Published in final edited form as:
Science. 2010 January 8; 327(5962): 206. doi:10.1126/science.1182015.

Structure of an RNA Polymerase II–TFIIB Complex and the Transcription Initiation Mechanism

Xin Liu, **David A. Bushnell**, **Dong Wang**, **Guillermo Calero**, and **Roger D. Kornberg***
Department of Structural Biology, Stanford University School of Medicine, Stanford, CA 94305, USA

Abstract

Previous X-ray crystal structures have given insight into the mechanism of transcription and the role of general transcription factors in the initiation of the process. A previous structure at 4.5 Å resolution of an RNA polymerase II–general transcription factor TFIIB complex revealed the N-terminal region of TFIIB, including a loop termed the "B-finger" reaching into the active center of the polymerase where it may interact with both DNA and RNA, but this structure showed little of the C-terminal region. A new crystal structure of the same complex at 3.8 Å resolution obtained under different solution conditions is complementary with the previous one, revealing the C-terminal region of TFIIB, located above the polymerase active center cleft, but showing none of the B-finger. In the new structure, the linker between the N- and C-terminal regions can also be seen, snaking down from above the cleft towards the active center. The two structures, taken together with others previously obtained, dispel longstanding mysteries of the transcription initiation process.

Cellular RNA polymerases require protein cofactors for promoter recognition and the initiation of transcription. In bacteria, this requirement is met by a single protein, the sigma factor (1). By contrast, RNA polymerase II (pol II) of eukaryotes depends on five "general" factors, comprising some 30 polypeptides, for promoter-dependent transcription. The general factors, known as TFIIB, -D, -E, -F, and -H, assemble with the polymerase and promoter DNA in a complex of approximately 2 MDa at every round of the initiation of transcription. Promoters containing a TATA box may be transcribed with only the TATA-binding protein (TBP) subunit of TFIID, whereas TATA-less promoters require the TBP-associated factor (TAF) subunits of TFIID as well. TFIIB and TBP/TFIID are primarily responsible for promoter recognition; indeed, TFIIB and TBP are alone sufficient for pol II transcription of a negatively supercoiled promoter *in vitro* (2). In the absence of supercoiling, TFIIE and TFIIH are required to introduce negative superhelical strain and unwind promoter DNA for the initiation of transcription. Structural studies of pol II, both alone and as an actively transcribing complex, have revealed a large conformational change between the "closed" complex containing entirely double stranded promoter DNA and the "open" complex containing an unwound region ("transcription bubble"). The promoter DNA is straight in the closed complex, but in the open complex it bends about 90° and descends some 30 A into the central polymerase cleft. The mechanism of this large conformational change has remained unclear. The general factors are believed to assist and to remain associated throughout the process, but following initiation, they are released and the polymerase escapes from the promoter. The challenge is to understand how protein-protein interactions can be formed during the assembly of the transcription initiation complex and then reversed during promoter escape.

Biochemical and genetic studies have implicated TFIIB in start site selection, and in the stabilization of the initial transcript. A previous structure of a pol II TFIIB complex, obtained

*To whom correspondence should be addressed. kornberg@stanford.edu.

from crystals grown in the presence of 800 mM ammonium sulfate, revealed the N-terminal region of TFIIB but little of the cyclin repeats that make up the C-terminal "core" of the protein (3). A notable finding was a loop, termed the "B-finger," reaching into the pol II active center. As the B-finger would clash with nascent RNA beyond about five residues, it might be involved in the decision between abortive transcription (release of a short nascent transcript and re-initiation) and promoter escape. We now report a significantly different structure of the pol II TFIIB complex, obtained from crystals grown from solution containing 1.2 M sodium/potassium phosphate. The structure was solved by molecular replacement with a model of pol II in the "clamp-closed" conformation. A difference (F_o-F_c) map between the pol II TFIIB structure determined here and that of pol II alone revealed the zinc ribbon domain in the N-terminal region of TFIIB, and additional electron density attributable to the linker and C-terminal core domains of TFIIB, but no density due to the B-finger (Fig. 1). The zinc ribbon domain was in essentially the same conformation as in the previous cocrystal structure (3), and was in the same location, in contact with the "dock" domain of pol II.

A homology model of the yeast TFIIB core was constructed from the published human TFIIB core structure (4,5) and was manually placed in the $F_o - F_c$ map (Fig. 1B), resulting in a good fit of the first cyclin repeat, consisting of a five-helix bundle (BH1-BH5). The location of the cyclin repeat was confirmed by a match to anomalous signals from Se-Met at four sites (Fig. 1B: M135, M168, M172 and M210), and was manually adjusted for a best fit to the anomalous signals. An additional α-helix (BH0) was then built connected to BH1 and was validated by anomalous signals from Se-Met at one intrinsic (M104) and one mutant (L110M) site (Fig. 1B). The rest of the linker domain was built with the aid of anomalous signals from Se-Met at two additional mutant sites (S83M and R95M).

Mutations in the B-finger (F66M, D69M, N72M, D75M and V79M) gave no Se-Met anomalous signals, indicating a high degree of mobility of the B-finger. The disordered region (residues 67-80) corresponds to the tip of the B-finger, and an analogous region in bacterial σ factor (σ_3-σ_4 linker or $\sigma_{3.2}$) is also partially disordered at a comparable resolution (6). Residues of the linker domain adjacent to the disordered region are located at the ends of the previous B-finger electron density (fig. S1), showing consistency with the previous results and suggesting that the B finger, although mobile, projects into the pol II active center as previously observed.

The N-terminal half of the linker domain winds past loops of Rpb1 and Rpb2 including the "lid," "rudder," and "fork loop 1" (Fig. 2A). Turn regions connecting on the N-terminal side to the B-finger and on the C-terminal side to the other half of the linker (BH0 helix) contain conserved glycines and prolines (fig. S2). The BH0 helix is the most exposed part of the linker, lying on the surface of the Rpb1 clamp (Fig. 2B). The association of the linker domain with Rpb1 and Rpb2 loops and clamp may contribute to the stabilization of the loop conformations and to clamp closure (the loops are disordered in the structure of pol II alone).

The TFIIB core domain is located above the pol II cleft, about 50 A from the N-terminal zinc-ribbon domain (Fig. 1A). The first cyclin repeat interacts with both the "protrusion" and the "wall" of Rpb2 (Fig. 2B), consistent with previous cross-linking and FeBABE cleavage mapping results (7,8). A large part of the protrusion is ordered by this interaction (fig. S3), but due to the presence of β-strands and loops, a model of the ordered region could not be built. Interaction of BH2 with the top of the wall (fig. S3) may explain why yeast TFIIB is specific for yeast pol II (9). A mutation (C149Y) in the BH2-wall interaction interface (fig. S3B) causes a shift of the transcription start site at *ADH1* in yeast (10). There was no evidence of the second cyclin repeat in the electron density map, presumably due to motion in absence of TBP and promoter DNA, which may be pronounced for yeast TFIIB because it contains a long inter-repeat linker (11) (fig. S2). The two cyclin repeats may perform distinctive roles, as a TFIIB-

like protein, which interacts with pol II and is required for transcription of sliced leader RNA in the human parasite *Trypanosoma brucei*, shows sequence homology to the first but not the second cyclin repeat (12). A recent publication that shows a structure of the pol II - TFIIB complex at 4.3 A resolution in general agrees with the structure reported here (13).

The topography of the pol II TFIIB complex is similar to that of a bacterial RNA polymerase sigma factor complex (6,14). As previously noted, the zinc ribbon and B-finger interact with similar polymerase surfaces as the σ_4 domain and the σ_3 - σ_4 linker. We now find that the TFIIB linker and core interact with similar polymerase surfaces as σ_2 and σ_3. There are similarities of the TFIIB and sigma factor structures as well. The first cyclin repeat of TFIIB aligns with the σ_3 domain, in particular with the $\sigma_{3.1}$ helix (Fig. 3), which contacts DNA downstream of the -10 element of bacterial promoters, facilitating transcription bubble formation (15,16). The TFIIB core makes sequence-specific contacts with promoter DNA, both upstream (BRE^u) and downstream (BRE^d) of the TATA box (17,18), and the solvent-exposed BH0-BH1 region may interact with the DNA downstream of BRE^d, in a manner analogous to the $\sigma_{3.1}$ helix - DNA interaction. Thus TFIIB may recapitulate two important functions of sigma factor in the initiation of transcription, that of the σ_3-σ_4 linker, which contacts the start site, and that of the $\sigma_{3.1}$ helix, which fixes the upstream edge of the transcription bubble. TFIIB may also serve as a scaffold for binding other transcription factors that perform functions attributed to sigma factor in bacteria. For example, some regions of the TFIIF subunit Tfg2, which shows sequence similarity to sigma and has been suggested to perform analogous functions, have been placed by cryo-electron microscopy (cryo-EM) in locations near that of the first cyclin repeat of TFIIB shown here (fig. S4) (19).

The first cyclin repeat in the pol II TFIIB cocrystal structure could be aligned with the same repeat in a previous structure of C-terminal fragments of human TFIIB and TBP bound to TATA box DNA (fig. S5). Docking the structures resulted in a model of a “minimal” pre-initiation complex (Fig. 4A, B). The second cyclin repeat in the model interacts with the Rpb2 protrusion and Rpb12, consistent with cross-linking and FeBABE cleavage mapping studies (7,8). When the TATA box DNA fragment in the model was extended with straight B-form DNA, to simulate the “closed” form of the promoter, only minor adjustments were needed, to avoid steric clashes with the clamp (Fig. 4A, B; see SOM). The DNA follows a path above the central cleft leading to the polymerase active center. The structure of the closed promoter complex could be modified with the nucleic acid region of a transcribing complex (20-23) to model an open promoter complex (Fig. 4C, fig. S6; see SOM). Extension of the template strand upstream of the transcription bubble in this model leads through a “tunnel” formed by TFIIB and pol II to the location of duplex DNA in the closed promoter model (dashed cyan line in Fig. 4C, location of tunnel indicated in Fig. 4D).

The formation of the template strand tunnel by TFIIB gives insight into the closed to open promoter transition. It suggests a straightforward mechanism for the process. Transient promoter melting may be captured by interaction of the template strand with the tunnel. The orientation and proximity of the promoter DNA to the upstream end of the tunnel provide a pathway whereby progressive melting and short-range interactions can lead, sequentially, to a large conformational change.

Biochemical studies in a mammalian system have demonstrated transcription bubble formation beginning about 20 bp downstream of the TATA box and extending, as transcription proceeds, to a bubble size of 18 bp and transcript length of about 7 residues, upon which the upstream 8 bp of the bubble reanneal (“bubble collapse”) (24-26). The template single strand between the growing DNA-RNA hybrid and the TFIIB tunnel presumably loops out, or “scrunches,” as previously described (dashed orange line in Fig. 4D) (27,28). Bubble collapse involves the region passing through the TFIIB tunnel (dashed cyan line in Fig. 4D), as well as the scrunched

strand. Following the reannealing of 8 bp, the upstream edge of the bubble is located adjacent to the Rpb1 rudder and lid, as observed in the structure of a transcribing complex. The open promoter model thus accounts for bubble expansion and collapse.

Biochemical studies have further demonstrated a role for the B-finger in bubble collapse and promoter escape (24,29). Interaction with the B-finger has been shown by the stabilizing effect of TFIIB on a 5-residue transcribing complex (3). Mutation of the B-finger eliminates a pause in transcription at 7 residues, observed with a “premelted” DNA template (24). The B finger may play multiple roles, both stabilizing a 5-residue transcribing complex and clashing with RNA in a 7-residue complex. These roles of the B finger may relate to the alternative states revealed by crystallography, one with a well localized B-finger but no linker or cyclin repeat, and the alternative containing a well localized linker and cyclin repeat but no B finger. These states are evidently similar in energy and may interconvert during the initiation of transcription: initiation occurs in the state depicted in the open promoter model, with the template strand bound by the linker and cyclin repeat; upon reaching about 5 residues in length, the nascent transcript contacts the B-finger, forming a stable complex (3), tipping the balance between states, flipping from that reported here to that obtained previously, in which the linker and cyclin repeat are mobile and the B finger is observed; extension of the transcript beyond 7 residues results in a clash with the B finger, displacing it as well, thereby completing the process of promoter escape. The requirement for disruption of protein-protein interactions that would seem to impede promoter escape is solved by substituting one set of contacts (B-finger pol II and B-finger RNA) for another (linker pol II and cyclin repeat pol II). Disruption of TFIIB-pol II interaction occurs sequentially. Meanwhile, the TFIIB tunnel maintains the transcription bubble until the DNA-RNA hybrid is long enough to persist on its own. The stability conferred upon the open promoter complex by the TFIIB tunnel is ultimately replaced by the stability of an RNA DNA hybrid. The entire process of transcription initiation, from initial promoter melting to promoter escape, may be explained in this way.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References and Notes

1. Gross CA, et al. Cold Spring Harb Symp Quant Biol 1998;63:141. [PubMed: 10384278]
2. Parvin JD, Sharp PA. Cell May 7;1993 73:533. [PubMed: 8490964]
3. Bushnell DA, Westover KD, Davis RE, Kornberg RD. Science Feb 13;2004 303:983. [PubMed: 14963322]
4. Tsai FT, Sigler PB. Embo J Jan 4;2000 19:25. [PubMed: 10619841]
5. Nikolov DB, et al. Nature Sep 14;1995 377:119. [PubMed: 7675079]
6. Murakami KS, Masuda S, Darst SA. Science May 17;2002 296:1280. [PubMed: 12016306]
7. Chen HT, Hahn S. Cell Oct 15;2004 119:169. [PubMed: 15479635]
8. Chen HT, Warfield L, Hahn S. Nat Struct Mol Biol Aug;2007 14:696. [PubMed: 17632521]
9. Shaw SP, Wingfield J, Dorsey MJ, Ma J. Mol Cell Biol Jul;1996 16:3651. [PubMed: 8668181]
10. Wu WH, Pinto I, Chen BS, Hampsey M. Genetics Oct;1999 153:643. [PubMed: 10511545]
11. Hayashi F, et al. Biochemistry Jun 2;1998 37:7941. [PubMed: 9609687]
12. Schimanski B, Brandenburg J, Nguyen TN, Caimano MJ, Gunzl A. Nucleic Acids Res 2006;34:1676. [PubMed: 16554554]
13. Kostrewa D, et al. Nature. 200910.1038/nature08548
14. Vassylyev DG, et al. Nature Jun 13;2002 417:712. [PubMed: 12000971]
15. Murakami KS, Darst SA. Curr Opin Struct Biol Feb;2003 13:31. [PubMed: 12581657]

16. Murakami KS, Masuda S, Campbell EA, Muzzin O, Darst SA. Science May 17;2002 296:1285. [PubMed: 12016307]
17. Lagrange T, Kapanidis AN, Tang H, Reinberg D, Ebright RH. Genes Dev Jan 1;1998 12:34. [PubMed: 9420329]
18. Deng W, Roberts SG. Genes Dev Oct 15;2005 19:2418. [PubMed: 16230532]
19. Chung WH, et al. Mol Cell Oct;2003 12:1003. [PubMed: 14580350]
20. Gnatt AL, Cramer P, Fu J, Bushnell DA, Kornberg RD. Science Jun 8;2001 292:1876. [PubMed: 11313499]
21. Westover KD, Bushnell DA, Kornberg RD. Cell Nov 12;2004 119:481. [PubMed: 15537538]
22. Westover KD, Bushnell DA, Kornberg RD. Science Feb 13;2004 303:1014. [PubMed: 14963331]
23. Korzheva N, et al. Science Jul 28;2000 289:619. [PubMed: 10915625]
24. Pal M, Ponticelli AS, Luse DS. Mol Cell Jul 1;2005 19:101. [PubMed: 15989968]
25. Holstege FC, Fiedler U, Timmers HT. EMBO J Dec 15;1997 16:7468. [PubMed: 9405375]
26. Giardina C, Lis JT. Science Aug 6;1993 261:759. [PubMed: 8342041]
27. Revyakin A, Liu C, Ebright RH, Strick TR. Science Nov 17;2006 314:1139. [PubMed: 17110577]
28. Kapanidis AN, et al. Science Nov 17;2006 314:1144. [PubMed: 17110578]
29. Tran K, Gralla JD. J Biol Chem Jun 6;2008 283:15665. [PubMed: 18411280]
30. This research was supported by NIH grants GM049985 and AI21144 to R.D.K. X.L. was supported by the Jane Coffin Childs Memorial Fund fellowship. D.W. was supported by the NIH Pathway to Independence Award (K99 GM085136). Portions of this research were carried out at the Stanford Synchrotron Radiation Laboratory, a national user facility operated by Stanford University on behalf of the U.S. Department of Energy, Office of Basic Energy Sciences. The SSRL Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological and Environmental Research, and by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program, and the National Institute of General Medical Sciences. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. Use of the Advanced Photon Source was supported by the U. S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357. Coordinates and structure factors have been deposited at the Protein Data Bank under accession code 3K7A.

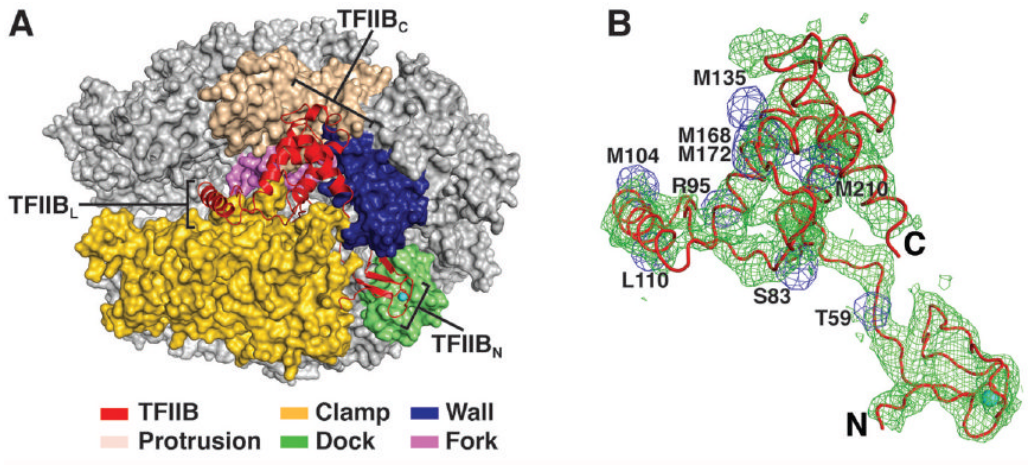

Fig. 1.

Structure of pol II TFIIB complex. (**A**) "Top" view of pol II in a surface representation, with previously identified domains in the colors indicated, and with TFIIB in ribbon representation. TFIIB zinc ribbon (TFIIB$_N$), linker (TFIIB$_L$), first cyclin repeat (TFIIB$_C$) are indicated. (**B**) Difference (F_o-F_c) electron density map between pol II TFIIB and pol II alone, contoured at 2.0σ, shown in green mesh, and Se-Met anomalous peaks, contoured at 6-10σ, shown in blue mesh.

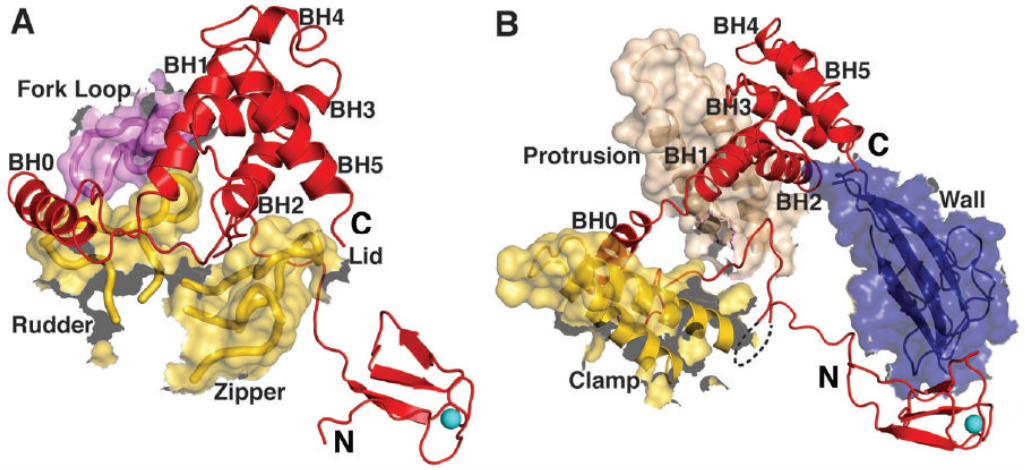

Fig. 2.

Structure of TFIIB and pol II TFIIB interactions. (**A**) Expanded view from Fig. 1A, with pol II fork loop 1, rudder, lid and zipper shown in both transparent surface and ribbon representation, and with TFIIB helices labeled. (**B**) Same as (A) rotated 45° around the X-axis, except with parts of pol II clamp, protrusion and wall shown in both transparent surface and ribbon representation. The B-finger from the previous cocrystal structure is shown as a dashed black line.

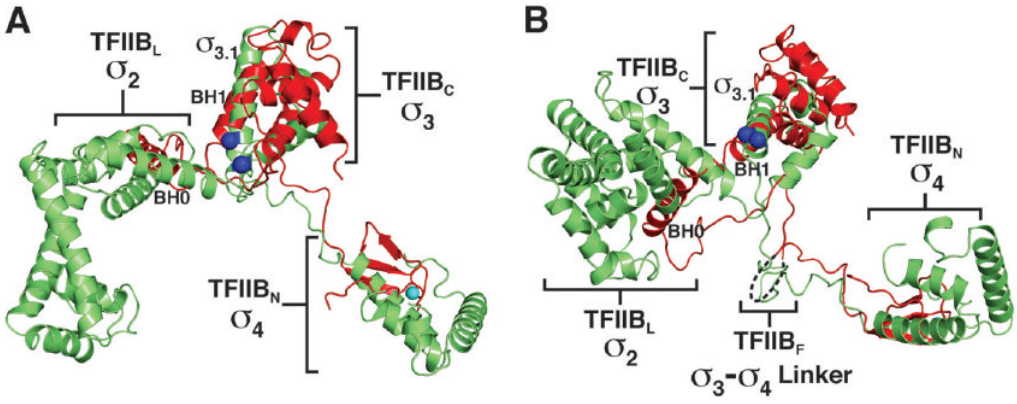

Fig. 3.

Comparison of TFIIB and bacterial σ factor structures. (**A**) Superposition of TFIIB (red) and σ factor (green) structures. Conserved residues H455 and E458 of σ factors that bind to the -10 element and mark the start of transcription bubble formation are highlighted as blue spheres. Corresponding domains from TFIIB and σ factor are labeled. (**B**) Same as (A) rotated 45° around the X-axis. The B-finger (TFIIB$_F$) from the previous cocrystal structure is shown as a dashed black line.

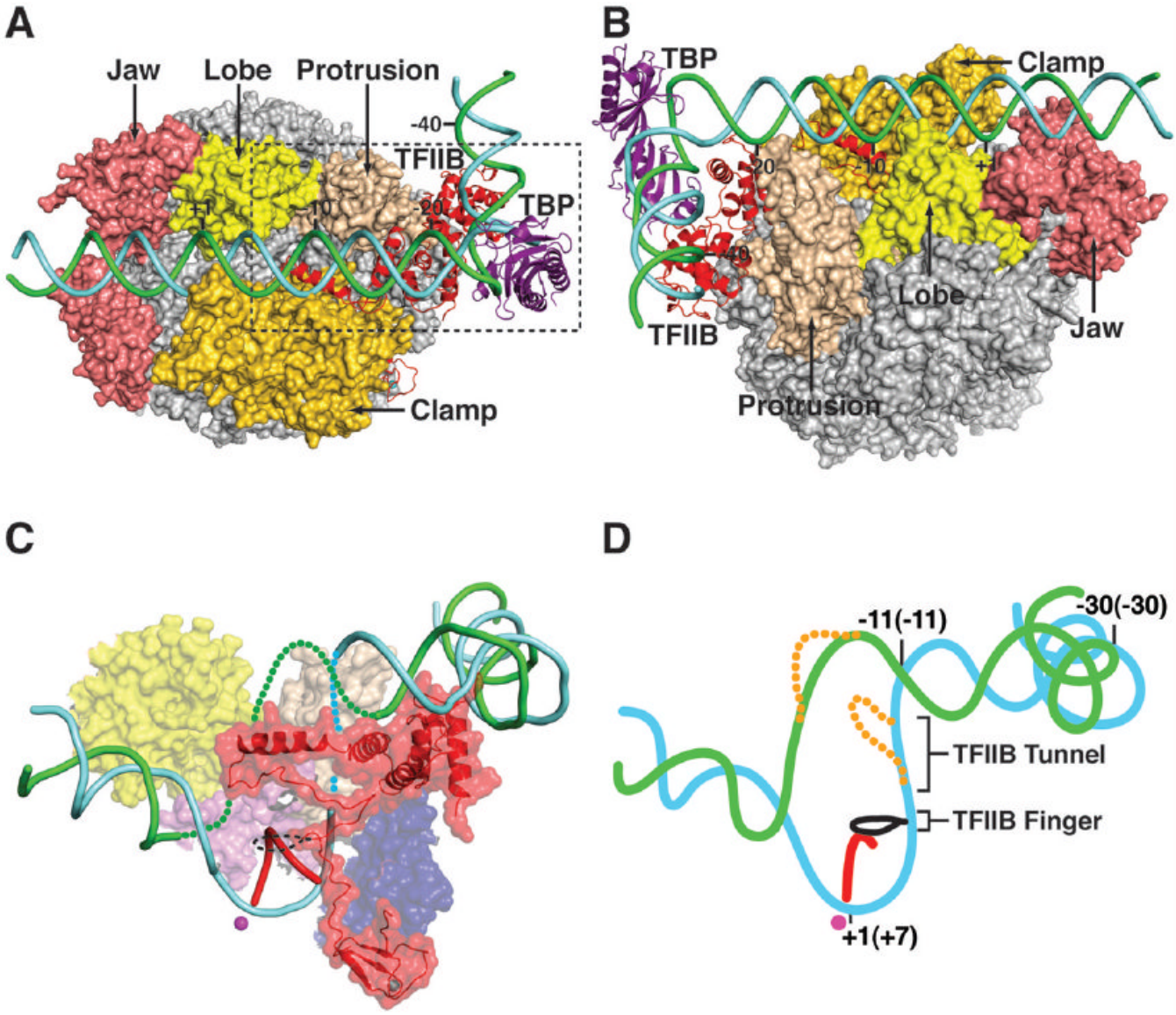

Fig. 4.

Models of promoter complexes. (**A**) Model of closed promoter complex based upon the alignment in fig. S4, viewed as in Fig. 1A rotated 45° around the Y-axis. Yeast $TFIIB_C$ and the second cyclin repeat of human $TFIIB_C$ are shown in red and TBP_C is shown in purple. DNA upstream and downstream of the TATA box was extended with B-form double helix. A slight distortion was introduced in the template strand to avoid clashes with the pol II clamp. Template and nontemplate strands are shown in cyan and green. (**B**) Same as (A) rotated 180° about Y-axis and 120° about X-axis. (**C**) Model of open promoter complex. Expected path of template strand, connecting the upstream edge of the transcription bubble seen in the structure of a transcribing complex (24) with duplex promoter DNA from the closed promoter complex, is shown as a dashed cyan line. Possible path of nontemplate strand is shown as a dashed green line. View is of the dashed boxed area in fig. S6A, rotated 45° around the X-axis. The pol II clamp has been removed for clarity. TFIIB is shown as a red transparent surface with the location of the B-finger represented by a black dashed line. The nascent RNA transcript and active site Mg^{2+} are shown as a red ribbon and a magenta sphere. (**D**) Cartoon of transcription

initiation complex, based on nucleic acid structure in (C), with template DNA positions at the start of transcription indicated, and those following synthesis of a 7-residue transcript in parentheses. The “scrunched” template strands resulting from synthesis of the 7-residue transcript are shown as dashed orange lines.