



Review

RNA polymerase II transcription: Structure and mechanism[☆]

Xin Liu¹, David A. Bushnell, Roger D. Kornberg^{*}

Department of Structural Biology, Stanford University School of Medicine, Stanford, CA 94305, USA

ARTICLE INFO

Article history:

Received 8 August 2012

Accepted 7 September 2012

Available online 18 September 2012

Keywords:

TFIIB

Trigger loop

Bridge helix

Backtracked state

Hybrid helix melting

ABSTRACT

A minimal RNA polymerase II (pol II) transcription system comprises the polymerase and five general transcription factors (GTFs) TFIIB, -D, -E, -F, and -H. The addition of Mediator enables a response to regulatory factors. The GTFs are required for promoter recognition and the initiation of transcription. Following initiation, pol II alone is capable of RNA transcript elongation and of proofreading. Structural studies reviewed here reveal roles of GTFs in the initiation process and shed light on the transcription elongation mechanism. This article is part of a Special Issue entitled: RNA Polymerase II Transcript Elongation.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Structural studies began with pol II itself, an assembly of a dozen polypeptides, with a mass in excess of 500 kD. Two-dimensional crystallography showed the way to growing large single crystals of the enzyme lacking two subunits, Rpb4 and Rpb7, suitable for X-ray analysis [1,2]. An initial 5 Å phase set was obtained with the use of a large heavy atom cluster [3], leading to structures at 2.8 Å resolution of the enzyme alone [4] and at 3.3 Å of an actively transcribing complex [5]. Structures were subsequently obtained of the complete 12-subunit enzyme as well [6,7].

2. Initiation of transcription

The initiation of pol II transcription is of particular interest because of the regulation of the process. Transcriptional activators and repressors exert their effects at this early stage in gene expression to influence cellular physiology and development. It has long been known that pol II assembles with the GTFs to form a giant preinitiation complex (PIC), larger than a ribosome, prior to every round of transcription. The PIC undergoes a series of transformations as the nascent RNA grows to a length of about 25 residues, whereupon a stable elongation complex is formed and transcription of the gene ensues. A key transformation of the PIC is from a “closed” promoter complex, in which the DNA is entirely double stranded, to an “open” promoter

complex, in which some 15 base pairs (bp) of the promoter DNA are unwound to form a so-called “transcription bubble.” Subsequent events include transcription start site (TSS) selection, de novo RNA synthesis (formation of the first phosphodiester bond), abortive initiation, and promoter escape.

Crystal structures of pol II–TFIIB complexes have been informative about the initiation mechanism. The structures have revealed an N-terminal zinc-ribbon domain contacting the pol II surface, and either a loop termed the “B finger” [8] or a C-terminal “core” domain [9,10], but not both the B finger and core at the same time. There are evidently two conformations of the pol II–TFIIB complex, whose interconversion is proposed to underlie the transcription mechanism. The B-finger projects into the pol II active center cleft, whereas the core domain is bound on the surface. The B-finger is evidently flexible, not even revealed by some refinement programs [10]. A portion of the B-finger, which may form a short α -helix, has been termed the “B-reader,” referring to interaction with the template DNA [10], but in all likelihood the entire B-finger is involved in TSS recognition and the initiation process. In addition to the B-finger and core domains, the crystal structures have revealed a linker loop winding from the pol II active center cleft to the surface (Fig. 1). Despite limited sequence homology, comparison of crystal structures of pol II–TFIIB and bacterial RNA polymerase– σ factor complexes has disclosed a remarkable similarity of domain structure and topography [9–12] (Fig. 2).

The crystal structure of a complex of the TFIIB core domain, TATA-box binding protein (TBP) core domain, and a TATA-box DNA fragment was previously determined [13,14]. This structure could be docked to that of the pol II–TFIIB complex by superposition of TFIIB core domains. Extension of the TATA-box DNA with straight B-form DNA led to a model of a “minimal” closed promoter complex [9,10] (Fig. 3). Such a minimal complex is relevant to transcription, inasmuch

[☆] This article is part of a Special Issue entitled: RNA Polymerase II Transcript Elongation.

^{*} Corresponding author. Tel.: +1 650 723 6167.

E-mail address: kornberg@stanford.edu (R.D. Kornberg).

¹ Present address: Department of Biophysics and Cecil H. and Ida Green Center for Reproductive Biology Sciences, Department of Obstetrics and Gynecology, University of Texas Southwestern Medical Center, Dallas, TX, 75390, USA.

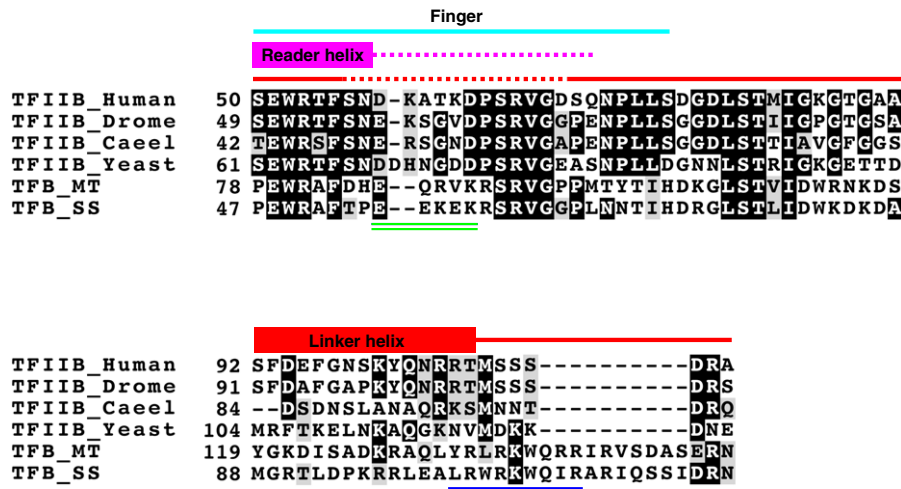


Fig. 1. Sequence alignments of the TFIIB finger and linker regions. Primary sequences of TFIIB/TFB proteins from *Homo sapiens* (TFIIB_Human), *Drosophila melanogaster* (TFIIB_Drome), *Caenorhabditis elegans* (TFIIB_Caeel), *Saccharomyces cerevisiae* (TFIIB_Yeast), *Methanococcus thermolithotrophicus* (TFB_MT) and *Sulfolobus shibatae* (TFB_SS) were aligned with the use of ClustalW2 and manually edited. Aligned sequences were highlighted with BOXSHADE. Cartoons representing the secondary structures are drawn based on the 3.8 Å pol II–TFIIB crystal structure [9] and are shown above the sequences in red. The B-finger/B-reader regions of the other two pol II–TFIIB crystal structures determined at 4.3 Å [10] and 4.5 Å [8] are also shown in purple and cyan. A fragment implicated in binding the initiating nucleotide is indicated by double green lines under the sequence alignment. A tryptophan-containing fragment in TFB located close to the upstream edge of the transcription bubble is indicated by double blue lines.

as TFIIB and TBP are alone sufficient for the TATA-directed pol II transcription with a negatively supercoiled DNA template [15]. Moreover, transcription in archaeobacteria requires only TFB and TBP, homologs of eukaryotic TFIIB and TBP [16,17]. Archaeal TFE, homologous to the N-terminal region of the TFIIE α subunit, is dispensable for transcription [18,19], and archaeal homologs of eukaryotic TFIIF and TFIH have not been identified.

A notable feature of DNA bound to TBP is bending of the double helix by nearly 90°. In the minimal closed promoter model, this bend has the consequence of wrapping the DNA around pol II. It may be for this reason that TBP is required even for transcription of genes lacking a TATA box, and for transcription by polymerases I and III as well [20,21].

In the minimal closed promoter model, the DNA downstream of the TATA box, in the direction of the TSS, lies directly above the pol II active center cleft, consistent with results of DNA–protein cross-linking [22]. Other general transcription factors may also be involved in positioning the closed promoter DNA. TFIIE, harboring a double-stranded DNA binding surface [23,24], has been located by cross-linking studies on the Rpb1 side of the active center cleft [25], in proximity to the DNA 10–20 bp downstream of the TATA box [22,26]. TFIIF, also implicated in DNA binding [27–29], has been placed in a similar location on the basis of cryo-EM [30], cross-linking, and hydroxyl radical cleavage mapping [25,31,32]. TFIH has been suggested to bind promoter DNA more than 30 bp downstream of the TATA box [22,33].

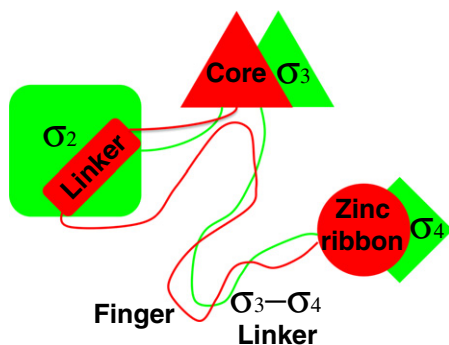


Fig. 2. Schematic alignment of TFIIB and σ factor structures. TFIIB and σ are shown in red and green. Corresponding domains of the two proteins are indicated.

TFIIB recognition elements upstream and downstream of the TATA box (BRE^u and BRE^d) serve, at least in part, as determinants for the directionality of transcription [13,34–37]. The divergent TFIIB from the parasite *Trypanosoma brucei* lacks an extensive basic surface for promoter binding, and it is also devoid of the structural and sequence features for interaction with BRE^u, which may account for the observed bidirectional transcription [38–41]. Widespread divergent transcription has recently been reported for active promoters in other species including mammals and yeast [42,43]. Neither the physiologic role nor the mechanism of this divergent transcription has been determined.

The transition from a closed to an “open” promoter complex entails melting of the double helix and descent of the template strand some 30 Å from its initial position above the active center cleft to the base of the cleft [5]. Pol II is unique among RNA polymerases (pol I, pol III, archaeal RNAP and bacterial RNAP) in its requirement for ATP-dependent helicase activity for promoter melting. Although σ^{54} -directed transcription by bacterial RNAP also requires ATP-hydrolysis, an AAA + activator rather than a helicase is involved. The activator is thought to induce domain rearrangements of σ^{54} , rather than to catalyze the melting of the promoter DNA. Domain rearrangements are necessary because σ^{54} blocks the path of the template strand to the active center [44].

The crystal structure of the pol II–TFIIB complex suggests a role for TFIIB in defining the path of the template DNA strand from atop the active center cleft to the base of the cleft. The TFIIB linker and core, as well as elements of pol II (clamp and loops), form a “tunnel,” completely enclosing the proposed path of the template strand [9,10]. This tunnel converges with the downstream duplex DNA channel and the ribonucleotide entry tunnel around the catalytic site [9,10] (Fig. 4). The first residues of the template strand liberated by promoter melting may be captured at the top of the TFIIB tunnel, followed by additional residues as the melting proceeds. Consistent with the proposed role of the TFIIB tunnel, a number of conserved lysine and arginine residues from pol II and TFIIB are located along the tunnel, apparently to facilitate DNA binding. Inasmuch as the tunnel is completely surrounded by TFIIB and pol II, movement of the TFIIB linker or core, as well as the pol II loops or clamp [4], may be required for entry of the DNA.

In bacteria, a set of conserved aromatic residues in the $\sigma_{2.3}$ region participates in promoter opening, through binding the non-template single strand around the promoter -10 element, the upstream edge of the transcription bubble [45,46]. In the absence of corresponding

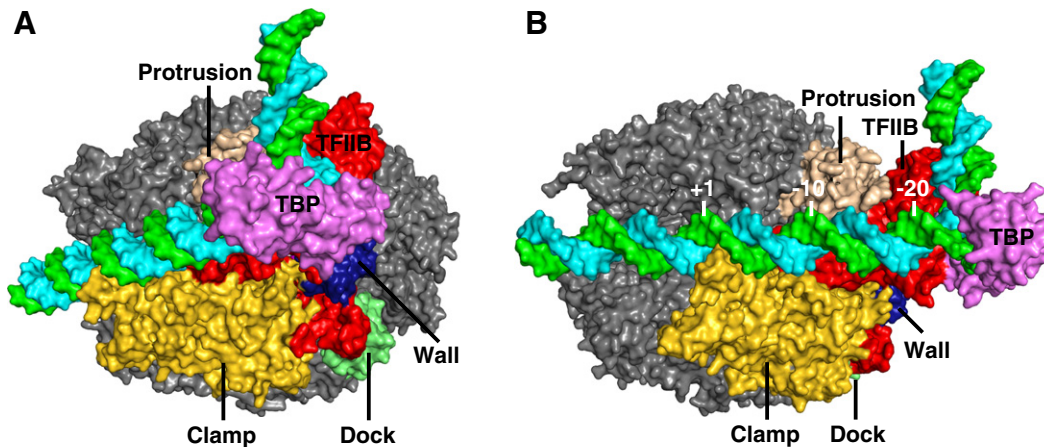


Fig. 3. Structural model of the “minimal” closed promoter complex. (A) “Top” view [9] of the closed promoter complex. Pol II, TFIIB, TBP and promoter DNA are all shown in surface representation. The pol II clamp is shown in gold, dock in lime, wall in blue, protrusion in wheat and the rest of pol II in gray. TFIIB is red and TBP is violet. The template and non-template strands are cyan and green respectively. (B) Same as (A) rotated approximately 45° around the y-axis. DNA registers are also indicated. All figures are rendered in PyMOL.

residues in TFIIB, other transcription factors including TFIIE and TFIIH may assist promoter melting through binding the non-template single strand [23,24,27–29]. In hyperthermophilic archaea, TFB and TBP are sufficient for promoter opening at elevated temperature [47], and the same set of proteins from *Methanococcus thermolithotrophicus* supports promoter opening on a linearized template DNA at 25 °C [48]. It may be relevant that the TFB linker helix (BH0) of *Methanococcus*, which is in close proximity to the promoter at the upstream edge of the transcription bubble, contains tryptophan, tyrosine and several basic residues, reminiscent of the bacterial $\sigma_{2.3}$ region, and conserved in archaea but not in eukaryotes [45] (Fig. 1). Moreover, mutagenesis of *Pyrococcus furiosus* polymerase and TFB has indicated that the TFB linker and its binding surface on the “clamp coiled coil” are required for promoter melting [10]. The clamp coiled coil is a conserved structural feature of bacterial RNAP, archaeal RNAP, and pol II, which provides a major interaction site for bacterial σ_2 and the TFIIB linker [49].

The distance between the TATA box and TSS is almost always about 30 bp for TATA box-containing genes [50]. Consistent with this, the

structural model of the minimal open promoter complex constructed from the pol II–TFIIB crystal structure defines a template strand path of approximately 30 residues from the TATA box to the catalytic site [9,10]. The notable exception is *Saccharomyces cerevisiae*, in which the distance from the TATA box to the TSS is 40–120 bp [51]. It has been suggested that promoter melting starts at the same location in *S. cerevisiae* as in other eukaryotes, about 12 bp downstream of the TATA box, but pol II transcription then scans downstream for the TSS [52]. Alterations in TFIIB, TFIIH, and pol II subunits Rpb1, 2, and 9 change the location of the TSS in *S. cerevisiae* [53–55]. The pol II–TFIIB pair appears to play a critical role [56], and may recognize an initiator element responsible for TSS selection [50,57,58] and subsequent RNA synthesis. In contrast, the mechanism of the TSS selection for TATA box-less genes is less well understood, although other promoter elements and transcription factors as well as nucleosome positioning may play a role in such a process [50,59].

Genetic studies have implicated highly conserved regions of the B-finger in TSS selection [53,60] (Fig. 1). Structural studies have indicated a close proximity of the B-finger to the initiator in the pol II active center [8–10]. A region of the B-finger and of the structurally and functionally related σ_3 – σ_4 linker is required for the efficient synthesis of the first phosphodiester bond, possibly through binding the initiating nucleotide [61,62] (Fig. 1). Indeed, a cross-link can be formed between the σ_3 – σ_4 linker and the initiating nucleotide [62,63]. Alterations of the B-finger and σ_3 – σ_4 linker have been shown to influence the distribution of abortive transcripts, raising the possibility of interaction of these regions with the nascent transcript as well [11,64]. The stimulation of both transcription and abortive initiation by the B-finger has been recapitulated in the archaeon *Methanocaldococcus jannaschii* [65,66].

The two crystal structures of pol II–TFIIB complexes showing alternative conformations of TFIIB, one with a core domain but no B-finger and the other with the reverse [8–10], suggest a pathway for promoter escape [9]. The conformation may switch from that with the core domain to that with the B-finger upon growth of the transcript to about 5 residues, at which point a persistent B-finger–transcript interaction occurs. The core domain is released in this transition, and further growth of the transcript is expected to dislodge the B-finger as well [9]. As TFIIB leaves the complex, TBP and promoter DNA will also dissociate, completing the process of promoter escape.

Biochemical studies have revealed a further correlate of TFIIB release. During the initiation of transcription, the upstream edge of the transcription bubble is fixed (probably by TBP and TFIIB binding) while the downstream edge moves forward, resulting in expansion of

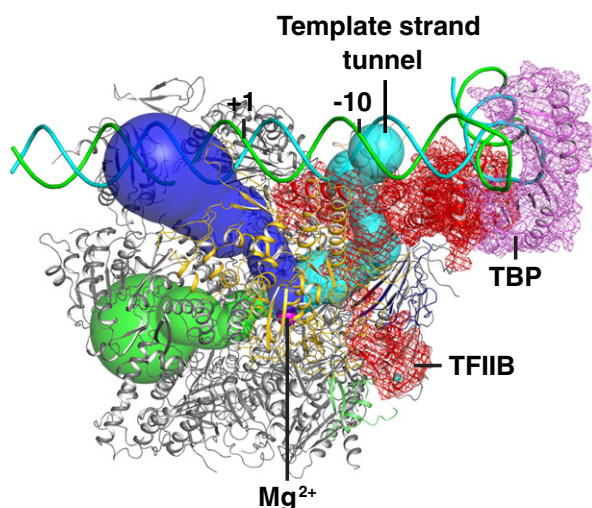


Fig. 4. Tunnels/channels in the pol II–TFIIB complex. Proteins and DNA are colored as in Fig. 3, except that pol II and DNA are shown as ribbons, TFIIB and TBP are shown as meshes, and the catalytic magnesium ion is shown as a magenta sphere. View is that of Fig. 2B rotated 45° around the x-axis. Tunnels/channels were calculated using the program MOLE [88] and are shown as semi-transparent surfaces. The template strand tunnel is shown in cyan, the downstream duplex DNA channel in blue and the ribonucleotide entry tunnel in green.

the bubble. When the nascent transcript grows beyond a critical length (about 7 residues for human pol II) the upstream region of the transcription bubble abruptly reanneals (“bubble collapse”) [64,67]. The passage of the upstream region of the template strand through the TFIIB tunnel provides a straightforward explanation for this otherwise mysterious event. Beyond a transcript length of about 7 residues, TFIIB is released and the tunnel is lost, exposing a region of the template strand, which is then free to reanneal [9].

3. Transcription elongation

The transition from initiation to a stable elongating complex occurs when the transcript reaches a length of about 25 residues. At this point, the RNA is capped, and beyond this point most, if not all, of the GTFs are released. The elongating enzyme is believed to oscillate by simple diffusion between three states (Fig. 5): a “pre-translocation” state, in which a nucleotide has been added to the growing RNA chain, with no other change in structure; a “post-translocation” state, in which the polymerase has moved forward (in the direction of transcription) one nucleotide step along the DNA template, making the active center available for entry of the next NTP; and a “backtracked” state, in which the polymerase has moved backward, extruding the nucleotide just added to the transcript from the active center. Forward movement of the transcribing polymerase is driven by NTP binding, which captures the complex in the post-translocation state [68].

The structures of transcribing complexes in all three states have been determined. Whereas pre-translocation complexes were formed by initiation on a “tailed template” (which does not require GTFs; [5]), post-translocation and backtracked complexes were formed by simple binding of pol II to nucleic acid “scaffolds,” containing a DNA oligonucleotide hybridized with a complementary RNA oligonucleotide, whose affinity for pol II was previously demonstrated [69]. Structures of post-translocation complexes with NTPs in the active center have given insight into the fidelity of RNA synthesis, the accuracy of readout of the genetic code, the essence of transcription. The first structure of a transcribing complex [5] was paradoxical in this regard, as it showed only hydrogen bonding between NTP in the active center and the coding base in the DNA strand. There was no evidence of interaction of NTP with the polymerase. The energy of hydrogen bonding is orders of magnitude too small to account for the fidelity of transcription. The paradox was resolved by screening of hundreds of crystals, leading to higher resolution and improved data quality [70]. A protein feature termed the trigger loop was observed in contact with NTP in the active center (Fig. 6). In many structures of complexes in the absence of NTP, the trigger loop was seen some 30 Å away from the active center (“open” conformation). Only in the presence of NTP matched with the coding base in the DNA template was the trigger loop in proximity to the active center (“closed” conformation). The trigger loop is evidently a mobile element that swings beneath an NTP in the active center. His 1085 of the trigger loop is then in position to serve as a proton

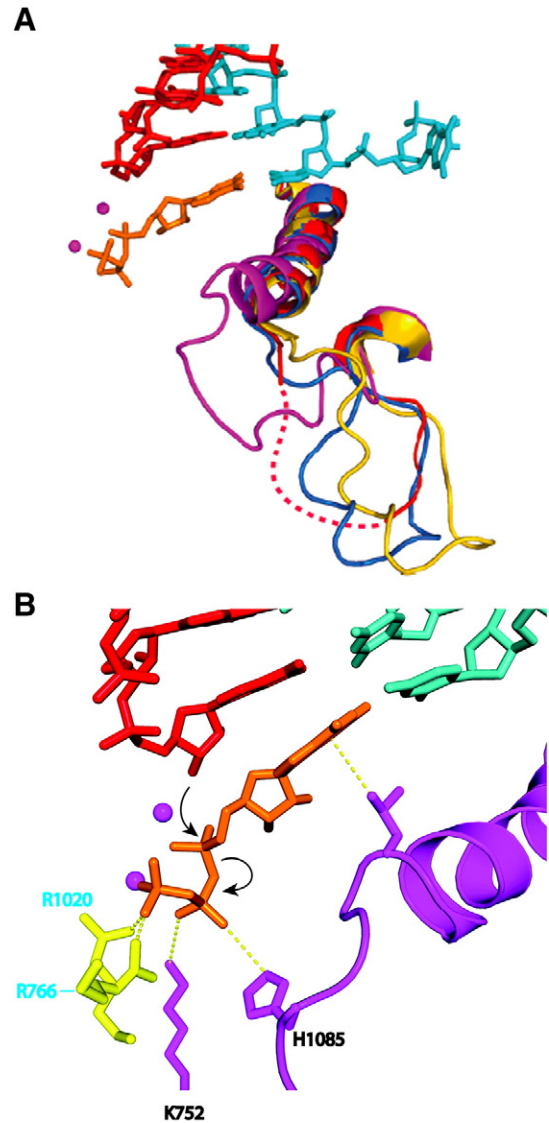


Fig. 6. The trigger loop and its interaction with substrate NTP. (A) Structure of the active center region of transcribing complexes with (trigger loop shown in purple) or without (trigger loops shown in red, blue, or yellow) nucleoside triphosphate (orange). RNA is red, DNA is cyan, and magnesium ions are depicted as purple spheres. (B) Expanded view of structure in (A) with NTP bound and with trigger loop in purple. Interactions of trigger loop with NTP are indicated by dashed yellow lines. In addition to the trigger loop, other interacting residues of pol II are shown in yellow and purple. Arrows indicate the flow of electrons during nucleophilic attack of the 3'-OH of the RNA chain terminus upon the α -phosphate of the NTP, for phosphodiester bond formation, phosphoanhydride bond breakage, and pyrophosphate release.

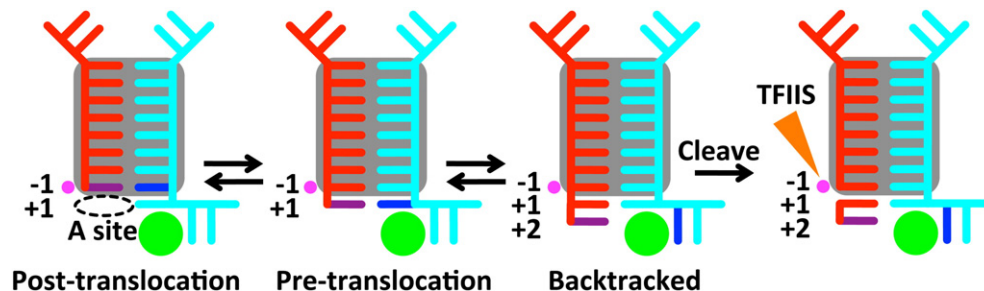


Fig. 5. The multiple states of pol II transcribing complexes. DNA and RNA strands are in cyan and red, with the coding base in the DNA highlighted in blue and the matched base in the RNA strand in purple. The polymerase is symbolized by a gray rectangle and the bridge helix of the polymerase by a green disk. The solid magenta circle represents the catalytic magnesium ion.

donor to the β -phosphate during phosphoanhydride bond breakage and phosphodiester bond formation. In this way, selection of the correct NTP is coupled to catalysis.

The underlying principle of NTP selection by trigger loop interaction is not base recognition but rather helical geometry. The trigger loop does contact the base of the NTP, but this is for the purpose of alignment with respect to the catalytic histidine. If a correct RNA–DNA hybrid base pair is formed, then alignment is precise, leading to catalysis. Discrimination between a ribo and a deoxyribo NTP is a case in point. Interaction with a single hydroxyl group cannot provide the many orders of magnitude of specificity observed, but the smaller diameter of a DNA–DNA than an RNA–DNA base pair, about 2 Å less, will result in a profound misalignment with respect to the trigger loop histidine, and therefore a lack of stimulation by the trigger loop of catalysis.

The trigger loop is a conserved feature of all multisubunit, DNA-directed RNA polymerases, either observed in structures, such as that of the *T. thermophilus* RNAP [71], or indicated by sequence analysis, as for *E. coli* and archaeal (*M. jannaschii*) RNAPs [72,73]. The importance of the trigger loop is shown by deletion, resulting in a reduction in transcription rate [71,73–75], by 60,000-fold in the case of *T. aquaticus* RNAP [75]. Deletion of the trigger loop also impairs discrimination between ribo and deoxyribo NTPs [75]. The importance of the trigger loop is further shown by the toxin α -amanitin, which binds beneath the trigger loop of pol II [76,77], limiting its mobility and preventing it from adopting the closed conformation required for NTP recognition and catalysis. Similarly, the antibiotic streptolydigin has been proposed to act by limiting motion of the trigger loop, and transcription in the presence of streptolydigin resembles that of a trigger loop deletion mutant of RNAP [74].

Point mutations in the trigger loop have given additional insight into its role in transcription. Mutations that likely affect the balance between open (inactive) and closed (active) conformations of the trigger loop have been shown to affect the transcription rate. “Superactive” forms of RNA polymerases can be generated by mutations in the trigger loop thought to favor the closed conformation [72,78].

Mutations of His 1085 support its proposed role as a proton donor. H1085A and F mutants are lethal in *S. cerevisiae* [76], whereas H1085Y, Q, and S mutants are viable but with maximal elongation rates approximately 10-fold less than wild type [76,78]. The H1085Y mutant shows a 5–8 fold reduction in selectivity for ribo over deoxyribo NTPs [76], and the corresponding mutation in *E. coli* RNAP shows a similar reduction in selectivity [73]. The H1085A mutant is viable in the presence of additional trigger loop mutations that confer “superactivity,” showing that a proton donor, such as His 1085, is not absolutely required for transcription. A general acid will increase the reaction rate, dependent on the pKa of the proton donor. For example, in the case of poliovirus RNA-directed RNA polymerase, mutation of the putative proton donor Lys 359 to His with its lower pKa reduces the reaction rate by a factor of ten [79]. Measurements on pol II mutants suggest that *S. cerevisiae* His 1085 contributes a factor of 5–10 to the rate of transcription. The additional enhancement of the transcription rate attributed to the trigger loop is likely due to the stabilization of the NTP in a conformation suitable for efficient bond formation [80].

The error rate in transcription with the wild type trigger loop is on the order of 10^{-6} , of which about 10^{-4} may be accounted for by the fidelity of RNA synthesis. The remaining two orders of magnitude are gained by proofreading, in a three-step process. First, the polymerase backtracks to extrude the misincorporated nucleotide. Then, in a reaction assisted by TFIIIS, the transcript is cleaved in the active center, releasing a dinucleotide containing the misincorporated residue (Fig. 5). Finally, fresh NTP enters the active center and synthesis resumes, with a chance of only one in 10^4 of repeating the original error.

It may be asked why is backtracking favored over forward translocation in case of misincorporation, and why does it stop at one residue, rather than proceeding further with the extrusion of additional

residues? The fit of the RNA–DNA hybrid helix to pol II, shown by the many crystal structures, explains the high affinity of the nucleic acid–protein interaction, and implies that a distortion of the helix due to nucleotide misincorporation will reduce the affinity, shifting the equilibrium of the transcribing enzyme away from the pre- and post-translocated states and in the direction of the backtracked state. The crystal structure of the backtracked state [81] explains why backtracking tends to occur by a single residue. The structure shows a binding site for the first but not for additional backtracked residues. The effect is to trap the one-residue-backtracked complex long enough for excision of the dinucleotide to occur. In the absence of cleavage, there may be further backtracking, as has been shown by both biochemical and structural analyses [82,83].

Although translocation is diffusional in nature, it may be viewed as sliding rather than dissociation and rebinding. The basis may reside, in part, in the interaction of the coding base in the DNA with the “bridge helix,” which crosses the central cleft between the Rpb1 and Rpb2 subunits (Fig. 7). Although the sequence is largely conserved between pol II and bacterial RNAP, the helix in pol II structures is straight, whereas that in RNAP structures is bent. In consequence of such a bend, a threonine side chain that contacts the coding base in a pol II transcribing complex is shifted 3.4 Å, or one base pair step, in the direction of translocation in RNAP. This led to the idea that the bridge helix serves as a molecular ratchet, bending to maintain contact with the coding base during forward translocation, and then snapping back to engage with the next base for another round of the process. In this way, pol II may release the DNA and RNA for translocation while preserving the register for transcription. Some support for the idea has come from biochemical and genetic studies of RNAP. For example, the 3'-end of the RNA can be cross-linked to residues of both bridge helix and trigger loop, which may be explained by two

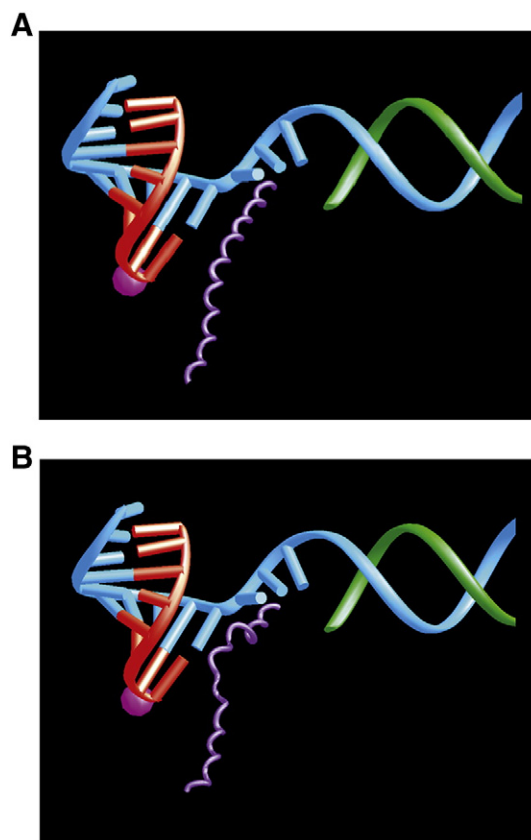


Fig. 7. The bridge helix in straight and bent states. Structures of nucleic acids and the bridge helix in the vicinity of the active center of (A) pol II and (B) RNAP are shown, with the DNA in cyan, the RNA in red, and the bridge helix in purple.

alternative conformations of the bridge helix, one bent and one straight [84]. Mutational and biochemical analyses of the transcribing complex further suggested a “two-pawl” Brownian ratchet mechanism, in which bridge helix bending is coupled to NTP binding to promote forward translocation [85].

In the course of translocation, and to complete a round of nucleotide addition, pol II must disrupt an RNA–DNA base pair at the upstream end of the hybrid helix, for release of the transcript into solution. The first structure of a transcribing complex in the post-translocation state [86] revealed how this strand separation is achieved (Fig. 8). The first six RNA–DNA hybrid base pairs from the active center show normal planar geometry with a separation appropriate for hydrogen bonding between the bases. Starting at the seventh base pair and beyond, the bases are non-coplanar and increasingly splayed apart. Three protein loops are involved: “rudder” contacts the DNA and “lid” contacts the RNA, maintaining the separation of strands; a phenylalanine side chain of the lid serves as a wedge between the strands; lysine residues of the “fork loop” contact the sugar phosphate backbones of the DNA and RNA to stabilize the sixth base pair and prevent unwinding of the hybrid helix from extending back to the active center. These findings were subsequently confirmed by a complete transcribing complex structure [87].

4. Next steps

Studies summarized here represent first steps towards unraveling the mysteries of transcription. The inferences regarding the role of TFIIB in initiation, the function of the trigger loop in fidelity, the basis for translocation, and the mechanism of transcript release, are based on crystal structures, but are nevertheless hypothetical. There is even less information regarding the roles of the additional GTFs in initiation, the transition from initiation to RNA chain elongation, regulation by Mediator and by elongation factors, and so forth. These issues are intertwined, inasmuch as all GTFs are required for initiation, the initiation process involves RNA chain elongation, and regulation occurs at multiple stages of both initiation and elongation processes. A definitive solution of the transcription problem will only come with the isolation of a complete PIC and its transformation to a transcribing complex in a controlled fashion. Then biochemical and biophysical studies may fully elucidate the mechanism and regulation of transcription.

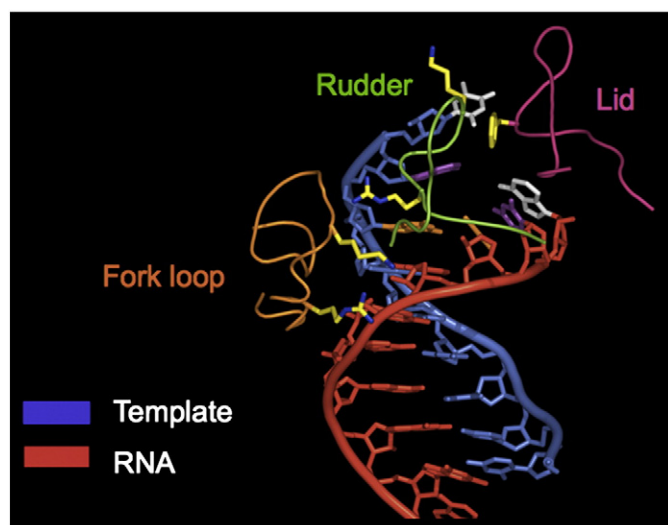


Fig. 8. The structural basis of transcript release. Structure of the RNA–DNA hybrid helix and interacting protein loops in a post-translocation complex.

References

- [1] S.A. Darst, A.M. Edwards, E.W. Kubalek, R.D. Kornberg, Three-dimensional structure of yeast RNA polymerase II at 16 Å resolution, *Cell* 66 (1) (1991) 121–128.
- [2] A.M. Edwards, S.A. Darst, S.A. Hemming, Y. Li, R.D. Kornberg, Epitaxial growth of protein crystals on lipid layers, *Nat. Struct. Biol.* 1 (3) (1994) 195–197.
- [3] J. Fu, A.L. Gnatt, D.A. Bushnell, G.J. Jensen, N.E. Thompson, R.R. Burgess, P.R. David, R.D. Kornberg, Yeast RNA polymerase II at 5 Å resolution, *Cell* 98 (6) (1999) 799–810.
- [4] P. Cramer, D.A. Bushnell, R.D. Kornberg, Structural basis of transcription: RNA polymerase II at 2.8 Å resolution, *Science* 292 (5523) (2001) 1863–1876.
- [5] A.L. Gnatt, P. Cramer, J. Fu, D.A. Bushnell, R.D. Kornberg, Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution, *Science* 292 (5523) (2001) 1876–1882.
- [6] D.A. Bushnell, R.D. Kornberg, Complete, 12-subunit RNA polymerase II at 4.1-Å resolution: implications for the initiation of transcription, *Proc. Natl. Acad. Sci. U. S. A.* 100 (12) (2003) 6969–6973.
- [7] K.J. Armache, H. Kettenberger, P. Cramer, Architecture of initiation-competent 12-subunit RNA polymerase II, *Proc. Natl. Acad. Sci. U. S. A.* 100 (12) (2003) 6964–6968.
- [8] D.A. Bushnell, K.D. Westover, R.E. Davis, R.D. Kornberg, Structural basis of transcription: an RNA polymerase II–TFIIB cocrystal at 4.5 Å resolution, *Science* 303 (5660) (2004) 983–988.
- [9] X. Liu, D.A. Bushnell, D. Wang, G. Calero, R.D. Kornberg, Structure of an RNA polymerase II–TFIIB complex and the transcription initiation mechanism, *Science* 327 (5962) (2010) 206–209.
- [10] D. Kostrewa, M.E. Zeller, K.J. Armache, M. Seizl, K. Leike, M. Thomm, P. Cramer, RNA polymerase II–TFIIB structure and mechanism of transcription initiation, *Nature* 462 (7271) (2009) 323–330.
- [11] K.S. Murakami, S. Masuda, S.A. Darst, Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution, *Science* 296 (5571) (2002) 1280–1284.
- [12] D.G. Vassylyev, S. Sekine, O. Laptchenko, J. Lee, M.N. Vassylyeva, S. Borukhov, S. Yokoyama, Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution, *Nature* 417 (6890) (2002) 712–719.
- [13] F.T. Tsai, P.B. Sigler, Structural basis of preinitiation complex assembly on human pol II promoters, *EMBO J.* 19 (1) (2000) 25–36.
- [14] D.B. Nikolov, H. Chen, E.D. Halay, A.A. Usheva, K. Hisatake, D.K. Lee, R.G. Roeder, S.K. Burley, Crystal structure of a TFIIB–TBP–TATA-element ternary complex, *Nature* 377 (6545) (1995) 119–128.
- [15] J.D. Parvin, P.A. Sharp, DNA topology and a minimal set of basal factors for transcription by RNA polymerase II, *Cell* 73 (3) (1993) 533–540.
- [16] W. Hausner, J. Wettach, C. Hethke, M. Thomm, Two transcription factors related with the eucaryal transcription factors TATA-binding protein and transcription factor IIB direct promoter recognition by an archaeal RNA polymerase, *J. Biol. Chem.* 271 (47) (1996) 30144–30148.
- [17] S.A. Qureshi, S.D. Bell, S.P. Jackson, Factor requirements for transcription in the archaeon *Sulfolobus shibatae*, *EMBO J.* 16 (10) (1997) 2927–2936.
- [18] B.L. Hanzelka, T.J. Darcy, J.N. Reeve, TFE, an archaeal transcription factor in *Methanobacterium thermoautotrophicum* related to eucaryal transcription factor TFIIE α , *J. Bacteriol.* 183 (5) (2001) 1813–1818.
- [19] S.D. Bell, S.P. Jackson, Mechanism and regulation of transcription in archaea, *Curr. Opin. Microbiol.* 4 (2) (2001) 208–213.
- [20] N. Hernandez, TBP, a universal eukaryotic transcription factor? *Genes Dev.* 7 (7B) (1993) 1291–1308.
- [21] R.J. White, S.P. Jackson, The TATA-binding protein: a central role in transcription by RNA polymerases I, II and III, *Trends Genet.* 8 (8) (1992) 284–288.
- [22] G. Miller, S. Hahn, A DNA-tethered cleavage probe reveals the path for promoter DNA in the yeast preinitiation complex, *Nat. Struct. Mol. Biol.* 13 (7) (2006) 603–610.
- [23] M. Okuda, Y. Watanabe, H. Okamura, F. Hanaoka, Y. Ohkuma, Y. Nishimura, Structure of the central core domain of TFIIE β with a novel double-stranded DNA-binding surface, *EMBO J.* 19 (6) (2000) 1346–1356.
- [24] T. Okamoto, S. Yamamoto, Y. Watanabe, T. Ohta, F. Hanaoka, R.G. Roeder, Y. Ohkuma, Analysis of the role of TFIIE in transcriptional regulation through structure–function studies of the TFIIE β subunit, *J. Biol. Chem.* 273 (31) (1998) 19866–19876.
- [25] H.T. Chen, L. Warfield, S. Hahn, The positions of TFIIF and TFIIE in the RNA polymerase II transcription preinitiation complex, *Nat. Struct. Mol. Biol.* 14 (8) (2007) 696–703.
- [26] F. Robert, D. Forget, J. Li, J. Greenblatt, B. Coulombe, Localization of subunits of transcription factors IIE and IIF immediately upstream of the transcriptional initiation site of the adenovirus major late promoter, *J. Biol. Chem.* 271 (15) (1996) 8517–8520.
- [27] C.M. Groft, S.N. Uljon, R. Wang, M.H. Werner, Structural homology between the Rap30 DNA-binding domain and linker histone H5: implications for preinitiation complex assembly, *Proc. Natl. Acad. Sci. U. S. A.* 95 (16) (1998) 9117–9122.
- [28] K. Kamada, J. De Angelis, R.G. Roeder, S.K. Burley, Crystal structure of the C-terminal domain of the RAP74 subunit of human transcription factor IIF, *Proc. Natl. Acad. Sci. U. S. A.* 98 (6) (2001) 3115–3120.
- [29] S. Tan, K.P. Garrett, R.C. Conaway, J.W. Conaway, Cryptic DNA-binding domain in the C terminus of RNA polymerase II general transcription factor RAP30, *Proc. Natl. Acad. Sci. U. S. A.* 91 (21) (1994) 9808–9812.
- [30] W.H. Chung, J.L. Craighead, W.H. Chang, C. Ezeokonkwo, A. Bareket-Samish, R.D. Kornberg, F.J. Asturias, RNA polymerase II/TFIIF structure and conserved organization of the initiation complex, *Mol. Cell* 12 (4) (2003) 1003–1013.
- [31] J. Eichner, H.T. Chen, L. Warfield, S. Hahn, Position of the general transcription factor TFIIF within the RNA polymerase II transcription preinitiation complex, *EMBO J.* 29 (4) (2010) 706–716.

- [32] Z.A. Chen, A. Jawhari, L. Fischer, C. Buchen, S. Tahir, T. Kamenski, M. Rasmussen, L. Lariviere, J.C. Bukowski-Wills, M. Nilges, et al., Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry, *EMBO J.* 29 (4) (2010) 717–726.
- [33] T.K. Kim, R.H. Ebright, D. Reinberg, Mechanism of ATP-dependent promoter melting by transcription factor IIH, *Science* 288 (5470) (2000) 1418–1422.
- [34] S.A. Qureshi, S.P. Jackson, Sequence-specific DNA binding by the *S. shibatae* TFIIB homolog, TFB, and its effect on promoter strength, *Mol. Cell* 1 (3) (1998) 389–400.
- [35] T. Lagrange, A.N. Kapanidis, H. Tang, D. Reinberg, R.H. Ebright, New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB, *Genes Dev.* 12 (1) (1998) 34–44.
- [36] W. Deng, S.G. Roberts, A core promoter element downstream of the TATA box that is recognized by TFIIB, *Genes Dev.* 19 (20) (2005) 2418–2423.
- [37] O. Littlefield, Y. Korkhin, P.B. Sigler, The structural basis for the oriented assembly of a TBP/TFB/promoter complex, *Proc. Natl. Acad. Sci. U. S. A.* 96 (24) (1999) 13668–13673.
- [38] B.S. Ibrahim, N. Kanneganti, G.E. Rieckhof, A. Das, D.V. Laurents, J.B. Palenchar, V. Bellofatto, D.A. Wah, Structure of the C-terminal domain of transcription factor IIB from *Trypanosoma brucei*, *Proc. Natl. Acad. Sci. U. S. A.* 106 (32) (2009) 13242–13247.
- [39] S. Martinez-Calvillo, S. Yan, D. Nguyen, M. Fox, K. Stuart, P.J. Myler, Transcription of *Leishmania major* Friedlin chromosome 1 initiates in both directions within a single region, *Mol. Cell* 11 (5) (2003) 1291–1299.
- [40] P. Respuela, M. Ferella, A. Rada-Iglesias, L. Aslund, Histone acetylation and methylation at sites initiating divergent polycistronic transcription in *Trypanosoma cruzi*, *J. Biol. Chem.* 283 (23) (2008) 15884–15892.
- [41] T.N. Siegel, D.R. Hekstra, L.E. Kemp, L.M. Figueiredo, J.E. Lowell, D. Fenyo, X. Wang, S. Dewell, G.A. Cross, Four histone variants mark the boundaries of polycistronic transcription units in *Trypanosoma brucei*, *Genes Dev.* 23 (9) (2009) 1063–1076.
- [42] A.C. Seila, J.M. Calabrese, S.S. Levine, G.W. Yeo, P.B. Rahl, R.A. Flynn, R.A. Young, P.A. Sharp, Divergent transcription from active promoters, *Science* 322 (5909) (2008) 1849–1851.
- [43] L.S. Churchman, J.S. Weissman, Nascent transcript sequencing visualizes transcription at nucleotide resolution, *Nature* 469 (7330) (2011) 368–373.
- [44] D. Bose, T. Pape, P.C. Burrows, M. Rappas, S.R. Wigneshwararaj, M. Buck, X. Zhang, Organization of an activator-bound RNA polymerase holoenzyme, *Mol. Cell* 32 (3) (2008) 337–346.
- [45] A. Feklistov, S.A. Darst, Promoter recognition by bacterial alternative sigma factors: the price of high selectivity? *Genes Dev.* 23 (20) (2009) 2371–2375.
- [46] A. Feklistov, S.A. Darst, Structural basis for promoter-10 element recognition by the bacterial RNA polymerase sigma subunit, *Cell* 147 (6) (2011) 1257–1269.
- [47] S.D. Bell, C. Jaxel, M. Nadal, P.F. Kosa, S.P. Jackson, Temperature, template topology, and factor requirements of archaeal transcription, *Proc. Natl. Acad. Sci. U. S. A.* 95 (26) (1998) 15218–15222.
- [48] W. Hausner, M. Thomm, Events during initiation of archaeal transcription: open complex formation and DNA–protein interactions, *J. Bacteriol.* 183 (10) (2001) 3025–3031.
- [49] W.J. Lane, S.A. Darst, Molecular evolution of multisubunit RNA polymerases: structural analysis, *J. Mol. Biol.* 395 (4) (2010) 686–704.
- [50] S.T. Smale, J.T. Kadonaga, The RNA polymerase II core promoter, *Annu. Rev. Biochem.* 72 (2003) 449–479.
- [51] K. Struhl, Molecular mechanisms of transcriptional regulation in yeast, *Annu. Rev. Biochem.* 58 (1989) 1051–1077.
- [52] C. Giardina, J.T. Lis, DNA melting on yeast RNA polymerase II promoters, *Science* 261 (5122) (1993) 759–762.
- [53] I. Pinto, W.H. Wu, J.G. Na, M. Hampsey, Characterization of sua7 mutations defines a domain of TFIIB involved in transcription start site selection in yeast, *J. Biol. Chem.* 269 (48) (1994) 30569–30573.
- [54] M.A. Ghazy, S.A. Brodie, M.L. Ammerman, L.M. Ziegler, A.S. Ponticelli, Amino acid substitutions in yeast TFIIF confer upstream shifts in transcription initiation and altered interaction with RNA polymerase II, *Mol. Cell. Biol.* 24 (24) (2004) 10975–10985.
- [55] M.W. Hull, K. McKune, N.A. Woychik, RNA polymerase II subunit RPB9 is required for accurate start site selection, *Genes Dev.* 9 (4) (1995) 481–490.
- [56] Y. Li, P.M. Flanagan, H. Tschochner, R.D. Kornberg, RNA polymerase II initiation factor interactions and transcription start site selection, *Science* 263 (5148) (1994) 805–807.
- [57] Z. Zhang, F.S. Dietrich, Mapping of transcription start sites in *Saccharomyces cerevisiae* using 5' SAGE, *Nucleic Acids Res.* 33 (9) (2005) 2838–2851.
- [58] J.N. Kuehner, D.A. Brow, Quantitative analysis of in vivo initiator selection by yeast RNA polymerase II supports a scanning model, *J. Biol. Chem.* 281 (20) (2006) 14119–14128.
- [59] H.S. Rhee, B.F. Pugh, Genome-wide structure and organization of eukaryotic pre-initiation complexes, *Nature* 483 (7389) (2012) 295–301.
- [60] S.L. Faltier, S.A. Brodie, A.S. Ponticelli, Promoter-specific shifts in transcription initiation conferred by yeast TFIIB mutations are determined by the sequence in the immediate vicinity of the start sites, *Mol. Cell. Biol.* 21 (14) (2001) 4427–4440.
- [61] K. Tran, J.D. Gralla, Control of the timing of promoter escape and RNA catalysis by the transcription factor IIb fingertip, *J. Biol. Chem.* 283 (23) (2008) 15665–15671.
- [62] A. Kulbachinskiy, A. Mustaev, Region 3.2 of the sigma subunit contributes to the binding of the 3'-initiating nucleotide in the RNA polymerase active center and facilitates promoter clearance during initiation, *J. Biol. Chem.* 281 (27) (2006) 8273–18276.
- [63] K. Severinov, D. Fenyo, E. Severinova, A. Mustaev, B.T. Chait, A. Goldfarb, S.A. Darst, The sigma subunit conserved region 3 is part of "5'-face" of active center of *Escherichia coli* RNA polymerase, *J. Biol. Chem.* 269 (33) (1994) 20826–20828.
- [64] M. Pal, A.S. Ponticelli, D.S. Luse, The role of the transcription bubble and TFIIB in promoter clearance by RNA polymerase II, *Mol. Cell* 19 (1) (2005) 101–110.
- [65] F. Werner, R.O. Weinzierl, Direct modulation of RNA polymerase core functions by basal transcription factors, *Mol. Cell. Biol.* 25 (18) (2005) 8344–8355.
- [66] S.C. Wiesler, R.O. Weinzierl, The linker domain of basal transcription factor TFIIB controls distinct recruitment and transcription stimulation functions, *Nucleic Acids Res.* 39 (2) (2011) 464–474.
- [67] F.C. Holstege, U. Fiedler, H.T. Timmers, Three transitions in the RNA polymerase II transcription complex during initiation, *EMBO J.* 16 (24) (1997) 7468–7480.
- [68] K.M. Herbert, W.J. Greenleaf, S.M. Block, Single-molecule studies of RNA polymerase: motoring along, *Annu. Rev. Biochem.* 77 (2008) 149–176.
- [69] M.L. Kireeva, N. Komissarova, D.S. Waugh, M. Kashlev, The 8-nucleotide-long RNA:DNA hybrid is a primary stability determinant of the RNA polymerase II elongation complex, *J. Biol. Chem.* 275 (9) (2000) 6530–6536.
- [70] D. Wang, D.A. Bushnell, K.D. Westover, C.D. Kaplan, R.D. Kornberg, Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis, *Cell* 127 (5) (2006) 941–954.
- [71] D.G. Vassilyev, M.N. Vassilyeva, J. Zhang, M. Palangat, I. Artsimovitch, R. Landick, Structural basis for substrate loading in bacterial RNA polymerase, *Nature* 448 (7150) (2007) 163–168.
- [72] L. Tan, S. Wiesler, D. Trzaska, H.C. Carney, R.O. Weinzierl, Bridge helix and trigger loop perturbations generate superactive RNA polymerases, *J. Biol.* 7 (10) (2008) 40.
- [73] J. Zhang, M. Palangat, R. Landick, Role of the RNA polymerase trigger loop in catalysis and pausing, *Nat. Struct. Mol. Biol.* 17 (1) (2009) 99–104.
- [74] D. Temiakov, N. Zenkin, M.N. Vassilyeva, A. Perederina, T.H. Tahirov, E. Kashkina, M. Savkina, S. Zorov, V. Nikiforov, N. Igarashi, et al., Structural basis of transcription inhibition by antibiotic streptolydigin, *Mol. Cell* 19 (5) (2005) 655–666.
- [75] Y. Yuzenkova, A. Bochkareva, V.R. Tadigotla, M. Roghianian, S. Zorov, K. Severinov, N. Zenkin, Stepwise mechanism for transcription fidelity, *BMC Biol.* 8 (1) (2010) 54.
- [76] C.D. Kaplan, K.-M. Larsson, R.D. Kornberg, The RNA polymerase II trigger loop functions in substrate selection and is directly targeted by α -amanitin, *Mol. Cell* 30 (5) (2008) 547–556.
- [77] D.A. Bushnell, P. Cramer, R.D. Kornberg, Structural basis of transcription: alpha-amanitin-RNA polymerase II cocrystal at 2.8 Å resolution, *Proc. Natl. Acad. Sci. U. S. A.* 99 (3) (2002) 1218–1222.
- [78] C.D. Kaplan, H. Jin, I.L. Zhang, A. Belyanin, Dissection of Pol II trigger loop function and Pol II activity-dependent control of start site selection in vivo, *PLoS Genet.* 8 (4) (2012) e1002627.
- [79] C. Castro, E.D. Smidansky, J.J. Arnold, K.R. Maksimchuk, I. Moustafa, A. Uchida, M. Götte, W. Konigsberg, C.E. Cameron, Nucleic acid polymerases use a general acid for nucleotidyl transfer, *Nat. Struct. Mol. Biol.* 16 (2) (2009) 212–218.
- [80] X. Huang, D. Wang, D.R. Weiss, D.A. Bushnell, R.D. Kornberg, M. Levitt, RNA polymerase II trigger loop residues stabilize and position the incoming nucleotide triphosphate in transcription, *Proc. Natl. Acad. Sci. U. S. A.* 107 (36) (2010) 15745–15750.
- [81] D. Wang, D.A. Bushnell, X. Huang, K.D. Westover, M. Levitt, R.D. Kornberg, Structural basis of transcription: backtracked RNA polymerase II at 3.4 Å resolution, *Science* 324 (5931) (2009) 1203–1206.
- [82] M.L. Kireeva, B. Hancock, G.H. Cremona, W. Walter, V.M. Studitsky, M. Kashlev, Nature of the nucleosomal barrier to RNA polymerase II, *Mol. Cell* 18 (1) (2005) 97–108.
- [83] A.C. Cheung, P. Cramer, Structural basis of RNA polymerase II backtracking, arrest and reactivation, *Nature* 471 (7337) (2011) 249–253.
- [84] V. Epshtein, A. Mustaev, V. Markovtsov, O. Bereshchenko, V. Nikiforov, A. Goldfarb, Swing-gate model of nucleotide entry into the RNA polymerase active center, *Mol. Cell* 10 (3) (2002) 623–634.
- [85] G. Bar-Nahum, V. Epshtein, A.E. Ruckenstein, R. Rafikov, A. Mustaev, E. Nudler, A ratchet mechanism of transcription elongation and its control, *Cell* 120 (2) (2005) 183–193.
- [86] K.D. Westover, D.A. Bushnell, R.D. Kornberg, Structural basis of transcription: separation of RNA from DNA by RNA polymerase II, *Science* 303 (5660) (2004) 1014–1016.
- [87] H. Kettenberger, K.J. Armache, P. Cramer, Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS, *Mol. Cell* 16 (6) (2004) 955–965.
- [88] M. Petrek, P. Kosinova, J. Koca, M. Otyepka, MOLE: a Voronoi diagram-based explorer of molecular channels, pores, and tunnels, *Structure* 15 (11) (2007) 1357–1363.