JB Minireview

J. Biochem. **141,** 601–608 (2007)
doi:10.1093/jb/mvm090

Phosphorylation of the C-terminal Domain of RNA Polymerase II Plays Central Roles in the Integrated Events of Eucaryotic Gene Expression

Yutaka Hirose and Yoshiaki Ohkuma*

Laboratory of Gene Regulation, Graduate School of Medical and Pharmaceutical Sciences, University of Toyama, 2630 Sugitani, Toyama 930-0194, Japan

Received December 18, 2006; accepted March 15, 2007; published online April 3, 2007

RNA polymerase II (Pol II) is the only polymerase to possess heptapeptide repeats in the C-terminal domain (CTD) of its large subunit. During transcription, CTD phopshorylation occurs and is maintained from initiation to termination. To date, among the three known CTD kinases possessing CDK–cyclin pairs, TFIIH is the only one that forms a preinitiation complex. The Mediator complex plays essential roles in transcription initiation and during the transition from initiation to elongation by transmitting signals from transcriptional activators to Pol II. P-TEFb specifically plays a role in transcription elongation. TFIIH and mediator phosphorylate serine 5 (Ser5) of the CTD heptapeptide repeat sequence, whereas P-TEFb phosphorylates serine 2 (Ser2). Recently, it has become clear that CTD phosphorylation is not only essential for transcription, but also as a platform for RNA processing and chromatin regulation. In this review, we discuss the central role of Pol II phosphorylation in these nuclear events.

Key words: CTD phosphorylation, mediator, phosphorylated CTD-interacting factor 1 (PCIF1), RNA polymerase II, TFIIH.

Abbreviations: CTD, carboxy-terminal domain; Pol II, RNA polymerase II; P-TEFb, positive transcription elongation factor b.

In eucaryotes, RNA polymerase II (Pol II), a member of the RNA polymerase family made up of three RNA polymerases (Pol I, Pol II and Pol III), plays a central role in transcription of protein-coding genes (*1*). However, contrary to bacterial RNA polymerase which is unique and transcribes all bacterial genes by itself, Pol II requires five general transcription factors (TFIIB, -D, -E, -F and -H) for accurate transcription initiation (*2*, *3*). In addition to this auxiliary factor requirement, Pol II has characteristic heptapeptide (YSPTSPS) repeats, referred to the CTD (C-terminal domain), in the carboxy (C)-terminus of its largest subunit Rpb1 (*4*). The function of this domain remained an enigma until the late 1980s when it was found that deletion of yeast CTD from 27 repeats to less than 10 caused severe lethality (*5*).

The first indication that the region was functionally relevant was evidence showing that it was involved in transcription. Pol II CTD was found to be phosphorylated by TFIIH upon completion of active preinitiation complex formation on a promoter located adjacent to the transcription initiation site. Phosphorylation is also dependent on the recruitment of TFIIE (*6*). TFIIH possesses 10 subunits and its CTD kinase activity is supported by the CDK/cyclin pair, CDK7 and cyclin H (*7*). At present, two other CTD kinases, the Mediator complex containing CDK8 and cyclin C as well as the transcription elongation factor P-TEFb containing CDK9 and cyclin T, have been reported (*8*, *9*). In addition to its functions in transcription, the phosphorylated CTD was found to control both steps of chromatin remodelling and RNA processing by providing a platform for the recruitment of factors involved in these two steps (*10*–*12*). Here we review these integrated nuclear events in gene expression and their coupling to CTD phosphorylation.

RNA POLYMERASE II CHANGES ITS STRUCTURE AND FUNCTION AT TRANSCRIPTION INITIATION

As described earlier, Pol II cannot initiate transcription alone. Five general transcription factors assist Pol II in initiating transcription by forming a preinitiation complex on the core promoter of protein-coding genes. The core promoters are mainly classified into two types on the basis of whether they contain a TATA-containing element or not. Promoters that lack the TATA element contain instead an initiator element and/or a downstream promotor element (DPE). In the case of TATA-containing inducible promoters, preinitiation complex formation is initiated upon the binding of the TATA-binding protein (TBP) of TFIID binding to the TATA element (*2*, *3*). This induces bending of the promoter which assists TFIIB binding to TBP as well as to the TATA proximal downstream sequence of the promoter known as the BRE (TFIIB-recognition element) (*13*). TFIIB then binds to the promoter just upstream

*To whom correspondence should be addressed. Tel: +81-76-434-7540, Fax: +81-76-434-5171,
E-mail: ohkumay@pha.u-toyama.ac.jp

of the TATA element as well as to Pol II, which induces binding of the active centre of Pol II to the initiation site for the initiation of accurate mRNA synthesis. Recently, progress in structural biology and molecular biology has made it possible to observe the transcription machinery at the molecular and even at the atomic level. The Nobel Prize in Chemistry was conferred on Dr Roger Kornberg, a Professor at Stanford University (*14*), largely for progress in this field. His work showed clearly that TFIIB binds also to TFIIF-prebound Pol II and inserts its N-terminal half containing a Zn-ribbon motif into the pore reaching to the active centre of Pol II which is utilized as the exit site for RNA transcripts (*15*). By performing this function, TFIIB assists Pol II to initiate transcription accurately. Prebinding of TFIIF to Pol II assists the launch of Pol II at the transcription initiation site (+1). From the X-ray crystal structure of the Pol II elongation form, Pol II also appears to undergo structural opening within its clamp and jaw regions and bind to the initiation site in mutual collaboration with TFIIB and TFIIF (*16*). Following the above sequence of events, TFIIE next stabilizes the complex by binding to nearly all the general transcription factors within the preinitiation complex, including Pol II (*3*). TFIIE is known to function with Pol II on three occasions; (i) TFIIEβ binds to DNA at the promoter adjacent region (approximately from −9 to +2), assists TFIIH in opening the promoter through its DNA-helicase activity, and helps Pol II make contact with the single-stranded initiation site in its active centre (*17*), (ii) TFIIEα binds to TFIIH and stimulates TFIIH-mediated CTD phosphorylation during transcription initiation (*6*) and (iii) TFIIEβ binds to the Rpb2 subunit at the jaw region of Pol II before initiation, which initiates TFIIEα binding to the Rpb5 subunit at the clamp region of Pol II, resulting in a change in Pol II conformation from the open to the closed configuration (Fig. 1) (*18*). The Pol II X-ray crystal structure also shows that the bridge helix is essential for this conformational change (*19*). As mentioned earlier, TFIIH contributes last to preinitiation complex formation on DNA and to Pol II function through its three enzyme activities (DNA helicase, ATPase and CTD kinase) (*20*).

Mediator complexes are known to consist of more than 20 subunits and to play key roles in transmitting gene activation signals to transcriptional activators, including co-activators, histone modifying factors and chromatin remodellers. They mediate both the opening of chromatin structure and transcriptional activation (*21*). Recently, three Mediator subcomplexes were observed in human and two of them were found to contain the CDK8

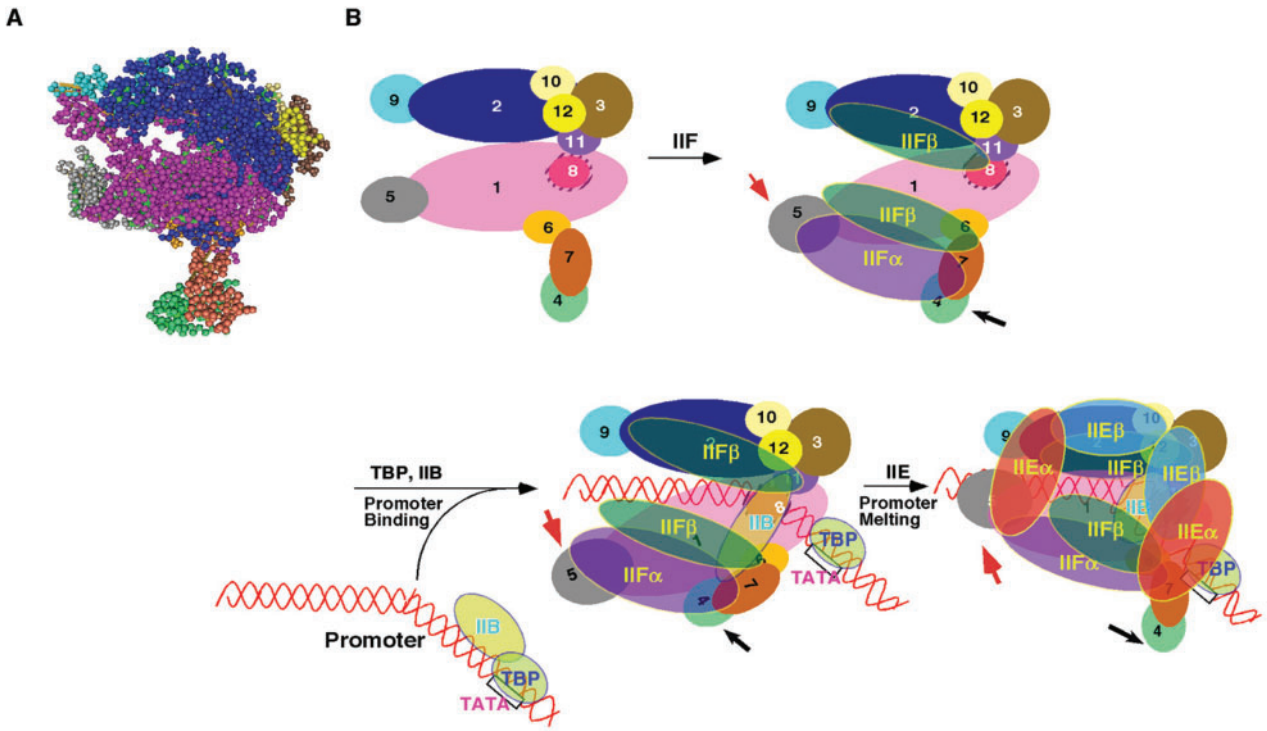

Fig. 1. **Pol II changes its conformation upon binding to various general transcription factors.** (A) Crystal structure of the whole Pol II. A space-filled structure model was drawn by the program Cn3D according to the structural information available at the PDB (accession number: 1NT9). (B) Schematic model of Pol II conformational changes. The positions of Pol II subunits were based on the structure in A. Twelve Pol II subunits (Rpb1-Rpb12) are numbered from 1 to 12, respectively. In solution, free Pol II binds to TFIIF (IIF) and changes its conformation. From the cryo-electron microscopy study, it is suggested that both TFIIF subunits (IIFa and IIFb) bind predominantly to Rpb2, Rpb4 and Rpb5 subunits, makes Pol II open up the clamp region (red arrow) and, at the same time, bends the Rpb4-Rpb7 subdomain (black arrow). Then the promoter together with the TATA box-bound TBP (or whole TFIID) and TFIIB (IIB) settles onto the active centre of Pol II. When TFIIE (IIE) is recruited, this binds to Rpb5, causes Pol II to close its structure and makes Pol II bind stably to the transcription initiation site of the promoter.

subunit (*22*). Mediators bind to the back side of the active centre of Pol II and to the CDK8 subunit, which together with cyclin C phosphorylates CTD in a manner similar to that of TFIIH (*23*).

Pol II undergoes conformational changes in two distinct ways. One is through physical interactions with general transcription factors and Mediators. TFIIB, TFIIF and TFIIE are definitely involved in Pol II conformational changes since they have been shown to associate physically with Pol II (*18*), and Mediators affect Pol II conformation in the same way (*23*). The other way is through CTD phosphorylation. As described earlier, TFIIH and CDK8-containing Mediators are CTD kinases and both phosphorylate Ser5 within Pol II (*24*, *25*). From structural analyses, Pol II consists of two characteristic modules; one is the core module consisting of the jaw and clamp submodules connected by the wall region, which altogether form a cleft with an active centre, and the other is the wedge module consisting of Rpb4 and Rpb7 subunits (*26*). CTD is located between the clamp and wedge and is connected to the Rpb1 core domain through the linker region.

CTD PHOSPHORYLATION OF POL II IS ESSENTIAL FOR TRANSCRIPTION

As noted earlier and shown in Fig. 2, the Pol II CTD is phosphorylated at Ser2 and Ser5 of the heptapeptide repeat sequence (YSPTSPS) during transcription. At initiation, Ser5 is the main target for CTD kinases, TFIIH and Mediators (*24*, *25*). During elongation, Ser2 is the main target for the third CTD kinase, P-TEFb (*27*). The importance of CTD phosphorylation was demonstrated by both *in vivo* and *in vitro* studies. From *in vivo* chromatin immunoprecipitation (ChIP) analyses in yeast,

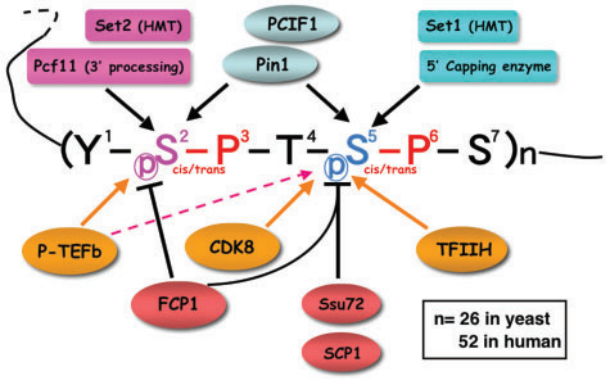

Fig. 2. **Regulation and recognition of the phosphorylated CTD.** Rpb1, the largest subunit of Pol II, has a unique C-terminal domain consisting of heptapeptide (YSPTSPS) repeats. The repeat number varies among different organisms, ranging from 26 in yeast to 52 in human. The CTD is mostly phosphorylated at Ser2 and Ser5 within the heptapeptide repeat during transcription. The kinases and phosphatases with specificity for Ser2 and Ser5 and the factors that bind to the CTD phosphorylated at Ser2, Ser5 or both, are indicated. Those factors include pre-mRNA processing factors, histone methyltransferases (HMT), nuclear peptidylprolyl *cis/trans* isomerase Pin1 and a novel WW domain containing protein PCIF1.

Pol II at the neighbouring transcription initiation site was shown to be heavily phosphorylated at Ser5 but much less at Ser2 (*24*). On the other hand, Pol II located within gene-coding regions or at neighbouring transcription termination sites was more phosphorylated at Ser2 than at Ser5. We could demonstrate the importance of CTD phosphorylation at Ser5 during early transcription by using *in vitro* biochemical studies: (i) TFIIE assisted TFIIH-dependent CTD phosphorylation only when the active transcription preinitiation complex was formed (*6*); (ii) in a chimeric TFIIE study using co-expressed human and nematode TFIIE subunits, we demonstrated that a TFIIE chimera made-up of human TFIIEα and nematode TFIIEβ had defects both in the transition from transcription initiation to elongation and in CTD phosphorylation at Ser5 (*28*). This strongly suggests that TFIIE is involved in the transcription transition stage and Ser5 phosphorylation of the CTD and, more importantly, that Ser5 phosphorylation might be essential for transition activity. In the case of Ser2, specific inhibition of P-TEFb kinase activity by a highly specific P-TEFb kinase inihibitor, flavopiridol, caused a reduction in Ser2 phosphorylation and a defect in the transcription elongation of heat shock genes (*27*). In combination with the *in vivo* results, it is obvious that Ser2 phosphorylation is tightly associated with transcription elongation.

Recently, it has become apparent that CTD phosphorylation at Ser2 and Ser5 is not only essential for transcription but also for chromatin remodelling at the gene downstream from the transcription initiation site as well as for RNA processing following transcription (*11*, *12*). As described earlier, Pol II is initially phosphorylated at Ser5 within CTD heptapeptide repeats during transcription initiation. Yeast studies showed that the recruitment of the histone methyltransferase Set1-containing complex, the capping enzyme and splicing enzymes after transcription initiation depended on this phosphorylation (*11*, *12*). Set1 tri-methylates histone H3 Lysine 4 (H3K4) in nucleosomes. Once Pol II has proceeded to the elongation phase, Ser2 becomes heavily phosphorylated. Additionally, phosphorylated Ser2 specifically recruits another histone methyltransferase Set2 and polyadenylation enzymes.

It has been suggested that Pol II is recycled to carry out multi-round transcription. Once Pol II reaches the termination signal at the end of transcription, it could be released from the gene at the same time as CTD dephosphorylation, resulting in its recycling for further rounds of transcription. This hypothesis is supported by the fact that only unphosphorylated Pol II is proficient for preinitiation complex formation (*29*). To date, several CTD phosphatases (FCP1, SCP1 and Ssu72) are known to preferentially function in Pol II recycling (*31*). There roles will be precisely explained in the latter part of this review.

RNA POLYMERASE II COORDINATES TRANSCRIPTION WITH RNA PROCESSING

Production of a functional mRNA in eucaryotes requires multistep processes including transcription of mRNA

precursors (pre-mRNAs) by Pol II and subsequent pre-mRNA processing. The Pol II CTD, especially in its phosphorylated form has been shown to play a central role in these processes by functioning as a loading platform for factors involved in pre-mRNA processing (*10*, *31*). The CTD is physically connected to the Pol II core complex through its linker region, which locates near the RNA exit pore of Pol II (*26*). This arrangement of the CTD may allow immediate and efficient packaging of a nascent transcript into an RNA–protein complex (RNP) with factors that have been recruited to the CTD during transcription. As described earlier, Ser2 and Ser5 phosphorylation and dephosphorylation of the CTD heptapeptide are dynamically regulated during the transcription cycle (Fig. 3). In the following sections, we will review recent studies that show how the phosphorylated CTD coordinates nuclear events for gene expression.

5′ CAPPING AND CTD

Addition of a m7GpppN cap structure to the 5′ end of all eucaryotic mRNAs (5′ capping) is the first pre-mRNA processing event and is catalysed by a series of three

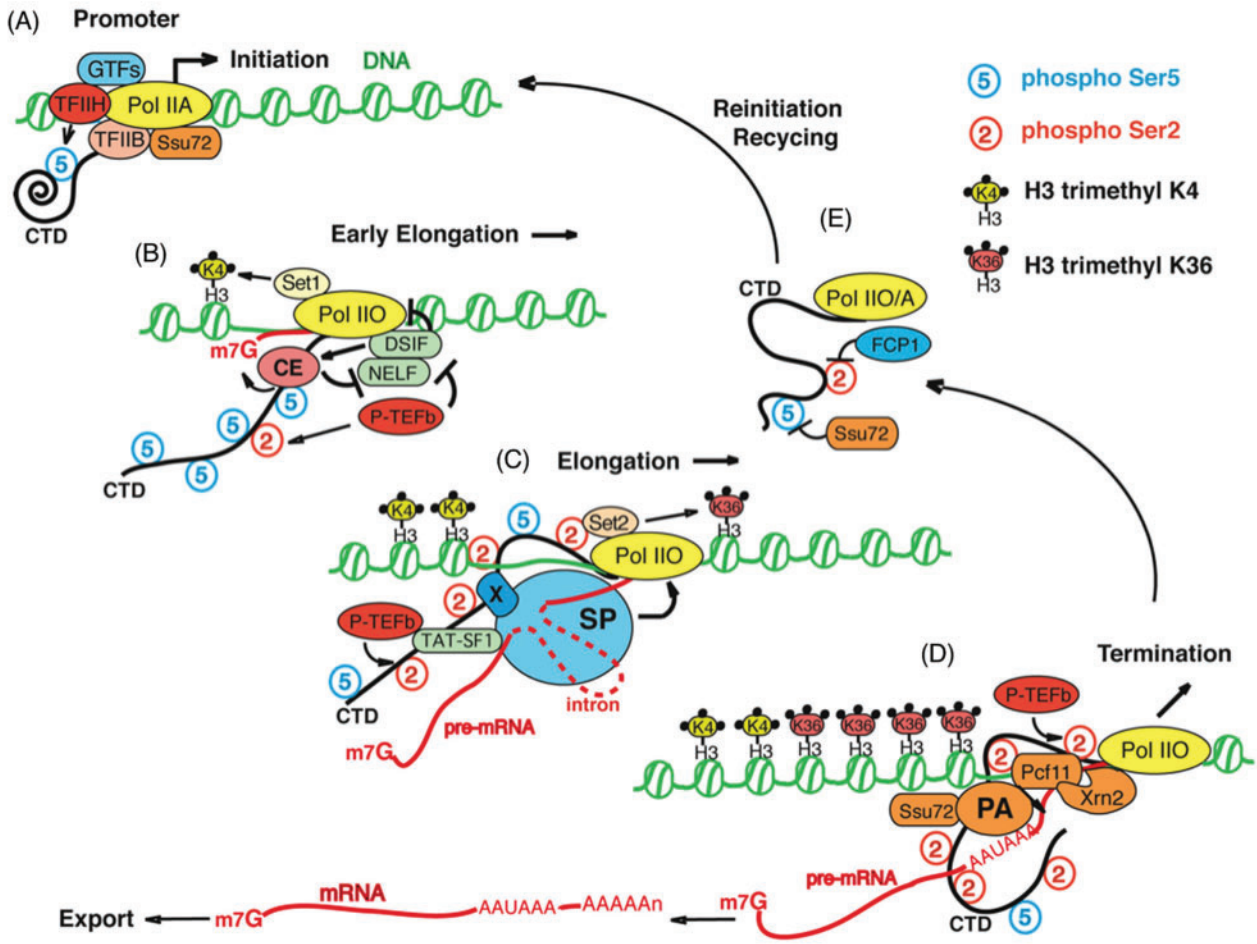

Fig. 3. **Dynamic changes in the CTD phosphorylation profile coordinate the Pol II transcription cycle with pre-mRNA processing and histone modification.** (A) The general transcription factors (GTFs) form a complex with initiation-competent hypo-phosphorylated Pol II (Pol IIA) at the promoter. Transcription starts at the same time as Ser5 phosphorylation of the CTD (thick black line) by TFIIH. (B) Shortly after transcription initiation, capping enzyme (CE) is recruited to the phosphorylated Pol II (Pol IIO) through its direct binding to Ser5-phosphorylated CTD. The histone methyltransferase Set1-containing complex is also recruited and tri-methylates histone H3 Lysine 4 (H3K4). Transcription pausing induced by DSIF/NELF is relieved by P-TEFb-mediated CTD phosphorylation. (C) Elongating Pol IIO is increasingly phosphorylated at Ser2 by P-TEFb and associated with histone methyltransferase Set2, which tri-methylates histone H3 Lysine 36 (H3K36). Pol IIO also helps the recruitment of the splicing machinery (SP), which splices sites in the pre-mRNA (red line). This step is mediated by an unknown phosphorylated CTD-binding factor (X) that facilitates the efficient excision of introns (red broken line). (D) Near the 3′ end of the gene, 3′ end processing factors (PA) are increasingly recruited to Pol IIO through direct interaction between Pcf11 and the Ser2-phosphorylated CTD. After transcribing the poly(A) signal (AATAAA), 3′ end processing factors possibly transfer to RNA to catalyse endonucleolytic cleavage (black arrow) and induce subsequent transcription termination, which is presumably helped by the 5′–3′ exonucleases Xrn2 and Pcf11. (E) After dissociating from the DNA template, Pol IIO is possibly dephosphorylated by the action of the CTD phosphatases, FCP1 and Ssu72, before recycling or reinitiation.

enzymatic activities, RNA 5′-triphosphatase (RT), guanylyltransferase (GT) and RNA (guanine-N7) methyltransferase (MT). In metazoans, the capping enzyme (CE) is bifunctional with both RT and GT activities, while in budding yeast, the capping enzyme consists of a heterodimer of RT (Cet1) and GT (Ceg1). 5′ capping occurs co-transcriptionally shortly after transcription initiation when the nascent RNA reaches 20–30 nt in length (*10*). Unlike other pre-mRNA processing steps, such as splicing or polyadenylation, this does not require specific RNA sequences or structures in the substrate. Nonetheless, it only occurs on RNAs transcribed by Pol II; RNAs transcribed by Pol I and Pol III are not polyadenylated. Both biochemical and genetic studies have revealed that specific and evolutionary conserved interactions between the CE and the phosphorylated CTD likely provide a basis for the specific and rapid targeting of the capping enzyme to Pol II transcripts. Early studies on CTD function in mammalian cells have shown that RNAs transcribed by CTD-truncated Pol II undergo inefficient 5′ capping (*32*).

The capping enzyme is recruited to the early transcription complex through specific interactions between the Ser5-phosphorylated CTD and the GT domain of mammalian CE. GT activity in turn is allosterically activated by binding to the Ser5-phosphorylated CTD peptide (*33*). In budding yeast, both GT (Cet1) and MT (Abd1) directly bind to the phosphorylated CTD. Chromatin immunoprecipitation (ChIP) analyses showed that Ser5 phosphorylation by the yeast Cdk7 (Kin28) subunit of TFIIH was required for recruitment of CE to the promoter region and that the CTD phosphatase Fcp1 was required for release of CE from the transcription elongation complex *in vivo* (*24*, *34*). Therefore, temporally and spatially regulated phosphorylation and dephosphorylation of the CTD are crucial for tight coupling of 5′ capping with early transcription.

Other protein–protein contacts are also involved in the specific recruitment of CE to the transcription machinery near promoters. In mammalian cells, the transcription elongation factor DSIF (DRB sensitivity-inducing factor), which is composed of human homologs of yeast Spt4 and Spt5, forms a complex with negative elongation factor (NELF) to induce transcription arrest at the promoter-proximal region. The negative effect of DSIF/NELF on transcription is relieved by P-TEFb-mediated CTD phosphorylation (*35*). CE is recruited to the DSIF/NELF complex through direct interaction with Spt5, which in turn stimulates GT activity, leading to attenuation of the negative effect of NELF on transcription (*36*). Thus, this network of protein–protein interactions may serve as an early transcription checkpoint to provide the appropriate time frame for the capping reaction and subsequent chain elongation of only capped nascent RNA (*31*) (Fig. 3).

SPLICING AND TRANSCRIPTION

In higher eucaryotes, many protein-coding genes are interrupted by one or more intervening sequences (introns). To make a functional mRNA, introns have to be accurately excised from an RNA precursor before being transported from the nucleus to the cytoplasm. Splicing occurs in a large complex called the spliceosome, which is composed of small nuclear ribonucleoprotein particles (snRNPs) and non-snRNP proteins including members of the serine/arginine-rich (SR) protein family (*37*). Both biochemical and *in vivo* studies have provided support for the existence of functional interactions between the Pol II CTD and the splicing apparatus. Like capping, splicing of RNAs transcribed by CTD-truncated Pol II is inefficient in mammalian cells (*38*). The CTD is required for targeting splicing factors to transcription sites *in vivo* (*39*). Furthermore, an *in vitro* study has demonstrated that purified phosphorylated Pol II stimulates splicing in the absence of transcription by accelerating the rate of one of the first steps in spliceosomal assembly (*40*). In *Xenopus* oocytes, CTD phosphorylation is required for co-transcriptional splicing (*41*). These observations support a model in which the phosphorylated CTD of elongating Pol II functions as a platform for the processing factors involved in splicing.

How does the CTD function in splicing and which factors serve to connect the CTD with the splicing machinery? Although some candidate proteins, including SCAFs (SR-like CTD-associated factors), yeast splicing factor Prp40, and CA150, have been shown to directly bind to the phosphorylated CTD (*42*), it still remains unclear whether these proteins actually function to connect the CTD to the splicing machinery. Recently, two independent studies using an *in vitro* system that detects co-transcriptional splicing demonstrated that immediate and specific spliceosome assembly on nascent RNA transcribed by Pol II results in efficient splicing by protecting RNA from degradation as well as by preventing the assembly of a non-specific complex (*43*, *44*). These *in vitro* systems may provide a useful tool for elucidating the molecular mechanisms involved in the coupling of these factors, including the identification of the factors that link splicing to transcription.

Transcription affects not only constitutive splicing but also alternative splicing, which allows the joining of different combinations of exons, potentially resulting in a multitude of diverse proteins from a single transcript. Recent studies support the idea that the transcription elongation rate or processivity contributes to splice site selection. The introduction of a Pol II-pause site near the alternative splice site or transcription by a mutant form of Pol II with a lower elongation rate stimulates the inclusion of alternative exons (*42*). Interestingly, an extension of the kinetic link between transcription and alternative splicing has been recently reported. Human Brm, a subunit of the chromatin remodelling complex SWI/SNF, has been shown to regulate alternative exon inclusion, perhaps by creating a ‘roadblock’ for Pol II elongation at the alternative exon-coding region. Intriguingly, Pol II pausing at this region is accompanied by a change in the phosphorylation pattern of the CTD (*45*). A reverse relationship in which splicing factors affect transcription has also been reported. Recruitment of spliceosomal snRNPs to elongating Pol II tethered by the transcription elongation factor TAT-SF1, which associates with the CTD kinase

P-TEFb, stimulated transcription elongation depending on whether or not splicing signals were present on the transcribed template (*46*).

TRANSCRIPTION AND 3′ END FORMATION

In eucaryotes, the pre-mRNA 3′ end is formed by a two-step reaction: endonucleolytic cleavage 10–30 nt downstream of a well-conserved poly(A) signal sequence followed by poly(A) addition to the 3′ end of the upstream cleavage product. This 3′ end processing requires multiple protein factors, including a cleavage/polyadenylation specificity factor (CPSF), a cleavage stimulation factor (CstF), two cleavage factors, CFIm and CFIIm, and poly(A) polymerase (PAP) in mammals and a cleavage-polyadenylation factor CPF, cleavage factors CF1A and CF1B, and PAP in yeast (*47*). Like other pre-mRNA processing events, 3′ end processing is intimately coupled to transcription and the CTD plays a central role in the linking of these two processes. RNAs transcribed by CTD-truncated Pol II are not efficiently polyadenylated in mammalian cells (*39*). A biochemical study has revealed that purified Pol II and its CTD activated 3′ cleavage in the absence of transcription (*48*). Several 3′ end processing factors both in mammals and in yeast have shown to directly interact with the CTD (*47*). These results suggest that the CTD stimulates the formation of a stable, catalytically active processing complex through direct interaction with the 3′ end processing factor, at least in mammalian cells.

Recent studies have provided evidence that Ser2 phosphorylation plays an important role in pre-mRNA 3′ end processing (*28*, *49*). A loss of Ser2 phosphorylation through inactivation of the CTD kinase Ctk1 in yeast or Cdk9 in *Drosophila* leads to a defect in processing but, unexpectedly, not in transcription elongation. Yeast ChIP analysis has shown that crosslinking of 3′ end processing factors to the 3′ end region of genes is dramatically decreased in Ctk1 deleted strains (*49*). These observations are consistent with another study showing that the yeast cleavage factor CF1A subunit, Pcf11, can preferentially bind to the Ser2-phosphorylated CTD through its CTD interacting domain (CID) (*50*). These results suggest that the increase in phosphorylation at Ser2 by the CTD kinase as transcription proceeds is required for efficient recruitment of 3′ end processing factors to the elongating Pol II.

3′ end processing is intimately coupled with transcription termination, in which the Pol II elongation complex dissociates from the DNA template and releases the nascent RNA. Transcription termination is necessary for the recycling of Pol II and for preventing inappropriate transcription of downstream genes. Early studies have shown that a functional poly(A) signal is required for termination in both mammals and yeast (*10*). Subsequent studies have revealed that RNAs transcribed by CTD-truncated Pol II are not normally terminated (*38*), and yeast 3′ end processing factors Rna14, Rna15, Pcf11 and Yhh1, all of which directly bind to the CTD, are also required for termination (*51*). Furthermore, a recent study has shown that yeast and *Drosophila* Pcf11 are able to dismantle an *in vitro* assembled elongation complex by bridging the CTD to the nascent RNA (*52*). These studies indicate that complex communication among 3′ end processing factors, the CTD, and RNAs containing poly(A) signals may mediate termination. It has been recently reported that the RNA 5′–3′ exonucleases, yeast Rat1 and human Xrn2 are necessary for efficient termination probably because they degrade the downstream cleaved products to catch up with the elongating Pol II (*53*, *54*). Rat1 associates with Rtt103 which possesses a CID domain and binds to the Ser2-phosphorylated CTD (*53*). Although the 5′–3′ exonuclease activity of Rat1 is necessary for termination, it is not sufficient (*54*). The precise mechanisms involved in transcriptional termination remain to be elucidated.

Surprisingly, several 3′ end processing factors are not only located at the 3′ ends of genes but also in promoter regions (*47*). Recently, the yeast 3′ end processing factor Ssu72, which is a component of CPF, has been shown to function in multiple stages of gene expression, including transcription initiation, cleavage/polyadenylation and the transcriptional termination of specific mRNAs and small nucleolar RNA (snoRNA) genes (*47*). Interestingly, Ssu72 has a CTD phosphatase activity with specificity for phosphorylated Ser5, which is required for transcription initiation but not for 3′ end processing (*55*). A recent study has shown that yeast CPF components, Ssu72 and Pta1, play roles in the formation of a transcription-dependent gene loop, in which the transcription terminator region physically interacts with the initiation region for the stimulation of Pol II recycling (*56*). The CTD phosphatase activity of Ssu72 is required for the gene looping (*56*).

REGULATING THE CTD STRUCTURE

As described earlier, dynamic changes in CTD phosphorylation during transcription are crucial for the coordination of transcription with pre-mRNA processing. Recent studies on the structures of the phosphorylated CTD-binding domain bound to differentially phosphorylated CTD peptides have revealed that the phosphorylated CTD can adapt different conformations depending on its binding partner, probably through an induced fit mechanism (*30*). In all structures, only the *trans*-conformations of Pro3 and Pro6 in the CTD heptapeptides were observed in the complex, indicating that proline isomerization may be critical for regulating CTD structure and recognition of its binding partners (*30*).

A nuclear peptidylprolyl *cis/trans* isomerase (PPIase), referred to as Ess1 in yeast and Pin1 in human, has been shown to bind to the phosphorylated CTD through its WW domain (*10*). Interestingly, Ess1 was originally identified in a genetic screen of mutants for defects in 3′ end formation (*10*). Although Ess1/Pin1 has been proposed to cause efficient 3′ end processing and/or subsequent transcription termination by affecting the local conformation of the phosphorylated CTD, the exact role of Ess1/Pin1 in this processing step remains unclear. Pin1 has recently been shown to negatively regulate Pol II activity at mitosis (*57*). Another human WW domain-containing nuclear protein, PCIF1, has been

recently reported to bind to the phosphorylated CTD (*58*). The WW domain of PCIF1 exhibits considerable homology to that of Pin1. Although PCIF1 associates with phosphorylated Pol II *in vivo* and may function to modulate Pol II activity, the precise roles of PCIF1 remain to be elucidated.

PERSPECTIVES

Recent progresses in this field have made it possible to visualize the mechanism of transcription from preinitiation complex assembly to initiation, and ultimately to elongation. Remaining questions often concern the mechanisms of transcriptional activation within the chromatin context and further the coordinated mechanisms for gene expression after transcription initiation including mechanisms for spatially and timely regulated recruitment of mRNA processing, surveillance and export factors to an elongating Pol II. As described, we know now that the phosphorylated CTD of Pol II gives a platform for recruitment of the factors that are involved in transcription, chromatin regulation as well as RNA processing. We are convinced that the crues for elucidation of the global mechanisms for gene expression will be given by the studies of the CTD phosphorylation and its related events including histone modifications and chromatin remodelling.

We wish to thank Aki Tanaka for critical reading of the manuscript and all the members of our laboratory for helpful discussions. This work was supported in part by grants-in-aid for priority areas from the Ministry of Education, Science, Sports and Culture of Japan (Y.H. and Y.O.).

REFERENCES

1. Young, R.A. (1991) RNA polymerase II. *Annu. Rev. Biochem.* **60**, 689–715
2. Roeder, R.G. (1996) The role of general initiation factors in transcription by RNA polymerase II. *Trends Biol. Sci.* **21**, 327–335
3. Ohkuma, Y. (1997) Multiple functions of general transcription factors TFIIE and TFIIH in transcription: possible points of regulation by *trans*-acting factors. *J. Biochem. (Tokyo)* **122**, 481–489
4. Corden, J.L. and Ingles, C.J. (1992) Carboxy-terminal domain of the largest subunit of eukaryotic RNA polymerase II. in *Transcriptional Regulation* (McKnight, S.L. and Yamamoto, K.R., eds.) pp. 81–107 Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
5. Nonet, M., Sweetser, D., and Young, R.A. (1987) Functional redundancy and structural polymorphism in the large subunit of RNA polymerase II. *Cell* **50**, 909–915
6. Ohkuma, Y. and Roeder, R.G. (1994) Regulation of TFIIH ATPase and kinase activities by TFIIE during active initiation complex formation. *Nature* **368**, 160–163
7. Giglia-Mari, G., Coin, F., Ranish, J.A., Hoogstraten, D., Theil, A., Wijgers, N., Jaspers, N.G., Raams, A., Argentini, M., van der Spek, P.J., Botta, E., Stefanini, M., Egly, J.-M., Aebersold, R., Hoeijmakers, J.H., and Vermeulen, W. (2004) A new, tenth subunit of TFIIH is responsible for the DNA repair syndrome trichothiodystrophy group A. *Nat. Genet.* **36**, 714–719
8. Liu, Y., Kung, C., Fishburn, J., Ansari, A.Z., Shokat, K.M., and Hahn, S. (2004) Two cyclin-dependent kinases promote RNA polymerase II transcription and formation of the scaffold complex. *Mol. Cell. Biol.* **24**, 1721–1735
9. Meinhart, A., Kamenski, T., Hoeppner, S., Baumli, S., and Cramer, P. (2005) A structural perspective of CTD function. *Genes Dev.* **19**, 1401–1415
10. Hirose, Y. and Manley, J.L. (2000) RNA polymerase II and the integration of nuclear events. *Genes Dev.* **14**, 1415–1429
11. Orphanides, G. and Reinberg, D. (2002) A unified theory of gene expression. *Cell* **108**, 439–451
12. Hampsey, M. and Reinberg, D. (2003) Tails of intrigue: phosphorylation of RNA polymerase II mediates histone methylation. *Cell* **113**, 429–432
13. Lagrange, T., Kapanidis, A.N., Tang, H., Reinberg, D., and Ebright, R.H. (1998) New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. *Genes Dev.* **12**, 34–44
14. Service, R.F. (2006) Nobel Prize: Chemistry. Solo winner detailed path from DNA to RNA. *Science* **314**, 236
15. Bushnell, D.A., Westover, K.D., Davis, R.E., and Kornberg, R.D. (2004) Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 Å. *Science* **303**, 983–988
16. Gnatt, A.L., Cramer, P., Fu, J., Bushnell, D.A., and Kornberg, R.D. (2001) Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science* **292**, 1844–1846
17. Douziech, M., Coin, F., Chipoulet, J.M., Arai, Y., Ohkuma, Y., Egly, J.-M., and Coulombe, B. (2000) Mechanism of promoter melting by the xeroderma pigmentosum complementation group B helicase of transcription factor IIH revealed by protein-DNA photo-cross-linking. *Mol. Cell Biol.* **20**, 8168–8177
18. Hayashi, K., Watanabe, T., Tanaka, A., Furumoto, T., Sato-Tsuchiya, C., Kimura, M., Yokoi, M., Ishihama, A., Hanaoka, F., and Ohkuma, Y. (2005) Studies of *Schizosaccharomyces pombe* TFIIE indicate conformational and functional changes in RNA polymerase II at transcription initiation. *Genes Cells* **10**, 207–224
19. Westover, K.D., Bushnell, D.A., and Kornberg, R.D. (2004) Structural basis of transcription: nucleotide selection by rotation in the RNA polymerase II active center. *Cell* **119**, 481–489
20. Svejstrup, J.Q., Vichi, P., and Egly, J.M. (1996) The multiple roles of transcription/repair factor TFIIH. *Trends Biochem. Sci.* **21**, 346–350
21. Malik, S. and Roeder, R.G. (2005) Dynamic regulation of pol II transcription by the mammalian Mediator complex. *Trends Biochem. Sci.* **30**, 256–263
22. Furumoto, T., Tanaka, A., Ito, M., Malik, S., Hirose, Y., Hanaoka, F., and Ohkuma, Y. (2006) A kinase subunit of the human Mediator complex, CDK8, positively regulates transcriptional activation. *Genes Cells*(in press)
23. Davis, J.A., Takagi, Y., Kornberg, R.D., and Asturias, F.A. (2002) Structure of the yeast RNA polymerase II holoenzyme: Mediator conformation and polymerase interaction. *Mol. Cell* **10**, 409–415
24. Komarnitsky, P., Cho, E.J., and Buratowski, S. (2000) Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev.* **14**, 2452–2460
25. Hengartner, C.J., Myer, V.E., Liao, S.M., Wilson, C.J., Koh, S.S., and Young, R.A. (1998) Temporal regulation of RNA polymerase II by Srb10 and Kin28 cyclin-dependent kinases. *Mol. Cell* **2**, 43–53
26. Armache, K.J., Kettenberger, H., and Cramer, P. (2003) Architecture of initiation- competent 12-subunit RNA polymerase II. *Proc. Natl Acad. Sci. USA* **100**, 6964–6968
27. Ni, Z., Schwartz, B.E., Werner, J., Suarez, J.R., and Lis, J.T. (2004) Coordination of transcription, RNA processing,

and surveillance by P-TEFb kinase on heat shock genes. *Mol. Cell* **13**, 55–65
28. Yamamoto, S., Watanabe, Y., van der Spek, P.J., Watanabe, T., Fujimoto, H., Hanaoka, F., and Ohkuma, Y. (2001) Studies of nematode TFIIE function reveal a link between Ser-5 phosphorylation of RNA polymerase II and the transition from transcription initiation to elongation. *Mol. Cell Biol.* **21**, 1–15
29. Lu, H., Flores, O., Weinmann, R., and Reinberg, D. (1991) The nonphosphorylated form of RNA polymerase II preferentially associates with the preinitiation complex. *Proc. Natl Acad. Sci. USA* **88**, 10004–10008
30. Meinhart, A., Kamenski, T., Hoeppner, S., Baumli, S., and Cramer, P. (2005) A structural perspective of CTD function. *Genes Dev.* **19**, 1401–1415
31. Sims III, R.J., Belotserkovskaya, R., and Reinberg, D. (2004) Elongation by RNA polymerase II: the short and long of it. *Genes Dev.* **18**, 2437–2468
32. McCracken, S., Fong, N., Rosonina, E., Yankulov, K., Brothers, G., Siderovski, D., Hessel, A., Foster, S., Shuman, S., and Bentley, D.L. (1997) 5′-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes Dev.* **11**, 3306–3318
33. Ho, C.K. and Shuman, S. (1999) Distinct roles for CTD Ser-2 and Ser-5 phosphorylation in the recruitment and allosteric activation of mammalian mRNA capping enzyme. *Mol. Cell* **3**, 405–411
34. Schroeder, S.C., Schwer, B., Shuman, S., and Bentley, D.L. (2000) Dynamic association of capping enzymes with transcribing RNA polymerase II. *Genes Dev.* **14**, 2435–2440
35. Yamaguchi, Y., Takagi, T., Wada, T., Yano, K., Furuya, A., Sugimoto, S., Hasegawa, J., and Handa, H. (1999) NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. *Cell* **97**, 41–51
36. Mandal, S.S., Chu, C., Wada, T., Handa, H., Shatkin, A.J., and Reinberg, D. (2004) Functional interactions of RNA-capping enzyme with factors that positively and negatively regulate promoter escape by RNA polymerase II. *Proc. Natl Acad. Sci. USA* **101**, 7572–7577
37. Moore, M.J., Query, CC., and Sharp, P.A. (1993) Splicing of precursors to mRNAs by the spliceosome. in *The RNA World* (Gesteland, R.F. and Atkins, J.F., eds.) pp. 303–358 Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
38. McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S.D., Wickens, M., and Bentley, D.L. (1997) The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature* **385**, 357–361
39. Misteli, T. and Spector, D.L. (1999) RNA polymerase II targets pre-mRNA splicing factors to transcription sites in vivo. *Mol. Cell* **3**, 697–705
40. Hirose, Y., Tacke, R., and Manley, J.L. (1999) Phosphorylated RNA polymerase II stimulates pre-mRNA splicing. *Genes Dev.* **13**, 1234–1239
41. Bird, G., Zorio, D.A., and Bentley, D.L. (2004) RNA Polymerase II carboxy-terminal domain phosphorylation is required for cotranscriptional pre-mRNA splicing and 3'- end formation. *Mol. Cell. Biol.* **24**, 8963–8969
42. Kornblihtt, A.R., de la Mata, M., Fededa, J.P., Munoz, M.J., and Nogues, G. (2004) Multiple links between transcription and splicing. *RNA* **10**, 1489–1498
43. Das, R., Dufu, K., Romney, B., Feldt, M., Elenko, M., and Reed, R. (2006) Functional coupling of RNAP II transcription to spliceosome assembly. *Genes Dev.* **20**, 1100–1109
44. Hicks, M.J., Yang, C.R., Kotlajich, M.V., and Hertel, K.J. (2006) Linking splicing to Pol II transcription stabilizes pre-mRNAs and influences splicing patterns. *PLoS Biol.* **4**, e147
45. Batsche, E., Yaniv, M., and Muchardt, C. (2006) The human SWI/SNF subunit Brm is a regulator of alternative splicing. *Nat. Struct. Mol. Biol.* **13**, 973–980
46. Fong, Y. and Zhou, Q. (2001) Stimulatory effect of splicing factors on transcriptional elongation. *Nature* **414**, 929–932
47. Proudfoot, N.J., Furger, A., and Dye, M.J. (2002) Integrating mRNA processing with transcription. *Cell* **108**, 501–512
48. Hirose, Y. and Manley, J.L. (1998) RNA polymerase II is an essential mRNA polyadenylation factor. *Nature* **395**, 93–96
49. Ahn, S.H., Kim, M., and Buratowski, S. (2004) Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3'-end processing. *Mol. Cell* **13**, 67–76
50. Licatalosi, D.D., Geiger, G., Minet, M., Schroeder, , S. Cilli, K., McNeil, J.B., and Bentley, D.L. (2002) Functional interaction of yeast pre-mRNA 3'-end processing factors with RNA polymerase II. *Mol. Cell* **9**, 1101–1111
51. Emanuel, R., Kaneko, S., and Manley, J.L. (2006) Terminating the transcript: breaking up is hard to do. *Genes Dev.* **20**, 1050–1056
52. Zhang, Z. and Gilmour, D.S. (2006) Pcf11 is a termination factor in *Drosophila* that dismantles the elongation complex by bridging the CTD of RNA polymerase II to the nascent transcript. *Mol. Cell* **21**, 65–74
53. Kim, M., Krogan, N.J., Vasiljeva, L., Rando, O.J., Nedea, E., Greenblatt, J.F., and Buratowski, S. (2004) The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature* **432**, 517–522
54. West, S., Gromak, N., and Proudfoot, N.J. (2004) Human 5'→3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. *Nature* **432**, 522–525
55. Krishnamurthy, S., He, X., Reyes-Reyes, M., Moore, C., and Hampsey, M. (2004) Ssu72 is an RNA polymerase II CTD phosphatase. *Mol. Cell* **14**, 387–394
56. Ansari, A. and Hampsey, M. (2005) A role for the CPF 3'-end processing machinery in RNAP II-dependent gene looping. *Genes Dev.* **19**, 2969–2978
57. Xu, Y.X., Hirose, Y., Zhou, X.Z., Lu, K.P., and Manley, J.L. (2003) Pin1 modulates the structure and function of human RNA polymerase II. *Genes Dev.* **17**, 2765–2776
58. Fan, H., Sakuraba, K., Komuro, A., Kato, S., Harada, F., and Hirose, Y. (2003) PCIF1, a novel human WW domain-containing protein, interacts with the phosphorylated RNA polymerase II. *Biochem. Biophys. Res. Commun.* **301**, 378–385