Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Conserved RNA polymerase II initiation complex structure

Merle Hantsche and Patrick Cramer

Recent cryo-electron microscopic studies have arrived at atomic models of the core transcription initiation complex comprising RNA polymerase (Pol) II and the basal transcription factors TBP, TFIIA, TFIIB, TFIIE, and TFIIF. A detailed comparison of two independently derived yeast and human core initiation complex structures reveals that they are virtually identical, demonstrating the conservation of the basic transcription machinery amongst eukaryotes. The additional factors TFIID, TFIIH, and Mediator have been located on the periphery of the core initiation complex, providing the topology of the entire initiation assembly, which comprises approximately 70 polypeptides with a molecular weight of ~4 Megadalton.

Address
Max Planck Institute for Biophysical Chemistry, Department of Molecular Biology, Am Fassberg 11, 37077 Göttingen, Germany

Corresponding author: Cramer, Patrick (pcramer@mpibpc.mpg.de)

Current Opinion in Structural Biology 2017, **47**:17–22

This review comes from a themed issue on **Protein-nucleic acid interactions**

Edited by **Stephen Cusack** and **Christoph Mueller**

http://dx.doi.org/10.1016/j.sbi.2017.03.013

Introduction

Gene transcription by RNA polymerase (Pol) II begins with the assembly of a pre-initiation complex that comprises the basal transcription factors (TFs) IIA, -B, -D, -E, -F, and -H, and the coactivator complex Mediator [1–3]. The entire initiation complex comprises approximately 70 polypeptides and has a molecular weight of around 4 MDa (Table 1). The structural basis for Pol II initiation has been studied for over a decade, and progress has been reviewed recently [4–7].

Here we present a detailed comparison of the very recently reported and independently derived atomic models of the yeast [8••] and human [9••] core transcription initiation complexes. These core complexes lack TFIIH (10 subunits) and Mediator (25–30 subunits), and contain the TATA box-binding protein (TBP) instead of the entire 14-subunit TFIID factor. The core structures comprise 20 polypeptides and nucleic acids, and are virtually identical, forming a solid basis for future studies.

Transcription initiation involves three major steps, formation of a closed promoter complex (CC), DNA opening and formation of an open promoter complex (OC), and initiation of the RNA chain, leading to an initially transcribing complex (ITC). During CC formation, promoter DNA is first recognized, and this is achieved by TFIIA, -B, and -D, which form an upstream promoter assembly. The upstream promoter assembly then recruits the Pol II-TFIIF complex. Binding of TFIIE and TFIIH then facilitates promoter opening and OC formation, whereby TFIIH consumes ATP. RNA synthesis converts the OC to an ITC. When the RNA grows beyond a critical length, the initiation complex disassembles and an elongation complex is formed, which is responsible for processive RNA synthesis.

Here we summarize the most recent structural studies that appeared since our last review [6], and compare the two high-resolution core initiation complex structures determined in the yeast and human systems. This analysis reveals the high conservation of the core initiation machinery and suggests minor movements that accompany DNA loading into the polymerase cleft. We also describe the locations of additional factors on the periphery of the core initiation complex and future directions for research.

Conserved core initiation complex structures

Prior work provided the architecture of the initiation complex at medium resolution [6,10••,11,12••,13••]. The recent cryo-EM structures of core initiation complexes [8••,9••] now provided atomic models of the core initiation complex in yeast and human and in different functional states (Table 2). These structures were obtained with different approaches. In one case (yeast), the peripheral TFIIH was omitted, and core Mediator (cMed) was included to enhance complex assembly and stability, although cMed dissociated upon grid preparation and was not observed. In the other case (human), the PIC including TFIIH was assembled and its structure determined, but TFIIH was flexible and was excluded from a focused refinement that led to a high-resolution structure of the core initiation complex.

These independent approaches using different species led to highly similar, defined core initiation complex structures

Table 1

Components of the Pol II initiation complex

Factor Name	Yeast		Human	
	Subunits	MW (kDa)	Subunits	Size (kDa)
Pol II	12	513	12	516
TFIIA	2	46	2	54
TFIIB	1	38	1	35
TFIID				
TBP	1	27	1	38
TAFs[b]	14[a]	1170	13	1243
TFIIE	2	92	2	83
TFIIF	3[a]	129	2	87
TFIIH				
Core	7	415	7	377
Kinase subcomplex (TFIIK)	3	119	3	113
Mediator				
Core (cMed)	16	459	17	574
Tail module	5	575	7	667
Kinase module	4[c]	428	4[c]	569
Unassigned subunits[d]	–	–	3	163
TOTAL SUM	69	4011	74	4519

[a] TFG3 (TAF14) is a non-essential yeast-specific protein that is a subunit of TFIID, TFIIF and chromatin remodeling complexes [28].
[b] Individual TAFs may be present in various copy numbers [28].
[c] The Mediator kinase module has 4 subunits and is likely not a bona fide component of the initiation complex.
[d] For 3 Mediator subunits the module assignment is not yet clear [23].

(Figure 1). The structures reveal details about how promoter DNA is positioned and retained within the complex. The upstream promoter assembly resembled that described in the 1990s [14] and TFIIB was positioned essentially as in the binary Pol II-TFIIB complex [15••]. The upstream promoter assembly containing TBP resides over the Pol II wall and positions DNA along the upper polymerase cleft, passing between the tips of the clamp and the protrusion that line the cleft. TFIIF and TFIIE bind to opposite sides of Pol II and flank promoter DNA. TFIIF binds to the Pol II lobe as described [16,17], and TFIIE binds between the Pol II clamp and the Rpb4-Rpb7 stalk. The recently determined crystal structure of TFIIE [18] generally confirmed the models of TFIIE within the core initiation complexes obtained by cryo-EM.

Mobility and structural differences

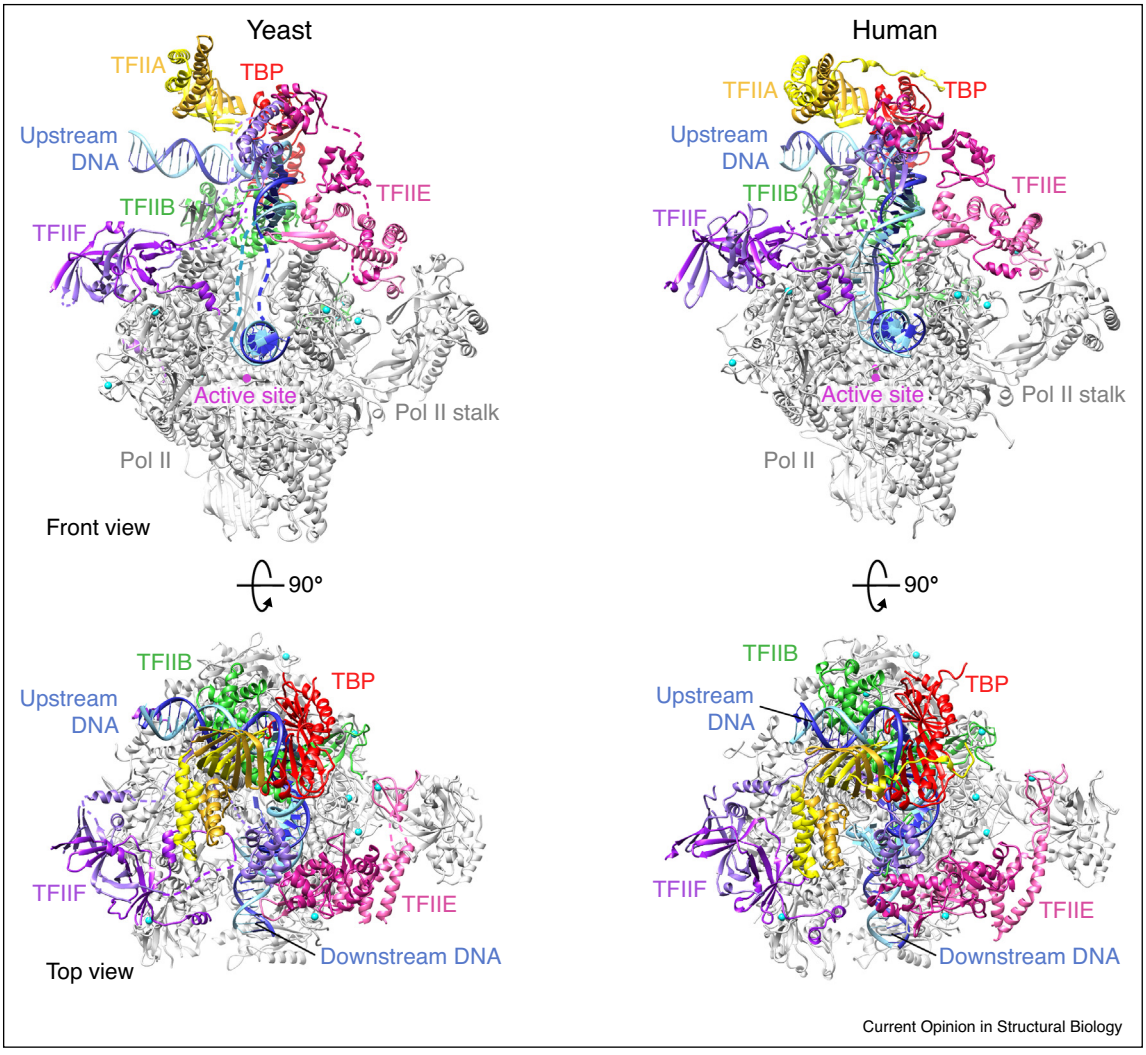
The high-resolution structures however not only confirmed the previously determined location of factors on Pol II and with respect to DNA. They also revealed previously unobserved elements of the factors, in particular mobile domains and domain linkers that adopt

Table 2

Recent cryo-EM structures of Pol II initiation complexes

Complex	Species	Resolution (Å)	EMDB code	PDB code	Reference
Core OC	Yeast	3.6	EMD-3378	5FYW	[8••]
Core CC	Yeast	8.8	EMD-3383	5FZ5	[8••]
Core ITC	Yeast	7.8	EMD-2785	4V1N	[13••]
Core ITC-cMed	Yeast	9.7	EMD-2786	4V1O	[13••]
PIC	Yeast	11	EMD-3114	5FMF	[12••]
PIC-Mediator	Yeast	15.3	EMD-8305	5SVA	[20••]
Core TFIIH	Human	10.0	EMD-8131	5IVW	[9••]
Core CC	Human	5.4	EMD-8135	5IYA	[9••]
Core OC	Human	3.9	EMD-8136	5IYB	[9••]
Core ITC	Human	3.9	EMD-8137	5IYC	[9••]
Core ITC (no IIS)	Human	3.9	EMD-8138	5IYD	[9••]
CC	Human	7.2	EMD-3307	5IY6	[9••]
OC	Human	8.6	EMD-8132	5IY7	[9••]
ITC	Human	7.9	EMD-8133	5IY8	[9••]
ITC (no IIS)	Human	6.3	EMD-8134	5IY9	[9••]
TFIID-IIA-DNA	Human	8.5	EMD-3305	5FUR	[19••]

Figure 1

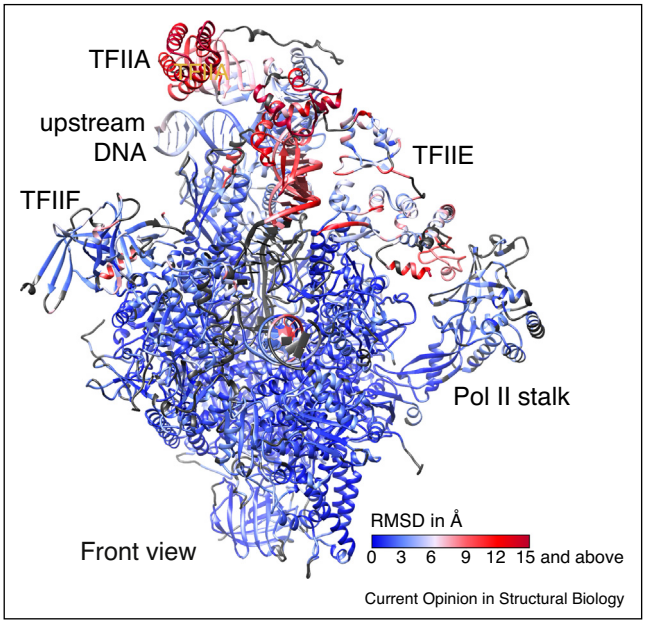


Conserved structure of the core RNA polymerase II open promoter initiation complex. Comparison of ribbon models of yeast (left, [8••]) and human (right, [9••]) complexes in front (top) and top view (bottom). Pol II is in silver and general factors are in different colors as indicated. The active site metal ion A is depicted as a magenta sphere.

a defined location only within the context of the core initiation complex. Generally, regions that are essential for cell viability are observed in the structures, whereas non-essential, non-conserved regions are often mobile. The structures also informed on details of factor interactions amongst each other and with the polymerase and DNA. The transcription factors adopt intricate structures in the context of the core complex. We speculate that these context-dependent folding transitions befit the transient nature of the complex, because their reversal is entropically favorable and can facilitate disassembly of the complex during the initiation-elongation transition.

Whereas the location of factors with respect to each other and with respect to Pol II is highly similar in the yeast and human structures, some differences in the details can be observed after superposition. These differences seem to indicate conformational flexibility of the complex because the differing elements are generally well conserved (Figure 2). There are only three notable differences between both structures. First, the exact position of the upstream complex, especially TFIIA, differs slightly, consistent with a high level of flexibility for this region in the yeast complex [8••]. Second, promoter DNA upstream of the transcription start site is moved up to 9 Å further into the polymerase cleft. As a result, the winged helix (WH) domains of TFIIEβ and TFIIFβ that contact this DNA region reside nearer the cleft in the human complex. The different DNA positions may be caused by the different promoter DNA sequences used, the

Figure 2



Structural comparison of human and yeast open complexes. Structures of the yeast and human OC [8••,9••] were superimposed by alignment of the active center region. The structure of the human OC is shown, colored by the r.m.s.d of the yeast and human models. Protein and DNA regions that are only modeled in the human structure are in grey.

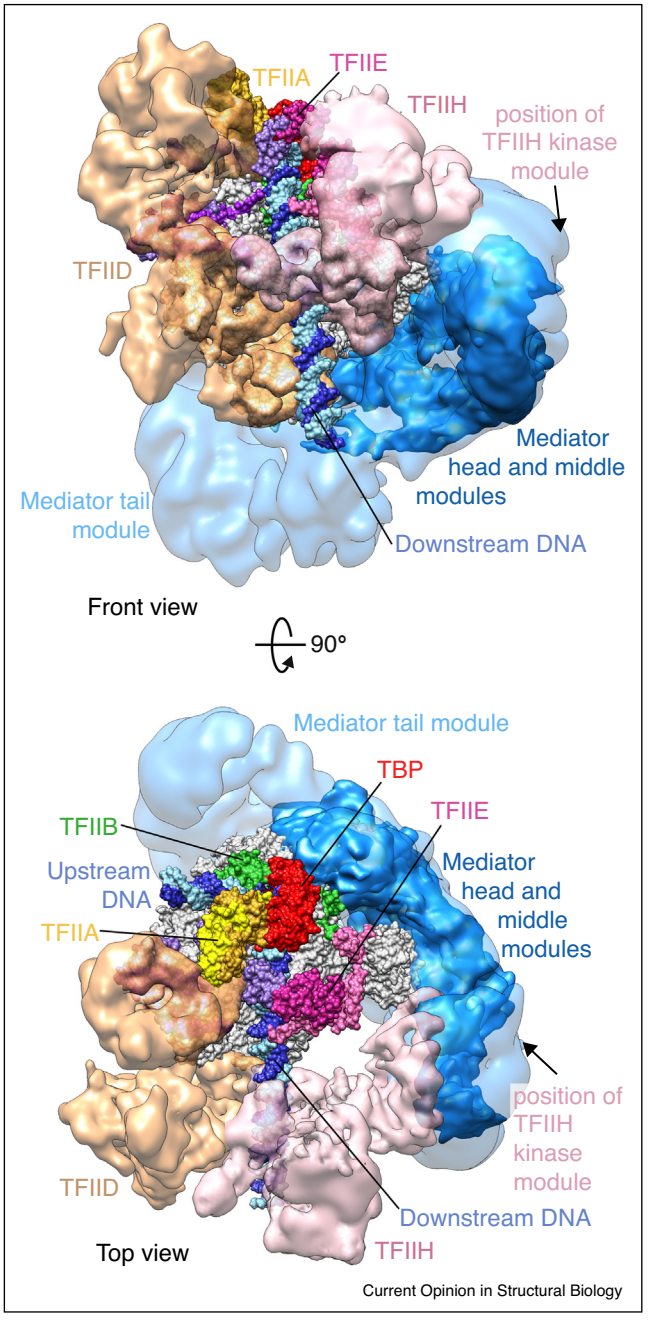
presence of TFIIH in the human complex, or different buffer conditions. Third, the TFIIB reader and linker [15••], two regions inside the Pol II cleft that are involved in transcription start site (TSS) selection and DNA positioning, are visible in the human OC, whereas they are mobile in the yeast OC. This may be due to the use of a larger mismatch bubble in the yeast DNA scaffold, which resembles the situation during TSS scanning.

Location of TFIID, TFIIH, and Mediator

The location of the multiprotein factors TFIID, TFIIH, and Mediator on the surface of the core complex has also emerged, although atomic details are still largely lacking (Figure 3). Whereas TFIID participates in promoter recognition, TFIIH is involved in DNA opening and Pol II phosphorylation, and Mediator functions in initiation complex stabilization and TFIIH kinase stimulation. The location of TFIID on promoter DNA was obtained by cryo-EM [19••], and superposition of this structure onto the core initiation complex results in a location of TFIID on the side of the Pol II lobe and protrusion. Contacts of TFIID with downstream DNA help to explain how TFIID can contribute to promoter recognition.

TFIIH was included in medium-resolution cryo-EM reconstructions of the yeast [12••,20••] and human

Figure 3



Topology of the entire initiation complex assembly. Low-resolution cryo-EM densities for TFIID (brown, [19••]), TFIIH (pink, [9••]), and Mediator (blue, [13••] and [20••]) are positioned onto the core initiation complex (surface model, [8••]). The position of the TFIIH kinase module is indicated [20]. The two views are as in Figure 1.

[9••,10••] initiation complexes. Fitting of the core initiation complex structure into these densities locates TFIIH to the downstream DNA near the polymerase jaws and clamp head. This location of TFIIH is consistent with its role in opening the DNA by translocating on downstream

DNA, thereby generating torsion that facilitates and/or maintains DNA bubble formation [21••].

Finally, Mediator was localized on the Pol II initiation complex in a medium-resolution cryo-EM reconstruction containing the core Mediator (cMed), which includes the head and middle modules of Mediator [13••]. This revealed that Mediator binds to and around the Pol II 'stalk' subcomplex Rpb4-Rpb7. Mediator also forms contacts with the TFIIB ribbon domain located on the Pol II dock domain, and with the Pol II foot domain. Contacts of Mediator to TFIIB and to Pol II regions that are required for initiation, such as the stalk, explain how Mediator stabilizes the initiation complex. A recent low-resolution cryo-EM reconstruction of a Mediator-PIC complex [20••] confirmed the location of Mediator [13••] and the proposed location of the tail module [13••,22••,23] that was predicted based on superposition of free Mediator reconstructions [24••,25••]. The work additionally revealed the TFIIH kinase module near the hook submodule in Mediator, although it remains to be investigated how exactly Mediator cooperates with the TFIIH kinase during stimulated phosphorylation of the Pol II C-terminal domain (CTD).

Stabilization of melted promoter DNA

Since structures of the conserved core initiation complex are now available in form of the CC, OC, and ITC in both species, the movements of promoter DNA during DNA opening may be inferred. For this, we assume that the basic mechanism of DNA opening is conserved, a fair assumption provided that the structure of the core initiation complex is so highly conserved.

Comparing the human and yeast CC reveals that closed promoter DNA can adopt different positions, on top of the cleft (yeast) or slightly penetrating the cleft (human). The more penetrating DNA position in the human system is possible because the Pol II clamp region assumes a slightly more open conformation thereby widening the cleft. Previous studies in the bacterial and archaeal systems have shown that clamp opening is involved in promoter melting [26,27]. The opening of the clamp brings the tip of the clamp helices closer to the α3-helix of the extended WH (eWH) of TFIIE. This contact is only observed after DNA has been melted. In both OC structures the eWH domain moves towards the cleft compared to the CC structures. In this new position the eWH domain contacts the tip of the clamp helices and may stabilize the upstream junction of the DNA bubble, thereby counteracting DNA re-closure.

Conclusions

During the last years, rapid and exciting progress has been made in understanding the structure of the initiation complex and the mechanisms that Pol II uses to start gene transcription. This was in part possible due to improved methods to obtain multiprotein factors by co-expression of recombinant subunits and purification from natural sources as endogenous complexes. In addition, the dramatic improvement in cryo-EM technology, in particular the advent of direct electron detectors, enabled high-resolution structure determination without the enormous challenge of growing crystals of these very transient assemblies. We note however that despite these advances, the resolution obtained by cryo-EM for many regions in initiation complexes is still insufficient for building reliable atomic models *de novo*, and rather required fitting and extension of known high-resolution crystal structures for Pol II and domains and subcomplexes of the transcription factors.

As a consequence of these advances, within only 7 years from the initial model of the CC and OC based on the Pol II-TFIIB crystal structure [15••] we have learned where all the other initiation factors are located. The community also obtained details on factor and nucleic acid interactions within the core initiation complex, and we came up with proposals for how different factors exert their function during initiation and the initiation–elongation transition. Structural biologists will however not stop to work on this question until complete atomic models of Pol II initiation intermediates will be available, and the mechanisms are understood at a chemical level.

Acknowledgements

We would like to thank Sarah Sainsbury, Christian Dienemann and Paulina Seweryn for discussion. MH was supported by the Deutsche Forschungsgemeinschaft (GRK1721). PC was supported by the Deutsche Forschungsgemeinschaft (SFB860, SPP1935) the European Research Council Advanced Investigator Grant TRANSREGULON (grant agreement No 693023), and the Volkswagen Foundation. We apologize to all those scientists whose publications could not be cited due to the narrow focus of this short review.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- •• of outstanding interest

1. Grünberg S, Hahn S: **Structural insights into transcription initiation by RNA polymerase II**. *Trends Biochem. Sci.* 2013, **38** http://dx.doi.org/10.1016/j.tibs.2013.1009.1002.
2. Liu X, Bushnell DA, Kornberg RD: **RNA polymerase II transcription: structure and mechanism**. *Biochim. Biophys. Acta* 2013, **1829**:2-8.
3. Roeder R: **The role of general initiation factors in transcription by RNA polymerase II**. *Trends Biochem. Sci.* 1996, **21**:327-335.
4. Hantsche M, Cramer P: **The structural basis of transcription: 10years after the nobel prize in chemistry**. *Angew. Chem. Int. Ed. Engl.* 2016, **55**:15972-15981.
5. Nogales E, Louder RK, He Y: **Cryo-EM in the study of challenging systems: the human transcription pre-initiation complex**. *Curr. Opin. Struct. Biol.* 2016, **40**:120-127.
6. Sainsbury S, Bernecky C, Cramer P: **Structural basis of transcription initiation by RNA polymerase II**. *Nat. Rev. Mol. Cell Biol.* 2015, **16**:129-143.
7. Han Y, He Y: **Eukaryotic transcription initiation machinery visualized at molecular level**. *Transcription* 2016, **7**:203-208.

8. •• Plaschka C, Hantsche M, Dienemann C, Burzinski C, Plitzko J, Cramer P: **Transcription initiation complex structures elucidate DNA opening**. *Nature* 2016, **533**:353-358.
Cryo-EM study providing near-atomic structures of the yeast CC and OC and a model for promoter opening.

9. •• He Y, Yan C, Fang J, Inouye C, Tjian R, Ivanov I, Nogales E: **Near-atomic resolution visualization of human transcription promoter opening**. *Nature* 2016, **533**:359-365.
Cryo-EM structures of the human CC, OC and ITC provide near-atomic models for the individual states of transcription initiation and elucidate DNA opening.

10. •• He Y, Fang J, Taatjes DJ, Nogales E: **Structural visualization of key steps in human transcription initiation**. *Nature* 2013, **495**:481-486.
Cryo-EM studies revealing the first three-dimensional reconstruction of the human initiation complex

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

12. •• Murakami K, Tsai K-L, Kalisman N, Bushnell DA, Asturias FJ, Kornberg RD: **Structure of an RNA polymerase II preinitiation complex**. *Proc. Natl. Acad. Sci. U. S. A.* 2015, **112**:13543-13548.
Medium-resolution yeast preinitiation complex cryo-EM reconstruction reveals location of TFIIH on downstream DNA that is consistent with the human counterpart.

13. •• Plaschka C, Lariviere L, Wenzeck L, Seizl M, Hemann M, Tegunov D, Petrotchenko EV, Borchers CH, Baumeister W, Herzog F *et al.*: **Architecture of the RNA polymerase II-Mediator core initiation complex**. *Nature* 2015, **518**:376-380.
Cyro-EM study reveals how the core Mediator, comprising the head and middle modules, contact the PIC.

14. Burley SK: **X-ray crystallographic studies of eukaryotic transcription initiation factors**. *Philos. Trans. R. Soc. Lond. Ser. B: Biol. Sci.* 1996, **351**:483.

15. •• Kostrewa D, Zeller ME, Armache K-J, Seizl M, Leike K, Thomm M, Cramer P: **RNA polymerase II-TFIIB structure and mechanism of transcription initiation**. *Nature* 2009, **462**:323-330.
X-ray crystallographic study of Pol II-TFIIB leads to models for closed and open promoter complexes.

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

17. Eichner J, Chen HT, Warfield L, Hahn S: **Position of the general transcription factor TFIIF within the RNA polymerase II transcription preinitiation complex**. *EMBO J.* 2009, **29**:706.

18. Miwa K, Kojima R, Obita T, Ohkuma Y, Tamura Y, Mizuguchi M: **Crystal structure of human general transcription factor TFIIE at atomic resolution**. *J. Mol. Biol.* 2016, **428**:4258-4266.

19. •• Louder RK, He Y, López-Blanco JR, Fang J, Chacón P, Nogales E: **Structure of promoter-bound TFIID and model of human pre-initiation complex assembly**. *Nature* 2016, **531**:604-609.
Cryo-EM reconstruction of TFIID together with promoter DNA explains how TFIID interacts with promoter elements.

20. •• Robinson Philip J, Trnka Michael J, Bushnell David A, Davis Ralph E, Mattei P-J, Burlingame Alma L, Kornberg Roger D: **Structure of a complete Mediator-RNA polymerase II pre-initiation complex**. *Cell* 2016, **166** 1411–1422.e16.
Cryo-EM study and protein cross-linking of the yeast initiation complex including TFIIH and Medaitor locate th the TFIIH kinase subcomplex.

21. •• Fishburn J, Tomko E, Galburt E, Hahn S: **Double-stranded DNA translocase activity of transcription factor TFIIH and the mechanism of RNA polymerase II open complex formation**. *Proc. Natl. Acad. Sci. U. S. A.* 2015, **112**:3961-3966.
Biochemical experiments indicate that TFIIH contains a translocase anthat acts on downstream DNA to push DNA into the active center cleft of Pol II.

22. •• Jeronimo C, Langelier M-F, Bataille Alain R, Pascal John M, Pugh BF, Robert F: **Tail and kinase modules differently regulate core Mediator recruitment and function in vivo**. *Mol. Cell.* 2016, **64**:455-466.
Genome-wide analysis of Mediator modules shows that the tail and kinase modules have regulatory functions.

23. Plaschka C, Nozawa K, Cramer P: **Mediator architecture and RNA polymerase II interaction**. *J. Mol. Biol.* 2016, **428**:2569-2574.

24. •• Tsai K-L, Tomomori-Sato C, Sato S, Conaway Ronald C, Conaway Joan W, Asturias Francisco J: **Subunit architecture and functional modular rearrangements of the transcriptional Mediator complex**. *Cell* 2014, **157**:1430-1444.
Revised architecture of yeast and human Mediator based on cryo-EM.

25. •• Wang X, Sun Q, Ding Z, Ji J, Wang J, Kong X, Yang J, Cai G: **Redefining the modular organization of the core Mediator complex**. *Cell Res.* 2014, **24**:796-808.
Revised yeast Mediator architecture based on cryo-EM.

26. Chakraborty A, Wang D, Ebright YW, Korlann Y, Kortkhonjia E, Chowdhury S, Wigneshweraraj S, Irschik H, Jansen R, Nixon BT *et al.*: **Opening and closing of the bacterial RNA Polymerase clamp: single-molecule fluorescence experiments define RNA polymerase clamp conformation in transcription initiation and elongation**. *Science* 2012, **337**:591-595.

27. Schulz S, Gietl A, Smollett K, Tinnefeld P, Werner F, Grohmann D: **TFE and Spt4/5 open and close the RNA polymerase clamp during the transcription cycle**. *Proc. Natl. Acad. Sci. U. S. A.* 2016, **113**:E1816-E1825.

28. Sanders SL, Garbett KA, Weil PA: **Molecular characterization of *Saccharomyces cerevisiae* TFIID**. *Mol. Cell Biol.* 2002, **22**:6000-6013.