REVIEW

RNA polymerase II and the integration of nuclear events

Yutaka Hirose[2] and James L. Manley[1,3]

[1]Department of Biological Sciences, Columbia University, New York, New York 10027 USA; [2]Department of Molecular and Cellular Biology, Cancer Research Institute, Kanazawa University, Ishikawa 920-0934, Japan

The synthesis of a messenger RNA in the nucleus of a eukaryotic cell is an immensely complex undertaking. Each step in the pathway requires an enormous number of protein factors and identifying them and figuring out how they work has been a major goal of molecular biologists for the last two decades. Based on in vitro assays showing that each of the major steps, that is, transcription, capping, splicing, and polyadenylation, can be carried out in isolation, and because intuitively each of these reactions seemed quite distinct from the others, it had been widely assumed that the machinery responsible for each step was distinct and functioned essentially independently. However, numerous studies during the last few years have provided considerable evidence that this is not the case. In retrospect, this conclusion had been foreshadowed by earlier experiments pointing to the possibility that any one of these reactions could enhance some aspect of another. For example, evidence was presented consistent with the idea that the mRNA 5′ cap could play a role in allowing efficient transcription (Jove and Manley 1982), splicing (Edery and Sonenberg 1985), and even polyadenylation (Hart et al. 1985). Subsequently, it was shown in several labs that an intact polyadenylation signal could be required for transcription termination by RNA polymerase II (RNAP II) (Whitelaw and Proudfoot 1986; Logan et al. 1987; Connelly and Manley 1988), and that the presence of splicing signals on a pre-mRNA could enhance polyadenylation and vice versa (Niwa et al. 1990; Niwa and Berget 1991). However, none of these interactions really suggested just how intimate these associations might be, especially the emergence of RNAP II as an important component of all these reactions: capping, splicing, polyadenylation, as well as of course transcription.

The largest subunit of RNAP II has a unique domain, not related to regions in any known protein, at its carboxyl terminus, termed the carboxy-terminal domain (CTD). The CTD consists of multiple repeats of an evolutionary conserved heptapeptide with the consensus sequence Tyr-Ser-Pro-Thr-Ser-Pro-Ser (for review, see Corden 1990). The number of the repeats varies among different organisms, ranging from 26–27 in yeast to 52 in mammals. In metazoans, there can be significant degeneracy at some positions in the CTD, in mammals this is most apparent in the most carboxy-terminal repeats. The significance of this degeneracy is currently unknown. The CTD is rich in potential phosphoacceptor amino acid residues and, in keeping with this, is subject to reversible phosphorylation during the transcription cycle (for review, see Dahmus 1996). RNAP II with a hypophosphorylated CTD (RNAP IIA) is included preferentially in the transcription preinitiation complex formed at the promoter, whereas RNAP II with a hyperphosphorylated CTD (RNAP IIO) is associated with elongation complexes. Not unexpectedly, the CTD plays an important role in transcription, especially transcription initiation (for review, see Carlson 1997).

In this review, we discuss recent progress relating to what might be called the integration of nuclear events. Our focus will be on studies aimed at deciphering how RNAP II functions in the various RNA processing reactions needed to synthesize a mature mRNA. Much of what we will discuss is illustrated in the model shown in Figure 1. The reader is also referred to several excellent related reviews that have appeared recently (Neugebauer and Roth 1997; Steinmetz 1997; Bentley 1999; Minvielle-Sebastian and Keller 1999).

mRNA capping

It has been known for some time that the cap structure found at the 5′ end of all eukaryotic mRNAs is formed shortly after transcription initiation, when nascent RNA chains are about 25–30 nucleotides in lengths (see, for example, Coppola et al. 1983; Jove and Manley 1984). Capping is carried out by a series of three enzymatic activities (for review, see Shuman 1995). RNA triphosphatase removes the γ-phosphate of the first nucleotide of the pre-mRNA, followed by the transfer of GMP to the resulting diphosphate end by RNA guanylyltransferase. RNA (guanine-7-) methyltransferase then adds a methyl group to the N7 position of the cap guanine to form the m7G(5′)ppp(5′)N cap. In metazoans, the capping enzyme is bifunctional with both RNA 5-triphosphatase and RNA guanylyltransferase activities, while in *Saccharomyces cerevisiae*, capping enzyme consists of a hetero-

[3]Corresponding author.
E-MAIL jlm2@columbia.edu; FAX (212) 865-8246.
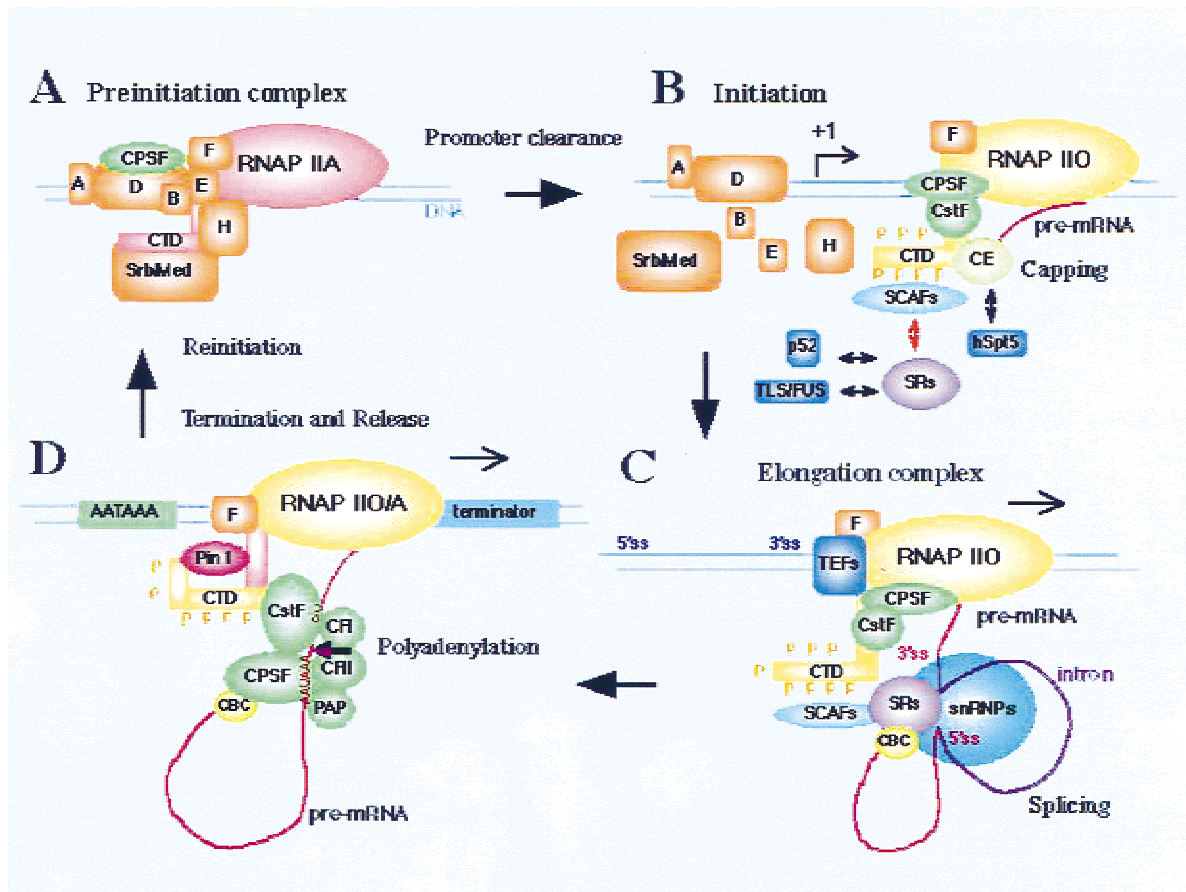
Hirose and Manley

Figure 1. Linking pre-mRNA processing to the RNAP II transcription cycle. (*A*) The general transcription factors (A–H) and srb/mediator (SrbMed) complex, represented by orange squares, form the preinitiation initiation complex with RNAP IIA at the promoter (for review, see Orphanides et al. 1996). The polyadenylation factor CPSF can also be found in this complex. (*B*) Shortly after transcription initiation, capping enzyme (CE) is recruited to and activated by the phosphorylated CTD of RNAP IIO. SCAF proteins can also associate with the phosphorylated CTD and may mediate the recruitment of SR proteins to RNAP IIO. However, interactions between SCAFs and SR proteins (indicated by a red double-headed arrow) and a role for SCAFs in splicing have not been experimentally demonstrated. Specific, apparently functional interactions between certain transcription factors (blue squares) and/or CE or SR proteins are indicated by double-headed arrows (see text for details). (*C*) Elongating RNAP IIO is associated with transcription elongation factors (TEFs: blue square; for review, see Reines et al. 1999) and helps in the recruitment of the splicing machinery (SR proteins, snRNPs) to splice sites in the pre-mRNA to facilitate efficient excision of introns (purple line). CBC represents the cap-binding complex, which has been suggested to be capable of stimulating both splicing and polyadenylation (Flaherty et al. 1997 and references therein). (*D*) After transcribing the poly(A) signal (AATAAA), polyadenylation factors (green ovals) associated with the CTD form a functional complex on the pre-mRNA to catalyze endonucleolytic cleavage (indicated by a purple arrow). Pin1 may stimulate a conformational change of the phosphorylated CTD, enhancing the efficiency of 3'-end formation and/or subsequent transcription termination. Whether or not TFIIF is present in termination complexes is unknown.

dimer of RNA triphosphatase (Cet1) and RNA guanylyl-transferase (Ceg1). Although capping has been known to occur only on RNAs made by RNAP II, the mechanism for this specificity and what insures rapid, efficient capping has been elucidated only recently.

We now know that RNAP II0, and specifically the CTD, plays a direct role in the capping reaction (Fig. 2). Mammalian capping enzyme can directly and selectively bind RNAP IIO (Yue et al. 1997) and the phosphorylated form of a recombinant glutathione-S transferase-CTD fusion protein (GST–CTD; McCracken et al. 1997a; Ho et al. 1998a) through its guanylyltransferase domain. In budding yeast, guanylyltransferase activity can be re-

cruited by phosphorylated CTD, but not by unphos-phorylated CTD, from a purified capping enzyme preparation (Cho et al. 1997), and recombinant Ceg1 and methyltransferase (Abd1) directly and independently bind to the phosphorylated CTD (McCracken et al. 1997a). Supporting the functional significance of these interactions, RNAs transcribed by RNAP II with a shortened CTD undergo inefficient capping in transiently transfected mammalian cultured cells (McCracken et al. 1997a). In addition, a viable truncation mutant of the yeast CTD was found to be synthetically lethal in combination with a capping enzyme mutant (Cho et al. 1997). Thus, these phosphorylation-dependent interac-
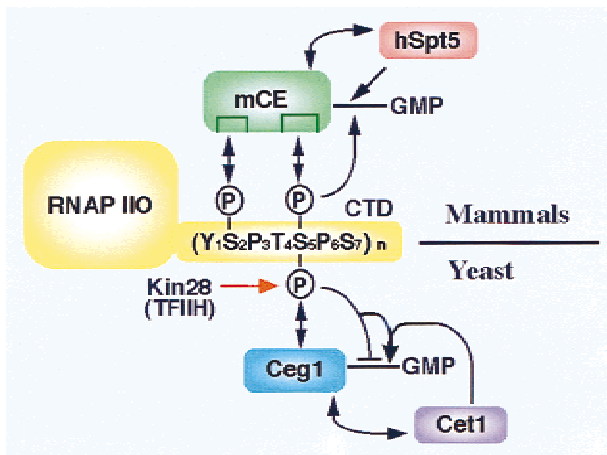
Figure 2. Allosteric activation of guanylyltransferase activity by the phosphorylated CTD. Distinctions in the mechanism by which guanylylation activity of the capping enzyme is enhanced by the CTD in mammals and yeast are shown in the top and bottom of the figure, respectively. CTD repeats phosphorylated on serine 5 stimulate guanylylation in mammals by interacting directly with the capping enzyme (mCE). Guanylylation can also be stimulated by the transcription factor hSPT5. In yeast, interaction between the guanylyltransferase subunit of CE (Ceg1) and phosphorylated CTD is, by itself, inhibitory, but is stimulatory in the presence of the triphosphatase subunit, Cet1. See text for details. Phosphorylation of serine position 5 specifically by the TFIIH component Kin28 (red arrow) enhances interaction between the CTD and Ceg1. Covalent linkage between CE and GMP is indicated by a single bar. Double-headed arrows indicate the physical interactions. (For simplicity, the 7-methyltransferase is not shown.)

tions between the capping apparatus and the CTD are evolutionarily conserved and likely provide a basis for the specific and rapid targeting of the capping enzyme to RNAP II transcripts.

Important extensions of the recruiting model of capping enzyme by the phosphorylated CTD have been made in the last two years. In the first step of the guanylyltransferase reaction, the enzyme itself is guanylylated to form a covalent enzyme–GMP intermediate. Despite the fact that Ceg1 can bind to phosphorylated GST–CTD (McCracken et al. 1997a), guanylylation activity, as measured by formation of the enzyme–GMP intermediate, could not be detected associated with GST–CTD (Cho et al. 1997). These seemingly conflicting findings were explained in an interesting way by Cho et al. (1998), who found first that phosphorylated, but not unphosphorylated, GST–CTD can actually inhibit Ceg1 activity, and second that the inhibition could be reversed and guanylylation activity actually enhanced, by addition of the Cet1 triphosphatase (Fig. 2). These findings and others suggest that Ceg1 activity can be allosterically regulated by interaction with both the Cet1 triphosphatase (see also Ho et al. 1998b) and RNAP IIO. The authors suggest that the observed inhibition of guanylyltransferase by CTD is designed to prevent spurious enzyme activity and to coordinate guanylylation with triphosphatase activity.

Mammalian capping enzyme activity is also allosterically regulated by interaction with the phosphorylated CTD (Ho and Shuman 1999). The isolated carboxy-terminal guanylyltransferase domain of mouse capping enzyme was shown to bind synthetic CTD peptides containing phosphoserine at either position 2 or 5 of the heptad YSPTSPS repeat, but not to unphosphorylated CTD peptides. The CTD phosphopeptides containing phosphoserine at position 5 stimulated formation of the enzyme–GMP intermediate by enhancing the enzyme's affinity for GTP. However, CTD peptides containing Ser-2-phosphorylation had no effect, either on basic enzyme guanylylation or on guanylylation activated by the Ser-5-phosphorylated peptide, indicating that the guanylyltransferase domain of the mammalian enzyme has two independent binding sites for the phosphorylated CTD: one is specific for the Ser-2-phosphorylated peptide and the other is an allosteric activator site recognized by the Ser-5-phosphorylated peptide (Fig. 2; Ho and Shuman 1999). Why in yeast Ceg1 is inhibited by phosphorylated CTD (in the absence of Cet1), whereas the mammalian guanylyltransferase domain is activated is unclear, but it may reflect the fact that transferase and triphosphatase activities are contained in the same polypeptide in mammals but not yeast. Thus the need to coordinate guanylyltransferase and triphosphatase activities does not exist in mammals. Together, these findings have suggested that the phosphorylated CTD functions not only as a simple landing pad for capping enzyme but also as an important regulator of enzyme activity, with activation correlated with position-specific phosphorylation (serine 5) within the CTD heptapeptide.

Genetics studies have revealed that a specific CTD-kinase, Kin28, is likely required for recruiting capping enzyme to RNAP II in yeast (Rodriguez et al. 2000). Three kinases, which have all been implicated in phosphorylation of the CTD in *S. cerevisiae*, were tested for their ability to allow recruitment of capping enzyme to the CTD. These included the Kin28–Ccl1 complex, a component of general transcription initiation factor TFIIH (Svejstrup et al. 1996); the Srb10–Srb11 complex, which is associated with RNAP II holoenzyme (Liao et al. 1995), and CTDK-I (Sterner et al. 1995). Combinations of mutant alleles of the genes encoding these kinases were tested with a *ceg1* temperature-sensitive (ts) mutant, which previously had been shown to exhibit synthetic lethality with a viable CTD truncation mutant (Cho et al. 1997). Although all of the kinases were able to phosphorylate GST–CTD to allow recruitment of capping enzyme in vitro, only *kin28* mutant alleles exhibited a genetic interaction with the *ceg1* mutant. The level of CTD phosphorylation and, intriguingly, Ceg1 protein levels were reduced in both the CTD truncation mutant and *kin28* mutants, raising the possibility that Ceg1 associated with CTD phosphorylated by Kin28 may be stabilized relative to unbound Ceg1. Furthermore, conditional mutants in which serine 5, but not serine 2, residues were replaced with alanines in either the first or second half of the CTD were synthetically lethal in combination with a *ceg1* mutant. These data

are in good agreement with the CTD phosphorylation requirements of mammalian capping enzyme (Ho and Shuman 1999), and strongly suggest that TFIIH-associated Kin28 phosphorylates serine 5 of the CTD repeat, at least in part, to target and activate the capping apparatus.

An interesting extension to these findings was suggested by the discovery that a protein implicated in transcriptional elongation, hSPT5, interacts physically and functionally with human capping enzyme (Fig. 2; Wen and Shatkin 1999). SPT5 was initially uncovered genetically in *S. cerevisiae* (Swanson and Winston 1992), and functions together with SPT4 to modulate an early step in transcription elongation in both yeast and humans (Hartzog et al. 1998; Wada et al. 1998a). hSPT5 was isolated in a yeast two-hybrid screen with human capping enzyme as bait, and the two proteins were shown to interact directly in vitro. hSPT5 strongly stimulates the guanylyltransferase but not the RNA triphosphatase activity. Intriguingly, no stimulation of capping was detected when hSPT5 was added together with a phosphorylated GST–CTD protein, raising the possibility that the two capping activators function redundantly. Given that Spt5 interacts preferentially with RNAP IIA (Wada et al. 1998b), two models can be suggested to explain the role of hSPT5 in capping. In one, suggested by the authors, SPT4/5 dissociates from the CTD upon phosphorylation, but remains associated with the transcription complex and functions to enhance capping upon recruitment of capping enzyme to the phosphorylated CTD. In a second model, hSPT5 could function to recruit capping enzyme to the holoenzyme in some cases prior to, or independent of, CTD phosphorylation. Capping enzyme could then be transferred to the CTD upon phosphorylation and transcription, or be activated by hSPT5 to ensure rapid and efficient capping of transcripts that may be initiated and elongated prior to, or in the absence of, CTD phosphorylation. RNAP IIA has in fact been implicated in the elongation of a small number of genes (Weeks et al. 1993).

Splicing and transcription

Splicing of mRNA precursors takes place in a large macromolecular 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 (for reviews, see Moore et al. 1993; Kramer 1996; Manley and Tacke 1996). Although cytological studies have suggested that splicing can occur cotranscriptionally (see, for example, Beyer and Osheim 1988; Bauren and Wieslander 1994), and factors required for splicing can be found localized at sites of active transcription (Zhang et al. 1994), functional coupling between transcription and splicing is not obligatory because splicing can be reconstituted in vitro with pretranscribed RNA and splicing-competent cell extracts. Both biochemical and in vivo studies have provided support for the existence of functional interactions between RNAP II, especially the CTD, and the splicing apparatus (see Figs. 1 and 3). RNAP IIO, but not RNAP
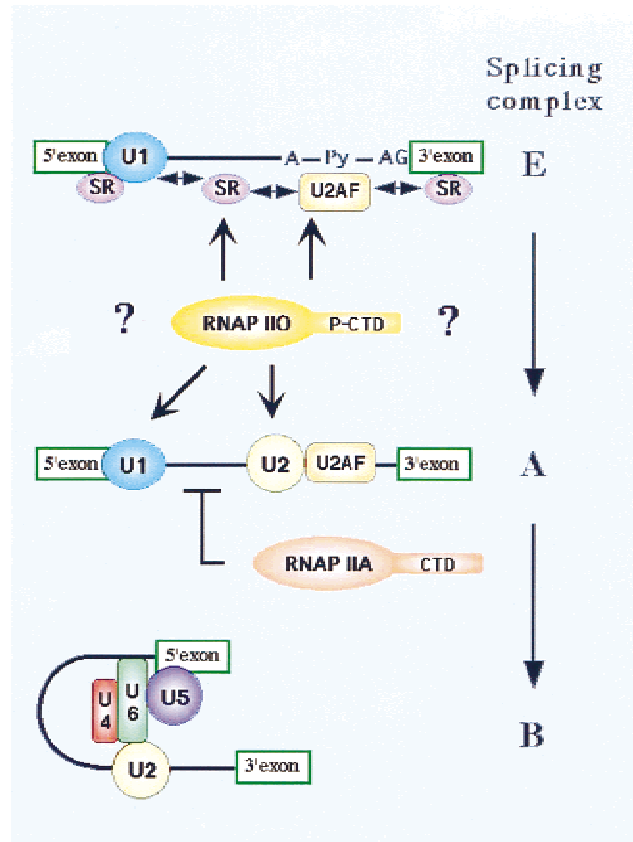

Figure 3. Effects of RNAP II on mammalian pre-mRNA splicing. Possible interactions between RNAP IIO and the splicing machinery during splicing complex formation are indicated by arrows. However, exactly how RNAP II0 stimulates splicing is unknown, as reflected by the question marks. Splicing complexes formed on the pre-mRNA splicing substrate in a stepwise manner (from E to B) are indicated on the right. The five snRNPs (U1, U2, U4/U6, and U5) are represented by ovals. U2AF indicates U2 snRNP auxiliary factor, which binds to the polypyrimidine tract (Py) near the 3′ splice site (AG). SR proteins, which have multiple functions in spliceosome assembly, are shown. Thick solid line indicates the intron and "A" in the intron indicates the branch point adenosine. The arrows show stimulation or stabilization of complex assembly by RNAP IIO, while inhibition or disruption of a complex A by RNAP IIA is indicated by the line with crossbar. See text for details.

IIA, has been found to associate with splicing factors, and this isoform has also been detected in active spliceosomes (Chabot et al. 1995; Mortillaro et al. 1996; Yuryev et al. 1996; Kim et al. 1997). Antibodies directed against the CTD and CTD peptides have been shown to inhibit splicing in vitro (Chabot et al. 1995; Yuryev et al. 1996). Like capping, splicing (and polyadenylation) of RNAs transcribed in transient transfection assays by CTD-truncated RNAP II was inefficient (McCracken et al. 1997b), and overexpression of phosphorylated CTD peptides was shown to inhibit splicing in cultured mammalian cells (Du and Warren 1997). These observations provided the initial evidence that the hyperphosphorylated CTD of elongating RNAP II may function in splicing,

perhaps serving as a platform upon which processing factors bind, thus helping to promote efficient and accurate splicing by targeting necessary factors to transcription sites.

Microscopic observations in mammalian cells have provided visual support for this targeting model (Misteli and Spector 1999). Following activation of a reporter gene in cells expressing either full-length or CTD-truncated RNAP II as the only source of active enzyme, sites of accumulation of both the newly synthesized reporter transcripts and splicing factors were simultaneously visualized by immunohistochemistry techniques. Although both sites colocalized well in cells expressing wild-type RNAP II, the transcription sites did not colocalize above random levels with either SR proteins or snRNP particles in cells expressing the CTD-truncated RNAP II. Estimation of splicing levels by in situ hybridization with short exon-spanning probes suggested that truncation of the CTD prevented accumulation of spliced products despite the presence of significant amounts of unspliced pre-mRNAs. These results support the idea that the CTD is required for targeting splicing factors to transcription sites and that this can be important for efficient splicing. The same authors also showed by immunoprecipitation experiments using transiently transfected cells expressing wild-type or mutant SR proteins that the RS domain of several SR proteins was necessary and sufficient for associating with the phosphorylated RNAP II largest subunit, although it is not clear whether these associations were direct or indirect.

The above findings are consistent with the view that the CTD is required for targeting splicing factors to transcription sites to ensure efficient splicing. However, is the efficient splicing in vivo due only to the increased local concentration of processing factors? Or does the CTD participate more directly in the actual splicing reaction? Like in other processing reactions such as capping or polyadenylation (see below), recent experiments have shown that RNAP II also plays a direct and active role in splicing in vitro in the absence of transcription (Fig. 3; Hirose et al. 1999). Purified RNAP IIO was found to strongly activate the splicing of several different pre-mRNAs in reconstituted splicing assays. RNAP IIO significantly increased formation of spliceosomal complexes, and the pre-spliceosomal A complex was notably increased at very early times of the reaction. These results indicate that RNAP IIO stimulates splicing by accelerating the rate of one of the first steps in spliceosomal assembly, probably by facilitating in some way binding of U1 and/or U2 snRNP particles to the pre-mRNA 5′ splice site and/or branch site, respectively. RNAP IIA, on the other hand, was capable of inhibiting splicing of some but not all of the pre-mRNA substrates tested, apparently by disrupting early pre-splicing complexes. The CTD was necessary for these effects on splicing because the CTD-less form of RNAP II (IIB) was without significant effect, but, unlike with capping and polyadenylation (see below), was not sufficient. These results provide additional support for the idea that RNAP II not only controls spatial distribution of pre-

mRNA processing factors in the cell nucleus to couple transcription to processing, but also can function directly with splicing factors to enhance the efficiency of splicing. Furthermore, this study also suggested that differential CTD phosphorylation has the potential to play a significant role in splicing regulation.

The concept of a physical and functional coupling between transcription and splicing has been extended by the finding that promoter structure can contribute to splice site selection (Cramer et al. 1997, 1999). One of the exons in the *fibronectin* (*FN*) gene, called EDI, has long been known to be subject to alternative splicing (inclusion or exclusion). Inclusion of this exon depends on the presence of specific *cis*-acting sequences located within EDI that function by binding specific SR proteins (Lavigueur et al. 1993). Various reporter constructs containing EDI under the control of different promoters were transiently transfected into human cultured cells and the extent of EDI inclusion was measured by RT–PCR and Northern blot analysis. The authors found that varying the promoter resulted in significant changes in the ratio of EDI inclusion versus exclusion (Cramer et al. 1997). Important controls showed that the effects were not due to differences in promoter strength or to differences in the site of transcription initiation, but rather to the identity of the promoter (Cramer et al. 1999). The authors further demonstrated that overexpression of specific SR proteins markedly stimulated EDI inclusion, but the effects of these SR proteins depended on the promoter from which the transcript was produced. These results have suggested that the transcription machinery can affect the recruitment of specific SR proteins to exonic *cis*-acting elements on the newly transcribed RNA.

How might promoter structure modulate alternative splicing? One possible explanation is that the precise nature of the initiation complex assembled on a particular promoter may affect the extent of phosphorylation of the RNAP II-CTD, and this in turn modulates the recruitment by the CTD of specific SR proteins to *cis*-acting regulatory sequences in the nascent RNA. Alternatively, different degrees of phosphorylation could affect the ability of the CTD to participate directly in early spliceosome assembly, as discussed above. Another possible mechanism, not related to CTD phosphorylation, involves recruitment of promoter-specific factors that may in turn function with the CTD to influence splicing (see below).

Splicing-transcription connectors

The above discussion suggests that RNAP II, via the CTD, can play a significant role in stimulating splicing. But it is not clear precisely how it does so, or even what proteins directly interact with the enzyme to influence splicing. As mentioned above, SR proteins and snRNPs have been found to associate with RNAP II0 or phosphorylated GST–CTD, but it is unclear if these interactions are direct. Thus it is worth considering what molecules might serve to connect RNAP II0 with the splicing machinery. Although nothing definitive is known

Hirose and Manley

regarding such hypothetical protein linkers, several proteins, or protein families, merit some discussion (see Fig. 1).

A group of at least four proteins now known as SCAFs (SR-like CTD-associated factors) are reasonable candidates to function in coupling transcription and splicing. These proteins were initially discovered in a yeast two-hybrid screen for CTD-interacting proteins (Yuryev et al. 1996), and features of their primary structure suggest they could be involved in linking RNAP II to splicing. These include an RS-like domain and an RNA binding domain (RBD), which are similar to domains found in SR proteins. It is intriguing that RS domain-containing proteins were found to interact with the CTD, as Greenleaf (1993) had earlier predicted an RS domain–CTD interaction that would link transcription and RNA processing. This was certainly a prescient suggestion, despite the fact that the details seem to be incorrect: SCAFs interact with the CTD via a distinct domain (the CTD interaction domain; CID), not their RS-like region. A subsequent study showed that one SCAF, SCAF8, interacts only with phosphorylated CTD, or RNAP II0, and can be localized in cells to foci that overlap with sites of transcription and processing (Patturajan et al. 1998). These results together all support a role for SCAFs in linking RNAP II transcription to processing, but direct support for this hypothesis is not yet available.

Yeast contains a protein, called Nrd1, with similarity to SCAFs, including a CID, an RS-like region (which is uncommon in yeast) and an RBD (Steinmetz and Brow 1996, 1998). Nrd1 is an essential protein that appears to function in the nucleus and to influence accumulation of certain pre-mRNAs, perhaps at the level of transcription elongation or 3′-end formation. A recent genetic analysis has provided support for the idea that Nrd1 may serve to link transcription with some aspect of pre-mRNA processing (Conrad et al. 2000). A *nrd1* ts allele was shown to be synthetically lethal with viable CTD truncation mutants, and genetic interactions were also observed with two genes encoding hnRNP-like proteins. However, it remains unclear exactly how Nrd1 functions, and there is no evidence that it is involved in splicing.

An interesting question is whether the link between transcription and splicing in fact extends to yeast. All the data supporting this interaction comes from studies in mammalian systems, and there are several factors suggesting yeast may not require such a coupling. For example, yeast genes are relatively simple and the vast majority lack intervening sequences, which when present are usually small and located near the 5′ end of the transcript. Perhaps reflecting a role limited to transcription (and capping), the CTD is half the size of its mammalian counterpart. However, it is noteworthy that although only ~3% of yeast genes contain intervening sequences, ~30% of the total primary transcript pool contains introns (Ares et al. 1999; Lopez and Seraphin 1999). This is because intervening sequences are overrepresented in highly expressed genes, such as those encoding ribosomal proteins. Although the reasons for this are unclear, an intriguing possibility is that introns facilitate

efficient mRNA synthesis, perhaps suggesting some sort of splicing–transcription link after all.

A second intriguing group of potential transcription-splicing coupling proteins are the TLS/FUS-related proteins, which also includes EWS and TAF$_{II}$68. The genes encoding all three proteins are known to be involved in chromosomal translocations that are associated with different sarcomas and leukemias (Attwooll et al. 1999; Sjogren et al. 1999 and references therein). In all cases, chimeric proteins consisting of the amino terminus of the TLS/FUS-related protein fused to the DNA binding domain of a transcriptional activator are created (Delattre et al. 1992; Crozat et al. 1993). While it seems likely that oncogenic transformation results at least in part from the altered transcriptional properties of the chimeric transcription factors, properties of the TLS/FUS proteins themselves are consistent with roles in both transcription and splicing. The primary structure of the proteins indicates that they all contain an RBD in their central region and a so-called RGG domain at their carboxyl terminus, both of which are features of hnRNP proteins (for review, see Burd and Dreyfuss 1994). The TLS/FUS and EWS proteins have indeed been isolated in complexes containing RNA and hnRNP A and C proteins (Zinszner et al. 1994). Furthermore, yeast two-hybrid screens found that the amino terminus of a TLS/FUS-related protein (the region retained in the chimeric oncoproteins) interacts with splicing factor SF1 (Zhang et al. 1998) while the carboxy-terminal RGG domain interacts with SR proteins (Yang et al. 1998). Further supporting a possible role in splicing, transient overexpression of TLS/FUS can affect the relative accumulation of alternatively spliced mRNAs from a cotransfected reporter (Hallier et al. 1998; Yang et al. 1998). On the other hand, TAF$_{II}$68 was isolated, as the name suggests, as a TATA binding protein (TBP)-associated factor; that is, as a component of transcription factor TFIID (Bertolotti et al. 1996). Although not present in stoichiometric amounts, both TAF$_{II}$68 and TLS/FUS could be found in distinct TFIID complexes, and at least TAF$_{II}$68 also copurifies extensively with RNAP II. Together, all these properties are consistent with the idea that TLS/FUS-related proteins function in the coupling of transcription and splicing. However, as with the SCAFs, direct support for this idea is lacking, and indeed biochemical evidence suggesting an alternative role for TLS/FUS in homologous recombination has been presented (Baechtold et al. 1999).

The human papillomavirus (HPV) E2 protein provides an example of a sequence-specific DNA binding protein that may also serve to link transcription and splicing (Lai et al. 1999). The E2 protein is well known to participate in control of viral transcription and replication, and contains well-conserved DNA binding and activation domains separated by a hinge region that is less well conserved between HPV subtypes. In certain types, though, the hinge consists of multiple RS dipeptide repeats, similar to SR proteins. One of these proteins, HPV-5 E2, interacts, via this RS domain-like region, with splicing factors, including SR proteins, and can colocalize in trans-

fected cells with splicing factors. Importantly, full activation of reporter genes was found to require the RS-rich hinge region, and this hinge region-dependent activation reflected enhanced splicing of the reporter transcripts (Lai et al. 1999). HPV-5 E2 thus provides an example of a promoter-bound transcription factor that can stimulate both transcription and splicing.

A final possible linking protein we will discuss is known as p52. p75, a variant of p52 arising from alternative splicing, was initially discovered by copurification with the transcriptional coactivator PC4, and both proteins were themselves shown to be capable of functioning as coactivators in in vitro transcription assays (Ge et al. 1998a). Coactivator function was suggested to reflect the ability of the proteins to interact with both transcriptional activators and with general transcription factors. Remarkably, p52 was also shown to interact with the SR protein ASF/SF2 in vitro and in vivo, and was suggested to influence alternative splicing of a model pre-mRNA in in vitro splicing assays (Ge et al. 1998b). Based on these properties, it is conceivable that p52 not only helps link RNAP II transcription to splicing, but also might contribute to the promoter-specific effects on alternative splicing described above.

Polyadenylation and the CTD

Polyadenylation of mRNA takes place in two steps: endonucleolytic cleavage of the mRNA precursor followed by poly(A) addition to the 3′ end of the upstream cleavage product. mRNA polyadenylation requires multiple protein factors, including, in mammalian cells cleavage/polyadenylation specificity factor (CPSF), cleavage stimulation factor (CstF), two cleavage factors, CFI and CFII, and poly(A) polymerase (for review, see Colgan and Manley 1997; Zhao et al. 1999a). Early studies showing

that transcription termination by RNAP II was dependent on an intact poly(A) signal (for review, see Proudfoot 1989) predicted a possible link between the transcription process and polyadenylation, and there is now solid evidence that these two processes are indeed intimately coupled (see Fig. 4).

As with capping and splicing, RNAs transcribed by CTD-truncated RNAP II were not efficiently polyadenylated in transiently transfected cells (McCracken et al. 1997b). It was also shown that CPSF and CstF present in unfractionated nuclear extracts could bind GST–CTD, and both were present in an RNAP II holoenzyme preparation. It appears that the 50 kD subunit of the heterotrimeric CstF (CstF-50) interacts directly with the CTD, based on binding experiments with GST–CTD and in vitro translated proteins. Unlike capping and splicing, the phosphorylation status of the CTD was found not to affect binding.

An initial interpretation of these results was that the CTD functions to help recruit polyadenylation factors to sites of RNAP II transcription, increasing their local concentration and thereby facilitating efficient processing. But how and when do poly(A) factors associate with the polymerase? The initially unexpected (but perhaps not so surprising in light of the capping and splicing connections described above) answer is that at least some of the action occurs at the promoter. While studying the general transcription factor TFIID, Dantonel et al. (1997) found that an extensively purified preparation contained in good yield at least three of the four subunits of CPSF. Furthermore, in a reconstituted transcription assay, CPSF was shown to transfer from TFIID to RNAP II concomitant with initiation. Together with the results of McCracken et al. (1997b), these findings suggest that at least some factors associate early with RNAP II and remain associated with it during elongation.
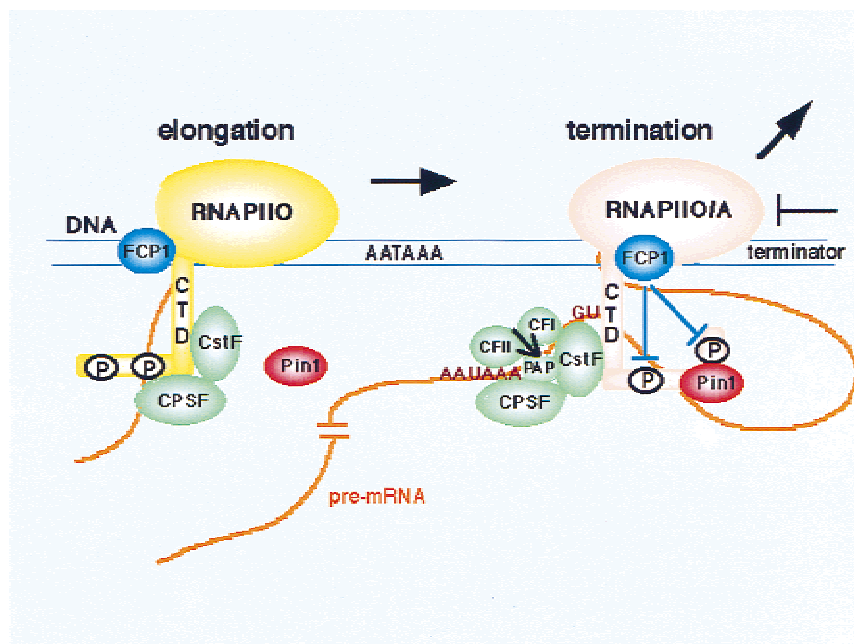

Figure 4. Coupling polyadenylation and transcription. After RNAP IIO passes the poly(A) site (AATAAA), polyadenylation factors (green ovals) and the CTD form a functional complex on the pre-mRNA to catalyze endonucleolytic cleavage (indicated by a purple arrow), possibly with the help of Pin1. In many cases, there may be lag in processing until RNAP II encounters a pause site (line with crossbar), which facilitates 3′ cleavage and in turn signals the polymerase to terminate and the complex to dissociate. Given that the CTD is required for the cleavage reaction, the transcribed pre-mRNA may form a large loop (see text for details). FCP1, the apparent CTD phosphatase (Cho et al. 1999), dephosphorylates the CTD to allow reinitiation and the next round of transcription. The timing of CTD dephosphorylation, and whether this contributes to polyadenylation/termination, is not known.

The above findings were extended by a biochemical study demonstrating a more direct role of RNAP II in polyadenylation (Hirose and Manley 1998). These authors provided evidence that RNAP II, and specifically the CTD, is required for the cleavage reaction in vitro in the absence of transcription. Purified RNAP IIA and II0 were both found to activate the first step of polyadenylation, 3′ cleavage, in a reconstituted system containing all the other polyadenylation factors. In addition, both unphosphorylated and hyperphosphorylated GST–CTD proteins activated cleavage just as efficiently as did RNAP II, although in this case the hyperphosphorylated CTD was more active than the nonphosphorylated form. In addition, 3′ cleavage in nuclear extracts could be inhibited by immunodepletion of RNAP II and rescued by add-back of the purified enzyme. These results suggest that the CTD of RNAP II participates directly in the formation and/or function of a stable, catalytically active processing complex through direct interaction with polyadenylation factor(s).

The factors required for mRNA polyadenylation have been intensively studied for over a decade. How had the involvement of RNAP II in 3′ processing escaped attention? Hirose and Manley (1998) suggest that it was masked by an apparent requirement for the small molecule creatine phosphate (CP) in the cleavage reaction. CP is usually employed as part of an ATP regeneration system for ATP requiring reactions and has been routinely added to 3′-processing reactions, where conflicting early experiments had suggested a possible ATP requirement. But Hirose and Manley (1997) showed that CP plays a different role in 3′ cleavage: ATP was found in fact not to be required for cleavage (at least in mammals), but CP seemed to be, although hydrolysis of its high-energy bond does not occur. These and other results led to the hypothesis that CP functions by mimicking a phosphoprotein, and this in turn resulted in the discovery that RNAP II0 can activate cleavage in the absence of CP (Hirose and Manley 1998). If indeed CP and the CTD function similarly, this has significant implications regarding mechanism. It would disfavor a model in which the CTD functions to facilitate complex assembly by making contacts with two or more factors, and favor an allosteric activation model. Or perhaps it does both, which seems to be the case in capping (see above). But the CP–CTD connection is not perfect. For example, the concentration of CP required to support optimal cleavage is ~10^7 fold greater than the required CTD concentration, and the CP–CTD model does not offer an explanation for why unphosphorylated CTD can activate cleavage.

The direct involvement of RNAP II in 3′ cleavage could indicate that the reaction must occur very shortly after the poly(A) site is transcribed, or that RNAP II might play an essential but transient role, perhaps facilitating formation of a conformation necessary for processing. Some support for the former hypothesis comes from studies examining cotranscriptional polyadenylation using a coupled in vitro transcription–polyadenylation system (Yonaha and Proudfoot 1999). Specific G-rich sequences originally identified as protein binding sites and proposed to contribute to transcription termination in vivo were found to promote RNAP II pausing in vitro and, importantly, to significantly enhance polyadenylation at an upstream poly(A) site in a transcription-dependent manner. A plausible explanation for these findings is that pausing the polymerase near the site of processing allows more time for the CTD to function in cleavage before RNAP II moves downstream or terminates (see Fig. 4).

Another recent study extended further the potential significance of the association between polyadenylation factors and RNAP II into the realm of DNA repair and tumor suppressors (Kleiman and Manley 1999). These authors discovered in a yeast two-hybrid screen an interaction between Cst-50 (the CTD-interacting subunit of CstF; see above) and BRCA1-associated ring domain protein (BARD1). BRCA1 is a breast/ovarian cancer-associated tumor suppressor protein of unknown function, although it has been suggested to be associated with the RNAP II holoenzyme (Scully et al. 1997a) and to participate in the response to DNA damage (Scully et al. 1997b). BARD1 is tightly associated with BRCA1 in vivo (Wu et al. 1996) and likely contributes to its function. CstF-50 and BARD1 were found to associate in vitro and a fraction of intact CstF is associated with the BARD1/BRCA1 complex in vivo. In functional assays, recombinant BARD1 was found to inhibit the 3′-cleavage reaction in vitro. Given the apparent association of BARD1/BRCA1 and CstF with elongating RNAP II, an intriguing model is that the function of the BARD1 inhibitory interaction is to prevent premature polyadenylation of nascent RNAs at sites of RNAP II pausing, for example, at sites of DNA damage.

The intimate coupling between transcription and polyadenylation mediated by the CTD is likely a general feature in mammalian cells, as we have just seen. Does it also occur in yeast cells? When the *S. cerevisiae Cup1* gene was transcribed by a CTD-deleted RNAP II, 3′ processing of the Cup1 transcripts became inefficient (McNeil et al. 1998). On the other hand, yeast Cyc1 and Yhr54C transcripts made by CTD-deleted RNAP II (McNeil et al. 1998), as well as His4 transcript made by RNAP I (Lo et al. 1998), were efficiently polyadenylated in vivo. These observations indicate that the CTD-mediated coupling between transcription and polyadenylation might be a gene specific feature in yeast. Nonetheless, additional studies have indicated that the CTD could mediate a physical and mechanistic link between transcription and polyadenylation in yeast. First, the yeast polyadenylation factor Pta1, which may function as an assembly factor (Zhao et al. 1999b), as has been suggested for a related human protein, symplekin (Takagaki and Manley 2000), could be selected from unfractionated yeast extract by phosphorylated GST–CTD (Rodriguez et al. 2000). These authors also showed that *kin28* mutant alleles defective for CTD phosphorylation resulted in a reduction in Pta1 levels. A second intriguing finding originated from a genetic screen for new transactivating factors involved in mRNA 3′-end forma-

tion. The screen uncovered mutations in an essential gene, *Ess1*, which encodes a peptidylprolyl-*cis-trans*-isomerase (PPIase), that led to a defect in 3′-end formation of a plasmid-derived pre-mRNA (Hani et al. 1999). Importantly, Ess1 has recently been found to associate specifically and directly with the hyperphosphorylated CTD (Morris et al. 1999). We will discuss Ess1 and other PPIases in the next section in relation to their possible functions in remodeling CTD-associated protein complexes. But together these results support the existence of an RNAP II CTD–polyadenylation link in yeast as well as mammals.

Remodeling the CTD

The RNA polymerase II CTD appears to interact with a number of multiprotein complexes involved in both transcription and pre-mRNA processing to produce the mature mRNA (see Fig. 1). The assembly and disassembly of processing complexes on the CTD likely occurs in a highly dynamic and temporally and spatially regulated manner during the transcription cycle. An important question then is how such transitions might be coordinated. In this regard, two newly discovered proteins, each of which possesses a different class of PPIase domain, are intriguing. PPIases catalyze *cis–trans* isomerization of the peptide bond on the amino-terminal side of proline residue in peptides and proteins. PPIases are classified into three distinct families. The cyclophilins and the FK506-binding proteins (FKBPs) families are sensitive to the immunosuppressant drugs cyclosporin A (CsA) and FK506, respectively. A third group is the parvulin family, which is not inhibited by those drugs. PPIases are thought to function in protein folding, trafficking, assembly/disassembly, and direct regulation of protein function (for review, see Hunter 1998; Gothel and Marahiel 1999).

Human SRcyp/SCAF10 was originally identified in a yeast two-hybrid screen as a CTD-interacting protein (Bourquin et al. 1997). SRcyp has a cyclophilin-like PPIase domain in its amino terminus and an RS-rich region similar to those found in SR proteins in its carboxy-terminal half, which also contains a CTD-interacting domain (CID; see above). Overall, this domain organization is related to that of the SCAFs, hence the name SCAF10. The rat homolog of the SRcyp, matrin CYP, was shown to possess PPIase activity in vitro (Mortillaro and Berezney 1998). SRcyp and matrin CYP were shown to co-localize in cells with the SR protein splicing factor SC35 and the snRNP protein U1-70K, respectively. Based on these findings, it was suggested that the CTD, via its proline-containing heptad repeats, may be induced to undergo conformational changes by the PPIase activity of SCAF10, and that this may help facilitate the assembly or disassembly of splicing factors on the CTD. Although these properties of SCAF10 are highly suggestive of an interesting role for SCAF10 in linking the RNAPII CTD to splicing, functional support for this idea is not yet available.

Another interesting nuclear PPIase is Ess1/Pin1,

which is an evolutionarily conserved member of the parvulin family. *Ess1* was initially discovered in yeast and mutations gave rise to defects in cell division (Hanes et al. 1989). The human counterpart of Ess1, Pin1, was discovered in a yeast two-hybrid screen with a cell-cycle kinase, NIMA, and depletion (by antisense) and overexpression in human cells both resulted in cell-cycle defects in G_2/M phase (Lu et al. 1996). In keeping with a possible direct role in cell cycle control, Pin1 has been shown to bind in vitro to a number of mitotic regulators (Shen et al. 1998). Ess1/Pin1 possesses two distinct domains: an amino-terminal WW domain, which is involved in protein–protein interactions, and a PPIase catalytic domain. A remarkable feature of the protein is its unique substrate specificity. Ess1/Pin1 dramatically enhances (1000–3000 fold) isomerization of prolines preceded by phosphorylated serine (pSer) or threonine (pThr) compared with a proline preceded by a nonphosphorylated residue (Yaffe et al. 1997; Hani et al. 1999). The WW domain of Pin1 has been shown to be responsible for interacting with pSer/pThr-Pro motifs in target proteins (Lu et al. 1999). Ser/Thr-Pro is the core of the target sequence recognized by cyclin-dependent kinases, which fits nicely with the possibility that Ess/Pin1 functions by binding to and altering the conformation of mitotic regulators phosphorylated on Ser/Thr-Pro.

But compelling evidence also suggests that Ess1/Pin1 functions in mRNA 3′-end formation by linking the processing reaction to transcription (see Fig. 4). As mentioned above, *Ess1* was uncovered in a genetic screen for ts mutations affecting proteins that function in mRNA 3′-end formation. The screen was designed to be very specific for 3′-end formation, and indeed mRNAs with unprocessed 3′ ends could be detected at the nonpermissive temperature (Hani et al. 1995, 1999). Pointing to a link with the CTD, a biochemical selection employing phosphorylated GST–CTD and a yeast extract identified Ess1 as the major interacting protein (Morris et al. 1999). The interaction is specific for the phosphorylated form of the CTD, and RNAP II0 appears to be the major Ess1/Pin1 interacting protein, in humans (Albert et al. 1999) as well as yeast, likely reflecting the abundance of the pSer-Pro dipeptides in the phosphorylated CTD. Especially given that the genetic screen specifically identified mutants defect in PPIase activity, an attractive model is that Ess1/Pin1 effects a conformational change in the phosphorylated CTD that leads to enhanced efficiency of 3′-end formation. It is also possible that changes in the CTD and/or complexed proteins affect subsequent transcription termination (see below).

The above discussion suggests that Ess1/Pin1 may well be a bifunctional protein, participating directly in cell cycle control and mRNA 3′-end formation. However, it is also conceivable that at least some of the cell cycle arrest phenotype could be indirect, reflecting changes in gene expression as a result of defects in 3′-end formation. For example, mutations in certain splicing factors in fission yeast have been shown to display cell cycle phenotypes (Potashkin et al. 1998), and genetic depletion of the polyadenylation factor CstF-64 in

chicken DT40 cells causes cell-cycle arrest (Takagaki and Manley 1998).

Transcription termination and RNA processing: the final link

As we mentioned at the beginning of this review, evidence has existed for over a decade suggesting that a functional polyadenylation signal is required for RNA polymerase II to terminate transcription (for review, see Proudfoot 1989). These experiments, which employed transient transfection assays with plasmids containing mutated poly(A) signals and nuclear run-on analysis of transcription, led to two models: In one, the polymerase is modified in some way as it passes the polyadenylation signal, causing it to become less processive and more likely to terminate. In the other, the act of 3' cleavage is directly signaled to elongating RNAP II, by the action of a 5'-to-3' exonuclease that rapidly degrades the downstream product of the cleavage reaction, and this then causes the polymerase to become termination prone. Although it still remains unclear which, if either, of these models are correct, a number of different approaches have strengthened the polyadenylation–termination link, and at least placed some limitations on the possible mechanism.

Given that mutations in *cis*-acting 3' processing signals prevent termination, an interesting question has been whether mutations in *trans*-acting polyadenylation factors would have similar effects. This was studied by Birse et al. (1998), who examined termination (by nuclear run-on) in yeast cells harboring mutations in different biochemically characterized polyadenylation factors. Intriguingly, cells with mutations in factors implicated in the cleavage step were defective in termination at the nonpermissive temperature, but those in which factors thought to be involved only in the second step were affected (e.g., poly(A) polymerase, or PAP) showed no termination defects. These findings provide both genetic evidence linking 3'-end formation and transcription termination and also support for the idea that RNA cleavage is required, consistent with the second model above. However, the demonstration in mammalian systems that certain factors can associate with RNAP II (although not PAP; McCracken et al. 1997b) weakens somewhat this second conclusion, as it's possible that the mutations alter in some way these interactions so that required changes in the RNAP II complex, and hence termination, do not occur.

Two additional studies, one using modified run-on assays with transfected mammalian cells (Dye and Proudfoot 1999), the other analyzing by PCR nascent transcripts from the heavily transcribed Balbani ring 1 gene isolated from salivary glands of the diptern *Chironomous tentans* (Bauren et al. 1998), provided additional, and surprisingly similar, insights into termination. Consistent with previous work, both studies provided evidence for termination (i.e., the lack of run-on or PCR signals) ~1000 bp downstream of the polyadenylation signal. However, in each case the large majority of nascent RNAs detected right up to the apparent termina-

tion site were uncleaved. This suggests that perhaps 3' cleavage is not essential for termination. However, another possibility is that processing and termination are actually temporally coupled. Perhaps there is a lag in processing until RNAP II encounters a pause site, which activates 3' cleavage, and this in turn signals the polymerase to terminate and the complex to dissociate (Fig. 4). These results are consistent with the requirement of RNAP II and specifically the CTD in 3' processing (McCracken et al. 1997b; Hirose and Manley 1998) and with the demonstration that a pause site can enhance polyadenylation in a coupled processing/transcription reaction (Yonaha and Proudfoot 1999). Intriguingly, both of these studies also implicated processing, or at least recognition, of the 3' most intron in subsequent termination. This likely reflects the enhancement of polyadenylation by splicing of the upstream intron (Niwa et al. 1990), which in turn activates termination.

The above studies all analyzed termination somewhat indirectly, by assaying for the presence or absence of nascent RNAs. It is thus reassuring that a study examining the terminating polymerases more directly reached very similar conclusions (Osheim et al. 1999). These authors visualized by electron microscopy (EM) transcribing polymerases on templates isolated from microinjected *Xenopus laevis* oocytes. As expected, they observed termination occurring downstream of polyadenylation signals. Significantly, a functional poly(A) signal was required for termination, and the strength of the signal was directly proportional to the efficiency of termination. Consistent with the experiments just described, a majority of the transcripts approaching the termination site were uncleaved, consistent with either a lack of a cleavage requirement for termination or a temporal linkage of the two processes. Interestingly, the nascent transcripts observed on different plasmids from the same oocyte showed differences in processing and termination efficiency, but all the RNAs on a single template behaved similarly. One explanation for this is that events at the promoter can dictate subsequent 3' processing efficiencies, which is consistent with the observations, discussed above, that CPSF can be recruited to promoters and transferred to elongating RNAP II (Dantonel et al. 1997). In any event, these studies together significantly strengthen the view that polyadenylation is functionally linked to subsequent transcription termination, but the exact mechanism remains elusive.

Cell biology: the emergence of the transcriptosome

The use of increasingly sophisticated methods for visualizing subcellular structures and for localizing individual proteins within them, has given rise to a picture of nuclear interactions remarkably similar to that suggested by the largely biochemical experiments described above. Considerable attention has focused on nuclear structures called speckles, which appear to correspond to interchromatin granule clusters that had initially been observed with the electron microscope (for review, see Spector 1993). Speckles, of which there are 20–40 in a

typical mammalian nucleus, have been visualized by immunofluorescent staining with antibodies against splicing factors, frequently SR proteins. It now seems clear that speckles represent storage sites, or perhaps sites of assembly or recycling of splicing complexes, rather than sites of active splicing (for review, see Singer and Green 1997; Misteli and Spector 1998). Both hyperphosphorylated RNAP II0 (Bregman et al. 1995) and polyadenylation factors (Schul et al. 1998) can also be observed associated with speckles, although apparently at the periphery. Transcriptional activation seems to result in a redistribution of factors from the speckles, and indeed RNAP II0 (Bregman et al. 1995) and the SR protein ASF/SF2 (Misteli et al. 1997) appear to migrate from speckles to sites of transcription. In cells expressing an RNAP II with a truncated CTD, however, relocalization of splicing factors to transcription sites does not occur (Misteli and Spector 1999), in keeping with the biochemistry describing interactions between the CTD and splicing factors. Thus the localization and dynamics of transcription and processing factors within the nucleus is consistent with the functional interactions observed in vitro.

Combining biochemistry with cell biology, Mintz et al. (1999) recently described a purification protocol for speckles. Although it was not possible to quantitate purification, EM visualization was consistent with a significant enrichment. The preparation contained RNAP II0 and was significantly enriched in splicing factors, especially SR proteins. A number of previously unknown proteins were also identified, and one, of 137 kD, seemed especially interesting. The protein is the apparent mammalian homolog of a recently described yeast splicing factor, Rse1 (Caspary et al. 1999), but shows similarity across its entire length to a polyadenylation factor (CPSF-160) and to a protein implicated in DNA repair (UV-DDB). Although the significance of these similarities is currently unknown, all three proteins are conserved from yeast to human and suggest potentially interesting relationships between these processes.

The above studies have provided evidence that RNAP II0 can colocalize with certain RNA processing factors, supporting the view that transcription and processing are indeed linked physically and functionally in the nuclei of mammalian cells. But this idea has recently been significantly extended by studies employing *X. laevis* oocytes. Gall and colleagues have for many years analyzed nuclear organization by studying the oocyte germinal vesicle, taking advantage of the large size of the organelle. In a recent study (Gall et al. 1999), they provided evidence for a remarkable convergence of the transcriptional and processing machineries. Specifically, by following the localization and movement of numerous required factors, it was shown that not only RNAP II and other components of the transcription, splicing and polyadenylation machineries, but also RNAP I and III and related factors accumulate initially following transport into the nucleus in structures known as Cajal (or coiled) bodies. Here it seems that factors needed for RNAP I-, II-, and III-mediated synthesis associate into massive holocomplexes, or transcriptosomes, and are subsequently transported to sites of RNA synthesis. In the case of RNAP II, "pol II transcriptosomes" are exported from Cajal bodies as the previously described B snurposomes, which consist of multiple pol II transcriptosomes and are likely identical to the interchromatin granule clusters, or speckles, characterized in somatic cells. The B snurposomes would then provide a reservoir of pol II transcriptosomes to active genes, and these particles in turn contain all the factors necessary for complete synthesis of the mature mRNA. Supporting the idea that processing factors are indeed associated with RNAP II throughout transcription, staining of highly active lampbrush chromosomes with antibodies directed against both splicing and polyadenylation factors revealed uniform staining along the chromosome's entire length, coincident with RNAP II staining (Gall et al. 1999). All in all, this picture resembles rather closely that emerging from the biochemical studies.

The idea of a multifunctional pol II transcriptosome provides a neat mechanism for insuring efficient and accurate processing of RNAP II-generated transcripts. But it is not without some conceptual challenges. For example, how would such a massive structure translocate along the DNA? Might we want to reconsider the idea that the DNA is mobile, with the transciptosome being stationary? Most genes have multiple intervening sequences, and considerable data from some twenty years ago indicates that introns are not necessarily removed in a 5'-to-3' order, and in some cases not until transcription is complete. How might the transcriptosome deal with this? One possibility is that the transcriptosome doesn't necessarily contain (a) complete spliceosome(s), but perhaps only factors necessary to define splice sites, or intron boundaries, and for committing the RNA to subsequent splicing. And what about regulation? Are there gene-specific transcriptosomes, or are regulatory factors not part of the transcriptosome? This is an especially vexing problem when considering regulated processing, where it is thought that changes in the relative concentrations of certain essential factors (e.g., SR proteins) can contribute to changing patterns of alternative processing. Despite these questions, the concept of the transcriptosome provides an exciting new way to think about how genes are expressed in the cell nucleus.

Summary and perspectives

The idea that RNAP II, and specifically the CTD of its largest subunit, participates in mRNA processing was unexpected and controversial just a few years ago, but now seems quite solidly established. In part because it's the simplest of the processing reactions, the case is tightest, and the details best understood, for capping: Biochemical and genetic data support a direct, functionally significant interaction between capping enzyme and a specifically phosphorylated form of the CTD that results in an allosteric activation of capping, and this mechanism seems conserved from yeast to humans. But the evidence that RNAP II, via the CTD, actively functions in splicing and polyadenylation is nearly as compelling.

In both these cases, the evidence so far comes largely from biochemical and cell biological studies, and principally from vertebrate systems. In keeping with the greater complexity of these two reactions, it is less clear exactly how the CTD functions. In both, though, it appears that it can act directly, independent of transcription. RNAP II functions very early in the splicing reaction, likely by interacting, directly or indirectly, with key splicing factors to facilitate splice site recognition. The CTD interacts with polyadenylation factors and can play a required role in the cleavage reaction. But beyond that, the mechanisms are unclear, and an important goal now is to elucidate the molecular mechanisms involved. It will be interesting, for example, to learn if there are similarities in how the CTD works in capping, splicing, and 3'-end formation.

Aside from mechanism, a significant issue will concern the importance of these interactions to gene regulation. A number of factors suggest that targeting the transcription-processing link should provide an important avenue of regulation. For example, given the large number of kinases that can phosphorylate the CTD and the enormous number of potential sites, the possibility that differential CTD phosphorylation might influence the constellation of factors associated with it, which could in turn influence alternative processing events, is attractive. The observations that polyadenylation factors can associate with the transcription pre-initiation complex, and that promoter structure can influence alternative splicing patterns, suggest that events that occur at the promoter might help dictate subsequent processing efficiency. For example, could sequence-specific DNA binding proteins, or perhaps transcriptional coactivators, contribute to recruitment of distinct processing factors to promoter-bound RNAP II? It will be critical now to obtain biochemical and/or genetic evidence that some of the potential regulators we have described here indeed function in gene control.

The implications of the transcriptosome theory have the potential to be far reaching. But first we must learn if transcriptosomes are found generally throughout metazoa, and whether they do indeed reflect preassembled, organized sites of transcription and processing. If so, is this an essential pathway of mRNA synthesis? Or might it be designed, for example, to enhance the efficiency with which certain highly expressed mRNAs are produced? Whatever the answers to these and other questions, it is remarkable how cell biology, biochemistry, and in some cases genetics are all providing evidence that processes within the cell nucleus are coordinated to a remarkable and unanticipated degree. The coming decade promises to be a fascinating one for understanding the role that RNA polymerase II plays in orchestrating these events, and the significance of this integration to the mechanisms and regulation of gene expression.

Acknowledgments

References

Albert, A., S. Lavoie, and M. Vincent. 1999. A hyperphosphorylated form of RNA polymerase II is the major interphase antigen of the phosphoprotein antibody MPM-2 and interacts with the peptidyl-prolyl isomerase Pin1. *J. Cell. Sci.* **112:** 2493–2500.

Ares, Jr., M., L. Grate, and M.H. Pauling. 1999. A handful of intron-containing genes produces the lion's share of yeast mRNA. *RNA* **5:** 1138–1139.

Attwooll, C., M. Tariq, M. Harris, J.D. Coyne, N. Telford, and J.M. Varley. 1999. Identification of a novel fusion gene involving hTAFII68 and CHN from a t(9;17)(q22;q11.2) translocation in an extraskeletal myxoid chondrosarcoma. *Oncogene* **18:** 7599–7601.

Baechtold, H., M. Kuroda, J. Sok, D. Ron, B.S. Lopez, and A.T. Akhmedov. 1999. Human 75-kDa DNA-pairing protein is identical to the pro-oncoprotein TLS/FUS and is able to promote D-loop formation. *J. Biol. Chem.* **274:** 34337–34342.

Bauren, G., and L. Wieslander. 1994. Splicing of Balbiani ring 1 gene pre-mRNA occurs simultaneously with transcription. *Cell* **76:** 183–192.

Bauren, G., S. Belikov, and L. Wieslander. 1998. Transcriptional termination in the Balbiani ring 1 gene is closely coupled to 3'-end formation and excision of the 3'-terminal intron. *Genes & Dev.* **12:** 2759–2769.

Bentley, D. 1999. Coupling RNA polymerase II transcription with pre-mRNA processing. *Curr. Opin. Cell. Biol.* **11:** 347–351.

Bertolotti, A., Y. Lutz, D.J. Heard, P. Chambon, and L. Tora. 1996. hTAF(II)68, a novel RNA/ssDNA-binding protein with homology to the pro-oncoproteins TLS/FUS and EWS is associated with both TFIID and RNA polymerase II. *EMBO J.* **15:** 5022–5031.

Beyer, A.L. and Y.N. Osheim. 1988. Splice site selection, rate of splicing, and alternative splicing on nascent transcripts. *Genes & Dev.* **2:** 754–765.

Birse, C.E., L. Minvielle-Sebastia, B.A. Lee, W. Keller, and N.J. Proudfoot. 1998. Coupling termination of transcription to messenger RNA maturation in yeast. *Science* **280:** 298–301.

Bourquin, J.P., I. Stagljar, P. Meier, P. Moosmann, J. Silke, T. Baechi, O. Georgiev, and W. Schaffner. 1997. A serine/arginine-rich nuclear matrix cyclophilin interacts with the C-terminal domain of RNA polymerase II. *Nucleic Acids Res.* **25:** 2055–2061.

Bregman, D.B., L. Du, S. van der Zee, and S.L. Warren. 1995. Transcription-dependent redistribution of the large subunit of RNA polymerase II to discrete nuclear domains. *J. Cell. Biol.* **129:** 287–298.

Burd, C.G. and G. Dreyfuss. 1994. Conserved structures and diversity of functions of RNA-binding proteins. *Science* **265:** 615–621.

Carlson, M. 1997. Genetics of transcriptional regulation in yeast: connections to the RNA polymerase II CTD. *Annu. Rev. Cell. Dev. Biol.* **13:** 1–23.

Caspary, F., A. Shevchenko, M. Wilm, and B. Seraphin. 1999. Partial purification of the yeast U2 snRNP reveals a novel yeast pre-mRNA splicing factor required for pre-spliceosome assembly. *EMBO J.* **18:** 3463–3474.

Chabot, B., S. Bisotto, and M. Vincent. 1995. The nuclear matrix phosphoprotein p255 associates with splicing complexes as part of the [U4/U6.U5] tri-snRNP particle. *Nucleic Acids Res.* **23:** 3206–3213.

Cho, E.J., T. Takagi, C.R. Moore, and S. Buratowski. 1997. mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain. *Genes & Dev.* **11:** 3319–3326.

Cho, E.J., C.R. Rodriguez, T. Takagi, and S. Buratowski. 1998. Allosteric interactions between capping enzyme subunits and the RNA polymase II carboxy-terminal domain. *Genes & Dev.* **12:** 3482–3487.

Cho, H., T.K. Kim, H. Mancebo, W.S. Lane, O. Flores, and D. Reinberg. 1999. A protein phosphatase functions to recycle RNA polymerase II. *Genes & Dev.* **13:** 1540–1552.

Colgan, D.F. and J.L. Manley. 1997. Mechanism and regulation of mRNA polyadenylation. *Genes & Dev.* **11:** 2755–2766.

Connelly, S. and J.L. Manley. 1988. A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. *Genes & Dev.* **2:** 440–452.

Conrad, N.K., S.M. Wilson, E.J. Steinmetz, M. Patturajan, D.A. Brow, M.S. Swanson, and J.L. Corden. 2000. A yeast heterogeneous nuclear ribonucleoprotein complex associated with RNA polymerase II. *Genetics* **154:** 557–571.

Coppola, J.A., A.S. Field, and D.S. Luse. 1983. Promoter-proximal pausing by RNA polymerase II in vitro: transcripts shorter than 20 nucleotides are not capped. *Proc. Natl. Acad. Sci.* **80:** 1251–1255.

Corden, J.L. 1990. Tails of RNA polymerase II. *Trends Biochem. Sci.* **15:** 383–387.

Cramer, P., C.G. Pesce, F.E. Baralle, and A.R. Kornblihtt. 1997. Functional association between promoter structure and transcript alternative splicing. *Proc. Natl. Acad. Sci.* **94:** 11456–11460.

Cramer, P., J.F. Caceres, D. Cazalla, S. Kadener, A.F. Muro, F.E. Baralle, and A.R. Kornblihtt. 1999. Coupling of transcription with alternative splicing: RNA pol II promoters modulate SF2/ASF and 9G8 effects on an exonic splicing enhancer. *Mol. Cell* **4:** 251–258.

Crozat, A., P. Aman, N. Mandahl, and D. Ron. 1993. Fusion of CHOP to a novel RNA-binding protein in human myxoid liposarcoma. *Nature* **363:** 640–644.

Dahmus, M.E. 1996. Reversible phosphorylation of the C-terminal domain of RNA polymerase II. *J. Biol. Chem.* **271:** 19009–19012.

Dantonel, J.C., K.G. Murthy, J.L. Manley, and L. Tora. 1997. Transcription factor TFIID recruits factor CPSF for formation of 3′ end of mRNA. *Nature* **389:** 399–402.

Delattre, O., J. Zucman, B. Plougastel, C. Desmaze, T. Melot, M. Peter, H. Kovar, I. Joubert, P. de Jong, and G. Rouleau. 1992. Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature* **359:** 162–165.

Du, L. and S.L. Warren. 1997. A functional interaction between the carboxy-terminal domain of RNA polymerase II and pre-mRNA splicing. *J. Cell. Biol.* **136:** 5–18.

Dye, M.J. and N.J. Proudfoot. 1999. Terminal exon definition occurs cotranscriptionally and promotes termination of RNA polymerase II. *Mol. Cell* **3:** 371–378.

Edery, I. and N. Sonenberg. 1985. Cap-dependent RNA splicing in a HeLa nuclear extract. *Proc. Natl. Acad. Sci.* **82:** 7590–7594.

Flaherty, S.M., P. Fortes, E. Izaurralde, I.W. Mattaj, and G.M. Gilmartin. 1997. Participation of the nuclear cap binding complex in pre-mRNA 3′ processing. *Proc. Natl. Acad. Sci.* **94:** 11893–11898.

Gall, J.G., M. Bellini, Z. Wu, and C. Murphy. 1999. Assembly of the nuclear transcription and processing machinery: Cajal bodies (coiled bodies) and transcriptosomes. *Mol. Biol. Cell* **10:** 4385–4402.

Ge, H., Y. Si, and R.G. Roeder. 1998a. Isolation of cDNAs encoding novel transcription coactivators p52 and p75 reveals an alternate regulatory mechanism of transcriptional activation. *EMBO J.* **17:** 6723–6729.

Ge, H., Y. Si, and A.P. Wolffe. 1998b. A novel transcriptional coactivator, p52, functionally interacts with the essential splicing factor ASF/SF2. *Mol. Cell* **2:** 751–759.

Gothel, S.F. and M.A. Marahiel. 1999. Peptidyl-prolyl cis-trans isomerases, a superfamily of ubiquitous folding catalysts. *Cell. Mol. Life Sci.* **55:** 423–436.

Greenleaf, A.L. 1993. Positive patches and negative noodles: linking RNA processing to transcription? *Trends Biochem. Sci.* **18:** 117–119.

Hallier, M., A. Lerga, S. Barnache, A. Tavitian, and F. Moreau-Gachelin. 1998. The transcription factor Spi-1/PU.1 interacts with the potential splicing factor TLS. *J. Biol. Chem.* **273:** 4838–4842.

Hanes, S.D., P.R. Shank, and K.A. Bostian. 1989. Sequence and mutational analysis of ESS1, a gene essential for growth in Saccharomyces cerevisiae. *Yeast* **5:** 55–72.

Hani, J., G. Stumpf, and H. Domdey. 1995. PTF1 encodes an essential protein in *Saccharomyces cerevisiae*, which shows strong homology with a new putative family of PPIases. *FEBS Lett.* **365:** 198–202.

Hani, J., B. Schelbert, A. Bernhardt, H. Domdey, G. Fischer, K. Wiebauer, and J.U. Rahfeld. 1999. Mutations in a peptidyl-prolyl-cis/trans-isomerase gene lead to a defect in 3′-end formation of a pre-mRNA in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **274:** 108–116.

Hart, R.P., M.A. McDevitt, and J.R. Nevins. 1985. Poly(A) site cleavage in a HeLa nuclear extract is dependent on downstream sequences. *Cell* **43:** 677–683.

Hartzog, G.A., T. Wada, H. Handa, and F. Winston. 1998. Evidence that Spt4, Spt5, and Spt6 control transcription elongation by RNA polymerase II in *Saccharomyces cerevisiae*. *Genes & Dev.* **12:** 357–369.

Hirose, Y. and J.L. Manley. 1997. Creatine phosphate, not ATP, is required for 3′-end cleavage of mammalian pre-mRNA in vitro. *J. Biol. Chem.* **272:** 29636–29642.

———. 1998. RNA polymerase II is an essential mRNA polyadenylation factor. *Nature* **395:** 93–96.

Hirose, Y., R. Tacke, and J.L. Manley. 1999. Phosphorylated RNA polymerase II stimulates pre-mRNA splicing. *Genes & Dev.* **13:** 1234–1239.

Ho, C.K. and S. Shuman. 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.

Ho, C.K., V. Sriskanda, S. McCracken, D. Bentley, B. Schwer, and S. Shuman. 1998a. The guanylyltransferase domain of mammalian mRNA capping enzyme binds to the phosphorylated carboxyl-terminal domain of RNA polymerase II. *J. Biol. Chem.* **273:** 9577–9585.

Ho, C.K., B. Schwer, and S. Shuman. 1998b. Genetic, physical, and functional interactions between the triphosphatase and guanylyltransferase components of the yeast mRNA capping apparatus. *Mol. Cell. Biol.* **18:** 5189–5198.

Hunter, T. 1998. Prolyl isomerases and nuclear function. *Cell* **92:** 141–143.

Jove, R. and J.L. Manley. 1982. Transcription initiation by RNA polymerase II is inhibited by S-adenosylhomocysteine. *Proc. Natl. Acad. Sci.* **79:** 5842–5846.

———. 1984. In vitro transcription from the adenovirus 2 major late promoter utilizing templates truncated at promoter-proximal sites. *J. Biol. Chem.* **259:** 8513–8521.

Kim, E., L. Du, D.B. Bregman, and S.L. Warren. 1997. Splicing

factors associate with hyperphosphorylated RNA polymerase II in the absence of pre-mRNA. *J. Cell. Biol.* **136:** 19–28.

Kleiman, F.E. and J.L. Manley. 1999. Functional interaction of BRCA1-associated BARD1 with polyadenylation factor CstF-50. *Science* **285:** 1576–1579.

Kramer, A. 1996. The structure and function of proteins involved in mammalian pre-mRNA splicing. *Annu. Rev. Biochem.* **65:** 367–409.

Lai, M.C., B.H. Teh, and W.Y. Tarn. 1999. A human papillomavirus E2 transcriptional activator. The interactions with cellular splicing factors and potential function in pre-mRNA processing. *J. Biol. Chem.* **274:** 11832–11841.

Lavigueur, A., H. LaBranche, A.R. Kornblihtt, and B. Chabot. 1993. A splicing enhancer in the human fibronectin alternate ED1 exon interacts with SR proteins and stimulates U2 snRNP binding. *Genes & Dev.* **7:** 2405–2417.

Liao, S.M., J. Zhang, D.A. Jeffery, A.J. Koleske, C.M. Thompson, D.M. Chao, M. Viljoen, H.J. van Vuuren, and R.A. Young. 1995. A kinase-cyclin pair in the RNA polymerase II holoenzyme. *Nature* **374:** 193–196.

Lo, H.J., H.K. Huang, and T.F. Donahue. 1998. RNA polymerase I-promoted HIS4 expression yields uncapped, polyadenylated mRNA that is unstable and inefficiently translated in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **18:** 665–675.

Logan, J., E. Falck-Pedersen, J.E. Darnell, Jr., and T. Shenk. 1987. A poly(A) addition site and a downstream termination region are required for efficient cessation of transcription by RNA polymerase II in the mouse beta maj-globin gene. *Proc. Natl. Acad. Sci.* **84:** 8306–8310.

Lopez, P.J. and B. Seraphin. 1999. Genomic-scale quantitative analysis of yeast pre-mRNA splicing: implications for splice-site recognition. *RNA* **5:** 1135–1137.

Lu, K.P., S.D. Hanes, and T. Hunter. 1996. A human peptidyl-prolyl isomerase essential for regulation of mitosis. *Nature* **380:** 544–547.

Manley, J.L. and R. Tacke. 1996. SR proteins and splicing control. *Genes & Dev.* **10:** 1569–1579.

McCracken, S., N. Fong, E. Rosonina, K. Yankulov, G. Brothers, D. Siderovski, A. Hessel, S. Foster, S. Shuman, and D.L. Bentley. 1997a. 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.

McCracken, S., N. Fong, K. Yankulov, S. Ballantyne, G. Pan, J. Greenblatt, S.D. Patterson, M. Wickens, and D.L. Bentley. 1997b. The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature* **385:** 357–361.

McNeil, J.B., H. Agah, and D. Bentley. 1998. Activated transcription independent of the RNA polymerase II holoenzyme in budding yeast. *Genes & Dev.* **12:** 2510–2521.

Minvielle-Sebastian, L. and W. Keller. 1999. mRNA polyadenylation and its coupling to other RNA processing reactions and to transcription. *Curr. Opin. Cell. Biol.* **11:** 352–357.

Mintz, P.J., S.D. Patterson, A.F. Neuwald, C.S. Spahr, and D.L. Spector. 1999. Purification and biochemical characterization of interchromatin granule clusters. *EMBO J.* **18:** 4308–4320.

Misteli, T. and D.L. Spector. 1998. The cellular organization of gene expression. *Curr. Opin. Cell. Biol.* **10:** 323–331.

———. 1999. RNA polymerase II targets pre-mRNA splicing factors to transcription sites in vivo. *Mol. Cell* **3:** 697–705.

Misteli, T., J.F. Caceres, and D.L. Spector. 1997. The dynamics of a pre-mRNA splicing factor in living cells. *Nature* **387:** 523–527.

Moore, M.J., C.C. Query and P.A. Sharp. 1993. Splicing of precursors to mRNAs by the spliceosome. In *The RNA World* (ed. R.F. Gesteland et al.), pp. 303–358. Cold Spring Harbor

Laboratory Press, Cold Spring Harbor, NY.

Mortillaro, M.J. and R. Berezney. 1998. Matrin CYP, an SR-rich cyclophilin that associates with the nuclear matrix and splicing factors. *J. Biol. Chem.* **273:** 8183–8192.

Mortillaro, M.J., B.J. Blencowe, X. Wei, H. Nakayasu, L. Du, S.L. Warren, P.A. Sharp, and R. Berezney. 1996. A hyperphosphorylated form of the large subunit of RNA polymerase II is associated with splicing complexes and the nuclear matrix. *Proc. Natl. Acad. Sci.* **93:** 8253–8257.

Morris, D.P., H.P. Phatnani, and A.L. Greenleaf. 1999. Phospho-carboxyl-terminal domain binding and the role of a prolyl isomerase in pre-mRNA 3′-end formation. *J. Biol. Chem.* **274:** 31583–31587.

Neugebauer, K.M. and M.B. Roth. 1997. Transcription units as RNA processing units. *Genes & Dev.* **11:** 3279–3285.

Niwa, M. and S.M. Berget. 1991. Mutation of the AAUAAA polyadenylation signal depresses in vitro splicing of proximal but not distal introns. *Genes & Dev.* **5:** 2086–2095.

Niwa, M., S.D. Rose, and S.M. Berget. 1990. In vitro polyadenylation is stimulated by the presence of an upstream intron. *Genes & Dev.* **4:** 1552–1559.

Orphanides, G., T. Lagrange, and D. Reinberg. 1996. The general transcription factors of RNA polymerase II. *Genes & Dev.* **10:** 2657–2683.

Osheim, Y.N., N.J. Proudfoot, and A.L. Beyer. 1999. EM visualization of transcription by RNA polymerase II: downstream termination requires a poly (A) signal but not transcript cleavage. *Mol. Cell* **3:** 379–387.

Patturajan, M., X. Wei, R. Berezney, and J.L. Corden. 1998. A nuclear matrix protein interacts with the phosphorylated C-terminal domain of RNA polymerase II. *Mol. Cell. Biol.* **18:** 2406–2415.

Potashkin, J., D. Kim, M. Fond, T. Humphrey, and D. Frendewey. 1998. Cell-division-cycle defects associated with fission yeast pre-mRNA splicing mutants. *Curr. Genet.* **34:** 153–163.

Proudfoot, N.J. 1989. How RNA polymerase II terminates transcription in higher eukaryotes. *Trends Biochem. Sci.* **14:** 105–110.

Reines, D., R.C. Conaway, and J.W. Conaway. 1999. Mechanism and regulation of transcriptional elongation by RNA polymerase II. *Curr. Opin. Cell. Biol.* **11:** 342–346.

Rodriguez, C.R., E.J. Cho, M.C. Keogh, C.L. Moore, A.L. Greenleaf, and S. Buratowski. 2000. Kin28, the TFIIH-associated carboxy-terminal domain kinase, facilitates the recruitment of mRNA processing machinery to RNA polymerase II. *Mol. Cell. Biol.* **20:** 104–112.

Schul, W., R. van Driel, and L. de Jong. 1998. A subset of poly(A) polymerase is concentrated at sites of RNA synthesis and is associated with domains enriched in splicing factors and poly(A) RNA. *Exp. Cell. Res.* **238:** 1–12.

Scully, R., S.F. Anderson, D.M. Chao, W. Wei, L. Ye, R.A. Young, D.M. Livingston, and J.D. Parvin. 1997a. BRCA1 is a component of the RNA polymerase II holoenzyme. *Proc. Natl. Acad. Sci.* **94:** 5605–5610.

Scully, R., J. Chen, R.L. Ochs, K. Keegan, M. Hoekstra, J. Feunteun, and D.M. Livingston. 1997b. Dynamic changes of BRCA1 subnuclear location and phosphorylation state are initiated by DNA damage. *Cell* **90:** 425–435.

Shen, M., P.T. Stukenberg, M.W. Kirschner, and K.P. Lu. 1998. The essential mitotic peptidyl-prolyl isomerase Pin1 binds and regulates mitosis-specific phosphoproteins. *Genes & Dev.* **12:** 706–720.

Shuman, S. 1995. Capping enzyme in eukaryotic mRNA synthesis. *Prog. Nucleic Acid Res. Mol. Biol.* **50:** 101–129.

Singer, R.H. and M.R. Green. 1997. Compartmentalization of

eukaryotic gene expression: causes and effects. *Cell* **91:** 291–294.

Sjogren, H., J. Meis-Kindblom, L.G. Kindblom, P. Aman, and G. Stenman. 1999. Fusion of the EWS-related gene TAF2N to TEC in extraskeletal myxoid chondrosarcoma. *Cancer Res.* **59:** 5064–5067.

Spector, D.L. 1993. Macromolecular domains within the cell nucleus. *Annu. Rev. Cell. Biol.* **9:** 265–315.

Steinmetz, E.J. 1997. Pre-mRNA processing and the CTD of RNA polymerase II: the tail that wags the dog? *Cell* **89:** 491–494.

Steinmetz, E.J. and D.A. Brow. 1996. Repression of gene expression by an exogenous sequence element acting in concert with a heterogeneous nuclear ribonucleoprotein-like protein, Nrd1, and the putative helicase Sen1. *Mol. Cell. Biol.* **16:** 6993–7003.

———. 1998. Control of pre-mRNA accumulation by the essential yeast protein Nrd1 requires high-affinity transcript binding and a domain implicated in RNA polymerase II association. *Proc. Natl. Acad. Sci.* **95:** 6699–6704.

Sterner, D.E., J.M. Lee, S.E. Hardin, and A.L. Greenleaf. 1995. The yeast carboxyl-terminal repeat domain kinase CTDK-I is a divergent cyclin-cyclin-dependent kinase complex. *Mol. Cell. Biol.* **15:** 5716–5724.

Svejstrup, J.Q., W.J. Feaver, and R.D. Kornberg. 1996. Subunits of yeast RNA polymerase II transcription factor TFIIH encoded by the CCL1 gene. *J. Biol. Chem.* **271:** 643–645.

Swanson, M.S. and F. Winston. 1992. SPT4, SPT5 and SPT6 interactions: effects on transcription and viability in saccharomyces cerevisiae. *Genetics* **132:** 325–336.

Takagaki, Y. and J.L. Manley. 1998. Levels of polyadenylation factor CstF-64 control IgM heavy chain mRNA accumulation and other events associated with B cell differentiation. *Mol. Cell* **2:** 761–771.

———. 2000. Complex protein interactions within the human polyadenylation machinery identify a novel component. *Mol. Cell. Biol.* **20:** 1515–1525.

Wada, T., T. Takagi, Y. Yamaguchi, A. Ferdous, T. Imai, S. Hirose, S. Sugimoto, K. Yano, G.A. Hartzog, F. Winston, S. Buratowski, and H. Handa. 1998a. DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. *Genes & Dev.* **12:** 343–356.

Wada, T., T. Takagi, Y. Yamaguchi, D. Watanabe, and H. Handa. 1998b. Evidence that P-TEFb alleviates the negative effect of DSIF on RNA polymerase II-dependent transcription in vitro. *EMBO J.* **17:** 7395–7403.

Weeks, J.R., S.E. Hardin, J. Shen, J.M. Lee, and A.L. Greenleaf. 1993. Locus-specific variation in phosphorylation state of RNA polymerase II in vivo: Correlations with gene activity and transcript processing. *Genes & Dev.* **7:** 2329–2344.

Wen, Y. and A.J. Shatkin. 1999. Transcription elongation factor hSPT5 stimulates mRNA capping. *Genes & Dev.* **13:** 1774–1779.

Whitelaw, E. and N. Proudfoot. 1986. Alpha-thalassaemia caused by a poly(A) site mutation reveals that transcriptional termination is linked to 3′ end processing in the human alpha 2 globin gene. *EMBO J.* **5:** 2915–2922.

Wu, L.C., Z.W. Wang, J.T. Tsan, M.A. Spillman, A. Phung, X.L. Xu, M.C. Yang, L.Y. Hwang, A.M. Bowcock, and R. Baer. 1996. Identification of a RING protein that can interact in vivo with the BRCA1 gene product. *Nat. Genet.* **14:** 430–440.

Yaffe, M.B., M. Schutkowski, M. Shen, X.Z. Zhou, P.T. Stukenberg, J.U. Rahfeld, J. Xu, J. Kuang, M.W. Kirschner, G. Fischer, L.C. Cantley, and K.P. Lu. 1997. Sequence-specific

and phosphorylation-dependent proline isomerization: a potential mitotic regulatory mechanism. *Science* **278:** 1957–1960.

Yang, L., L.J. Embree, T. Schickwann, and D.D. Hickstein. 1998. Oncoprotein TLS interacts with serine-arginine proteins involved in RNA splicing. *J. Biol. Chem.* **273:** 27761–27764.

Yonaha, M. and N.J. Proudfoot. 1999. Specific transcriptional pausing activates polyadenylation in a coupled in vitro system. *Mol. Cell* **3:** 593–600.

Yue, Z., E. Maldonado, R. Pillutla, H. Cho, D. Reinberg, and A.J. Shatkin. 1997. Mammalian capping enzyme complements mutant *Saccharomyces cerevisiae* lacking mRNA guanylyltransferase and selectively binds the elongating form of RNA polymerase II. *Proc. Natl. Acad. Sci.* **94:** 12898–12903.

Yuryev, A., M. Patturajan, Y. Litingtung, R.V. Joshi, C. Gentile, M. Gebara, and J.L. Corden. 1996. The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins. *Proc. Natl. Acad. Sci.* **93:** 6975–6980.

Zhang, D., A.J. Paley, and G. Childs. 1998. The transcriptional repressor ZFM1 interacts with and modulates the ability of EWS to activate transcription. *J. Biol. Chem.* **273:** 18086–18091.

Zhang, G., K.L. Taneja, R.H. Singer, and M.R. Green. 1994. Localization of pre-mRNA splicing in mammalian nuclei. *Nature* **372:** 809–812.

Zhao, J., L. Hyman, and C. Moore. 1999a. Formation of mRNA 3′ ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol. Mol. Biol.* **63:** 405–445.

Zhao, J., M. Kessler, S. Helmling, J.P. O'Connor, and C. Moore. 1999b. Pta1, a component of yeast CF II, is required for both cleavage and poly(A) addition of mRNA precursor. *Mol. Cell. Biol.* **19:** 7733–7740.

Zinszner, H., R. Albalat, and D. Ron. 1994. A novel effector domain from the RNA-binding protein TLS or EWS is required for oncogenic transformation by CHOP. *Genes & Dev.* **8:** 2513–2526.

RNA polymerase II and the integration of nuclear events

Yutaka Hirose and James L. Manley

References	This article cites 118 articles, 70 of which can be accessed free at: **http://genesdev.cshlp.org/content/14/12/1415.full.html#ref-list-1**
License	
Email Alerting Service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**