Control of elongation by RNA polymerase II

Joan W. Conaway, Ali Shilatifard, Arik Dvir and Ronald C. Conaway

The elongation stage of eukaryotic mRNA synthesis can be regulated by transcription factors that interact directly with the RNA polymerase II (pol II) elongation complex and by activities that modulate the structure of its chromatin template. Recent studies have revealed new elongation factors and have implicated the general initiation factors TFIIE, TFIIF and TFIIH, as well as the C-terminal domain (CTD) of the largest subunit of pol II, in elongation. The recently reported high-resolution crystal structure of RNA polymerase II, which provides insight into the architecture of the elongation complex, marks a new era of investigation into transcription elongation.

EUKARYOTIC mRNA SYNTHESIS is a complex biochemical process that is governed by the concerted action of a diverse collection of transcription factors that control the activity of RNA polymerase II (pol II) at both the initiation and elongation stages of transcription. Over the past 30 years, attention has focused largely on the initiation stage of eukaryotic mRNA synthesis. During this time, components of the pol II initiation machinery, including the Mediator complex and the general initiation factors, have been defined biochemically and their roles in transcription firmly established.

Over the past few years, increasing attention has focused on the mechanism and regulation of the elongation stage of eukaryotic mRNA synthesis. As we describe below, a major advance in our understanding of elongation has come with the recent report by Kornberg and co-workers1 of a 3.5-Å-resolution crystal structure of pol II. The new pol II structure is providing insights into the molecular mechanisms underlying the remarkable processivity of the elongation complex, as well as its propensity to pause and arrest. In addition, biochemical studies have continued to bring to light new information on the mechanism of elongation by pol II and on the transcription factors that control this process (Table 1). Unexpected post-initiation roles for the general initiation factors TFIIE, TFIIF and TFIIH in early elongation and promoter escape have been brought to light by studies investigating the transition of pol II from the initiation to the elongation stage of transcription. Direct evidence supporting a role for phosphorylation of the C-terminal domain (CTD) of the largest subunit of pol II in regulation of elongation has been obtained. New elongation factors that control the activity of pol II on naked DNA and on chromatin templates have been discovered and their mechanisms of action characterized in increasing detail. Finally, insights into the physiological roles of elongation factors in normal cell function and in disease are beginning to emerge.

Structure and properties of the RNA polymerase II elongation complex

Pol II is a ~600-kDa, 12-subunit enzyme that is highly conserved from yeast to mammals. Understanding how pol II elongates RNA chains requires answering, at a molecular level, two basic questions concerning its catalytic mechanism. First, what gives pol II its remarkable processivity; that is, how does the enzyme grip the DNA template and nascent RNA so tenaciously that it can synthesize transcripts the length of the two-million-nucleotide dystrophin precursor mRNA in a single reaction, without dissociating from the DNA template? And, second, why is elongating pol II so susceptible to transient pausing and arrest?

The crystal structure of a ten-subunit catalytically active form of yeast pol II has provided information concerning the architecture of the elongation complex and its catalytic properties1. As illustrated in the model of Fig. 1a, the pol II catalytic site, with its single essential Mg2+ ion, resides at the end of a deep cleft formed at the interface of the Rpb1 and Rpb2 subunits. The size of this cleft is sufficient to accommodate the predicted 8–9 basepairs of an RNA–DNA hybrid in a configuration that ensures proper alignment of the 3′-OH of the nascent transcript and the catalytic site Mg2+ ion. The remainder of the cleft between the Rpb1 and Rpb2 subunits is of sufficient size to accommodate ~20 basepairs of template DNA downstream of the catalytic site.

The pol II crystal structure has shed light on the extraordinary stability and processivity of the elongation complex. A pair of ‘jaws’ formed by portions of the Rpb1, Rpb5 and Rpb9 subunits are positioned to stabilize the pol II elongation complex by encircling the DNA-binding cleft downstream of the catalytic site and gripping the DNA template as it enters the enzyme. A hinged domain formed by portions of the Rpb1, Rpb2 and Rpb6 subunits is positioned to stabilize the elongation complex by acting as a ‘sliding clamp’ that locks template DNA near the catalytic site in place. A pair of potential RNA-binding ‘grooves’ are positioned properly to stabilize the elongation complex by binding to the nascent transcript approximately 10–20 nucleotides downstream of the catalytic site. Notably, it is likely that a transcript occupying one of the potential RNA-binding grooves would lock the sliding clamp in the closed position, further stabilizing the elongation complex.

Evidence from biochemical studies suggests that pausing and arrest result from aberrant backward movement of pol II on the DNA template and displacement of the 3′-end of the transcript from the catalytic site in a process that is spontaneously reversible (i.e. pausing) or not (i.e. arrest) (Fig. 1b)2. The elongation factor SII reactivates arrested pol II by promoting endonucleolytic cleavage of the nascent transcript, so that the

J.W. Conaway and R.C. Conaway are at the Program in Molecular and Cell Biology, Oklahoma Medical Research Foundation, Oklahoma City, OK 73104, USA; A. Shilatifard is at the Edward A. Doisy Dept of Biochemistry, St Louis University School of Medicine, St Louis, MO 63104, USA; and A. Dvir is at the Dept of Biological Sciences, Oakland University, Rochester, MI 48309, USA.

Email: conaway@omrf.ouhsc.edu

© 2000, Elsevier Science Ltd. All rights reserved. PII: S0968-0004(00)01615-7

9968 – 0004/00/8 – See front matter © 2000, Elsevier Science Ltd. All rights reserved. PII: S0968-0004(00)01615-7

375
new 3'-end of the RNA is again properly aligned in the catalytic site. A 'pore' located directly beneath the catalytic site Mg²⁺ ion might function not only as a portal for entry of NTPs into the catalytic site but also as a passage for exit of the 3'-end of the nascent transcript during backward movement of pol II on the DNA. This pore might also provide SII access to the catalytic site, so that it can activate the nascent transcript cleavage reaction.

**Roles for general initiation factors TFIIE, TFIIF and TFIH in early elongation and promoter escape**

Transcription by pol II is a multistep process that requires minimally the five general initiation factors TFIIB, TFIID, TFIIE, TFIIF and TFIH. Transcription begins with assembly of pol II and all five initiation factors into a closed preinitiation complex at the promoter and culminates in ATP-dependent formation of an open complex by the TFIIH XPB DNA helicase (TFIIH is composed of nine subunits, two of which, designated XPB and XPD, are DNA helicases) and synthesis of the first phosphodiester bond of nascent transcripts. Following transcription initiation, pol II moves away from the promoter and dissociates from the general initiation factors after synthesizing 10–15-nucleotide-long transcripts, in a step referred to as promoter escape (Fig. 2).

Over the past few years, biochemical studies have revealed the striking complexity of early elongation and promoter escape. These studies have shown that very early pol II elongation complexes containing transcripts less than approximately nine nucleotides in length are unstable and prone to aborting transcription. As a consequence, at each step of nucleotide addition, there is a competition between dissociation of the elongation complex from the DNA template and formation of the next phosphodiester bond, until pol II has synthesized approximately nine-nucleotide-long transcripts. Notably, recent evidence suggests that general initiation factor TFIIF, which is both required for initiation and capable of stimulating the rate of elongation by pol II (Ref. 7), acts during very early elongation to decrease the frequency of abortive transcription, perhaps simply by increasing the rate of nucleotide addition, so that the nascent transcript grows to approximately nine nucleotides before pol II aborts transcription.

Before escaping the promoter, very early pol II elongation complexes are also susceptible to transcriptional arrest. Arrest by these early elongation complexes is suppressed by TFIIE and TFIIH in a reaction that, like formation of the open complex, requires ATP-dependent action of the TFIIH XPB DNA helicase.

Although the mechanisms of action of the TFIIH XPB DNA helicase in transcription initiation and promoter escape are not yet clear, evidence suggests they might be related. First, TFIIH activity in both transcription initiation and promoter escape is dependent on TFIIE (Ref. 9–12). Second, mutations of the TFIIH XPB DNA helicase have very similar effects on the efficiencies of both transcription initiation and promoter escape. A fraction of polymerases that have initiated transcription in the presence of ATP are capable of escaping the promoter, even after removal of ATP from reaction mixtures. This observation suggests that a fraction of preinitiation complexes are converted by the TFIIH XPB DNA helicase to both an initiation- and escape-competent transcriptional intermediate in a single ATP-dependent step that results in open complex formation and, perhaps, additional structural rearrangements that remove impediments to promoter escape. The ATP-activated pol II preinitiation complex decays very rapidly from an active to an inactive conformation in the absence of ATP (Ref. 13). Thus, very early pol II elongation complexes might suffer arrest and require additional ATP-dependent action of the TFIIH XPB DNA helicase if they fail to escape the promoter before they have decayed to an inactive conformation.

These findings are consistent with the model that the very early pol II elongation complex suffers transcriptional arrest unless it undergoes a critical ATP-dependent structural transition to an escape-competent conformation; however, the molecular basis of this transition is presently unknown. One potential clue to the nature of this transition has come from findings indicating that promoter escape by pol II is

---

**Table 1. RNA polymerase II elongation factors**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Activities</th>
<th>Comments</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA150</td>
<td>Stimulates Tat-dependent elongation</td>
<td></td>
<td>41</td>
</tr>
<tr>
<td>CSB</td>
<td>Suppresses pausing</td>
<td>Mutated in Cockayne syndrome, required for transcription-coupled DNA repair</td>
<td>39</td>
</tr>
<tr>
<td>DSIF (Spt4/Spt5)</td>
<td>Inhibits elongation</td>
<td>Stimulates elongation under certain conditions</td>
<td>24</td>
</tr>
<tr>
<td>ELL</td>
<td>Suppresses pausing</td>
<td>Forms complex with hSfn8, EAP20, EAP45, interacts with p53</td>
<td>42,52</td>
</tr>
<tr>
<td>ELL2</td>
<td>Suppresses pausing</td>
<td></td>
<td>53</td>
</tr>
<tr>
<td>Elongator</td>
<td>Histone acetyl-transferase</td>
<td>Associated with elongating pol II</td>
<td>49</td>
</tr>
<tr>
<td>Elongin A</td>
<td>Suppresses pausing</td>
<td>Binds to and is stimulated by Elongin BC complex</td>
<td>54</td>
</tr>
<tr>
<td>Elongin A2</td>
<td>Suppresses pausing</td>
<td>Binds to Elongin BC complex</td>
<td>55</td>
</tr>
<tr>
<td>FACT</td>
<td>Promotes elongation on chromatin template</td>
<td></td>
<td>48</td>
</tr>
<tr>
<td>Factor 2</td>
<td>Promotes ATP-dependent transcript release</td>
<td></td>
<td>56</td>
</tr>
<tr>
<td>HMG14</td>
<td>Promotes elongation on chromatin template</td>
<td></td>
<td>57</td>
</tr>
<tr>
<td>NELF</td>
<td>Inhibits elongation</td>
<td>Binds to RNA?</td>
<td>23</td>
</tr>
<tr>
<td>P-TEFb</td>
<td>CTD kinase, antagonizes NELF and DSIF, and needed for Tat-dependent elongation</td>
<td>Cdk9 and cyclin T1, T2 or K subunits, transcription activity depends on kinase</td>
<td>26</td>
</tr>
<tr>
<td>SII</td>
<td>Prevents arrest and promotes nascent transcript cleavage</td>
<td></td>
<td>2,35</td>
</tr>
<tr>
<td>Tat-Sf1</td>
<td>Suppresses pausing and stimulates Tat-dependent elongation</td>
<td></td>
<td>40,58</td>
</tr>
<tr>
<td>TFIIF</td>
<td>Suppresses pausing</td>
<td>RAP30 and RAP74 subunits</td>
<td>7</td>
</tr>
</tbody>
</table>

*Abbreviations: CSB, Cockayne syndrome group B; CTD, C-terminal domain; DSIF, DRB-sensitivity-inducing factor; FACT, facilitates chromatin transcription; HMG14, high-mobility-group 14; NELF, negative elongation factor complex; P-TEFb, positive transcription elongation factor b; SSRP1, structure-specific recognition protein 1.*
substantially less dependent on TFIIE, TFIIF and ATP in transcription reconstituted with TFIIF mutants that are defective in supporting transcription initiation but can efficiently stimulate elongation. These findings have led to the proposals, first, that a TFIIF function required for formation of the preinitiation complex and transcription initiation might actually present an impediment to promoter escape and, second, that this impediment can be overcome by ATP-dependent action of the TFIIF XPB DNA helicase.

How might a TFIIF activity required for initiation present an impediment to promoter escape? Based on results of crosslinking experiments and analysis of electron micrographs of the preinitiation complex, it has been proposed that promoter DNA is tightly wrapped around pol II and the general initiation factors in the preinitiation complex. Notably, evidence suggests that the TFIIF RAP30 and RAP74 subunits can be crosslinked in the preinitiation complex to promoter DNA both upstream and downstream of the TATA box. It has been proposed that these TFIIF-DNA contacts are largely responsible for formation and maintenance of the tightly wrapped structure. TFIIF-dependent bending of the DNA within the preinitiation complex might facilitate unwinding of promoter DNA during formation of the open complex.

The initiation-defective TFIIF mutant described above contains a RAP30 subunit with a deletion that disrupts a C-terminal DNA-binding domain with sequence similarity to a highly conserved promoter-binding domain found in the C terminus of the bacterial transcription initiation factor σ (Ref. 18). Recent NMR studies have revealed that this RAP30 DNA-binding domain is structurally similar to the winged helix–turn–helix DNA-binding domains of linker histone H5 and hepatocyte nuclear transcription factor 3 (HNF3)/forkhead. Therefore, it is possible that interaction of TFIIF and promoter DNA downstream of the transcriptional start site could present an impediment to promoter escape. If this model is correct, the TFIIF XPB DNA helicase could increase the efficiency of promoter escape by disrupting this interaction.

Additional evidence indicates, first, that premelting a region of promoter DNA downstream of the transcriptional start site largely removes the TFIIF-induced impediment to promoter escape and, second, that the RAP30 C-terminal DNA-binding domain interacts most tightly with double-stranded DNA (Ref. 18). Given this evidence, it is possible that the TFIIF XPB DNA helicase could disrupt TFIIF interactions with the promoter simply by unwinding downstream DNA.

### Elongation and phosphorylation of the RNA polymerase II CTD

Elegant studies carried out more than 15 years ago by Dahmus and co-workers first revealed that the transition of pol II from initiation to elongation is accompanied by hyperphosphorylation of the heptapeptide repeats in the C-terminal domain (CTD) of the largest polymerase subunit. Efforts to understand the role of CTD phosphorylation in transcription by pol II were aided greatly by protein kinase inhibitors such as 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB), which block CTD phosphorylation and induce arrest of elongating pol II in vivo and in crude transcription systems. These studies predicted the existence of two classes of elongation factors responsible for DRB-sensitive transcription: a class of negatively acting factors that inhibit transcription and a class of positively acting, DRB-sensitive factors that overcome this inhibition. Indeed, fractionation of crude, DRB-sensitive transcription systems have now led to the identification of three novel elongation factors: one positively acting, DRB-sensitive factor, P-TEFb (Refs 23,24), and two negatively acting factors, NELF (Ref. 25) and DSIF (Ref. 26) (Table 1).

P-TEFb is a DRB-sensitive, cyclin-dependent CTD kinase composed of Cdk9 (also referred to as PITALRE) and one of several cyclins including T1, T2 and K (Refs 27–31). P-TEFb cannot prevent arrest of polymerases lacking a CTD (Ref. 27), and CTD-kinase-deficient P-TEFb point mutants are transcriptionally inactive. In addition, a close correlation between inhibition of P-TEFb CTD kinase and transcriptional activities by a large panel of drugs has been observed. Taken together, these observations suggest that CTD phosphorylation by P-TEFb is required to prevent arrest of elongating pol II.

Mammalian NELF is a multiprotein complex composed of novel -66, -61, -59, -58 and -46 kDa polypeptides with potential RNA-binding activity. Mammalian DSIF is a heterodimer of 160 kDa and 16 kDa subunits, which are
homologs of the *S. cerevisiae* Spt4 and Spt5 proteins. The exact mechanism by which DSIF, NELF and P-TEFb act together to regulate elongation by pol II is presently unknown, but evidence suggests that DSIF and NELF negatively regulate elongation through interactions with polymerase containing a hypophosphorylated CTD (Ref. 25). Thus, phosphorylation of the pol II CTD by P-TEFb might promote elongation by blocking interactions of DSIF and NELF with the elongation complex.

Evidence supporting roles for P-TEFb, DSIF and NELF in transcription *in vivo* has come from genetic studies in yeast and investigations of transcription of the HIV-1 long terminal repeat (LTR) in mammalian cells. First, some *spt4* and *spt5* mutations in *S. cerevisiae* render yeast sensitive to the drug 6-azauracil, which reduces intracellular ribonucleoside triphosphate concentrations and thus is expected to decrease pol II elongation rates *in vivo*. Second, some conditional *spt5* mutations are suppressed when combined with pol II mutations that impair elongation. Third, some *spt4* and *spt5* mutations produce synthetic lethal phenotypes when combined with pol II or elongation factor SII mutations. Finally, both P-TEFb and its CTD kinase activity are required for Tat-dependent elongation by pol II from the HIV-1 LTR, where P-TEFb is recruited to the HIV-1 *polyprotein* gene through a specific interaction of its cyclin T1 subunit with RNA-bound Tat (Refs 28, 29, 33). Tat-dependent transcription requires additional elongation factors, including Tat-SF. Interestingly, P-TEFb, DSIF and Tat-SF are all components of a larger complex, Tat-SF (Tat stimulatory factor), which can be recruited to the HIV-1 promoter, along with pol II, to provide elongation factors needed for Tat-dependent transcription (Ref. 34 and references therein).

**RNA polymerase II elongation factors**

The hunt for elongation factors has led to the discovery of a collection of cellular proteins that directly target transcribing pol II (Table 1). These proteins fall into two broad functional classes based on their abilities to either reactivate arrested pol II or prevent transient pausing by the enzyme. To date, the first class of elongation factors is composed solely of members of the SII family. As discussed above and described in detail in recent reviews, SII reactivates arrested pol II by triggering a polymerase-associated endoribonuclease that cleaves the nascent transcript upstream of its 3'-OH terminus, thereby creating a new 3'-OH terminus that is correctly positioned with respect to the polymerase catalytic site and can be re-extended. Although the precise function of SII family members in cells is not yet clear, Reines and co-workers have shown that SII mutations can lead to a dramatic reduction in total poly(A)
RNA, as well as in a number of specific transcripts in *S. cerevisiae*28. Whether SII directly regulates expression of many genes or whether its primary targets are a few important transcriptional regulators remains to be determined.

A diverse collection of proteins, including TFIIJ, ELL and Elongin, are capable of suppressing transient pausing by pol II (Refs 35,37,38). Recently, several additional elongation factors that suppress pausing have been discovered; these include the Cockayne syndrome B (CSB) protein29, which was originally identified by its requirement in transcription-coupled nucleotide excision repair (NER) and which is mutated in individuals suffering from the genetic disorder Cockayne syndrome32; the Tat-Sf1 (Ref. 40) and CA150 proteins41, which were originally identified by their abilities to promote elongation by pol II from the HIV-1LTR and the ELL complex32,43, which is composed of ELL, two novel proteins and a mammalian homolog of *S. cerevisiae* Snf8, which is required for efficient derepression of glucose-repressed genes.

Studies on the basic mechanics of elongation by RNA polymerases suggest that elongation factors which suppress pausing will have crucial roles in mRNA synthesis. First, purified pol II is unable to extend RNA chains in vitro at rates sufficient to account for the observed rates of mRNA synthesis in vivo; whereas mRNA synthesis is carried out efficiently in eukaryotic cell cultures at rates of 1200–1500 nucleotides/min, elongation by purified mammalian pol II proceeds optimally in vitro at rates of only 300–400 nucleotides/min44. Second, mechanistic studies suggest that pausing forms the rate-limiting step(s) in elongation45. Transcribing RNA polymerases are susceptible to pausing for varying lengths of time at each step of nucleotide addition. The duration of pausing is often greater than $K_{cat}$, arguing that elongating polymerase cycles between active and inactive conformations at each step of nucleotide addition. Thus, the rate of nucleotide addition is limited, not by $K_{cat}$ but by the fraction of time polymerase spends in an inactive conformation. Elongation factors like TFIIJ, Elongin and ELL are capable of stimulating the rate of elongation by RNA polymerase II similarly, both at saturating ribonucleoside triphosphate concentrations, when the time required for nucleotide addition varies from milliseconds to seconds, and at extremely low ribonucleoside triphosphate concentrations, when the time required for nucleotide addition can be >30 min. These factors therefore appear to increase the rate of elongation by decreasing the fraction of time RNA polymerase II spends in an inactive, paused conformation37,45.

**Elongation and chromatin**

In addition to elongation factors that interact directly with pol II and suppress pausing or arrest, a collection of proteins that promote elongation by modifying chromatin structure have been identified (Table 1). An ATP-dependent activity in fractions containing the SWI/SNF complex can promote elongation by pol II from the *HSP70* promoter by remodelling nucleosome(s) just downstream of the transcriptional start site46. HMG14 can stimulate elongation by pol II on histone H1-containing SV40 minichromosomes and might function by counteracting H1 activity, so that the local chromatin structure becomes less compact47. FACT, composed of Spt16 and the HMG1-like protein SSRP, stimulates elongation by pol II on chromatin templates that lack H1. FACT interacts specifically with nucleosomes and with H2A-H2B dimers and thus might function by promoting nucleosome disassembly during transcription48.

The three-subunit *S. cerevisiae* Elongator complex interacts with elongating pol II containing a hyperphosphorylated CTD. Like mutations in Spt4, Spt5 and SII, mutations in Elongator subunits render yeast sensitive to 6-azauracil. In addition, genetic interactions between Elongator subunits and SII have been observed. Finally, a variety of genes including *GAL1-10*, *PHO5* and *INO2* exhibit delays in activator-dependent expression in yeast containing Elongator mutations. In light of the recent observation that Elongator has intrinsic HAT activity, it is possible that Elongator can promote efficient elongation by pol II by modifying and destabilizing nucleosomes in the path of polymerase49.

**Future prospects**

Our understanding of the mechanism and regulation of elongation by pol II is improving rapidly. Over the past few years, novel elongation factors have been discovered and their mechanisms of action characterized in increasing detail. In the future, we expect that biochemical studies will focus on identifying the complete repertoire of elongation factors, defining their mechanisms of action and illuminating their roles in cell function. In addition, in light of growing evidence that DNA-binding transcriptional activators affect not only the rate of initiation but also the efficiency of elongation, we expect that future studies will focus on elucidating the role of known elongation factors in transcriptional activation and on identifying and characterizing additional proteins that participate in this process. In light of intriguing evidence that the pol II elongation complex interacts functionally with DNA-repair proteins, like CSB (Ref. 39), and DNA-recombination proteins, like Hrp1 andTho2 (Ref. 50), we expect that future studies will focus on establishing how interactions of these proteins with the elongation complex affects transcription, and what significance the integration of transcription, repair and recombination has for cell function. Finally, in light of very recent evidence that the pol II elongation complex interacts functionally with proteins involved in mRNA capping, splicing and polyadenylation51, it is clear that future investigations will be pursued in the context of a new paradigm that portrays the elongation complex not only as a target of regulation by upstream signals but also as a critical regulator of downstream events in expression of eukaryotic genes.

**Acknowledgements**

Work in the authors’ laboratories is supported by ACS Grant RP69921801 (A.S.), NSF Grant MCB-9817004 (A.D.), NIH Grant GM41628 and funds provided to the Oklahoma Medical Foundation by the H.A. and Mary K. Chapman Charitable Trust (J.W.C. and R.C.C.). J.W.C. is an Associate Investigator of the Howard Hughes Medical Institute and A.S. is an Edward Mallinckrodt Young Investigator.

**References**

8 Yan, Q. et al. (1999) Dual roles for TFIIH in promoter escape by RNA polymerase II. J. Biol. Chem. 274, 35668–35675
In the July issue of *TIBS*, we published an article by Ibba *et al.* (*TIBS* 25, 311–316). The authors noticed an error in the table in Box 2. The aminoacylation site of Class I and Class II should read ‘2′-OH of tRNA’ and ‘3′-OH of tRNA (except PheRS)’, respectively.