

## Review Article

# Updating the CTD Story: From Tail to Epic

**Bartłomiej Bartkowiak, April L. MacKellar, and Arno L. Greenleaf**

*Department of Biochemistry and Center for RNA Biology, Duke University Medical Center, Durham, NC 27710, USA*

Correspondence should be addressed to Arno L. Greenleaf, arno@biochem.duke.edu

Received 27 June 2011; Accepted 10 August 2011

Academic Editor: Carles Sune

Copyright © 2011 Bartłomiej Bartkowiak et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Eukaryotic RNA polymerase II (RNAPII) not only synthesizes mRNA but also coordinates transcription-related processes via its unique C-terminal repeat domain (CTD). The CTD is an RNAPII-specific protein segment consisting of repeating heptads with the consensus sequence  $Y_1S_2P_3T_4S_5P_6S_7$  that has been shown to be extensively post-transcriptionally modified in a coordinated, but complicated, manner. Recent discoveries of new modifications, kinases, and binding proteins have challenged previously established paradigms. In this paper, we examine results and implications of recent studies related to modifications of the CTD and the respective enzymes; we also survey characterizations of new CTD-binding proteins and their associated processes and new information regarding known CTD-binding proteins. Finally, we bring into focus new results that identify two additional CTD-associated processes: nucleocytoplasmic transport of mRNA and DNA damage and repair.

## 1. Introduction

Since its discovery by Fischer and Krebs in 1955 [1], the reversible phosphorylation of proteins has been implicated in the regulation of almost every aspect of cellular function, including metabolism, cell division, differentiation, signaling, and countless others. A particularly fascinating form of this regulation is employed during the transcription of DNA by RNA Polymerase II (RNAPII). Eukaryotic transcription and the concomitant pre-mRNA processing require the precise coordination between, and recruitment of, specific sets of factors at specific stages of the transcription cycle. This coupling of transcription and associated processes has been shown to be dependent on a particular feature of RNAPII, the C-terminal repeat domain or CTD [2]. Distinguishing RNAPII from its prokaryotic and eukaryotic (RNAPIII and RNAPI) counterparts, the CTD is an extension of the polymerase's largest subunit, Rpb1, and is composed of a tandem array of seven amino acid repeats with the consensus sequence  $Y_1S_2P_3T_4S_5P_6S_7$ . The number of these heptad repeats varies from organism to organism and appears to correlate with genomic complexity; there are 26 repeats in yeast, 44 in *Drosophila*, and 52 in humans [3, 4]. Despite being dispensable for the catalytic activity of RNAPII, the

CTD is conserved through evolution and is essential for life; for example, removing two thirds of the CTD repeats results in inviability [5, 6]. Although the function of the CTD remained elusive for several years after its discovery, research over the last three decades has confirmed its role as a selective and flexible scaffold for numerous factors involved in transcription (for reviews see [2, 7, 8]). The plastic (both in terms of conformation and susceptibility to post-translational modification) and repetitive nature of the CTD allows it to undergo a relatively well-characterized sequence of phosphorylation and dephosphorylation events during the transcription cycle (initiation, elongation, and termination), linking transcription with transcription-associated processes in a temporal manner [2, 7, 9]. The CTD has been shown to play a role in a wide variety of transcription-associated functions, and its repertoire of binding partners, modifications, and associated processes has grown rapidly over the last few years.

The CTD's unique structure, functional characteristics, and fundamental role in transcription have generated a substantial amount of interest and have made it the subject of considerable study. This research has greatly expanded our understanding of the CTD, its interacting factors, and the process of transcription in general, but as demonstrated

by recent discoveries of hitherto uncharacterized CTD modifications, kinases, and binding factors/modes, there is still much that remains to be learned. As the study of gene expression and transcriptional control mechanisms expands from preinitiation into elongation, a deeper and more nuanced understanding of the CTD will likely become essential in order to deconvolute the relationships between various aspects of gene control. This is especially pertinent in terms of understanding the crosstalk between transcriptional elongation and co/post-transcriptional events, such as splicing, export, and translation. In this review, we will discuss recent developments and emerging paradigms in the study of the CTD, its modifications, binding partners, and associated processes. To further expand on an especially relevant (in terms of transcriptional control) and newly emerging CTD-associated process, we will also present an in-depth discussion of the nuclear export of RNA, with a particular focus on the interactions between the nuclear export machinery, the CTD, and transcription.

## 2. Phosphorylation of the CTD of RNAPII and the Transcription Cycle: A General Model and Its Limitations

The binding specificity of the CTD, and therefore the recruitment of particular factors, is determined by the CTD's phosphorylation state, which undergoes a series of alterations throughout the transcription cycle. Serine 2, serine 5, and more recently serine 7 of the heptad repeat have been identified as the primary targets for this transcriptionally regulated phosphorylation. In the general model of the "phosphoCTD cycle," RNAPII is recruited for assembly at the promoter with an unphosphorylated CTD; moreover, it appears that CTD phosphorylation prior to preinitiation complex (PIC) formation has an inhibitory effect on transcription [10]. Upon PIC formation, the CTD is phosphorylated at the Ser5 and Ser7 positions; this is followed by an increase in Ser2 phosphorylation during elongation [11], yielding a CTD that contains a mix of doubly (and when considering Ser7, perhaps triply) phosphorylated repeats at Ser2 and Ser5 in the center of the gene (reviewed in [2, 9]). As the polymerase elongates towards the 3' end of the gene, Ser5-specific phosphatases decrease the Ser5 phosphorylation levels (on noncoding genes, Ser7P levels drop as well), leaving the CTD phosphorylated at Ser2 to terminate transcription (Figure 1(a)). However, the CTD of terminating RNAPII may also be phosphorylated at Ser7 positions, as the Ser7 mark has been reported to be present at high levels throughout the entire transcription unit on many protein-coding genes [12].

Despite being highly intuitive, this gradient model of the "phosphoCTD cycle" (i.e., high-to-low Ser5P levels and low-to-high Ser2P levels as RNAPII moves across a gene, with Ser7P throughout) is oversimplified for several reasons. First, although the model accounts for variations in the general phosphorylation pattern of the CTD (recognizing the presence of gene segments with high Ser5P, high Ser5/2P, and high Ser2P levels), it fails to fully capture the highly

dynamic nature of the process. The phosphorylation state of the CTD is likely to be in continuous flux throughout the entire transcription cycle, with multiple kinases and phosphatases working together to maintain specific phosphorylation patterns on particular subsets of heptads. For example, it has been recently shown that the Ser7 phosphates at the 3' end of the gene are actually placed anew, after being removed by an unidentified phosphatase [12] (more on this below). Another complication is the existence of specific patterns of phosphorylated repeats within the CTD. Our understanding of this particular aspect of CTD phosphorylation is entangled with some of the limitations of chromatin immunoprecipitation (ChIP), the method used to characterize the phosphorylation state of the CTD at different positions within the transcription unit; thus it is convenient to discuss both of these topics together. Although ChIP can be used to measure the relative level of each phosphoCTD mark (Ser2P, Ser5P, or Ser7P) at a specific position within a gene, the repetitive nature of the CTD makes it impossible to determine exactly which of the 52 heptads (in mammals) are phosphorylated. In addition, the particular pattern of phosphorylated residues within each heptad and within sequential heptad units is also unknown; therefore, the Ser2, 5, and 7 phosphorylated CTD of elongating RNAPII is likely to be composed of heptads phosphorylated at one, both, all, and none of the relevant positions. Among the many unanswered questions concerning the detailed phosphorylation state of the CTD, an especially interesting one concerns determinism: is the phosphorylation pattern of a specific heptad at a particular position within a gene exactly the same during each transcription cycle, or does phosphorylation occur in a more stochastic manner? The answer to these questions may lie within the processivity of the CTD kinases, the extent to which they stimulate each other's activity, and the specifics of their antagonistic relationship with the CTD phosphatases.

Discussion of CTD phosphorylation patterns also raises an interesting point regarding the antibodies used to identify the post-translationally modified heptads (both in ChIP and Western blotting). It is worth reiterating that the Ser5P to Ser2P gradient model presented above is almost entirely based on the reactivity of these antibodies [13]. The most common phosphoCTD antibodies are the Ser2P-specific H5 and the Ser5P-specific H14 [14]; however, there are a multitude of other antibodies available against the CTD, including the recent and well-characterized anti-Ser2P 3E10 and anti-Ser5P 3E8 antibodies [15]. As might be expected, the reactivity of these antibodies with their specific phospho-epitopes is affected by other modifications within the heptad in question and the heptads around it. For example, the Ser2P-specific H5 antibody binds to a triheptad peptide phosphorylated on both Ser2 and Ser5 with greater affinity than to a triheptad containing only Ser2P [16]. More impressive examples are the complete lack of H14 reactivity to a single S5P in the first repeat of a diheptad peptide and the inability of 3E10 to detect an S2P followed by an S5P in the next heptad [15]. Thus, the phosphoCTD antibodies are multiheptad pattern specific, and it appears that many CTD-binding proteins are as

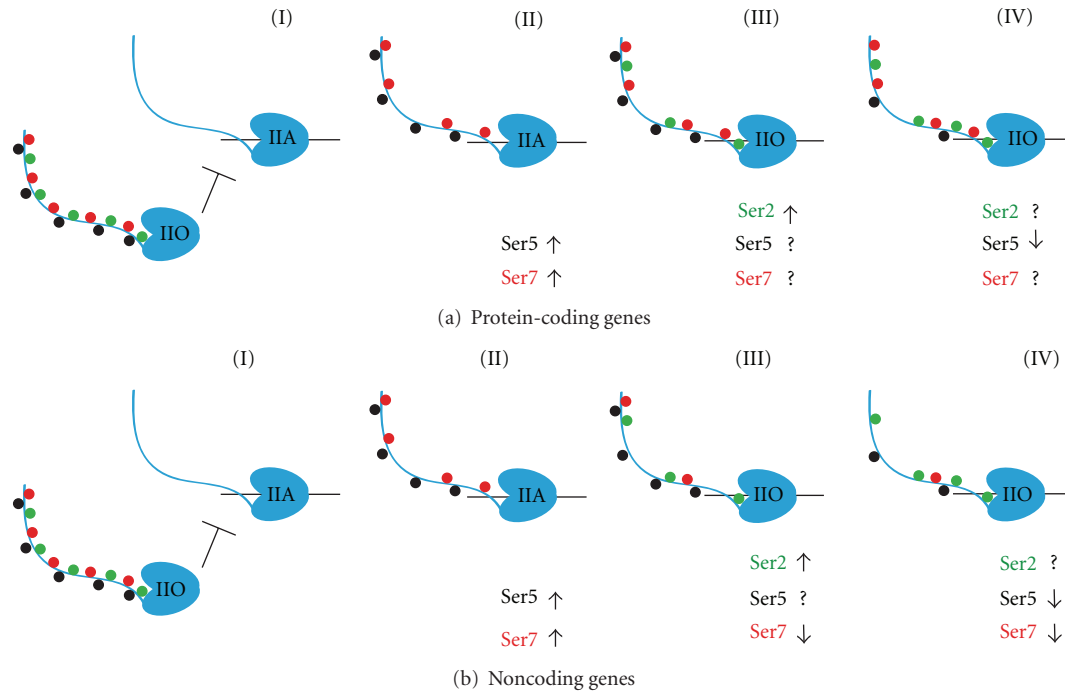


FIGURE 1: Revised “phospho-CTD cycle.” (a) (I) For protein-coding genes, RNAPII is recruited to the promoter with an unphosphorylated CTD (IIA form); moreover, it appears that CTD phosphorylation prior to preinitiation complex (PIC) formation has an inhibitory effect on transcription. (II) Upon preinitiation complex formation, the CTD is phosphorylated at the Ser5 (black) and Ser7 (red) positions. (III) During elongation, an increase in Ser2 phosphorylation (green) produces the hyperphosphorylated form of the CTD, which is probably an ensemble of singly, doubly, and triply phosphorylated heptads. (IV) As the polymerase elongates towards the 3’ end of the gene, the activity of Ser5-specific phosphatases decreases the Ser5 phosphorylation levels, while the Ser2 and Ser7 phosphate levels remain largely unchanged. (b) (I) For noncoding genes, RNAPII is also recruited to the promoter with an unphosphorylated CTD (IIA form), and CTD phosphorylation prior to preinitiation complex (PIC) formation seems to inhibit transcription. (II) Upon preinitiation complex formation, the CTD phosphorylation of Ser5 (black) and Ser7 (red) increases. (III) During elongation, Ser2 phosphorylation (green) increases, while Ser7 phosphorylation begins to decline, presumably due to the activity of a yet unidentified Ser7 phosphatase. (IV) At the 3’ end of the gene, the activity of Ser5-specific phosphatases decreases the Ser5 phosphorylation levels, Ser7 phosphorylation levels continue to decrease, and Ser2 phosphate levels remain largely unchanged.

well (see below); this correlates nicely with evolutionary studies that implicate heptad pairs as the functional unit of the CTD [17, 18]. These observations suggest that the in-depth characterization of CTD modification patterns will be important for a comprehensive understanding of factor recruitment/binding and is not a purely academic exercise. Although the ramifications discussed above (along with a few others, such as epitope masking (for a more detailed discussion see [2])) do not in any way invalidate phospho-CTD ChIP, they should be kept in mind when interpreting such data.

The general model of the phospho-CTD cycle also fails to account for exceptions to the canonical patterns of phosphorylation. The recent publication of three genome-wide studies of phosphoCTD RNAPII occupancies in yeast [12, 19, 20] has allowed for the verification of the S5P to S2P gradient model at high resolution across the entire genome. Overall, the general model appears to hold for the majority of genes [19, 20]; however, there seem to be a number of exceptions to the defined norm. The extent of these exceptions has led the authors of one of the studies to call for a reanalysis of the accepted paradigm [12], while

the other two groups find that the general pattern occurs globally [20] or near globally with some stipulations [19]. It should be noted that the discrepancies between these studies might be due to the different methods used to bin/cluster genes into “average transcription units” for analysis. This binning, which is often limited by polymerase occupancy and complicated by gene length and the presence of neighboring transcription units, is especially challenging due to the compact nature of the yeast genome. One consensus that appears to emerge from the genome-wide studies is that different classes of RNAPII-transcribed genes have different CTD phosphorylation profiles. Examples include the high levels of Ser7P throughout the length of protein-coding genes (as compared to its decrease 5’ to 3’ on noncoding genes), markedly lower levels of Ser2P on noncoding genes, and enrichment/preference for specific modifications on genes for snRNAs, cryptic unstable transcripts (CUTs), stable untranslated transcripts (SUTs), genes of different lengths, and genes with different promoter classes [19, 20]. These phosphoCTD variations correlate with the different requirements for the termination and processing of distinct transcript types and make intuitive sense in the light of the

CTD's role during transcription. Therefore, future versions of the "phosphoCTD cycle" will need to take such class-specific differences into account. In our revised model, we chose to separate the general cycle into two broad, but distinct gene classes: protein-coding and noncoding genes (Figures 1(a) and 1(b)). However, this revised model is still an oversimplified representation of a complex process and should be viewed as such.

### 3. Other Post-Translational Modifications of the CTD

Lastly (in terms of complications), it has not escaped notice that modification of the CTD is not theoretically limited to the phosphorylation of Ser2, Ser5, and Ser7. Many other post-translational modifications have been observed, including phosphorylation of Tyr1 [21] and Thr4 [22] and glycosylation [23]; however, the extent and transcriptional functions of these modifications are currently unknown (for more discussion, see [8]). These "noncanonical" modifications, once fully characterized as functionally significant, have the potential to expand the CTD code further and redefine aspects of the general model; Ser7, a relative newcomer to the general paradigm, is a good example of how this can occur.

Yet another important CTD modification, which until recently had not been directly observed to play a direct role in factor binding, is the enzymatic isomerization of the heptad repeat's peptidyl-proline bonds. Although multiple studies have suggested a role for the CTD-interacting peptidyl-prolyl *cis/trans* isomerases (Ess1 in yeast and Pin1 in humans) in transcription and CTD phosphorylation [24–26], all of the structures of CTD-substrates/CTD-binding protein complexes revealed the CTD proline residues to be exclusively in the more energetically stable, and therefore predominant, *trans* state. This changed last year when two structural studies found that the Ser5-specific CTD phosphatase Ssu72 bound to the *cis* conformation of an Ser5-Pro6 motif within the heptad repeat [27, 28]. Concordantly, the activity of the proline isomerase Ess1 was found to facilitate the rapid dephosphorylation of the CTD by Ssu72 *in vitro*, suggesting that this *cis/trans* interconversion plays a role in the fine-tuning of the phosphorylation state of the CTD [27]. These findings have broad implications for CTD biology, both by increasing the number of distinct CTD states and serving as a regulatory mechanism for CTD phosphorylation. However, it still remains to be determined whether proline isomerization is a general property of RNAPII transcription or if it is gene specific [27], a distinction that may apply to other types of modifications as well.

A good example of a transcript class-specific CTD modification is the newly discovered methylation of an arginine (R1810) in heptad 31 of the human CTD [29]. As an apology for the arginine (one of two in the human CTD), it should be noted that while the first 26 repeats of the human CTD conform strongly to the consensus sequence (YSPTSPS), there is significant divergence from the consensus in the C-terminal half of the CTD [30]. It has

been previously postulated that the various noncanonical heptads (and even particular segments of the CTD; such as the N- and C-termini [31]) may have specific functions, and this arginine methylation seems to be a case in point (for further discussion, please see [30]). Mediated by the methyltransferase CARM1 and inhibited by Ser5 and Ser2 phosphorylation, the methylation appears to repress the expression of snRNAs and snoRNAs in a general manner [29]. This and other modifications of the noncanonical heptads may serve as a discriminatory mark for RNAPII recruited to particular genes or transcript classes. It should also be noted that Ser7P is currently thought to be transcript class-specific CTD modification, as Ser7 to alanine mutations in the CTD cause a defect in snRNA transcription while having little effect on protein-coding genes [32]. However, the ubiquitous nature of Ser7P on protein-coding genes, along with the finding that Ser7 is enriched on RNAPII within introns [19], argues for some (perhaps more subtle) functional role for Ser7P on most transcription units.

Thus, the general "phosphoCTD cycle" has given way to a "CTD code" of staggering complexity, one that we are just beginning to explore in detail. This complexity reflects the vast number of different genes, processing events, and transcriptional programs that RNAPII must coordinate. Although the segmented gradient model has proven to be very useful for conveying the CTD's principal function during RNAPII transcription, as our understanding of the CTD and associated processes improves, it is likely to undergo drastic changes in the near future. Understanding the nuances of this CTD code will be imperative to understanding the link between transcription and cotranscriptional events and to perhaps eventually unlock the therapeutic potential of the CTD.

### 4. The CTD Kinases

The specific phosphorylation events within the heptad repeat are mediated through the activity of a transcription-associated subset of cyclin-dependent kinases (CDKs) known as the CTD kinases. Unlike their cell cycle counterparts, the CTD kinases form complexes with members of the noncycling "transcription cyclin" family and are active throughout the cell cycle. Nevertheless, CTD kinase activity is tightly regulated through a variety of mechanisms, including selective recruitment, binding by kinase-associated factors, and sequestration by inhibitory factors. Although somewhat promiscuous *in vitro* (for example, the Ser2-specific CTD kinase Ctk1 can phosphorylate both Ser2 and Ser5 *in vitro* [16]), *in vivo* the CTD kinases are selective for particular heptad residues (Ser2, 5, and 7) and stages of transcription. Thus, the various CTD kinases are most conveniently presented in the context of a segmented transcription cycle; however, it should be made clear that in reality the various kinase activities lack the clearly delineated boundaries that such a presentation suggests.

Although we will not be discussing them in detail, it should be kept in mind that the CTD kinases function in conjunction with the more enigmatic CTD phosphatases

(several of which have been characterized in yeast, including Fcp1 [33], Ssu72 [34, 35], and Rtr1 [36]; see [7] for a review and [9] for further discussion).

**4.1. Initiation and the Promoter: CDK7 and CDK8.** The phosphorylation of CTD Ser5 and Ser7 residues during the formation of the preinitiation complex is mediated by the CTD kinase subunit of the general transcription factor TFIIF: Kin28/Ccl1 in yeast and CDK7/CyclinH in metazoa [11, 13, 37–39]. In an elegant interplay, the kinase activity of Kin28 is stimulated by the mediator coactivator complex, which binds to, and delivers, unphosphorylated RNAPII to the promoter [40]. The resulting phosphorylation of the CTD leads to the dissociation of mediator [41]; thus after fulfilling its function, mediator is able to use the CTD and Kin28 to induce its own release from transcriptionally active RNAPII. Intriguingly, a subpopulation of the mediator complex has been found to include an extra module that contains the CDK8/CyclinC (Srb10/Srb11 in yeast) kinase/cyclin pair. Initially discovered as a suppressor of a CTD truncation [42], CDK8 has emerged as the only CTD kinase to be implicated in the repression of transcription. Concordant with the process of mediator release, one of the mechanisms by which CDK8 has been proposed to execute its inhibitory activity is through the premature phosphorylation of the CTD (prior to PIC formation) and inactivation of the CDK7/CyclinH complex [10, 43]. Although historically the focus has been on CDK8's role as negative regulator of transcription, increasing numbers of studies are finding that CDK8 can also play a positive role in transcriptional activation. Therefore, a complete understanding of the function of CDK8 in CTD phosphorylation and transcription remains elusive and is likely to be a topic of much research in the near future (for more details and a comprehensive review, see [7, 44]).

Once thought to be essential for promoter clearance, the activity of Kin28 has been shown to be dispensable for global gene transcription [45, 46]. Despite this, Kin28 has been found to enhance polymerase progression through long genes in yeast (over 2 kb), suggesting that it plays a role in transcriptional elongation or in the inhibition of premature termination [19]. CDK7 also takes part in the phosphorylation and activation of other CDKs (see [47] for a review); however, its (and Ser5Ps) most clearly defined transcriptional role is the recruitment of the 5' end capping machinery. Not only does this ensure the proper processing of the nascent mRNA, it has also been shown to mediate the recruitment of the Ser2 CTD kinases in some organisms (either directly or through recruitment of the capping machinery) [48–50]; this suggests that phosphorylation of Ser5 plays a role in triggering the onset of Ser2 phosphorylation.

**4.2. The Elongation Phase: Ctk1, Bur1, and Their Metazoan Counterparts.** Subsequent to Kin28 activity at the promoter, phosphorylation of Ser2 of the CTD heptad occurs downstream of the transcription start site (TSS) and coincides with RNAPII entry into productive elongation. Coupled with the activity of Ser5P-specific phosphatases (Rtr1 in yeast [36]), this leads to a transition from high Ser5P to high Ser2P (as

characterized by ChIP). The Ser5P to Ser2P crossover point, defined as the point at which the ChIP signals for the two CTD marks cross, is on average ~450 bp downstream of the TSS and appears to be independent of the overall gene length [19]. This implies that the dynamics of Ser2 and Ser5 phosphorylation are not scaled to gene length; however, the significance of the crossover point in terms of the actual phosphorylation state of the CTD is obscure.

In *Saccharomyces cerevisiae* (*Sc*), the phosphorylation of Ser2 is primarily mediated by CTDK-I, a three subunit enzyme (consisting of Ctk1, a CDK homologue; Ctk2, a cyclin homologue; and Ctk3, whose function is unknown) [51, 52]. Although it is responsible for the bulk of Ser2 phosphorylation *in vivo*, Ctk1 is not essential for viability or for transcriptional elongation. The CTD kinase activity of Ctk1 has been linked to several transcription-associated processes, including the recruitment of the Set2 histone methyltransferase [53, 54], 3' end processing [19, 55, 56], and termination (reviewed in [9]). In addition to its role as a CTD kinase, Ctk1 (independent of its kinase activity) has also been shown to be involved in the dissociation of basal transcription factors from RNAPII [57]. Despite having a principal role in CTD Ser2 phosphorylation, Ctk1 is not the only Ser2 kinase in yeast; it coexists with the essential Bur1 kinase (which consists of the CDK homologue Bur1 and the cyclin Bur2) [58]. While it has been proposed that Bur1's primary transcription-related substrate is the elongation factor Spt4/5 [59, 60], rather than the CTD, recent evidence indicates that Bur1 binds to the Ser5P CTD and contributes to Ser2 phosphorylation during early elongation, possibly stimulating subsequent Ctk1 activity [49, 61]. As mentioned previously, Bur1 has also been found to exhibit an elongation phase Ser7 kinase activity, which appears to counteract the activity of a yet unidentified Ser7 phosphatase [12]. Another pair of Ser2 elongation kinases is also present in the fission yeast *Saccharomyces pombe* (*Sp*), where Lsk1, the *Sp* orthologue of Ctk1, has been shown to be responsible for the bulk of Ser2 phosphorylation, while *Sp* CDK9, the *Sp* orthologue of Bur1, is able to phosphorylate both the CTD and Spt5 [50, 62].

Until recently, higher eukaryotes appeared to have only one Ser2 CTD kinase: P-TEFb (which is composed of CDK9 and cyclinT). P-TEFb is able to phosphorylate both the Ser2 position of the CTD and the elongation factor Spt5 and is essential for transcriptional elongation [63–65] (for detailed discussions, see [66–68]). The substrate specificity of P-TEFb, coupled with its equal sequence similarity to both Bur1 and Ctk1, has led to the proposal that P-TEFb reconstitutes the activities of both yeast kinases in higher eukaryotes [69]. However, there was some evidence that this may not be the case; two evolutionary studies concluded that while *Sc* Bur1 is the closest *Sc* relative of metazoan CDK9 proteins, *Sc* Ctk1 is actually more closely related to another set of relatively little-studied metazoan CDK proteins [70, 71]. Based on these evolutionary studies, work in our lab has characterized the previously unstudied *Drosophila* CDK12 (dCDK12) and little-studied human CDK12 (hCDK12) as elongation phase CTD kinases and the metazoan orthologues of yeast Ctk1 [72]. Unlike most

other cell cycle and transcriptional CDKs, CDK12 (and CDK13, a highly related paralogue absent in *Drosophila* but present in many “higher” organisms) contains splicing factor-related structural features (RS domains) and has been previously implicated in the regulation of alternative pre-mRNA splicing [73–77]. Although both CDK12 and CDK13 manifest CTD kinase activity in *in vitro* kinase assays, only CDK12 seems to have an effect on global CTD phosphorylation *in vivo*; thus CDK13’s role in transcription (assuming one exists) remains elusive [72]. In terms of the cyclin partner of CDK12, our lab found that endogenous dCDK12 associates with cyclinK, a Ctk2-like cyclin that has been previously characterized as an alternative partner for CDK9 [78]. These findings are inconsistent with previous reports that CDK12 and CDK13 interact with the L class cyclins [74, 75]; thus, whether cyclinK is the cyclin partner of human CDK12 and CDK13 remains to be determined. As of this paper, other than our initial characterization, there have been no published studies of CDK12 and CDK13 in the context of transcription and transcriptional elongation, thus much remains to be learned about these kinases. Intriguingly, the depletion of dCDK12 affects the phosphorylation state of the CTD without affecting RNAPII occupancy (BB and ALG, unpublished); therefore, CDK12 might prove to be a useful tool for studying the links between CTD phosphorylation patterns and transcription elongation-associated processes in higher eukaryotes.

A final point regarding the CTD kinases relates to their therapeutic potential. As the CTD kinases are involved in the coupling of various signaling pathways to transcription and RNA processing events, they play important roles in the regulation of cell growth, proliferation, and survival. Thus, targeting these kinase activities may be potentially useful for the treatment of human diseases and cancer [79]. In fact, the CDK inhibitor flavopiridol, which targets P-TEFb, is used for the treatment of some forms of leukemia, and P-TEFb has been implicated in HIV replication [80, 81]. The emerging links between the CTD and DNA repair/genomic stability (see below), combined with the fact that ~15% of disease-causing mutations are a consequence of the misregulation of alternative splicing [82] (a function associated with both the CTD and CDK12/13, see below), suggest that a more comprehensive understanding of the CTD and its kinases could have broad medical implications in the future.

## 5. CTD Functions

The CTD has been implicated in a broad spectrum of transcription-associated functions, and its collection of binding partners has continued to expand over the last few years. Important target processes include mRNA (and snRNA) capping, splicing, 3’ end processing, termination, and more recently nuclear export (discussed below). In terms of non-RNA processing-associated events, the CTD has been shown to play roles in transcriptional activation, cotranscriptional chromatin modification, chromatin remodeling, and genome stability. Of course, the analysis of the CTD’s role in each specific function is very challenging, as many of the

CTD-mediated transcriptional processes are interlinked. For example, capping has been shown to influence both the splicing of the first intron and 3’ end processing [83–87]; splicing of the last intron affects 3’ end processing and vice versa [88–90]; and alternative splice site choices are affected by the cotranscriptional histone modifications at splice site junctions [91–93]. While these interactions demonstrate the high degree of coordination involved in mRNA synthesis, they unfortunately complicate the interpretation of functional studies. In a broad sense, it is fair to state that exactly how S2, S5, and S7 phosphorylation affect initiation, elongation, and termination remains poorly understood. Despite the lack of a universal understanding, particular aspects of CTD function have been characterized to an impressive level of detail, and many of the more enigmatic functions are becoming better understood through continued investigation. In the next few sections, we present a rather cursory overview of the functions of the CTD (for other reviews, see [2, 7, 8]) before moving on to a discussion of two newly emergent CTD-related functions: mRNA export and DNA repair/genomic stability.

**5.1. 5’ Capping.** As mentioned previously, one of the most clearly recognized functions of the CTD is its involvement in the 5’ end capping of mRNA through the recruitment of the capping machinery. The modification of the 5’ end of the RNA with the 5’ 7-methyl guanosine cap is unique to RNAPII transcripts and occurs just after the transcript clears the polymerases exit channel [94, 95]. Transcripts made by a CTD-less RNAPII were found to be inefficiently capped, leading to the characterization of the physical interaction between the capping enzymes and the phospho-CTD [96–98]. Subsequent studies showed that the capping enzyme associates with the 5’ end of genes *in vivo*, which correlates with the enzyme’s function, and that this association is dependent on phosphorylation of Ser5 of the CTD [11, 37, 46]. In addition to raising the local concentration of the capping enzyme near the exit channel, the CTD has also been shown to stimulate its activity. An example of a phosphorylation state-specific function, mammalian guanylyltransferase was found to bind to both Ser2P and Ser5P synthetic heptad repeats but was allosterically activated only by Ser5P [99]. The interaction between the 5’ capping machinery and the CTD has also been investigated structurally, resulting in some interesting insights. The crystal structure of the *Candida albicans* guanylyltransferase Cgt1 complexed with a synthetic Ser5P four heptad repeat peptide revealed that the CTD binds within an extended docking site on the enzymes surface using two nonadjacent heptads, and a full-heptad repeat was looped out away from the interaction site [100]. This looping not only demonstrates the inherent flexibility of the CTD but also suggests that by binding two remote heptads, CTD binding factors may be able to loop out large portions of the CTD. This looping could potentially result in the formation of novel structural motifs, which could in turn serve as binding sites for other CTD binding factors, leading to organized, sequential binding [7]. Whether this actually occurs is still an open question; however, recent

studies have shown that the binding of some well-known 3' end processing factors to the CTD appears to be cooperative in nature [101].

**5.2. 3' End Processing.** Another well-recognized function of the CTD is its role in 3' end processing and termination (for reviews, see [102, 103]). Analogous to capping, transcription by a CTD-less RNAPII was shown to affect both processes, and cleavage and polyadenylation factors were found to bind to the phospho-CTD [104–110]. Accordingly, inhibition of Ctk1 in yeast has been shown to decrease the efficiency of cleavage at poly(A) sites [56] and result in the disruption of polyadenylation factor recruitment to the 3' end of the gene [55]. A genome-wide analysis has also shown that depletion of Ctk1 using a tetracycline-repressible degron mutant causes a “pileup” of polymerases at the poly(A) site in a subset of genes with good consensus poly(A) sequences [19]. This increase in RNAPII occupancy suggests that improper CTD phosphorylation at these sites can result in a strong transcriptional pause that is perhaps due to the rate-limiting recruitment of a specific factor. Strongly linked to polyadenylation, termination has also been reported to be affected by the CTD [108]; in addition, Rtt103, a component of the termination complex, has been shown to bind Ser2P CTD [111]. However, the role of CTD modification in termination is not yet well understood (see [9, 103, 112] for further discussion). Intriguingly, it has been observed that the recruitment of one well-recognized CTD binding 3' end processing factor, Pcf11 (which preferentially binds to Ser2 phosphorylated CTD), does not directly correlate with the level of Ser2 CTD phosphorylation [20, 55, 106]; analysis by ChIP indicates that Pcf11 is recruited mainly at the poly(A) site, while Ser2P levels rise throughout the coding region. Potential explanations for this phenomenon highlight some of the interesting complications surrounding phospho-CTD factor recruitment. Perhaps Pcf11 requires a certain threshold of Ser2P or both Ser2P and an external signal, such as the presence of the newly synthesized polyadenylation site, the unmasking of particular CTD epitopes, or the presence additional factors (Pcf11 CTD binding was recently shown to be cooperative [101]). Other considerations include the remodeling of the CTD via a pattern-specific change (such as the formation of Ser2P only heptads) or modifications that are undetectable by ChIP. With regards to the latter and the previous discussion on proline isomerization, Pcf11 was reported to specifically recognize three neighboring *trans* prolines within a mixed population of *cis-trans* isomers [113]. It is likely that the recruitment of many other CTD-binding factors is also mediated through multiple mechanisms and dependent on the satisfaction of particular sets of conditions.

**5.3. snRNA Processing.** One of the relatively more recently discovered functions of the CTD is its role in the transcription and 3' end processing of snRNAs [114, 115]. Even though they are transcribed by RNAPII, snRNAs are unlike most coding transcripts; they do not undergo splicing or polyadenylation and instead rely on a conserved 3' box RNA

processing element downstream of the coding region for proper 3' end processing and termination [116]. The 3' end processing of snRNAs has received a lot of recent attention, as it is currently the only specific function attributed to the phosphorylation of Ser7 of the heptad repeat [32]. Seemingly dispensable for viability and expression of protein-coding genes, Ser7 has been found to be essential for endogenous snRNA gene expression. This requirement for Ser7 phosphorylation was subsequently linked to the integrator complex [32], a large CTD-associated multiprotein complex involved in snRNA 3' end processing [117]. Further characterization of the CTD-integrator interaction not only revealed that both Ser2P and Ser7P were required but also that efficient binding required a specific arrangement of the modifications [118]. Screening of synthetic diheptad repeats revealed that although maximal binding was achieved with both Ser2 and Ser7 phosphorylated repeats, the minimal interaction domain consisted of a Ser7P on the first heptad followed by a Ser2P on the second heptad; any other combination of two phosphates was insufficient for integrator binding [118]. These findings lend further support for the pattern-specific binding of CTD-associated factors and, coupled with the fact that integrator is specific to snRNA genes, also suggest that the appropriate Ser2P/Ser7P patterns may be snRNA gene specific (although many other discriminatory mechanisms could be at play, such as the previously discussed methylation of R1810).

**5.4. Histone Modifications.** Tying together two important aspects of gene expression and transcriptional coordination, the CTD has also been shown to be involved in the co-transcriptional modification of histones and remodeling of chromatin structure. Although a full discussion of histone modifications and the histone code hypothesis is beyond the scope of this paper, the significance of these processes cannot be understated: they play integral functions in almost every aspect of gene expression and regulation (for a review, see [119]). Here, we will only give a very brief overview of some CTD-related functions and a short list of new developments.

Providing the first clear link between the CTD and histone modification, the yeast Set1 and Set2 methyltransferases were found to be recruited to actively transcribed genes at specific stages of the transcription cycle (5' end versus interior of the gene, resp.) through interactions with the Ser5P (via the PAF complex for Set1) and Ser2/5P CTD [120–123] (see [124] for a review). Thus, Set1's activity, the methylation of histone H3 at the K4 position, peaks near the promoter, while Set2's methylation of H3 at K36 occurs downstream in the coding region. The H3K36 trimethylation mark can be used to identify transcriptionally active genes and has been shown to suppress inappropriate transcription from cryptic promoters, which was initially thought to occur through the recruitment of the histone deacetylase Rpd3S [125, 126]. Surprisingly, recent studies have found that Rpd3S is actually recruited via direct binding to the phospho-CTD; although H3K36 trimethylation is dispensable for Rpd3S recruitment, it appears to be required for activation

of its deacetylation activity [127, 128]. Intriguingly, Set2 is capable of H3K36 dimethylation independent of its CTD-interacting SRI domain or Ctk1 [54], implying that only one specific aspect (H3K36 trimethylation) of Set2's activity is regulated by interaction with the phospho-CTD. Another interesting chromatin-related CTD binding factor is the histone H3 chaperone and transcription elongation factor Spt6 [129]. Spt6 has been shown to bind the Ser2P CTD through a tandem SH2 domain [130] and interact with the multifunctional elongation factor IWS1/Spn1, an Spt6-interacting factor that associates with the nuclear RNA export factor REF1/Aly (and Yra1 in yeast (ALM and ALG, unpublished)) and possibly facilitates nuclear export [129]. In addition to its association with REF1/Aly, IWS1/Spn1 has been recently found to be required for the optimal loading of the mammalian Set2 (HYPB/Setd2) in the coding regions of several genes [131], linking nucleosome reassembly with elongation-coupled H3K36 trimethylation *in vivo*. Several other recently characterized chromatin-related CTD interactions include the recruitment of the chromatin-remodeling factor CHD8 [132] and the FACT histone chaperone via HP1 [133]. The number of recognized chromatin-associated CTD interacting factors is likely to grow rapidly over the next few years as our understanding of both the processes of, and the relationship between, CTD phosphorylation and histone modification improves.

**5.5. Splicing.** One of the more intriguing but relatively poorly characterized functions of the CTD is its involvement in cotranscriptional splicing (for reviews, see [134, 135]). Although neither active transcription nor the CTD is absolutely required for splicing (presynthesized mRNAs can be spliced by injection into *Xenopus* oocytes or by incubation with nuclear extracts [136]), experimental evidence accumulated over the last three decades implicates the CTD as a key player in the coupling between the two processes. The link between the CTD and splicing was first proposed in the early 90s [137], and in the mid-to-late 90s, the hyperphosphorylated RNAPII was shown to associate with the SR (Serine/Arginine rich) family of splicing factors and with components of the splicing machinery [138–140]. Concordantly transcription by a CTD-less RNAPII was shown to result in low splicing efficiency *in vivo* [108], and the addition of an anti-CTD antibody or exogenous expression of CTD peptides resulted in the accumulation of unspliced transcripts [140] and the nuclear reorganization of splicing factors [141]. In what could be considered the reciprocal experiment, it was also shown that isolated CTD fragments and purified phosphorylated RNAPII were able to activate splicing reactions *in vitro* [142, 143]. In addition to the SR-like CTD-associated factors (SCAFs) [140, 144], several other CTD-binding splicing factors have been identified, including the yeast U1 snRNP component Prp40 [145] and the mammalian splicing factors CA150 (TCERG1) [146], PSF, and p54/NRB [147]. A recent addition has been the splicing factor U2AF, which in an *in vitro* complementation assay was shown to be recruited to the CTD in complex with another spliceosome component, Prp19, in order to overcome a weak

polypyrimidine-binding tract in an IgMA3 substrate [148]. Although alternative splicing has been directly demonstrated to be affected by the presence of the CTD [149], the fact that RNAPII elongation rate has also been implicated in splice site choice [150, 151] has made it difficult to determine how much of the effect is due to the specific recruitment of splicing factors to the CTD and how much is due to changes in the elongation rate (kinetic coupling) or other processes. One study that merits specific mention is that of Batsche et al., who reported that the inclusion of a set of alternative exons in the middle of the CD44 gene was dependent on the site-specific accumulation of RNAPII occupancy induced by the catalytic subunit of the SWI/SNF chromatin-remodeling complex [152]. Surprisingly, this “stalled” RNAPII exhibited a SWI/SNF-dependent switch of the RNAPII phosphorylation status from the elongation characteristic Ser2P to the more promoter characteristic Ser5P, which perhaps creates a barrier to further elongation (for a detailed discussion regarding kinetic coupling please, see [134]). In addition to its general role in the recruitment of splicing factors, it has also been suggested that the CTD may function as a molecular tether for distant splice sites within the mRNA. This would be accomplished through the binding of the nascent 3' splice site to a CTD-associated splicing factor, thus effectively immobilizing it near the polymerase mRNA exit channel in anticipation of the cognate 5' splice site. Enhancement of the local concentration of the splice sites would then dramatically increase the efficacy of the splicing reaction (see [135]). This tethering model is especially attractive in higher eukaryotes in which the intron lengths of many genes exceed thousands of base pairs. Although studies have demonstrated that mRNA tethering is likely [153], it has yet to be directly linked to the CTD and its phosphorylation.

Although it is not the only CDK kinase to have been implicated in the regulation of both transcription and alternative splicing (CDK11/cyclinL has been shown to affect both processes but lacks a reported CTD kinase activity [154, 155]), CDK12's CTD kinase activity coupled with its structural and functional characteristics places it directly at the juncture between transcription, the CTD, and splicing. The N-terminal domains of both CDK12 and its more enigmatic paralogue CDK13 contain arginine/serine (RS) dipeptide-rich segments, which are characteristic of splicing factors and splicing factor regulators and are believed to be important for protein-protein interactions [156, 157]. Much like other RS domain-containing proteins and most factors involved in pre-mRNA splicing, CDK12 and CDK13 were found to exhibit a punctate pattern of localization in the nucleus, superimposed on a more even distribution throughout the nucleoplasm. The punctate localization points appear to represent “nuclear speckles” [76, 77], commonly thought to be sites of splicing factor storage [158]. In accordance with these structural features and localization, *in vivo* splicing assays using reporter genes have demonstrated that the ectopic overexpression or depletion of CDK12 and CDK13 modulates alternative splice site selection [74, 75, 77]. In addition, CDK13 was observed to affect HIV splicing in a Tat-dependent manner [73]. These splicing effects



were postulated to be a consequence of CDK13-mediated phosphorylation of the canonical SR splicing factor ASF/SF2 [76], but all of the extant studies utilized ectopic overexpression of CDK12 and CDK13, usually without their potential cyclin partners, in assay systems that can detect splice site changes but are unable to provide mechanistic insights. Moreover, many of the results appear to argue against the direct phosphorylation of specific splicing factors by CDK12 and 13; for example, the N-termini of CDK12 and 13 are able to affect splicing independently of the kinase domains; and one study has reported that the phosphorylation of ASF/SF2 by CDK13 appears to be indirect [74]. Complicating analyses of this type is the fact that the overexpression of virtually any SR protein will have effects on splicing via sequestration of other SR proteins (and SR protein-binding partners), competition for SR phosphorylating and dephosphorylating factors, and the occurrence of other nonspecific and unforeseen events; thus, it is difficult to amalgamate the current data into an overall consensus. Despite these caveats, the unusual structures of CDK12 and CDK13, coupled with their ability to modulate splice site choices in *in vitro* assays, make it tempting to speculate that the two kinases may serve as a central link between the processes of transcription, CTD phosphorylation, and splicing; however, whether such a link exists is still undetermined.

**5.6. The CTD and mRNA Export.** While the connection between mRNA processing and the CTD has been established, recent studies have begun to investigate CTD involvement with the last step in mRNA production: formation of an export competent messenger ribonucleoprotein particle (mRNP). While this area of research has yet to mature, it is known that mRNAs are exported through a Ran-GTP-independent pathway that involves a specific set of conserved export receptor and adaptor proteins. There appears to be one universal receptor for mRNA export, Mex67/Mtr2 in yeast and TAP/p15 in mammals, which interacts with the mature mRNP and the nuclear pore complex (NPC) to facilitate export. The export receptor functions in conjunction with the export adaptor proteins, which cotranscriptionally associate with the nascent mRNA. There are two main mRNA export adaptor proteins, Yra1/ALY in the transcription and mRNA export (TREX) complex and Sac3 in the TREX-2 complex. While export mediated by TREX-2 and Sac3 has been shown to be coupled to chromatin modification through Sus1, a common factor in both TREX-2 and the Spt-Ada-Gcn5 acetyltransferase (SAGA) complex (reviewed in [159, 160]), we will focus on the novel connection between mRNA export by Yra1 and the CTD.

TREX was the first characterized transcription export complex and consists of the THO complex of elongation and hyperrecombination-related factors, including Hpr1, Tho2, Thp1, and Mft1; the export adaptor protein Yra1/ALY; the RNA helicase Sub2/UAP56; a protein of unknown function, Tex1 [161]. Yra1/ALY has been proposed to link aspects of mRNA splicing and processing to mRNA export based on its interactions with Sub2/UAP56 [162, 163] and Pcf11 [164] and the observation that ALY may be associated with

the exon junction complex in mammals [165–167]. The Hpr1 subunit of THO has been hypothesized to play an important role in Yra1 and Sub2 recruitment because  $\Delta hpr1$  yeast displays a decrease in the levels of both Sub2 and Yra1 occupancy on certain genes [168]; however, as deletion of *HPR1* has been shown to affect transcription elongation, this decrease in Yra1 and Sub2 occupancy may be due to an elongation defect rather than a direct effect of Hpr1 on the export proteins [169].

In a previous study from our lab, a proteomics screen in yeast identified Yra1 as a putative phospho-CTD-associated protein (PCAP) [170]. We have since characterized the phospho-CTD- (PCTD-) binding activity of this export adaptor and demonstrated (*via* ChIP) that partial deletion of its phospho-CTD interaction domain (PCID) leads to a near loss of Yra1 association with active genes (MacKellar and Greenleaf, In press (*J. Biol. Chem.*)). We therefore think that either Yra1 is responsible for recruiting the rest of TREX to active genes via the CTD or Yra1 is recruited to active genes independently of THO. Further chromatin immunoprecipitation studies using additional Yra1 variants that do not bind the CTD are needed to further clarify the role of the CTD in Yra1 function and the dependence of the rest of TREX on Yra1.

We also found that the PCID of Yra1 contains an RNA recognition motif (RRM). While RRMs are known to be versatile-binding domains that mediate both protein-nucleic acid and protein-protein interactions (see [171, 172] and references therein), until this year no protein had been found to use an RRM for PCTD binding. However, a recent study on spliceosomal factor U2AF65 indicated that its noncanonical RRM (also called a U2A homology motif (UHM)) might also mediate an interaction between the splicing factor and the PCTD [148]. We propose that there is a class of factors that use its RRM as a dual-purpose domain first for associating with the transcription elongation complex *via* PCTD binding and then for associating with the nascent transcript through RNA binding. Structural studies on Yra1 in complex with the PCTD are necessary to examine the role and binding mode of the RRM in this interaction.

Based on our discovery that the CTD is involved in recruiting Yra1 to genes, we have revised the model of Yra1 cotranscriptional recruitment to include the CTD (Figure 2). We propose that Yra1 (possibly in complex with THO and Sub2) is recruited to active genes during elongation when the CTD is doubly phosphorylated on Ser2 and Ser5. Because we do not yet know whether Yra1 can bind both the CTD and TREX, it is modeled in both contexts. As the nascent mRNA grows, Yra1 remains bound to the CTD, and this may prevent Yra1 from blocking the activities of other mRNA processing factors, averting premature export. At the 3' end of the gene, Yra1 dissociates from the CTD, because either the phosphorylation pattern changes to predominantly Ser2 phosphorylation (which Yra1 does not bind) or Yra1 binds the nascent mRNA with higher affinity. This model is still largely speculative, and further experimentation is required to test the hypotheses it represents.

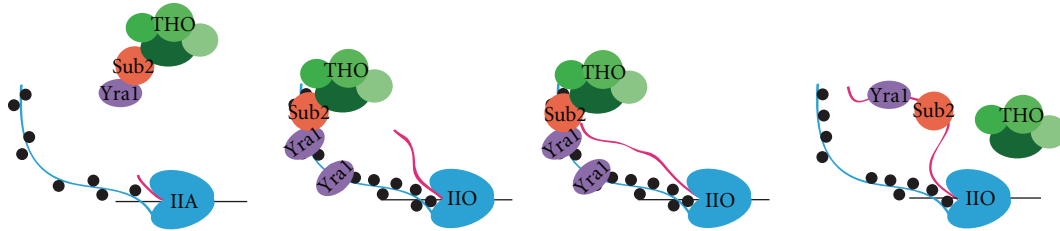


FIGURE 2: CTD-mediated recruitment of Yra1. Yra1 (possibly in complex with THO and Sub2) does not bind to Ser5P containing CTD and is not observed at high level at the 5' end of genes. Yra1 is recruited to active genes during elongation when the CTD is doubly phosphorylated on Ser2 and Ser5. As it is not known whether Yra1 can bind both the CTD and TREX, it is modeled in both contexts. Yra1 remains bound to the CTD until the nascent mRNA reaches a certain length or reaches the CTD-bound Yra1; this would avert premature export attempts. At the 3' end of the gene, Yra1 dissociates from the CTD, either because the phosphorylation pattern changes to predominant Ser2 phosphorylation (in which Yra1 does not bind) or because Yra1 binds the nascent mRNA with higher affinity.

**5.7. The Next Frontier in CTD Research: DNA Damage and Repair.** Results connecting transcription/RNA processing with recombination and DNA damage repair have been obtained in bacteria, yeast, and mammals (reviewed in [173]). For example, defects in mRNA splicing (by ASF/SF2 in mammals) and packaging (THO/TREX and TREX-2 in yeast) have been linked to genomic instability and hyper-recombination via R-loop formation [174–176]. A potential involvement of the CTD in repair/recombination is suggested by the observations that mutations in the subunits of the CTDK-I kinase render yeast sensitive to DNA-damaging agents and that DNA damage leads to alterations in the phosphorylation pattern of the CTD [177]. Moreover, recent work hints that the CTD and its associated proteins play a role in sensing DNA damage and promoting repair.

For example, Bennett and colleagues used the diploid yeast deletion strain collection to identify a large number of genes whose homozygous deletion leads to ionizing radiation (IR) sensitivity [178, 179]. We have found that a significant number of these genes encode phospho-CTD-associating proteins (PCAPs), thus linking IR damage repair to the phospho-CTD (Winsor et al., *in prep*). A different kind of damage repair, as signaled by mitotic recombination, also appears PCTD linked since diploid yeast deleted for *CTK1* (*ctk1Δ/ctk1Δ* strains) displays reduced rates of spontaneous mitotic recombination at several loci (Winsor et al., *in prep*). Thus, the PCTD appears to be involved in processes that maintain genome stability.

In a related vein, it was shown recently that mammalian RecQ5 protein, a putative “antirecombinase,” is associated with RNAPII on active genes [180, 181]. *In vitro* experiments showed that RecQ5 binds directly to the elongation-associated phospho-CTD *via* a Set2 Rpb1-interaction (SRI) domain; moreover, deletion of the SRI domain resulted in loss of RecQ5 protein at multiple loci [182]. While the function of RecQ5 in RNAPII elongation complexes is not yet known, we favor a model in which it remains poised on the PCTD, ready to act if the polymerase encounters a situation that might induce inappropriate transcription-linked recombination (e.g., [174–179]). It will be extremely informative to analyze the involvement of transcribing RNAPII, and proteins associated with its PCTD, in repair/recombination events that contribute to genome stability.

## 6. Conclusions

After 25 years of research, much is still unknown about the CTD of RNAPII and its role in coordinating a surprising number of nuclear events with transcription. In addition to transcript elongation (RNAPII movement along the template), mRNA processing, and chromatin modification, the collection of CTD-interacting processes is now thought to also include mRNA export and DNA repair; future investigations into the links between these events and the CTD should be remarkably informative. New information on CTD phosphorylation patterns, which modulate its interactions with nuclear factors, has recently been generated through genome-wide ChIP experiments, and multiple new insights into global CTD phosphorylation have emerged. On the other hand, the complexity and nuance of the patterns of post-translational modifications are such that the actual distribution of phosphate groups along the CTD is not known for even one transcription elongation complex. Thus, while the inventory of CTD-modifying enzymes continues to expand, it is clear that we have much to learn about the ways in which they collaborate to produce modification patterns as found *in vivo*. Even so, the list of proteins known to interact with a specifically modified form of the CTD also continues to grow, expanding our knowledge of the roles played by the CTD in coordinating transcription-related processes. As our knowledge of CTD modifications and interactions expands and becomes more refined, our understanding of the “phospho-CTD cycle” and the manner in which the PCTD orchestrates the numerous events connected to the process of DNA-dependent RNA synthesis will continue to evolve.

## Acknowledgments

This work was supported by Bridging Funds from the Duke University School of Medicine and by NIH Grant no. GM040505.

## Author's Contribution

B. Bartkowiak and A. L. MacKellar contributed equally to this work.

## References

- [1] E. H. Fischer and E. G. Krebs, "Conversion of phosphorylase b to phosphorylase a in muscle extracts," *Journal of Biological Chemistry*, vol. 216, no. 1, pp. 121–132, 1955.
- [2] H. P. Phatnani and A. L. Greenleaf, "Phosphorylation and functions of the RNA polymerase II CTD," *Genes and Development*, vol. 20, no. 21, pp. 2922–2936, 2006.
- [3] L. A. Allison, J. K. Wong, V. D. Fitzpatrick, M. Moyle, and C. J. Ingles, "The C-terminal domain of the largest subunit of RNA polymerase II of *Saccharomyces cerevisiae*, *Drosophila melanogaster*, and mammals: a conserved structure with an essential function," *Molecular and Cellular Biology*, vol. 8, no. 1, pp. 321–329, 1988.
- [4] J. L. Corden, "Tails of RNA polymerase II," *Trends in Biochemical Sciences*, vol. 15, no. 10, pp. 383–387, 1990.
- [5] M. Nonet, D. Sweetser, and R. A. Young, "Functional redundancy and structural polymorphism in the large subunit of RNA polymerase II," *Cell*, vol. 50, no. 6, pp. 909–915, 1987.
- [6] W. A. Zehring, J. M. Lee, J. R. Weeks, R. S. Jokerst, and A. L. Greenleaf, "The C-terminal repeat domain of RNA polymerase II largest subunit is essential in vivo but is not required for accurate transcription initiation in vitro," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 85, no. 11, pp. 3698–3702, 1988.
- [7] A. Meinhart, T. Kamenski, S. Hoepfner, S. Baumli, and P. Cramer, "A structural perspective of CTD function," *Genes and Development*, vol. 19, no. 12, pp. 1401–1415, 2005.
- [8] S. Egloff and S. Murphy, "Cracking the RNA polymerase II CTD code," *Trends in Genetics*, vol. 24, no. 6, pp. 280–288, 2008.
- [9] S. Buratowski, "Progression through the RNA Polymerase II CTD Cycle," *Molecular Cell*, vol. 36, no. 4, pp. 541–546, 2009.
- [10] C. J. Hengartner, V. E. Myer, S. M. Liao, C. J. Wilson, S. S. Koh, and R. A. Young, "Temporal regulation of RNA polymerase II by Srb10 and Kin28 cyclin-dependent kinases," *Molecular Cell*, vol. 2, no. 1, pp. 43–53, 1998.
- [11] P. Komarnitsky, E. J. Cho, and S. Buratowski, "Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription," *Genes and Development*, vol. 14, no. 19, pp. 2452–2460, 2000.
- [12] R. Tietjen, D. W. Zhang, J. B. Rodriguez-Molina et al., "Chemical-genomic dissection of the CTD code," *Nature Structural and Molecular Biology*, vol. 17, no. 9, pp. 1154–1161, 2010.
- [13] M. Kim, H. Suh, E. J. Cho, and S. Buratowski, "Phosphorylation of the yeast Rpb1 C-terminal domain at serines 2, 5, and 7," *Journal of Biological Chemistry*, vol. 284, no. 39, pp. 26421–26426, 2009.
- [14] M. Patturajan, R. J. Schulte, B. M. Sefton et al., "Growth-related changes in phosphorylation of yeast RNA polymerase II," *Journal of Biological Chemistry*, vol. 273, no. 8, pp. 4689–4694, 1998.
- [15] R. D. Chapman, M. Heidemann, T. K. Albert et al., "Transcribing RNA polymerase II is phosphorylated at CTD residue serine-7," *Science*, vol. 318, no. 5857, pp. 1780–1782, 2007.
- [16] J. C. Jones, H. P. Phatnani, T. A. Haystead, J. A. MacDonald, S. M. Alam, and A. L. Greenleaf, "C-terminal repeat domain kinase I phosphorylates Ser2 and Ser5 of RNA polymerase II C-terminal domain repeats," *Journal of Biological Chemistry*, vol. 279, no. 24, pp. 24957–24964, 2004.
- [17] P. Liu, J. M. Kenney, J. W. Stiller, and A. L. Greenleaf, "Genetic organization, length conservation, and evolution of RNA polymerase II carboxyl-terminal domain," *Molecular Biology and Evolution*, vol. 27, no. 11, pp. 2628–2641, 2010.
- [18] J. W. Stiller and M. S. Cook, "Functional unit of the RNA polymerase II C-terminal domain lies within heptapeptide pairs," *Eukaryotic Cell*, vol. 3, no. 3, pp. 735–740, 2004.
- [19] H. Kim, B. Erickson, W. Luo et al., "Gene-specific RNA polymerase II phosphorylation and the CTD code," *Nature Structural and Molecular Biology*, vol. 17, no. 10, pp. 1279–1286, 2010.
- [20] A. Mayer, M. Lidschreiber, M. Siebert, K. Leike, J. Söding, and P. Cramer, "Uniform transitions of the general RNA polymerase II transcription complex," *Nature Structural and Molecular Biology*, vol. 17, no. 10, pp. 1272–1278, 2010.
- [21] R. Baskaran, M. E. Dahmus, and J. Y. J. Wang, "Tyrosine phosphorylation of mammalian RNA polymerase II carboxyl-terminal domain," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 90, no. 23, pp. 11167–11171, 1993.
- [22] J. Zhang and J. L. Corden, "Identification of phosphorylation sites in the repetitive carboxyl-terminal domain of the mouse RNA polymerase II largest subunit," *Journal of Biological Chemistry*, vol. 266, no. 4, pp. 2290–2296, 1991.
- [23] W. G. Kelly, M. E. Dahmus, and G. W. Hart, "RNA polymerase II is a glycoprotein. Modification of the COOH-terminal domain by O-GlcNAc," *Journal of Biological Chemistry*, vol. 268, no. 14, pp. 10416–10424, 1993.
- [24] S. Krishnamurthy, M. A. Ghazy, C. Moore, and M. Hampsey, "Functional interaction of the Ess1 prolyl isomerase with components of the RNA polymerase II initiation and termination machineries," *Molecular and Cellular Biology*, vol. 29, no. 11, pp. 2925–2934, 2009.
- [25] Y. X. Xu and J. L. Manley, "Pin1 modulates RNA polymerase II activity during the transcription cycle," *Genes and Development*, vol. 21, no. 22, pp. 2950–2962, 2007.
- [26] C. B. Wilcox, A. Rossetini, and S. D. Hanes, "Genetic interactions with C-terminal domain (CTD) kinases and the CTD of RNA Pol II suggest a role for ESS1 in transcription initiation and elongation in *Saccharomyces cerevisiae*," *Genetics*, vol. 167, no. 1, pp. 93–105, 2004.
- [27] J. W. Werner-Allen, C. J. Lee, P. Liu et al., "cis-Proline-mediated Ser(P)5 dephosphorylation by the RNA polymerase II C-terminal domain phosphatase Ssu72," *Journal of Biological Chemistry*, vol. 286, pp. 5717–5726, 2011.
- [28] K. Xiang, T. Nagaike, S. Xiang et al., "Crystal structure of the human symplekin-Ssu72-CTD phosphopeptide complex," *Nature*, vol. 467, no. 7316, pp. 729–733, 2010.
- [29] R. J. Sims III, L. A. Rojas, D. Beck et al., "The C-terminal domain of RNA polymerase II is modified by site-specific methylation," *Science*, vol. 332, pp. 99–103, 2011.
- [30] R. D. Chapman, M. Heidemann, C. Hintermair, and D. Eick, "Molecular evolution of the RNA polymerase II CTD," *Trends in Genetics*, vol. 24, no. 6, pp. 289–296, 2008.
- [31] N. Fong and D. L. Bentley, "Capping, splicing, and 3' processing are independently stimulated by RNA polymerase II: different functions for different segments of the CTD," *Genes and Development*, vol. 15, no. 14, pp. 1783–1795, 2001.
- [32] S. Egloff, D. O'Reilly, R. D. Chapman et al., "Serine-7 of the RNA polymerase II CTD is specifically required for snRNA gene expression," *Science*, vol. 318, no. 5857, pp. 1777–1779, 2007.
- [33] E. J. Cho, M. S. Kobor, M. Kim, J. Greenblatt, and S. Buratowski, "Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser 2 of the RNA polymerase II C-terminal

- domain," *Genes and Development*, vol. 15, no. 24, pp. 3319–3329, 2001.
- [34] S. Krishnamurthy, X. He, M. Reyes-Reyes, C. Moore, and M. Hampsey, "Ssu72 Is an RNA polymerase II CTD phosphatase," *Molecular Cell*, vol. 14, no. 3, pp. 387–394, 2004.
- [35] M. Reyes-Reyes and M. Hampsey, "Role for the Ssu72 C-terminal domain phosphatase in RNA polymerase II transcription elongation," *Molecular and Cellular Biology*, vol. 27, no. 3, pp. 926–936, 2007.
- [36] A. L. Mosley, S. G. Pattenden, M. Carey et al., "Rtr1 is a CTD phosphatase that regulates RNA polymerase II during the transition from serine 5 to serine 2 phosphorylation," *Molecular Cell*, vol. 34, no. 2, pp. 168–178, 2009.
- [37] S. C. Schroeder, B. Schwer, S. Shuman, and D. Bentley, "Dynamic association of capping enzymes with transcribing RNA polymerase II," *Genes and Development*, vol. 14, no. 19, pp. 2435–2440, 2000.
- [38] M. S. Akhtar, M. Heidemann, J. R. Tietjen et al., "TFIIH kinase places bivalent marks on the carboxy-terminal domain of RNA polymerase II," *Molecular Cell*, vol. 34, no. 3, pp. 387–393, 2009.
- [39] K. Glover-Cutter, S. Larochele, B. Erickson et al., "TFIIH-associated Cdk7 kinase functions in phosphorylation of C-terminal domain Ser7 residues, promoter-proximal pausing, and termination by RNA polymerase II," *Molecular and Cellular Biology*, vol. 29, no. 20, pp. 5455–5464, 2009.
- [40] B. W. Guidi, G. Bjornsdottir, D. C. Hopkins et al., "Mutual targeting of mediator and the TFIIH kinase Kin28," *Journal of Biological Chemistry*, vol. 279, no. 28, pp. 29114–29120, 2004.
- [41] T. Max, M. Sogaard, and J. Q. Svestrup, "Hyperphosphorylation of the C-terminal repeat domain of RNA polymerase II facilitates dissociation of its complex with mediator," *Journal of Biological Chemistry*, vol. 282, no. 19, pp. 14113–14120, 2007.
- [42] S. M. Liao, J. Zhang, D. A. Jeffery et al., "A kinase-cyclin pair in the RNA polymerase II holoenzyme," *Nature*, vol. 374, no. 6518, pp. 193–196, 1995.
- [43] S. Akoulitchev, S. Chuiikov, and D. Reinberg, "TFIIH is negatively regulated by cdk8-containing mediator complexes," *Nature*, vol. 407, no. 6800, pp. 102–106, 2000.
- [44] M. D. Galbraith, A. J. Donner, and J. M. Espinosa, "CDK8: a positive regulator of transcription," *Transcription*, vol. 1, no. 1, pp. 4–12, 2010.
- [45] S. W. Hong, S. M. Hong, J. W. Yoo et al., "Phosphorylation of the RNA polymerase II C-terminal domain by TFIIH kinase is not essential for transcription of *Saccharomyces cerevisiae* genome," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 34, pp. 14276–14280, 2009.
- [46] E. I. Kanin, R. T. Kipp, C. Kung et al., "Chemical inhibition of the TFIIH-associated kinase Cdk7/Kin28 does not impair global mRNA synthesis," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 14, pp. 5812–5817, 2007.
- [47] R. P. Fisher, "Secrets of a double agent: CDK7 in cell-cycle control and transcription," *Journal of Cell Science*, vol. 118, no. 22, pp. 5171–5180, 2005.
- [48] Y. Pei, H. Du, J. Singer et al., "Cyclin-dependent kinase 9 (Cdk9) of fission yeast is activated by the CDK-activating kinase Csk1, overlaps functionally with the TFIIH-associated kinase Mcs6, and associates with the mRNA cap methyltransferase Pcm1 in vivo," *Molecular and Cellular Biology*, vol. 26, no. 3, pp. 777–788, 2006.
- [49] H. Qiu, C. Hu, and A. G. Hinnebusch, "Phosphorylation of the Pol II CTD by KIN28 enhances BUR1/BUR2 recruitment and Ser2 CTD phosphorylation near promoters," *Molecular Cell*, vol. 33, no. 6, pp. 752–762, 2009.
- [50] L. Viladevall, C. V. S. Amour, A. Rosebrock et al., "TFIIH and P-TEFb coordinate transcription with capping enzyme recruitment at specific genes in fission yeast," *Molecular Cell*, vol. 33, no. 6, pp. 738–751, 2009.
- [51] J. M. Lee and A. L. Greenleaf, "CTD kinase large subunit is encoded by CTK1, a gene required for normal growth of *Saccharomyces cerevisiae*," *Gene Expression*, vol. 1, no. 2, pp. 149–167, 1991.
- [52] D. E. Sterner, Jae Moon Lee, S. E. Hardin, and A. L. Greenleaf, "The yeast carboxyl-terminal repeat domain kinase CTDK-I is a divergent cyclin-cyclin-dependent kinase complex," *Molecular and Cellular Biology*, vol. 15, no. 10, pp. 5716–5724, 1995.
- [53] T. Xiao, H. Hall, K. O. Kizer et al., "Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast," *Genes and Development*, vol. 17, no. 5, pp. 654–663, 2003.
- [54] M. L. Youdell, K. O. Kizer, E. Kisseleva-Romanova et al., "Roles for Ctk1 and Spt6 in regulating the different methylation states of histone H3 lysine 36," *Molecular and Cellular Biology*, vol. 28, no. 16, pp. 4915–4926, 2008.
- [55] S. H. Ahn, M. Kim, and S. Buratowski, "Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing," *Molecular Cell*, vol. 13, no. 1, pp. 67–76, 2004.
- [56] D. A. Skaar and A. L. Greenleaf, "The RNA polymerase II CTD kinase CTDK-I affects pre-mRNA 3' cleavage/polyadenylation through the processing component Pti1p," *Molecular Cell*, vol. 10, no. 6, pp. 1429–1439, 2002.
- [57] S. H. Ahn, M. C. Keogh, and S. Buratowski, "Ctk1 promotes dissociation of basal transcription factors from elongating RNA polymerase II," *EMBO Journal*, vol. 28, no. 3, pp. 205–212, 2009.
- [58] S. Yao, A. Neiman, and G. Prelich, "BUR1 and BUR2 encode a divergent cyclin-dependent kinase-cyclin complex important for transcription in vivo," *Molecular and Cellular Biology*, vol. 20, no. 19, pp. 7080–7087, 2000.
- [59] M. C. Keogh, V. Podolny, and S. Buratowski, "Bur1 kinase is required for efficient transcription elongation by RNA polymerase II," *Molecular and Cellular Biology*, vol. 23, no. 19, pp. 7005–7018, 2003.
- [60] Y. Pei and S. Shuman, "Characterization of the *Schizosaccharomyces pombe* Cdk9/Pch1 protein kinase: Spt5 phosphorylation, autophosphorylation, and mutational analysis," *Journal of Biological Chemistry*, vol. 278, no. 44, pp. 43346–43356, 2003.
- [61] Y. Liu, L. Warfield, C. Zhang et al., "Phosphorylation of the transcription elongation factor Spt5 by yeast Bur1 kinase stimulates recruitment of the PAF complex," *Molecular and Cellular Biology*, vol. 29, no. 17, pp. 4852–4863, 2009.
- [62] J. Karagiannis and M. K. Balasubramanian, "A cyclin-dependent kinase that promotes cytokinesis through modulating phosphorylation of the carboxy terminal domain of the RNA Pol II Rpb1p sub-unit," *PLoS One*, vol. 2, no. 5, article e433, 2007.
- [63] N. F. Marshall, J. Peng, Z. Xie, and D. H. Price, "Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase," *Journal of Biological Chemistry*, vol. 271, no. 43, pp. 27176–27183, 1996.
- [64] N. F. Marshall and D. H. Price, "Control of formation of two distinct classes of RNA polymerase II elongation complexes,"

- Molecular and Cellular Biology*, vol. 12, no. 5, pp. 2078–2090, 1992.
- [65] N. F. Marshall and D. H. Price, “Purification of P-TEFb, a transcription factor required for the transition into productive elongation,” *Journal of Biological Chemistry*, vol. 270, no. 21, pp. 12335–12338, 1995.
- [66] T. Lenasi and M. Barboric, “P-TEFb stimulates transcription elongation and pre-mRNA splicing through multilateral mechanisms,” *RNA Biology*, vol. 7, no. 2, pp. 145–150, 2010.
- [67] B. M. Peterlin and D. H. Price, “Controlling the elongation phase of transcription with P-TEFb,” *Molecular Cell*, vol. 23, no. 3, pp. 297–305, 2006.
- [68] S. Nechaev and K. Adelman, “Pol II waiting in the starting gates: regulating the transition from transcription initiation into productive elongation,” *Biochimica et Biophysica Acta*, vol. 1809, pp. 34–45, 2011.
- [69] A. Wood and A. Shilatifard, “Bur1/Bur2 and the Ctk complex in yeast: the split personality of mammalian P-TEFb,” *Cell Cycle*, vol. 5, no. 10, pp. 1066–1068, 2006.
- [70] Z. Guo and J. W. Stiller, “Comparative genomics of cyclin-dependent kinases suggest co-evolution of the RNAP II C-terminal domain and CTD-directed CDKs,” *BMC Genomics*, vol. 5, article 69, 2004.
- [71] J. Liu and E. T. Kipreos, “Evolution of cyclin-dependent kinases (CDKs) and CDK-activating kinases (CAKs): differential conservation of CAKs in yeast and metazoa,” *Molecular Biology and Evolution*, vol. 17, no. 7, pp. 1061–1074, 2000.
- [72] B. Bartkowiak, P. Liu, H. P. Phatnani et al., “CDK12 is a transcription elongation-associated CTD kinase, the metazoan ortholog of yeast Ctk1,” *Genes and Development*, vol. 24, no. 20, pp. 2303–2316, 2010.
- [73] R. Berro, C. Pedati, K. Kehn-Hall et al., “CDK13, a new potential human immunodeficiency virus type 1 inhibitory factor regulating viral mRNA splicing,” *Journal of Virology*, vol. 82, no. 14, pp. 7155–7166, 2008.
- [74] H. H. Chen, Y. C. Wang, and M. J. Fann, “Identification and characterization of the CDK12/cyclin L1 complex involved in alternative splicing regulation,” *Molecular and Cellular Biology*, vol. 26, no. 7, pp. 2736–2745, 2006.
- [75] H. H. Chen, Y. H. Wong, A. M. Geneviere, and M. J. Fann, “CDK13/CDC2L5 interacts with L-type cyclins and regulates alternative splicing,” *Biochemical and Biophysical Research Communications*, vol. 354, no. 3, pp. 735–740, 2007.
- [76] T. K. Ko, E. Kelly, and J. Pines, “CrkRS: a novel conserved Cdc2-related protein kinase that colocalises with SC35 speckles,” *Journal of Cell Science*, vol. 114, no. 14, pp. 2591–2603, 2001.
- [77] Y. Even, S. Durieux, M. L. Escande et al., “CDC2L5, a Cdk-like kinase with RS domain, interacts with the ASF/SF2-associated protein p32 and affects splicing in vivo,” *Journal of Cellular Biochemistry*, vol. 99, no. 3, pp. 890–904, 2006.
- [78] T. J. Fu, J. Peng, G. Lee, D. H. Price, and O. Flores, “Cyclin K functions as a CDK9 regulatory subunit and participates in RNA polymerase II transcription,” *Journal of Biological Chemistry*, vol. 274, no. 49, pp. 34527–34530, 1999.
- [79] D. J. Matthews and M. E. Gerritsen, *Targeting Protein Kinases for Cancer Therapy*, John Wiley & Sons, Hoboken, NJ, USA, 2010.
- [80] H. S. Y. Mancebo, G. Lee, J. Flygare et al., “P-TEFb kinase is required for HIV Tat transcriptional activation in vivo and in vitro,” *Genes and Development*, vol. 11, no. 20, pp. 2633–2644, 1997.
- [81] Y. Zhu, T. Pe’ery, J. Peng et al., “Transcription elongation factor P-TEFb is required for HIV-1 tat transactivation in vitro,” *Genes and Development*, vol. 11, no. 20, pp. 2622–2632, 1997.
- [82] A. V. Philips and T. A. Cooper, “RNA processing and human disease,” *Cellular and Molecular Life Sciences*, vol. 57, no. 2, pp. 235–249, 2000.
- [83] K. Inoue, M. Ohno, H. Sakamoto, and Y. Shimura, “Effect of the cap structure on pre-mRNA splicing in *Xenopus* oocyte nuclei,” *Genes and Development*, vol. 3, no. 9, pp. 1472–1479, 1989.
- [84] J. D. Lewis, E. Izaurralde, A. Jarmolowski, C. McGuigan, and L. W. Mattaj, “A nuclear cap-binding complex facilitates association of U1 snRNP with the cap-proximal 5’ splice site,” *Genes and Development*, vol. 10, no. 13, pp. 1683–1698, 1996.
- [85] C. Cooke and J. C. Alwine, “The cap and the 3’ splice site similarly affect polyadenylation efficiency,” *Molecular and Cellular Biology*, vol. 16, no. 6, pp. 2579–2584, 1996.
- [86] S. M. Flaherty, P. Fortes, E. Izaurralde, I. W. Mattaj, and G. M. Gilmartin, “Participation of the nuclear cap binding complex in pre-mRNA 3’ processing,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 22, pp. 11893–11898, 1997.
- [87] R. P. Hart, M. A. McDevitt, and J. R. Nevins, “Poly(A) site cleavage in a HeLa nuclear extract is dependent on downstream sequences,” *Cell*, vol. 43, no. 3, pp. 677–683, 1985.
- [88] G. Baurén, S. Belikov, and L. Wieslander, “Transcriptional termination in the Balbiani ring 1 gene is closely coupled to 3’-end formation and excision of the 3’-terminal intron,” *Genes and Development*, vol. 12, no. 17, pp. 2759–2769, 1998.
- [89] M. Niwa and S. M. Berget, “Mutation of the AAUAAA polyadenylation signal depresses in vitro splicing of proximal but not distal introns,” *Genes and Development*, vol. 5, no. 11, pp. 2086–2095, 1991.
- [90] S. Vagner, C. Vagner, and I. W. Mattaj, “The carboxyl terminus of vertebrate poly(A) polymerase interacts with U2AF 65 to couple 3’-end processing and splicing,” *Genes and Development*, vol. 14, no. 4, pp. 403–413, 2000.
- [91] M. Alló, V. Buggiano, J. P. Fededa et al., “Control of alternative splicing through siRNA-mediated transcriptional gene silencing,” *Nature Structural and Molecular Biology*, vol. 16, no. 7, pp. 717–724, 2009.
- [92] R. F. Luco, Q. Pan, K. Tominaga, B. J. Blencowe, O. M. Pereira-Smith, and T. Misteli, “Regulation of alternative splicing by histone modifications,” *Science*, vol. 327, no. 5968, pp. 996–1000, 2010.
- [93] I. E. Schor, N. Rascovan, F. Pelisch, M. Alió, and A. R. Kornblihtt, “Neuronal cell depolarization induces intragenic chromatin modifications affecting NCAM alternative splicing,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 11, pp. 4325–4330, 2009.
- [94] R. Jove and J. L. Manley, “In vitro transcription from the adenovirus 2 major late promoter utilizing templates truncated at promoter-proximal sites,” *Journal of Biological Chemistry*, vol. 259, no. 13, pp. 8513–8521, 1984.
- [95] E. B. Rasmussen and J. T. Lis, “In vivo transcriptional pausing and cap formation on three *Drosophila* heat shock genes,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 90, no. 17, pp. 7923–7927, 1993.
- [96] E. J. Cho, T. Takagi, C. R. Moore, and S. Buratowski, “mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain,” *Genes and Development*, vol. 11, no. 24, pp. 3319–3326, 1997.

- [97] S. McCracken, N. Fong, E. Rosonina et al., "5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II," *Genes and Development*, vol. 11, no. 24, pp. 3306–3318, 1997.
- [98] Z. Yue, E. Maldonado, R. Pillutla, H. Cho, D. Reinberg, and A. J. Shatkin, "Mammalian capping enzyme complements mutant *Saccharomyces cerevisiae* lacking mRNA guanylyltransferase and selectively binds the elongating form of RNA polymerase II," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 24, pp. 12898–12903, 1997.
- [99] C. K. Ho and S. Shuman, "Distinct roles for CTD Ser-2 and Ser-5 phosphorylation in the recruitment and allosteric activation of mammalian mRNA capping enzyme," *Molecular Cell*, vol. 3, no. 3, pp. 405–411, 1999.
- [100] C. Fabrega, V. Shen, S. Shuman, and C. D. Lima, "Structure of an mRNA capping enzyme bound to the phosphorylated carboxy-terminal domain of RNA polymerase II," *Molecular Cell*, vol. 11, no. 6, pp. 1549–1561, 2003.
- [101] B. M. Lunde, S. L. Reichow, M. Kim et al., "Cooperative interaction of transcription termination factors with the RNA polymerase II C-terminal domain," *Nature Structural and Molecular Biology*, vol. 17, no. 10, pp. 1195–1201, 2010.
- [102] N. Proudfoot, "New perspectives on connecting messenger RNA 3' end formation to transcription," *Current Opinion in Cell Biology*, vol. 16, no. 3, pp. 272–278, 2004.
- [103] S. Buratowski, "Connections between mRNA 3' end processing and transcription termination," *Current Opinion in Cell Biology*, vol. 17, no. 3, pp. 257–261, 2005.
- [104] D. Barilla, B. A. Lee, and N. J. Proudfoot, "Cleavage/polyadenylation factor IA associates with the carboxyl-terminal domain of RNA polymerase II in *Saccharomyces cerevisiae*," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, pp. 445–450, 2001.
- [105] A. Kyburz, M. Sadowski, B. Dichtl, and W. Keller, "The role of the yeast cleavage and polyadenylation factor subunit Ydh1p/Cft2p in pre-mRNA 3'-end formation," *Nucleic Acids Research*, vol. 31, no. 14, pp. 3936–3945, 2003.
- [106] D. D. Licatalosi, G. Geiger, M. Minet et al., "Functional interaction of yeast pre-mRNA 3' end processing factors with RNA polymerase II," *Molecular Cell*, vol. 9, no. 5, pp. 1101–1111, 2002.
- [107] T. Maniatis and R. Reed, "An extensive network of coupling among gene expression machines," *Nature*, vol. 416, no. 6880, pp. 499–506, 2002.
- [108] S. McCracken, N. Fong, K. Yankulov et al., "The C-terminal domain of RNA polymerase II couples mRNA processing to transcription," *Nature*, vol. 385, no. 6614, pp. 357–361, 1997.
- [109] N. J. Proudfoot, A. Furger, and M. J. Dye, "Integrating mRNA processing with transcription," *Cell*, vol. 108, no. 4, pp. 501–512, 2002.
- [110] Y. Hirose and J. L. Manley, "RNA polymerase II is an essential mRNA polyadenylation factor," *Nature*, vol. 395, no. 6697, pp. 93–96, 1998.
- [111] M. Kim, N. J. Krogan, L. Vasiljeva et al., "The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II," *Nature*, vol. 432, no. 7016, pp. 517–522, 2004.
- [112] J. N. Kuehner, E. L. Pearson, and C. Moore, "Unravelling the means to an end: RNA polymerase II transcription termination," *Nature Reviews Molecular Cell Biology*, vol. 12, pp. 283–294, 2011.
- [113] C. G. Noble, D. Hollingworth, S. R. Martin et al., "Key features of the interaction between Pcf11 CID and RNA polymerase II CTD," *Nature Structural and Molecular Biology*, vol. 12, no. 2, pp. 144–151, 2005.
- [114] E. Y. Jacobs, I. Ogiwara, and A. Weiner, "Role of the C-terminal domain of RNA polymerase II in U2 snRNA transcription and 3' processing," *Molecular and Cellular Biology*, vol. 24, no. 2, pp. 846–855, 2004.
- [115] J. E. Medlin, P. Uguen, A. Taylor, D. L. Bentley, and S. Murphy, "The C-terminal domain of pol II and a DRB-sensitive kinase are required for 3' processing of U2 snRNA," *EMBO Journal*, vol. 22, no. 4, pp. 925–934, 2003.
- [116] S. Eglhoff, D. O'Reilly, and S. Murphy, "Expression of human snRNA genes from beginning to end," *Biochemical Society Transactions*, vol. 36, no. 4, pp. 590–594, 2008.
- [117] D. Baillat, M. A. Hakimi, A. M. Naar, A. Shilatifard, N. Cooch, and R. Shiekhattar, "Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II," *Cell*, vol. 123, no. 2, pp. 265–276, 2005.
- [118] S. Eglhoff, S. A. Szczepaniak, M. Dienstbier, A. Taylor, S. Knight, and S. Murphy, "The integrator complex recognizes a new double mark on the RNA polymerase II carboxyl-terminal domain," *Journal of Biological Chemistry*, vol. 285, no. 27, pp. 20564–20569, 2010.
- [119] B. Li, M. Carey, and J. L. Workman, "The role of chromatin during transcription," *Cell*, vol. 128, no. 4, pp. 707–719, 2007.
- [120] K. O. Kizer, H. P. Phatnani, Y. Shibata, H. Hall, A. L. Greenleaf, and B. D. Strahl, "A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation," *Molecular and Cellular Biology*, vol. 25, no. 8, pp. 3305–3316, 2005.
- [121] N. J. Krogan, M. Kim, A. Tong et al., "Methylation of histone H3 by Set2 in *Saccharomyces cerevisiae* is linked to transcriptional elongation by RNA polymerase II," *Molecular and Cellular Biology*, vol. 23, no. 12, pp. 4207–4218, 2003.
- [122] N. J. Krogan, J. Dover, A. Wood et al., "The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation," *Molecular Cell*, vol. 11, no. 3, pp. 721–729, 2003.
- [123] H. H. Ng, F. Robert, R. A. Young, and K. Struhl, "Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity," *Molecular Cell*, vol. 11, no. 3, pp. 709–719, 2003.
- [124] M. Hampsey and D. Reinberg, "Tails of intrigue: phosphorylation of RNA polymerase II mediates histone methylation," *Cell*, vol. 113, no. 4, pp. 429–432, 2003.
- [125] A. A. Joshi and K. Struhl, "Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation," *Molecular Cell*, vol. 20, no. 6, pp. 971–978, 2005.
- [126] M. C. Keogh, S. K. Kurdistani, S. A. Morris et al., "Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex," *Cell*, vol. 123, no. 4, pp. 593–605, 2005.
- [127] S. Drouin, L. Laramee, P. E. Jacques, A. Forest, M. Bergeron, and F. Robert, "DSIF and RNA polymerase II CTD phosphorylation coordinate the recruitment of Rpd3S to actively transcribed genes," *PLoS Genet*, vol. 6, Article ID e1001173, 2010.
- [128] C. K. Govind, H. Qiu, D. S. Ginsburg et al., "Phosphorylated Pol II CTD recruits multiple HDACs, including Rpd3C(S), for methylation-dependent deacetylation of ORF nucleosomes," *Molecular Cell*, vol. 39, no. 2, pp. 234–246, 2010.

- [129] S. M. Yoh, H. Cho, L. Pickle, R. M. Evans, and K. A. Jones, "The Spt6 SH2 domain binds Ser2-P RNAPII to direct Iws1-dependent mRNA splicing and export," *Genes and Development*, vol. 21, no. 2, pp. 160–174, 2007.
- [130] M. Sun, L. Larivière, S. Dengl, A. Mayer, and P. Cramer, "A tandem SH2 domain in transcription elongation factor Spt6 binds the phosphorylated RNA polymerase II C-terminal repeat domain (CTD)," *Journal of Biological Chemistry*, vol. 285, no. 53, pp. 41597–41603, 2010.
- [131] S. M. Yoh, J. S. Lucas, and K. A. Jones, "The Iws1:Spt6:CTD complex controls cotranscriptional mRNA biosynthesis and HYPB/Setd2-mediated histone H3K36 methylation," *Genes and Development*, vol. 22, no. 24, pp. 3422–3434, 2008.
- [132] M. Rodriguez-Paredes, M. Ceballos-Chavez, M. Esteller, M. Garcia-Dominguez, and J. C. Reyes, "The chromatin remodeling factor CHD8 interacts with elongating RNA polymerase II and controls expression of the cyclin E2 gene," *Nucleic Acids Research*, vol. 37, pp. 2449–2460, 2009.
- [133] S. H. Kwon, L. Florens, S. K. Swanson, M. P. Washburn, S. M. Abmayr, and J. L. Workman, "Heterochromatin protein 1 (HP1) connects the FACT histone chaperone complex to the phosphorylated CTD of RNA polymerase II," *Genes and Development*, vol. 24, no. 19, pp. 2133–2145, 2010.
- [134] M. J. Munoz, M. de la Mata, and A. R. Kornblihtt, "The carboxy terminal domain of RNA polymerase II and alternative splicing," *Trends in Biochemical Sciences*, vol. 35, no. 9, pp. 497–504, 2010.
- [135] A. C. Goldstrohm, A. L. Greenleaf, and M. A. Garcia-Blanco, "Co-transcriptional splicing of pre-messenger RNAs: considerations for the mechanism of alternative splicing," *Gene*, vol. 277, no. 1-2, pp. 31–47, 2001.
- [136] G. Bird, D. A. R. Zorio, and D. L. Bentley, "RNA polymerase II carboxy-terminal domain phosphorylation is required for cotranscriptional pre-mRNA splicing and 3'-end formation," *Molecular and Cellular Biology*, vol. 24, no. 20, pp. 8963–8969, 2004.
- [137] A. L. Greenleaf, "Positive patches and negative noodles: linking RNA processing to transcription?" *Trends in Biochemical Sciences*, vol. 18, no. 4, pp. 117–119, 1993.
- [138] E. Kim, L. Du, D. B. Bregman, and S. L. Warren, "Splicing factors associate with hyperphosphorylated RNA polymerase II in the absence of pre-mRNA," *Journal of Cell Biology*, vol. 136, no. 1, pp. 19–28, 1997.
- [139] M. J. Mortillaro, B. J. Blencowe, X. Wei et al., "A hyperphosphorylated form of the large subunit of RNA polymerase II is associated with splicing complexes and the nuclear matrix," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, no. 16, pp. 8253–8257, 1996.
- [140] A. Yuryev, M. Patturajan, Y. Litingtung et al., "The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, no. 14, pp. 6975–6980, 1996.
- [141] L. Du and S. L. Warren, "A functional interaction between the carboxy-terminal domain of RNA polymerase II and pre-mRNA splicing," *Journal of Cell Biology*, vol. 136, no. 1, pp. 5–18, 1997.
- [142] Y. Hirose, R. Tacke, and J. L. Manley, "Phosphorylated RNA polymerase II stimulates pre-mRNA splicing," *Genes and Development*, vol. 13, no. 10, pp. 1234–1239, 1999.
- [143] C. Zeng and S. M. Berget, "Participation of the C-terminal domain of RNA polymerase II in exon definition during pre-mRNA splicing," *Molecular and Cellular Biology*, vol. 20, no. 21, pp. 8290–8301, 2000.
- [144] M. Patturajan, X. Wei, R. Berezney, and J. L. Corden, "A nuclear matrix protein interacts with the phosphorylated C-terminal domain of RNA polymerase II," *Molecular and Cellular Biology*, vol. 18, no. 4, pp. 2406–2415, 1998.
- [145] D. P. Morris and A. L. Greenleaf, "The splicing factor, Prp40, binds the phosphorylated carboxyl-terminal domain of RNA polymerase II," *Journal of Biological Chemistry*, vol. 275, no. 51, pp. 39935–39943, 2000.
- [146] S. M. Carty, A. C. Goldstrohm, C. Suñé, M. A. Garcia-Blanco, and A. L. Greenleaf, "Protein-interaction modules that organize nuclear function: FF domains of CA150 bind the phosphoCTD of RNA polymerase II," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 16, pp. 9015–9020, 2000.
- [147] A. Emili, M. Shales, S. McCracken et al., "Splicing and transcription-associated proteins PSF and p54nrb/nonO bind to the RNA polymerase II CTD," *RNA*, vol. 8, no. 9, pp. 1102–1111, 2002.
- [148] C. J. David, A. R. Boyne, S. R. Millhouse, and J. L. Manley, "The RNA polymerase II C-terminal domain promotes splicing activation through recruitment of a U2AF65-Prp19 complex," *Genes and Development*, vol. 25, pp. 972–983, 2011.
- [149] M. de la Mata and A. R. Kornblihtt, "RNA polymerase II C-terminal domain mediates regulation of alternative splicing by SRp20," *Nature Structural and Molecular Biology*, vol. 13, no. 11, pp. 973–980, 2006.
- [150] M. de la Mata, C. R. Alonso, S. Kadener et al., "A slow RNA polymerase II affects alternative splicing in vivo," *Molecular Cell*, vol. 12, no. 2, pp. 525–532, 2003.
- [151] S. Kadener, P. Cramer, G. Nogués et al., "Antagonistic effects of T-Ag and VP16 reveal a role for RNA pol II elongation on alternative splicing," *EMBO Journal*, vol. 20, no. 20, pp. 5759–5768, 2001.
- [152] E. Batsche, M. Yaniv, and C. Muchardt, "The human SWI/SNF subunit Brm is a regulator of alternative splicing," *Nature Structural & Molecular Biology*, vol. 13, pp. 22–29, 2006.
- [153] M. J. Dye, N. Gromak, and N. J. Proudfoot, "Exon tethering in transcription by RNA polymerase II," *Molecular Cell*, vol. 21, no. 6, pp. 849–859, 2006.
- [154] P. Loyer, J. H. Trembley, J. A. Grenet et al., "Characterization of cyclin L1 and L2 interactions with CDK11 and splicing factors: influence of cyclin L isoforms on splice site selection," *Journal of Biological Chemistry*, vol. 283, no. 12, pp. 7721–7732, 2008.
- [155] J. H. Trembley, D. Hu, L. C. Hsu et al., "PITSLSRE p110 protein kinases associate with transcription complexes and affect their activity," *Journal of Biological Chemistry*, vol. 277, no. 4, pp. 2589–2596, 2002.
- [156] S. Lin and X. D. Fu, "SR proteins and related factors in alternative splicing," *Advances in Experimental Medicine and Biology*, vol. 623, pp. 107–122, 2007.
- [157] J. C. Long and J. F. Caceres, "The SR protein family of splicing factors: master regulators of gene expression," *Biochemical Journal*, vol. 417, no. 1, pp. 15–27, 2009.
- [158] X. Y. Zhong, P. Wang, J. Han, M. G. Rosenfeld, and X. D. Fu, "SR proteins in vertical integration of gene expression from transcription to RNA processing to translation," *Molecular Cell*, vol. 35, no. 1, pp. 1–10, 2009.
- [159] S. Rodriguez-Navarro, "Insights into SAGA function during gene expression," *EMBO Reports*, vol. 10, pp. 843–850, 2009.

- [160] P. Pascual-Garcia and S. Rodriguez-Navarro, "A tale of coupling, Sus1 function in transcription and mRNA export," *RNA Biology*, vol. 6, no. 2, pp. 141–144, 2009.
- [161] K. Strasser, S. Masuda, P. Mason et al., "TREX is a conserved complex coupling transcription with messenger RNA export," *Nature*, vol. 417, pp. 304–308, 2002.
- [162] K. Strasser and E. Hurt, "Splicing factor Sub2p is required for nuclear mRNA export through its interaction with Yra1p," *Nature*, vol. 413, no. 6856, pp. 648–652, 2001.
- [163] F. Stutz, A. Bachi, T. Doerks et al., "REF, an evolutionary conserved family of hnRNP-like proteins, interacts with TAP/Mex67p and participates in mRNA nuclear export," *RNA*, vol. 6, pp. 638–650, 2000.
- [164] S. A. Johnson, G. Cubberley, and D. L. Bentley, "Cotranscriptional recruitment of the mRNA export factor Yra1 by direct interaction with the 3' end processing factor Pcf11," *Molecular Cell*, vol. 33, no. 2, pp. 215–226, 2009.
- [165] H. Le Hir, E. Izaurralde, L. E. Maquat, and M. J. Moore, "The spliceosome deposits multiple proteins 20–24 nucleotides upstream of mRNA exon-exon junctions," *EMBO Journal*, vol. 19, no. 24, pp. 6860–6869, 2000.
- [166] T. Ø. Tange, T. Shibuya, M. S. Jurica, and M. J. Moore, "Biochemical analysis of the EJC reveals two new factors and a stable tetrameric protein core," *RNA*, vol. 11, no. 12, pp. 1869–1883, 2005.
- [167] F. Lejeune, Y. Ishigaki, X. Li, and L. E. Maquat, "The exon junction complex is detected on CBP80-bound but not eIF4E-bound mRNA in mammalian cells: dynamics of mRNP remodeling," *EMBO Journal*, vol. 21, no. 13, pp. 3536–3545, 2002.
- [168] D. Zenklusen, P. Vinciguerra, Y. Strahm, and F. Stutz, "The yeast hnRNP-Like proteins Yra1p and Yra2p participate in mRNA export through interaction with Mex67p," *Molecular and Cellular Biology*, vol. 21, no. 13, pp. 4219–4232, 2001.
- [169] A. G. Rondon, S. Jimeno, M. Garcia-Rubio, and A. Aguilera, "Molecular evidence that the eukaryotic THO/TREX complex is required for efficient transcription elongation," *Journal of Biological Chemistry*, vol. 278, pp. 39037–39043, 2003.
- [170] H. P. Phatnani, J. C. Jones, and A. L. Greenleaf, "Expanding the functional repertoire of CTD kinase I and RNA polymerase II: novel phosphoCTD-associating proteins in the yeast proteome," *Biochemistry*, vol. 43, no. 50, pp. 15702–15719, 2004.
- [171] C. Maris, C. Dominguez, and F. H. T. Allain, "The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression," *FEBS Journal*, vol. 272, no. 9, pp. 2118–2131, 2005.
- [172] A. Clery, M. Blatter, and F. H. Allain, "RNA recognition motifs: boring? Not quite," *Current Opinion in Structural Biology*, vol. 18, pp. 290–298, 2008.
- [173] A. Aguilera, "The connection between transcription and genomic instability," *EMBO Journal*, vol. 21, no. 3, pp. 195–201, 2002.
- [174] X. Li and J. L. Manley, "Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability," *Cell*, vol. 122, no. 3, pp. 365–378, 2005.
- [175] C. Gonzalez-Aguilera, C. Tous, B. Gomez-Gonzalez, P. Huertas, R. Luna, and A. Aguilera, "The THP1-SAC3-SUS1-CDC31 complex works in transcription elongation-mRNA export preventing RNA-mediated genome instability," *Molecular Biology of the Cell*, vol. 19, pp. 4310–4318, 2008.
- [176] S. Jimeno, A. G. Rondán, R. Luna, and A. Aguilera, "The yeast THO complex and mRNA export factors link RNA metabolism with transcription and genome instability," *EMBO Journal*, vol. 21, no. 13, pp. 3526–3535, 2002.
- [177] D. Ostapenko and M. J. Solomon, "Budding yeast CTDK-I is required for DNA damage-induced transcription," *Eukaryotic Cell*, vol. 2, no. 2, pp. 274–283, 2003.
- [178] C. B. Bennett, L. K. Lewis, G. Karthikeyan et al., "Genes required for ionizing radiation resistance in yeast," *Nature Genetics*, vol. 29, no. 4, pp. 426–434, 2001.
- [179] T. J. Westmoreland, J. R. Marks, J. A. Olson Jr., E. M. Thompson, M. A. Resnick, and C. B. Bennett, "Cell cycle progression in G1 and S phases is CCR4 dependent following ionizing radiation or replication stress in *Saccharomyces cerevisiae*," *Eukaryotic Cell*, vol. 3, no. 2, pp. 430–446, 2004.
- [180] O. Aygun, J. Svejstrup, and Y. Liu, "A RECQ5-RNA polymerase II association identified by targeted proteomic analysis of human chromatin," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 25, pp. 8580–8584, 2008.
- [181] K. Izumikawa, M. Yanagida, T. Hayano et al., "Association of human DNA helicase RecQ5 $\beta$  with RNA polymerase II and its possible role in transcription," *Biochemical Journal*, vol. 413, no. 3, pp. 505–516, 2008.
- [182] R. Kanagaraj, D. Huehn, A. MacKellar et al., "RECQ5 helicase associates with the C-terminal repeat domain of RNA polymerase II during productive elongation phase of transcription," *Nucleic Acids Research*, vol. 38, no. 22, pp. 8131–8140, 2010.