNA polymerase complexes reveal the mechanism of DNA loading and transcription initiation. *Mol. Cell* 70, 1111–1120.
- Goodrich, J.A., and Tjian, R. (1994). Transcription factors IIE and IIH and ATP hydrolysis direct promoter clearance by RNA polymerase II. *Cell* 77, 145–156.
- Grünberg, S., and Hahn, S. (2013). Structural insights into transcription initiation by RNA polymerase II. *Trends Biochem. Sci.* 38, 603–611.
- Han, Y., Yan, C., Nguyen, T.H.D., Jackobel, A.J., Ivanov, I., Knutson, B.A., and He, Y. (2017). Structural mechanism of ATP-independent transcription initiation by RNA polymerase I. *eLife* 6, e27414.
- Haruki, H., Nishikawa, J., and Laemmli, U.K. (2008). The anchor-away technique: rapid, conditional establishment of yeast mutant phenotypes. *Mol. Cell* 31, 925–932.
- He, Y., Fang, J., Taatjes, D.J., and Nogales, E. (2013). Structural visualization of key steps in human transcription initiation. *Nature* 495, 481–486.
- He, Y., Yan, C., Fang, J., Inouye, C., Tjian, R., Ivanov, I., and Nogales, E. (2016). Near-atomic resolution visualization of human transcription promoter opening. *Nature* 533, 359–365.
- Holstege, F.C., Fiedler, U., and Timmers, H.T. (1997). Three transitions in the RNA polymerase II transcription complex during initiation. *EMBO J.* 16, 7468–7480.
- Holstege, F.C., Tantin, D., Carey, M., van der Vliet, P.C., and Timmers, H.T. (1995). The requirement for the basal transcription factor IIE is determined by the helical stability of promoter DNA. *EMBO J.* 14, 810–819.
- Holstege, F.C., van der Vliet, P.C., and Timmers, H.T. (1996). Opening of an RNA polymerase II promoter occurs in two distinct steps and requires the basal transcription factors IIE and IIH. *EMBO J.* 15, 1666–1677.
- Kannan, S., Kohlhoff, K., and Zacharias, M. (2006). B-DNA under stress: over- and untwisting of DNA during molecular dynamics simulations. *Biophys. J.* 91, 2956–2965.
- Kashkina, E., Anikin, M., Brueckner, F., Lehmann, E., Kochetkov, S.N., McAllister, W.T., Cramer, P., and Temiakov, D. (2007). Multisubunit RNA polymerases melt only a single DNA base pair downstream of the active site. *J. Biol. Chem.* 282, 21578–21582.
- Kastner, B., Fischer, N., Golas, M.M., Sander, B., Dube, P., Boehringer, D., Hartmuth, K., Deckert, J., Hauer, F., Wolf, E., et al. (2008). GraFix: sample preparation for single-particle electron cryomicroscopy. *Nat. Methods* 5, 53–55.
- Kostrewa, D., Zeller, M.E., Armache, K.J., Seizl, M., Leike, K., Thomm, M., and Cramer, P. (2009). RNA polymerase II-TFIIB structure and mechanism of transcription initiation. *Nature* 462, 323–330.

- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079.
- Li, X., Mooney, P., Zheng, S., Booth, C.R., Braunfeld, M.B., Gubbens, S., Agard, D.A., and Cheng, Y. (2013). Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat. Methods* 10, 584–590.
- Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550.
- Martinez-Rucobo, F.W., and Cramer, P. (2013). Structural basis of transcription elongation. *Biochim. Biophys. Acta* 1829, 9–19.
- Moss, T., Langlois, F., Gagnon-Kugler, T., and Stefanovsky, V. (2007). A housekeeper with power of attorney: the rRNA genes in ribosome biogenesis. *Cell. Mol. Life Sci.* 64, 29–49.
- Murakami, K., Tsai, K.L., Kalisman, N., Bushnell, D.A., Asturias, F.J., and Kornberg, R.D. (2015). Structure of an RNA polymerase II preinitiation complex. *Proc. Natl. Acad. Sci. USA* 112, 13543–13548.
- Murakami, K.S., Masuda, S., Campbell, E.A., Muzzin, O., and Darst, S.A. (2002). Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science* 296, 1285–1290.
- Nogales, E., Louder, R.K., and He, Y. (2017). Structural insights into the eukaryotic transcription initiation machinery. *Annu. Rev. Biophys.* 46, 59–83.
- Pan, G., and Greenblatt, J. (1994). Initiation of transcription by RNA polymerase II is limited by melting of the promoter DNA in the region immediately upstream of the initiation site. *J. Biol. Chem.* 269, 30101–30104.
- Parvin, J.D., and Sharp, P.A. (1993). DNA topology and a minimal set of basal factors for transcription by RNA polymerase II. *Cell* 73, 533–540.
- Plaschka, C., Hantsche, M., Dienemann, C., Burzinski, C., Plitzko, J., and Cramer, P. (2016). Transcription initiation complex structures elucidate DNA opening. *Nature* 533, 353–358.
- Plaschka, C., Larivière, L., Wenzel, L., Seizl, M., Hemann, M., Tegunov, D., Petrotchenko, E.V., Borchers, C.H., Baumeister, W., Herzog, F., et al. (2015). Architecture of the RNA polymerase II-Mediator core initiation complex. *Nature* 518, 376–380.
- Ramstein, J., and Lavery, R. (1988). Energetic coupling between DNA bending and base pair opening. *Proc. Natl. Acad. Sci. USA* 85, 7231–7235.
- Rhee, H.S., and Pugh, B.F. (2012). Genome-wide structure and organization of eukaryotic pre-initiation complexes. *Nature* 483, 295–301.
- Rohou, A., and Grigorieff, N. (2015). CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* 192, 216–221.
- Sadian, Y., Tafur, L., Kosinski, J., Jakobi, A.J., Wetzel, R., Buczak, K., Hagen, W.J., Beck, M., Sachse, C., and Müller, C.W. (2017). Structural insights into transcription initiation by yeast RNA polymerase I. *EMBO J.* 36, 2698–2709.
- Sainsbury, S., Bernecky, C., and Cramer, P. (2015). Structural basis of transcription initiation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* 16, 129–143.
- Sainsbury, S., Niesser, J., and Cramer, P. (2013). Structure and function of the initially transcribing RNA polymerase II-TFIIB complex. *Nature* 493, 437–440.
- SantaLucia, J., Jr. (1998). A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl. Acad. Sci. USA* 95, 1460–1465.
- Scheres, S.H. (2012). A Bayesian view on cryo-EM structure determination. *J. Mol. Biol.* 415, 406–418.
- Schilbach, S., Hantsche, M., Tegunov, D., Dienemann, C., Wigge, C., Urlaub, H., and Cramer, P. (2017). Structures of transcription pre-initiation complex with TFIH and Mediator. *Nature* 551, 204–209.
- Schulz, D., Schwalb, B., Kiesel, A., Baejen, C., Torkler, P., Gagneur, J., Soeding, J., and Cramer, P. (2013). Transcriptome surveillance by selective termination of noncoding RNA synthesis. *Cell* 155, 1075–1087.
- Schulz, S., Gietl, A., Smollett, K., Tinnefeld, P., Werner, F., and Grohmann, D. (2016). TFE and Spt4/5 open and close the RNA polymerase clamp during the transcription cycle. *Proc. Natl. Acad. Sci. USA* 113, E1816–E1825.
- Schwalb, B., Michel, M., Zacher, B., Frühauf, K., Demel, C., Tresch, A., Gagneur, J., and Cramer, P. (2016). TT-seq maps the human transient transcriptome. *Science* 352, 1225–1228.
- Spitalny, P., and Thomm, M. (2003). Analysis of the open region and of DNA-protein contacts of archaeal RNA polymerase transcription complexes during transition from initiation to elongation. *J. Biol. Chem.* 278, 30497–30505.
- Sydow, J.F., Brueckner, F., Cheung, A.C., Damsma, G.E., Dengli, S., Lehmann, E., Vassilyev, D., and Cramer, P. (2009). Structural basis of transcription: mismatch-specific fidelity mechanisms and paused RNA polymerase II with frayed RNA. *Mol. Cell* 34, 710–721.
- Tyree, C.M., George, C.P., Lira-DeVito, L.M., Wampler, S.L., Dahmus, M.E., Zawel, L., and Kadonaga, J.T. (1993). Identification of a minimal set of proteins that is sufficient for accurate initiation of transcription by RNA polymerase II. *Genes Dev.* 7 (7A), 1254–1265.
- van Dijk, M., and Bonvin, A.M. (2009). 3D-DART: a DNA structure modelling server. *Nucleic Acids Res.* 37, W235–W239.
- Vannini, A., and Cramer, P. (2012). Conservation between the RNA polymerase I, II, and III transcription initiation machineries. *Mol. Cell* 45, 439–446.
- Vorländer, M.K., Khatler, H., Wetzel, R., Hagen, W.J.H., and Müller, C.W. (2018). Molecular mechanism of promoter opening by RNA polymerase III. *Nature* 553, 295–300.
- Warfield, L., Luo, J., Ranish, J., and Hahn, S. (2016). Function of conserved topological regions within the *Saccharomyces cerevisiae* basal transcription factor TFIH. *Mol. Cell. Biol.* 36, 2464–2475.
- Zheng, G., Lu, X.J., and Olson, W.K. (2009). Web 3DNA—a web server for the analysis, reconstruction, and visualization of three-dimensional nucleic-acid structures. *Nucleic Acids Res.* 37, W240–W246.

## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Antibodies</b>		
Rabbit polyclonal anti-Ssl2	<a href="#">Warfield et al., 2016</a>	3079
H3-pan monoclonal antibody HRP conjugate	Abcam	ab21054, RRID: AB_880437
<b>Bacterial and Virus Strains</b>		
<i>E.coli</i> BL21 CodonPlus (DE3) RIL	Agilent	Cat#230245
<b>Chemicals, Peptides, and Recombinant Proteins</b>		
<i>S. cerevisiae</i> RNA polymerase II	<a href="#">Sydow et al., 2009</a>	N/A
<i>S. cerevisiae</i> TBP	<a href="#">Plaschka et al., 2016</a>	N/A
<i>S. cerevisiae</i> TFIID95	<a href="#">Plaschka et al., 2016</a>	N/A
<i>S. cerevisiae</i> TFIIB	<a href="#">Sainsbury et al., 2013</a>	N/A
<i>S. cerevisiae</i> TFIIF	<a href="#">Plaschka et al., 2016</a>	N/A
<i>S. cerevisiae</i> TFIIE	<a href="#">Plaschka et al., 2016</a>	N/A
<i>S. cerevisiae</i> core Mediator	<a href="#">Plaschka et al., 2016</a>	N/A
Glutaraldehyde 25%	EMS	Cat#16200
4-thiouracil	Sigma	Cat#440736
Rapamycin	Enzo	Cat#BML-A275
<b>Critical Commercial Assays</b>		
Yeast Nuclei extraction kit	BioVision	K289
<b>Deposited Data</b>		
Yeast core CC1 structure	This work	PDB: 6GYK
Yeast core CC <sup>dist</sup> structure	This work	PDB: 6GYL
Yeast CC <sup>dist</sup> structure	This work	PDB: 6GYM
coreCC1 map	This work	EMD: 0090
coreCC <sup>dist</sup> map	This work	EMD: 0091
CC <sup>dist</sup> map	This work	EMD:0092
4tU-sequencing data with and without Ssl2 depletion	This work	GEO: GSE111777
<b>Experimental Models: Organisms/Strains</b>		
<i>S. cerevisiae</i> Y40343	<a href="#">Haruki et al., 2008</a>	Euroscarf Y40343
<i>S. cerevisiae</i> Y40343 Ssl2-FRB	This work	N/A
<b>Oligonucleotides</b>		
HIS4-nt 5'-GCA CGC TGT GTA TAT AAT AGC TAT GGA ACG TTC GAT TC(2-AP) CCT CCG ATG TGT GTT GTA CAT ACA TAA AAA TAT CA-3'	This work	IDT
HIS4-t 5'-TGA TAT TTT TAT GTA TGT ACA ACA CAC ATC GGA GGT GAA TCG AAC GTT CCA TAG CTA TTA TAT ACA CAG CGT GC-3'	This work	IDT
GAT1-nt 5'- CCC AGC CAC ATA TAT ATA GGT GTG TGC CAC TCC CGG CC(2-AP) CGG TAT TAG CAT GCA CGT TTT CTT TCC TTT GCT TTT-3'	This work	IDT
GAT1-t 5'- AAA AGC AAA GGA AAG AAA ACG TGC ATG CTA ATA CCG TGG CCG GGA GTG GCA CAC ACC TAT ATA TAT GTG GCT GGG-3'	This work	IDT
HIS4-mm-t 5'-TGA TAT TTT TAT GTA TGT ACA ACA CAC ATC GCT TAG AGG CTC TGC GTT CCA TAG CTA TTA TAT ACA CAG CGT GC-3'	This work	IDT
HIS4-nt+17 5'-GCA CGC TGT GTA TAT AAT AGC TAT GGA (2-AP)CG TTC GAT TCA CCT CCG ATG TGT GTT GTA CAT ACA TAA AAA TAT CA-3'	This work	IDT
HIS4-nt+31 5'-GCA CGC TGT GTA TAT AAT AGC TAT GGA ACG TTC GAT TCA CC(2-AP) CCG ATG TGT GTT GTA CAT ACA TAA AAA TAT CA-3'	This work	IDT

(Continued on next page)



## Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Recombinant DNA		
pOPINE-TBP	<a href="#">Plaschka et al., 2016</a>	N/A
pOPINE-TFIAd95	<a href="#">Plaschka et al., 2016</a>	N/A
pAHS3C-TFIIF	<a href="#">Plaschka et al., 2016</a>	N/A
pOPINE-TFIIB	<a href="#">Sainsbury et al., 2013</a>	N/A
pET21-TFIIE	<a href="#">Plaschka et al., 2016</a>	N/A
pETduet-Med17-Med22/11-Med6-Med8/20/18	<a href="#">Plaschka et al., 2016</a>	N/A
pET28b-Med19-Med14	<a href="#">Plaschka et al., 2016</a>	N/A
pCDFDuet-Med31-Med10-Med7-Med21	<a href="#">Plaschka et al., 2016</a>	N/A
pCP99-Med1/4/9	<a href="#">Plaschka et al., 2016</a>	N/A
Software and Algorithms		
Relion 1.4	<a href="#">Scheres et al., 2012</a>	RRID: SCR_016274
PHENIX suite	<a href="#">Adams et al., 2010</a>	RRID: SCR_014224
COOT v0.8.3	<a href="#">Emsley et al., 2010</a>	RRID: SCR_014222
Other		
Titan Krios G2	FEI/Thermo Fischer	N/A
BioQuantum Imaging Filter	Gatan	N/A
K2 Summit Direct Electron Detector	Gatan	N/A

## CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact Patrick Cramer ([patrick.cramer@mpibpc.mpg.de](mailto:patrick.cramer@mpibpc.mpg.de)).

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

For recombinant expression the *Escherichia coli* strain BL21CodonPlus (DE3) RIL (Agilent) (Merck Millipore) was used. Cells were grown at 37°C as described in the STAR Methods Details section. *In vivo* yeast experiments were done using an Ssl2-anchor away strain based on Y40343 ([Haruki et al., 2008](#)).

## METHOD DETAILS

### Ssl2 nuclear depletion and 4sU-sequencing

The Ssl2 anchor away yeast strain was created from the rapamycin-resistant strain Y40343 ([Haruki et al., 2008](#)) by homologous recombination after PCR-amplifying a Ssl2-FRB-KanMX6 fragment with gene-specific primers. Clones were selected on G418 plates and confirmed by colony-PCR and sequencing. The Ssl2 anchor away strain shows a strong growth phenotype under depletion conditions ([Figure S1F](#)). The anchor away was additionally validated by anti-FRB chromatin immunoprecipitation (ChIP) and qPCR with primers against specific gene promoter regions ([Baejen et al., 2017](#)). Before metabolic 4sU-sequencing (4sU-seq), Ssl2 depletion was induced at OD 0.6 by the addition of 1 µg/ml rapamycin for 60 minutes. Both anchor away and 4sU-seq, RNA extraction and library preparation were performed as described ([Schulz et al., 2013](#)).

### Nuclei isolation and western blotting

Yeast cells were grown to OD 0.6 under the same conditions as for 4sU sequencing, treated with 1 µg/ml rapamycin or DMSO as control and harvested after 60 minutes. Nuclei were purified from rapamycin treated and untreated samples using a Yeast Nuclei Isolation kit (BioVision). Purified nuclei were run on SDS-PAGE and blotted onto a PVDF-membrane. Ssl2 was detected using a polyclonal rabbit antibody against yeast Ssl2 ([Warfield et al., 2016](#)). H3 was used as loading control and detected by HRP-conjugated by an H3pan antibody (Abcam).

### Bioinformatics data analysis

Data analysis was performed as described ([Schulz et al., 2013](#)), with modifications. Briefly, paired-end 50 bp reads with additional 6 bp of barcodes were obtained for labeled RNA. Reads were demultiplexed and aligned to the *S. cerevisiae* genome (sacCer3, version 64.2.1) using STAR (version 2.3.0) ([Dobin et al., 2013](#)). SAMTools was used to quality filter SAM files ([Li et al., 2009](#)).

Alignments with MAPQ smaller than 7 ( $-q \geq 7$ ) were skipped and only proper pairs ( $-f2$ ) were selected. Further processing of the 4sU-seq data was carried out using R/Bioconductor. We used a spike-in (RNAs) normalization strategy essentially as described (Schwalb et al., 2016) to allow observation of global shifts and antisense bias determination (ratio of spurious reads originating from the opposite strand introduced by the RT reactions). Read counts for all features were calculated using HTSeq (Anders et al., 2015) and corrected for antisense bias using antisense bias ratios calculated as described (Schwalb et al., 2016). Gene expression fold changes upon rapamycin treatment were calculated using DESeq2 (Love et al., 2014). Differentially expressed genes were identified using a fold change of at least 1.5 and an adjusted P value of maximal 0.1. Gene-wise DNA duplex free energies were calculated over a window of 8 nucleotides based on nearest-neighbor thermodynamics (SantaLucia, 1998).

### DNA opening assay

*S. cerevisiae* Pol II, TFIIB, TFIIF, TFIIE were purified as described (Plaschka et al., 2016; Schilbach et al., 2017). 2-amino purine (2-AP)-labeled promoter DNAs were synthesized (IDT) and reannealed on a slow temperature gradient from 95–20°C. *HIS4* scaffolds were based on the native sequence with A+27 replaced by 2-AP (non-template: 5'-GCA CGC TGT GTA TAT AAT AGC TAT GGA ACG TTC GAT TCA\* CCT CCG ATG TGT GTT GTA CAT ACA TAA AAA TAT CA-3' and template: 5'-TGA TAT TTT TAT GTA TGT ACA ACA CAC ATC GGA GGT GAA TCG AAC GTT CCA TAG CTA TTA TAT ACA CAG CGT GC-3'). Base pair +27 of the native *GAT1* promoter sequence was mutated from C:G to 2-AP:T for label incorporation (non-template: 5'-CCC AGC CAC ATA TAT ATA GGT GTG TGC CAC TCC CGG CCA\* CGG TAT TAG CAT GCA CGT TTT CTT TCC TTT GCT TTT-3' and template: 5'-AAA AGC AAA GGA AAG AAA ACG TGC ATG CTA ATA CCG TGG CCG GGA GTG GCA CAC ACC TAT ATA TAT GTG GCT GGG-3'). The miss-matched DNA control was done using the +27 2-AP labeled *HIS4* non-template strand and a miss-matched template strand (5'-TGA TAT TTT TAT GTA TGT ACA ACA CAC ATC GCT TAG AGG CTC TGC GTT CCA TAG CTA TTA TAT ACA CAG CGT GC-3'). Positional experiments with 2-AP labels at +17 and +31 were done using differently labeled non-template strands with complementary template strands (5'-GCA CGC TGT GTA TAT AAT AGC TAT GGA A\*CG TTC GAT TCA CCT CCG ATG TGT GTT GTA CAT ACA TAA AAA TAT CA-3' and 5'-GCA CGC TGT GTA TAT AAT AGC TAT GGA ACG TTC GAT TCA CCA\* CCG ATG TGT GTT GTA CAT ACA TAA AAA TAT CA-3'). 25 pmol of each component was mixed in 25  $\mu$ L reactions in 10 mM HEPES pH 7, 65 mM KCl, 5% glycerol, 1 mM DTT. Reactions were incubated for 90 minutes or 22 hours at 25°C and 2-AP fluorescence was measured with  $\lambda_{ex}/\lambda_{em}$  of 307/370 nm. Normalization of background fluorescence from protein and nucleic acids was done with unlabelled samples. Measurements were done in triplicates.

### Preparation of cPICs

Proteins were purified as above, and TFIIA and core Mediator (cMed) were purified as described (Plaschka et al., 2016; Schilbach et al., 2017). Promoter DNA was synthesized according to the native *GAT1* sequence (non-template: 5'- GCG GTG CCC GGC CCA GCC ACA TAT ATA TAG GTG TGT GCC ACT CCC GGC CCC GGT ATT AGC ATG CAC GTT TTC TTT CCT TTG CTT T-3' and template: 5'-AAA GCA AAG GAA AGA AAA CGT GCA TGC TAA TAC CGG GGC CGG GAG TGG CAC ACA CCT ATA TAT ATG TGG CTG GGC CGG GCA CCG C-3'). Pol II (0.32 nmol) and TFIIF (1.6 nmol) were added to a premixed complex of *GAT1* promoter DNA (0.8 nmol), TBP (1.6 nmol) and TFIIA (3.2 nmol). TFIIE (3.2 nmol) and cMed (0.38 nmol) were added and the complex was incubated for 90 minutes at 25°C in 10 mM HEPES pH 7, 200 mM KCl, 5% glycerol, 1 mM DTT.

Complex purification was carried out by sucrose-gradient centrifugation and crosslinking (Kastner et al., 2008). The gradient was prepared from a 10% sucrose light solution and a 30% sucrose heavy solution containing 0.075% glutaraldehyde. Ultracentrifugation was done for 16 h at 4°C and 175,000 g. Fractions containing the core PIC-cMed were collected and remaining cross linker was quenched with 15 mM lysine, pH 8.0. Samples were then dialyzed for 5–7 h against 10 mM HEPES pH 7.5, 150 mM KOAc, 2 mM DTT until glycerol and sucrose were removed as determined by refractive index measurements. Dialysed samples were then concentrated to 0.4 mg/ml in a Vivaspinn 500 MWCO 100,000 centrifugal filter and directly used for cryo-EM grid preparation.

### Cryo-EM and image processing of cCC1 and cCC<sup>dist</sup>

For cryo-EM grid preparation, 3.5  $\mu$ L sample were applied to glow-discharged R1.2/1.3 UltrAuFoil grids (Quantifoil, Germany), blotted and plunge-frozen in liquid ethane (Vitrobot (FEI, USA) at 95% humidity, 4°C, 8.5 s blotting time, blot force 13). Data collection was performed on a Titan Krios (FEI) using a K2 direct electron detector (Gatan) in EFTM mode. Data was collected with defocus ranging from  $-0.8$  to  $-3.0 \mu$ m and at a magnification of 130,000x yielding a calibrated pixel size of 1.07 Å/px. The total electron dose was 37 e<sup>-</sup>/Å<sup>2</sup> distributed over 33 movie frames. About 3,800 micrographs were selected manually and dose-weighted and motion-corrected using an in-house developed software based on the MotionCorr algorithms (Li et al., 2013). The CTF was estimated by CTFFIND4 (Rohou and Grigorieff, 2015). If not stated otherwise, all data processing steps were done in RELION (Scheres, 2012).

Initial particle picking was performed using 2D projections of a yeast closed cPIC (Plaschka et al., 2016) filtered to 20 Å and particles were extracted in 300x300 pixel boxes. False positively picked particles were removed manually and by several rounds of 2D classification yielding a clean 385,000 particles dataset (Figure S3B). This dataset was refined in 3D with the yeast OC filtered to 40 Å as a reference. To improve 3D classification particle polishing was performed. From this reconstruction, a mask for the DNA was isolated and used for 3D classification yielding CC and OC complexes. CC particles were pooled and reclassified using a mask containing the Tfg2 WH, TFIIE and the Pol II stalk yielding cCC1 and cCC<sup>dist</sup>. Final classes were refined in 3D and B-factor sharpened with RELION. cCC1 was refined to 5.1 Å resolution (0.143 FSC) with a B-factor ranging from  $-80$  to  $-150 \text{ Å}^2$  and cCC<sup>dist</sup> was refined

to 4.8 Å resolution (0.143 FSC) applying B-factors of -100-130 Å<sup>2</sup>. Local resolution estimation was done by a combination of local FSC-based filtering and local B-factor sharpening. Decreasing local resolution of the map periphery is apparent in the locally filtered maps and local resolution histogram for resolution bins (Figure S3D).

### Data processing and CC<sup>dist</sup> reconstruction

A set of 255,000 particles was extracted from raw data used to reconstruct a previously published PIC (Schilbach et al., 2017). Closed and open complex particles were separated using a mask on promoter DNA. CC particles were reclassified using a global mask and TFIID-containing classes were subjected to a final round of 3D classification. The two best classes were merged and refined in 3D to 6.7 Å with a fixed B-factor of -100 Å<sup>2</sup>. Locally filtered maps were calculated as for cCC1 and cCC<sup>dist</sup>, revealing a well-resolved cPIC at ~4 Å and TFIID at lower resolution.

### Model building and refinement

The protein models for cCC1 and cCC<sup>dist</sup> were built based on the previous PIC structure (Schilbach et al., 2017) and real-space fitted as rigid body domains with Phenix (Adams et al., 2010). DNA was fitted manually to the map in COOT (Emsley and Cowtan, 2004). Iterative rounds of real space refinement and geometry optimization in Phenix (Adams et al., 2010) yielded a model for cCC1 and cCC<sup>dist</sup> DNA. Because the local resolution of the DNA was ~5 Å, strong B-DNA geometry restraints were used during refinement of cCC1 DNA. For cCC<sup>dist</sup> DNA, deviations of helical twist and base pair shift from the standard B-DNA values were required to obtain a good fit. CC<sup>dist</sup> was also built from the existing PIC model (Schilbach et al., 2017) and DNA from the cCC<sup>dist</sup> structure. Promoter DNA was then extended by B-DNA pieces generated with 3D-DART (van Dijk and Bonvin, 2009) that contained map-matching bends in the DNA. Proteins were real-space refined as rigid body modules and DNA was refined using the same strategy as for cCC1 and cCC<sup>dist</sup>.

## QUANTIFICATION AND STATISTICAL ANALYSIS

Data for promoter opening assays are mean values of at least three technical replicates.

## DATA AND SOFTWARE AVAILABILITY

Coordinates for cCC1, cCC<sup>dist</sup> and CC<sup>dist</sup> structures were deposited with the Protein Data Bank with accession codes PDB: 6GYK, 6GYL, and 6GYM, respectively. The associated electron microscopy maps were deposited with the Electron Microscopy Data Bank with accession codes EMD: 0090, EMD: 0091, and EMD: 0092.

**Molecular Cell, Volume 73**

**Supplemental Information**

**Promoter Distortion and Opening  
in the RNA Polymerase II Cleft**

**Christian Dienemann, Björn Schwalb, Sandra Schilbach, and Patrick Cramer**



Figure S1

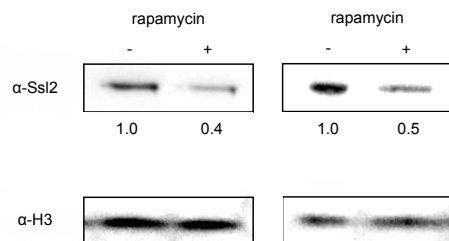
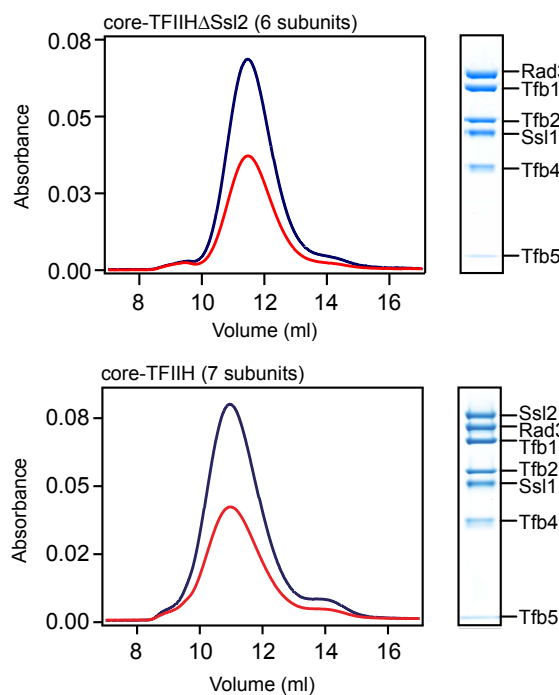
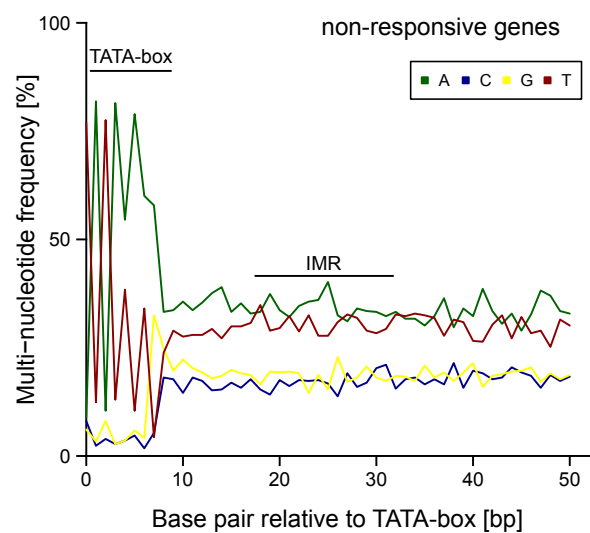
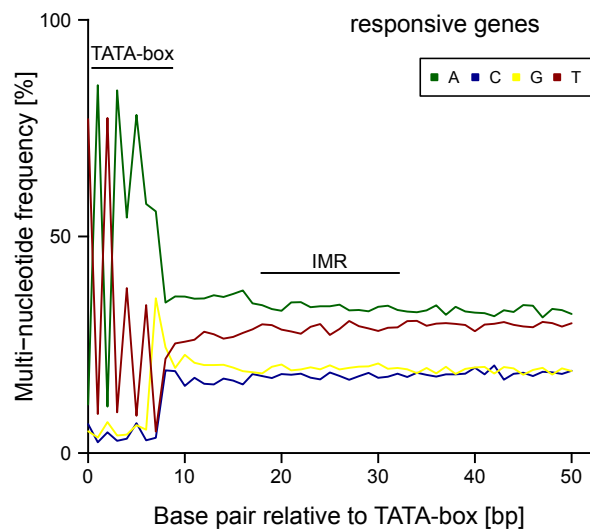
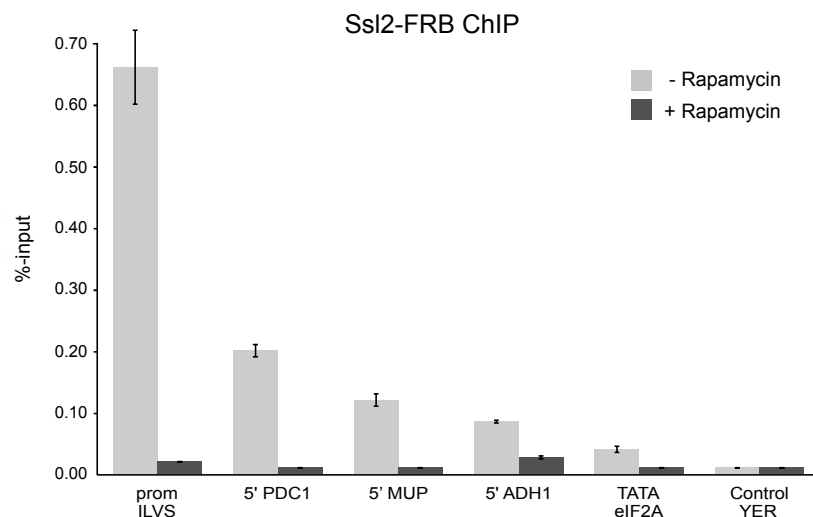
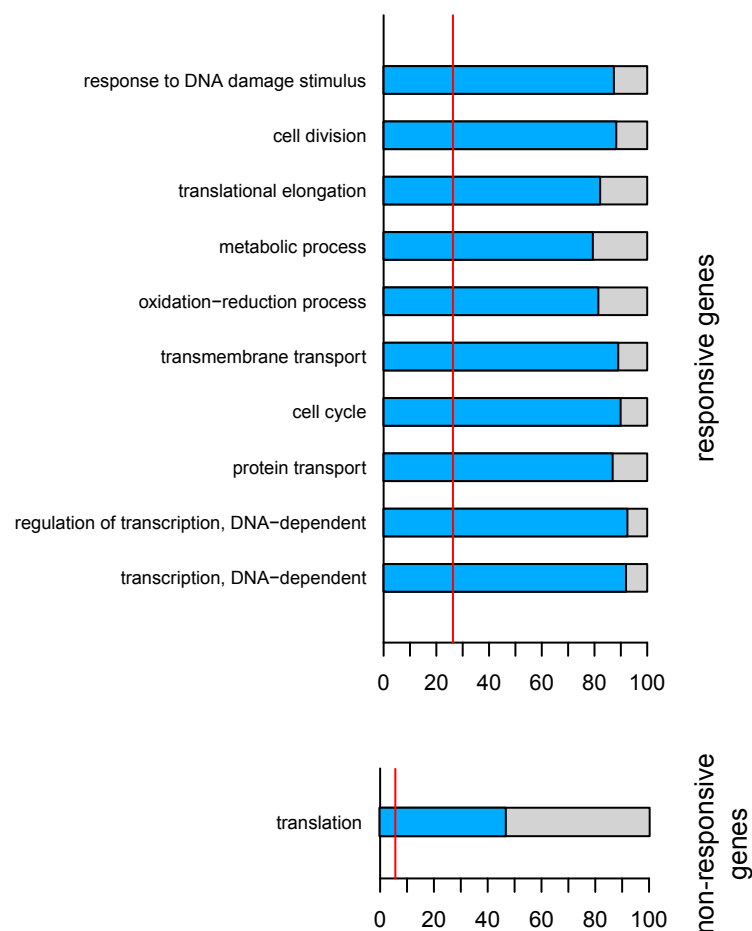
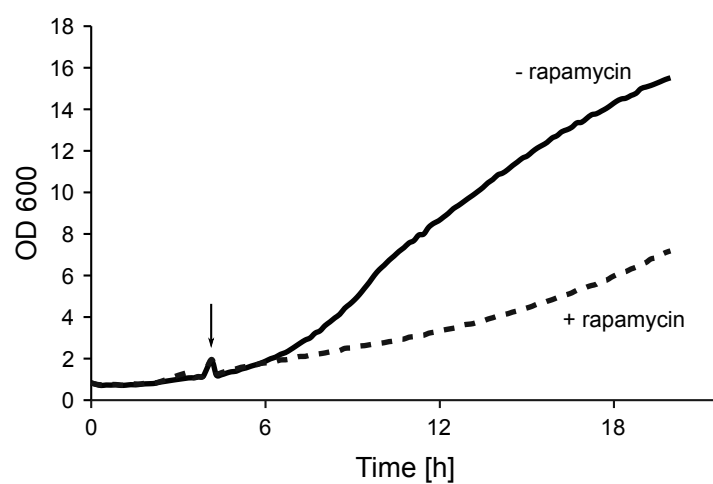
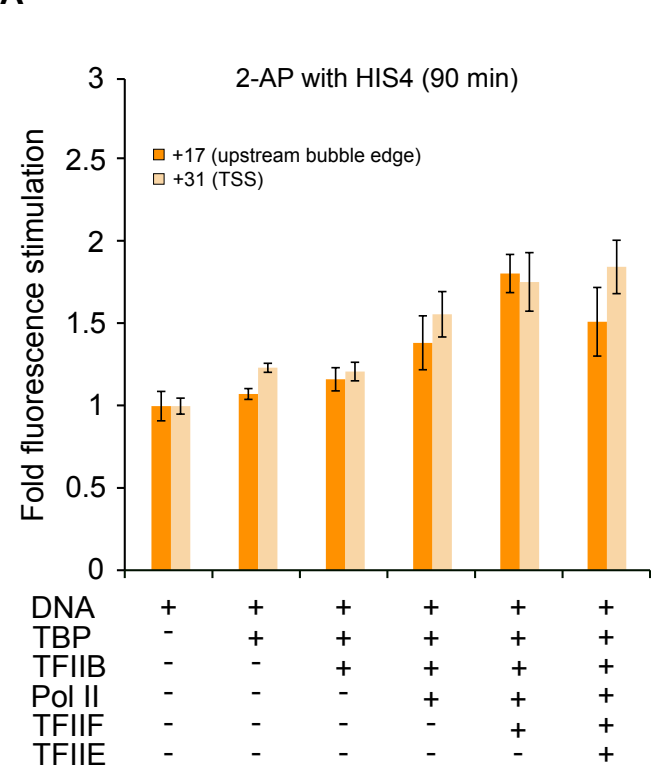
**A****B****E****C****D****F**

Figure S2

**A**

TATA Box  
 HIS4 GTGTATATAAATAGCTATGGAACGTTTCGATTCACCTCCGATGTGTGTTG  
 GAT1 ACATATATATAGGTGTGTGCCACTCCCGGCCACGGTATTAGCATGCAC  
 GAT1mut ACATATATATAGGTGTGTGCCACTTCCATTACGGTATTAGCATGCAC

0 +10 +20 +30 +40

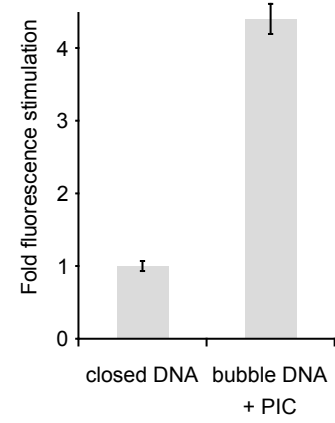
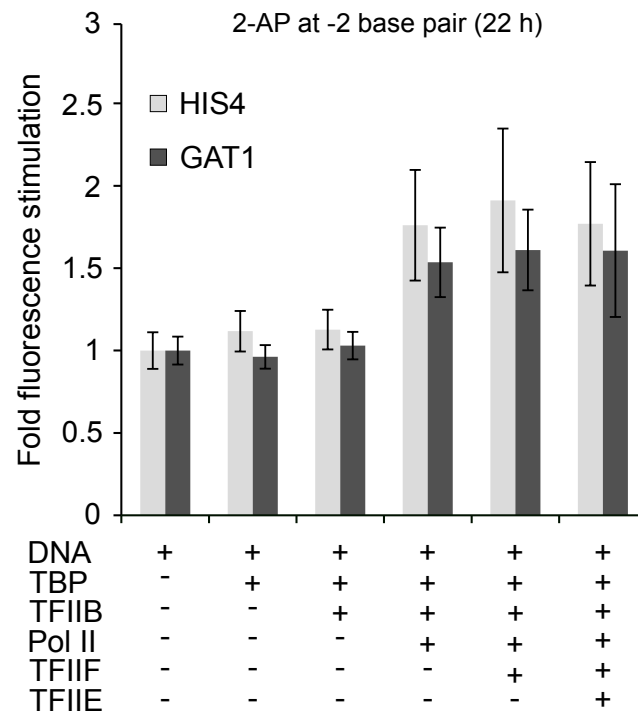
**B****C**

Figure S3

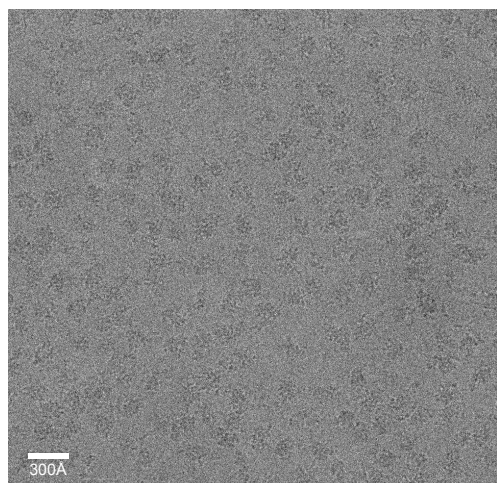
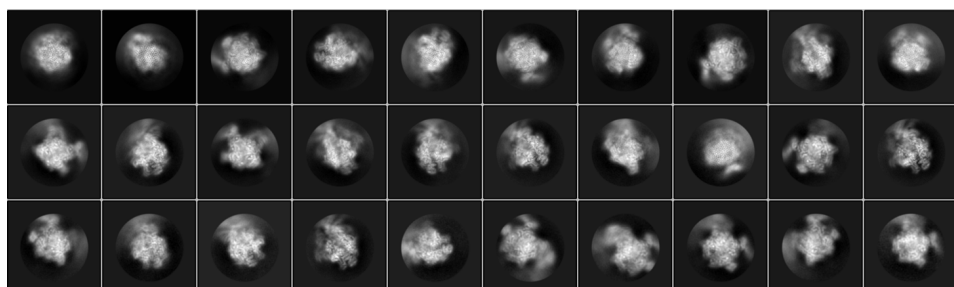
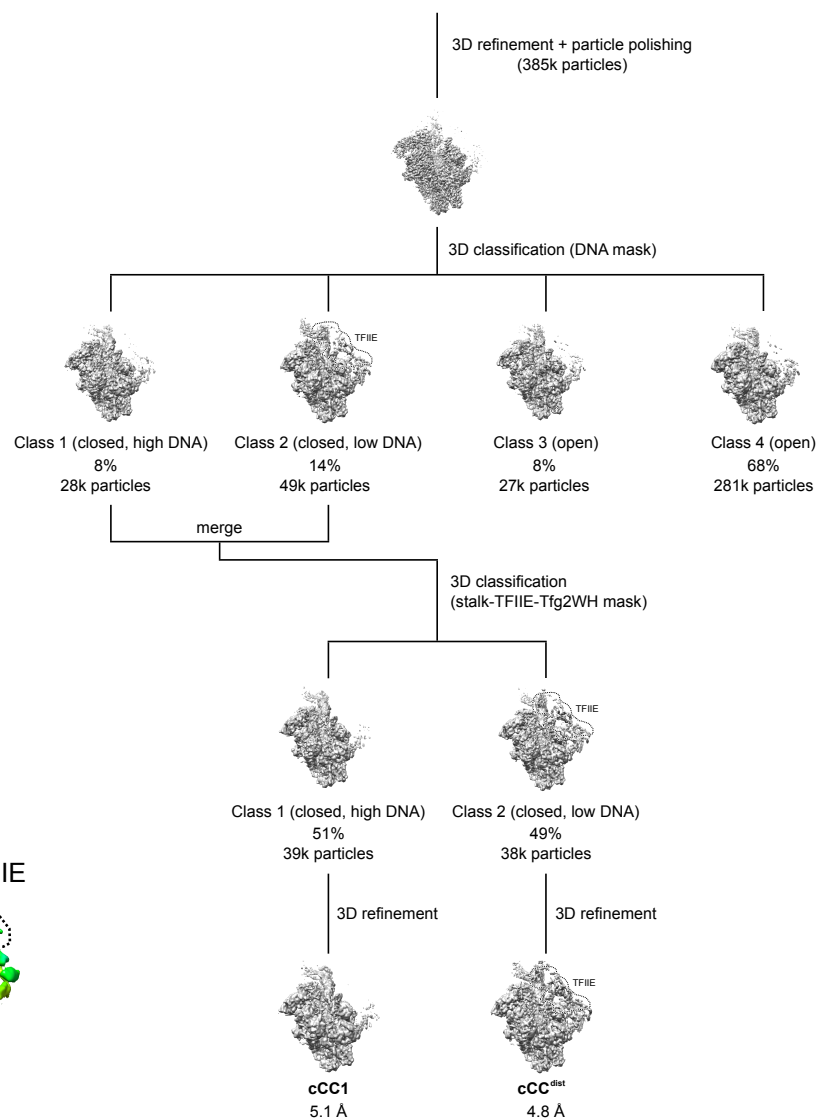
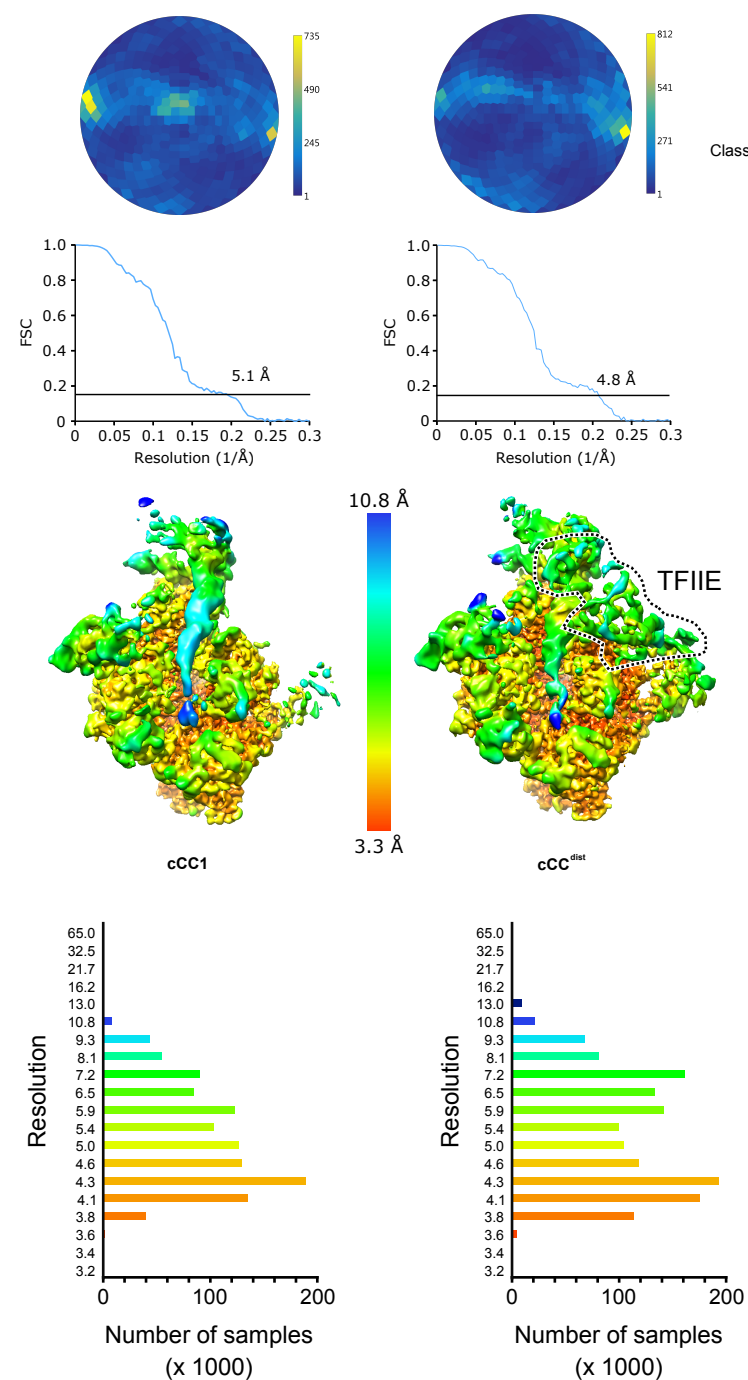
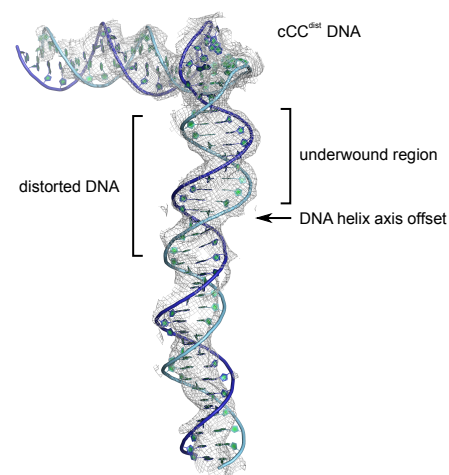
**A****B****C****D****E**

Figure S4

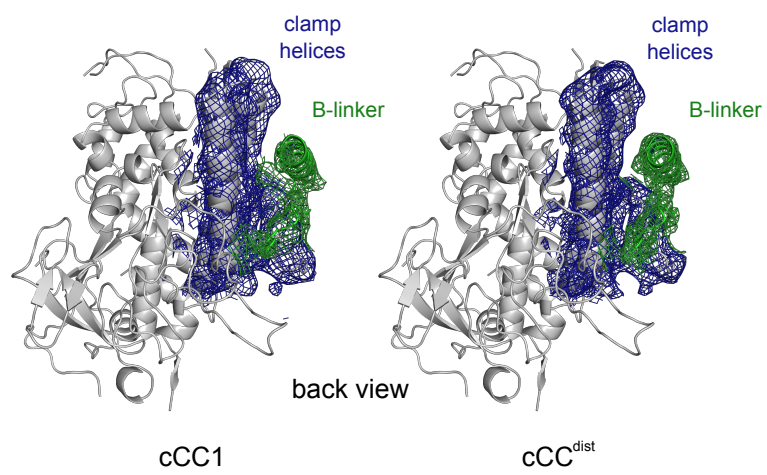
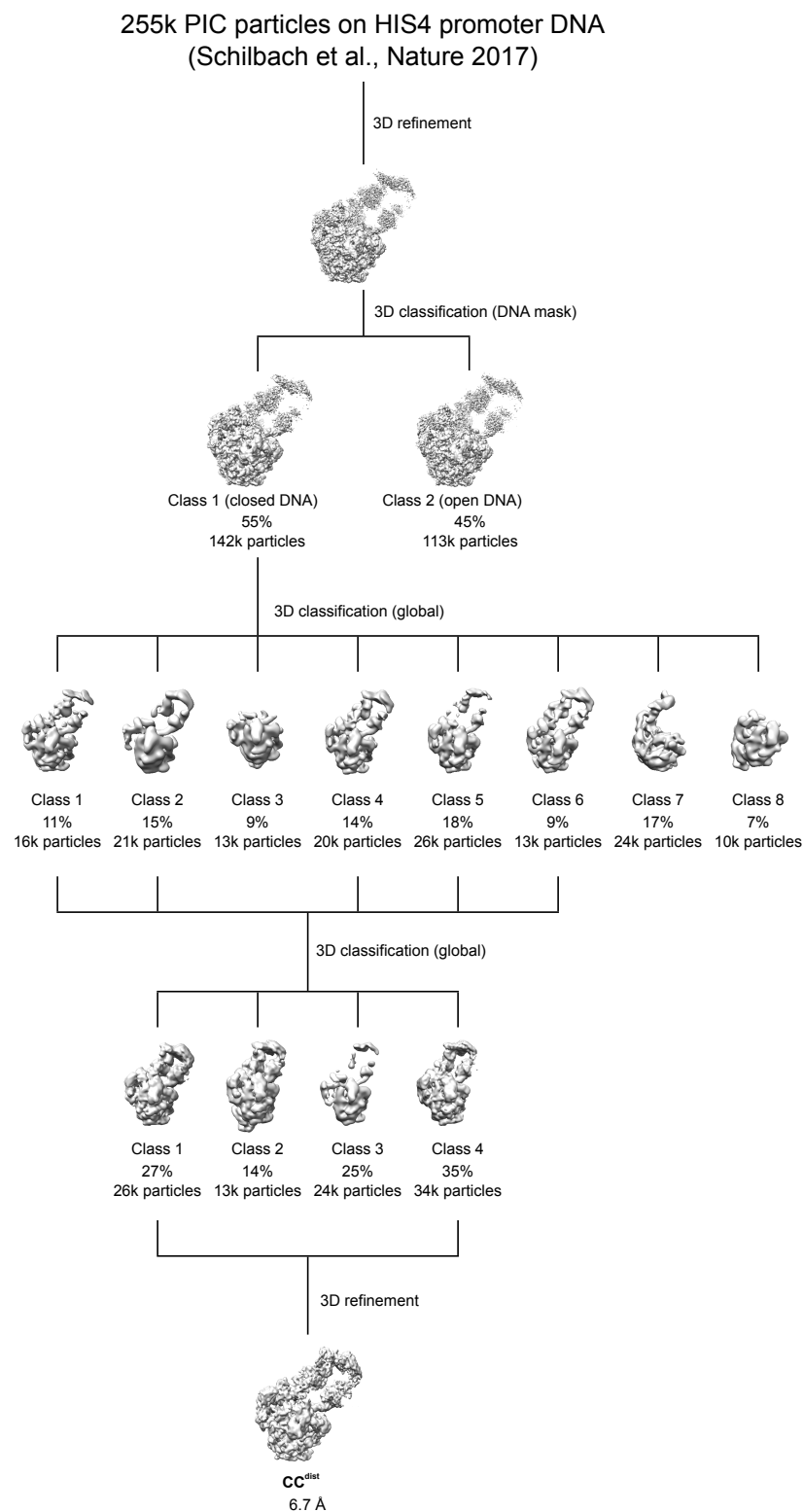


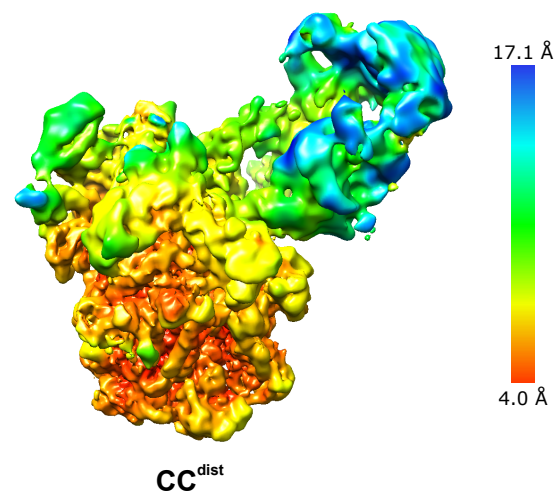
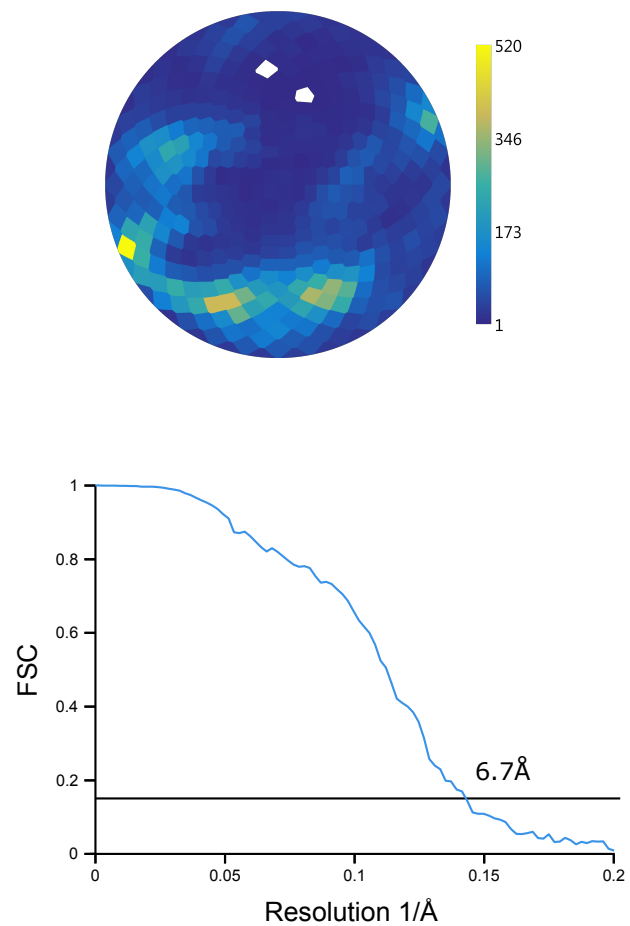


Figure S5

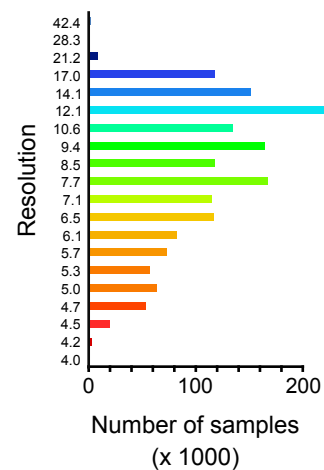
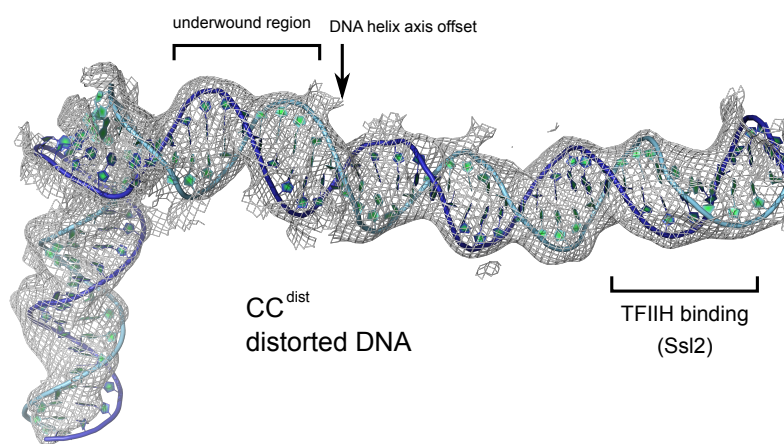
A



B



C



### Figure S1 | Validations and controls for Ssl2 anchor away, Related to Figure 1

- (A) Western blot analysis of purified yeast nuclei with and without rapamycin treatment. Ssl2-depletion was detected by specific antibodies. H3 was used as loading control. Two replicates are shown. Quantification shows the factor of band intensity reduction.
- (B) Yeast TFIIH can be purified without Ssl2. Chromatograms of analytical gel filtrations are shown for wild type TFIIH and TFIIH lacking the Ssl2 subunit. The integrity of both complexes was detected by SDS-PAGE.
- (C) Ssl2-FRB ChIP qPCR without and with rapamycin treatment. Genes were chosen from Ssl2-independent (PDC1) and Ssl2-dependent genes (others). Primers for qPCR were chosen in the 5' region of the respective gene close to the promoter. YER is the random chromatin control.
- (D) GO-term enrichment of Ssl2-dependent and independent genes. Enrichment analysis was done using Fisher tests. Top GO categories are sorted by p-values. The blue bars depict the percent fraction of genes that are enriched from each GO term. The red line indicates the proportion of each category that is expected by chance.
- (E) Ssl2-dependent and -independent genes do not show base enrichment in their IMR. Average base content of promoter regions downstream of TATA for Ssl2-dependent genes are shown in % for each base position. Genes were aligned at their TATA box. TATA box and IMR are marked with solid lines.
- (F) Rapamycin treatment induces growth phenotype in Ssl2 anchor away yeast. The growth curve of the Ssl2 anchor away strain is shown under normal (solid) and Ssl2 depletion (dashed) conditions. Growth curves were collected over 20 hours in YPD medium at 30°C.

### Figure S2 | TFIIH independent opening of HIS4 and GAT1 promoter DNA, Related to Figure 2

- (A) TFIIH independent opening of HIS4 promoter DNA can be detected across the IMR. 2-aminopurine (2-AP) labels were incorporated at the upstream bubble edge (+17 bases from the TATA box) and around the region where the TSS would be optimally placed (+31). Bar plots show the fluorescence increase normalized to the DNA only reaction.
- (B) 2-AP Fluorescence stimulation with miss-matched HIS4 promoter DNA. The artificial bubble template DNA was created by introducing mismatches in the template DNA strand.
- (C) The GAT1 promoter melts spontaneously *in vitro* after 22 hours. Spontaneous melting of the GAT1 promoter is achieved after incubation time with cPIC components was extended to 22 hours. Bar plots show the fluorescence increase normalized to the DNA only reaction. 2-AP labelled HIS4 promoter DNA after 22 hours is shown for comparison. Data is shown as mean  $\pm$  SEM in all panels.

### Figure S3 | Data processing of cCC1 and cCC<sup>dist</sup>, Related to Figure 3

- (A) A micrograph representative for the collected data is shown. The micrographs have been contrast enhanced for clarity.
- (B) Selected 2D classes that were used after manual and 2D-classification based particle cleanup. 2D classes are ordered by number of particles in descending order.
- (C) Schematic for processing strategy of the cCC1/cCC<sup>dist</sup> dataset. Reconstructions of each processing step are shown as grey surfaces. Particle numbers are rounded. Resolution is

given according to the 0.143 FSC threshold and with a B-factor of  $-100\text{\AA}^2$ .

**(D)** Angular distribution and Fourier-shell correlation (FSC) plots for cCC1 and cCC<sup>dist</sup> reconstructions. Angular distribution was calculated in 7.5° orientation bins. The 0.143 FSC threshold is marked in FSC plots as solid line. Local resolution was calculated by a combination of local FSC weighting and B-factor sharpening. The resolution histogram represents the number of samples per resolution bin. Locally filtered volumes of the reconstructions for cCC1 and cCC<sup>dist</sup> are shown in front view (Figure 3A and B).  
**(E)** Model and map of the distorted cCC<sup>dist</sup> DNA. The model is shown in cartoon representation, the locally filtered map as grey mesh.

#### **Figure S4 | Clamp and B-linker of cCC1 and cCC<sup>dist</sup>, Related to Figure 3**

The B-linker is well ordered in cCC1 and cCC<sup>dist</sup>. The Pol II clamp and B-linker of cCC1 and cCC<sup>dist</sup> are shown in cartoon representation. The map for clamp helices and B-linker is shown as blue and green mesh, respectively.

#### **Figure S5 | Data processing of CC<sup>dist</sup>, Related to Figure 5**

**(A)** 3D sorting of PIC particles leads to the CC<sup>dist</sup> reconstruction. A schematic of the 3D sorting strategy for the CC<sup>dist</sup> reconstruction is shown. Intermediate reconstructions are shown as grey volumes. Particle numbers are rounded. Resolution is given according to the 0.143 FSC threshold and with a B-factor of  $-100\text{\AA}^2$ .

**(B)** Angular distribution and Fourier-shell correlation (FSC) plots for the CC<sup>dist</sup> reconstruction. Angular distribution was calculated in 7.5° orientation bins. The 0.143 FSC threshold is indicated as solid line. Local resolution was calculated by a combination of local FSC weighting and B-factor sharpening. Locally filtered volume of CC<sup>dist</sup> is shown in side view (Figure 5A). The resolution histogram represents the number of samples per resolution bin.

**(C)** Model and map of the distorted CC<sup>dist</sup> DNA. The model is shown in cartoon representation, the locally filtered map as grey mesh.

**Table S1 | Eukaryotic closed complex structures**

			<i>Pol II</i>				<i>Pol I*</i>	<i>Pol III</i>
<b>State</b>	<b>cCC</b>	<b>cCC1</b>	<b>cCC2</b>	<b>CC2</b>	<b>cCC<sup>dist</sup></b>	<b>CC<sup>dist</sup></b>	<b>CC2</b>	<b>CC1</b>
<b>Reference</b>	Plaschka et al., 2016	this study	He et al., 2016	He et al., 2016	this study	this study	Engel et al., 2017	Vorländer et al., 2018
<b>Resolution</b>	7.8 Å	5.1 Å	5.4 Å	7.2 Å	4.8 Å	8.0 Å	modeled	5.5/4.2 Å
<b>DNA distance from active site</b>	72.9 Å	71.5 Å	55.9 Å	55.9 Å	64.3 Å	64.3 Å	44.9 Å	68.4 Å
<b>DNA conformation</b>	canonical	canonical	canonical	canonical	distorted	distorted	canonical	canonical
<b>Clamp</b>	closed	closed	open	open	closed	closed	open	closed
<b>B-linker</b>	weak	+	-	-	+	+		
<b>B-reader</b>	+	+	-	-	+	+		
<b>TFIIE</b>	weak	-	+	+	+	+		
<b>Tfg2WH</b>	weak	-	+	+	+	+		

\* based on structural modeling.