Structure and mechanism of the RNA polymerase II transcription machinery

Allison C. Schier and Dylan J. Taatjes

Department of Biochemistry, University of Colorado, Boulder, Colorado 80303, USA

RNA polymerase II (Pol II) transcribes all protein-coding genes and many noncoding RNAs in eukaryotic genomes. Although Pol II is a complex, 12-subunit enzyme, it lacks the ability to initiate transcription and cannot consistently transcribe through long DNA sequences. To execute these essential functions, an array of proteins and protein complexes interact with Pol II to regulate its activity. In this review, we detail the structure and mechanism of over a dozen factors that govern Pol II initiation (e.g., TFIID, TFIIH, and Mediator), pausing, and elongation (e.g., DSIF, NELF, PAF, and P-TEFb). The structural basis for Pol II transcription regulation has advanced rapidly in the past decade, largely due to technological innovations in cryoelectron microscopy. Here, we summarize a wealth of structural and functional data that have enabled a deeper understanding of Pol II transcription mechanisms; we also highlight mechanistic questions that remain unanswered or controversial.

Transcription by the RNA polymerase II (Pol II) enzyme occurs not only at annotated protein-coding genes but throughout the genome, and is fundamentally important for most physiological processes. The 12-subunit Pol II enzyme is conserved throughout eukaryotes, and is distinguished from other RNA polymerase enzymes (e.g., Pol I or Pol III) (for review, see Khatter et al. 2017; Engel et al. 2018) by the types of genes/sequences that it transcribes and by the factors and mechanisms that control its function. Many Pol II regulatory factors have been identified over the years (Thomas and Chiang 2006; Grünberg and Hahn 2013; Kwak and Lis 2013; Sainsbury et al. 2015; Roeder 2019), but a detailed understanding of their structure and function has been challenging because of the large size and conformationally flexible nature of the Pol II transcription machinery. Recent technological advances in cryoelectron microscopy (cryoEM) have enabled the structural characterization of these complexes at increas-

ingly high resolution, which has rapidly advanced understanding of the molecular basis of Pol II transcription.

Structural biology continues to transform our understanding of complex biological processes because it allows visualization of proteins and protein complexes at or near atomic-level resolution. Combined with mutagenesis and functional assays, structural data can at once establish how enzymes function, justify genetic links to human disease, and drive drug discovery. In the past few decades, workhorse techniques such as NMR and X-ray crystallography have been complemented by cryoEM, cross-linking mass spectrometry (CXMS), and other methods. Recent improvements in data collection and imaging technologies have transformed cryoEM into a powerhouse structural technique that rivals X-ray crystallography in terms of resolution but does not require crystals (Kuhlbrandt 2014).

In the past, cryoEM has been limited by sample contrast, radiation damage, radiation-induced motions in the sample, and insufficient software for image processing. Advancements in these areas have allowed researchers to determine cryoEM structures to near angstrom resolution. One of the most important technological advancements was in the cameras that are used to acquire electron micrographs. Past methods acquired data with film or CCD cameras, which consisted of a single exposure that would be subject to blurring due to beam-induced sample movement and/or radiation damage. With new, state-of-the-art direct detection cameras, many images are acquired over several seconds, generating "movies" with many individual micrograph "frames" (Scheres 2014). This allows individual frames to be aligned, correcting for beam-induced motions. Additionally, early frames can be removed to reduce blurring that results from initial exposure of the sample, and later frames can be removed to reduce the impact of radiation damage that accrues during data acquisition (Scheres 2014). Improvements to image processing software have kept pace (Punjani et al. 2017; Zivanov et al. 2018; Wagner et al. 2019), and with more processing power and pipelined approaches, it is now faster and easier to generate 3D

models. Taken together, these innovations have improved the resolution of cryoEM reconstructions to the near-atomic range and allowed the analysis of increasingly smaller proteins or protein complexes.

In this review, we describe the structure and function of the Pol II transcription machinery, with an emphasis on the structural data that have provided key mechanistic insights. We highlight basic functions and structural interfaces among over a dozen proteins or protein complexes, most of which directly interact with the Pol II enzyme. Throughout, we also highlight some controversial or unanswered questions about Pol II transcription mechanisms.

The preinitiation complex (PIC): a universal intermediate in Pol II transcription?

Although other factors can be considered PIC components, we define the PIC to consist of eight factors: TFIIA, TFIIB, TFIID, TFIIE, TFIIF, TFIIH, Pol II, and Mediator (Fig. 1A). Of these eight factors, some are relatively small in size (e.g., TFIIB is a single subunit, TFIIE is a dimeric complex) whereas others are large, multisubunit assemblies (TFIID, TFIIH, Mediator, and Pol II itself). In general, all eight PIC factors assemble at protein-coding genes, but the composition of the PIC may be distinct at different classes of Pol II transcribed genes (e.g., lncRNAs or snRNAs) (Sadowski et al. 1993) or in different cell types (Deato and Tjian 2007).

Like other macromolecular assemblies in biology (e.g., the proteasome or the ribosome), the PIC functions like a machine: It does work and has moving parts and components that fit precisely together. The PIC assembles on regions of genomic DNA called promoters. Once bound, the PIC can "open" the promoter DNA to initiate transcription (Fig. 1B; see below). The PIC also appears to be re-

quired for bidirectional transcription that is commonly observed at enhancers in mammalian organisms (Core et al. 2014; Duttke et al. 2015; Scruggs et al. 2015). Eukaryotic promoters contain DNA motifs that are distinguished by their sequence and their position relative to the transcription start site (TSS). Promoter DNA sequences are recognized by PIC factors (see below), and as a consequence, promoter DNA represents a "template" that directs the proper assembly of the PIC. In this review, we touch upon the role of promoter sequence motifs, and we define a promoter to represent DNA sequences 250 bp upstream of and 50 bp downstream from the TSS. Different promoter sequence motifs have been characterized and are reviewed in detail elsewhere (Vo Ngoc et al. 2017).

In the following sections, we outline the structural organization of the eukaryotic PIC and then describe the conformational changes that accompany promoter melting and transcription initiation.
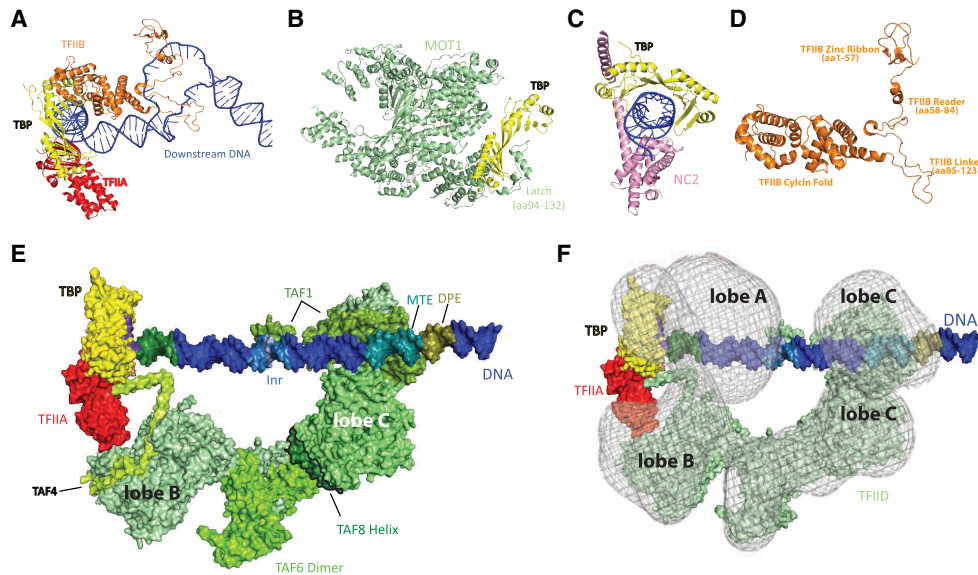
TBP: a regulatory hub within TFIID

TBP (TATA binding protein) has ancient origins as a transcription factor; in eukaryotes, TBP regulates Pol II transcription as a component of the TFIID complex (see below). Although only a single protein of intermediate size (human TBP has 339 residues), TBP is a major interaction hub within the PIC. TBP adopts a crescent shape and binds DNA through its concave surface (Fig. 2A). Factors that inhibit TBP–DNA binding interact with this concave surface and include TFIID subunits TAF11/13 (Gupta et al. 2017) and the TAF1 TAND (TAF1 N-terminal domain) (Liu et al. 1998; Anandapadamanaban et al. 2013). The TAF11/13 interaction with TBP blocks TFIIA-TBP binding (Gupta et al. 2017) and TFIIB-TBP binding, which will prevent PIC assembly. MOT1 and NC2 are other factors that interact with the concave surface of TBP (Fig. 2B,
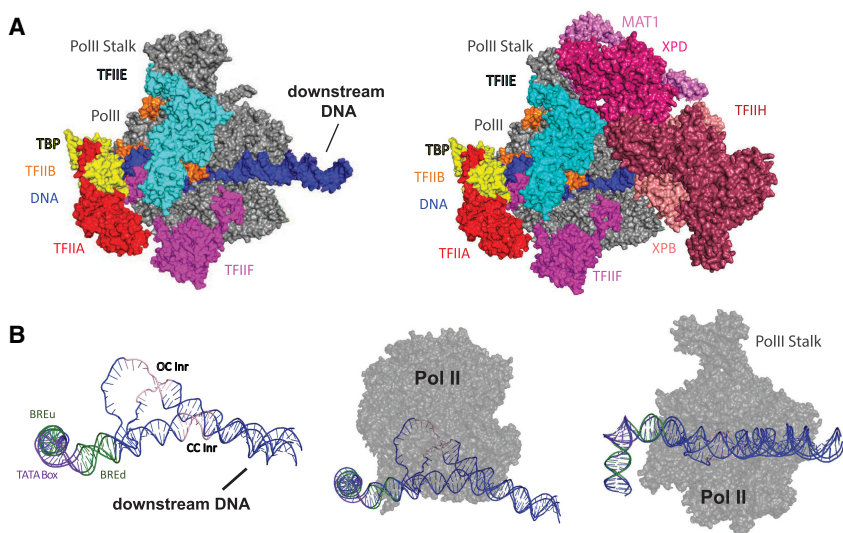
Figure 1. Overview of the PIC and DNA path in closed and open complex. (*A*) The preinitiation complex (PIC) consists of TFIIA (red), TFIIB (orange), TFIID (not pictured), TFIIE (cyan), TFIIF (magenta), TFIIH (maroon), RNA polymerase II (Pol II, gray), and promoter DNA (blue). Upstream promoter DNA is bound by TFIIB, TBP, TFIIA, and TFIIF. Downstream DNA is bound by TFIIH and TFIID. After promoter opening, TFIIB, TFIIE, and TFIIF interact with and stabilize the ssDNA in the Pol II cleft. PIC is shown with TFIIH (*left*) and without TFIIH (right). Adapted from PDB 5IY7 (He et al. 2016). (*B*) Promoter DNA before and after opening. Closed complex (CC) DNA is mostly linear with the characteristic 90° bend at the TATA box (purple). Upon promoter opening, the strands separate at the transcription start site (denoted Inr, light pink). *Center* and *right* images show Pol II overlaid with the DNA to show location in the active site. The center view shows the top of Pol II, with ssDNA extending into the active site. The *right* view shows the front of Pol II with the DNA. Adapted from PDB 5IY6 and PDB 5IY7 (He et al. 2016).

Figure 2. TBP is a regulatory hub within the larger TFIID complex. (*A*) Structure of TBP (yellow), TFIIA (red), and TFIIB (orange) bound to promoter DNA in an open PIC complex. The TFIIB zinc ribbon (amino acids 1–57, amino acids 7–57 pictured) is important for recruiting Pol II to the TSS. Amino acids 56–60 in the TFIIB zinc ribbon and B-reader (amino acids 58–84) domains stabilize the open template strand, whereas amino acids 86, 98–100, and 103 in the linker domain (amino acids 85–123) interact with the nontemplate strand (Kostrewa et al. 2009). TFIIA residues important for TBP interaction (TBP residues 187–208): 345–349, 375, and 376 from TFIIA subunit 1, and 65–67 in TFIIA subunit 2. TFIIB residues that interact with TBP (residues 271, 274, 278, 283–287, 306, and 337): 169, 177, 188, 195, 205, 208, 243, 246, 247, and 249. Adapted from PDB 5IY7 (He et al. 2016). (*B*) MOT1 (light green) binds TBP (yellow) and displaces in an ATP-dependent manner. MOT1 contains 16 HEAT repeats that bind multiple regions of TBP. MOT1 also binds upstream DNA; MOT1 contains a "latch" (amino acids 94–132) that blocks TBP–DNA reassociation. Adapted from PDB 3OC3 (Wollmann et al. 2011). (*C*) NC2 (pink) binds TBP (yellow) and DNA to negatively regulate transcription. NC2 binds the DNA major groove and an NC2 α helix (amino acids 180–210) sterically blocks TFIIB–TBP interactions. Adapted from PDB 1JFI (Kamada et al. 2001). (*D*) Overall structure of human TFIIB in the PIC. Structural domains include the TFIIB core (with two cyclin folds), the linker (amino acids 85–123), the reader (amino acids 58–84), and the zinc ribbon (amino acids 1–57). These domains are important for stabilizing single-stranded promoter DNA in the open complex. Adapted from PDB 5IY7 (He et al. 2016). (*E*) Structure of TFIID bound to promoter DNA, along with TFIIA and TBP. TFIIA–TFIIB–TBP bind upstream of the TSS. In this example, the supercore promoter (SCP) was used (Juven-Gershon et al. 2006), which has multiple promoter elements, some or all of which are not found in promoters genome-wide. However, the protein–DNA interactions shown here are likely to occur at promoters with different sequences. SCP upstream DNA elements are the BREu (−37 to −32), TATA (−31 to −24), and BREd (−23 to −17). TFIIA residues 68–71 and 35–27 interact with the TAF4/12 dimer in TFIID lobe B (TAF4 residues 1002-1007; TAF12 residue 75). TAF1 residues 972, 996, 1022, and 1023 bind the Inr sequence (−3 to +3, with +1 shown in light blue); residues 797, 843, 844, 852, and 862 bind the motif ten element (MTE; +18 to +27), and TAF1 residues 839, 843, 852, 856, and 858 bind the downstream promoter element (DPE; +28 to +34). TAF2 (residue 543) also interacts with the MTE. Another interaction involves the TAF4 hairpin (amino acids 966–1000) and upstream DNA at the end of the BREd. Adapted from PDB 5IY7 (He et al. 2016) and PDB 6MZM (Patel et al. 2018). (*F*) The structure in *E* includes only structured domains; here, we show a rendering of the "rearranged" free TFIID structure (semiopaque), which shows the entire TFIID density (i.e., including disordered regions). The free TFIID structure (adapted from EMD-2284) (Cianfrocco et al. 2013) was visually aligned and superimposed onto the structure shown in *E*. This rendering highlights additional TFIID density downstream from the MTE/DPE promoter elements and shows an approximate position of the mobile lobe A, behind lobe B. Adapted from PDB 6MZM (Patel et al. 2018). Amino acid interactions were determined using the "find any contacts" function in PyMol, set to 4 Å. Amino acids listed correspond to the human proteins.

C). MOT1 and NC2 are conserved from yeast to humans and appear to help regulate TBP interactions with promoter DNA and with PIC factors (Auble et al. 1994; Wollmann et al. 2011). For instance, NC2 blocks TFIIA or TFIIB binding to TBP (Kamada et al. 2001; Gilfillan et al. 2005), at least partially through competition for the same binding sites on TBP (Kim et al. 1995).

Based on recent structural data from the Nogales laboratory (Patel et al. 2018), it appears that TFIID functions as a delivery vehicle for TBP, and once TBP is deposited it may function independently of TFIID (i.e., the TAFs of TFIID could dissociate from the promoter). Under such a sce-nario, factors that regulate TBP occupancy could have enormous regulatory importance. For instance, NC2 and MOT1 could be instrumental for TBP removal and shutting down active transcription. Mot1 is essential in yeast and its ATPase function impacts TBP occupancy across the genome (Sprouse et al. 2006; Venters et al. 2011). Efforts to define the regulatory importance of MOT1 or NC2 have shown modest effects in yeast or mammals (van Werven et al. 2008), primarily suppressing cryptic transcription (e.g., transcription initiation within gene bodies) (Koster et al. 2014; Xue et al. 2017). This suggests auxiliary or redundant roles for other factors.

Interestingly, although NC2 and MOT1 inhibit TBP binding to TATA or TATA-like sequences, this appears to promote transcription of genes that lack such sequences in *Drosophila* (Hsu et al. 2008); moreover, promoters containing TATA + Inr elements are refractory to NC2 repression (Malecová et al. 2007). Thus, factors that block TBP–DNA interactions are not universally repressive but instead may direct transcription from specific types of promoters. The MYC transcription factor was also shown to interact with the concave surface of TBP (Wei et al. 2019). MYC is a well-studied transcription factor that activates transcription through mechanisms that remain incompletely understood (Rahl et al. 2010). Potentially, MYC–TBP interactions may alleviate repressive interactions with the TAF1 TAND or TAF11/13 to promote transcription at MYC-responsive genes.

TBP–TFIIA–TFIIB

Once deposited at its DNA-binding site (e.g., a TATA or TATA-like sequence) upstream of the TSS, TBP bends DNA by ~90° (Kim et al. 1993b; Geiger et al. 1996; Tan et al. 1996). Within the PIC, this bent DNA structure is further stabilized by TFIIA, TFIIB, and TFIIF (see below). TBP-induced DNA bending helps shed repressive interactions between TBP and TAF11/13 (Patel et al. 2018) but may also set up a specific 3D architecture for active promoters: DNA bending will reposition promoter-bound factors to enable interactions that would otherwise not be possible on linear DNA. Although TBP binding affinity for TATA sequences is high (ca. 2 nM in absence of other factors), its DNA binding is blocked within TFIID (see below) to prevent promiscuous interactions with genomic DNA. TBP-DNA binding is further stabilized by TFIIA and TFIIB (Imbalzano et al. 1994), which bind TBP through opposite ends of its crescent structure (Fig. 2A). TFIIA and TFIIB, in turn, establish an orientation preference for TBP (Kays and Schepartz 2000) and form a network of interactions within a fully assembled PIC (see below). The ability of TFIIA and TFIIB to stabilize DNA-bound TBP (Imbalzano et al. 1994; Hieb et al. 2007) may be especially important at so-called TATA-less genes, which predominate in mammalian genomes (Vo Ngoc et al. 2017). In contrast, almost all yeast protein-coding genes contain a TATA or TATA-like sequence upstream of the TSS (Rhee and Pugh 2012).

In addition to its TBP interaction, TFIIA interacts directly with the TAF4/12 dimer within TFIID lobe B (see below), and TFIIA binding favors TFIID lobe A rearrangement to a transcriptionally-competent state (Cianfrocco et al. 2013; Louder et al. 2016; Patel et al. 2018). TFIIB binds the opposite side of TBP (relative to TFIIA) and directly interacts with Pol II near the RPB1 dock (Kostrewa et al. 2009; Liu et al. 2010), thus physically linking DNA-bound TBP to the Pol II enzyme. In fact, TFIIB extends from TBP to the Pol II active site, but physically blocks an RNA exit channel (see below). During transcription initiation, the TFIIB linker helix and the B-reader loop play a role in stabilizing open promoter DNA (Fig. 2D), which is described in a later section.

TFIID

Human TFIID is ~1.3 MDa in size and contains TBP + 13 TAFs (TBP-associated factors) that are present in one or two copies each (Table 1). TFIID has a horseshoe-shaped architecture (Fig. 2E) with three lobes: A, B, and C (Louder et al. 2016). The yeast counterpart shares basic structural and functional aspects (Papai et al. 2009; Kolesnikova et al. 2018), but differs in some important ways (see below). Cryo-EM data have revealed human TFIID to be an unusually flexible and dynamic complex, and the functional relevance of these characteristics is just beginning to be understood. The structural dynamics of human TFIID is reflected in its mobile "lobe A," which appears to spontaneously sample interactions with lobe C or lobe B that are separated by about 100 Å (Cianfrocco et al. 2013). Lobe A contains TAF5, TAF6/9, TAF4/12, and TAF3/10 (Patel et al. 2018), in which histone-fold containing dimerized TAFs are listed together. Lobe A also contains a TAF11/13 dimer and TBP, which represents an important regulatory module within TFIID (see above).

TFIID lobe B contains many of the same subunits as lobe A, except that it lacks TBP and TAF11/13 and lobe B contains a TAF8/10 dimer instead of TAF3/10 (Table 1). Lobe B is important for binding upstream DNA, and it binds TFIIA through its TAF4/12 subunits (Patel et al. 2018). The C lobe contains TAF1 and TAF2—the largest subunits of TFIID—and TAF7, which interacts through a central domain of TAF1 (Bhattacharya et al. 2014; Wang et al. 2014). The "BC core" of TFIID is held together by a TAF6 homodimer, which connects lobe B and C. A portion of TAF8 (residues 130–235) also stabilizes the connection between lobes B and C, linking TAF2 and TAF6 (Patel et al. 2018). The TFIID BC core serves as a molecular ruler because lobe B interacts with upstream DNA (e.g., the TATA box), whereas the C lobe contains TAF1 and TAF2, which recognize sequence elements at the TSS (Inr) and downstream from the TSS (MTE, DPE). These TFIID–DNA interactions are outlined in Figure 2E and anchor TFIID to promoter DNA and help nucleate PIC assembly (Louder et al. 2016; Patel et al. 2018).

A working model for TFIID binding to promoter DNA, based on structural data from the Nogales group (Louder et al. 2016; Patel et al. 2018) and supported by other cellular and biochemical studies, is as follows: Step 1: TFIID lobe C (TAF1, TAF2, and TAF7) binds DNA downstream from the TSS. Step 2: Lobe A moves away from lobe C toward lobe B; lobe A contains TBP and therefore this is considered a means by which TBP can be delivered to its binding site ~30 bp upstream of the TSS. Lobe A also contains TAF11/13 and the N-terminal domain (TAND) of TAF1, which is flexibly tethered to structured domains of TAF1 in lobe C (Patel et al. 2018). Either TAF11/13 or the TAF1 TAND can bind the concave surface of TBP to block its binding to DNA (Liu et al. 1998; Gupta et al. 2017). Step 3: Upon structural rearrangement of lobe A, TBP is now positioned ~30 bp upstream of the TSS; at this location, it can bind TATA-containing or even TATA-less promoters. Because of the fixed distance of

Table 1. *Overview of human and yeast (*S. cerevisiae*) TFIID, TFIIH, and Mediator subunits*

Complex	Subunit	Lobe/module	Yeast	Human
TFIID	TAF1	Lobe C	1011	1872
	TAF2	Lobe C	1572	1199
	TAF3	Lobe A	315	929
	TAF4*	Lobes A and B	411	1085
	TAF5*	Lobes A and B	723	800
	TAF6*	Lobes A and B	486	677
	TAF7	Lobe C	461	349
	TAF8	Lobe B	518	310
	TAF9*	Lobes A and B	154	265
	TAF10*	Lobes A and B	218	218
	TAF11	Lobe A	202	211
	TAF12*	Lobes A and B	610	161
	TAF13	Lobe A	140	124
	Taf14	*Lobe C*	*248*	—
	TBP	Lobe A	244	339
TFIIH	XPB	Core	843	782
	XPD	Core	778	760
	p62	Core	642	548
	p52	Core	523	462
	p44	Core	461	395
	p34	Core	338	308
	p8	Core	72	71
	MAT1	Kinase module	321	309
	CCNH	Kinase module	393	323
	CDK7	Kinase module	306	346
Mediator	MED1	Core	566	1581
	MED2/29	Core	431	200
	MED3/27	Core	397	311
	MED4	Core	284	270
	MED5/24	Core	1132	989
	MED6	Core	295	246
	MED7	Core	222	233
	MED8	Core	223	268
	MED9	Core	149	146
	MED10	Core	157	135
	MED11	Core	115	117
	MED14	Core	1082	1454
	MED15	Core	1081	788
	MED16	Core	974	877
	MED17	Core	687	651
	MED18	Core	307	208
	MED19	Core	220	244
	MED20	Core	210	212
	MED21	Core	140	144
	MED22	Core	121	200
	MED23	**Core**	—	**1368**
	MED25	**Core**	—	**747**
	MED26	**Core**	—	**600**
	MED28	**Core**	—	**178**
	MED30	**Core**	—	**178**
	MED31	Core	127	131
	MED12	Kinase module	1427	2177
	MED12L	**Kinase module**	—	**2145**
	MED13	Kinase module	1420	2174
	MED13L	**Kinase module**	—	**2210**
	CDK8	Kinase module	555	464
	CDK19	**Kinase module**	—	**502**
	CCNC	Kinase module	323	283

The right columns, designated "yeast" and "human," indicate the number of residues in each protein. Starred TAFs are present in two copies in TFIID. Italics indicate yeast-specific subunits, whereas bold indicates human subunits not present in yeast.

the BC core, TAF1 and TAF2 binding to downstream sequences sets the location of TBP delivery to ~30 bp upstream. TFIIA–TBP binding will enable TBP to bind DNA by displacing the TAF1 TAND or TAF11/13 interactions with TBP. Perhaps as a consequence, TFIIA binding also favors this so-called "rearranged" TFIID structural state (Cianfrocco et al. 2013). Step 4: TBP binds upstream DNA, inserting Phe residues into the minor groove to bend at a 90° angle (Nikolov et al. 1992; Kim et al. 1993a,b). Step 5: TFIIB binds to TBP opposite TFIIA (Bleichenbacher et al. 2003), at a site formerly blocked by TAF11/13; TFIIB can then recruit Pol II–TFIIF. Step 6: Recruitment of Pol II–TFIIF to the promoter displaces the TAF4 contact with upstream DNA (Patel et al. 2018).

In addition to its ability to recognize common core promoter sequence motifs, human TFIID possesses tandem bromodomains within TAF1 (Jacobson et al. 2000) and a PHD finger domain in TAF3 (Vermeulen et al. 2007; van Ingen et al. 2008), which can bind acetylated or trimethylated histones, respectively. Each of these domains is flexibly tethered to TAF1 and TAF3 (Patel et al. 2018), perhaps to facilitate scanning of promoter-associated chromatin while TFIID is DNA-bound. The bromodomains of TAF1 and the PHD finger of TAF3 suggest that the genomic occupancy of TFIID and/or its function is regulated by chromatin marks. TAF3 binds H3K4me3 with high affinity (160 nM) (Vermeulen et al. 2007), suggesting that this mark could help recruit TFIID to gene promoters. The tandem bromodomains of TAF1 were shown to preferentially bind H4 acetylated peptides (e.g., H4K5ac/K12ac) with moderate affinity (1–5 µM) (Jacobson et al. 2000). Histone acetylation and H3K4me3 are each associated with active transcription and their ability to bind TFIID may contribute to this function. The histone-fold domains in the TAF subunits lack patches of positive charge that are present in histone proteins (Patel et al. 2018), thus within TFIID the histone folds serve as dimerization domains only and do not appear to bind nucleic acids.

In its canonical structural state, TFIID binding to promoter DNA would not be compatible with transcription initiation. Lobe C subunits could bind downstream elements, but TBP would be incapable of DNA binding due to inhibitory interactions with the TAF1 TAND and/or TAF11/13. However, even in the rearranged structural state, which deposits TBP at the appropriate upstream site, TFIID structure is not compatible with transcription initiation (Patel et al. 2018). For instance, TAF1 and TAF2 would clash with Pol II, and TAF4 would clash with the TFIIF WH domain based on cryo-EM structural data from partial PICs with yeast or human factors (He et al. 2013, 2016; Plaschka et al. 2016). Thus, TFIID structural rearrangements must occur during PIC assembly and transcription initiation, but these remain to be characterized. Moreover, because different genes possess different promoter sequence motifs and potentially distinct chromatin marks, TFIID structure and function may not be universal but instead could vary at different genomic loci. In support of this, partial TFIID complexes have been reconstituted that are stable (Bieniossek et al. 2013; Gupta et al. 2017), and partial TFIID assemblies have been isolated in human

cells (Deato and Tjian 2007; Trowitzsch et al. 2015; Antonova et al. 2018).

Although a complete understanding of TFIID structure and function is lacking, it is evident that human TFIID differs from yeast TFIID in fundamental ways. The yeast structure is more compact and it remains unclear whether it undergoes similar structural rearrangements compared with human TFIID (Kolesnikova et al. 2018); for instance, it is not evident that TFIIA and promoter DNA binding favors a specific structural transition in yeast TFIID, as observed with human TFIID (Cianfrocco et al. 2013). Other structural differences involve the yeast Taf8 and Taf6 subunits. In yeast (*K. phaffii*) TFIID, Taf8 interacts with the yeast-specific Taf14 protein and a C-terminal region of Taf2 that is not observed in human TFIID structures (Kolesnikova et al. 2018). As in humans, Taf6 forms a dimer in yeast TFIID; however, the dimerization interface is structurally distinct in that Taf6 forms a heterotetramer with Taf5, which contacts all three lobes in the more compact yeast TFIID structure. A more complete comparison of the human and yeast TFIID structures awaits further progress in resolving disordered regions within the complex. Analysis of TFIID sequences throughout evolution shows that yeast TFIID is somewhat of an outlier based on the fact that no chromatin "reader" domains (e.g., PHD or bromodomains) are represented (Antonova et al. 2019). However, yeast genomes possess auxiliary bromodomain-containing proteins that associate with TFIID and may replicate specific chromatin-targeting functions (Rhee and Pugh 2012). Yeast also lack discernable Inr elements or downstream promoter elements that are recognized by TAF1 and TAF2 in metazoans. In fact, compared with metazoan protein-coding genes, yeast promoters have an increased spacing of their TATA or TATA-like elements relative to the TSS. In *S. pombe*, initiation occurs 30–70 bp from the TATA, and 40–200 bp downstream from the TATA in *S. cerevisiae* (Yang and Ponticelli 2012). Such increased spacing could preclude yeast TFIID from binding DNA sequences downstream from the TSS, at least in some cases. Nevertheless, ChIP-exo data from *S. cerevisiae* show evidence that Taf1 occupies promoter regions downstream from the TSS, as observed in metazoans (Rhee and Pugh 2012). Although structural details of human TFIID have advanced dramatically in the past 20 yr (Andel et al. 1999), it is notable that many regions of the complex remain unresolved in the cryoEM maps, most likely due to structural disorder. These unresolved regions are typically poorly conserved in yeast TFIID (Patel et al. 2018).

The SAGA complex

Although not considered a PIC factor, the SAGA complex contains TAF subunits that are also present in TFIID. Moreover, SAGA plays key roles in the regulation of Pol II transcription, but its precise mechanistic contributions remain to be fully elucidated (Fischer et al. 2019; Donczew et al. 2020). SAGA contains 20 subunits (including TBP), organized into four modules: (1) the core module, which contains Tafs (Taf5, Taf6, Taf9, Taf10, and Taf12) as well as other subunits; (2) the Tra1 module, which contains only the large Tra1 protein (TRRAP in humans); (3) the HAT module, which contains the acetyltransferase Gcn5; and (4) the DUB module, which contains the deubiquitinase enzyme Ubp8. Recent cryoEM studies have uncovered new structural details of yeast SAGA complexes (Papai et al. 2020; Wang et al. 2020). Here, we briefly outline structural features that relate to TBP and the Taf subunits present in TFIID. Notably, SAGA can deliver TBP to promoter DNA. As in TFIID, TBP is inhibited from promiscuous DNA binding by the SAGA complex; however, the mechanisms are distinct and reflect its different subunit composition. For example, the concave DNA-binding surface of TBP is directed toward the SAGA core structure to block access to genomic DNA (Papai et al. 2020). In SAGA, Spt3 replaces the Taf11/13 dimer in its interaction with the TBP C terminus, whereas Spt8 occupies an N-terminal TBP site that can also interact with TFIIA (Papai et al. 2020; Wang et al. 2020). In fact, TFIIA appears to regulate TBP delivery to TATA-containing DNA sequences, perhaps through displacement of Spt8 (Papai et al. 2020). The shared Taf subunits are present in two copies in TFIID, but only one copy each for SAGA. This results in a distinct structural organization of the Taf subunits in SAGA. Most notable are Taf5 and Taf6, which form a heterodimer that differentially orients the seven-blade WD40 domain (Taf5) and the HEAT repeats of Taf6. This Taf5-Taf6 heterodimer serves as a major structural scaffold in SAGA, contacting about a dozen different protein domains within the complex (Papai et al. 2020; Wang et al. 2020).

Pol II and TFIIF

The Pol II enzyme has many functionally relevant domains that were first revealed at high resolution through X-ray crystallography (Cramer et al. 2000, 2001; Gnatt et al. 2001). We focus on some key Pol II structural domains here (Fig. 3A), and describe their basic role in transcription (Cheung and Cramer 2012).

The trigger loop and bridge helix The trigger loop moves between an open and closed state with each NTP addition cycle (Kaplan et al. 2008). The loop closes onto the incoming RNA base and helps detect base pair mismatches, thereby contributing to the fidelity of Pol II transcription (Kaplan et al. 2008; Gout et al. 2017). The bridge helix helps separate the DNA duplex at the active site and undergoes cooperative structural transitions with the trigger loop during nucleotide incorporation and translocation (Cramer et al. 2001; Brueckner and Cramer 2008).

The cleft, clamp, and stalk The cleft and stalk are the most prominent Pol II structural features. The cleft is a deep, positively charged groove along one face of the enzyme complex, and the clamp controls opening and closing of the cleft. Double-stranded promoter DNA resides above the cleft prior to template melting. Upon melting, the single-stranded template DNA can descend into the
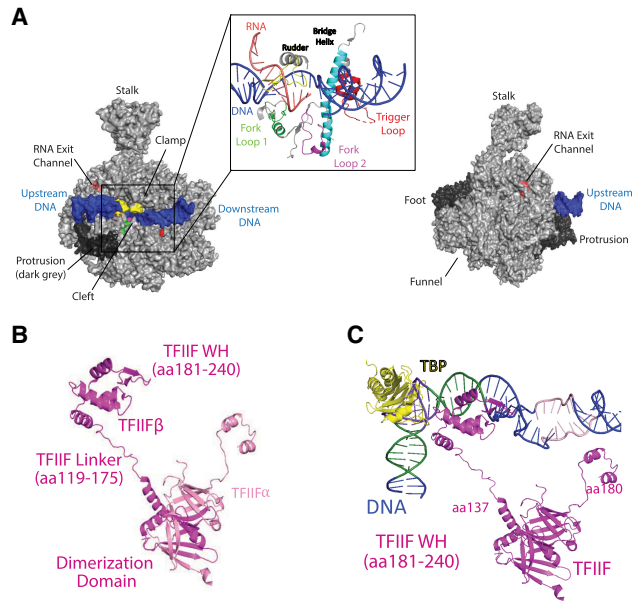
Figure 3. Pol II and TFIIF. (*A*) Bovine Pol II (gray) shown in two orientations (rotated 180°). DNA is colored in blue, and the Pol II stalk, clamp, foot, funnel, and RNA exit channel are marked; the protrusion and foot domains are shown in dark gray. The *inset* shows a zoomed-in view of the Pol II active site, with the trigger loop (Rpb1 amino acids 1095–1130) shown in red, bridge helix (Rpb1 amino acids 833–869) in cyan, rudder (Rpb1 amino acids 318–338) in yellow, fork loop 1 (Rpb2 amino acids 461–480) in green, and fork loop 2 (Rpb2 amino acids 499–520) in magenta (Cramer et al. 2001). Adapted from PDB 5FLM (Bernecky et al. 2016). (*B*) Structure of TFIIF in the PIC. RAP30 (TFIIFβ) is shown in light pink, and RAP74 (TFIIFα) is shown in light pink. (*C*) The same view of TFIIF is shown, with open promoter DNA and TBP added. The RAP30 winged helix (WH) domain (amino acids 181–240) interacts with the upstream DNA (−37 to −32) to help stabilize promoter DNA. The RAP30 linker (amino acids 119–175) interacts with TBP and may aid in positioning the WH domain. TBP residues 194 and 195 are implicated in the interaction, with RAP30 residues 172, 174, and 176. Adapted from PDB 5IY7 (He et al. 2016). Amino acid interactions were determined using the "find any contacts" function in PyMol, set to 4 Å. Amino acid residues listed correspond to bovine (panel A) or human proteins.

cleft to reach the active site, and this can occur with the clamp in an open (He et al. 2013) or closed state (Dienemann et al. 2019); if open, template melting is accompanied by closing of the clamp (He et al. 2013, 2016). The stalk extends from the foot domain at the base of the Pol II enzyme (Bushnell and Kornberg 2003). The stalk is contacted by various initiation and elongation factors and its movement helps coordinate opening and closing of the clamp.

The wall and protrusion The wall resides in the cleft, near the active site, and represents the site at which the RNA:DNA hybrid separates. At this point, upstream DNA makes a 90° turn to exit Pol II (Cramer et al. 2001). The RNA:DNA hybrid also separates at the wall.

The protrusion is an exterior, positively charged domain that is situated above the wall, at the site of DNA exit from the cleft. Reannealing of transcribed DNA occurs as it exits the enzyme, and the protrusion may participate in this process.

The funnel and RNA exit channel The funnel is the site of NTP entry and extends from the active site to the Pol II exterior, whereas the RNA exit channel initiates near the wall and directs RNA along the dock domain (Kettenberger et al. 2004), to exit adjacent to the RPB1 linker domain (Bernecky et al. 2016).

The rudder and lid These RPB1 domains reside near the active site and act to separate the RNA:DNA hybrid at the wall (Gnatt et al. 2001) and direct the RNA to its exit channel and the DNA out toward the protrusion.

The RPB1 C-terminal domain (CTD) Although not a structured domain, the Pol II CTD is central to the regulation of Pol II transcription (Harlen and Churchman 2017). The CTD sequence generally consists of heptad repeats of YSPTSPS; the yeast Pol II CTD contains 26 (*S. cerevisiae*) or 29 repeats (*S. pombe*) of this sequence, whereas 52 repeats are present in human RPB1. Prior to transcription initiation, the Pol II CTD likely helps recruit Mediator to the promoter (Kim et al. 1994) through direct, high-affinity interactions (Näär et al. 2002; Robinson et al. 2016). During transcription initiation, the CTD becomes phosphorylated by transcription-associated kinases, including CDK7 (TFIIH kinase) and CDK9 (P-TEFb kinase). Among other things, these phospho-marks help direct binding of RNA processing factors (e.g., capping enzymes, splicing and termination factors) as Pol II leaves the promoter-proximal region and transcribes through gene bodies. As a long, highly disordered sequence, the Pol II CTD also enables the formation of molecular condensates, another key aspect of transcription regulation (see below).

Within the core of the Pol II enzyme, sequence conservation is high between yeast and human complexes, which reflects identical mechanisms of RNA polymerization from a DNA template (Cramer et al. 2001). Sequences are more divergent toward the exterior/surface residues, which may reflect biochemically distinct interfaces with other factors.

TFIIF consists of two subunits (Fig. 3B) that form a dimerization interface (Gaiser et al. 2000) near the RPB2 lobe and the RPB9 jaw, first shown through cross-linking assays (Chen et al. 2010; Eichner et al. 2010; Mühlbacher et al. 2014) and later via cryoEM (He et al. 2013; Plaschka et al. 2015). The TFIIF subunits were originally given the gene names RNA polymerase-associated protein 74 and RNA polymerase-associated protein 30 (RAP74, RAP30; Tfg1, and Tfg2 in yeast). As PIC factors were being discovered through biochemical purification and in vitro transcription, the TFIIF subunits were considered components of Pol II itself (Sopta et al. 1985). TFIIF binds Pol II near the RPB2 lobe and protrusion domains, which reside along the Pol II cleft. Consistent with this location, TFIIF prevents Pol II from nonspecifically interacting with DNA

(Conaway et al. 1991) and initiating transcription at inappropriate sites. Within promoter-bound PICs, TFIIF helps orient and stabilize DNA both upstream of and downstream from the TSS (He et al. 2013). TFIIF may also promote PIC assembly through dephosphorylation of the Pol II CTD. The RAP74 subunit interacts with several CTD phosphatases (Friedl et al. 2003; Yeo et al. 2003) and structural details of RAP74-FCP1 have been obtained (Kamada et al. 2003; Nguyen et al. 2003). Pol II with an unphosphorylated CTD is required for interaction with the Mediator complex (Näär et al. 2002; Max et al. 2007; Robinson et al. 2012). Furthermore, TFIIF itself stabilizes the Pol II–Mediator interaction, although the structural basis remains unclear (Bernecky et al. 2011).

TFIIF binding opens the Pol II clamp and stabilizes double-stranded DNA above the Pol II cleft (He et al. 2013). This is done through protein–DNA interactions mediated by TFIIF and Pol II itself. Upon binding a promoter-bound Pol II–TBP–TFIIB–TFIIA complex, TFIIF induces structural changes that enable the RPB2 clamp head and a two-helix bundle in RPB5 to bind DNA upstream of and downstream from the TSS, respectively (He et al. 2016). Furthermore, in the PIC, the RAP30 WH domain contacts upstream DNA (the BREd element in the promoter used to assemble human PICs) (Fig. 3C) positioned above the RPB2 protrusion, which helps anchor the DNA along the cleft to facilitate promoter opening (He et al. 2013, 2016). TFIIF also forms an interface with TFIIA that completes a protein chain that extends from the Pol II stalk across the Pol II cleft and to TBP bound to the upstream TATA box sequence (TFIIB–TBP–TFIIA–TFIIF–TFIIE–RPB4/7 stalk). The protein bridge across the Pol II cleft, formed by TFIIE and TFIIF, traps the double-stranded DNA above the cleft (He et al. 2013, 2016; Plaschka et al. 2016; Schilbach et al. 2017). Collectively, this sets the stage for promoter melting and transcription initiation, which requires the coordinated functions of additional factors TFIIE and TFIIH (see below).

TFIIE and TFIIH

Biochemical studies have shown that TFIIE and TFIIH function cooperatively within the PIC (Goodrich and Tjian 1994; Ohkuma and Roeder 1994; Holstege et al. 1996), and cryoEM data has revealed multiple contact points between these complexes (Schilbach et al. 2017). TFIIE activates the translocase function of XPB and the kinase function of CDK7 in ways that remain incompletely understood (Ohkuma et al. 1995; Watanabe et al. 2003; Lin and Gralla 2005) but are likely to occur through TFIIE–TFIIH interactions that stabilize and properly orient TFIIH within the PIC.

CryoEM data from the Cramer laboratory (Plaschka et al. 2016; Schilbach et al. 2017) has revealed numerous interfaces between TFIIE and TFIIH in the context of the yeast PIC (S. cerevisiae), and similar interfaces have been observed (He et al. 2016) or modeled (Yan et al. 2019) in the human PIC. These include several interactions with the TFIIH p62 (Tfb1 in yeast) subunit, whose pleckstrin homology (PH) domain binds the E-bridge helix

(α8 in human IIEα) and whose BSD1 and three-helix bundle interacts with the E-floater helix (Tfa1 residues 351–373; α9 in human IIEα). The RING domain of TFIIH subunit MAT1 (residues 1–70 in yeast Tfb3; 1–65 in human MAT1) (Fig. 4A) also forms an interface with N-terminal linker helices in TFIIEα (yeast Tfa1); the TFIIH RING domain also interacts with the RPB7 subunit of Pol II stalk (He et al. 2016; Plaschka et al. 2016; Schilbach et al. 2017). One distinction between yeast and human TFIIE–TFIIH interactions within the PIC involves the E-dock (a partially disordered region between residues 200–250; yeast Tfa1) and the p62 PH domain. In the human PIC, the conserved E-dock helix (α7) instead interacts with p62 BSD2. Because the yeast PIC was assembled in the presence of a core Mediator complex (Schilbach et al. 2017), however, it remains possible that structural distinctions between yeast and human PICs may derive from different compositions of the PIC complexes being studied (He et al. 2013, 2016; Plaschka et al. 2016; Schilbach et al. 2017).

TFIIE has multiple anchor points to Pol II within the PIC. One is at the Pol II stalk (Fig. 4A), in which the N-terminal zinc ribbon (ZR) domain of TFIIEα interacts with RPB7 (He et al. 2016; Schilbach et al. 2017). Another contact point is with the clamp coiled-coil domain of RPB1, which interacts with a C-terminal extension of TFIIEβ WH2 domain. Starting from the Pol II stalk (i.e., the IIEα ZR-RPB7 interaction), a series of TFIIE winged helix (WH) domains forms a bridge over the Pol II cleft: IIEα WH, IIEβ WH2, and IIEβ WH1. This bridge is completed through an interaction with the WH domain from the RAP30 subunit of TFIIF, which traps duplex DNA above the cleft and stabilizes the PIC (He et al. 2013, 2016; Plaschka et al. 2016; Schilbach et al. 2017).

TFIIH is a 10-subunit complex (Rimel and Taatjes 2018) that is conformationally flexible and appears to undergo major structural rearrangements during PIC assembly (Nogales and Greber 2019). TFIIH consists of two modules: the core module and the kinase module (Table 1). The core adopts a circular structure in which p62 (yeast Tfb1) snakes through the complex and two enzymatic subunits, XPB and XPD, are connected at one end (Fig. 4B; Greber et al. 2017, 2019). Whereas the enzymatic function of XPD plays no role in transcription (but is important for DNA repair), its physical presence is important for stable association of the TFIIH kinase module, through the XPD arch domain (Abdulrahman et al. 2013). XPB is critical for transcription initiation. XPB functions as an ATP-dependent translocase (Fishburn et al. 2015) that acts to separate the DNA strands at the promoter to enable the single-stranded template to enter the Pol II active site (see below).

Recent reports have suggested that despite its importance for promoter opening genome-wide, the XPB subunit of TFIIH may not be absolutely required for transcription initiation, at least at some genes (Alekseev et al. 2017; Dienemann et al. 2019). These findings have been controversial, because earlier biochemical studies supported a requirement for TFIIH/XPB in promoter opening; however, these earlier studies were primarily completed in the absence of TFIID (TBP instead), TFIIA, and
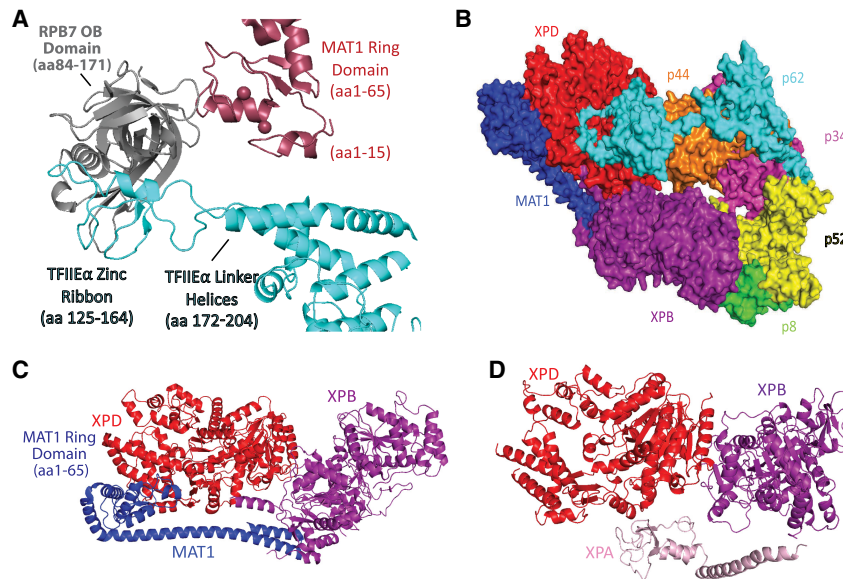
Figure 4. Structural details for TFIIE and TFIIH. (*A*) TFIIE and TFIIH converge at the Pol II stalk. The Ring domain in the TFIIH subunit MAT1 interacts with the OB domain of RPB7 and the TFIIE linker helices. Amino acids involved in the interactions are RPB7 164–168 and MAT1 40–45, 55, and 56. RPB7 also interacts with TFIIE, through residues 91–96, 105–107, 111, 151, 153, 158, and 160 (RPB7) and 124, 137–143, 145, 150–152, 161, and 162 (TFIIEα). PyMol was unable to detect any contacts within 4 Å between the IIE linker helix and the MAT1 ring domain in the human PIC, but they were identified in the yeast PIC (Schilbach et al. 2017). Adapted from PDB 5IY7 (He et al. 2016). (*B*) The structure of the free human TFIIH core complex, with MAT1. MAT1 helps link the two ATPase subunits XPB and XPD. MAT1 is in blue, XPB is in purple, XPD is in red, p8 is in green, p62 is in cyan, p34 is in magenta, and p44 is in orange. Not shown are the kinase module subunits CDK7 and CCNH. Adapted from PDB 6NMI (Greber et al. 2019). (*C*) Additional detail for MAT1 interactions with XPB and XPD. The XPD–MAT1 interaction involves over a dozen residues between 250 and 370 (XPD), plus residues 641 and 642, and about a dozen residues between 1 and 161 in MAT1. The XPB–MAT1 interaction involves residues 174, 183, 186, 189, 190, 195, 198, and 200–203 (XPB) and MAT1 residues 174, 177, 181, 184, 185, 188, and 194. Adapted from PDB 6NMI (Greber et al. 2019). (*D*) The DNA repair protein XPA displaces MAT1 at its XPB and XPD interfaces, and causes structural changes in XPD and XPB. XPA binding also allows rearrangement of core TFIIH subunits to fully engage the translocase and helicase functions of the complex (Kokic et al. 2019). Key residues involved in the interaction are 421, 422, 425, 714, 718, and 720 (XPB) and XPA residues 153, 157–159, 232, and 235–237; residues 634, 638, 641, 645, 647, and 648 (XPD) and XPA residues 164–166, 168, 174, 177, and 179. Adapted from PDB 6RO4 (Kokic et al. 2019). Amino acid interactions were determined using the "find any contacts" function in PyMol, set to 4 Å. Amino acids listed correspond to the human proteins.

Mediator. Even with partial PIC assemblies, however, it was evident that TFIIE and TFIIH were required for maximum Pol II activity in vitro (Holstege et al. 1995; Kumar et al. 1998). A compelling result was obtained through structural studies involving the biochemical reconstitution of promoter-bound *S. cerevisiae* PICs (Plaschka et al. 2016). While attempting to isolate "closed" PIC complexes that lacked TFIIH, the Cramer group (Plaschka et al. 2016) nevertheless observed a substantial population of complexes in the open state. Unlike past biochemical experiments, these PIC complexes contained Mediator, which we speculate may promote template opening in coordination with TFIIE. Although further studies will be needed to address this hypothesis, it is notable that human TFIIE has been shown to possess an ATP-dependent helicase function that could potentially contribute to template opening in the absence of TFIIH (Ayoubi et al. 2019).

The TFIIH XPD and XPB subunits serve as contact points for the MAT1 subunit (Fig. 4C), which connects the TFIIH core with the kinase module. Upon binding promoter DNA, XPB breaks its contact with XPD (Greber et al. 2017, 2019); thus, core TFIIH now adopts a horseshoe shape. This conformational shift does not impact the MAT1–XPD interaction (mediated through the MAT1 ARCH anchor domain) but may release the MAT1–XPB interaction. Release of the MAT1–XPB interaction would untether the kinase module because MAT1 connects the

TFIIH core to the CDK7/CCNH kinase/cyclin dimer (Luo et al. 2015). This untethering may be essential to reposition CDK7 for phosphorylation of the C-terminal domain (CTD) of the Pol II RPB1 subunit. Based on data from yeast PICs, this repositioning of the TFIIH kinase module may be triggered when the Mediator complex assembles into the PIC (Plaschka et al. 2016; Schilbach et al. 2017). Whether the MAT1–XPD interaction is also released during transcription remains unknown, but this would completely dissociate the kinase module from the TFIIH core. Incidentally, this dissociation must occur for the TFIIH core to function in DNA repair, and an exchange of MAT1 for the DNA repair factor XPA occurs along the same XPD–XPB interface (Fig. 4D; Kokic et al. 2019).

Mediator and sequence-specific DNA-binding transcription factors (TFs)

The Mediator complex was the last of the PIC factors to be discovered, initially in yeast through genetics (Thompson et al. 1993; Koleske and Young 1994) and biochemistry (Flanagan et al. 1991; Kim et al. 1994). The biochemical experiments included in vitro transcription assays with known PIC factors and partially purified extracts. A key aspect was that activated transcription—that is, increased transcriptional output in response to sequence-specific DNA-binding transcription factors (TFs; also known as activators)—could not be reconstituted with the so-called

general transcription factors alone. An additional activity was needed, called Mediator (Table 1), to enable activator-dependent transcription (Kim et al. 1994). The human complex was similarly discovered through biochemical assays by several laboratories (Fondell et al. 1996; Boyer et al. 1999; Näär et al. 1999; Rachez et al. 1999; Ryu et al. 1999).

Sequence-specific, DNA-binding TFs drive all biological processes (Lee and Young 2013) and they function in part by binding enhancer or promoter sequences and subsequently recruiting the PIC to specific genomic loci. TF binding to enhancers or promoters correlates with activation of Pol II transcription (Heinz et al. 2015; Haberle and Stark 2018); however, TFs do not bind directly to the Pol II enzyme. Instead, TFs communicate their activation signals to Pol II through the Mediator complex (Fig. 5), which interacts extensively with Pol II (Bernecky et al. 2011; Plaschka et al. 2015; Robinson et al. 2016; Schilbach et al. 2017). The correlation between TF-Mediator binding and Pol II activation implicates TF-Mediator interfaces as high-impact targets for molecular therapeutics. Among the few structurally characterized TF-Mediator interfaces (Yang et al. 2006; Milbradt et al. 2011; Currie et al. 2017), each displays high-affinity binding (150 nM or better) that results from TF interactions with hydrophobic pockets on Mediator subunits. Structured, hydrophobic pockets represent druggable targets that could be exploited for therapeutic purposes. Along these lines, Arthanari and coworkers (Nishikawa et al. 2016) identified a small molecule that blocked a Mediator–TF interaction in yeast (*C. glabrata*) and yielded a physiological outcome that mimicked inhibition of the TF itself. Furthermore, several laboratories have designed small molecules that mimic transcriptional activators, presumably through binding Mediator–TF interfaces (Rowe et al. 2007; Jung et al. 2009).

The mechanisms by which Mediator activates Pol II function remain unclear, but likely involve TF-induced structural changes in Mediator (Taatjes et al. 2002), which appear to remodel Mediator–Pol II interactions to promote initiation and/or promoter escape (Meyer et al. 2010). Such structural transitions may be highly dependent on the MED14 subunit (Cevher et al. 2014; Plaschka et al. 2015; Tsai et al. 2017), but the precise mechanisms remain unknown. A recent cryoEM analysis of Mediator isolated from murine B cells expands upon these concepts and has provided the highest resolution (5.9 Å) data for a metazoan Mediator complex to date (El Khattabi et al. 2019). As expected, the mouse Mediator complex showed some structural distinctions with yeast Mediator based on its larger size. For instance, the various mobile domains appeared to be more interconnected in the mouse Mediator complex, suggesting that potential conformational changes may require more extensive remodeling of protein–protein interfaces compared with yeast. Subunits comprising the tail segment (Med15, Med16, Med23–25, and Med27–30) were also shown to be more structurally integrated (vs. yeast) and appeared to be more interconnected with other Mediator structural domains (El Khattabi et al. 2019).
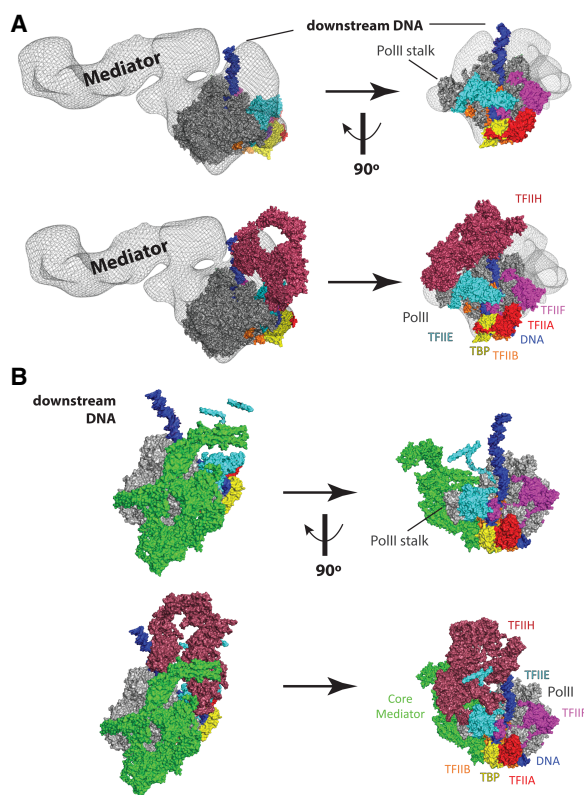
Figure 5. PIC structural models that include the Mediator complex. (*A*) Model of a partial human PIC that includes Mediator. Figure was prepared by rendering a Mediator–Pol II–TFIIF cryoEM density (Bernecky et al. 2011) in PyMol as a semiopaque black mesh, then visually aligned to a human PIC structure (He et al. 2016). Colors for each PIC factor are identical to Figure 1. The *top* row shows two views of the complex without TFIIH, and the *bottom* row shows the same views with TFIIH. Structural remodeling is likely upon binding TFIIH, as clashes are evident in this artificially docked model. The differences between yeast (see *B*) and human Mediator reflect the much larger size of the human Mediator complex (Table 1). However, the orientation of the human Mediator complex modeled in the human PIC is distinct from the yeast PIC. These differences could result from true differences in PIC structure (yeast vs. human) or could simply result from the fact that the human PIC model is not derived from a single complete structural assembly, as done for the yeast PIC (Schilbach et al. 2017). Adapted from PDB 5IY7 (He et al. 2016) and EMD-5343 (Bernecky et al. 2011). (*B*) Structure of a yeast PIC (*S. cerevisiae*), shown in identical orientations with *A*, based on alignment of Pol II. Here, a core Mediator complex is shown in green, whereas all other PIC factors are shown in the same colors as *A*. The *top* row shows two views of the complex without TFIIH, and the *bottom* row shows the same views with TFIIH. The different orientation of downstream DNA (vs. *A*) reflects potential structural differences between yeast and human PICs. Note, however, that *A* is a hypothetical model that merges two different structures, whereas *B* represents cryoEM data from a single structure (Schilbach et al. 2017). Adapted from PDB 5oqm (Schilbach et al. 2017).

Whereas structures of *S. pombe* (Larivière et al. 2012; Nozawa et al. 2017; Tsai et al. 2017) and *S. cerevisiae* (Imasaki et al. 2011; Tsai et al. 2014; Robinson et al. 2015)

Mediator complexes have been resolved to high or intermediate resolution, structural details of the entire yeast Mediator complex remain elusive, in part because of the highly flexible "tail" region, which consists of the Med2, Med3, Med5, Med15, and Med16 subunits. A review focused on Mediator structure and function, based largely on cryoEM structural data, was recently published (Harper and Taatjes 2018); below, we highlight new cryoEM results with a yeast complex associated with the PIC (Fig. 5B). The Cramer laboratory (Schilbach et al. 2017) was able to determine a PIC-cMED structure to 5.8 Å resolution (cMED = core Mediator, which contains 15 of the 21 Mediator subunits in *S. cerevisiae*) (Fig. 5B). This structure (Schilbach et al. 2017), which is ~2 MDa in size and contained 46 proteins, includes TFIIA, TFIIB, TBP, TFIIE, TFIIF, TFIIH, Pol II, and a 15-subunit core Mediator complex. This represents the most complete PIC structure to date and is only lacking TFIID and a subset of Mediator subunits. Comparison of the free cMED structure from *S. pombe* (Nozawa et al. 2017) with the cMED structure in the *S. cerevisiae* PIC revealed conformational changes in Mediator upon PIC association (Schilbach et al. 2017). This observation is consistent with lower-resolution cryoEM data with human Mediator (Fig. 5A; Bernecky et al. 2011), and likely results from the unusually high percentage of intrinsically disordered sequences in Mediator subunits (Tóth-Petróczy et al. 2008).

The PIC–cMED structure revealed numerous contact points between Mediator and the PIC (Plaschka et al. 2015; Schilbach et al. 2017). Specifically, the Med18 subunit contacts the B-ribbon of TFIIB and the Rpb1 dock domain, and Med20 contacts the Rpb3/11 dimer. Med18 and Med20 reside in the "movable jaw" of yeast Mediator (Larivière et al. 2012), and the Rpb1 dock and the Rpb3/11 dimer reside along the back of the Pol II enzyme, roughly opposite the entry site of downstream DNA. Additional Mediator–Pol II contacts involved Med8 and Med22, which contact Rpb4 (Pol II stalk), and a Med9 interaction with the foot domain of Rpb1, which is located at the base of the stalk. Med8 and Med22 occupy the spine and arm domains of yeast Mediator (Larivière et al. 2012), and the plank domain includes Med9 (Nozawa et al. 2017; Tsai et al. 2017). Cross-linking mass spectrometry (CXMS) data further revealed Med19 (a component of the hook domain) cross-links to the disordered CTD of Rpb1. Finally, cross-links were detected between Mediator and TFIIH subunits: Med7 to Rad3/XPD and Med6 to Tfb3/MAT1 (Schilbach et al. 2017). Med6 and Med7 occupy the shoulder and knob regions of Mediator, respectively (Nozawa et al. 2017; Tsai et al. 2017). Taken together, cryoEM and CXMS data from the *S. cerevisiae* PIC–cMED assembly revealed eight separate structural interfaces between Mediator and the PIC, with most involving Pol II but one with TFIIB and two with TFIIH (Schilbach et al. 2017). Previous studies also established interactions between the Pol II CTD and Med6, Med8, and Med17 in *S. cerevisiae* (Robinson et al. 2012). These results are consistent with established roles for Mediator in stabilizing PICs in yeast (Eyboulet et al. 2015).

The set of interactions between yeast Mediator, Pol II, and other PIC factors identified through cryoEM and CXMS studies are consistent with studies involving mammalian Mediator complexes. Because intermediate- to high-resolution structural data are lacking for mammalian Mediator-PIC assemblies, it remains to be determined whether specific molecular interfaces will be conserved.

Structural transitions during promoter opening and transcription initiation

Biochemical and single-molecule biophysics experiments revealed that yeast TFIIH subunit Ssl2 (human XPB) functions as a 5′-to-3′ translocase on duplex DNA (Fishburn et al. 2015). CryoEM data from yeast and human PICs indicate that XPB (yeast Ssl2) binds about 25–30 bp downstream from the TSS (He et al. 2016; Plaschka et al. 2016), with its two ATPase lobes on either side of the minor groove (Schilbach et al. 2017). Provided that the upstream DNA is anchored by the TBP–TFIIA–TFIIB–TFIIF assembly within the PIC, a 5′-to-3′ translocation on the nontemplate strand will reel the downstream DNA in the upstream direction, toward the Pol II active site (Fishburn et al. 2015). (Translocation in the 3′-to-5′ direction on the template strand has also been reported [Lin et al. 2005], and would yield the same result.) In this way, torsional strain will increase, ultimately melting the duplex DNA around the TSS (Fig. 1B). XPB must translocate ~12 bp to melt the template and position the TSS at the active site (He et al. 2016). At this point, the Pol II clamp closes slightly (He et al. 2016; Plaschka et al. 2016) and transcription initiation can occur.

This cooperative mechanism, involving PIC–DNA interactions both upstream of and downstream from the TSS, may have evolved to ensure that correct PIC assembly is a prerequisite for template opening, a first step in activation of Pol II transcription. A complete PIC will also be able to stabilize the so-called "open complex" by trapping the open template DNA to prevent its reannealing. The template DNA strand is stabilized by interactions with the TFIIB B-reader (Fig. 2A) and the rudder (RPB1), wall (RPB2), and fork loops (RPB2) near the Pol II active site (Fig. 3A), whereas the nontemplate strand is stabilized by the TFIIB B-linker and RPB2 fork loop 2 (Figs. 2A, 3A; Kostrewa et al. 2009; He et al. 2013, 2016; Plaschka et al. 2016; Schilbach et al. 2017). Domains within TFIIE and TFIIF also appear to stabilize the open complex. In the yeast PIC, the Tfa1 (human TFIIEα) eWH and E-wing domains interact with the upstream edge of the separated DNA strands (Plaschka et al. 2016). The "arm domain" of Tfg1 (human RAP74) also converges at this site and forms a β-strand with the RPB2 protrusion, which projects into the cleft (Plaschka et al. 2016). In the human PIC, the RAP74 arm domain can be disordered (He et al. 2016) but may similarly stabilize the separated DNA strands, perhaps through interaction with the TFIIB B-linker helix that may thread through the template and nontemplate DNA strands

(He et al. 2013). Collectively, these interactions between the PIC and promoter DNA require precise positioning at the time of promoter melting. The large size of the PIC likely provides the structural stability necessary to properly orient each of these domains in 3D space, along and within the Pol II cleft.

After transcription initiation, additional structural transitions must occur once the nascent RNA reaches a length of 12–13 nt. At this point, the RNA clashes with the TFIIB B-ribbon (He et al. 2016), the B-reader (Kostrewa et al. 2009), and the wall (RPB2) (Sainsbury et al. 2013). TFIIB also blocks the RNA exit channel; thus, release of TFIIB (or large-scale structural rearrangement) is required for further extension of the RNA. This structural rearrangement coincides with formation of a stable 8-bp RNA:DNA hybrid in the Pol II active site, as observed in the elongation complex (Bernecky et al. 2016). The structural transition involving TFIIB is considered the "promoter escape" stage of transcription initiation. Throughout the stages of open complex formation, transcription initiation, and until promoter escape, the upstream DNA remains stably engaged with TBP–TFIIA–TFIIB–TFIIF (He et al. 2016). The role of Mediator during these transitions remains poorly understood; however, Mediator binds the Pol II CTD with high affinity (Robinson et al. 2016), and Mediator–Pol II interactions are probably disrupted during promoter escape. The TFIIH-associated kinase CDK7 phosphorylates the CTD during transcription initiation and this is known to disrupt Mediator–CTD binding (Max et al. 2007). In fact, CDK7 (Kin28 in yeast) phosphorylation of the Pol II CTD is stimulated by human (Boeing et al. 2010; Meyer et al. 2010) or yeast Mediator (Kim et al. 1994) through unknown mechanisms.

In addition to TFIIB, TFIIF and TFIIE must dissociate after promoter escape to allow DSIF and NELF to bind the Pol II enzyme. DSIF and NELF help establish Pol II pausing, which is a common intermediate in metazoans and occurs ~20–60 bp downstream from the TSS. Pol II pausing appears to be important to (1) ensure proper 5'-capping of the nascent RNA (Rasmussen and Lis 1993; Tome et al. 2018), (2) prevent reinitiation of transcription by another Pol II enzyme (Gressel et al. 2017; Shao and Zeitlinger 2017), and (3) maintain the promoter in a nucleosome-free state (Gilchrist et al. 2010). Structural data from the Cramer laboratory (Vos et al. 2018b) has revealed that DSIF and NELF bind surfaces on the Pol II enzyme that are also bound by TFIIB, TFIIE, and TFIIF (Fig. 6A). Dissociation of TFIIB, TFIIE, and TFIIF therefore ensures that DSIF and NELF binding is temporally regulated and coincides with their distinct roles in transcription: PIC factors TFIIB, TFIIE, and TFIIF for initiation through promoter escape and DSIF and NELF for promoter-proximal pausing and elongation. DSIF and NELF are considered Pol II elongation factors and are described further below.

Pol II elongation factors

Many of the factors described below directly bind the Pol II enzyme but are not components of the PIC. These so-called elongation factors do not represent an exhaustive list of proteins and protein complexes that control Pol II transcription, but each relates back to PIC factors in various ways. Moreover, the factors described below are implicated in early stages of transcription (i.e., toward gene 5' ends), although some also regulate downstream events.

TFIIS

The structure of TFIIS (Dst1 in *S. cerevisiae*) bound to yeast Pol II was determined by X-ray crystallography (Kettenberger et al. 2003) and is shown in Figure 6B. TFIIS contains three domains (I, II, and III), and domain II (residues 148–238 in *S. cerevisiae* TFIIS/Dst1) binds the surface of Pol II at the Rpb1 jaw domain. An interdomain linker (residues 239–264) then adopts an α-helical structure and extends into the Pol II funnel. This positions TFIIS domain III (residues 265–309) at the active site, where two highly conserved acidic residues (D290 and E291) are positioned to promote nucleophilic cleavage of a backtracked RNA substrate. Nucleophilic cleavage of the RNA is likely mediated by an activated water molecule, with the acidic residues involved in positioning metal ions to promote cleavage of the phosphodiester bond. This cleavage results in a new RNA 3' end and an open site for an incoming NTP to hybridize with the DNA template. Importantly, TFIIS insertion into the funnel and active site maintains space for NTP entry (Kettenberger et al. 2003).

Although TFIIS can be considered a PIC factor (Kim et al. 2007), it appears to be most important after promoter escape (Adelman et al. 2005; Sigurdsson et al. 2010), within 2 kb of the TSS, and at gene 3' ends (Sheridan et al. 2019). If Pol II pauses and backtracks, the active site lacks an unhybridized DNA template base. Furthermore, RNA backtracks into the funnel, and elongation is blocked. Such backtracked Pol II enzymes represent stably paused intermediates and may require polyubiquitination and degradation to remove from the DNA template (Sigurdsson et al. 2010). Thus, TFIIS acts as an "antipausing factor" because of its ability to stimulate cleavage of backtracked RNAs. Although Pol II enzymes have an intrinsic ability to cleave backtracked RNAs, this activity is enhanced by TFIIS and complete loss of this intrinsic activity is lethal in yeast (Sigurdsson et al. 2010).

DSIF and NELF

DSIF is a dimer consisting of a small SPT4 subunit and a larger SPT5 subunit. SPT5 contains an NGN domain, which is conserved even in bacterial genomes (Werner 2012), and four (yeast) or six (humans) Kyrpides, Ouzounis, Woese (KOW) domains (Kyrpides et al. 1996). NELF consists of four subunits: NELFA, NELFB, NELFC/D, and NELFE. NELFC and NELFD are nearly identical and associate with NELF in a mutually exclusive fashion. Whereas DSIF is conserved in yeast, NELF is absent from yeast genomes. DSIF and NELF bind the Pol II enzyme at sites that overlap with PIC factors TFIIB, TFIIE, and TFIIF (Fig. 6A). Based on in vitro (Missra and Gilmour 2010; Li et al. 2013) and cellular data, DSIF and NELF
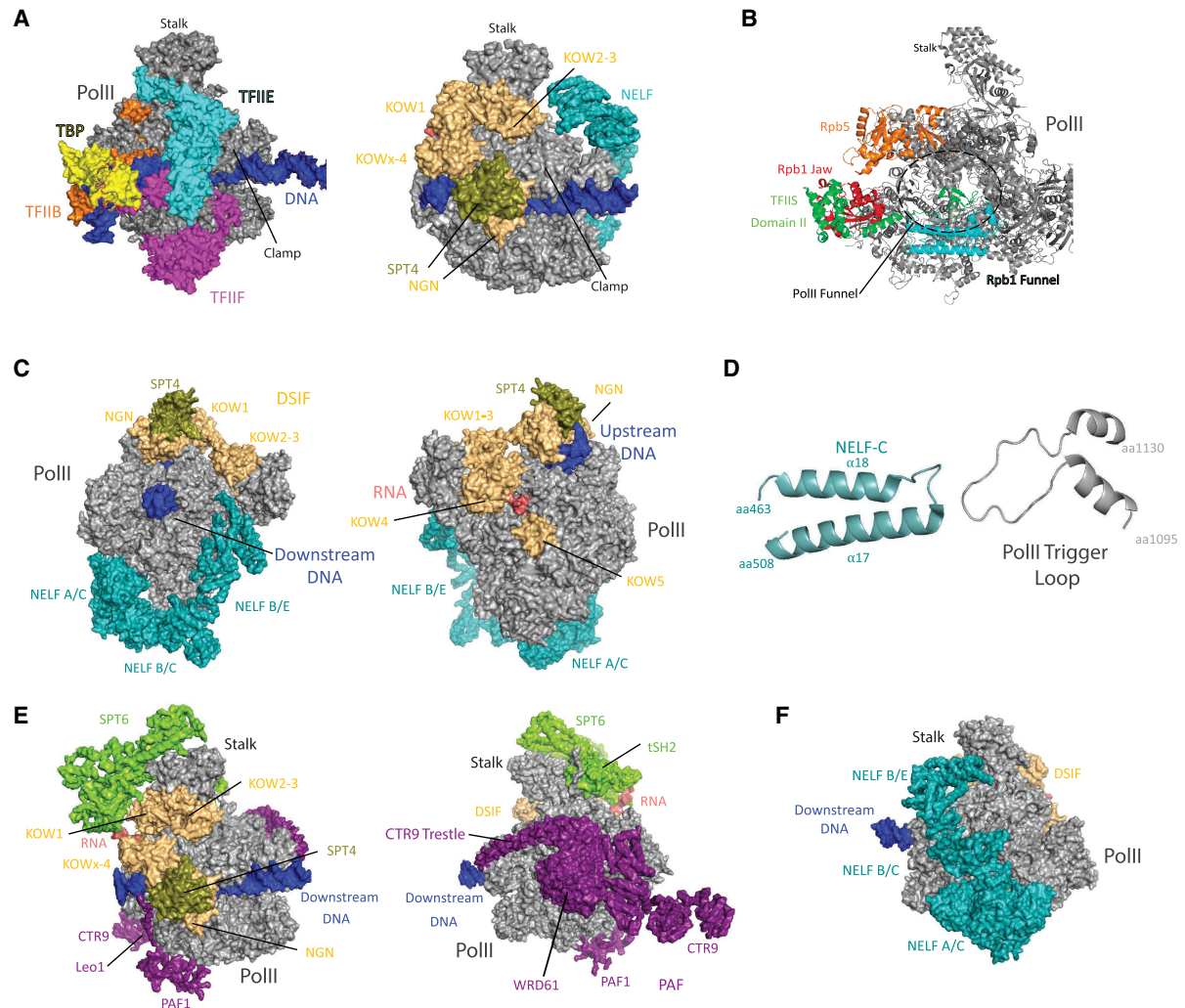
Figure 6. Pol II elongation complexes bound to pausing (NELF) or elongation factors (DSIF, PAF, SPT6). (*A*) Structures of partial PICs that emphasize how DSIF binds Pol II surfaces occupied by TFIIB, TFIIE, and TFIIF. At *left* is a Pol II structure with TBP, TFIIB, TFIIE, and TFIIF. At *right* is a Pol II structure bound to DSIF and NELF. Adapted from PDB 6GML (Vos et al. 2018b) and PDB 5IY7 (He et al. 2016). (*B*) Structure of TFIIS bound to Pol II. Pol II is shown in gray and TFIIS is shown in green; the Rpb1 jaw is shown in red, the Rpb1 funnel is shown in cyan, and Rpb5 is shown in orange. TFIIS amino acids 230–301 extend into the Pol II funnel, which positions TFIIS residues D290 and E291 near a catalytic zinc ion, which helps catalyze cleavage of backtracked RNA. Adapted from PDB 5IY7 (He et al. 2016). (*C*) Structure of NELF and DSIF bound to a transcribing/paused Pol II. DNA is shown in blue, nascent RNA is shown in salmon, NELF is shown in teal, SPT5 is shown in yellow, and SPT4 is shown in olive. Select NELF/DSIF domains or subunits are shown, with two views rotated 180°. Adapted from PDB 6GML (Vos et al. 2018b). (*D*) Detail from the DSIF/NELF–Pol II structure, showing the interaction between NELFC and the Pol II trigger loop (RPB1 amino acids 1095–1130). Adapted from PDB 6GML (Vos et al. 2018b). (*E*) Two views (rotated 180°) of the PAF complex, SPT6, and DSIF bound to Pol II. Pol II is shown in gray, DSIF is shown in yellow, SPT4 is shown in olive, DNA is shown in blue, RNA is shown in salmon, PAF is shown in purple, and SPT6 is shown in neon green. Adapted from PDB 6GMH (Vos et al. 2018a). (*F*) NELF–Pol II binding is mutually exclusive with PAF–Pol II binding. The NELF–DSIF–Pol II structure is shown in the same orientation as the right-hand image in *E*. The WRD61 and CTR9 subunits of PAF directly clash with NELFA/C, NELFB/C, and NELFB/E. Adapted from PDB 6GML (Vos et al. 2018b).

associate with Pol II at promoters and downstream from the TSS (Rahl et al. 2010). This location coincides with promoter-proximal pause sites, and DSIF and NELF have been shown to promote Pol II pausing (Core and Adelman 2019).

The structure of an elongating mammalian Pol II enzyme bound to human NELF and DSIF (Fig. 6C) showed how NELF binds Pol II and negatively regulates transcrip-

tion elongation (Vos et al. 2018b). The NELFA/C dimer binds at the RPB1 funnel, which could restrict NTPs from entering the active site. A different region of the NELFA/C dimer also contacts the trigger loop in its open state, which could prevent RNA chain extension by restricting translocation (Fig. 6D). Furthermore, through comparison with cryoEM data from DSIF–Pol II complexes (Bernecky et al. 2017), it was apparent that

NELF binding stabilizes a nonproductive intermediate in which the RNA:DNA hybrid in the active site is tilted by 15° (Vos et al. 2018b). This modest structural change would nevertheless prevent any incoming NTP from hybridizing with the DNA template. The cryoEM structure also showed that NELF binding blocks the TFIIS–Pol II interaction site, and biochemical experiments confirmed a mutually exclusive Pol II association for NELF or TFIIS (Vos et al. 2018b). By preventing TFIIS binding, NELF also favors Pol II pausing and backtracking, which may extend the lifetime of paused Pol II complexes. It appears that the CDK9 kinase, as part of the P-TEFb complex, removes NELF by phosphorylation (see below). DSIF is also phosphorylated by P-TEFb but DSIF remains associated with the Pol II enzyme to promote elongation.

CryoEM structures show that DSIF associates across the surface of Pol II, which reflects its many distinct structured domains that are joined through flexible tethers (Fig. 6C). The SPT4 subunit and the SPT5 NGN domain form a bridge over the Pol II cleft (Bernecky et al. 2017; Ehara et al. 2017), which likely stabilizes the elongating complex and promotes Pol II processivity. This is supported by cell-based studies that suggest reduced capacity to transcribe long genes if DSIF function is disrupted (Shetty et al. 2017; Fitz et al. 2018). The DSIF bridge over the Pol II cleft contacts the RPB1 clamp helices on one side and the RPB2 protrusion on the other. This site is also adjacent to the upstream DNA exit and therefore may promote reannealing of the DNA template behind the transcribing polymerase (Bernecky et al. 2017; Vos et al. 2018b). This could be important to prevent formation of R-loops, which could otherwise disrupt transcription and contribute to genome instability (Sollier et al. 2014). DSIF also binds around the RNA exit channel and we speculate that this could help regulate RNA folding or cotranscriptional RNA processing events.

PAF and SPT6

The human PAF complex contains five subunits (CTR9, LEO1, PAF1, CDC73, and WDR61) and is about 400 kDa in size. PAF binds surfaces on Pol II that are shared with NELF (Fig. 6E), and NELF dissociation is required for PAF binding (Vos et al. 2018a). NELF dissociation is triggered by P-TEFb phosphorylation (see below). The PAF1 and LEO1 subunits form a dimer in the PAF complex and help anchor the complex to the so-called external domains of RPB2. The CTR9 subunit contains several interesting structural features. A set of 19 tetratricopeptide repeats extends from RPB11 to RPB8 to the RPB1 funnel and foot. A 100-Å-long α-helix (called the trestle; residues 807–892) then extends from the foot to RPB5, which is located at the site of downstream DNA entry into the Pol II cleft. An additional ∼300 amino acids are disordered and C-terminal to the trestle (Vos et al. 2018a). These may play important roles in promoting transcription through chromatin, consistent with the function for the PAF complex (Pavri et al. 2006; Hou et al. 2019).

Human SPT6 is about 200 kDa in size and binds around the Pol II stalk (RPB4/7), with direct interactions to RPB7

(Vos et al. 2018a). Another key SPT6–Pol II interaction involves the SPT6 tSH2 domain (Sun et al. 2010) with the linker region of RPB1. The tSH2 domain (residues 1328–1516) is required for SPT6 association with the elongating Pol II complex and its interaction requires phosphorylation of the Pol II RPB1 linker domain (Sdano et al. 2017; Vos et al. 2018a), which can be deposited by the P-TEFb kinase (see below). The SPT6 core (residues 284–1287) also interacts with DSIF through its SPT5 KOWx–KOW4 and KOW1 domains (Fig. 6E), as revealed by cryoEM and CXMS data (Vos et al. 2018a). An N-terminal region of SPT6 (residues 1–284) was disordered in the complex but is positioned such that it could potentially interact with nucleosomes in front of or behind the transcribing polymerase, consistent with its biological roles in RNA processing and as a histone chaperone (Bortvin and Winston 1996; Kaplan et al. 2003; Yoh et al. 2008; Dronamraju et al. 2018; Jeronimo et al. 2019).

Comparison of cryoEM structures representing different Pol II functional intermediates allowed the Cramer group (Vos et al. 2018a) to identify allosteric mechanisms that enable PAF, DSIF, and SPT6 to enhance the rate of Pol II elongation. Compared with the DSIF–Pol II complex (Bernecky et al. 2017), the addition of PAF and SPT6 repositioned the Pol II stalk and opened the RNA clamp formed by SPT5 (Vos et al. 2018a). Specifically, the KOW2–KOW3 interaction with KOW1 was disrupted and the KOWx–KOW4 domain rotated ∼50° and moved away from the exiting RNA. Furthermore, the SPT5 KOW1 domain rotated and moved away from the upstream DNA. This coincided with a movement of upstream DNA (i.e., DNA exiting behind elongating Pol II) away from the RPB2 protrusion and insertion of the C-terminal extension from LEO1 (residues 503–529) (Vos et al. 2018a). Collectively, these structural transitions may increase the rate of Pol II elongation by facilitating RNA exit and promoting DNA reannealing behind the transcribing polymerase.

P-TEFb

P-TEFb consists of a kinase/cyclin pair: CDK9 and CCNT1 (or CCNT2). Crystal structures of the complex have been determined (Baumli et al. 2008; Tahirov et al. 2010), which reveal an organization and interfaces common among CDK:Cyclin pairs but also distinct features. Substrates for the CDK9 kinase have been identified from proteomics experiments in human cell extracts (Sansó et al. 2016) or from analog-sensitive cell lines (Decker et al. 2019). These studies revealed dozens of high-confidence targets, with many representing transcription cofactors or RNA processing factors. CDK9 can also function as part of the larger super elongation complex (SEC) (Luo et al. 2012), which contains other proteins known to regulate transcription elongation.

Through its CDK9 kinase, P-TEFb may control the activity of many transcription regulatory factors. Using biochemical assays and cryoEM, the Cramer laboratory (Vos et al. 2018b) recently demonstrated that P-TEFb can (1) phosphorylate the NELFA tentacle domain, which may

contribute to release of NELF from Pol II; (2) phosphorylate an SPT5 linker that may help open the RNA clamp to promote transcription elongation, and (3) phosphorylate the RPB1 linker to enable binding of SPT6 through its tSH2 domain (Sdano et al. 2017). P-TEFb also phosphorylates the Pol II RPB1 CTD, which promotes CTD association with various RNA processing factors and chromatin modifying complexes (Kizer et al. 2005; Lee and Skalnik 2008; Ebmeier et al. 2017). RNA processing (e.g., splicing, cleavage, polyadenylation) and chromatin modification represent additional levels of transcription regulation that are reviewed elsewhere (Venkatesh and Workman 2015; Saldi et al. 2016; Herzel et al. 2017).

Mediator kinase module

The Mediator kinase module is a large complex (600 kDa; Table 1) that contains four subunits (Fant and Taatjes 2019). In the yeast *S. cerevisiae*, the kinase module consists of Srb8-11, which are orthologs of the human genes MED12, MED13, CDK8, and CCNC. Low-resolution cryoEM structures of the entire yeast and human kinase modules have been determined (Knuesel et al. 2009a; Tsai et al. 2013), but high-resolution data exist only for yeast Srb11 (Hoeppner et al. 2005) and the human CDK8: CCNC dimer (Schneider et al. 2011). Interestingly, paralogs of CDK8, MED12, and MED13 exist in mammalian genomes as CDK19, MED12L, and MED13L, respectively. No structural data exist for these paralogs.

The MED12 subunit has been shown to be important for activation of CDK8 kinase function, along with CCNC (Knuesel et al. 2009b). Biochemical studies suggested that the N terminus of MED12, which is a hot spot for oncogenic mutations (Makinen et al. 2011; Lim et al. 2014), interacted with CCNC as part of the MED12-dependent activation mechanism (Turunen et al. 2014). However, these findings were contradicted by CXMS data obtained with partial assemblies of the human CDK8 module (Klatt et al. 2020). The CXMS results supported an interaction between the MED12 N terminus (residues 30–42) and a disordered CDK8 activation loop (residues 173–203). Similar activation loops (also known as T-loops) exist in other CDKs, but the CDK8 sequence contains a D instead of a typical T at residue 191, suggesting a phosphorylation-independent activation mechanism. The MED12 N terminus (residues 19–50) is predicted to adopt an α-helical structure, and a model was proposed in which negatively charged residues (e.g., E33) in this MED12 "activation helix" mimicked CDK8 activation loop phosphorylation as a means to activate CDK8 (Klatt et al. 2020). The structural discrepancies between these models for MED12-dependent activation of CDK8 could reflect two distinct mechanisms of activation. However, each study tested different sets of mutations and/or evaluated incomplete assemblies of the four-subunit, 600-kDa CDK8 module. Future experiments will benefit from analysis of complete CDK8 module assemblies or evaluation with knock-in cell lines that ensure expression of mutant subunits at physiologically relevant levels.

The kinase module reversibly associates with Mediator through its MED13 subunit (Knuesel et al. 2009a; Tsai et al. 2013), but structural details about the key interfaces remain unclear. Although Mediator binds Pol II within the PIC (Bernecky et al. 2011; Plaschka et al. 2015; Robinson et al. 2016; Schilbach et al. 2017), Mediator association with the kinase module prevents this interaction (Elmlund et al. 2006; Knuesel et al. 2009a; Ebmeier and Taatjes 2010). These results suggest that the Mediator kinase module may function at postinitiation stages of Pol II transcription, and this is supported by cellular and biochemical data (Donner et al. 2010; Galbraith et al. 2013; Steinparzer et al. 2019). Furthermore, SILAC-MS experiments reveal several postinitiation regulatory factors as high-confidence Mediator kinase substrates (Poss et al. 2016), including NELFA. In support, inhibition of CDK8 kinase activity increases Pol II promoter-proximal pausing in mouse and human cells (Steinparzer et al. 2019). In addition to their enzymatic functions, it appears that CDK8 and its paralog CDK19 have key structural/scaffolding roles in mammals, although molecular details remain unclear. Gene expression changes are markedly different upon CDK8 kinase inhibition compared with subunit knockdown (Poss et al. 2016), and CDK19 in particular appears to function as a structural scaffold, whereas its kinase activity is less consequential (Audetat et al. 2017; Steinparzer et al. 2019).

Concluding remarks and future directions

Structural data from cryoEM and complementary methods, combined with functional studies, has advanced our understanding of Pol II transcription. Twenty years ago, the set of PIC factors had been identified but the basic structural architecture of the PIC was not known. At present, we have a greatly improved understanding of PIC structure, function, and dynamics. Many regulatory interfaces have been identified at high resolution, which allows rational design of molecular probes for mechanistic studies or as lead compounds for molecular therapeutics. However, key details remain to be uncovered and we focus on a number of outstanding questions below.

Among the PIC factors, TFIID and Mediator remain the most enigmatic in terms of their structure and function within the PIC; many questions remain regarding TFIIH as well. The size (Table 1) and flexibility of these factors contributes to the difficulties in understanding their structural and functional roles. For instance, the Nogales laboratory (Greber et al. 2019) has shown that TFIIH adopts a ring-like structure that is broken upon binding promoter DNA (XPB–XPD contacts are disrupted). Furthermore, the kinase associated with TFIIH, CDK7, is part of a kinase module that has the potential to become highly mobile (Yan et al. 2019). CryoEM data of the yeast PIC–cMED complex revealed that the kinase module (Kin28, Ccl1, and Tfb3) moves away from the TFIIH core toward the periphery of the PIC, between the hook, knob, and shoulder of the Mediator complex (Schilbach et al. 2017). This positions the kinase module more

Schier and Taatjes

favorably for Pol II CTD phosphorylation, but it remains unclear how this repositioning is triggered or whether the TFIIH kinase module remains anchored at this site during transcription.

It is also unclear how transcription reinitiation is regulated by the PIC. In vitro, reinitiation has been shown to occur more rapidly compared with de novo initiation (Hawley and Roeder 1987; Jiang and Gralla 1993). To reinitiate transcription, a second Pol II enzyme must engage the promoter at the TSS, and this may be facilitated by a PIC scaffold complex that remains following Pol II promoter escape (Yudkovsky et al. 2000). It is not established whether TFIID lobe C subunits and/or XPB are required to rebind their downstream DNA sequences for reinitiation to occur. Potentially, TAF1, TAF2, and XPB bind downstream DNA only to initiate a pioneering round of transcription, and their dissociation could facilitate reinitiation. It was proposed by Nogales et al. (Patel et al. 2018) that TFIID binding to downstream promoter elements may be important to accurately position TBP upstream of the TSS, but the TAF subunits of TFIID may dissociate after TBP deposition. We speculate that another regulatory purpose for promoter-proximal Pol II pausing could be to prevent TFIID lobe C or TFIIH XPB interactions with downstream sequences, which may otherwise be inhibitory to transcription reinitiation. Conceptually, this is similar to known roles for paused Pol II in the maintenance of nucleosome-free regions at active gene promoters (Gilchrist et al. 2010). In cells, a phenomenon called transcriptional bursting has been described (Fukaya et al. 2016; Tantale et al. 2016), based on data from live-cell imaging experiments. Transcriptional bursting appears to involve rapid reinitiation by multiple Pol II enzymes, but the molecular mechanisms remain unknown (Lenstra et al. 2016).

Throughout the stages of PIC assembly, initiation, promoter escape, pausing, and elongation, numerous factors compete for the same binding surfaces on Pol II. Consequently, regulation of these interactions is paramount. Although the mutually exclusive binding of initiation versus elongation factors provides a biological rationale for their exchange during different transcriptional stages, Pol II transcription will not be efficient if initiation, pausing, or elongation factors continually compete for Pol II binding. Precisely how these interactions are controlled remains incompletely understood. One possibility is simply through phase separation (Cramer 2019); the Pol II CTD itself can undergo liquid phase separation (Kwon et al. 2013) and Pol II CTD condensates may possess altered biophysical properties based on the CTD phosphorylation state (Boehning et al. 2018). At the TSS, the CTD is primarily unphosphorylated, whereas it becomes highly phosphorylated within gene bodies. Consistent with this, CTD phosphorylation promotes formation of condensates that exclude Mediator (Guo et al. 2019) but instead incorporate elongation factors such as P-TEFb (Lu et al. 2018) or splicing components (Guo et al. 2019). Another means of regulation is through posttranslational modifications, such as phosphorylation by transcription-associated kinases. Although this is complicated by the array of kinases and phosphatases that can converge at sites of active transcription, it is well-established that phosphorylation can increase or decrease protein-protein or protein-nucleic acid binding affinities (Pufall et al. 2005; Lee et al. 2010; Mylona et al. 2016). Collectively, dynamic and reversible modification of proteins through posttranslational modifications or segregation of initiation versus elongation factors into biophysically distinct molecular condensates could help ensure that initiation, pausing, and elongation factors interact with Pol II at the appropriate stages of transcription. Note that these potential regulatory mechanisms are not mutually exclusive and may function cooperatively throughout transcription initiation, elongation, and termination.

Much remains to be discovered about the structural transitions that the PIC undergoes during transcription initiation. For instance, sequence-specific DNA-binding TFs can activate Pol II transcription, but the molecular mechanisms remain incompletely understood. TF–Mediator binding induces structural changes that correlate with activation of Pol II transcription, perhaps by remodeling Mediator–Pol II interactions (Meyer et al. 2010; Tsai et al. 2014, 2017). Potentially, these TF-induced structural changes could contribute to transcriptional bursting, given that transient TF–DNA binding has been shown to correlate with bursting (Mir et al. 2018; Donovan et al. 2019; Stavreva et al. 2019). Details about the molecular mechanisms await higher-resolution information for the Mediator–Pol II structural transitions that result in transcription activation. In addition, both TFIID (lobe C subunits TAF1 and TAF2) and TFIIH (XPB subunit) bind DNA downstream from the TSS and must dissociate to allow Pol II to transcribe through the promoter-proximal region (Schilbach et al. 2017; Patel et al. 2018). In fact, TFIID binding to the promoter appears to be inhibitory to Pol II–DNA binding and transcription initiation (Patel et al. 2018). Evidence for TFIID conformational changes during transcription initiation have been obtained through biochemical studies (Yakovchuk et al. 2010; Zhang et al. 2015), but structural details remain to be determined. Similarly, several distinct Mediator–Pol II structural intermediates are likely to have functional relevance during Pol II initiation and promoter escape, based on the extensive interaction between these complexes and the demonstrated conformational flexibility (Bernecky et al. 2011; Bernecky and Taatjes 2012; Schilbach et al. 2017; Tsai et al. 2017; El Khattabi et al. 2019). CryoEM is suited to address these challenges and multiple functionally distinct intermediates could be characterized through a combination of biochemical and computational approaches.

Finally, a detailed mechanistic understanding of Pol II transcription will require characterization of PIC dynamics in real time. In vitro single molecule studies (Tomko and Galburt 2019) can augment structural data to better define how Pol II and associated regulatory factors work together to transcribe from a DNA template. Furthermore, advances in live cell imaging will complement the continually improving structural and mechanistic models of Pol II transcription. Despite recent progress (Liu and Tjian 2018), many basic questions remain

unanswered, such as (1) how genomes are organized in the three-dimensional space of the nucleus (Furlong and Levine 2018), (2) how transcriptional bursting occurs (Donovan et al. 2019; Rodriguez et al. 2019; Stavreva et al. 2019), (3) how enhancer–promoter interactions are controlled (Lim et al. 2018; Li et al. 2019), (4) how enhancers actually work to activate gene expression (Benabdallah et al. 2019; Heist et al. 2019), (5) how gene expression patterns are maintained (i.e., active vs. repressed) through mitosis (Teves et al. 2018), and so on. Also, what set of cofactors are essential for these processes, and which are redundant, context-specific, or cell type-specific? Addressing these questions will be important but challenging, especially in mammals, which have larger genomes, more elaborate enhancer–promoter regulatory networks (Levine et al. 2014), more potential regulatory inputs (including noncoding RNAs), and a greater diversity of cell types. Fortunately, given the technological and methodological advances over the past few decades, we have reached a point at which most experimental questions can be rigorously addressed.

Acknowledgments

References

Abdulrahman W, Iltis I, Radu L, Braun C, Maglott-Roth A, Giraudon C, Egly JM, Poterszman A. 2013. ARCH domain of XPD, an anchoring platform for CAK that conditions TFIIH DNA repair and transcription activities. *Proc Natl Acad Sci* **110:** E633–E642. doi:10.1073/pnas.1213981110

Adelman K, Marr MT, Werner J, Saunders A, Ni Z, Andrulis ED, Lis JT. 2005. Efficient release from promoter-proximal stall sites requires transcript cleavage factor TFIIS. *Mol Cell* **17:** 103–112. doi:10.1016/j.molcel.2004.11.028

Alekseev S, Nagy Z, Sandoz J, Weiss A, Egly JM, Le May N, Coin F. 2017. Transcription without XPB establishes a unified helicase-independent mechanism of promoter opening in eukaryotic gene expression. *Mol Cell* **65:** 504–514.e4. doi:10.1016/j.molcel.2017.01.012

Anandapadamanaban M, Andresen C, Helander S, Ohyama Y, Siponen MI, Lundström P, Kokubo T, Ikura M, Moche M, Sunnerhagen M. 2013. High-resolution structure of TBP with TAF1 reveals anchoring patterns in transcriptional regulation. *Nat Struct Mol Biol* **20:** 1008–1014. doi:10.1038/nsmb .2611

Andel F 3rd, Ladurner AG, Inouye C, Tjian R, Nogales E. 1999. Three-dimensional structure of the human TFIID–IIA–IIB complex. *Science* **286:** 2153–2156. doi:10.1126/science.286 .5447.2153

Antonova SV, Haffke M, Corradini E, Mikuciunas M, Low TY, Signor L, van Es RM, Gupta K, Scheer E, Vos HR, et al. 2018. Chaperonin CCT checkpoint function in basal transcription factor TFIID assembly. *Nat Struct Mol Biol* **25:** 1119–1127. doi:10.1038/s41594-018-0156-z

Antonova SV, Boeren J, Timmers HTM, Snel B. 2019. Epigenetics and transcription regulation during eukaryotic diversifica-

tion: the saga of TFIID. *Genes Dev* **33:** 888–902. doi:10 .1101/gad.300475.117

Auble DT, Hansen KE, Mueller CG, Lane WS, Thorner J, Hahn S. 1994. Mot1, a global repressor of RNA polymerase II transcription, inhibits TBP binding to DNA by an ATP-dependent mechanism. *Genes Dev* **8:** 1920–1934. doi:10.1101/gad.8.16 .1920

Audetat KA, Galbraith MD, Odell AT, Lee T, Pandey A, Espinosa JM, Dowell RD, Taatjes DJ. 2017. A kinase-independent role for cyclin-dependent kinase 19 in p53 response. *Mol Cell Biol* **37:** e00626-16. doi:10.1128/MCB.00626-16

Ayoubi LE, Dumay-Odelot H, Chernev A, Boissier F, Minvielle-Sébastia L, Urlaub H, Fribourg S, Teichmann M. 2019. The hRPC62 subunit of human RNA polymerase III displays helicase activity. *Nucleic Acids Res* **47:** 10313–10326. doi:10 .1093/nar/gkz788

Baumli S, Lolli G, Lowe ED, Troiani S, Rusconi L, Bullock AN, Debreczeni JE, Knapp S, Johnson LN. 2008. The structure of P-TEFb (CDK9/cyclin T1), its complex with flavopiridol and regulation by phosphorylation. *EMBO J* **27:** 1907–1918. doi:10.1038/emboj.2008.121

Benabdallah NS, Williamson I, Illingworth RS, Kane L, Boyle S, Sengupta D, Grimes GR, Therizols P, Bickmore WA. 2019. Decreased enhancer–promoter proximity accompanying enhancer activation. *Mol Cell* **76:** 473–484.e7. doi:10.1016/j .molcel.2019.07.038

Bernecky C, Taatjes DJ. 2012. Activator–mediator binding stabilizes RNA polymerase II orientation within the human Mediator–RNA polymerase II–TFIIF assembly. *J Mol Biol* **417:** 387–394. doi:10.1016/j.jmb.2012.02.014

Bernecky C, Grob P, Ebmeier CC, Nogales E, Taatjes DJ. 2011. Molecular architecture of the human mediator-RNA polymerase II–TFIIF assembly. *PLoS Biol* **9:** e1000603. doi:10 .1371/journal.pbio.1000603

Bernecky C, Herzog F, Baumeister W, Plitzko JM, Cramer P. 2016. Structure of transcribing mammalian RNA polymerase II. *Nature* **529:** 551–554. doi:10.1038/nature16482

Bernecky C, Plitzko JM, Cramer P. 2017. Structure of a transcribing RNA polymerase II–DSIF complex reveals a multidentate DNA–RNA clamp. *Nat Struct Mol Biol* **24:** 809–815. doi:10 .1038/nsmb.3465

Bhattacharya S, Lou X, Hwang P, Rajashankar KR, Wang X, Gustafsson JA, Fletterick RJ, Jacobson RH, Webb P. 2014. Structural and functional insight into TAF1–TAF7, a subcomplex of transcription factor II D. *Proc Natl Acad Sci* **111:** 9103–9108. doi:10.1073/pnas.1408293111

Bienossek C, Papai G, Schaffitzel C, Garzoni F, Chaillet M, Scheer E, Papadopoulos P, Tora L, Schultz P, Berger I. 2013. The architecture of human general transcription factor TFIID core complex. *Nature* **493:** 699–702. doi:10.1038/nature11791

Bleichenbacher M, Tan S, Richmond TJ. 2003. Novel interactions between the components of human and yeast TFIIA/TBP/ DNA complexes. *J Mol Biol* **332:** 783–793. doi:10.1016/ S0022-2836(03)00887-8

Boehning M, Dugast-Darzacq C, Rankovic M, Hansen AS, Yu T, Marie-Nelly H, McSwiggen DT, Kokic G, Dailey GM, Cramer P, et al. 2018. RNA polymerase II clustering through carboxy-terminal domain phase separation. *Nat Struct Mol Biol* **25:** 833–840. doi:10.1038/s41594-018-0112-y

Boeing S, Rigault C, Heidemann M, Eick D, Meisterernst M. 2010. RNA polymerase II C-terminal heptarepeat domain Ser-7 phosphorylation is established in a mediator-dependent fashion. *J Biol Chem* **285:** 188–196. doi:10.1074/jbc.M109 .046565

Bortvin A, Winston F. 1996. Evidence that Spt6p controls chromatin structure by a direct interaction with histones. *Science* **272:** 1473–1476. doi:10.1126/science.272.5267.1473

Boyer TG, Martin MED, Lees E, Ricciardi RP, Berk AJ. 1999. Mammalian Srb/Mediator complex is targeted by adenovirus E1a protein. *Nature* **399:** 276–279. doi:10.1038/20466

Brueckner F, Cramer P. 2008. Structural basis of transcription inhibition by α-amanitin and implications for RNA polymerase II translocation. *Nat Struct Mol Biol* **15:** 811–818. doi:10.1038/nsmb.1458

Bushnell DA, Kornberg RD. 2003. Complete, 12-subunit RNA polymerase II at 4.1-Å resolution: implications for the initiation of transcription. *Proc Natl Acad Sci* **100:** 6969–6973. doi:10.1073/pnas.1130601100

Cevher MA, Shi Y, Li D, Chait BT, Malik S, Roeder RG. 2014. Reconstitution of active human core Mediator complex reveals a critical role of the MED14 subunit. *Nat Struct Mol Biol* **21:** 1028–1034. doi:10.1038/nsmb.2914

Chen ZA, Jawhari A, Fischer L, Buchen C, Tahir S, Kamenski T, Rasmussen M, Lariviere L, Bukowski-Wills J, Nilges M, et al. 2010. Architecture of the RNA polymerase II–TFIIF complex revealed by cross-linking and mass spectrometry. *EMBO J* **29:** 717–726. doi:10.1038/emboj.2009.401

Cheung AC, Cramer P. 2012. A movie of RNA polymerase II transcription. *Cell* **149:** 1431–1437. doi:10.1016/j.cell.2012.06.006

Cianfrocco MA, Kassavetis GA, Grob P, Fang J, Juven-Gershon T, Kadonaga JT, Nogales E. 2013. Human TFIID binds to core promoter DNA in a reorganized structural state. *Cell* **152:** 120–131. doi:10.1016/j.cell.2012.12.005

Conaway RC, Garrett KP, Hanley JP, Conaway JW. 1991. Mechanism of promoter selection by RNA polymerase II: mammalian transcription factors α and βγ promote entry of polymerase into the preinitiation complex. *Proc Natl Acad Sci* **88:** 6205–6209. doi:10.1073/pnas.88.14.6205

Core L, Adelman K. 2019. Promoter-proximal pausing of RNA polymerase II: a nexus of gene regulation. *Genes Dev* **33:** 960–982. doi:10.1101/gad.325142.119

Core LJ, Martins AL, Danko CG, Waters CT, Siepel A, Lis JT. 2014. Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. *Nat Genet* **46:** 1311–1320. doi:10.1038/ng.3142

Cramer P. 2019. Organization and regulation of gene transcription. *Nature* **573:** 45–54. doi:10.1038/s41586-019-1517-4

Cramer P, Bushnell DA, Fu J, Gnatt AL, Maier-Davis B, Thompson NE, Burgess RR, Edwards AM, David PR, Kornberg RD. 2000. Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* **288:** 640–649. doi:10.1126/science.288.5466.640

Cramer P, Bushnell DA, Kornberg RD. 2001. Structural basis of transcription: RNA polymerase II at 2.8 Å resolution. *Science* **292:** 1863–1876. doi:10.1126/science.1059493

Currie SL, Doane JJ, Evans KS, Bhachech N, Madison BJ, Lau DKW, McIntosh LP, Skalicky JJ, Clark KA, Graves BJ. 2017. ETV4 and AP1 transcription factors form multivalent interactions with three sites on the MED25 activator-interacting domain. *J Mol Biol* **429:** 2975–2995. doi:10.1016/j.jmb.2017.06.024

Deato MD, Tjian R. 2007. Switching of the core transcription machinery during myogenesis. *Genes Dev* **21:** 2137–2149. doi:10.1101/gad.1583407

Decker TM, Forné I, Straub T, Elsaman H, Ma G, Shah N, Imhof A, Eick D. 2019. Analog-sensitive cell line identifies cellular substrates of CDK9. *Oncotarget* **10:** 6934–6943. doi:10.18632/oncotarget.27334

Dienemann C, Schwalb B, Schilbach S, Cramer P. 2019. Promoter distortion and opening in the RNA polymerase II cleft. *Mol Cell* **73:** 97–106.e4. doi:10.1016/j.molcel.2018.10.014

Donczew R, Warfield L, Pacheco D, Erijman A, Hahn S. 2020. Two roles for the yeast transcription coactivator SAGA and a set of genes redundantly regulated by TFIID and SAGA. *Elife* **9:** e50109. doi:10.7554/eLife.50109

Donner AJ, Ebmeier CC, Taatjes DJ, Espinosa JM. 2010. CDK8 is a positive regulator of transcriptional elongation within the serum response network. *Nat Struct Mol Biol* **17:** 194–201. doi:10.1038/nsmb.1752

Donovan BT, Huynh A, Ball DA, Patel HP, Poirier MG, Larson DR, Ferguson ML, Lenstra TL. 2019. Live-cell imaging reveals the interplay between transcription factors, nucleosomes, and bursting. *EMBO J* **38:** e100809. doi:10.15252/embj.2018100809

Dronamraju R, Hepperla AJ, Shibata Y, Adams AT, Magnuson T, Davis IJ, Strahl BD. 2018. Spt6 association with RNA polymerase II directs mRNA turnover during transcription. *Mol Cell* **70:** 1054–1066.e4. doi:10.1016/j.molcel.2018.05.020

Duttke SH, Lacadie SA, Ibrahim MM, Glass CK, Corcoran DL, Benner C, Heinz S, Kadonaga JT, Ohler U. 2015. Human promoters are intrinsically directional. *Mol Cell* **57:** 674–684. doi:10.1016/j.molcel.2014.12.029

Ebmeier CC, Taatjes DJ. 2010. Activator-mediator binding regulates Mediator–cofactor interactions. *Proc Natl Acad Sci* **107:** 11283–11288. doi:10.1073/pnas.0914215107

Ebmeier CC, Erickson B, Allen BL, Allen MA, Kim H, Fong N, Jacobsen JR, Liang K, Shilatifard A, Dowell RD, et al. 2017. Human TFIIH kinase CDK7 regulates transcription-associated chromatin modifications. *Cell Rep* **20:** 1173–1186. doi:10.1016/j.celrep.2017.07.021

Ehara H, Yokoyama T, Shigematsu H, Yokoyama S, Shirouzu M, Sekine SI. 2017. Structure of the complete elongation complex of RNA polymerase II with basal factors. *Science* **357:** 921–924. doi:10.1126/science.aan8552

Eichner J, Chen H, Warfield L, Hahn S. 2010. Position of the general transcription factor TFIIF within the RNA polymerase II transcription preinitiation complex. *EMBO J* **29:** 706–716. doi:10.1038/emboj.2009.386

El Khattabi LE, Zhao H, Kalchschmidt J, Young N, Jung S, Van Blerkom P, Kieffer-Kwon P, Kieffer-Kwon KR, Park S, Wang X, et al. 2019. A pliable mediator acts as a functional rather than an architectural bridge between promoters and enhancers. *Cell* **178:** 1145–1158.e20. doi:10.1016/j.cell.2019.07.011

Elmlund H, Baraznenok V, Lindahl M, Samuelsen CO, Koeck PJB, Holmberg S, Hebert H, Gustafsson CM. 2006. The cyclin-dependent kinase 8 module sterically blocks mediator interactions with RNA polymerase II. *Proc Natl Acad Sci* **103:** 15788–15793. doi:10.1073/pnas.0607483103

Engel C, Neyer S, Cramer P. 2018. Distinct mechanisms of transcription initiation by RNA polymerases I and II. *Annu Rev Biophys* **47:** 425–446. doi:10.1146/annurev-biophys-070317-033058

Eyboulet F, Wydau-Dematteis S, Eychenne T, Alibert O, Neil H, Boschiero C, Nevers MC, Volland H, Cornu D, Redeker V, et al. 2015. Mediator independently orchestrates multiple steps of preinitiation complex assembly *in vivo*. *Nucleic Acids Res* **43:** 9214–9231. doi:10.1093/nar/gkv782

Fant CB, Taatjes DJ. 2019. Regulatory functions of the Mediator kinases CDK8 and CDK19. *Transcription* **10:** 76–90. doi:10.1080/21541264.2018.1556915

Fischer V, Schumacher K, Tora L, Devys D. 2019. Global role for coactivator complexes in RNA polymerase II

transcription. *Transcription* **10:** 29–36. doi:10.1080/21541264.2018.1521214

Fishburn J, Tomko E, Galburt E, Hahn S. 2015. Double-stranded DNA translocase activity of transcription factor TFIIH and the mechanism of RNA polymerase II open complex formation. *Proc Natl Acad Sci* **112:** 3961–3966. doi:10.1073/pnas.1417709112

Fitz J, Neumann T, Pavri R. 2018. Regulation of RNA polymerase II processivity by Spt5 is restricted to a narrow window during elongation. *EMBO J* **37:** e97965. doi:10.15252/embj.201797965

Flanagan PM, Kelleher RJ III, Sayre MH, Tschochner H, Kornberg R. 1991. A Mediator required for activation of RNA polymerase II transcription in vitro. *Nature* **350:** 436–438. doi:10.1038/350436a0

Fondell JD, Ge H, Roeder RG. 1996. Ligand induction of a transcriptionally active thyroid hormone receptor coactivator complex. *Proc Natl Acad Sci* **93:** 8329–8333. doi:10.1073/pnas.93.16.8329

Friedl EM, Lane WS, Erdjument-Bromage H, Tempst P, Reinberg D. 2003. The C-terminal domain phosphatase and transcription elongation activities of FCP1 are regulated by phosphorylation. *Proc Natl Acad Sci* **100:** 2328–2333. doi:10.1073/pnas.2628049100

Fukaya T, Lim B, Levine M. 2016. Enhancer control of transcriptional bursting. *Cell* **166:** 358–368. doi:10.1016/j.cell.2016.05.025

Furlong EEM, Levine M. 2018. Developmental enhancers and chromosome topology. *Science* **361:** 1341–1345. doi:10.1126/science.aau0320

Gaiser F, Tan S, Richmond TJ. 2000. Novel dimerization fold of RAP30/RAP74 in human TFIIF at 1.7 A resolution. *J Mol Biol* **302:** 1119–1127. doi:10.1006/jmbi.2000.4110

Galbraith MD, Allen MA, Bensard CL, Wang X, Schwinn MK, Qin B, Long HW, Daniels DL, Hahn WC, Dowell RD, et al. 2013. HIF1A employs CDK8-mediator to stimulate RNAPII elongation in response to hypoxia. *Cell* **153:** 1327–1339. doi:10.1016/j.cell.2013.04.048

Geiger JH, Hahn S, Lee S, Sigler PB. 1996. Crystal structure of the yeast TFIIA/TBP/DNA complex. *Science* **272:** 830–836. doi:10.1126/science.272.5263.830

Gilchrist DA, dos Santos G, Fargo DC, Xie B, Gao Y, Li L, Adelman K. 2010. Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. *Cell* **143:** 540–551. doi:10.1016/j.cell.2010.10.004

Gilfillan S, Stelzer G, Piaia E, Hofmann MG, Meisterernst M. 2005. Efficient binding of NC2.TATA-binding protein to DNA in the absence of TATA. *J Biol Chem* **280:** 6222–6230. doi:10.1074/jbc.M406343200

Gnatt AL, Cramer P, Fu J, Bushnell DA, Kornberg RD. 2001. Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science* **292:** 1876–1882. doi:10.1126/science.1059495

Goodrich JA, Tjian R. 1994. Transcription factors IIE and IIH and ATP hydrolysis direct promoter clearance by RNA polymerase II. *Cell* **77:** 145–156. doi:10.1016/0092-8674(94)90242-9

Gout JF, Li W, Fritsch C, Li A, Haroon S, Singh L, Hua D, Fazelinia H, Smith Z, Seeholzer S, et al. 2017. The landscape of transcription errors in eukaryotic cells. *Sci Adv* **3:** e1701484. doi:10.1126/sciadv.1701484

Greber BJ, Nguyen THD, Fang J, Afonine PV, Adams PD, Nogales E. 2017. The cryo-electron microscopy structure of human transcription factor IIH. *Nature* **549:** 414–417. doi:10.1038/nature23903

Greber BJ, Toso DB, Fang J, Nogales E. 2019. The complete structure of the human TFIIH core complex. *Elife* **8:** e44771. doi:10.7554/eLife.44771

Gressel S, Schwalb B, Decker TM, Qin W, Leonhardt H, Eick D, Cramer P. 2017. CDK9-dependent RNA polymerase II pausing controls transcription initiation. *Elife* **6:** e29736. doi:10.7554/eLife.29736

Grünberg S, Hahn S. 2013. Structural insights into transcription initiation by RNA polymerase II. *Trends Biochem Sci* **38:** 603–611. doi:10.1016/j.tibs.2013.09.002

Guo YE, Manteiga JC, Henninger JE, Sabari BR, Dall'Agnese A, Hannett NM, Spille JH, Afeyan LK, Zamudio AV, Shrinivas K, et al. 2019. Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. *Nature* **572:** 543–548. doi:10.1038/s41586-019-1464-0

Gupta K, Watson AA, Baptista T, Scheer E, Chambers AL, Koehler C, Zou J, Obong-Ebong I, Kandiah E, Temblador A, et al. 2017. Architecture of TAF11/TAF13/TBP complex suggests novel regulation properties of general transcription factor TFIID. *Elife* **6:** e30395. doi:10.7554/eLife.30395

Haberle V, Stark A. 2018. Eukaryotic core promoters and the functional basis of transcription initiation. *Nat Rev Mol Cell Biol* **19:** 621–637. doi:10.1038/s41580-018-0028-8

Harlen KM, Churchman LS. 2017. The code and beyond: transcription regulation by the RNA polymerase II carboxy-terminal domain. *Nat Rev Mol Cell Biol* **18:** 263–273. doi:10.1038/nrm.2017.10

Harper TM, Taatjes DJ. 2018. The complex structure and function of Mediator. *J Biol Chem* **293:** 13778–13785. doi:10.1074/jbc.R117.794438

Hawley DK, Roeder RG. 1987. Functional steps in transcription initiation and reinitiation from the major late promoter in a HeLa nuclear extract. *J Biol Chem* **262:** 3452–3461.

He Y, Fang J, Taatjes DJ, Nogales E. 2013. Structural visualization of key steps in human transcription initiation. *Nature* **495:** 481–486. doi:10.1038/nature11991

He Y, Yan C, Fang J, Inouye C, Tjian R, Ivanov I, Nogales E. 2016. Near-atomic resolution visualization of human transcription promoter opening. *Nature* **533:** 359–365. doi:10.1038/nature17970

Heinz S, Romanoski CE, Benner C, Glass CK. 2015. The selection and function of cell type-specific enhancers. *Nat Rev Mol Cell Biol* **16:** 144–154. doi:10.1038/nrm3949

Heist T, Fukaya T, Levine M. 2019. Large distances separate coregulated genes in living *Drosophila* embryos. *Proc Natl Acad Sci* **116:** 15062–15067. doi:10.1073/pnas.1908962116

Herzel L, Ottoz DSM, Alpert T, Neugebauer KM. 2017. Splicing and transcription touch base: co-transcriptional spliceosome assembly and function. *Nat Rev Mol Cell Biol* **18:** 637–650. doi:10.1038/nrm.2017.63

Hieb AR, Halsey WA, Betterton MD, Perkins TT, Kugel JF, Goodrich JA. 2007. TFIIA changes the conformation of the DNA in TBP/TATA complexes and increases their kinetic stability. *J Mol Biol* **372:** 619–632. doi:10.1016/j.jmb.2007.06.061

Hoeppner S, Baumli S, Cramer P. 2005. Structure of the mediator subunit cyclin C and its implications for CDK8 function. *J Mol Biol* **350:** 833–842. doi:10.1016/j.jmb.2005.05.041

Holstege FC, Tantin D, Carey M, van der Vliet PC, Timmers HT. 1995. The requirement for the basal transcription factor IIE is determined by the helical stability of promoter DNA. *EMBO J* **14:** 810–819. doi:10.1002/j.1460-2075.1995.tb07059.x

Holstege FCP, van der Vliet PC, Timmers HTM. 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. doi:10.1002/j.1460-2075.1996 .tb00512.x

Hou L, Wang Y, Liu Y, Zhang N, Shamovsky I, Nudler E, Tian B, Dynlacht BD. 2019. Paf1C regulates RNA polymerase II progression by modulating elongation rate. *Proc Natl Acad Sci* **116:** 14583–14592. doi:10.1073/pnas.1904324116

Hsu JY, Juven-Gershon T, Marr MT 2nd, Wright KJ, Tjian R, Kadonaga JT. 2008. TBP, Mot1, and NC2 establish a regulatory circuit that controls DPE-dependent versus TATA-dependent transcription. *Genes Dev* **22:** 2353–2358. doi:10.1101/gad.1681808

Imasaki T, Calero G, Cai G, Tsai KL, Yamada K, Cardelli F, Erdjument-Bromage H, Tempst P, Berger I, Kornberg GL, et al. 2011. Architecture of the mediator head module. *Nature* **475:** 240–243. doi:10.1038/nature10162

Imbalzano AN, Zaret KS, Kingston RE. 1994. Transcription factor (TF) IIB and TFIIA can independently increase the affinity of the TATA-binding protein for DNA. *J Biol Chem* **269:** 8280–8286.

Jacobson RH, Ladurner AG, King DS, Tjian R. 2000. Structure and function of a human TAF$_{II}$250 double bromodomain module. *Science* **288:** 1422–1425. doi:10.1126/science.288.5470.1422

Jeronimo C, Poitras C, Robert F. 2019. Histone recycling by FACT and Spt6 during transcription prevents the scrambling of histone modifications. *Cell Rep* **28:** 1206–1218.e8. doi:10.1016/j.celrep.2019.06.097

Jiang Y, Gralla JD. 1993. Uncoupling of initiation and reinitiation rates during HeLa RNA polymerase II transcription in vitro. *Mol Cell Biol* **13:** 4572–4577. doi:10.1128/MCB.13.8.4572

Jung D, Shimogawa H, Kwon Y, Mao Q, Sato S, Kamisuki S, Kigoshi H, Uesugi M. 2009. Wrenchnolol derivative optimized for gene activation in cells. *J Am Chem Soc* **131:** 4774–4782. doi:10.1021/ja900669k

Juven-Gershon T, Cheng S, Kadonaga JT. 2006. Rational design of a super core promoter that enhances gene expression. *Nat Methods* **3:** 917–922. doi:10.1038/nmeth937

Kamada K, Shu F, Chen H, Malik S, Stelzer G, Roeder RG, Meisterernst M, Burley SK. 2001. Crystal structure of negative cofactor 2 recognizing the TBP-DNA transcription complex. *Cell* **106:** 71–81. doi:10.1016/S0092-8674(01)00417-2

Kamada K, Roeder RG, Burley SK. 2003. Molecular mechanism of recruitment of TFIIF- associating RNA polymerase C-terminal domain phosphatase (FCP1) by transcription factor IIF. *Proc Natl Acad Sci* **100:** 2296–2299. doi:10.1073/pnas.262798199

Kaplan CD, Laprade L, Winston F. 2003. Transcription elongation factors repress transcription initiation from cryptic sites. *Science* **301:** 1096–1099. doi:10.1126/science.1087374

Kaplan CD, Larsson KM, Kornberg RD. 2008. The RNA polymerase II trigger loop functions in substrate selection and is directly targeted by α-amanitin. *Mol Cell* **30:** 547–556. doi:10.1016/j.molcel.2008.04.023

Kays AR, Schepartz A. 2000. Virtually unidirectional binding of TBP to the AdMLP TATA box within the quaternary complex with TFIIA and TFIIB. *Chem Biol* **7:** 601–610. doi:10.1016/S1074-5521(00)00009-0

Kettenberger H, Armache K, Cramer P. 2003. Architecture of the RNA polymerase II–TFIIS complex and implications for mRNA cleavage. *Cell* **114:** 347–357. doi:10.1016/S0092-8674(03)00598-1

Kettenberger H, Armache K, Cramer P. 2004. Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. *Mol Cell* **16:** 955–965. doi:10.1016/j.molcel.2004.11.040

Khatter H, Vorländer MK, Müller CW. 2017. RNA polymerase I and III: similar yet unique. *Curr Opin Struct Biol* **47:** 88–94. doi:10.1016/j.sbi.2017.05.008

Kim JL, Nikolov DB, Burley SK. 1993a. Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature* **365:** 520–527. doi:10.1038/365520a0

Kim Y, Geiger JH, Hahn S, Sigler PB. 1993b. Crystal structure of a yeast TBP/TATA-box complex. *Nature* **365:** 512–520. doi:10.1038/365512a0

Kim Y, Björklund S, Li Y, Sayre MH, Kornberg RD. 1994. A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. *Cell* **77:** 599–608. doi:10.1016/0092-8674(94)90221-6

Kim TK, Zhao Y, Ge H, Bernstein R, Roeder RG. 1995. TATA-binding protein residues implicated in a functional interplay between negative cofactor NC2 (Dr1) and general factors TFIIA and TFIIB. *J Biol Chem* **270:** 10976–10981. doi:10.1074/jbc.270.18.10976

Kim B, Nesvizhskii AI, Rani PG, Hahn S, Aebersold R, Ranish JA. 2007. The transcription elongation factor TFIIS is a component of RNA polymerase II preinitiation complexes. *Proc Natl Acad Sci* **104:** 16068–16073. doi:10.1073/pnas.0704573104

Kizer KO, Phatnani HP, Shibata Y, Hall H, Greenleaf AL, Strahl BD. 2005. A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. *Mol Cell Biol* **25:** 3305–3316. doi:10.1128/MCB.25.8.3305-3316.2005

Klatt F, Leitner A, Kim IV, Ho-Xuan H, Schneider EV, Langhammer F, Weinmann R, Müller MR, Huber R, Meister G, et al. 2020. A precisely positioned MED12 activation helix stimulates CDK8 kinase activity. *Proc Natl Acad Sci* **117:** 2894–2905. doi:10.1073/pnas.1917635117

Knuesel MT, Meyer KD, Bernecky C, Taatjes DJ. 2009a. The human CDK8 subcomplex is a molecular switch that controls Mediator co-activator function. *Genes Dev* **23:** 439–451. doi:10.1101/gad.1767009

Knuesel MT, Meyer KD, Donner AJ, Espinosa JM, Taatjes DJ. 2009b. The human CDK8 subcomplex is a histone kinase that requires Med12 for activity and can function independently of Mediator. *Mol Cell Biol* **29:** 650–661. doi:10.1128/MCB.00993-08

Kokic G, Chernev A, Tegunov D, Dienemann C, Urlaub H, Cramer P. 2019. Structural basis of TFIIH activation for nucleotide excision repair. *Nat Commun* **10:** 2885. doi:10.1038/s41467-019-10745-5

Koleske AJ, Young RA. 1994. An RNA polymerase II holoenzyme responsive to activators. *Nature* **368:** 466–469. doi:10.1038/368466a0

Kolesnikova O, Ben-Shem A, Luo J, Ranish J, Schultz P, Papai G. 2018. Molecular structure of promoter-bound yeast TFIID. *Nat Commun* **9:** 4666. doi:10.1038/s41467-018-07096-y

Koster MJ, Yildirim AD, Weil PA, Holstege FC, Timmers HT. 2014. Suppression of intragenic transcription requires the MOT1 and NC2 regulators of TATA-binding protein. *Nucleic Acids Res* **42:** 4220–4229. doi:10.1093/nar/gkt1398

Kostrewa D, Zeller ME, Armache K, Seizl M, Leike K, Thomm M, Cramer P. 2009. RNA polymerase II–TFIIB structure and mechanism of transcription initiation. *Nature* **462:** 323–330. doi:10.1038/nature08548

Kuhlbrandt W. 2014. Biochemistry. The resolution revolution. *Science* **343:** 1443–1444. doi:10.1126/science.1251652

Kumar KP, Akoulitchev S, Reinberg D. 1998. Promoter-proximal stalling results from the inability to recruit transcription factor IIH to the transcription complex and is a regulated event.

Proc Natl Acad Sci **95:** 9767–9772. doi:10.1073/pnas.95.17.9767

Kwak H, Lis JT. 2013. Control of transcriptional elongation. *Annu Rev Genet* **47:** 483–508. doi:10.1146/annurev-genet-110711-155440

Kwon I, Kato M, Xiang S, Wu L, Theodoropoulos P, Mirzaei H, Han T, Xie S, Corden JL, McKnight SL. 2013. Phosphorylation-regulated binding of RNA polymerase II to fibrous polymers of low-complexity domains. *Cell* **155:** 1049–1060. doi:10.1016/j.cell.2013.10.033

Kyrpides NC, Woese CR, Ouzounis CA. 1996. KOW: a novel motif linking a bacterial transcription factor with ribosomal proteins. *Trends Biochem Sci* **21:** 425–426. doi:10.1016/S0968-0004(96)30036-4

Larivière L, Plaschka C, Seizl M, Wenzeck L, Kurth F, Cramer P. 2012. Structure of the Mediator head module. *Nature* **492:** 448–451. doi:10.1038/nature11670

Lee JH, Skalnik DG. 2008. Wdr82 is a C-terminal domain-binding protein that recruits the Setd1A Histone H3-Lys4 methyltransferase complex to transcription start sites of transcribed human genes. *Mol Cell Biol* **28:** 609–618. doi:10.1128/MCB.01356-07

Lee TI, Young RA. 2013. Transcriptional regulation and its misregulation in disease. *Cell* **152:** 1237–1251. doi:10.1016/j.cell.2013.02.014

Lee CW, Ferreon JC, Ferreon ACM, Arai M, Wright PE. 2010. Graded enhancement of p53 binding to CREB-binding protein (CBP) by multisite phosphorylation. *Proc Natl Acad Sci* **107:** 19290–19295. doi:10.1073/pnas.1013078107

Lenstra TL, Rodriguez J, Chen H, Larson DR. 2016. Transcription dynamics in living cells. *Annu Rev Biophys* **45:** 25–47. doi:10.1146/annurev-biophys-062215-010838

Levine M, Cattoglio C, Tjian R. 2014. Looping back to leap forward: transcription enters a new era. *Cell* **157:** 13–25. doi:10.1016/j.cell.2014.02.009

Li J, Liu Y, Rhee HS, Ghosh SK, Bai L, Pugh BF, Gilmour DS. 2013. Kinetic competition between elongation rate and binding of NELF controls promoter-proximal pausing. *Mol Cell* **50:** 711–722. doi:10.1016/j.molcel.2013.05.016

Li J, Dong A, Saydaminova K, Chang H, Wang G, Ochiai H, Yamamoto T, Pertsinidis A. 2019. Single-molecule nanoscopy elucidates RNA polymerase II transcription at single genes in live cells. *Cell* **178:** 491–506.e28. doi:10.1016/j.cell.2019.05.029

Lim WK, Ong CK, Tan J, Thike AA, Ng CC, Rajasegaran V, Myint SS, Nagarajan S, Nasir ND, McPherson JR, et al. 2014. Exome sequencing identifies highly recurrent MED12 somatic mutations in breast fibroadenoma. *Nat Genet* **46:** 877–880. doi:10.1038/ng.3037

Lim B, Heist T, Levine M, Fukaya T. 2018. Visualization of transvection in living *Drosophila* embryos. *Mol Cell* **70:** 287–296.e6. doi:10.1016/j.molcel.2018.02.029

Lin YC, Gralla JD. 2005. Stimulation of the XPB ATP-dependent helicase by the β subunit of TFIIE. *Nucleic Acids Res* **33:** 3072–3081. doi:10.1093/nar/gki623

Lin YC, Choi WS, Gralla JD. 2005. TFIIH XPB mutants suggest a unified bacterial-like mechanism for promoter opening but not escape. *Nat Struct Mol Biol* **12:** 603–607. doi:10.1038/nsmb949

Liu Z, Tjian R. 2018. Visualizing transcription factor dynamics in living cells. *J Cell Biol* **217:** 1181–1191. doi:10.1083/jcb.201710038

Liu D, Ishima R, Tong KI, Bagby S, Kokubo T, Muhandiram DR, Kay LE, Nakatani Y, Ikura M. 1998. Solution structure of a TBP-TAF$_{II}$230 complex: protein mimicry of the minor groove

surface of the TATA box unwound by TBP. *Cell* **94:** 573–583. doi:10.1016/S0092-8674(00)81599-8

Liu X, Bushnell DA, Wang D, Calero G, Kornberg RD. 2010. Structure of an RNA polymerase II–TFIIB complex and the transcription initiation mechanism. *Science* **327:** 206–209. doi:10.1126/science.1182015

Louder RK, He Y, López-Blanco JR, Fang J, Chacón P, Nogales E. 2016. Structure of promoter-bound TFIID and model of human pre-initiation complex assembly. *Nature* **531:** 604–609. doi:10.1038/nature17394

Lu H, Yu D, Hansen AS, Ganguly S, Liu R, Heckert A, Darzacq X, Zhou Q. 2018. Phase-separation mechanism for C-terminal hyperphosphorylation of RNA polymerase II. *Nature* **558:** 318–323. doi:10.1038/s41586-018-0174-3

Luo Z, Lin C, Shilatifard A. 2012. The super elongation complex (SEC) family in transcriptional control. *Nat Rev Mol Cell Biol* **13:** 543–547. doi:10.1038/nrm3417

Luo J, Cimermancic P, Viswanath S, Ebmeier CC, Kim B, Dehecq M, Raman V, Greenberg CH, Pellarin R, Sali A, et al. 2015. Architecture of the human and yeast general transcription and DNA repair factor TFIIH. *Mol Cell* **59:** 794–806. doi:10.1016/j.molcel.2015.07.016

Makinen N, Mehine M, Tolvanen J, Kaasinen E, Li Y, Lehtonen HJ, Gentile M, Yan J, Enge M, Taipale M, et al. 2011. MED12, the mediator complex subunit 12 gene, is mutated at high frequency in uterine leiomyomas. *Science* **334:** 252–255. doi:10.1126/science.1208930

Malecová B, Gross P, Boyer-Guittaut M, Yavuz S, Oelgeschläger T. 2007. The initiator core promoter element antagonizes repression of TATA-directed transcription by negative cofactor NC2. *J Biol Chem* **282:** 24767–24776. doi:10.1074/jbc.M702776200

Max T, Søgaard M, Svejstrup JQ. 2007. Hyperphosphorylation of the C-terminal repeat domain of RNA polymerase II facilitates dissociation of its complex with mediator. *J Biol Chem* **282:** 14113–14120. doi:10.1074/jbc.M701345200

Meyer KD, Lin S, Bernecky C, Gao Y, Taatjes DJ. 2010. p53 activates transcription by directing structural shifts in Mediator. *Nat Struct Mol Biol* **17:** 753–760. doi:10.1038/nsmb.1816

Milbradt AG, Kulkarni M, Yi T, Takeuchi K, Sun ZY, Luna RE, Selenko P, Näär AM, Wagner G. 2011. Structure of the VP16 transactivator target in the mediator. *Nat Struct Mol Biol* **18:** 410–415. doi:10.1038/nsmb.1999

Mir M, Stadler MR, Ortiz SA, Hannon CE, Harrison MM, Darzacq X, Eisen MB. 2018. Dynamic multifactor hubs interact transiently with sites of active transcription in *Drosophila* embryos. *Elife* **7:** e40497. doi:10.7554/eLife.40497

Missra A, Gilmour DS. 2010. Interactions between DSIF (DRB sensitivity inducing factor), NELF (negative elongation factor), and the *Drosophila* RNA polymerase II transcription elongation complex. *Proc Natl Acad Sci* **107:** 11301–11306. doi:10.1073/pnas.1000681107

Mühlbacher W, Sainsbury S, Hemann M, Hantsche M, Neyer S, Herzog F, Cramer P. 2014. Conserved architecture of the core RNA polymerase II initiation complex. *Nat Commun* **5:** 4310. doi:10.1038/ncomms5310

Mylona A, Theillet FX, Foster C, Cheng TM, Miralles F, Bates PA, Selenko P, Treisman R. 2016. Opposing effects of Elk-1 multisite phosphorylation shape its response to ERK activation. *Science* **354:** 233–237. doi:10.1126/science.aad1872

Näär AM, Beaurang PA, Zhou S, Abraham S, Solomon W, Tjian R. 1999. Composite co-activator ARC mediates chromatin-directed transcriptional activation. *Nature* **398:** 828–832. doi:10.1038/19789

Näär AM, Taatjes DJ, Zhai W, Nogales E, Tjian R. 2002. Human CRSP interacts with RNA polymerase II CTD and adopts a specific CTD-bound conformation. *Genes Dev* **16:** 1339–1344. doi:10.1101/gad.987602

Nguyen BD, Abbott KL, Potempa K, Kobor MS, Archambault J, Greenblatt J, Legault P, Omichinski JG. 2003. NMR structure of a complex containing the TFIIF subunit RAP74 and the RNA polymerase II carboxyl-terminal domain phosphatase FCP1. *Proc Natl Acad Sci* **100:** 5688–5693. doi:10.1073/pnas.1031524100

Nikolov DB, Hu SH, Lin J, Gasch A, Hoffmann A, Horikoshi M, Chua NH, Roeder RG, Burley SK. 1992. Crystal structure of TFIID TATA-box binding protein. *Nature* **360:** 40–46. doi:10.1038/360040a0

Nishikawa JL, Boeszoermenyi A, Vale-Silva LA, Torelli R, Posteraro B, Sohn YJ, Ji F, Gelev V, Sanglard D, Sanguinetti M, et al. 2016. Inhibiting fungal multidrug resistance by disrupting an activator-Mediator interaction. *Nature* **530:** 485–489. doi:10.1038/nature16963

Nogales E, Greber BJ. 2019. High-resolution cryo-EM structures of TFIIH and their functional implications. *Curr Opin Struct Biol* **59:** 188–194. doi:10.1016/j.sbi.2019.08.002

Nozawa K, Schneider TR, Cramer P. 2017. Core Mediator structure at 3.4 Å extends model of transcription initiation complex. *Nature* **545:** 248–251. doi:10.1038/nature22328

Ohkuma Y, Roeder RG. 1994. Regulation of TFIIH ATPase and kinase activities by TFIIE during active initiation complex formation. *Nature* **368:** 160–163. doi:10.1038/368160a0

Ohkuma Y, Hashimoto S, Wang CK, Horikoshi M, Roeder RG. 1995. Analysis of the role of TFIIE in basal transcription and TFIIH-mediated carboxy-terminal domain phosphorylation through structure-function studies of TFIIE-α. *Mol Cell Biol* **15:** 4856–4866. doi:10.1128/MCB.15.9.4856

Papai G, Tripathi MK, Ruhlmann C, Werten S, Crucifix C, Weil PA, Schultz P. 2009. Mapping the initiator binding Taf2 subunit in the structure of hydrated yeast TFIID. *Structure* **17:** 363–373. doi:10.1016/j.str.2009.01.006

Papai G, Frechard A, Kolesnikova O, Crucifix C, Schultz P, Ben-Shem A. 2020. Structure of SAGA and mechanism of TBP deposition on gene promoters. *Nature* **577:** 711–716. doi:10.1038/s41586-020-1944-2

Patel AB, Louder RK, Greber BJ, Grünberg S, Luo J, Fang J, Liu Y, Ranish J, Hahn S, Nogales E. 2018. Structure of human TFIID and mechanism of TBP loading onto promoter DNA. *Science* **362:** eaau8872. doi:10.1126/science.aau8872

Pavri R, Zhu B, Li G, Trojer P, Mandal S, Shilatifard A, Reinberg D. 2006. Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. *Cell* **125:** 703–717. doi:10.1016/j.cell.2006.04.029

Plaschka C, Larivière L, Wenzeck L, Seizl M, Hemann M, Tegunov D, Petrotchenko EV, Borchers CH, Baumeister W, Herzog F, et al. 2015. Architecture of the RNA polymerase II–Mediator core initiation complex. *Nature* **518:** 376–380. doi:10.1038/nature14229

Plaschka C, Hantsche M, Dienemann C, Burzinski C, Plitzko J, Cramer P. 2016. Transcription initiation complex structures elucidate DNA opening. *Nature* **533:** 353–358. doi:10.1038/nature17990

Poss ZC, Ebmeier CC, Odell AT, Tangpeerachaikul A, Lee T, Pelish HE, Shair MD, Dowell RD, Old WM, Taatjes DJ. 2016. Identification of mediator kinase substrates in human cells using cortistatin A and quantitative phosphoproteomics. *Cell Rep* **15:** 436–450. doi:10.1016/j.celrep.2016.03.030

Pufall MA, Lee GM, Nelson ML, Kang HS, Velyvis A, Kay LE, McIntosh LP, Graves BJ. 2005. Variable control of Ets-1 DNA binding by multiple phosphates in an unstructured region. *Science* **309:** 142–145. doi:10.1126/science.1111915

Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA. 2017. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14:** 290–296. doi:10.1038/nmeth.4169

Rachez C, Lemon BD, Suldan Z, Bromleigh V, Gamble M, Näär AM, Erdjument-Bromage H, Tempst P, Freedman LP. 1999. Ligand-dependent transcription activation by nuclear receptors requires the DRIP complex. *Nature* **398:** 824–828. doi:10.1038/19783

Rahl PB, Lin CY, Seila AC, Flynn RA, McCuine S, Burge CB, Sharp PA, Young RA. 2010. c-Myc regulates transcriptional pause release. *Cell* **141:** 432–445. doi:10.1016/j.cell.2010.03.030

Rasmussen EB, Lis JT. 1993. In vivo transcriptional pausing and cap formation on three *Drosophila* heat shock genes. *Proc Natl Acad Sci* **90:** 7923–7927. doi:10.1073/pnas.90.17.7923

Rhee HS, Pugh BF. 2012. Genome-wide structure and organization of eukaryotic pre-initiation complexes. *Nature* **483:** 295–301. doi:10.1038/nature10799

Rimel JK, Taatjes DJ. 2018. The essential and multi-functional TFIIH complex. *Protein Sci* **27:** 1018–1037. doi:10.1002/pro.3424

Robinson PJ, Bushnell DA, Trnka MJ, Burlingame AL, Kornberg RD. 2012. Structure of the mediator head module bound to the carboxy-terminal domain of RNA polymerase II. *Proc Natl Acad Sci* **109:** 17931–17935. doi:10.1073/pnas.1215241109

Robinson PJ, Trnka MJ, Pellarin R, Greenberg CH, Bushnell DA, Davis R, Burlingame AL, Sali A, Kornberg RD. 2015. Molecular architecture of the yeast mediator complex. *Elife* **4:** e08719. doi:10.7554/eLife.08719

Robinson PJ, Trnka MJ, Bushnell DA, Davis RE, Mattei PJ, Burlingame AL, Kornberg RD. 2016. Structure of a complete mediator-RNA polymerase II pre-initiation complex. *Cell* **166:** 1411–1422.e16. doi:10.1016/j.cell.2016.08.050

Rodriguez J, Ren G, Day CR, Zhao K, Chow CC, Larson DR. 2019. Intrinsic dynamics of a human gene reveal the basis of expression heterogeneity. *Cell* **176:** 213–226.e18. doi:10.1016/j.cell.2018.11.026

Roeder RG. 2019. 50+ years of eukaryotic transcription: an expanding universe of factors and mechanisms. *Nat Struct Mol Biol* **26:** 783–791. doi:10.1038/s41594-019-0287-x

Rowe SP, Casey RJ, Brennan BB, Buhrlage SJ, Mapp AK. 2007. Transcriptional up-regulation in cells mediated by a small molecule. *J Am Chem Soc* **129:** 10654–10655. doi:10.1021/ja0736865

Ryu S, Zhou S, Ladurner AG, Tjian R. 1999. The transcriptional cofactor complex CRSP is required for activity of the enhancer-binding protein Sp1. *Nature* **397:** 446–450. doi:10.1038/17141

Sadowski CL, Henry RW, Lobo SM, Hernandez N. 1993. Targeting TBP to a non-TATA box *cis*-regulatory element: a TBP-containing complex activates transcription from snRNA promoters through the PSE. *Genes Dev* **7:** 1535–1548. doi:10.1101/gad.7.8.1535

Sainsbury S, Niesser J, Cramer P. 2013. Structure and function of the initially transcribing RNA polymerase II–TFIIB complex. *Nature* **493:** 437–440. doi:10.1038/nature11715

Sainsbury S, Bernecky C, Cramer P. 2015. Structural basis of transcription initiation by RNA polymerase II. *Nat Rev Mol Cell Biol* **16:** 129–143. doi:10.1038/nrm3952

Saldi T, Cortazar MA, Sheridan RM, Bentley DL. 2016. Coupling of RNA polymerase II transcription elongation with pre-

mRNA splicing. *J Mol Biol* **428:** 2623–2635. doi:10.1016/j.jmb.2016.04.017

Sansó M, Levin RS, Lipp JJ, Wang VY, Greifenberg AK, Quezada EM, Ali A, Ghosh A, Larochelle S, Rana TM, et al. 2016. P-TEFb regulation of transcription termination factor Xrn2 revealed by a chemical genetic screen for Cdk9 substrates. *Genes Dev* **30:** 117–131. doi:10.1101/gad.269589.115

Scheres SH. 2014. Beam-induced motion correction for sub-megadalton cryo-EM particles. *Elife* **3:** e03665. doi:10.7554/eLife.03665

Schilbach S, Hantsche M, Tegunov D, Dienemann C, Wigge C, Urlaub H, Cramer P. 2017. Structures of transcription pre-initiation complex with TFIIH and Mediator. *Nature* **551:** 204–209. doi:10.1038/nature24282

Schneider EV, Böttcher J, Blaesse M, Neumann L, Huber R, Maskos K. 2011. The structure of CDK8/CycC implicates specificity in the CDK/cyclin family and reveals interaction with a deep pocket binder. *J Mol Biol* **412:** 251–266. doi:10.1016/j.jmb.2011.07.020

Scruggs BS, Gilchrist DA, Nechaev S, Muse GW, Burkholder A, Fargo DC, Adelman K. 2015. Bidirectional transcription arises from two distinct hubs of transcription factor binding and active chromatin. *Mol Cell* **58:** 1101–1112. doi:10.1016/j.molcel.2015.04.006

Sdano MA, Fulcher JM, Palani S, Chandrasekharan MB, Parnell TJ, Whitby FG, Formosa T, Hill CP. 2017. A novel SH2 recognition mechanism recruits Spt6 to the doubly phosphorylated RNA polymerase II linker at sites of transcription. *Elife* **6:** e28723. doi:10.7554/eLife.28723

Shao W, Zeitlinger J. 2017. Paused RNA polymerase II inhibits new transcriptional initiation. *Nat Genet* **49:** 1045–1051. doi:10.1038/ng.3867

Sheridan RM, Fong N, D'Alessandro A, Bentley DL. 2019. Widespread backtracking by RNA Pol II is a major effector of gene activation, 5′ pause release, termination, and transcription elongation rate. *Mol Cell* **73:** 107–118.e4. doi:10.1016/j.molcel.2018.10.031

Shetty A, Kallgren SP, Demel C, Maier KC, Spatt D, Alver BH, Cramer P, Park PJ, Winston F. 2017. Spt5 plays vital roles in the control of sense and antisense transcription elongation. *Mol Cell* **66:** 77–88.e5. doi:10.1016/j.molcel.2017.02.023

Sigurdsson S, Dirac-Svejstrup AB, Svejstrup JQ. 2010. Evidence that transcript cleavage is essential for RNA polymerase II transcription and cell viability. *Mol Cell* **38:** 202–210. doi:10.1016/j.molcel.2010.02.026

Sollier J, Stork CT, García-Rubio ML, Paulsen RD, Aguilera A, Cimprich KA. 2014. Transcription-coupled nucleotide excision repair factors promote R-loop-induced genome instability. *Mol Cell* **56:** 777–785. doi:10.1016/j.molcel.2014.10.020

Sopta M, Carthew RW, Greenblatt J. 1985. Isolation of three proteins that bind to mammalian RNA polymerase II. *J Biol Chem* **260:** 10353–10360.

Sprouse RO, Brenowitz M, Auble DT. 2006. Snf2/Swi2-related ATPase Mot1 drives displacement of TATA-binding protein by gripping DNA. *EMBO J* **25:** 1492–1504. doi:10.1038/sj.emboj.7601050

Stavreva DA, Garcia DA, Fettweis G, Gudla PR, Zaki GF, Soni V, McGowan A, Williams G, Huynh A, Palangat M, et al. 2019. Transcriptional bursting and co-bursting regulation by steroid hormone release pattern and transcription factor mobility. *Mol Cell* **75:** 1161–1177.e11. doi:10.1016/j.molcel.2019.06.042

Steinparzer I, Sedlyarov V, Rubin JD, Eislmayr K, Galbraith MD, Levandowski CB, Vcelkova T, Sneezum L, Wascher F, Amman F, et al. 2019. Transcriptional responses to IFN-γ require mediator kinase-dependent pause release and mechanistically distinct CDK8 and CDK19 functions. *Mol Cell* **76:** 485–499.e8. doi:10.1016/j.molcel.2019.07.034

Sun M, Larivière L, Dengl S, Mayer A, Cramer P. 2010. A tandem SH2 domain in transcription elongation factor Spt6 binds the phosphorylated RNA polymerase II C-terminal repeat domain (CTD). *J Biol Chem* **285:** 41597–41603. doi:10.1074/jbc.M110.144568

Taatjes DJ, Näär AM, Andel F III, Nogales E, Tjian R. 2002. Structure, function, and activator-induced conformations of the CRSP coactivator. *Science* **295:** 1058–1062. doi:10.1126/science.1065249

Tahirov TH, Babayeva ND, Varzavand K, Cooper JJ, Sedore SC, Price DH. 2010. Crystal structure of HIV-1 Tat complexed with human P-TEFb. *Nature* **465:** 747–751. doi:10.1038/nature09131

Tan S, Hunziker Y, Sargent DF, Richmond TJ. 1996. Crystal structure of a yeast TFIIA/TBP/DNA complex. *Nature* **381:** 127–134. doi:10.1038/381127a0

Tantale K, Mueller F, Kozulic-Pirher A, Lesne A, Victor JM, Robert MC, Capozi S, Chouaib R, Bäcker V, Mateos-Langerak J, et al. 2016. A single-molecule view of transcription reveals convoys of RNA polymerases and multi-scale bursting. *Nat Commun* **7:** 12248. doi:10.1038/ncomms12248

Teves SS, An L, Bhargava-Shah A, Xie L, Darzacq X, Tjian R. 2018. A stable mode of bookmarking by TBP recruits RNA polymerase II to mitotic chromosomes. *Elife* **7:** e35621. doi:10.7554/eLife.35621

Thomas MC, Chiang CM. 2006. The general transcription machinery and general cofactors. *Crit Rev Biochem Mol Biol* **41:** 105–178. doi:10.1080/10409230600648736

Thompson CM, Koleske AJ, Chao DM, Young RA. 1993. A multisubunit complex associated with the RNA polymerase II CTD and TATA-binding protein in yeast. *Cell* **73:** 1361–1375. doi:10.1016/0092-8674(93)90362-T

Tome JM, Tippens ND, Lis JT. 2018. Single-molecule nascent RNA sequencing identifies regulatory domain architecture at promoters and enhancers. *Nat Genet* **50:** 1533–1541. doi:10.1038/s41588-018-0234-5

Tomko EJ, Galburt EA. 2019. Single-molecule approach for studying RNAP II transcription initiation using magnetic tweezers. *Methods* **159-160:** 35–44. doi:10.1016/j.ymeth.2019.03.010

Tóth-Petróczy Á, Oldfield CJ, Simon I, Takagi Y, Dunker AK, Uversky VN, Fuxreiter M. 2008. Malleable machines in transcription regulation: the mediator complex. *PLoS Comput Biol* **4:** e1000243. doi:10.1371/journal.pcbi.1000243

Trowitzsch S, Viola C, Scheer E, Conic S, Chavant V, Fournier M, Papai G, Ebong IO, Schaffitzel C, Zou J, et al. 2015. Cytoplasmic TAF2-TAF8-TAF10 complex provides evidence for nuclear holo-TFIID assembly from preformed submodules. *Nat Commun* **6:** 6011. doi:10.1038/ncomms7011

Tsai KL, Sato S, Tomomori-Sato C, Conaway RC, Conaway JW, Asturias FJ. 2013. A conserved mediator-CDK8 kinase module association regulates Mediator-RNA polymerase II interaction. *Nat Struct Mol Biol* **20:** 611–619. doi:10.1038/nsmb.2549

Tsai KL, Tomomori-Sato C, Sato S, Conaway RC, Conaway JW, Asturias FJ. 2014. Subunit architecture and functional modular rearrangements of the transcriptional mediator complex. *Cell* **157:** 1430–1444. doi:10.1016/j.cell.2014.05.015

Tsai KL, Yu X, Gopalan S, Chao TC, Zhang Y, Florens L, Washburn MP, Murakami K, Conaway RC, Conaway JW, et al. 2017. Mediator structure and rearrangements required for holoenzyme formation. *Nature* **544:** 196–201. doi:10.1038/nature21393

Turunen M, Spaeth JM, Keskitalo S, Park MJ, Kivioja T, Clark AD, Mäkinen N, Gao F, Palin K, Nurkkala H, et al. 2014. Uterine leiomyoma-linked MED12 mutations disrupt mediator-associated CDK activity. *Cell Rep* **7:** 654–660. doi:10.1016/j.celrep.2014.03.047

van Ingen H, van Schaik FM, Wienk H, Ballering J, Rehmann H, Dechesne AC, Kruijzer JA, Liskamp RM, Timmers HT, Boelens R. 2008. Structural insight into the recognition of the H3K4me3 mark by the TFIID subunit TAF3. *Structure* **16:** 1245–1256. doi:10.1016/j.str.2008.04.015

van Werven FJ, van Bakel H, van Teeffelen HA, Altelaar AF, Koerkamp MG, Heck AJ, Holstege FC, Timmers HT. 2008. Cooperative action of NC2 and Mot1p to regulate TATA-binding protein function across the genome. *Genes Dev* **22:** 2359–2369. doi:10.1101/gad.1682308

Venkatesh S, Workman JL. 2015. Histone exchange, chromatin structure and the regulation of transcription. *Nat Rev Mol Cell Biol* **16:** 178–189. doi:10.1038/nrm3941

Venters BJ, Irvin JD, Gramlich P, Pugh BF. 2011. Genome-wide transcriptional dependence on conserved regions of Mot1. *Mol Cell Biol* **31:** 2253–2261. doi:10.1128/MCB.01464-10

Vermeulen M, Mulder KW, Denissov S, Pijnappel WW, van Schaik FM, Varier RA, Baltissen MP, Stunnenberg HG, Mann M, Timmers HT. 2007. Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* **131:** 58–69. doi:10.1016/j.cell.2007.08.016

Vo Ngoc L, Wang YL, Kassavetis GA, Kadonaga JT. 2017. The punctilious RNA polymerase II core promoter. *Genes Dev* **31:** 1289–1301. doi:10.1101/gad.303149.117

Vos SM, Farnung L, Boehning M, Wigge C, Linden A, Urlaub H, Cramer P. 2018a. Structure of activated transcription complex Pol II–DSIF–PAF–SPT6. *Nature* **560:** 607–612. doi:10.1038/s41586-018-0440-4

Vos SM, Farnung L, Urlaub H, Cramer P. 2018b. Structure of paused transcription complex Pol II–DSIF–NELF. *Nature* **560:** 601–606. doi:10.1038/s41586-018-0442-2

Wagner T, Merino F, Stabrin M, Moriya T, Antoni C, Apelbaum A, Hagel P, Sitsel O, Raisch T, Prumbaum D, et al. 2019. SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. *Commun Biol* **2:** 218. doi:10.1038/s42003-019-0437-z

Wang H, Curran EC, Hinds TR, Wang EH, Zheng N. 2014. Crystal structure of a TAF1-TAF7 complex in human transcription factor IID reveals a promoter binding module. *Cell Res* **24:** 1433–1444. doi:10.1038/cr.2014.148

Wang H, Dienemann C, Stützer A, Urlaub H, Cheung ACM, Cramer P. 2020. Structure of the transcription coactivator SAGA. *Nature* **577:** 717–720. doi:10.1038/s41586-020-1933-5

Watanabe T, Hayashi K, Tanaka A, Furumoto T, Hanaoka F, Ohkuma Y. 2003. The carboxy terminus of the small subunit of TFIIE regulates the transition from transcription initiation to elongation by RNA polymerase II. *Mol Cell Biol* **23:** 2914–2926. doi:10.1128/MCB.23.8.2914-2926.2003

Wei Y, Resetca D, Li Z, Johansson-Åkhe I, Ahlner A, Helander S, Wallenhammar A, Morad V, Raught B, Wallner B, et al. 2019. Multiple direct interactions of TBP with the MYC oncoprotein. *Nat Struct Mol Biol* **26:** 1035–1043. doi:10.1038/s41594-019-0321-z

Werner F. 2012. A nexus for gene expression-molecular mechanisms of Spt5 and NusG in the three domains of life. *J Mol Biol* **417:** 13–27. doi:10.1016/j.jmb.2012.01.031

Wollmann P, Cui S, Viswanathan R, Berninghausen O, Wells MN, Moldt M, Witte G, Butryn A, Wendler P, Beckmann R, et al. 2011. Structure and mechanism of the Swi2/Snf2 remodeller Mot1 in complex with its substrate TBP. *Nature* **475:** 403–407. doi:10.1038/nature10215

Xue Y, Pradhan SK, Sun F, Chronis C, Tran N, Su T, Van C, Vashisht A, Wohlschlegel J, Peterson CL, et al. 2017. Mot1, Ino80C, and NC2 function coordinately to regulate pervasive transcription in yeast and mammals. *Mol Cell* **67:** 594–607.e4. doi:10.1016/j.molcel.2017.06.029

Yakovchuk P, Gilman B, Goodrich JA, Kugel JF. 2010. RNA polymerase II and TAFs undergo a slow isomerization after the polymerase is recruited to promoter-bound TFIID. *J Mol Biol* **397:** 57–68. doi:10.1016/j.jmb.2010.01.025

Yan C, Dodd T, He Y, Tainer JA, Tsutakawa SE, Ivanov I. 2019. Transcription preinitiation complex structure and dynamics provide insight into genetic diseases. *Nat Struct Mol Biol* **26:** 397–406. doi:10.1038/s41594-019-0220-3

Yang C, Ponticelli AS. 2012. Evidence that RNA polymerase II and not TFIIB is responsible for the difference in transcription initiation patterns between *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. *Nucleic Acids Res* **40:** 6495–6507. doi:10.1093/nar/gks323

Yang F, Vought BW, Satterlee JS, Walker AK, Jim Sun Z, Watts JL, DeBeaumont R, Mako Saito R, Hyberts SG, Yang S, et al. 2006. An ARC/Mediator subunit required for SREBP control of cholesterol and lipid homeostasis. *Nature* **442:** 700–704. doi:10.1038/nature04942

Yeo M, Lin PS, Dahmus ME, Gill GN. 2003. A novel RNA polymerase II C-terminal domain phosphatase that preferentially dephosphorylates serine 5. *J Biol Chem* **278:** 26078–26085. doi:10.1074/jbc.M301791200

Yoh SM, Lucas JS, Jones KA. 2008. The Iws1:Spt6:CTD complex controls cotranscriptional mRNA biosynthesis and HYPB/Setd2-mediated histone H3K36 methylation. *Genes Dev* **22:** 3422–3434. doi:10.1101/gad.1720008

Yudkovsky N, Ranish JA, Hahn S. 2000. A transcription reinitiation intermediate that is stabilized by activator. *Nature* **408:** 225–229. doi:10.1038/35041603

Zhang Z, Boskovic Z, Hussain MM, Hu W, Inouye C, Kim HJ, Abole AK, Doud MK, Lewis TA, Koehler AN, et al. 2015. Chemical perturbation of an intrinsically disordered region of TFIID distinguishes two modes of transcription initiation. *Elife* **4:** e07777. doi:10.7554/eLife.07777

Zivanov J, Nakane T, Forsberg BO, Kimanius D, Hagen WJ, Lindahl E, Scheres SH. 2018. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7:** e42166. doi:10.7554/eLife.42166