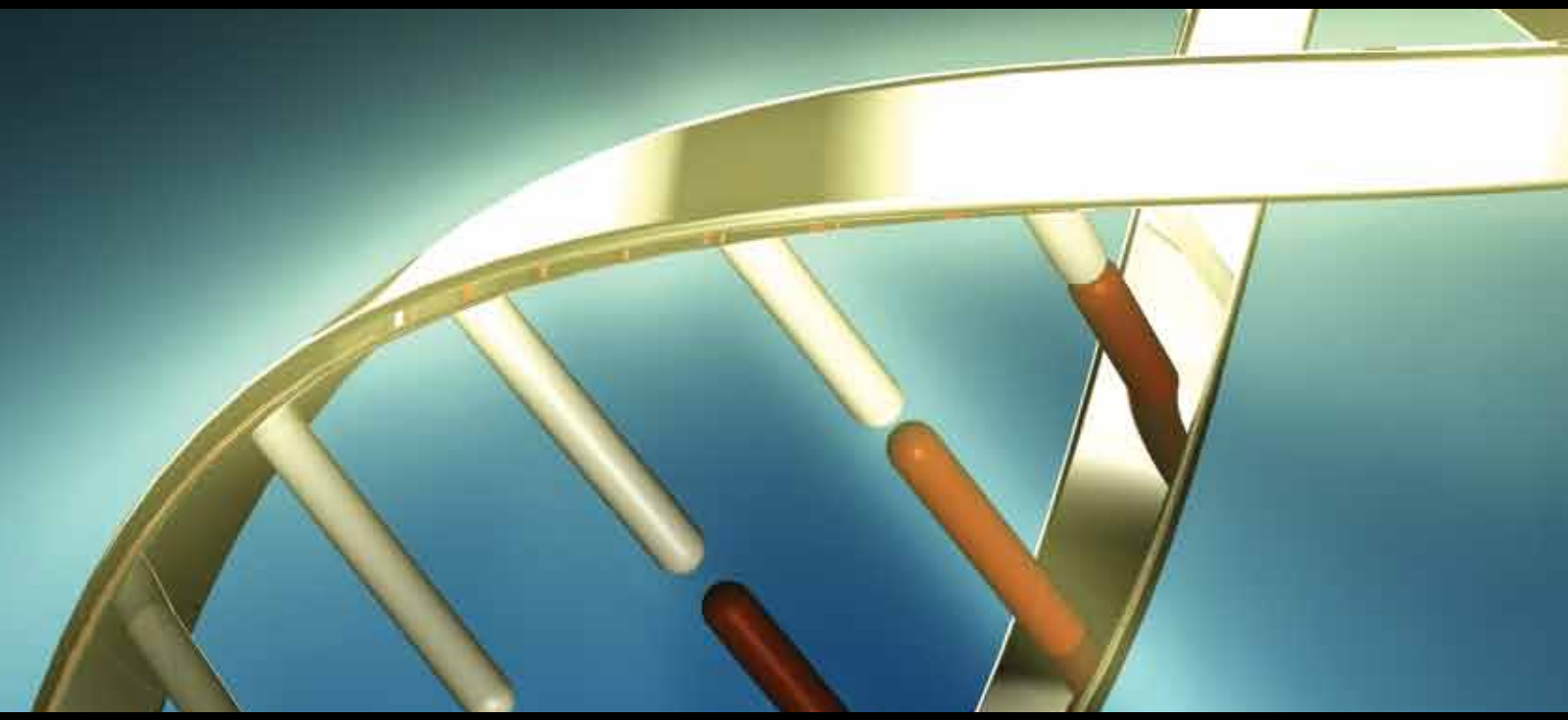
Genetics Research International

Gene Control during Transcription Elongation

Guest Editors: Sebastián Chávez, David S. Gross, Damien Hermand, and Carles Suñé



Gene Control during Transcription Elongation

Genetics Research International

Gene Control during Transcription Elongation

Guest Editors: Sebastián Chávez, David S. Gross,
Damien Hermand, and Carles Suñé

Editorial Board

Contents

Hindawi Publishing Corporation
Genetics Research International
Volume 2012, Article ID 758384, 2 pages
doi:10.1155/2012/758384

Editorial
Gene Control during Transcription Elongation

Sebastián Chávez,[1] David S. Gross,[2] Damien Hermand,[3] and Carlos Suñé[4]

[1] *Departamento de Genética, Facultad de Biología, Universidad de Sevilla, Seville 41012, Spain*
[2] *Department of Biochemistry and Molecular Biology, Louisiana State University Health Sciences Center, Shreveport, LA 71130-3932, USA*
[3] *Laboratoire de Genetique Moleculaire, Academie Universitaire Louvain, 5000 Namur, Belgium*
[4] *Departamento de Biología Molecular, Instituto de Parasitología y Biomedicina "López Neyra" (IPBLN-CSIC), Parque Tecnológico Ciencias de la Salud, 18100 Armilla, Spain*

Correspondence should be addressed to Sebastián Chávez, schavez@us.es

Received 7 December 2011; Accepted 7 December 2011

This special issue of Genetics Research International is dedicated to transcription elongation and the role that it plays in the control of gene expression. As K. Brannan and D. Bentley highlight in their retrospective review, 30 years after the first examples of gene control during transcription elongation were described, the importance of this step of the RNA biogenesis process has become increasingly clear, in parallel with myriad findings that connect transcription elongation to almost every relevant genome-related phenomenon. It is, therefore, gratifying to present the most current work from an array of leading scientists, who offer in one rich volume up-to-date review articles of this interesting field.

Addressed are the covalent modifications of the C-terminal domain of the largest subunit of RNA polymerase II, whose extensive studies have launched the interesting paradigm of a CTD code governing and integrating cotranscriptionally the different steps of mRNA biogenesis. The teams led by A. Greenleaf (B. Bartkowiak et al.) and A. Ansari (D. Zhang et al.) review this important aspect of RNA-polymerase-II-dependent transcription. They discuss the readers, writers, and erasers of the posttranslational modifications that occur on the CTD and that function as a binding platform for enzymatic complexes that regulate Pol II elongation, pre-mRNA splicing, RNA export from the nucleus, chromatin remodelling, and DNA repair.

Also featured is the way transcription is ruled by the other great code, that is, the histone code, a panoply of specific interactions between genome effectors and chromatin that is mediated by covalent modifications of the histones. E. M. Crisucci and K. M. Arndt explore the roles of the Paf1 complex (Paf1C) in regulating gene expression in both budding yeast and metazoans. They review evidence that Paf1C associates with elongating Pol II, and by doing so facilitates histone modifications to the underlying chromatin template as well as contributes to transcription termination and RNA 3′-end formation. A complementary view of chromatin dynamics during transcription elongation is provided by A. A. Duina, who provides a comprehensive review of the roles of the histone chaperones Spt6 and FACT in facilitating passage of Pol II on the chromatin template. He compares and contrasts their mechanisms, highlighting recent evidence indicating that they travel across transcribed regions in likely association with elongating Pol II, where they play an important role in the removal and redeposition of nucleosomes during polymerase elongation. Cotraversal of Pol II and its regulatory factors is also an important aspect of those regulatory phenomena affecting the transition between initiation and elongation and during early elongation. L. A. Stargell and colleagues (M. N. Yearling et al.) describe the complex interplay among the constellation of factors that govern the transition of poised RNA polymerases into active elongation.

The posttranscriptional fate of RNA is paradoxically dictated cotranscriptionally. mRNA export and alternative splicing are coupled to transcription elongation as the teams led by A. Kornblihtt (M. de la Mata et al.) and S. Rodriguez-Navarro (M. M. Molina-Navarro et al.) explain in their reviews. C. Suñé and colleagues (M. Sánchez-Álvarez et al.) show us that this coupling is not only temporal but also spatial and describe the increasing progress of light microscopy

at single-cell resolution to characterize transcription factories and interchromatin granule clusters within the nucleus.

mRNAs are not the only RNA species produced by RNA polymerase II. J. Colin et al. guide us through the fascinating world of cryptic transcripts and their connection to the control of gene expression by early transcription termination. The coexistence of canonical and cryptic transcription is a challenge for genomicists when addressing global transcription. J. E. Pérez-Ortín and colleagues address this and other problems in their review on how to measure the different steps of gene expression genomewide.

Control of transcription elongation is essential for understanding the regulatory behaviour of the genome; accordingly, the number of genes known to be regulated at the elongation level is exponentially increasing in the literature. HIV transcription is a classic model of gene regulation during elongation. K. A. Nilson and D. H. Price review it, with special attention to the roles of the viral Tat protein, the best characterized gene-specific effector of transcription elongation, and with a view on the therapeutic approaches to AIDS that targets HIV transcription. As E. de Nadal and F. Posas describe, stress-responsive genes provide another good example where transcription elongation is highly regulated. Moreover, functional contributions of transcription elongation reach fields beyond gene expression, like V(D)J recombination during the generation of antigen receptor diversity in T and B lymphocytes. The team of C. Hernández-Munain (B. del Blanco et al.) explains in their review how the elongation activity of RNA polymerase II is required for creating accessible chromatin for the RAG1 and RAG2 recombinases to initiate recombination.

Transcription elongation by RNA polymerase II is the focus of most reviews in this special issue, but elongation by other RNA polymerases is also attracting the interest of the scientific community. This is the case for O. Gadal and his colleagues (B. Albert et al.), who describe what is known of the role of RNA polymerase I elongation in regulating rRNA production.

Overall, the articles in this special issue describe the present state of transcription elongation. Ongoing advances in this field will certainly propel our capacity to more deeply understand gene expression and genome dynamics. In this regard, these reviews all provide a perspective into the future of gene control during transcription elongation.

Sebastián Chávez
David S. Gross
Damien Hermand
Carlos Suñé

Hindawi Publishing Corporation
Genetics Research International
Volume 2012, Article ID 347214, 19 pages
doi:10.1155/2012/347214

Review Article

Emerging Views on the CTD Code

David W. Zhang,[1] Juan B. Rodríguez-Molina,[1] Joshua R. Tietjen,[1] Corey M. Nemec,[1] and Aseem Z. Ansari[1,2]

[1] *Department of Biochemistry, University of Wisconsin-Madison, 433 Babcock Drive, Madison, WI 53706, USA*
[2] *Genome Center of Wisconsin, University of Wisconsin-Madison, 433 Babcock Drive, Madison, WI 53706, USA*

Correspondence should be addressed to Aseem Z. Ansari, ansari@biochem.wisc.edu

Received 13 September 2011; Accepted 3 November 2011

Academic Editor: David Gross

The C-terminal domain (CTD) of RNA polymerase II (Pol II) consists of conserved heptapeptide repeats that function as a binding platform for different protein complexes involved in transcription, RNA processing, export, and chromatin remodeling. The CTD repeats are subject to sequential waves of posttranslational modifications during specific stages of the transcription cycle. These patterned modifications have led to the postulation of the "CTD code" hypothesis, where stage-specific patterns define a spatiotemporal code that is recognized by the appropriate interacting partners. Here, we highlight the role of CTD modifications in directing transcription initiation, elongation, and termination. We examine the major readers, writers, and erasers of the CTD code and examine the relevance of describing patterns of posttranslational modifications as a "code." Finally, we discuss major questions regarding the function of the newly discovered CTD modifications and the fundamental insights into transcription regulation that will necessarily emerge upon addressing those challenges.

1. Introduction

The transcription of DNA to RNA in eukaryotes is catalyzed by three structurally related RNA polymerases, with each acting on a different class of genes [1]. RNA polymerase I synthesizes most of the ribosomal RNA (rRNA) subunits while RNA polymerase III synthesizes tRNAs, 5S rRNA, and other small RNAs [2–4]. These two polymerases account for 75% and 15% of transcription in the cell, respectively [5]. However, the most studied polymerase is RNA Polymerase II (Pol II), which is responsible for the transcription of protein-coding genes, small nuclear RNA (snRNA), and small nucleolar RNA (snoRNA) [6–8]. In higher eukaryotes, Pol II generates long noncoding RNA (lncRNA) and microRNA (miRNA) [9, 10]. Pol II also transcribes cryptic unstable transcripts (CUTs) and stable unannotated transcripts (SUTs), which are degraded after synthesis [11–13]. The suppression of CUTs is important to prevent inappropriate transcription within ORFs, to enhance processivity during transcription elongation, and to prevent gene silencing via histone deacetylation [14–18].

Of the twelve Pol II subunits, five are common between the three polymerases [1, 19–21]. It is believed that the specific functions attributed to each polymerase arise from the combined action of remaining nonidentical subunits and other factors that associate with them. An especially unique feature of Pol II is the carboxy-terminal domain (CTD) of its large subunit Rpb1 (Figure 1(a)). The CTD serves as the primary point of contact for a wide variety of molecular machines involved in RNA biogenesis during the transcription cycle (reviewed in [8, 22–32]). This domain consists of a highly conserved heptapeptide repeat: $Y_1S_2P_3T_4S_5P_6S_7$ [33–36]. The number of times this sequence is repeated varies among eukaryotic organisms, ranging from 15 repeats in amoeba, to 26 repeats in the budding yeast *Saccharomyces cerevisiae,* to 52 repeats in humans. When fully extended, the yeast CTD can span a distance of up to 650 Å, over 4 times the diameter of the core polymerase (Figure 1(b)) [24, 34, 35]. The ability of this repetitive sequence to interact with a wide range of nuclear factors stems from the dynamic plasticity of its structure and the diversity of binding surfaces generated by the multitude of post-translational

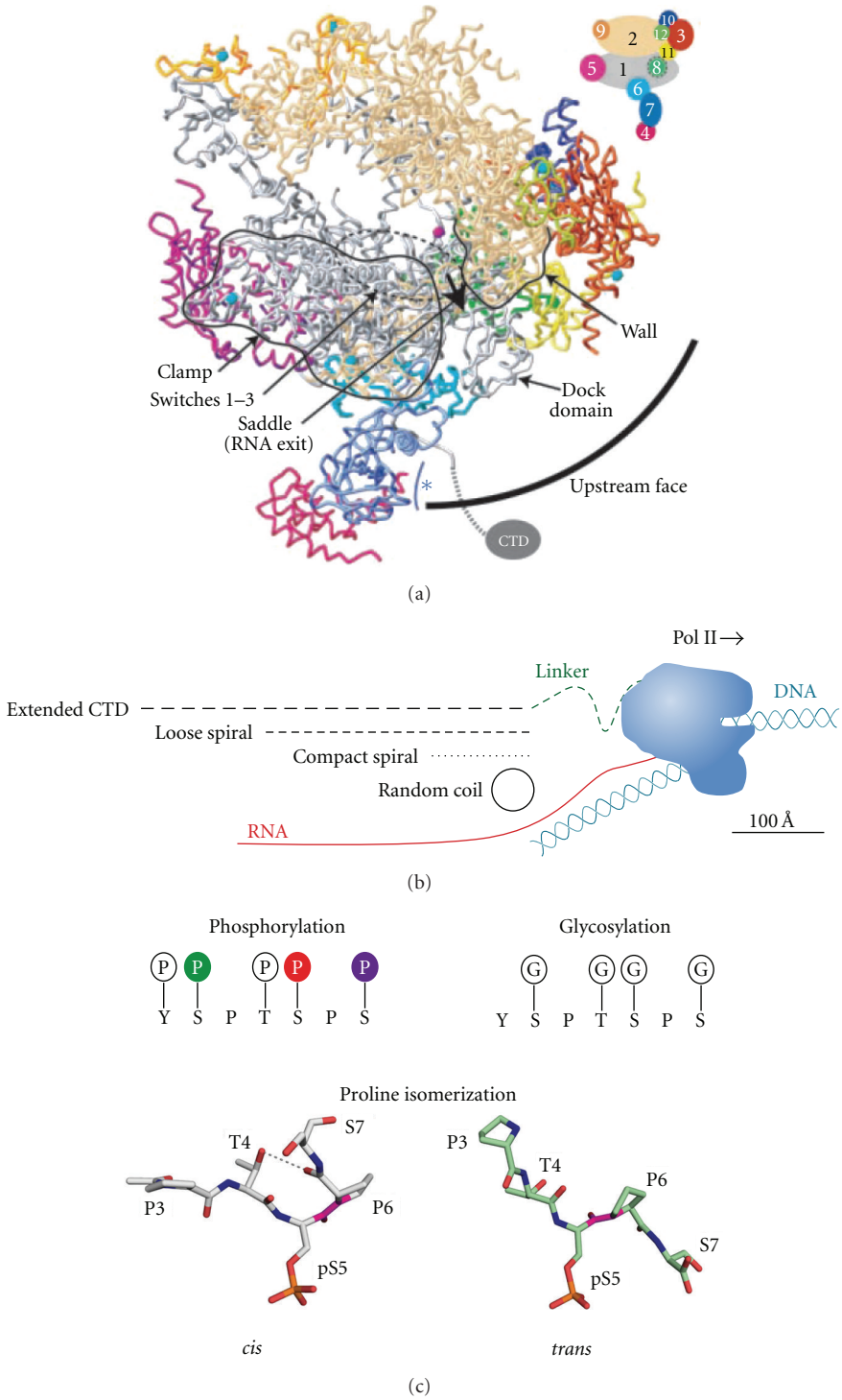


Figure 1: RNA polymerase II structure. (a) Side view of the core Pol II crystal structure containing all twelve subunits and displaying the RNA exit channel (bold arrow) and the positioning of the CTD adapted from Armache et al. [71]. Cartoon in the upper right displays the color coding for the Pol II subunits used in the crystal structure. (b) Illustration of the relative length(s) between the CTD in various conformations and the core Pol II adapted from Meinhart et al. [72]. RNA positioning (red) upon exit of the Pol II and the positioning of the DNA template (blue) upstream and downstream of the core Pol II are also displayed. (c) Known modifications possible on the Pol II CTD are displayed. Glycosylation and phosphorylation are mutually exclusive modifications. Structural images of a heptad repeat in the *cis*- and *trans*-conformation are also shown [73–75]. G: β-O-linked N-acetylglucosamine [76]; P: O-linked phosphate.

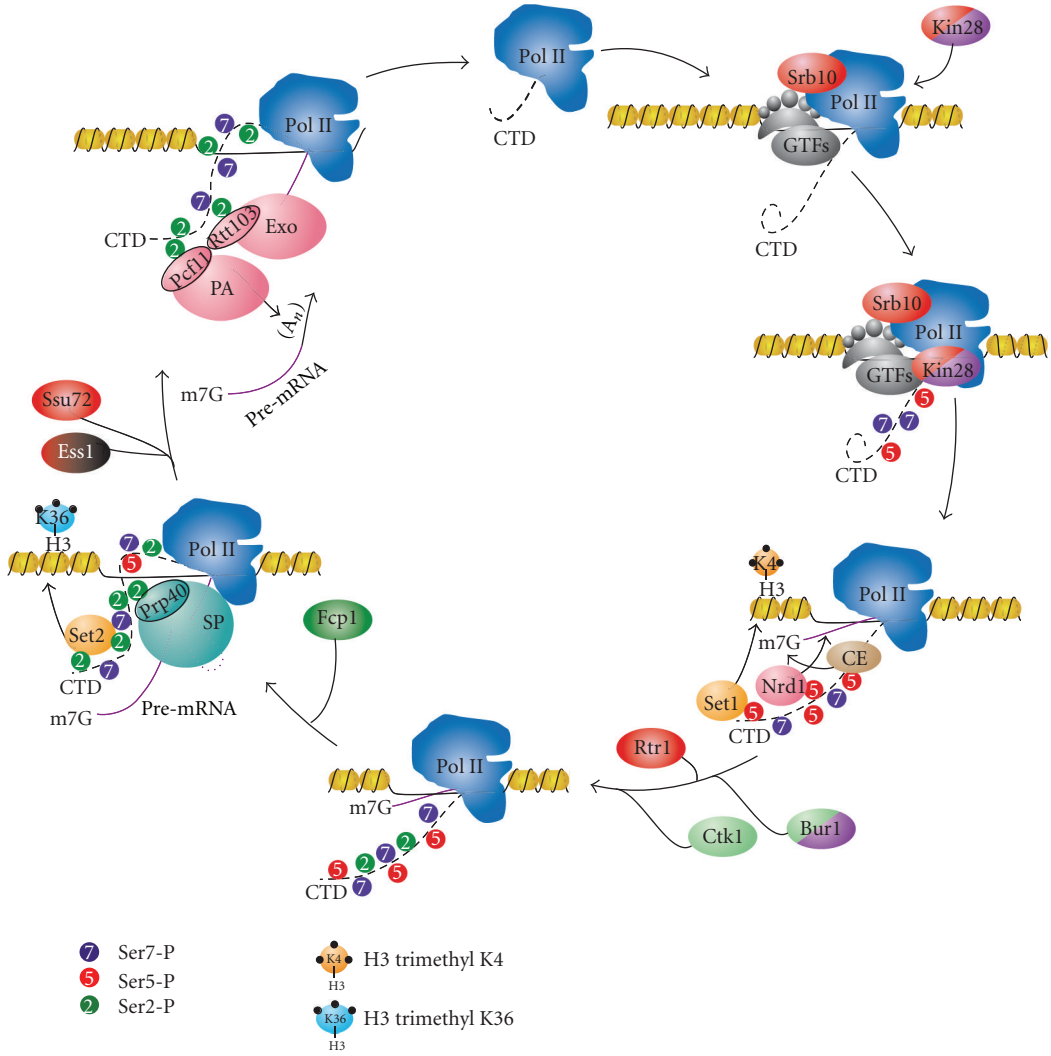


Figure 2: The primary components of the RNA biogenesis machinery and their interactions with the RNA polymerase II C-terminal domain (CTD). Briefly, hypophosphorylated Pol II assembles at the preinitiation complex (PIC) with the Mediator and general transcription factors (GTFs), with TFIIH associating last. The TFIIH-associated kinase Kin28 phosphorylates Ser5 (shown in red) and Ser7 (shown in purple) on the CTD. Mediator-associated kinase Srb10 also contributes to the phosphorylation of Ser5-P. This mark enables promoter release and mediates interactions with the capping enzyme (CE) complex, Nrd1 component of termination machinery, and Set1 histone methyltransferase, which places trimethyl marks on histone H3K4. The Ser5-P mark also facilitates recruitment of Bur1 kinase. Bur1 places initial Ser2-P marks, which facilitate recruitment of Ctk1 kinase, and continues to replenish Ser7-P marks during elongation. Ctk1 is the primary Ser2 kinase, and its phosphorylation recruits splicing machinery (SP) through Prp40, as well as Set2 histone methyltransferase, which places di- and trimethyl marks on histone H3K36. Cleavage and polyadenylation (PA) machinery are recruited through many factors associating with the CTD. One of the factors, Pcf11, binds cooperatively to Ser2-P with Rtt103. The exonuclease complex (Exo) is also recruited through interaction between CTD and Rtt103 and through cooperative interaction between Rtt103 and Pcf11. Finally, the hypophosphorylated CTD is regenerated through three CTD phosphatases. Ser2-P is removed by the phosphatase Fcp1, while two phosphatases, Rtr1 and Ssu72, combine to remove Ser5-P marks during elongation and at termination, respectively. Upon dephosphorylation, Pol II is released with the assistance of a mechanism involving Pcf11 and can begin another cycle of transcription.

modifications it can accommodate. Tyrosine, threonine, and three serines can all be phosphorylated, the threonine and serine can be glycosylated, and the prolines can undergo isomerization (Figure 1(c)) [27, 37, 38]. In humans, CTD repeats further away from core Pol II bear noncanonical repeats that can be methylated [39]. Taken together, at least 10^{59} unique modification patterns can occur on the CTD. The combinatorial nature of these modifications, which is reminiscent of the histone code, led to the hypothesis of a CTD code, where the patterns of modifications are read by the transcriptional machinery and these patterns dictate the association or disassociation of complexes [40, 41]. To date, much effort has been made towards characterizing these modifications and understanding the interactions between the CTD and components of various protein machines that play a role in RNA biogenesis. Our current knowledge of the integration of these events by Pol II CTD is summarized in Figure 2, and the known yeast CTD-interacting factors are

Table 1: Proteins known to bind RNA polymerase II C-terminal domain in *S. cerevisiae*.

Protein/complex	Role in RNA biogenesis	Phospho-CTD bound	References
TFIIE	Preinitiation complex	Hypophosphorylated CTD	[63, 77]
TFIIF	Preinitiation complex	Hypophosphorylated CTD	[77]
TBP	Preinitiation complex (TFIID)	Hypophosphorylated CTD	[78]
Mediator Complex	Transcription activation/repression	Hypophosphorylated CTD	[48, 79]
Ceg1	Capping	Ser5-P	[80–85]
Abd1	Capping	PCTD	[83]
Set1	Histone methylation	Ser5-P	[86]
Rpd3C(Rco1)	Histone deacetylation	Ser2-P + Ser5-P	[87, 88]
Spt6	Histone chaperone	Ser2-P	[89]
Nrd1	Transcription termination/processing	Ser5-P	[90]
Sen1	Transcription termination/processing	Unknown	[91]
Asr1	Pol II ubiquitylation (Rpb4/7 Ejection)	Ser5-P	[92]
Ess1	Proline isomerase	Ser2-P	[93, 94]
Set2	Histone methylation	Ser2-P + Ser5-P	[95, 96]
Prp40	Splicing	PCTD	[97]
Npl3	Promotes elongation/prevents polyadenylation	Ser2-P	[98]
Pcf11	Cleavage/polyadenylation (CF1A)	Ser2-P	[99, 100]
Rna14	Cleavage/polyadenylation (CF1A)	PCTD	[101]
Rna15	Cleavage/polyadenylation (CF1A)	PCTD	[101]
Ydh1	Cleavage/polyadenylation (CPF)	PCTD	[102]
Yhh1	Cleavage/polyadenylation (CPF)	PCTD	[103]
Pta1	Cleavage/polyadenylation (CPF)	Ser5-P	[104]
Rtt103	5′-3′ Exonuclease (Rat1)	Ser2-P	[105]
Sus1	mRNA export	Ser5-P	[106]
Yra1	mRNA export	Hyperphosphorylated CTD	[107]
Rsp5	Pol II ubiquitylation (DNA damage response)	Ser2-P	[108, 109]
Hrr25	DNA damage repair	PCTD	[24, 110]

CTD-interacting proteins, the processes they are involved in, the phosphorylation state of the CTD with which they associate, and where in the literature the interaction is documented. Ser2-P refers to phosphorylated serine 2, Ser5-P refers to phosphorylated serine 5, and PCTD refers to a mixed phosphorylation state generated by *in vitro* phosphorylation of a CTD peptide with cell extracts. Additional protein-CTD interactions are described [110] but have not been directly tested.

displayed in Table 1. The focus of this paper is to highlight the recent advances in our understanding of the role of CTD in the early stages of the Pol II transcription cycle, expand on the concept of the CTD code hypothesis, and address the current questions and challenges within the field.

1.1. RNA Pol II Transcription Cycle

1.1.1. Transcription Initiation. Initiation of transcription begins with the recruitment of gene-specific transcription factors (TFs), general transcription factors (GTFs), the Mediator complex, and Pol II. These factors self-assemble into a pre-initiation complex (PIC) at the promoters of Pol II-transcribed genes [29, 32]. Recognition of the promoter is only partially understood, but it is believed to occur via the recognition of the various *cis*-elements in the promoter region, such as the TATA box. Binding generally occurs within upstream nucleosome-free regions—the DNA centered over promoters flanked by well-positioned nucleosomes [42–45]. There are two main models for how these factors assemble at this region: the sequential model and the holoenzyme model (Figure 3). In both models, TFs first bind at the upstream activating/repressing sequences (UAS/URS) and recruit the transcriptional machinery. In the sequential model, TBP/TFIID/SAGA assembly at the promoter is accompanied by TFIIA, followed by TFIIB [46, 47]. Then, the Mediator complex arrives, connecting the PIC to transcription factors assembled at the UAS/URS [48–51]. This massive complex consists of three large modules known as the head, middle, and tail and an additional kinase module containing a cyclin-dependent kinase (Srb10 in yeast, Cdk8 in metazoans) [52–57]. The Mediator complex is important for basal transcription and plays a central role in facilitating communication between transcription factors bound to regulatory elements and the PIC [49–51, 56–60]. However, there are studies that suggest the Mediator is not present at most genes, and it only associates with a few UAS/URS in an activator- and stress-specific manner [61, 62]. Pol II is then recruited, followed by the last GTF, TFIIH, which is brought to the PIC by TFIIE [63]. It is possible that several pathways of ordered recruitment exist for GTFs. Other components, including Pol II, TFIIE, and TFIIH, may

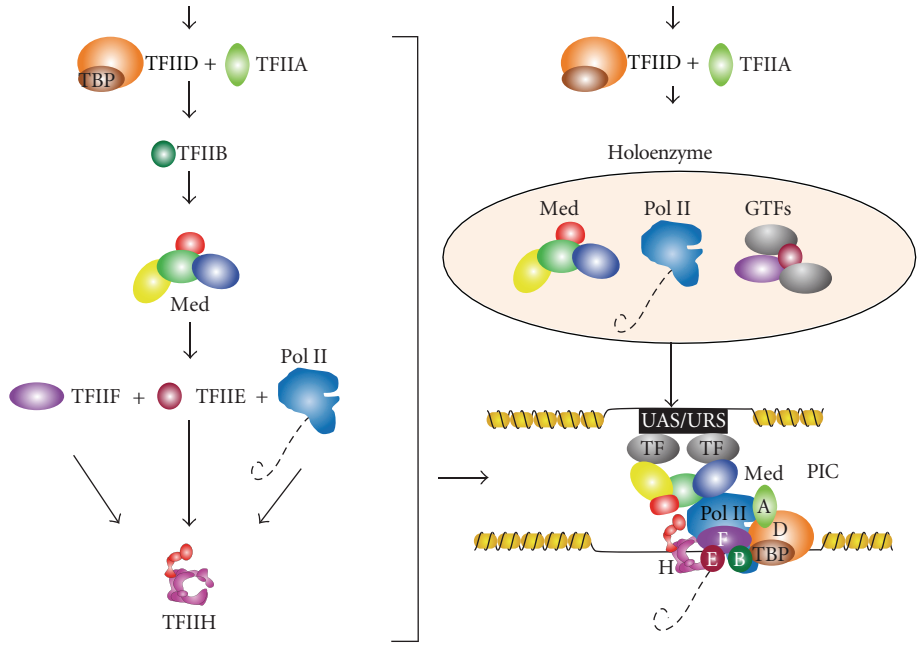


Figure 3: Recruitment and composition of PIC components. Sequential recruitment of the Mediator complex, GTFs, and Pol II (left) or the recruitment of the Pol II holoenzyme (top right), which assembles the pre-initiation complex (PIC) at promoters (bottom right).

be recruited via interactions with the Mediator [64]. The holoenzyme model originated from the observation that Srb proteins, which are components of the Mediator, are tightly associated with core Pol II in the absence of DNA [65]. In this model, Pol II is associated with the Mediator and other general transcription factors as a massive holoenzyme supercomplex that is recruited immediately after TBP binds [66–68]. These complexes have been identified in yeast and mammalian systems [69]. Importantly, Pol II is fully able to activate transcription upon arrival in this state [68, 70].

Two complexes of the PIC, TFIIH and the Mediator, contain important kinases that phosphorylate the CTD. TFIIH is a ten-subunit complex containing two helicases, an ATPase, a ubiquitin ligase, a neddylation regulator, and a cyclin-dependent kinase (Kin28 in yeast, Cdk7 in metazoans) [111–118]. Both Kin28/Cdk7 and Srb10/Cdk8 have been shown to phosphorylate Ser5 (Ser5-P) *in vivo*, with Kin28/Cdk7 being the dominant kinase [24, 113, 119–124]. The 5′ enriched Ser5-P mark has been linked to a variety of chromatin-modifying and RNA processing events.

1.1.2. Transcription Elongation. Phosphorylation of Ser5 is involved in coordinating the placement of several key post-translational modifications on chromatin that constitute the histone code [41] (reviewed in [125–127]). The structural properties of chromatin, such as the +1 nucleosome that resides immediately after gene promoters, are thought to provide a significant physical barrier to transcription. This barrier is weakened or removed through the combined action of posttranslational modifications on the flexible histone tails and chromatin remodeling complexes [127]. In this context, the Ser5-P mark recruits the yeast histone methyltransferase Set1. Trimethylation of histone H3K4 by Set1 and subsequent trimethylation of H3K79 by Dot1 are frequently associated with active transcription and have a reciprocal effect on H3K14 acetylation by SAGA and NuA3 [28, 86, 128, 129]. Ser5-P also recruits the histone deacetylase complexes Set3 and Rpd3C(S) [87], which are important for suppressing CUT initiation at promoters [87, 88].

An especially important role of Ser5-P is the recruitment of the capping enzyme complex. The capping complex places the m^7G cap on the nascent transcript as it exits the core polymerase, stabilizing the mRNA by preventing its degradation by 5′-3′ exonucleases. The CTD repeats proximal to the core Pol II are ideally placed near the RNA exit tunnel to facilitate this capping reaction [130, 131]. The guanylyltransferase (Ceg1 in *S. cerevisiae*) and possibly the methyltransferase (Abd1 in *cerevisiae*) directly interact with both the Ser5-P and the core polymerase [80–85, 132, 133]. Although the recognition of the CTD is structurally different between yeast and mammalian capping enzymes, both complexes require Ser5-P for binding [81, 131]. A parallel line of experiments showed that inhibition of Kin28 kinase activity using a small-molecule inhibitor leads to a severe reduction in Ser5-P and 5′-capping of transcripts at gene promoters [134, 135]. In agreement with this, tethering the mammalian capping enzyme to the CTD rescues the null Ser5 to alanine mutants in the fission yeast *Schizosaccharomyces pombe* [136]. Interestingly, inactivation of Kin28 does not eliminate transcription: neither steady-state mRNA levels nor the ability to initiate transcription at the inducible *GAL1* gene is significantly compromised by the inhibition [135]. A subsequent study using the same chemical inhibition system confirmed the earlier observations but incorrectly attributed small differences in transcript levels to inappropriate normalization of earlier microarray

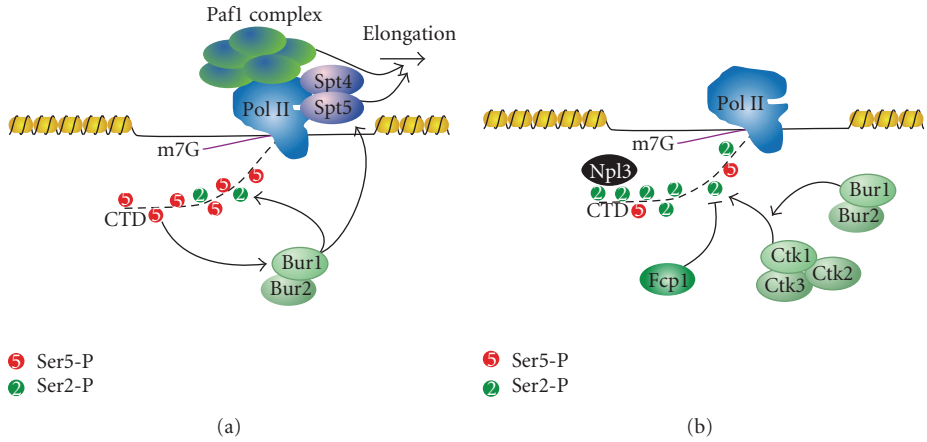


Figure 4: Bur1 phosphorylation of the CTD facilitates the transition from initiation to elongation. (a) Ser5-P enhances recruitment and subsequent phosphorylation of Ser2 by Bur1. Bur1 also phosphorylates Spt5, which acts with the Paf1 complex to promote elongation. (b) CTD phosphorylation by Bur1 enhances the activity of Ctk1 on Ser2. The majority of the Ser2-P is maintained by competition between phosphorylation by Ctk1 and dephosphorylation by Fcp1. This increase in Ser2-P facilitates recruitment of many Ser2-P-binding proteins, such as Npl3.

data [137]. No such global normalization was performed by Kanin et al. [135] and it is unclear why the subsequent study [137] made the unsubstantiated and erroneous claim that the data was treated incorrectly. Kanin et al. were quite cognizant of the consequences of inhibiting an enzyme that could have a role in global transcription. Moreover, quantitative PCR and northern blot assays, experiments that were not reliant on microarray normalization, showed little difference in expression (Hein and Ansari, 2007, unpublished data) [135]. These results strongly support the conclusion that inactivating Kin28 does not significantly impact global transcription. It is important to note that these studies only focused on chemical inhibition of Kin28 and that the inhibition is not an "all or none" phenomenon due to equilibrium binding of the small molecule to the kinase; it is possible that extremely low levels of Ser5 phosphorylation, by either Srb10 or residual Kin28, suffice for transcription initiation. Importantly, chemical inhibition of both Kin28 and Srb10 shows a drop in Pol II across the ORF, supporting the model where Ser5-P may help in promoter clearance [138].

We and others have recently demonstrated that Kin28/Cdk7 is also the primary kinase that phosphorylates Ser7 (Ser7-P) [139–141]. The phosphorylation occurs at protein-coding and noncoding genes and seems to be Mediator dependent [142]. Cyclin-dependent kinases are thought to prefer a substrate bearing Ser-Pro rather than Ser-Tyr dipeptides [143]. Additionally, while Kin28 has been localized to promoters [83], Ser7-P marks were thought to be found only at non-coding genes and at the 3′ end of protein coding genes [144, 145]. The role of Ser7-P at promoters remains an active area of investigation.

Following promoter clearance, transcription initiation factors are exchanged for transcription elongation factors required for RNA processing, passage through chromatin, and suppressing cryptic transcripts. In budding yeast, this exchange occurs immediately after the +1 nucleosome [146]. The association of these elongation factors, which include Paf1, Spt16, Spt4, Spt5, Spt6, Spn1, and Elf1, occurs concurrently on all Pol II genes and is independent of gene length, type, or expression [146]. The recruitment of these factors is essential for transcription processivity (Spt4/5) [147–149], histone regulation (Spt6/16, Spn1, Elf1) [150–156], and gene activation/3′ processing (Paf1) [157]. Similarly, mammalian P-TEFb complex is recruited to Pol II at this stage of transcription [158–161]. This complex contains a cyclin-dependent kinase (Cdk9) that phosphorylates the DRB-sensitivity-inducing factor (DSIF), which allows Pol II to overcome the promoter-proximal pausing induced by the negative elongation factor (NELF) complex [23, 159]. It is unclear if promoter-proximal pausing occurs in yeast, but it is known that Bur1 (the yeast homolog of Cdk9) promotes elongation through post-translational modification of Spt5 (DSIF) (Figure 4(a)) [162]. Bur1 also improves transcription elongation through the recruitment of histone-modifying enzymes and the phosphorylation of CTD. Bur1 activity promotes the ubiquitylation of H2BK123 by the ubiquitin conjugating enzyme Rad6 and Bre1 [129, 163]. H2BK123Ub promotes Set1 trimethylation of histone H3K4 and subsequent trimethylation of H3K79, both of which are important for transcription activation [28, 86, 128, 129]. Bur1 also promotes transcription elongation by coupling promoter-proximal CTD modifications with promoter-distal marks. Bur1 is recruited to the transcription complex by the Ser5-P marks placed at the promoter. It then phosphorylates Ser2 (Ser2-P), priming the CTD for the recruitment of Ctk1 (Cdk12), the major Ser2 kinase [164]. Initial CTD phosphorylation also increases the activity of Ctk1, thereby coupling sequential CTD modifications (Figure 4(b)) [23, 159, 165, 166]. Interestingly, Bur1 travels with Pol II and phosphorylates Ser7-P. Although the exact role of this modification is unclear, it is likely a mark that

promotes elongation, as genes with uniformly high levels of Ser7-P are transcribed at significantly higher levels [138].

Most Ser5-P marks are removed near the +1 nucleosome through the action of the newly characterized CTD phosphatase Rtr1 [167]. This phosphatase has been shown to specifically remove Ser5-P marks immediately after promoter clearance. The Ser2-P phosphatase Fcp1 also associates during elongation, but Ser2-P levels remain high across the transcript due to the opposing action of the Ser2-P kinase Ctk1 [168, 169]. It is thought that the Ubp8 component of SAGA travels with Pol II and promotes deubiquitylation of H2BK123Ub [170], which allows the association of Ctk1 and subsequent phosphorylation of Ser2 on the CTD [171].

Ser2-P is critically important for the interaction between the CTD and many histone modifying and RNA processing machines [75, 83, 132, 172–178]. Increasing levels of Ser2-P, in combination with the residual Ser5-P, promote the recruitment of the Set2 methyltransferase, which catalyzes the formation of H3K36me2 and H3K36me3 [95, 96, 179–181]. This leads to the recruitment of the histone deacetylase complex Rpd3C(S) and the removal of acetylation from histones H3 and H4, thereby resetting the transcription state of the nucleosomes and repressing cryptic transcription within ORFs [87, 182, 183]. Ser2-P is involved in the cotranscriptional and posttranscriptional processing of RNA. Cotranscriptional processing of introns via splicing involves the yeast protein Prp40, which preferentially associates with Ser2-P/Ser5-P marked CTD [97]. Ser2-P is also bound by the SR-like (serine/arginine rich) protein Npl3, which functions in elongation, 3′-end processing, hnRNP formation, and mRNA export [184–187]. Finally, increasing levels of Ser2-P, coupled with depletion of Ser5-P, leads to the recruitment of the termination and polyadenylation machinery (discussed below).

1.1.3. Transcription Termination. The role of CTD modifications in orchestrating transcription termination is better described in recent reviews [31, 188]. In essence, two models have been proposed to explain how Pol II termination occurs, with the emerging view being that it is likely a combination of the two models that best describes the mechanism. The first model, known as the "allosteric" or "antiterminator" model, proposes that transcription through the polyadenylation site leads to an exchange of elongation factors for termination factors, resulting in a conformational change of the elongation complex. Indeed, this model is supported by chromatin immunoprecipitation (ChIP) data of elongation factor exchange at the 3′ end of genes [146, 189]. The second model, known as the "torpedo" model, postulates that cleavage of the transcript at the cleavage and polyadenylation site (CPS) creates an entry site for the 5′-3′ exonuclease Rat1 (Xrn2 in mammals), which degrades the 3′ RNA and promotes Pol II release by "torpedoing" the complex [189–191]. In this model, recruitment of Rat1 is likely to be indirect, possibly through its partner Rtt103. Rtt103 has been shown to bind Ser2-P in a cooperative manner with Pcf11 [192], an essential component of the cleavage factor IA (CFIA) complex that also promotes Pol II release [193]. Interestingly, ChIP data shows Pcf11 at both protein-coding and noncoding genes, and mutating Pcf11 results in terminator read-through due to inefficient cleavage at both gene classes [75, 174, 193–196]. Pcf11 may play an important role in both the termination and processing of protein-coding and non-coding genes.

Processing of Pol II transcripts occurs via one of two distinct, gene class-specific pathways in yeast. Many small mRNAs (<550 bp), CUTs, snRNA, and snoRNAs (non-coding genes) are processed via the Nrd1-Nab3 pathway (Figure 5), while longer mRNAs (protein-coding genes) are processed in a polyadenylation-dependent process (Figure 6) [8, 11, 12, 27, 31, 178, 195, 197–199]. The decision to proceed down a certain processing path is modulated by the phosphorylation state of the CTD. Nrd1 preferentially associates with Ser5-P, and its recruitment is also enhanced via histone H3K4 trimethylation by Set1 [90, 200]. Nrd1 and Nab3 scan the nascent RNA for specific sequence elements (GUAA/G or UGGA for Nrd1, and UCUU or CUUG for Nab3) as it exits the core polymerase [90, 199, 201–207]. The helicase Sen1 (senataxin in humans), which exists in complex with Nrd1 and Nab3, resolves the DNA:RNA hybrids known as R-loops that form between the template DNA and the nascent RNA, keeping the specific sequence elements exposed and preserving genomic stability [208–210]. The involvement of Sen1 is dependent on the phosphatase Glc7, which dephosphorylates Sen1 and is essential for the proper termination of snRNA and snoRNA transcripts [211]. Upon detecting its consensus sequence elements, the Nrd1 complex and the Rnt1 endonuclease cleave these short transcripts [195, 212–214], which are then trimmed at the 3′ end by the TRAMP complex and the exosome [6, 215–217]. Nrd1 then disengages from the transcription complex, with help from antagonizing Ser2-P marks [198]. Unlike snRNA/snoRNAs, which have protective structural elements in the RNA, Nrd1-terminated CUTs have no protective elements at their 3′ ends and are thus fully degraded by TRAMP after cleavage [8, 11, 12]. Nrd1 has been mapped to the 5′ end of transcribed regions, but a recent study has demonstrated that Nrd1 occupancy is maintained across the open reading frame of genes [196]. Although no homolog of Nrd1 has been found in mammalian cells, the Integrator complex that is involved in 3′ processing of snRNA transcripts is recruited by Ser7-P [218]. The association of this complex with Ser7-P CTD was demonstrated by the abolishment of this interaction upon mutation of Ser7 to alanine [145]. Subsequent analysis using a panel of CTD peptides determined that the Integrator prefers to bind a diphosphorylated CTD substrate spanning two heptad repeats in the S7-P-S2-P conformation [219]. It is possible that Ser7-P may serve as a similar scaffold for snRNA and snoRNA processing machinery in yeast.

The second pathway, used for the processing of most mRNA transcripts, involves the cleavage and polyadenylation factor (CPF) complex, cleavage factor IA and IB (CFIA and CFIB) complexes, and the exosome (Figure 6) [31, 195, 197]. Many of the termination and 3′ processing factors involved in this process are known to preferentially associate with Ser2-P or Ser2-P/Ser5-P enriched CTD including: Npl3, Rtt103, Rna14, Rna15, Ydh1, Yhh1, Pta1, and Pcf11. In this pathway, Rna15 competes with Npl3 for recognition of a

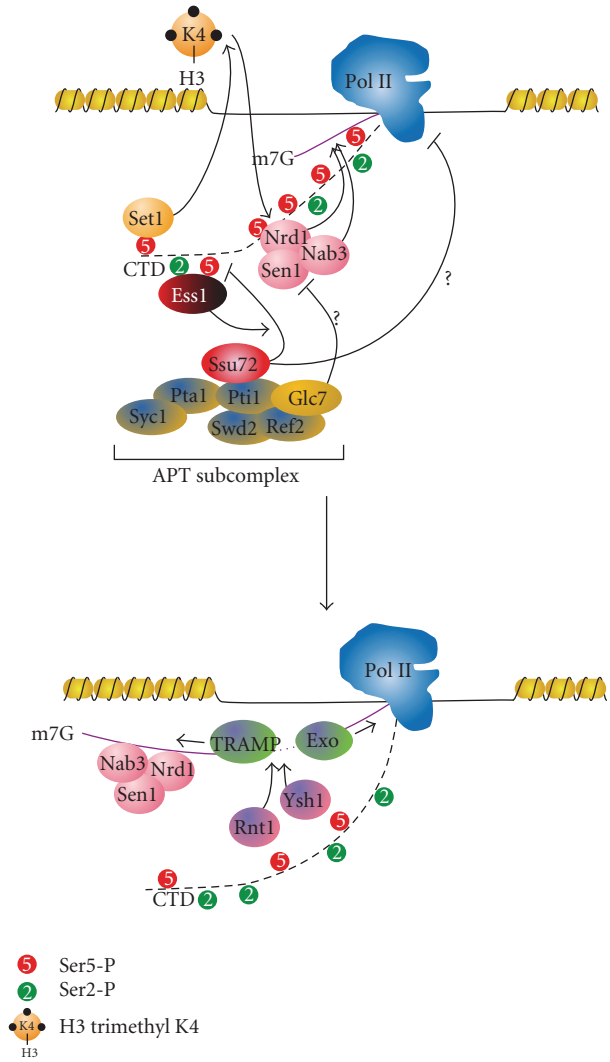


Figure 5: Nrd1-dependent termination pathway. The Nrd1-Nab3-Sen1 complex is recruited via interaction between Nrd1 and Ser5-P. This recruitment is facilitated by H3K4me3, which is placed by the Set1 histone methyltransferase. The mechanisms by which the Ssu72 and Glc7 phosphatases promote termination are still unclear, but it may be that the dephosphorylation of Sen1 by Glc7 and of the CTD by Ssu72 causes the polymerase to pause, and allowing the termination machinery to associate. During elongation, both Nrd1 and Nab3 scan the nascent RNA for their preferred sequences (see text for details). Upon finding their concensus sequences, Nrd1-Nab3-Sen1 complex is able to be associated with the RNA. The endonucleases Rnt1 and Ysh1 may contribute to the cleavage of the RNA, which is followed by 3′-5′ trimming the transcript by the TRAMP complex and by the degradation of the remaining RNA exiting Pol II by the 5′-3′ exonuclease Rat1 (Exo).

UA-rich site in the nascent RNA [98, 187]. This competition is removed upon phosphorylation of Npl3 by casein kinase 2 (CK2) [98]. Rna15 can then bind the nascent RNA and promote endonucleolytic cleavage followed by polyadenylation by the polyadenylate polymerase (Pap1) [197, 220]. Polyadenylation-binding proteins (PAB) then protect the mature transcript from exonucleolytic degradation (Figure 6) [221].

In both pathways, the CTD is hypophosphorylated by the combined action of two essential phosphatases at the end of transcription: Ssu72 and Fcp1. Ssu72 is a member of the Associated with Pta1 (APT) complex, which is present at both gene classes and is involved in 3′ processing of non-coding RNAs [222]. As such, Ssu72 is primarily localized at the 3′ end of transcripts [222], although there is one instance in which it has been found at promoters [223]. Temperature-sensitive mutants of Ssu72 exhibit read-through at both protein-coding and non-coding transcripts [224]. Ssu72 is the primary Ser5-P phosphatase [225], and its phosphatase activity is enhanced by the prolyl isomerase Ess1/Pin1 and by interacting with Pta1/symplekin [226–228]. Recently, crystal structures have shed light on the mechanism of Ssu72: the phosphatase binds to Ser5-P only when the adjacent Pro6 is in the *cis*-conformation [73, 74]. In contrast to Ssu72, Fcp1 associates with TFIIF during transcription and is found across the entire transcribed

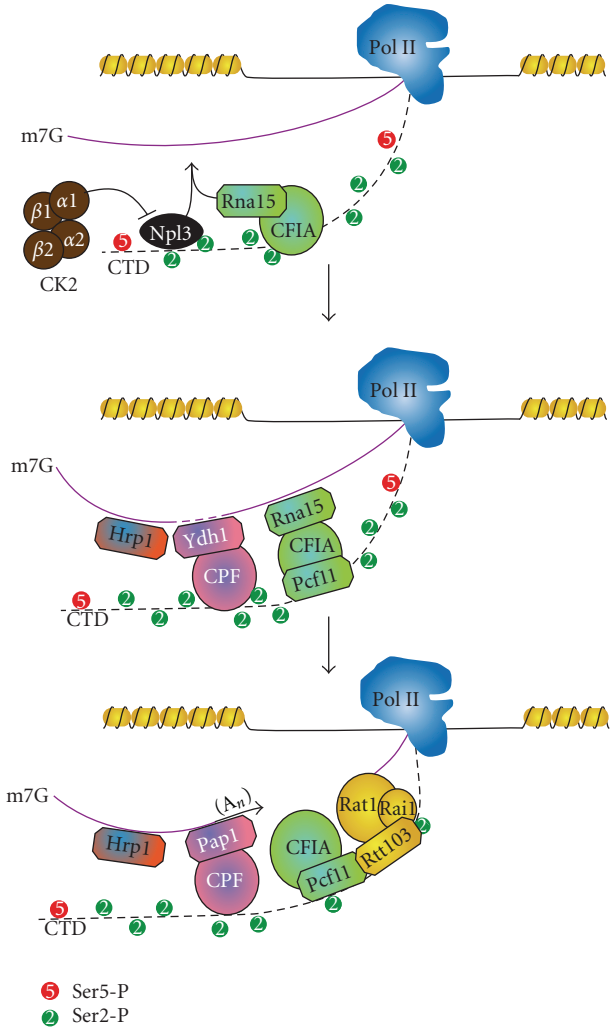


FIGURE 6: The mRNA termination pathway. Rna15 competes with Npl3 for binding to the nascent RNA. CK2 phosphorylates Npl3, allowing Rna15 to find its preferred binding site (an A/U-rich region) in the RNA. The CPF and CFIA components assemble through interactions with the CTD and the Yth1 component of CPF cleaves the nascent RNA at the polyadenylation site, followed by polyadenylation by Pap1. Then the Rat1 exonuclease complex associates via cooperative interaction between Pcf11 and Rtt103 and leads to termination and dissociation of Pol II.

region [168, 169, 229, 230]. Although it has Ser5-P and Ser2-P phosphatase activity *in vitro*, Fcp1 is considered a Ser2-P-specific phosphatase *in vivo* [231, 232]. Fcp1 activity is enhanced upon phosphorylation of Fcp1 by CK2 [233]. Defects in Fcp1 also result in transcription read-through at Nrd1-dependent transcripts [198]. Though it is unclear which phosphatase removes Ser7-P, new data from our lab suggest that Ssu72 may be the phosphatase that removes Ser7-P at both the 5′ and 3′ ends of genes [234]. Removal of this mark may be even more important than its placement as mutation of Ser7 to alanine slows growth while mutating Ser7 to the phosphomimic glutamate is lethal [144].

Global dephosphorylation of the CTD facilitates the release of Pol II from DNA, which can then recycle to promoters for the next cycle of transcription [224, 235, 236]. It has been proposed that transcription termination and subsequent dephosphorylation of the CTD is coupled to transcription reinitiation through gene looping, by which the promoter and terminator regions are brought together, allowing Pol II to associate more rapidly with the PIC [237, 238]. Intriguingly, Ssu72 and the GTF TFIIB have been shown to be essential in gene looping [223, 239]. Taken together, the phosphorylation and dephosphorylation of the CTD is intimately involved in every phase of transcription, from initiation, to elongation, to termination, and possibly reinitiation.

1.1.4. Other Regulatory Roles of the CTD. In addition to its many roles in transcription initiation, elongation, and termination, the CTD has been implicated in a variety of transcription-extrinsic processes, such as mRNA export and stress response. mRNA export (reviewed in [240–242]) requires the packaging of the mRNA into export-competent messenger ribonucleoprotein (mRNP) via association with the

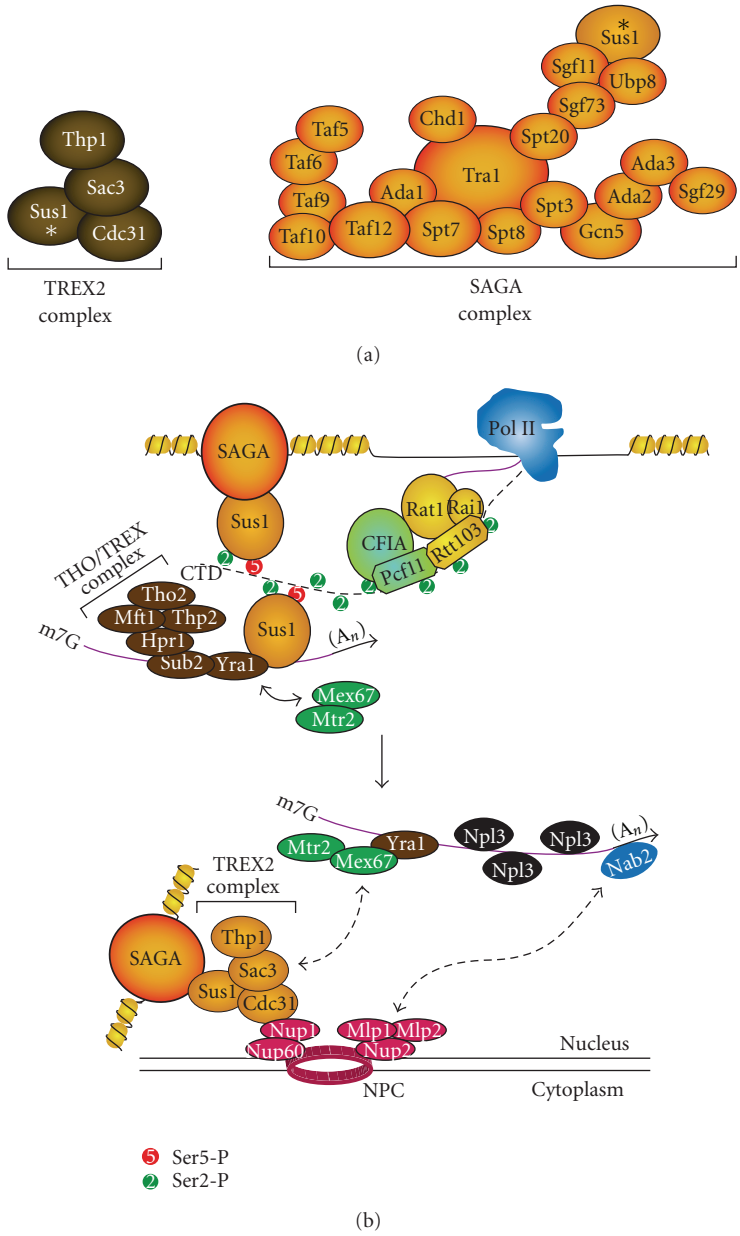


Figure 7: Sus1 in TREX2 and SAGA complexes coordinates mRNA export. (a) Subunit compositions of TREX2 and SAGA complexes are shown, highlighting Sus1 (asterisk). (b) mRNA export coordinated by Sus1. Sus1 binds Ser2-P and Ser2-P/Ser5-P CTD, connecting the CTD to the SAGA histone acetyltransferase complex. Sus1 also interacts with Yra1 component of the THO/TREX complex on the RNA. Mex67-Mtr2 are recruited by interaction with Yra1 and help form the export-competent mRNP. At the nuclear pore complex (NPC), Mlp1-Mlp2 interact with the polyA mRNA-binding protein Nab2 and Mex67 interacts with Sac3 of the TREX2 complex. This interaction brings the export-competent mRNP to the NPC in preparation for export to the cytoplasm. Sus1 is a component of both TREX2 and SAGA and serves to tether actively transcribed gene promoters to the NPC.

Mex67:Mtr2 heterodimer [243]. This heterodimer is brought to the mRNA by Yra1 and Sub2, components of the THO subunit of the TREX1 complex [242]. The process of mRNP export is coordinated by the protein Sus1. This central protein directly interacts with Ser5-P and Ser2-P/Ser5-P of the CTD, Ub8 subunit of the SAGA complex, Yra1 subunit of the TREX1 complex, and Sac3 subunit of the TREX2 complex at the nuclear pore (Figure 7) [106, 244].

The CTD is also involved in stress response. The ubiquitin ligase Rsp5 binds the CTD and ubiquitylates Pol II in response to DNA damage [245, 246]. Similarly, UV-induced DNA damage in mammalian fibroblasts results in hyperphosphorylation of the CTD by the mammalian positive transcription elongation factor b (P-TEFb), which then regulates Pol II ubiquitylation and subsequent degradation [247]. Under conditions not well understood, Ser5-P can

also recruit the Asr1 ubiquitin ligase, which ubiquitylates the Rpb1 and Rpb2 subunits of Pol II. This ubiquitylation promotes ejection of the Rpb4/7 heterodimer from the core polymerase and inactivates Pol II, which may provide a mechanism for stopping polymerases engaged in abortive or cryptic transcription [92].

2. The CTD Code Controversy: Is It a Code?

The concept of the CTD code was first proposed due to the enormous amount of information that can be encoded via post-translational modification of the CTD repeats [40, 248]. The code would coordinate the assembly of complexes that "read, write, and erase" the code during transcription. Historically, the Ser5-P and Ser2-P marks have been the best characterized, with the canonical distribution of Ser5-P being enriched at the 5′ end of genes and Ser2-P enriched towards the 3′ end. Recently, our lab and several others have been able to map the phospho-CTD occupancy profiles across the yeast genome [135, 138, 139, 146, 196]. There are interesting discrepancies between the observations made by various groups. For example, Mayer et al. find the canonical profile to be present at every gene with Ser7-P profiles overlapping with Ser5-P [146], while we find clusters of genes with noncanonical CTD profiles for Ser2-P, Ser5-P, and Ser7-P [138]. We observe gene-specific phosphorylation profiles, with Ser2-P levels being significantly lower at non-coding genes and Ser7-P profiles diverging from Ser5-P profiles only at protein-coding genes. The distinct patterns of CTD marks at these two gene classes reflect the different mechanisms of transcription termination and 3′ end processing machinery that act on these two classes of RNA. Similarly, Kim et al. also observe differences in phospho-CTD profiles at snoRNAs and at introns [196]. However, the positions of the Ser5-P and Ser7-P peaks in Kim et al. are offset from Tietjen et al. and Mayer et al. Importantly, all three genome-wide analyses reveal an unexpected degree of cooccurrence of CTD marks, suggesting a bivalent or even multivalent mode of recognition by docking partners. In support of this idea, the Set2 histone methyltransferase and the Integrator complex have been shown to prefer a bivalent mark rather than a single phosphorylated residue [95, 96, 219].

In addition to the various phosphorylation marks, the isomerization state of the CTD also contributes to the complexity of the code. For example, Pcf11 binds the CTD in the *trans*-conformation while Ssu72 prefers a *cis*-CTD as substrate [73–75]. Many in the transcription field have made the argument that the CTD code is not a true code because it does not convey biological information via a rigorous decoding key. However, research in the last several years has demonstrated that specific phosphorylation marks and proline isomerization are important for conveying information from *cis*-elements encountered by Pol II to the protein complexes necessary for successful progression through the transcription cycle. Further investigation into the mechanism of this information transfer will resolve the controversy over the existence of a CTD code.

3. Future Directions

Extraordinarily rapid progress has been made over the last several years in the field of CTD research; however, many important questions remain unanswered. Although the profiles of Ser7-P have been mapped and several of its kinases discovered, its function at protein coding genes remains unclear. Additionally, most of the kinases identified are established members of the transcription initiation or elongation complexes. One could expect to find new enzymes that could modulate the CTD in response to signals, as post-translational modifications are often used as a mechanism for cells to respond to external stimuli. The recent discovery of Ser7-P at elongating Pol II has also prompted the question of whether Tyr1 and Thr4 phosphorylation (Tyr1-P and Thr4-P) occurs? Tyr1 can be phosphorylated by c-Abl in mammals, but no homolog is present in yeast [249]. In addition, both Tyr1-P and Thr4-P has been detected in *S. pombe* [250]. Interestingly, Tyr1-P and Thr4-P were found in both the hyperphosphorylated and hypophosphorylated states of Pol II, opening the possibility of CTD function independent of transcription. However, neither the profile nor function of these potential modifications have been extensively characterized.

The role of non-canonical residues and their modification states on mammalian CTD remain to be explored. In mammals, the Ser7 residue is only weakly conserved in polymerase-distal repeats of the CTD, often changed to lysine or arginine [144]. Interestingly, Arg1810 of *rpb1* in the human CTD is methylated by the coactivator-associated methyltransferase1 (CARM1) [39]. This methylation occurs prior to both transcription initiation and phosphorylation of Ser2 or Ser5, and mutation of this residue results in the improper expression of a variety of snRNAs and snoRNAs. In addition to methylation, the CTD may also be subject to glycosylation. Recent studies suggest O-GlcNAc are transferred to Ser5 and Ser7 by O-GlcNac transferase and removed by O-GlcNAc aminidase during PIC assembly. This cycling of O-GlcNAc may be important for preventing aberrant CTD phosphorylation by TFIIH [251].

Besides the characterization of novel marks, significant structural challenges remain for understanding the known phosphomarks. One limitation of ChIP is its inability to identify the exact phosphorylation patterns across individual CTD repeats *in vivo* at different points during the transcription cycle. Recent mutational analysis suggests that the minimal functional unit of the CTD consists of three consecutive Ser-Pro dipeptide residues in a S2-S5-S2 configuration [36], but it is unclear if all three serines can be phosphorylated on one functional unit or if phosphorylation alternates between repeats. The lack of positively charged aminoacids makes the phospho-CTD patterns difficult to decipher via mass spectrometry. Additionally, the highly repetitive nature of the CTD makes it difficult to distinguish between the first repeat and the twenty-first. Consequently, the position along the CTD where interacting partners associate remains a mystery. Mutation of Ser2 to glutamate in the core-distal repeats and mutation of Ser5 to glutamate in the core-proximal repeats are lethal [252]. However, this

does not directly demonstrate whether the proteins that bind these phosphorylated residues are located at these repeats. Characterizing the phosphorylation patterns and protein occupancies at individual repeats will help determine the existence of a "CTD recognition" code, and this promises to be one of the most exciting and important challenges in the future of CTD research.

Acknowledgments

The authors gratefully acknowledge the support of the US National Science Foundation (MCB 07147), W. M. Keck, Shaw Scholar and Vilas Associate awards (to A. Z. Ansari). D. W. Zhang was supported by the Burris Pre-Doctoral Fellowship. J. R. Tietjen and J. B. Rodríguez-Molina were supported by a US National Human Genome Research Institute training grant to the Genomic Sciences Training Program (5T32HG002760). C. M. Nemec was supported by Chemistry Biology Interface Training Program (5T32GM008505-18).

References

[1] P. Cramer, K. J. Armache, S. Baumli et al., "Structure of eukaryotic RNA polymerases," *Annual Review of Biophysics*, vol. 37, pp. 337–352, 2008.

[2] I. Grummt, "Life on a planet of its own: regulation of RNA polymerase I transcription in the nucleolus," *Genes & Development*, vol. 17, no. 14, pp. 1691–1702, 2003.

[3] J. Russell and J. C. B. M. Zomerdijk, "RNA-polymerase-I-directed rDNA transcription, life and works," *Trends in Biochemical Sciences*, vol. 30, no. 2, pp. 87–96, 2005.

[4] G. Dieci, G. Fiorino, M. Castelnuovo, M. Teichmann, and A. Pagano, "The expanding RNA polymerase III transcriptome," *Trends in Genetics*, vol. 23, no. 12, pp. 614–622, 2007.

[5] M. Werner, P. Thuriaux, and J. Soutourina, "Structure-function analysis of RNA polymerases I and III," *Current Opinion in Structural Biology*, vol. 19, no. 6, pp. 740–745, 2009.

[6] F. Wyers, M. Rougemaille, G. Badis et al., "Cryptic Pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase," *Cell*, vol. 121, no. 5, pp. 725–737, 2005.

[7] C. A. Davis and M. Ares, "Accumulation of unstable promoter-associated transcripts upon loss of the nuclear exosome subunit Rrp6p in Saccharomyces cerevisiae," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 9, pp. 3262–3267, 2006.

[8] S. Lykke-Andersen and T. H. Jensen, "Overlapping pathways dictate termination of RNA polymerase II transcription," *Biochimie*, vol. 89, no. 10, pp. 1177–1182, 2007.

[9] M. Faller and F. Guo, "MicroRNA biogenesis: there's more than one way to skin a cat," *Biochimica et Biophysica Acta*, vol. 1779, no. 11, pp. 663–667, 2008.

[10] M. Guttman, I. Amit, M. Garber et al., "Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals," *Nature*, vol. 458, no. 7235, pp. 223–227, 2009.

[11] J. T. Arigo, D. E. Eyler, K. L. Carroll, and J. L. Corden, "Termination of cryptic unstable transcripts is directed by yeast RNA-binding proteins Nrd1 and Nab3," *Molecular Cell*, vol. 23, no. 6, pp. 841–851, 2006.

[12] M. Thiebaut, E. Kisseleva-Romanova, M. Rougemaille, J. Boulay, and D. Libri, "Transcription termination and nuclear degradation of cryptic unstable transcripts: a role for the nrd1-nab3 pathway in genome surveillance," *Molecular Cell*, vol. 23, no. 6, pp. 853–864, 2006.

[13] S. Y. Ying and S. L. Lin, "Intron-mediated RNA interference and microRNA biogenesis," *Methods in Molecular Biology*, vol. 487, pp. 387–413, 2009.

[14] C. D. Kaplan, L. Laprade, and F. Winston, "Transcription elongation factors repress transcription initiation from cryptic sites," *Science*, vol. 301, no. 5636, pp. 1096–1099, 2003.

[15] J. Camblong, N. Iglesias, C. Fickentscher, G. Dieppois, and F. Stutz, "Antisense RNA stabilization induces transcriptional gene silencing via histone deacetylation in S. cerevisiae," *Cell*, vol. 131, no. 4, pp. 706–717, 2007.

[16] F. De Santa, I. Barozzi, F. Mietton et al., "A large fraction of extragenic RNA Pol II transcription sites overlap enhancers," *PLoS Biology*, vol. 8, no. 5, Article ID e1000384, 2010.

[17] U. A. Ørom, T. Derrien, M. Beringer et al., "Long noncoding RNAs with enhancer-like function in human cells," *Cell*, vol. 143, no. 1, pp. 46–58, 2010.

[18] F. Sato, S. Tsuchiya, S. J. Meltzer, and K. Shimizu, "MicroRNAs and epigenetics," *FEBS Journal*, vol. 278, no. 10, pp. 1598–1609, 2011.

[19] R. A. Young, "RNA polymerase II," *Annual Review of Biochemistry*, vol. 60, pp. 689–715, 1991.

[20] N. Woychik and R. Young, "Exploring RNA polymerase II structure and function," in *Transcription: Mechanisms and Regulation*, R. C. Conaway and J. W. Conaway, Eds., pp. 227–242, Raven Press, New York, NY, USA, 1994.

[21] G. V. Shpakovski, J. Acker, M. Wintzerith, J. F. Lacroix, P. Thuriaux, and M. Vigneron, "Four subunits that are shared by the three classes of RNA polymerase are functionally interchangeable between Homo sapiens and Saccharomyces cerevisiae," *Molecular and Cellular Biology*, vol. 15, no. 9, pp. 4702–4710, 1995.

[22] Y. Hirose and J. L. Manley, "RNA polymerase II and the integration of nuclear events," *Genes & Development*, vol. 14, no. 12, pp. 1415–1429, 2000.

[23] R. J. Sims, R. Belotserkovskaya, and D. Reinberg, "Elongation by RNA polymerase II: the short and long of it," *Genes & Development*, vol. 18, no. 20, pp. 2437–2468, 2004.

[24] H. P. Phatnani and A. L. Greenleaf, "Phosphorylation and functions of the RNA polymerase II CTD," *Genes & Development*, vol. 20, no. 21, pp. 2922–2936, 2006.

[25] R. D. Kornberg, "The molecular basis of eukaryotic transcription," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 32, pp. 12955–12961, 2007.

[26] Y. Hirose and Y. Ohkuma, "Phosphorylation of the C-terminal domain of RNA polymerase II plays central roles in the integrated events of eucaryotic gene expression," *Journal of Biochemistry*, vol. 141, no. 5, pp. 601–608, 2007.

[27] S. Egloff and S. Murphy, "Cracking the RNA polymerase II CTD code," *Trends in Genetics*, vol. 24, no. 6, pp. 280–288, 2008.

[28] B. J. Venters and B. F. Pugh, "How eukaryotic genes are transcribed regulation of eukaroytic gene transcription," *Critical Reviews in Biochemistry and Molecular Biology*, vol. 44, no. 2-3, pp. 117–141, 2009.

[29] S. Buratowski, "Progression through the RNA Polymerase II CTD Cycle," *Molecular Cell*, vol. 36, no. 4, pp. 541–546, 2009.

[30] R. Perales and D. Bentley, "'Cotranscriptionality': the transcription elongation complex as a nexus for nuclear transactions," *Molecular Cell*, vol. 36, no. 2, pp. 178–191, 2009.

[31] P. Richard and J. L. Manley, "Transcription termination by nuclear RNA polymerases," *Genes & Development*, vol. 23, no. 11, pp. 1247–1269, 2009.

[32] 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, no. 1, pp. 34–45, 2010.

[33] J. L. Corden, D. L. Cadena, J. M. Ahearn, and M. E. Dahmus, "A unique structure at the carboxyl terminus of the largest subunit of eukaryotic RNA polymerase II," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 82, no. 23, pp. 7934–7938, 1985.

[34] J. L. Corden, "Tails of RNA polymerase II," *Trends in Biochemical Sciences*, vol. 15, no. 10, pp. 383–387, 1990.

[35] 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.

[36] 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.

[37] Q. Zeidan and G. W. Hart, "The intersections between O-GlcNAcylation and phosphorylation: implications for multiple signaling pathways," *Journal of Cell Science*, vol. 123, no. 1, pp. 13–22, 2010.

[38] S. M. Fuchs, R. N. Laribee, and B. D. Strahl, "Protein modifications in transcription elongation," *Biochimica et Biophysica Acta*, vol. 1789, no. 1, pp. 26–36, 2009.

[39] 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, no. 6025, pp. 99–103, 2011.

[40] S. Buratowski, "The CTD code," *Nature Structural Biology*, vol. 10, no. 9, pp. 679–680, 2003.

[41] T. Jenuwein and C. D. Allis, "Translating the histone code," *Science*, vol. 293, no. 5532, pp. 1074–1080, 2001.

[42] A. G. Pedersen, P. Baldi, Y. Chauvin, and S. Brunak, "The biology of eukaryotic promoter prediction—a review," *Computers and Chemistry*, vol. 23, no. 3-4, pp. 191–207, 1999.

[43] A. D. Basehoar, S. J. Zanton, and B. F. Pugh, "Identification and distinct regulation of yeast TATA box-containing genes," *Cell*, vol. 116, no. 5, pp. 699–709, 2004.

[44] G. C. Yuan, Y. J. Liu, M. F. Dion et al., "Genome-scale identification of nucleosome positions in S. cerevisiae," *Science*, vol. 309, no. 5734, pp. 626–630, 2005.

[45] B. J. Venters and B. F. Pugh, "A canonical promoter organization of the transcription machinery and its regulators in the Saccharomyces genome," *Genome Research*, vol. 19, no. 3, pp. 360–371, 2009.

[46] G. Orphanides, T. Lagrange, and D. Reinberg, "The general transcription factors of RNA polymerase II," *Genes & Development*, vol. 10, no. 21, pp. 2657–2683, 1996.

[47] S. J. Zanton and B. F. Pugh, "Full and partial genome-wide assembly and disassembly of the yeast transcription machinery in response to heat shock," *Genes & Development*, vol. 20, no. 16, pp. 2250–2265, 2006.

[48] J. Q. Svejstrup, Y. Li, J. Fellows, A. Gnatt, S. Bjorklund, and R. D. Kornberg, "Evidence for a mediator cycle at the initiation of transcription," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 12, pp. 6075–6078, 1997.

[49] L. C. Myers and R. D. Kornberg, "Mediator of transcriptional regulation," *Annual Review of Biochemistry*, vol. 69, pp. 729–749, 2000.

[50] R. D. Kornberg, "Mediator and the mechanism of transcriptional activation," *Trends in Biochemical Sciences*, vol. 30, no. 5, pp. 235–239, 2005.

[51] S. Malik and R. G. Roeder, "Dynamic regulation of pol II transcription by the mammalian Mediator complex," *Trends in Biochemical Sciences*, vol. 30, no. 5, pp. 256–263, 2005.

[52] J. S. Kang, S. H. Kim, M. S. Hwang, S. J. Han, Y. C. Lee, and Y. J. Kim, "The structural and functional organization of the yeast mediator complex," *Journal of Biological Chemistry*, vol. 276, no. 45, pp. 42003–42010, 2001.

[53] T. Borggrefe, R. Davis, H. Erdjument-Bromage, P. Tempst, and R. D. Kornberg, "A complex of the Srb8, -9, -10, and -11 transcriptional regulatory proteins from yeast," *Journal of Biological Chemistry*, vol. 277, no. 46, pp. 44202–44207, 2002.

[54] B. Guglielmi, N. L. van Berkum, B. Klapholz et al., "A high resolution protein interaction map of the yeast Mediator complex," *Nucleic Acids Research*, vol. 32, no. 18, pp. 5379–5391, 2004.

[55] J. Z. Chadick and F. J. Asturias, "Structure of eukaryotic Mediator complexes," *Trends in Biochemical Sciences*, vol. 30, no. 5, pp. 264–271, 2005.

[56] A. Casamassimi and C. Napoli, "Mediator complexes and eukaryotic transcription regulation: an overview," *Biochimie*, vol. 89, no. 12, pp. 1439–1446, 2007.

[57] Á. Toth-Petroczy, C. J. Oldfield, I. Simon et al., "Malleable machines in transcription regulation: the Mediator complex," *PLoS Computational Biology*, vol. 4, no. 12, Article ID e1000243, 2008.

[58] R. Biddick and E. T. Young, "Yeast Mediator and its role in transcriptional regulation," *Comptes Rendus Biologies*, vol. 328, no. 9, pp. 773–782, 2005.

[59] S. Bjorklund and C. M. Gustafsson, "The yeast Mediator complex and its regulation," *Trends in Biochemical Sciences*, vol. 30, no. 5, pp. 240–244, 2005.

[60] R. C. Conaway, S. Sato, C. Tomomori-Sato, T. Yao, and J. W. Conaway, "The mammalian Mediator complex and its role in transcriptional regulation," *Trends in Biochemical Sciences*, vol. 30, no. 5, pp. 250–255, 2005.

[61] X. Fan, D. M. Chou, and K. Struhl, "Activator-specific recruitment of Mediator in vivo," *Nature Structural & Molecular Biology*, vol. 13, no. 2, pp. 117–120, 2006.

[62] X. Fan and K. Struhl, "Where does mediator bind in vivo?" *PLoS One*, vol. 4, no. 4, Article ID e5029, 2009.

[63] M. E. Maxon, J. A. Goodrich, and R. Tjian, "Transcription factor IIE binds preferentially to RNA polymerase IIa and recruits TFIIH: a model for promoter clearance," *Genes & Development*, vol. 8, no. 5, pp. 515–524, 1994.

[64] C. Esnault, Y. Ghavi-Helm, S. Brun et al., "Mediator-dependent recruitment of TFIIH modules in preinitiation complex," *Molecular Cell*, vol. 31, no. 3, pp. 337–346, 2008.

[65] A. J. Koleske and R. A. Young, "An RNA polymerase II holoenzyme responsive to activators," *Nature*, vol. 368, no. 6470, pp. 466–469, 1994.

[66] 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.

[67] C. J. Hengartner, C. M. Thompson, J. Zhang et al., "Association of an activator with an RNA polymerase II holoenzyme," *Genes & Development*, vol. 9, no. 8, pp. 897–910, 1995.

[68] A. Barberis, J. Pearlberg, N. Simkovich et al., "Contact with a component of the polymerase II holoenzyme suffices for gene activation," *Cell*, vol. 81, no. 3, pp. 359–368, 1995.

[69] D. M. Chao, E. L. Gadbois, P. J. Murray et al., "A mammalian SRB protein associated with an RNA polymerase II holoenzyme," *Nature*, vol. 380, no. 6569, pp. 82–85, 1996.

[70] A. J. Koleske and R. A. Young, "The RNA polymerase II holoenzyme and its implications for gene regulation," *Trends in Biochemical Sciences*, vol. 20, no. 3, pp. 113–116, 1995.

[71] K. J. Armache, H. Kettenberger, and P. Cramer, "Architecture of initiation-competent 12-subunit RNA polymerase II," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 12, pp. 6964–6968, 2003.

[72] A. Meinhart, T. Kamenski, S. Hoeppner, S. Baumli, and P. Cramer, "A structural perspective of CTD function," *Genes & Development*, vol. 19, no. 12, pp. 1401–1415, 2005.

[73] 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.

[74] 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, no. 7, pp. 5717–5726, 2011.

[75] A. Meinhart and P. Cramer, "Recognition of RNA polymerase II carboxy-terminal domain by 3′-RNA-processing factors," *Nature*, vol. 430, no. 6996, pp. 223–226, 2004.

[76] 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.

[77] M. E. Kang and M. E. Dahmus, "The photoactivated cross-linking of recombinant C-terminal domain to proteins in a HeLa cell transcription extract that comigrate with transcription factors IIE and IIF," *Journal of Biological Chemistry*, vol. 270, no. 40, pp. 23390–23397, 1995.

[78] A. Usheva, E. Maldonado, A. Goldring et al., "Specific interaction between the nonphosphorylated form of RNA polymerase II and the TATA-binding protein," *Cell*, vol. 69, no. 5, pp. 871–881, 1992.

[79] L. C. Myers, C. M. Gustafsson, D. A. Bushnell et al., "The Med proteins of yeast and their function through the RNA polymerase II carboxy-terminal domain," *Genes & Development*, vol. 12, no. 1, pp. 45–54, 1998.

[80] 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 & Development*, vol. 11, no. 24, pp. 3319–3326, 1997.

[81] 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.

[82] 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.

[83] P. Komarnitsky, E. J. Cho, and S. Buratowski, "Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription," *Genes & Development*, vol. 14, no. 19, pp. 2452–2460, 2000.

[84] 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 & Development*, vol. 11, no. 24, pp. 3306–3318, 1997.

[85] S. C. Schroeder, B. Schwer, S. Shuman, and D. Bentley, "Dynamic association of capping enzymes with transcribing RNA polymerase II," *Genes & Development*, vol. 14, no. 19, pp. 2435–2440, 2000.

[86] 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.

[87] 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.

[88] S. Drouin, L. Laramée, 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 Genetics*, vol. 6, no. 10, Article ID e1001173, pp. 1–12, 2010.

[89] 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 & Development*, vol. 21, no. 2, pp. 160–174, 2007.

[90] L. Vasiljeva, M. Kim, H. Mutschler, S. Buratowski, and A. Meinhart, "The Nrd1-Nab3-Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain," *Nature Structural & Molecular Biology*, vol. 15, no. 8, pp. 795–804, 2008.

[91] J. S. Finkel, K. Chinchilla, D. Ursic, and M. R. Culbertson, "Sen1p performs two genetically separable functions in transcription and processing of U5 small nuclear RNA in Saccharomyces cerevisiae," *Genetics*, vol. 184, no. 1, pp. 107–118, 2010.

[92] A. Daulny, F. Geng, M. Muratani, J. M. Geisinger, S. E. Salghetti, and W. P. Tansey, "Modulation of RNA polymerase II subunit composition by ubiquitylation," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 50, pp. 19649–19654, 2008.

[93] D. P. Morris, H. P. Phatnani, and A. L. Greenleaf, "Phospho-carboxyl-terminal domain binding and the role of a prolyl isomerase in pre-mRNA 3'-end formation," *Journal of Biological Chemistry*, vol. 274, no. 44, pp. 31583–31587, 1999.

[94] X. Wu, C. B. Wilcox, G. Devasahayam et al., "The Ess1 prolyl isomerase is linked to chromatin remodeling complexes and the general transcription machinery," *The EMBO Journal*, vol. 19, no. 14, pp. 3727–3738, 2000.

[95] 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.

[96] E. Vojnic, B. Simon, B. D. Strahl, M. Sattler, and P. Cramer, "Structure and carboxyl-terminal domain (CTD) binding of the Set2 SRI domain that couples histone H3 Lys36 methylation to transcription," *Journal of Biological Chemistry*, vol. 281, no. 1, pp. 13–15, 2006.

[97] 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.

[98] J. L. Dermody, J. M. Dreyfuss, J. Villén et al., "Unphosphorylated SR-like protein Npl3 stimulates RNA polymerase II elongation," *PLoS One*, vol. 3, no. 9, Article ID e3273, 2008.

[99] 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 & Molecular Biology*, vol. 12, no. 2, pp. 144–151, 2005.

[100] D. Hollingworth, C. G. Noble, I. A. Taylor, and A. Ramos, "RNA polymerase II CTD phosphopeptides compete with RNA for the interaction with Pcf11," *RNA*, vol. 12, no. 4, pp. 555–560, 2006.

[101] D. Barillà, B. A. Lee, and N. J. Proudfoot, "Cleavage/polyadenylation factor IA associates with the carboxyl-terminal domain of RNA polymerase II in Saccharomyces cerevisae," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 2, pp. 445–450, 2001.

[102] 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.

[103] B. Dichtl, D. Blank, M. Sadowski, W. Hübner, S. Weiser, and W. Keller, "Yhh1p/Cft1p directly links poly(A) site recognition and RNA polymerase II transcription termination," *The EMBO Journal*, vol. 21, no. 15, pp. 4125–4135, 2002.

[104] C. R. Rodriguez, E. J. Cho, M. C. Keogh, C. L. Moore, A. L. Greenleaf, and S. Buratowski, "Kin28, the TFIIH-associated carboxy-terminal domain kinase, facilitates the recruitment of mRNA processing machinery to RNA polymerase II," *Molecular and Cellular Biology*, vol. 20, no. 1, pp. 104–112, 2000.

[105] 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.

[106] P. Pascual-García, C. K. Govind, E. Queralt et al., "Sus1 is recruited to coding regions and functions during transcription elongation in association with SAGA and TREX2," *Genes & Development*, vol. 22, no. 20, pp. 2811–2822, 2008.

[107] A. L. MacKellar and A. L. Greenleaf, "Cotranscriptional association of mRNA export factor Yra1 with C-terminal domain of RNA polymerase II," *Journal of Biological Chemistry*, vol. 286, no. 42, pp. 36385–36395, 2011.

[108] A. Chang, S. Cheang, X. Espanel, and M. Sudol, "Rsp5 WW domains interact directly with the carboxyl-terminal domain of RNA polymerase II," *Journal of Biological Chemistry*, vol. 275, no. 27, pp. 20562–20571, 2000.

[109] B. P. Somesh, J. Reid, W. F. Liu et al., "Multiple mechanisms confining RNA polymerase II ubiquitylation to polymerases undergoing transcriptional arrest," *Cell*, vol. 121, no. 6, pp. 913–923, 2005.

[110] 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.

[111] J. Q. Svejstrup, W. J. Feaver, J. LaPointe, and R. D. Kornberg, "RNA polymerase transcription factor IIH holoenzyme from yeast," *Journal of Biological Chemistry*, vol. 269, no. 45, pp. 28044–28048, 1994.

[112] J. Q. Svejstrup, Z. Wang, W. J. Feaver et al., "Different forms of TFIIH for transcription and DNA repair: holo-TFIIH and a nucleotide excision repairosome," *Cell*, vol. 80, no. 1, pp. 21–28, 1995.

[113] 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.

[114] V. E. Myer and R. A. Young, "RNA polymerase II holoenzymes and subcomplexes," *Journal of Biological Chemistry*, vol. 273, no. 43, pp. 27757–27760, 1998.

[115] F. Tirode, D. Busso, F. Coin, and J. M. Egly, "Reconstitution of the transcription factor TFIIH: assignment of functions for the three enzymatic subunits, XPB, XPD, and cdk7," *Molecular Cell*, vol. 3, no. 1, pp. 87–95, 1999.

[116] W. H. Chang and R. D. Kornberg, "Electron crystal structure of the transcription factor and DNA repair complex, core TFIIH," *Cell*, vol. 102, no. 5, pp. 609–613, 2000.

[117] Y. Takagi, C. A. Masuda, W. H. Chang et al., "Ubiquitin ligase activity of TFIIH and the transcriptional response to DNA damage," *Molecular Cell*, vol. 18, no. 2, pp. 237–243, 2005.

[118] G. Rabut, G. Le Dez, R. Verma et al., "The TFIIH subunit Tfb3 regulates cullin neddylation," *Molecular Cell*, vol. 43, no. 3, pp. 488–495, 2011.

[119] M. E. Dahmus, "Reversible phosphorylation of the C-terminal domain of RNA polymerase II," *Journal of Biological Chemistry*, vol. 271, no. 32, pp. 19009–19012, 1996.

[120] O. Bensaude, F. Bonnet, C. Cassé, M. F. Dubois, V. T. Nguyen, and B. Palancade, "Regulated phosphorylation of the RNA polymerase II C-terminal domain (CTD)," *Biochemistry and Cell Biology*, vol. 77, no. 4, pp. 249–255, 1999.

[121] B. Palancade and O. Bensaude, "Investigating RNA polymerase II carboxyl-terminal domain (CTD) phosphorylation," *European Journal of Biochemistry*, vol. 270, no. 19, pp. 3859–3870, 2003.

[122] W. J. Feaver, J. Q. Svejstrup, N. L. Henry, and R. D. Kornberg, "Relationship of CDK-activating kinase and RNA polymerase II CTD kinase TFIIH/TFIIK," *Cell*, vol. 79, no. 6, pp. 1103–1109, 1994.

[123] P. Rickert, J. L. Corden, and E. Lees, "Cyclin C/CDK8 and cyclin H/CDK7/p36 are biochemically distinct CTD kinases," *Oncogene*, vol. 18, no. 4, pp. 1093–1102, 1999.

[124] M. M. Gebara, M. H. Sayre, and J. L. Corden, "Phosphorylation of the carboxy-terminal repeat domain in RNA polymerase II by cyclin-dependent kinases is sufficient to inhibit transcription," *Journal of Cellular Biochemistry*, vol. 64, no. 3, pp. 390–402, 1997.

[125] A. Munshi, G. Shafi, N. Aliya, and A. Jyothy, "Histone modifications dictate specific biological readouts," *Journal of Genetics and Genomics*, vol. 36, no. 2, pp. 75–88, 2009.

[126] J. S. Lee, E. Smith, and A. Shilatifard, "The language of histone crosstalk," *Cell*, vol. 142, no. 5, pp. 682–685, 2010.

[127] J. L. Workman, "Nucleosome displacement in transcription," *Genes & Development*, vol. 20, no. 15, pp. 2009–2017, 2006.

[128] S. Nakanishi, B. W. Sanderson, K. M. Delventhal, W. D. Bradford, K. Staehling-Hampton, and A. Shilatifard, "A comprehensive library of histone mutants identifies nucleosomal residues required for H3K4 methylation," *Nature Structural & Molecular Biology*, vol. 15, no. 8, pp. 881–888, 2008.

[129] A. Wood, J. Schneider, J. Dover, M. Johnston, and A. Shilatifard, "The Paf1 complex is essential for histone monoubiquitination by the Rad6-Bre1 complex, which signals for histone methylation by COMPASS and Dot1p," *Journal of Biological Chemistry*, vol. 278, no. 37, pp. 34739–34742, 2003.

[130] P. Cramer, D. A. Bushnell, and R. D. Kornberg, "Structural basis of transcription: RNA polymerase II at 2.8 ångstrom resolution," *Science*, vol. 292, no. 5523, pp. 1863–1876, 2001.

[131] A. Ghosh, S. Shuman, and C. Lima, "Structural insights to how Mammalian capping enzyme reads the CTD code," *Molecular Cell*, vol. 43, no. 2, pp. 299–310, 2011.

[132] 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 & Development*, vol. 15, no. 14, pp. 1783–1795, 2001.

[133] S. Moteki and D. Price, "Functional coupling of capping and transcription of mRNA," *Molecular Cell*, vol. 10, no. 3, pp. 599–609, 2002.

[134] Y. Liu, C. Kung, J. Fishburn, A. Z. Ansari, K. M. Shokat, and S. Hahn, "Two cyclin-dependent kinases promote RNA polymerase II transcription and formation of the scaffold complex," *Molecular and Cellular Biology*, vol. 24, no. 4, pp. 1721–1735, 2004.

[135] 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.

[136] B. Schwer and S. Shuman, "Deciphering the RNA polymerase II CTD code in fission yeast," *Molecular Cell*, vol. 43, no. 2, pp. 311–318, 2011.

[137] S. W. Hong, M. H. Seong, W. Y. Jae 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.

[138] J. R. Tietjen, D. W. Zhang, J. B. Rodríguez-Molina et al., "Chemical-genomic dissection of the CTD code," *Nature Structural & Molecular Biology*, vol. 17, no. 9, pp. 1154–1161, 2010.

[139] 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.

[140] 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.

[141] K. Glover-Cutter, S. Larochelle, 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.

[142] S. Boeing, C. Rigault, M. Heidemann, D. Eick, and M. Meisterernst, "RNA polymerase II C-terminal heptarepeat domain Ser-7 phosphorylation is established in a mediator-dependent fashion," *Journal of Biological Chemistry*, vol. 285, no. 1, pp. 188–196, 2010.

[143] Z. Songyang, K. P. Lu, Y. T. Kwon et al., "A structural basis for substrate specificities of protein Ser/Thr kinases: primary sequence preference of casein kinases I and II, NIMA, phosphorylase kinase, calmodulin-dependent kinase II, CDK5, and Erk1," *Molecular and Cellular Biology*, vol. 16, no. 11, pp. 6486–6493, 1996.

[144] 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.

[145] 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.

[146] 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 & Molecular Biology*, vol. 17, no. 10, pp. 1272–1278, 2010.

[147] G. A. Hartzog, T. Wada, H. Handa, and F. Winston, "Evidence that Spt4, Spt5, and Spt6 control transcription elongation by RNA polymerase II in Saccharomyces cerevisiae," *Genes & Development*, vol. 12, no. 3, pp. 357–369, 1998.

[148] F. W. Martinez-Rucobo, S. Sainsbury, A. C.M. Cheung, and P. Cramer, "Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity," *The EMBO Journal*, vol. 30, no. 7, pp. 1302–1310, 2011.

[149] D. Grohmann, J. Nagy, A. Chakraborty et al., "The initiation factor TFE and the elongation factor Spt4/5 compete for the RNAP clamp during transcription initiation and elongation," *Molecular Cell*, vol. 43, no. 2, pp. 263–274, 2011.

[150] M. W. Adkins and J. K. Tyler, "Transcriptional activators are dispensable for transcription in the absence of Spt6-mediated chromatin reassembly of promoter regions," *Molecular Cell*, vol. 21, no. 3, pp. 405–416, 2006.

[151] 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.

[152] G. Orphanides, W. H. Wu, W. S. Lane, M. Hampsey, and D. Reinberg, "The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins," *Nature*, vol. 400, no. 6741, pp. 284–288, 1999.

[153] A. Jamai, A. Puglisi, and M. Strubin, "Histone chaperone Spt16 promotes redeposition of the original H3-H4 histones evicted by elongating RNA polymerase," *Molecular Cell*, vol. 35, no. 3, pp. 377–383, 2009.

[154] L. Zhang, A. G. L. Fletcher, V. Cheung, F. Winston, and L. A. Stargell, "Spn1 regulates the recruitment of Spt6 and the Swi/Snf complex during transcriptional activation by RNA polymerase II," *Molecular and Cellular Biology*, vol. 28, no. 4, pp. 1393–1403, 2008.

[155] S. M. McDonald, D. Close, H. Xin, T. Formosa, and C. P. Hill, "Structure and biological importance of the Spn1-Spt6 interaction, and its regulatory role in nucleosome binding," *Molecular Cell*, vol. 40, no. 5, pp. 725–735, 2010.

[156] D. Prather, N. J. Krogan, A. Emili, J. F. Greenblatt, and F. Winston, "Identification and characterization of Elf1, a conserved transcription elongation factor in Saccharomyces cerevisiae," *Molecular and Cellular Biology*, vol. 25, no. 22, pp. 10122–10135, 2005.

[157] J. A. Jaehning, "The Paf1 complex: platform or player in RNA polymerase II transcription?" *Biochimica et Biophysica Acta*, vol. 1799, pp. 379–388, 2010.

[158] V. Brès, S. M. Yoh, and K. A. Jones, "The multi-tasking P-TEFb complex," *Current Opinion in Cell Biology*, vol. 20, no. 3, pp. 334–340, 2008.

[159] 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.

[160] D. H. Price, "P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II," *Molecular and Cellular Biology*, vol. 20, no. 8, pp. 2629–2634, 2000.

[161] 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.

[162] K. Zhou, W. H. W. Kuo, J. Fillingham, and J. F. Greenblatt, "Control of transcriptional elongation and cotranscriptional histone modification by the yeast BUR kinase substrate Spt5," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 17, pp. 6956–6961, 2009.

[163] A. Wood, J. Schneider, J. Dover, M. Johnston, and A. Shilatifard, "The Bur1/Bur2 complex is required for histone H2B monoubiquitination by Rad6/Bre1 and histone methylation by COMPASS," *Molecular Cell*, vol. 20, no. 4, pp. 589–599, 2005.

[164] 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.

[165] J. B. Kim and P. A. Sharp, "Positive transcription elongation factor B phosphorylates hSPT5 and RNA polymerase II carboxyl-terminal domain independently of cyclin-dependent kinase-activating kinase," *Journal of Biological Chemistry*, vol. 276, no. 15, pp. 12317–12323, 2001.

[166] 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.

[167] 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.

[168] M. S. Kobor, J. Archambault, W. Lester et al., "An unusual eukaryotic protein phosphatase required for transcription by RNA polymerase II and CTD dephosphorylation in S. cerevisiae," *Molecular Cell*, vol. 4, no. 1, pp. 55–62, 1999.

[169] 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 & Development*, vol. 15, no. 24, pp. 3319–3329, 2001.

[170] K. W. Henry, A. Wyce, W. S. Lo et al., "Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8," *Genes & Development*, vol. 17, no. 21, pp. 2648–2663, 2003.

[171] A. Wyce, T. Xiao, K. A. Whelan et al., "H2B ubiquitylation acts as a barrier to Ctk1 nucleosomal recruitment prior to removal by Ubp8 within a SAGA-related complex," *Molecular Cell*, vol. 27, no. 2, pp. 275–288, 2007.

[172] D. Bentley, "Coupling RNA polymerase II transcription with pre-mRNA processing," *Current Opinion in Cell Biology*, vol. 11, no. 3, pp. 347–351, 1999.

[173] D. Bentley, "The mRNA assembly line: transcription and processing machines in the same factory," *Current Opinion in Cell Biology*, vol. 14, no. 3, pp. 336–342, 2002.

[174] 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.

[175] N. J. Proudfoot, A. Furger, and M. J. Dye, "Integrating mRNA processing with transcription," *Cell*, vol. 108, no. 4, pp. 501–512, 2002.

[176] 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.

[177] D. L. Bentley, "Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors," *Current Opinion in Cell Biology*, vol. 17, no. 3, pp. 251–256, 2005.

[178] S. Buratowski, "Connections between mRNA 3′ end processing and transcription termination," *Current Opinion in Cell Biology*, vol. 17, no. 3, pp. 257–261, 2005.

[179] J. Li, D. Moazed, and S. P. Gygi, "Association of the histone methyltransferase Set2 with RNA polymerase II plays a role in transcription elongation," *Journal of Biological Chemistry*, vol. 277, no. 51, pp. 49383–49388, 2002.

[180] 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.

[181] B. Li, L. Howe, S. Anderson, J. R. Yates, and J. L. Workman, "The Set2 histone methyltransferase functions through the phosphorylated carboxyl-terminal domain of RNA polymerase II," *Journal of Biological Chemistry*, vol. 278, no. 11, pp. 8897–8903, 2003.

[182] M. J. Carrozza, B. Li, L. Florens et al., "Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription," *Cell*, vol. 123, no. 4, pp. 581–592, 2005.

[183] 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.

[184] W. Gilbert, C. W. Siebel, and C. Guthrie, "Phosphorylation by Sky1p promotes Npl3p shuttling and mRNA dissociation," *RNA*, vol. 7, no. 2, pp. 302–313, 2001.

[185] E. P. Lei, H. Krebber, and P. A. Silver, "Messenger RNAs are recruited for nuclear export during transcription," *Genes & Development*, vol. 15, no. 14, pp. 1771–1782, 2001.

[186] M. E. Bucheli and S. Buratowski, "Npl3 is an antagonist of mRNA 3′ end formation by RNA polymerase II," *The EMBO Journal*, vol. 24, no. 12, pp. 2150–2160, 2005.

[187] M. E. Bucheli, X. He, C. D. Kaplan, C. L. Moore, and S. Buratowski, "Polyadenylation site choice in yeast is affected by competition between Npl3 and polyadenylation factor CFI," *RNA*, vol. 13, no. 10, pp. 1756–1764, 2007.

[188] N. J. Proudfoot, "Ending the message: poly(A) signals then and now," *Genes & Development*, vol. 25, no. 17, pp. 1770–1782, 2011.

[189] M. Kim, S. H. Ahn, N. J. Krogan, J. F. Greenblatt, and S. Buratowski, "Transitions in RNA polymerase II elongation complexes at the 3′ ends of genes," *The EMBO Journal*, vol. 23, no. 2, pp. 354–364, 2004.

[190] S. Connelly and J. L. Manley, "A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II," *Genes & Development*, vol. 2, no. 4, pp. 440–452, 1988.

[191] S. West, N. Gromak, and N. J. Proudfoot, "Human 5′ → 3′ exonuclease Xm2 promotes transcription termination at co-transcriptional cleavage sites," *Nature*, vol. 432, no. 7016, pp. 522–525, 2004.

[192] 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 & Molecular Biology*, vol. 17, no. 10, pp. 1195–1201, 2010.

[193] Z. Zhang, J. Fu, and D. S. Gilmour, "CTD-dependent dismantling of the RNA polymerase II elongation complex by the pre-mRNA 3′-end processing factor, Pcf11," *Genes & Development*, vol. 19, no. 13, pp. 1572–1580, 2005.

[194] M. Sadowski, B. Dichtl, W. Hübner, and W. Keller, "Independent functions of yeast Pcf11p in pre-mRNA 3′ end processing and in transcription termination," *The EMBO Journal*, vol. 22, no. 9, pp. 2167–2177, 2003.

[195] M. Kim, L. Vasiljeva, O. J. Rando, A. Zhelkovsky, C. Moore, and S. Buratowski, "Distinct pathways for snoRNA and mRNA termination," *Molecular Cell*, vol. 24, no. 5, pp. 723–734, 2006.

[196] H. Kim, B. Erickson, W. Luo et al., "Gene-specific RNA polymerase II phosphorylation and the CTD code," *Nature Structural & Molecular Biology*, vol. 17, no. 10, pp. 1279–1286, 2010.

[197] C. E. Birse, L. Minvielle-Sebastia, B. A. Lee, W. Keller, and N. J. Proudfoot, "Coupling termination of transcription to messenger RNA maturation in yeast," *Science*, vol. 280, no. 5361, pp. 298–301, 1998.

[198] R. K. Gudipati, T. Villa, J. Boulay, and D. Libri, "Phosphorylation of the RNA polymerase II C-terminal domain dictates transcription termination choice," *Nature Structural & Molecular Biology*, vol. 15, no. 8, pp. 786–794, 2008.

[199] E. J. Steinmetz, N. K. Conrad, D. A. Brow, and J. L. Corden, "RNA-binding protein Nrd1 directs poly(A)-independent 3′-end formation of RNA polymerase II transcripts," *Nature*, vol. 413, no. 6853, pp. 327–331, 2001.

[200] N. Terzi, L. S. Churchman, L. Vasiljeva, J. Weissman, and S. Buratowski, "H3K4 trimethylation by Set1 promotes efficient termination by the Nrd1-Nab3-Sen1 pathway," *Molecular and Cellular Biology*, vol. 31, no. 17, pp. 3569–3583, 2011.

[201] E. J. Steinmetz and D. A. Brow, "Repression of gene expression by an exogenous sequence element acting in concert with a heterogeneous nuclear ribonucleoprotein-like protein, Nrd1, and the putative helicase Sen1," *Molecular and Cellular Biology*, vol. 16, no. 12, pp. 6993–7003, 1996.

[202] E. J. Steinmetz and D. A. Brow, "Control of pre-mRNA accumulation by the essential yeast protein Nrd1 requires high-affinity transcript binding and a domain implicated in RNA polymerase II association," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 12, pp. 6699–6704, 1998.

[203] N. K. Conrad, S. M. Wilson, E. J. Steinmetz et al., "A yeast heterogeneous nuclear ribonucleoprotein complex associated with RNA polymerase II," *Genetics*, vol. 154, no. 2, pp. 557–571, 2000.

[204] K. L. Carroll, D. A. Pradhan, J. A. Granek, N. D. Clarke, and J. L. Corden, "Identification of cis elements directing termination of yeast nonpolyadenylated snoRNA transcripts," *Molecular and Cellular Biology*, vol. 24, no. 14, pp. 6241–6252, 2004.

[205] K. L. Carroll, R. Ghirlando, J. M. Ames, and J. L. Corden, "Interaction of yeast RNA-binding proteins Nrd1 and Nab3 with RNA polymerase II terminator elements," *RNA*, vol. 13, no. 3, pp. 361–373, 2007.

[206] F. Hobor, R. Pergoli, K. Kubicek et al., "Recognition of transcription termination signal by the nuclear polyadenylated RNA-binding (NAB) 3 protein," *Journal of Biological Chemistry*, vol. 286, no. 5, pp. 3645–3657, 2011.

[207] W. Wlotzka, G. Kudla, S. Granneman, and D. Tollervey, "The nuclear RNA polymerase II surveillance system targets polymerase III transcripts," *The EMBO Journal*, vol. 30, no. 9, pp. 1790–1803, 2011.

[208] D. Ursic, K. L. Himmel, K. A. Gurley, F. Webb, and M. R. Culbertson, "The yeast SEN1 gene is required for the processing of diverse RNA classes," *Nucleic Acids Research*, vol. 25, no. 23, pp. 4778–4785, 1997.

[209] K. Skourti-Stathaki, N. Proudfoot, and N. Gromak, "Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination," *Molecular Cell*, vol. 42, no. 6, pp. 794–805, 2011.

[210] H. E. Mischo, B. Gómez-González, P. Grzechnik et al., "Yeast Sen1 helicase protects the genome from transcription-associated instability," *Molecular Cell*, vol. 41, no. 1, pp. 21–32, 2011.

[211] E. Nedea, D. Nalbant, D. Xia et al., "The Glc7 phosphatase subunit of the cleavage and polyadenylation factor is essential for transcription termination on snoRNA genes," *Molecular Cell*, vol. 29, no. 5, pp. 577–587, 2008.

[212] G. Chanfreau, P. Legrain, and A. Jacquier, "Yeast RNase III as a key processing enzyme in small nucleolar RNAs metabolism," *Journal of Molecular Biology*, vol. 284, no. 4, pp. 975–988, 1998.

[213] G. Chanfreau, G. Rotondo, P. Legrain, and A. Jacquier, "Processing of a dicistronic small nucleolar RNA precursor by the RNA endonuclease Rnt1," *The EMBO Journal*, vol. 17, no. 13, pp. 3726–3737, 1998.

[214] C. Allmang, J. Kufel, G. Chanfreau, P. Mitchell, E. Petfalski, and D. Tollervey, "Functions of the exosome in rRNA, snoRNA and snRNA synthesis," *The EMBO Journal*, vol. 18, no. 19, pp. 5399–5410, 1999.

[215] J. LaCava, J. Houseley, C. Saveanu et al., "RNA degradation by the exosome is promoted by a nuclear polyadenylation complex," *Cell*, vol. 121, no. 5, pp. 713–724, 2005.

[216] D. E. Egecioglu, A. K. Henras, and G. F. Chanfreau, "Contributions of Trf4p- And Trf5p-dependent polyadenylation to the processing and degradative functions of the yeast nuclear exosome," *RNA*, vol. 12, no. 1, pp. 26–32, 2006.

[217] L. Vasiljeva and S. Buratowski, "Nrd1 interacts with the nuclear exosome for 3′ processing of RNA polymerase II transcripts," *Molecular Cell*, vol. 21, no. 2, pp. 239–248, 2006.

[218] D. Baillat, M. A. Hakimi, A. M. Näär, 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.

[219] S. Egloff, 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.

[220] L. Minvielle-Sebastia, P. J. Preker, and W. Keller, "RNA14 and RNA15 proteins as components of a yeast pre-mRNA 3'-end processing factor," *Science*, vol. 266, no. 5191, pp. 1702–1705, 1994.

[221] M. J. Moore, "From birth to death: the complex lives of eukaryotic mRNAs," *Science*, vol. 309, no. 5740, pp. 1514–1518, 2005.

[222] E. Nedea, X. He, M. Kim et al., "Organization and function of APT, a subcomplex of the yeast cleavage and polyadenylation factor involved in the formation of mRNA and small nucleolar RNA 3′-ends," *Journal of Biological Chemistry*, vol. 278, no. 35, pp. 33000–33010, 2003.

[223] A. Ansari and M. Hampsey, "A role for the CPF 3′ -end processing machinery in RNAP II-dependent gene looping," *Genes & Development*, vol. 19, no. 24, pp. 2969–2978, 2005.

[224] E. J. Steinmetz and D. A. Brow, "Ssu72 protein mediates both poly(A)-coupled and poly(A)-independent termination of RNA polymerase II transcription," *Molecular and Cellular Biology*, vol. 23, no. 18, pp. 6339–6349, 2003.

[225] 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.

[226] 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.

[227] M. A. Ghazy, X. He, B. N. Singh, M. Hampsey, and C. Moore, "The essential N terminus of the pta1 scaffold protein is required for snoRNA transcription termination and Ssu72 function but is dispensable for pre-mRNA 3'-end processing," *Molecular and Cellular Biology*, vol. 29, no. 8, pp. 2296–2307, 2009.

[228] N. Singh, Z. Ma, T. Gemmill et al., "The Ess1 prolyl isomerase is required for transcription termination of small noncoding RNAs via the Nrd1 pathway," *Molecular Cell*, vol. 36, no. 2, pp. 255–266, 2009.

[229] J. Archambault, R. S. Chambers, M. S. Kobor et al., "An essential component of a C-terminal domain phosphatase that interacts with transcription factor IIF in Saccharomyces cerevisiae," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 26, pp. 14300–14305, 1997.

[230] S. E. Kong, M. S. Kobor, N. J. Krogan et al., "Interaction of Fcp1 phosphatase with elongating RNA polymerase II holoenzyme, enzymatic mechanism of action, and genetic interaction with elongator," *Journal of Biological Chemistry*, vol. 280, no. 6, pp. 4299–4306, 2005.

[231] S. Hausmann, H. Erdjument-Bromage, and S. Shuman, "Schizosaccharomyces pombe carboxyl-terminal domain (CTD) phosphatase Fcp1," *Journal of Biological Chemistry*, vol. 279, no. 12, pp. 10892–10900, 2004.

[232] A. Ghosh, S. Shuman, and C. D. Lima, "The structure of Fcp1, an essential RNA polymerase II CTD phosphatase," *Molecular Cell*, vol. 32, no. 4, pp. 478–490, 2008.

[233] K. L. Abbott, M. B. Renfrow, M. J. Chalmers et al., "Enhanced binding of RNAP II CTD phosphatase FCP1 to RAP74 following CK2 phosphorylation," *Biochemistry*, vol. 44, no. 8, pp. 2732–2745, 2005.

[234] D. W. Zhang, A. L. Mosley, S. R. Ramisetty et al., "Ssu72 phosphatase dependent erasure of phospho-Ser7 marks on the RNA Polymerase II C-terminal domain is essential for viability and transcription termination," *Journal of Biological Chemistry*. In press.

[235] H. Cho, T. K. Kim, H. Mancebo, W. S. Lane, O. Flores, and D. Reinberg, "A protein phosphatase functions to recycle RNA polymerase II," *Genes & Development*, vol. 13, no. 12, pp. 1540–1552, 1999.

[236] B. Dichtl, D. Blank, M. Ohnacker et al., "A role for SSU72 in balancing RNA polymerase II transcription elongation and termination," *Molecular Cell*, vol. 10, no. 5, pp. 1139–1150, 2002.

[237] J. M. O'Sullivan, S. M. Tan-Wong, A. Morillon et al., "Gene loops juxtapose promoters and terminators in yeast," *Nature Genetics*, vol. 36, no. 9, pp. 1014–1018, 2004.

[238] B. N. Singh, A. Ansari, and M. Hampsey, "Detection of gene loops by 3C in yeast," *Methods*, vol. 48, no. 4, pp. 361–367, 2009.

[239] B. N. Singh and M. Hampsey, "A transcription-independent role for TFIIB in gene looping," *Molecular Cell*, vol. 27, no. 5, pp. 806–816, 2007.

[240] P. Vinciguerra and F. Stutz, "mRNA export: an assembly line from genes to nuclear pores," *Current Opinion in Cell Biology*, vol. 16, no. 3, pp. 285–292, 2004.

[241] A. Köhler and E. Hurt, "Exporting RNA from the nucleus to the cytoplasm," *Nature Reviews Molecular Cell Biology*, vol. 8, no. 10, pp. 761–773, 2007.

[242] M. Stewart, "Nuclear export of mRNA," *Trends in Biochemical Sciences*, vol. 35, no. 11, pp. 609–617, 2010.

[243] A. Segref, K. Sharma, V. Doye et al., "Mex67p, a novel factor for nuclear mRNA export, binds to both poly(A)+ RNA and nuclear pores," *The EMBO Journal*, vol. 16, no. 11, pp. 3256–3271, 1997.

[244] D. Jani, S. Lutz, N. J. Marshall et al., "Sus1, Cdc31, and the Sac3 CID region form a conserved interaction platform that promotes nuclear pore association and mRNA export," *Molecular Cell*, vol. 33, no. 6, pp. 727–737, 2009.

[245] J. M. Huibregtse, J. C. Yang, and S. L. Beaudenon, "The large subunit of RNA polymerase II is a substrate of the Rsp5 ubiquitin-protein ligase," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 8, pp. 3656–3661, 1997.

[246] S. L. Beaudenon, M. R. Huacani, G. Wang, D. P. McDonnell, and J. M. Huibregtse, "Rsp5 ubiquitin-protein ligase mediates DNA damage-induced degradation of the large subunit of RNA polymerase II in Saccharomyces cerevisiae," *Molecular and Cellular Biology*, vol. 19, no. 10, pp. 6972–6979, 1999.

[247] G. F. Heine, A. A. Horwitz, and J. D. Parvin, "Multiple mechanisms contribute to inhibit transcription in response to DNA damage," *Journal of Biological Chemistry*, vol. 283, no. 15, pp. 9555–9561, 2008.

[248] S. Buratowski, "The CTD code," *Protein Science*, vol. 6, pp. 249–253, 1997.

[249] R. Baskaran, S. R. Escobar, and J. Y. J. Wang, "Nuclear c-Abl is a COOH-terminal repeated domain (CTD)-tyrosine kinase-specific for the mammalian RNA polymerase II: possible role in transcription elongation," *Cell Growth & Differentiation*, vol. 10, no. 6, pp. 387–396, 1999.

[250] H. Sakurai and A. Ishihama, "Level of the RNA polymerase II in the fission yeast stays constant but phosphorylation of its carboxyl terminal domain varies depending on the phase and rate of cell growth," *Genes to Cells*, vol. 7, no. 3, pp. 273–284, 2002.

[251] B. A. Lewis, "O-GlcNAc modification of the CTD of human RNA Polymerase II is an intermediate step occurring prior to transcription initiation," unpublished data.

[252] M. L. West and J. L. Corden, "Construction and analysis of yeast RNA polymerase II CTD deletion and substitution mutations," *Genetics*, vol. 140, no. 4, pp. 1223–1233, 1995.

Hindawi Publishing Corporation
Genetics Research International
Volume 2012, Article ID 170173, 5 pages
doi:10.1155/2012/170173

Review Article

Control of Transcriptional Elongation by RNA Polymerase II: A Retrospective

Kris Brannan and David L. Bentley

Department of Biochemistry and Molecular Genetics, University of Colorado School of Medicine, Aurora, CO 80045, USA

Correspondence should be addressed to David L. Bentley, david.bentley@ucdenver.edu

Received 12 September 2011; Accepted 11 October 2011

Academic Editor: Sebastián Chávez

The origins of our current understanding of control of transcription elongation lie in pioneering experiments that mapped RNA polymerase II on viral and cellular genes. These studies first uncovered the surprising excess of polymerase molecules that we now know to be situated at the at the 5′ ends of most genes in multicellular organisms. The pileup of pol II near transcription start sites reflects a ubiquitous bottle-neck that limits elongation right at the start of the transcription elongation. Subsequent seminal work identified conserved protein factors that positively and negatively control the flux of polymerase through this bottle-neck, and make a major contribution to control of gene expression.

1. Introduction

The initiation phase of the RNA polymerase II (pol II) transcription cycle involves multiple events, including recruitment of general transcription factors and pol II to the promoter, melting of the DNA template, initiation of RNA synthesis, and pol II promoter clearance, which marks entry into the elongation phase. The stochastic nature of all of these steps poses a potential problem if it becomes necessary to mount a rapid activation of transcription. Following initiation pol II often encounters a rate-limiting barrier that appears to lie between early elongation and productive elongation. The transition between these two phases of the transcription cycle has now been characterized as a powerful regulatory switch used to increase or decrease gene expression in a signal-responsive fashion. Here we review the early discoveries that laid the foundation for a detailed understanding of transcriptional regulation at this transition.

2. Early Evidence of Polymerase Pausing and Premature Termination in DNA Viruses

Nearly 30 years ago it was reported by the late Yosef Aloni and colleagues that run-on transcripts made in nuclei from SV40 infected cells were strongly biased toward the 5′ end of the late transcription unit suggesting that pol II accumulated in the promoter-proximal region [1]. Analysis of labeled RNA extended on viral transcription complexes (VTCs) assembled *in vivo* and purified from infected cells revealed two additional unusual features of transcription from the late promoter. First, two pause sites were mapped around positions +15 and +40 relative to the start site by identifying the junctions between unlabelled RNA made *in vivo* and labeled RNA extended *in vitro* [2]. Second, a major product of transcription on VTCs is a discrete 93–95 base RNA, that is, prematurely terminated near a potential hairpin loop structure. Similar evidence for promoter-proximal stalling and/or premature termination were subsequently reported for the early and late promoters of polyoma virus [3]. These results prompted speculation that SV40 late transcription might be regulated by a mechanism [1] that regulates a decision between premature termination and productive elongation, analogous to attenuation on bacterial operons [4]. About the same time Luse and colleagues showed that transcription complexes assembled in HeLa nuclear extract on the adenovirus 2 major late promoters under NTP limiting conditions gave rise to uncapped transcripts about 20 nucleotides long that could be elongated into capped transcripts

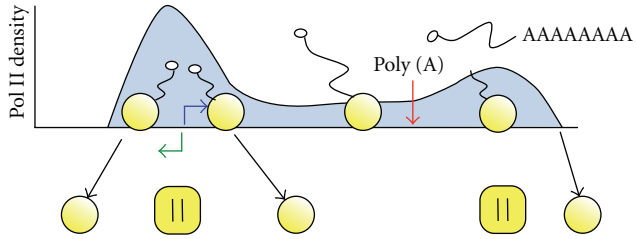


Figure 1: RNA pol II density profile across a typical metazoan protein-coding gene. Elevated density around the transcription start site (TSS) results from promoter-proximal pausing and possibly premature termination of transcription. Blue and green arrows denote divergent transcription from the TSS. A second peak of pol II accumulation downstream of the poly (A) site precedes termination coupled to cleavage/polyadenylation. Black arrows denote termination of transcription with eviction of pol II (yellow circles) from the DNA template downstream of the poly (A) site (red arrow) and possibly also in the promoter-proximal region. The mRNA cap structure is denoted by a white circle.

upon NTP addition [5]. The implication of this result is that pol II can pause at relatively discrete positions near the transcription start site and remain competent to resume elongation. They called this phenomenon "promoter-proximal pausing." Together these seminal early studies revealed quite unexpected patterns of stalling, pausing, and premature termination by host cell's pol II when it transcribes certain viral genes. The question posed by these studies was whether this unusual behavior by pol II was peculiar to viral genes or shared in common with cellular genes.

3. Pol II Pile-Ups on Cellular Genes

It was not long before the first evidence emerged that pol II also piles up near the transcription start sites of cellular genes. High levels of pol II were found to accumulate at the 5′ ends of the Drosophila heat shock gene *hsp70* [6, 7], and human *c-myc* genes even though the genes were not actively expressed [8, 9]. These 5′ polymerases were not only able to incorporate labeled NTPs in the nuclear run-on reaction but were also resistant to sarkosyl. Moreover, in some cases they were demonstrated to be associated with a single-stranded transcription bubble showing definitively that they were actively engaged on the template [10]. Subsequent run-on studies revealed that pol II was distributed with a similar strong bias in favor of the promoter-proximal region on Hsp26 and GAPDH in Drosophila [11] and adenosine deaminase, c-fos, DHFR and transthyretin genes in mammals [12–15]. As a footnote several of these early nuclear runon studies detected transcription proceeding in both directions from the start site, but the significance of this divergent transcription remained obscure [8, 16]. These results therefore showed that the pattern of pol II accumulation near start sites, first observed in DNA viruses, was common to a number of cellular genes. In fact it emerged from these early studies that pol II accumulated near the TSS of most or all cellular genes where it was localized in sufficient detail. Based on this evidence Krumm and colleagues suggested in 1995 that promoter-proximal pausing was a "general rate-limiting step" in the pol II transcription cycle [17]. Recently, this prediction has been largely borne out by ChIP-seq and Gro-Seq studies that localized pol II genome-wide and found high levels of pol II accumulation at the start sites of thousands of genes in Drosophila and human cells [18–20]. Indeed in human cell lines relatively few genes have a uniform distribution of pol II throughout their length compared to those with a promoter-proximal pol II pile-up (H. Kim, S. Kim, K. Brannan and D. Bentley unpublished observations). Promoter-proximal pol II accumulation likely involves sequence elements upstream and downstream of the start site as well as chromatin structure [21–23]. While the details of what makes pol II pile-up near start sites remain somewhat obscure, this is clearly a characteristic shared by numerous promoters (Figure 1).

4. Promoter-Proximal Pausing versus Premature Termination

What is the root cause for why pol II is so unevenly distributed across so many genes? The original *in vitro* pulse chase experiments of Coppola and colleagues showed that pol II can pause close to the start site and then resume elongation [5]. Since then, the most popular interpretation of *in vivo* polymerase mapping studies has been that they result from a similar "promoter-proximal pausing" phenomenon. That at least some promoter-proximal polymerase can resume elongation is demonstrated by nuclear runon experiments; indeed, these polymerases would not be detectable by this method if they could not elongate and incorporate labeled nucleotides. However, the possibility that some fraction of the promoter-proximal polymerases terminate prematurely and never enter the productive elongation phase cannot be eliminated. The evidence for premature termination is quite clear for the SV40 late and HIV viral genes [24, 25], but it is much less compelling for cellular genes. Prematurely terminated RNAs are a major product of *c-myc* transcription in microinjected Xenopus oocytes, but the physiological relevance of this phenomenon remains unproven [26]. Recently, short (20–90 bases) transcription start site-associated (TSS-a) sense and antisense transcripts present at very low levels in the nucleus were detected by high-throughput RNA sequencing [27]. Whether these transcripts are products of promoter-proximal premature termination or pol II pausing are interesting questions for future investigation.

5. The Function of Polymerase Accumulation at Start Sites

An important question to emerge from the early studies of pol II localization on viral and cellular genes was: "What is the purpose of pol II piling up at the start sites of genes even before they are activated?" One answer to this question quickly emerged from studies of three genes with regulated transcriptional output: the cellular *Hsp70* and *c-myc* genes [6, 8, 9] and a transfected reporter driven by the HIV1 LTR [24]. In each of these cases nuclear run-on transcription revealed

a key difference between the activated and nonactivated states: the ratio of polymerases within the gene body relative to the 5′ end increased when transcription was activated. The significance of these studies is that they showed regulation of gene expression can be exerted at the level of transcriptional elongation by controlling the fraction of polymerases that are permitted to travel beyond the promoter-proximal region. Furthermore at *Hsp70*, the amount of paused pol II prior to heat shock correlated with the amount of mRNA made after heat shock [23]. Therefore, a satisfying answer to the question of why pol II accumulates near start sites is that it provides a pool of engaged polymerases ready for rapid mobilization in response to a gene activation stimulus. A second way that localized pol II accumulation at the TSS may enhance rapid transcriptional responses is excluding nucleosomes, thereby providing a bookmark in the chromatin that can be easily accessed by the transcriptional machinery [22]. A third suggestion is that an extended pol II dwell time within the promoter proximal region allows for cotranscriptional capping of the nascent mRNA [28, 29], and could help to "license" productive elongation complexes by allowing time for recruitment of processing and elongation factors. On the other hand, there is no direct evidence that a pol II pile-up near the TSS is required for efficient capping.

6. Control of Elongation by Transcriptional Activators

How is the flux of pol II from the promoter-proximal region into the body of a gene controlled? The first important clue was again provided by a virus; in this case HIV1. Groundbreaking work of Kao and colleagues showed that the viral transactivator protein Tat had the novel ability stimulate elongation by pol II [24]. Without Tat, most polymerases that initiate from the HIV1 LTR terminate prematurely shortly downstream of the TAR hairpin loop sequence in a manner resembling the SV40 late transcription unit, but in the presence of Tat, pol II acquires the ability to extend transcripts all the way to the end of the provirus. To explain these surprising results, Kao et al. suggested that Tat regulates transcription by an antitermination mechanism similar to that exerted by the bacteriophage lambda N protein [30]. However, it remained possible that Tat also controlled transcriptional pausing, which is frequently a pre-requisite for termination.

HIV Tat is an unusual transactivator because it binds to the nascent RNA transcript. Therefore, the question remained open as to whether conventional DNA-bound activators can influence transcriptional elongation. Part of the answer to this question came with the demonstration that Tat could activate transcription when tethered to a DNA-binding site in the promoter [31]. Subsequent studies showed that enhancers and promoter-bound chimeric transcription factors comprising activation domains fused to a DNA-binding domain can stimulate elongation [32]. Furthermore a number of natural cellular activators stimulate elongation including heat-shock factor, NFkB, and *c-myc* [21, 33, 34]. Activation domains that enhance elongation and initiation, respectively, can synergize with one another and the most potent activation domains such as Herpes virus VP16 can stimulate both initiation and elongation [35, 36].

7. The Yin and Yang of Elongational Control

How do activators like HIV Tat and cellular transcription factors stimulate pol II transit away from the promoter-proximal region and into the downstream region of the gene for productive mRNA synthesis? The solution to this problem was provided by landmark studies that uncovered novel inhibitors of elongation and the factors that antagonize them. This story started with an early insight into how the ATP analogue 5, 6-dichloro-1-ß-D-ribofuranosylbenzimidazole (DRB) inhibits pol II transcription. Pulse labeling of RNA in adenovirus-infected cells revealed that DRB inhibited chain elongation but not initiation [37]. In a *tour de force* of classical biochemistry, the Handa and Price labs took advantage of this inhibitor to identify the core negative and positive factors that control the "yin and yang" of transcriptional elongation. Handa's lab identified the DRB-sensitivity-inducing factor (DSIF) as Spt4/5 a conserved pol II binding complex that is required for inhibition of elongation near 5′ ends [38]. Soon afterwards, these workers identified a second negative-elongation factor, NELF, that cooperates with DSIF [39]. The counterpart to these negative factors is positive transcription elongation factor b (PTEFb) discovered by Marshall and Price [40]. PTEFb was identified as the cyclin-dependent protein kinase complex Cdk9-CyclinT1 [41, 42] that is specifically inhibited by DRB. In a remarkable convergence of independent studies, it turned out that the negative-factors DSIF and NELF and the positive-factor PTEFb are all components of the same control system. Thus, a major function of PTEFb is to "alleviate" the negative effects of DSIF and NELF [43] which it does by phosphorylating them both as well as the pol II C-terminal domain [44, 45].

Elucidation of the interplay between positive- and negative-elongation factors provided a basis for understanding how transcription factors can regulate elongation. The vital missing piece of the puzzle was filled in with the discovery that Tat when bound to TAR in the nascent HIV1 transcript contacts PTEFb through Cyclin T1 and this interaction is required for stimulation of transcriptional elongation [41, 42, 46]. Tat-mediated recruitment of PTEFb permits modification of the paused pol II complex by phosphorylation of the pol II CTD, Spt5, and NELF resulting in a transition to productive elongation. A similar mechanism involving PTEFb-mediated antagonism of the negative-elongation factors DSIF and NELF is thought to regulate elongation at many cellular genes including *c-fos* and *NFkB* targets [45, 47]. PTEFb (Cdk9/CyclinT1) is found embedded in multiple complexes with different protein and RNA subunits [48, 49] and there are likely to be multiple ways that it can be recruited to genes. These include binding directly to transcription factors [33] and chromatin components [50].

8. Concluding Remarks

Tremendous advances have been made in understanding control of gene expression at the level of transcriptional

elongation since the early days when it was identified on a few viral and cellular genes. Now this mechanism is recognized to be at least as important as control of the initiation step in pol II transcription. Still, important questions remain unresolved about the nature of promoter-proximally accumulated pol II. It is still not clear how many of these paused polymerases have backtracked and are destined ultimately to resume elongation and how many are destined for premature termination. These scenarios suggest the possibility of distinct targets for regulation by controlled polymerase release into the body of the gene. It will be interesting to see how these targets might be used in various developmental and signal-responsive contexts.

Acknowledgments

Research in the authors' lab is supported by NIH Grant GM063873 to D. Bentley. K. Brannan was supported by NIHT32-GM08730.

References

[1] N. Hay, H. Skolnik-David, and Y. Aloni, "Attenuation in the control of SV40 gene expression," *Cell*, vol. 29, no. 1, pp. 183–193, 1982.

[2] H. Skolnik-David and Y. Aloni, "Pausing of RNA polymerase molecules during in vivo transcription of the SV40 leader region," *EMBO Journal*, vol. 2, no. 2, pp. 179–184, 1983.

[3] W. C. Skarnes, D. C. Tessier, and N. H. Acheson, "RNA polymerases stall and/or prematurely terminate nearby both early and late promoters on polyomavirus DNA," *Journal of Molecular Biology*, vol. 203, no. 1, pp. 153–171, 1988.

[4] C. Yanofsky, "Transcription attenuation: once viewed as a novel regulatory strategy," *Journal of Bacteriology*, vol. 182, no. 1, pp. 1–8, 2000.

[5] J. A. Coppola, A. S. Field, and D. S. Luse, "Promoter-proximal pausing by RNA polymerase II in vitro: transcripts shorter than 20 nucleotides are not capped," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 80, no. 5, pp. 1251–1255, 1983.

[6] D. S. Gilmour and J. T. Lis, "RNA polymerase II interacts with the promoter region of the noninduced hsp70 gene in Drosophila melanogaster cells," *Molecular and Cellular Biology*, vol. 6, no. 11, pp. 3984–3989, 1986.

[7] A. E. Rougvie and J. T. Lis, "The RNA polymerase II molecule at the 5′ end of the uninduced hsp70 gene of D. melanogaster is transcriptionally engaged," *Cell*, vol. 54, no. 6, pp. 795–804, 1988.

[8] D. L. Bentley and M. Groudine, "A block to elongation is largely responsible for decreased transcription of c-myc in differentiated HL60 cells," *Nature*, vol. 321, no. 6071, pp. 702–706, 1986.

[9] D. Eick and G. W. Bornkamm, "Transcriptional arrest within the first exon is a fast control mechanism in c-myc gene expression," *Nucleic Acids Research*, vol. 14, no. 21, pp. 8331–8346, 1986.

[10] C. Giardina, R. M. Perez, and J. T. Lis, "Promoter melting and TFIID complexes on Drosophila genes in vivo," *Genes and Development*, vol. 6, no. 11, pp. 2190–2200, 1992.

[11] A. E. Rougvie and J. T. Lis, "Postinitiation transcriptional control in Drosophila melanogaster," *Molecular and Cellular Biology*, vol. 10, no. 11, pp. 6041–6045, 1990.

[12] Z. Chen, M. L. Harless, D. A. Wright, and R. E. Kellems, "Identification and characterization of transcriptional arrest sites in exon 1 of the human adenosine deaminase gene," *Molecular and Cellular Biology*, vol. 10, no. 9, pp. 4555–4564, 1990.

[13] M. A. Collart, N. Tourkine, D. Belin, P. Vassalli, P. Jeanteur, and J. M. Blanchard, "c-fos Gene transcription in murine macrophages is modulated by a calcium-dependent block to elongation in intron 1," *Molecular and Cellular Biology*, vol. 11, no. 5, pp. 2826–2831, 1991.

[14] J. Mirkovitch and J. E. Darnell, "Mapping of RNA polymerase on mammalian genes in cells and nuclei," *Molecular Biology of the Cell*, vol. 3, no. 10, pp. 1085–1094, 1992.

[15] L. J. Schilling and P. J. Farnham, "Inappropriate transcription from the 5' end of the murine dihydrofolate reductase gene masks transcriptional regulation," *Nucleic Acids Research*, vol. 22, no. 15, pp. 3061–3068, 1994.

[16] G. F. Crouse, E. J. Leys, and R. N. McEwan, "Analysis of the mouse dhfr promoter region: existence of a divergently transcribed gene," *Molecular and Cellular Biology*, vol. 5, no. 8, pp. 1847–1858, 1985.

[17] A. Krumm, L. B. Hickey, and M. Groudine, "Promoter-proximal pausing of RNA polymerase II defines a general rate-limiting step after transcription initiation," *Genes and Development*, vol. 9, no. 5, pp. 559–572, 1995.

[18] M. G. Guenther, S. S. Levine, L. A. Boyer, R. Jaenisch, and R. A. Young, "A chromatin landmark and transcription initiation at most promoters in human cells," *Cell*, vol. 130, no. 1, pp. 77–88, 2007.

[19] G. W. Muse, D. A. Gilchrist, S. Nechaev et al., "RNA polymerase is poised for activation across the genome," *Nature Genetics*, vol. 39, no. 12, pp. 1507–1511, 2007.

[20] L. J. Core, J. J. Waterfall, and J. T. Lis, "Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters," *Science*, vol. 322, no. 5909, pp. 1845–1848, 2008.

[21] S. A. Brown, A. N. Imbalzano, and R. E. Kingston, "Activator-dependent regulation of transcriptional pausing on nucleosomal templates," *Genes and Development*, vol. 10, no. 12, pp. 1479–1490, 1996.

[22] D. A. Gilchrist, G. Dos Santos, D. C. Fargo et al., "Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation," *Cell*, vol. 143, no. 4, pp. 540–551, 2010.

[23] H. S. Lee, K. W. Kraus, M. F. Wolfner, and J. T. Lis, "DNA sequence requirements for generating paused polymerase at the start of hsp70," *Genes and Development*, vol. 6, no. 2, pp. 284–295, 1992.

[24] S. Y. Kao, A. F. Calman, P. A. Luciw, and B. M. Peterlin, "Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product," *Nature*, vol. 330, no. 6147, pp. 489–493, 1987.

[25] H. Skolnik-David, N. Hay, and Y. Aloni, "Site of premature termination of late transcription of simian virus 40 DNA: enhancement by 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 79, no. 9, pp. 2743–2747, 1982.

[26] D. L. Bentley and M. Groudine, "Sequence requirements for premature termination of transcription in the human c-myc gene," *Cell*, vol. 53, no. 2, pp. 245–256, 1988.

[27] A. C. Seila, J. M. Calabrese, S. S. Levine et al., "Divergent transcription from active promoters," *Science*, vol. 322, no. 5909, pp. 1849–1851, 2008.

[28] Y. Pei, B. Schwer, and S. Shuman, "Interactions between fission yeast Cdk9, its cyclin partner Pch1, and mRNA capping

enzyme Pct1 suggest an elongation checkpoint for mRNA quality control," *The Journal of Biological Chemistry*, vol. 278, no. 9, pp. 7180–7188, 2003.

[29] 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.

[30] J. Greenblatt, J. R. Nodwell, and S. W. Mason, "Transcriptional antitermination," *Nature*, vol. 364, no. 6436, pp. 401–406, 1993.

[31] C. D. Southgate and M. R. Green, "The HIV-1 Tat protein activates transcription from an upstream DNA-binding site: implications for Tat function," *Genes and Development*, vol. 5, no. 12, pp. 2496–2507, 1991.

[32] K. Yankulov, J. Blau, T. Purton, S. Roberts, and D. L. Bentley, "Transcriptional elongation by RNA polymerase II is stimulated by transactivators," *Cell*, vol. 77, no. 5, pp. 749–759, 1994.

[33] M. Barboric, R. M. Nissen, S. Kanazawa, N. Jabrane-Ferrat, and B. M. Peterlin, "NF-κB binds P-TEFb to stimulate transcriptional elongation by RNA polymerase II," *Molecular Cell*, vol. 8, no. 2, pp. 327–337, 2001.

[34] P. B. Rahl, C. Y. Lin, A. C. Seila et al., "C-Myc regulates transcriptional pause release," *Cell*, vol. 141, no. 3, pp. 432–445, 2010.

[35] W. S. Blair, R. A. Fridell, and B. R. Cullen, "Synergistic enhancement of both initiation and elongation by acidic transcription activation domains," *EMBO Journal*, vol. 15, no. 7, pp. 1658–1665, 1996.

[36] J. Blau, H. Xiao, S. McCracken, P. O'Hare, J. Greenblatt, and D. Bentley, "Three functional classes of transcriptional activation domains," *Molecular and Cellular Biology*, vol. 16, no. 5, pp. 2044–2055, 1996.

[37] N. W. Fraser, P. B. Sehgal, and J. E. Darnell, "DRB-induced premature termination of late adenovirus transcription," *Nature*, vol. 272, no. 5654, pp. 590–593, 1978.

[38] T. Wada, T. Takagi, Y. Yamaguchi et al., "DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs," *Genes and Development*, vol. 12, no. 3, pp. 343–356, 1998.

[39] Y. Yamaguchi, T. Takagi, T. Wada et al., "NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation," *Cell*, vol. 97, no. 1, pp. 41–51, 1999.

[40] 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.

[41] P. Wei, M. E. Garber, S. M. Fang, W. H. Fischer, and K. A. Jones, "A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA," *Cell*, vol. 92, no. 4, pp. 451–462, 1998.

[42] Y. Zhu, T. Peery, 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.

[43] T. Wada, T. Takagi, Y. Yamaguchi, D. Watanabe, and H. Handa, "Evidence that P-TEFb alleviates the negative effect of DSIF on RNA polymerase II-dependent transcription in vitro," *EMBO Journal*, vol. 17, no. 24, pp. 7395–7403, 1998.

[44] 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," *The Journal of Biological Chemistry*, vol. 271, no. 43, pp. 27176–27183, 1996.

[45] T. Yamada, Y. Yamaguchi, N. Inukai, S. Okamoto, T. Mura, and H. Handa, "P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation," *Molecular Cell*, vol. 21, no. 2, pp. 227–237, 2006.

[46] 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.

[47] L. Amir-Zilberstein, E. Ainbinder, L. Toube, Y. Yamaguchi, H. Handa, and R. Dikstein, "Differential regulation of NF-κB by elongation factors is determined by core promoter type," *Molecular and Cellular Biology*, vol. 27, no. 14, pp. 5246–5259, 2007.

[48] 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.

[49] E. Smith, C. Lin, and A. Shilatifard, "The super elongation complex (SEC) and MLL in development and disease," *Genes and Development*, vol. 25, no. 7, pp. 661–672, 2011.

[50] Z. Yang, J. H. N. Yik, R. Chen et al., "Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4," *Molecular Cell*, vol. 19, no. 4, pp. 535–545, 2005.

Hindawi Publishing Corporation
Genetics Research International
Volume 2012, Article ID 276948, 13 pages
doi:10.1155/2012/276948

Review Article

Regulation of Ribosomal RNA Production by RNA Polymerase I: Does Elongation Come First?

Benjamin Albert,[1,2] Jorge Perez-Fernandez,[1,2,3] Isabelle Léger-Silvestre,[1,2] and Olivier Gadal[1,2]

[1] *LBME du CNRS, 118 route de Narbonne, 31000 Toulouse, France*
[2] *Laboratoire de Biologie Moleculaire Eucaryote, Université de Toulouse, 118 route de Narbonne, 31000 Toulouse, France*
[3] *Universität Regensburg, Biochemie-Zentrum Regensburg (BZR), Lehrstuhl Biochemie III, 93053 Regensburg, Germany*

Correspondence should be addressed to Olivier Gadal, olivier.gadal@ibcg.biotoul.fr

Received 31 August 2011; Accepted 27 September 2011

Academic Editor: Sebastián Chávez

Ribosomal RNA (rRNA) production represents the most active transcription in the cell. Synthesis of the large rRNA precursors (35–47S) can be achieved by up to 150 RNA polymerase I (Pol I) enzymes simultaneously transcribing each rRNA gene. In this paper, we present recent advances made in understanding the regulatory mechanisms that control elongation. Built-in Pol I elongation factors, such as Rpa34/Rpa49 in budding yeast and PAF53/CAST in humans, are instrumental to the extremely high rate of rRNA production per gene. rRNA elongation mechanisms are intrinsically linked to chromatin structure and to the higher-order organization of the rRNA genes (rDNA). Factors such as Hmo1 in yeast and UBF1 in humans are key players in rDNA chromatin structure *in vivo*. Finally, elongation factors known to regulate messengers RNA production by RNA polymerase II are also involved in rRNA production and work cooperatively with Rpa49 *in vivo*.

1. Introduction

In cell nuclei, three RNA polymerases transcribe the genome. The most importance is placed on RNA polymerase II (Pol II), which is responsible for synthesizing mRNA and a large variety of noncoding RNAs. The vast majority of RNA production in growing cells is carried out by RNA polymerase I (Pol I), which transcribes the precursor of large rRNA, and by RNA polymerase III (Pol III), which transcribes 5S rRNA, tRNA, and some noncoding RNAs. Observation of cryofixed cryosubstituted other sections analyzed by electron microscopy reveals that exponentially growing budding yeast cells contain up to 10^4 ribosomes per μm^3 [1], which represents up to 10% of the cytoplasmic volume [2] (Figure 1(a)).

An early step in ribosome biogenesis is initiated by the extremely high transcriptional activity of Pol I and occurs in the largest nuclear domain, the nucleolus (Figure 1(b)). Electron microscopy of nuclear/nucleolar chromatin dispersed by Miller spreading allowed rRNA gene transcription and cotranscriptional assembly to be visualized directly at the single gene level [3] (Figure 1(d)). rDNA is organized in head-to-tail tandem arrays of rRNA genes [4] contained in budding yeast between 100 and 200 copies per cell [5], and from 200 to 300 per mammalian haploid genome [6]. Analysis of transcribed ribosomal DNA (rDNA) after Miller spreading revealed that up to 150 Pol I enzymes simultaneously transcribe rRNA genes in mutant with only 25 rRNA genes [1, 7] (Figures 1(c) and 1(d)). Importantly, despite being the most highly transcribed genes of the genome, rDNA is subject to epigenetic regulation, and only some rRNA genes are transcriptionally active [8]. In this paper, we will focus on recent advances made in understanding the regulation of Pol I activity, including elongation in the context of ribosome assembly.

2. Is Regulation of Pol I Initiation the Only Regulated Step in rRNA Production *In Vivo*?

Eukaryotic RNA polymerases are able to recognize promoters only when these sequence elements are associated with

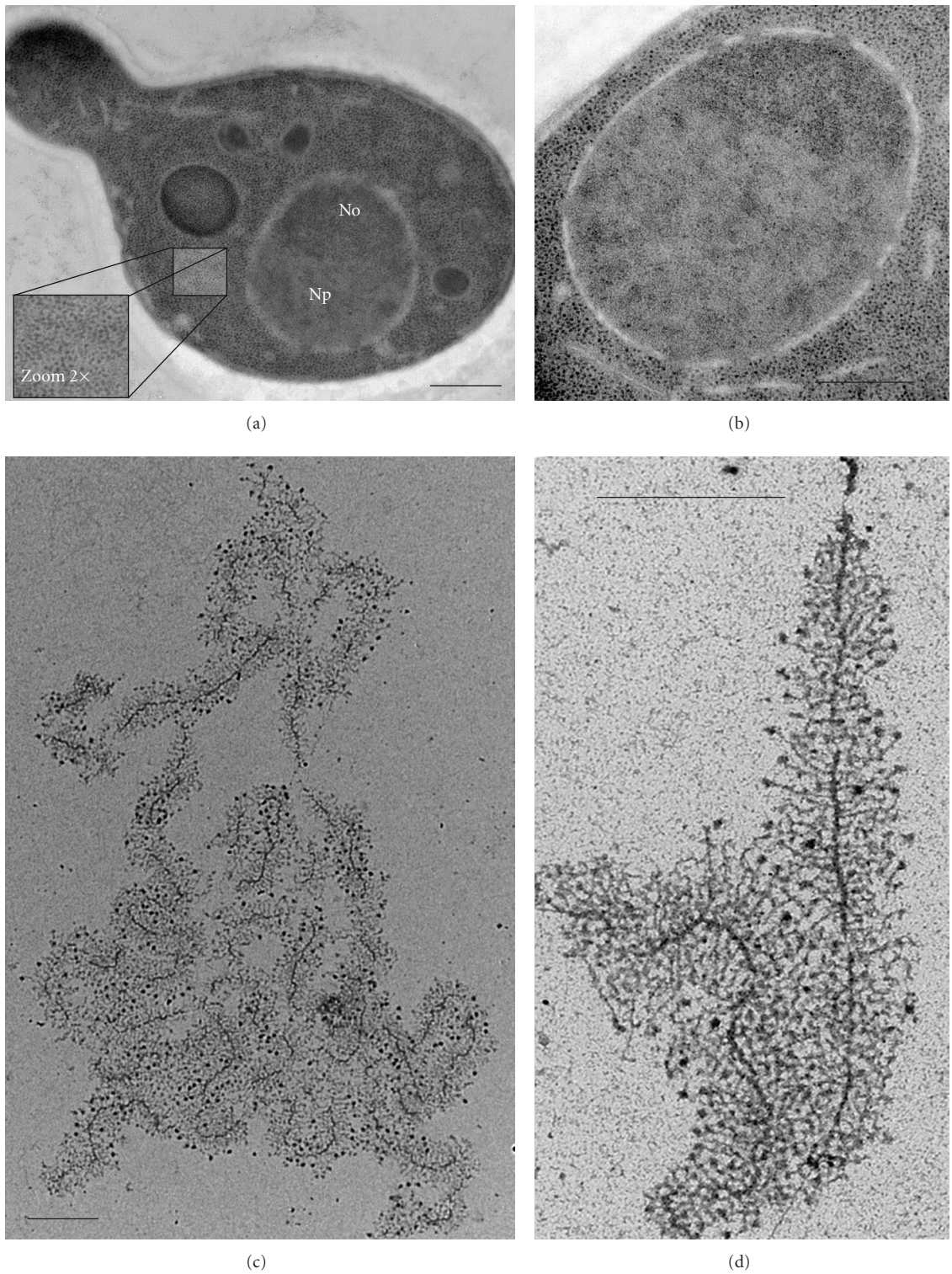


Figure 1: Budding yeast cells and ribosome production. (a) Morphology of *Saccharomyces cerevisiae* cells after cryofixation and freeze substitution. Ribosomes are individually localized in the cytoplasm (see individual ribosomes detected in the zoomed region). In the nucleus, the nucleolus (No) is detected as a large electron-dense region compared with low electron density of the nucleoplasm (Np). (b) Morphology of the nucleolus. The nucleus appears outlined by a double envelope with pores, and the nucleolus is in close contact with the nuclear envelope. In the nucleolus, a dense fibrillar network is visible throughout the nucleolar volume. Granular components are dispersed throughout the rest of the nucleolus. (c) Visualization of active genes in rDNA. Using a mutant strain with a reduced number of rDNA copies (strain NOY1071; 25 rDNA copies), Miller spreading of total nucleolar DNA allowed single-gene analysis of rRNA genes. (d) Quantification of actively transcribed rDNA. Using high magnification, we can detect individual polymerases associated with nascent rRNA. Bars represent 500 nm.

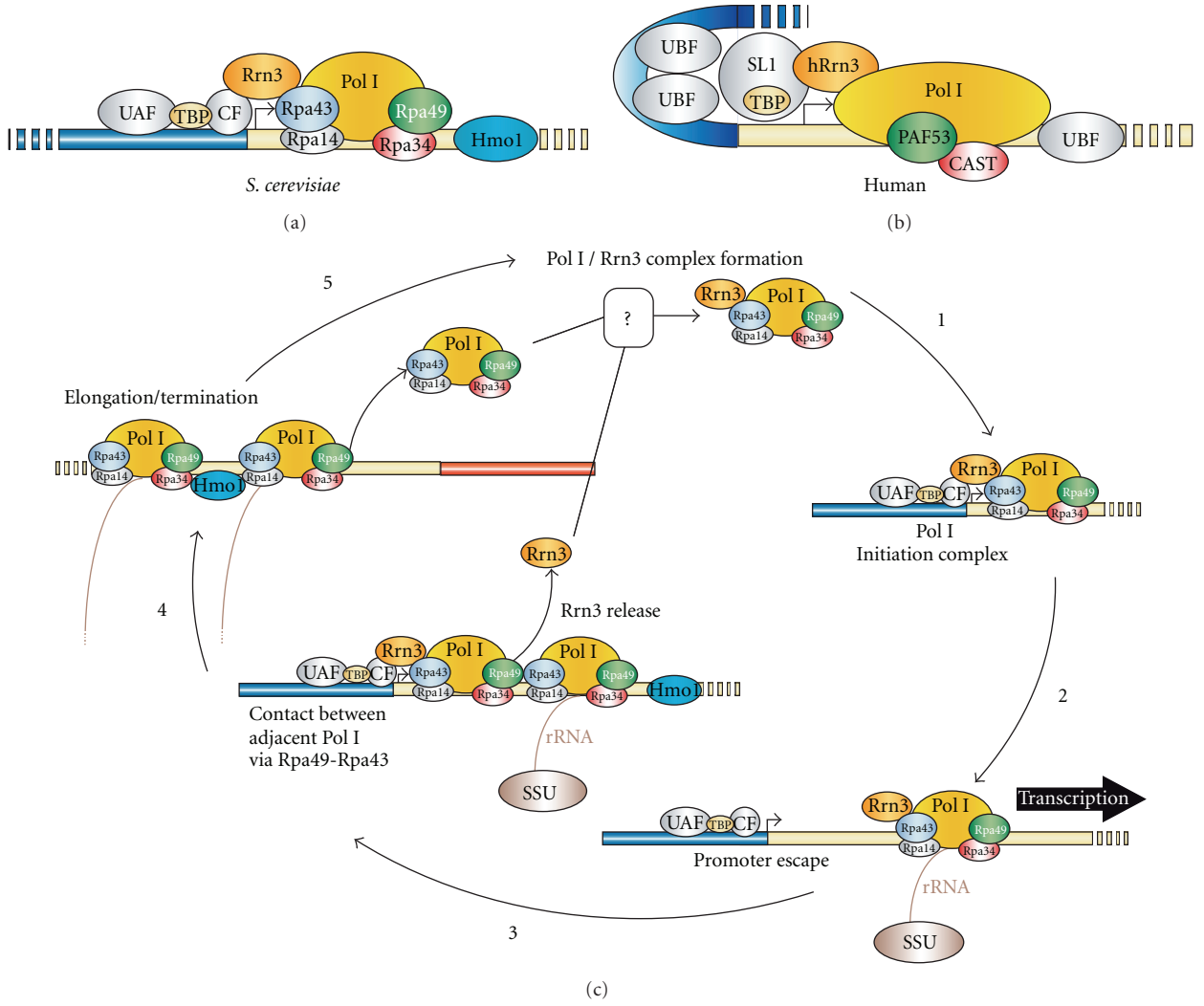


Figure 2: Schematic representation of the Pol I transcription cycle. Simplified composition of the Pol I preinitiation complex (PIC) in (a) budding yeast and (b) human cells. The Pol I transcription cycle in budding yeast. (1) Recruitment of Rrn3/Pol I onto an rDNA promoter associated with the SL1 and UAF complex allows PIC formation. (2) Promoter escape and rRNA synthesis are coupled with cotranscriptional recruitment of the SSU processome. (3) Rrn3 dissociation is achieved by the formation of an adjacent PIC. Pol I subunits Rpa49 and Rpa43 from the adjacent polymerases promote Rrn3 release from the transcribing Pol I. (4) Pol I transcription of rRNA is coupled with nascent rRNA processing and termination. (5) Pol I holoenzyme is recycled by reassociation with Rrn3, an as yet uncharacterized regulatory process. Hmo1 function during elongation remains to be clarified, but is revealed by a tight genetic interaction with Pol I elongation mutant *rpa49*Δ (4 and 5).

specific initiation factors. Pol I initiation factors have been characterized for both humans and yeast (Figure 2(a)). In mammals, selectivity factor 1 (SL1) in humans and TIF-1B in mice are composed of the TATA-binding protein (TBP) and four TBP-associated factors (TAFs), bound to the core promoter [9–14]. Upstream binding factor (UBF) acts as a dimer and induces a loop formation called the enhanceosome, which brings the activating sequence into close proximity with the core promoter element [15, 16]. UBF binding stabilizes the association of SL1/TIF-1B with promoter elements [9]. A recent study suggested that UBF bound after SL1 binding and during promoter escape by Pol I [17]. UBF is also bound to the transcribed region [18] and can regulate Pol I elongation [19]. Additionally, UBF and SL1 are regulated by posttranslational modifications. Active Pol I enzymes are associated with numerous other factors such as TFIIH, protein kinase CK2, nuclear actin, nuclear myosin 1 (NM1), chromatin modifiers G9a and SIRT7 and with proteins involved in replication and DNA repair: Ku70/80, proliferating cell nuclear antigen, and CSB. For most factors, mechanistical insights are lacking (see [20] for a recent review).

In budding yeast, a core factor (CF) associates with the Pol I promoter, and this binding is stabilized via TBP by an upstream-associated factor, or UAF [21–27]. CF and SL1 are likely to be functionally equivalent. In contrast, yeast UAF and mammalian UBF both interact with upstream stimulatory elements but have very different functions. UBF1

also regulates Pol I elongation [19]. The *S. cerevisiae* HMG-Box protein, Hmo1, is associated with the Pol I-transcribed region and is able to rescue growth of the Pol I elongation mutant *rpa49*Δ [28]. Therefore, UBF1 and Hmo1 might have a conserved function in stimulating Pol I elongation (Albert et al. submitted).

Surprisingly, both human and yeast Pol I enzymes are unable to initiate productive RNA synthesis with only promoter-bound factors [29, 30]. Only a minor fraction of free Pol I is associated with an additional initiation factor: Rrn3 in yeast, hRrn3 in humans, or TIF-1A in mice (Figure 2(b)). When Pol I associates with one of these factors, it recognizes the promoter-bound factors and forms a preinitiation complex (PIC) [30–35].

The amount of Pol I-Rrn3p complexes represents a limiting step in transcription initiation, but how this association is achieved and regulated remains a major research topic [36]. Numerous signaling pathways target Pol I activity *in vivo*. The target of rapamycin complex 1 (TORC1) regulates ribosome production in response to nutriment availability [37]. Upon inhibition of TORC1 by rapamycin or during stationary phase, the amount of Pol I-Rrn3 complex drops [31, 38, 39]. The regulatory function of Rrn3's association with Pol I was demonstrated by producing an artificial fusion protein of Rrn3 joined to its interacting Pol I subunit, Rpa43. In a partially purified *in vitro* system, this fusion, called CARA, led to a constitutively active Pol I even during stress, showing that Pol I complexed with Rrn3 is initiation competent even under conditions known to inhibit ribosome production [39]. This initial observation suggested that a deregulated initiation event is sufficient to generate constitutively active Pol I *in vivo*. However, other findings are now challenging this initial regulatory model based on the availability of an Rrn3-Pol I complex. Rrn3 function is not restricted to initiation only, and it is also involved in a postinitiation step of the Pol I transcription cycle. Rrn3 is released from Pol I during postinitiation, and this process requires Rpa49, another Pol I-specific subunit [40]. In the absence of Rpa49, the CARA mutant is not viable [40]. Therefore, when Rrn3 is physically tethered to Pol I, Rpa49 function becomes essential [40], which suggests the existence of a functional interaction between Rpa49 and Rrn3 after Pol I recruitment [1]. Initial studies suggested that the interaction between Rrn3 and another Pol I-specific subunit, Rpa43, is regulated by phosphorylation [34, 41]. A mutational analysis of Pol I did not reveal the specific residues involved in this regulation but did not exclude the involvement of phosphorylation [42]. Recent works from the Tschochner's Laboratory have demonstrated that Rrn3 is destabilized by a PEST domain, a peptide sequence rich in proline (P), glutamate (E), serine (S), and threonine (T), in its N-terminal domain [43, 44]. A nondegradable form of Rrn3, missing this PEST motif, attenuated the reduction in initiation competent Pol I-Rrn3p complexes observed upon nutrient depletion. Such a mutation should mimic the CARA mutant phenotype. Unfortunately, this non-degradable form of Rrn3 associated with Pol I has not been tested *in vitro* in a partially purified extract. Nevertheless, in this background, rRNA synthesis was downregulated *in vivo* upon nutrient depletion [44]. Additionally, although levels of the Rrn3-Pol I complex are depleted during stress, the amount remaining is sufficient to observe ongoing initiation events. Therefore, Pol I activity is not regulated only by the initiation competent Rrn3-Pol I complex, but is likely to be influenced by nascent ribosomal assembly. An elegant study suggested that down-regulation of ribosomal protein production could also result in a rapid decay of newly made rRNA *in vivo* [45]. Indeed, Sch9, which acts downstream of TORC1, targets ribosomal protein gene (RPG) transcription as well as rRNA production by Pol I [46, 47]. Rrn3 might also impact rRNA processing since a mutant of Cbf5, the pseudouridine synthetase that modifies rRNA, is rescued by Rrn3 overexpression [48]. Along the same lines, accumulation of RPG mRNA in the CARA mutant background was resistant to repression by TORC1 inhibition [39]. In fission yeast, a subunit of the Rrn7 core factor also binds RPG promoters, suggesting a coupling between rRNA production and RPG transcription [49]. In budding yeast, Hmo1, *bona fide* Pol I transcription factors, also bind most RPG promoters [50]. Stochiometric production of all ribosomal constituents is tightly controlled and is probably achieved at multiple levels [2]. *Is Pol I initiation the only regulated step in rRNA production in vivo?* Although the association of Rrn3 with Pol I is a very important regulatory step, it is only one of the numerous pathways that regulate Pol I activity. In this paper, we will extensively describe how the rRNA elongation step might be regulated to integrate all the complex processes necessary to achieve this early step of ribosome assembly.

3. Early Ribosome Assembly Occurs during rDNA Transcription

RNA synthesis in the nucleus is invariably coupled with the recruitment of specific proteins shortly after synthesis and leads to the formation of large ribonucleoproteins (RNPs). Strikingly, the fate of the RNA depends on the RNA polymerase synthesizing the transcript. Through their association with the transcribing polymerase appropriate RNA-interacting proteins are driven into the local proximity of the newly synthesized RNA. The COOH-terminal domain (CTD) of the largest subunit of Pol II is the best example of this mechanism [51]. The CTD can recruit the pre-mRNA capping, splicing, and 3′-processing machinery, which are then tethered together with the pre-mRNA. Often, the same factors affecting early RNP assembly, maturation, and export also regulate Pol II elongation, which effectively bridges these processes [52, 53]. Cotranscriptional assemblies of RNP particles are well known to impact the fate of the transcribed RNA. Even more important, polymerase elongation rates can determine the nature of mature mRNA products, as shown by alternative splicing that depends on the elongation rate [54].

Pol I transcription provided the first example of cotranscriptional RNP particle assembly [3]. Paradoxically, the intimate connection between early RNP assembly and Pol I elongation has been suggested only recently. From Oskar

Miller's original 1969 description of a transcribed rRNA gene as a Christmas tree, in which the nascent RNA cotranscriptionally assembled with maturation factors that appeared as decorating "terminal balls," 33 years elapsed before the molecular nature of the terminal balls in budding yeast was fully unveiled by the laboratories of Ann Beyer and Susan Baserga [3]. The terminal balls were renamed SSU processomes [55] and are early preribosomal particles, which contain U3 snoRNA and a set of proteins called the UTPs (U three proteins). An intimate relationship between early assembly and transcription was then suggested from a study of a UTP subgroup, the tUTPs, for transcription-UTP [56]. The tUTPs form a complex with a protein composition very similar to the UTP-A complex [57] with only one distinction: the presence of either Utp5 or Pol5, respectively. tUTP/UTP-A is recruited to the chromatin independently of transcription; is required for efficient accumulation of Pol I transcripts; has been suggested to be essential for Pol I transcription in a run-on assay [56]. Alternatively, tUTP was suggested to be required for early pre-rRNA stabilization. In the absence of being complexed with tUTP/UTP-A, nascent RNA transcripts are targeted by TRAMP (Trf4/Air2/Mtr4p Polyadenylation) complex and degraded by the nuclear exosome [58]. In human cells, recruitment of human tUTP orthologs to NORs (*nucleolus organizer regions*) occurs independently of transcription but depends on the protein UBF [59]. Miller spreads have revealed that nascent rRNAs are not only assembled cotranscriptionally with a large set of proteins, but are also cleaved cotranscriptionally in *E. coli* [60], *Dyctiostelium* [61], and budding yeast [62]. This cotranscriptional rRNA cleavage has been independently confirmed and quantified using *in vivo* labeling approaches [63]. Following tUTP/UTP-A recruitment, cotranscriptional assembly is a stepwise and highly hierarchical process [64] that utilizes preexisting autonomous building blocks. These entities sequentially interact with the pre-rRNA and exhibit different interdependencies with respect to each other. This model has been validated and extended by different groups [65].

4. Pol I Elongation Rate, rRNA Cotranscriptional Maturation, and Topological Stress

When driven by a strong Pol II promoter, ribosomal DNA can be transcribed by Pol II [66]. Functional ribosomes can ultimately be produced without Pol I, but this is accompanied by a drastic phenotype. Yeast lacking Pol I activity can survive with such an artificial rDNA construct but grow very poorly (i.e., have a doubling time 4 to 5 times longer than normal) and have a massively altered nucleolar and nuclear morphologies [67, 68]. Despite being not strictly essential, the Pol I elongation rate must be properly controlled for efficient ribosome production.

Biochemical purification of Pol I activity copurified a large fraction of the rRNA maturation machinery [69]. Early maturation factors such as Prp43, Nop56, Nop58, or Gno1 have been shown by two-hybrid assay to be physically tethered to the Pol I-specific subunit Rpa34 [40]. The loading of tUTP/UTP-A onto nascent transcripts is facilitated by its association with Pol I promoters prior to pre-rRNA [56, 59] and is required either for transcription or for stabilizing nascent transcripts [58]. The first direct evidence that elongation rate regulation is required for proper ribosomal assembly came from a study of an elongation mutant by David Schneider in the laboratory of Masayasu Nomura [70]. In a mutant bearing a single substitution (*rpa135-D784G)* in the second largest subunit of Pol I, near the active center of the enzyme, rRNA processing was affected with the greatest defect occurring in the 60S assembly pathway [70]. This suggested that the stochiometric production of the 40S and 60S subunits depends on the elongation rate of Pol I.

The structural analysis comparing the three nuclear RNA polymerases revealed one feature that distinguished Pol I from the other two (Pol II and Pol III): the presence of the heterodimer Rpa49/Rpa34 [71]. Rpa49 and Rpa34 form a heterodimer that is structurally similar to the one formed by two Pol II transcription factors, TFIIE and TFIIF. The hPAF53/CAST heterodimer is the human ortholog of Rpa49/Rpa34 [72]. We recently uncovered how Rpa49 releases Rrn3 from Pol I during elongation, despite the diametric opposition of these proteins in the Pol I complex [1, 71]. With such a configuration, Rpa49 and Rrn3 are unlikely to interact directly with each other in the Pol I complex. However, the very high loading rate of Pol I per rRNA gene leads to extensive contact between adjacent Pol I enzymes (Figure 1(d)). Interactions between adjacent Pol I enzymes along the rDNA place Rpa49 and Rpa43 in close proximity, thus allowing Rpa49 to interact with the Rpa43 subunit on a nearby Pol I to promote the release of Rrn3 from Rpa43 [1]. In our model, the Pol I/Rrn3 complex dissociates only after another Pol I starts transcribing rRNA and reflects cooperativity between Pol I enzymes (Figure 2(b)). When we measured distribution of Pol I along rDNA using Miller spreading, we observed a striking degree of enzyme clustering, with 60% of the enzymes in direct contact, which is compatible with our model. Additionally, this clustering was inhibited by deletion of the Rpa49 gene. *What are the consequences of this clustering?* By opening the DNA duplex, DNA in front of RNA polymerase becomes overwound, or positively supercoiled; the DNA behind the polymerase becomes underwound, or negatively supercoiled [73]. Importantly, when multiple polymerases are in close proximity, topological distortion of the DNA duplex in front of and behind each adjacent polymerase would be compensated, resulting in topological stress only in front of the first polymerase, and after the last. Therefore, the absence of clustering of adjacent polymerases in the *rpa49*Δ mutant should lead to massive rDNA supercoiling. Indeed, Rpa49 deletion has a tight genetic relationship with the two type I yeast topoisomerases: Top1 and Top3 [40, 74]. Top3, in complex with Sgs1, is required for the stability of rDNA *in vivo* and is genetically linked to Rpa49 [74, 75]. Top1 is also involved in rDNA stability [76] and seems to be directly involved in rRNA production. Rpa34 has also been found to interact directly with Top1 in two-hybrid assays [40].

The extremely high transcription rate of Pol I should lead to extensive torsional stress on the rDNA template.

Is topological stress a selective feature of Pol I transcription? A good answer to this question came from the study using actinomycin D, a DNA-intercalating agent widely used *in vivo* as a Pol I inhibitor in metazoan cells. The transcription of rRNA genes in mammalian cells is about 50–100-fold more sensitive to actinomycin D than the synthesis of small RNAs and heterogeneous nuclear RNA [77]. However, Pol I-selective inhibition by actinomycin D is not inherent to the Pol I transcriptional machinery. Actinomycin D has no effect at low concentrations *in vitro* or in a transfected reporter system [78]. In fact, low levels of actinomycin D stimulate Pol I rRNA initiation events *in vivo*; Pol I elongation, however, is strongly inhibited [79]. Such surprising activity can be explained by the ability of low actinomycin D concentrations to stabilize DNA/topoisomerase interactions [80]. The specific inhibition of rRNA synthesis is at least in part due to the close relationship between topoisomerase and Pol I activities *in vivo*.

In yeast, a *top1* deletion mutant led to the formation of R-loops (DNA-RNA hybrids) within transcribed rRNA genes [81]. These rRNA production defects, with accumulation of rRNA intermediates ending in G-rich stretches of the 18S rRNA, are similar to those observed in the Pol I elongation mutants *rpa135-D784G* and *rpa49*Δ (discussed in [70, 81]). Additionally, it was reported that the maximum transcription rate of rDNA by Pol I in a *top1* deletion mutant generates massive negative supercoiling of rDNA template, which was revealed by the presence of DNA template melting in an A-T-rich region of rDNA that was detectable in Miller spreads [82]. Top2 is essential for growth, but *top2* mutants primarily succumb to mitotic failure. Top2 function in rRNA transcription can be studied in Top2 mutants kept in G1 phase and prevented from entering mitosis. Top1 and Top2 cooperate in Pol I transcription [83] but in contrast to a *top1* deletion mutant, a *top2-ts* mutant accumulates positive supercoiling, which inhibits elongation by Pol I [82]. The difference between these Top1 and Top2 phenotypes is surprising since both topoisomerases are able to relax positively and negatively supercoiled DNA. The difference may be related to the intrinsic properties of each enzyme and to the ability of the chromatin to transiently absorb either negative or positive torsional stress [84]. Top1 is a torque-sensitive topoisomerase with a poor ability to relax chromatin. It functions mainly in the relaxation of the negative supercoiling produced in the wake of Pol I [85, 86]. In contrast, Top2 is likely to be more efficient in relaxing the positive supercoiling produced ahead of Pol I, because no sink for torsional stress, such as melting of the DNA template, exists in this case [84]. Finally, genetic interactions between Top1 and Trf4/Trf5, two members of the TRAMP complex [87], have been reported previously, and these might be explained by the presence of unresolved R-loops that accumulate in the absence of nuclear rRNA degradation machinery.

5. Epigenetic Regulation and Chromatin Status of rDNA

In eukaryotes, rRNA genes are either organized in a closed chromatin state in which they are transcriptionally inactive in transcription or are in an open chromatin state [8]. In mammals, inactive rRNA genes are subject to DNA methylation (for review, see [20]). In budding yeast, which lacks DNA methylation, a proportion of rRNA genes is also transcriptionally repressed.

Active rDNA with an open chromatin structure can be distinguished from inactive rRNA genes by psoralen crosslinking. Psoralen is an intercalating reagent which crosslinks to DNA under UV irradiation. The nucleosomes associated with inactive DNA lock the DNA topology and prevent psoralens from crosslinking to DNA. When rDNA is subjected to psoralen treatment and crosslinking, two types of bands are detected: a slow- and a fast-migrating band [8, 88]. The molecular identity of the proteins associated with both types of rDNA molecules was unambiguously established using psoralen combined with ChEC (chromatin endogenous cleavage) [89, 90]. In ChEC methods, rDNA-associated proteins are expressed as recombinant proteins fused to micrococcal nuclease (MNase). After formaldehyde crosslinking and psoralen treatment, MNase is activated and cleaves either the fast- or slow-migrating rDNA band. The fast-migrating band corresponds to transcriptionally inactive rDNA, which is enriched in nucleosomes [90]. The slow-migrating band is transcriptionally competent rDNA, enriched in Pol I, depleted of histones (compared to genomic DNA), and bound by Hmo1, a yeast protein similar to UBF [90]. In budding yeast, the number of rRNA genes in open chromatin seems not to be a major regulatory determinant for Pol I activity [7]. rDNA copy number can change from cell to cell as a result of unequal crossing over and ranges between 100 and 200 repeats in a population [91]. Fob1 is required for this chromosomal instability [92]. In the absence of Fob1, cell populations with a stable number of rDNA repeats could be generated [7]. With 42 copies, or even more so with only 25 copies, most rRNA genes were active and in an open chromatin state. With an rDNA copy number as low as 25 actively transcribed rRNA genes, growth was not affected, but more polymerases were loaded on each active gene [7].

During the cell cycle, the ratio of open to closed rDNA changes. Newly replicated rDNA becomes psoralen inaccessible and shows nucleosome assembly on both strands after the passage of the replication forks [93]. Following replication, the amount of open chromatin was found to steadily increase at all stages of the cell cycle, including during cell cycle arrest [94]. This increase required Pol I activity. The maintenance of the open chromatin state did not require Pol I activity, but Hmo1 inhibited replication-independent nucleosome assembly [94].

In sharp contrast to ChEC data, chromatin immunoprecipitation (ChIP) analyses targeting H3 and H2B histones in mutant strains harboring most rRNA genes in an open chromatin state (42 or 25 rDNA copies) strongly suggested that significant quantities of histone molecules are present on active rRNA genes [95] (Albert et al. Submitted).

Both ChIP and psoralen-ChEC have some intrinsic limitations. ChIP is known to be sensitive to background binding, which can lead to false-positive detection. Conversely, ChEC is prone to false-negative detection [89]. After MNase

activation, MNase fusion proteins are released from genomic DNA after cleavage of the DNA molecule. These released proteins are then able to cleave genomic DNA at nonspecific sites [89]. ChEC experiments must be performed in a carefully controlled time-course. Due to the abundance of histones molecules, cleavage time is kept bellow 15 minutes [90]. Open rRNA genes appear more resistant to MNase digestion than genes in closed rDNA or naked plasmid DNA, but cleavage is detectable [90]. With such experimental limitations, ChEC combined with psoralen can be used to compare relative levels of histone enrichment, but cannot be used to conclude whether histone is present or absent on open chromatin.

Due to such intrinsic technical limitations, the exact composition of rDNA in the open chromatin state is still widely debated [90, 95]. The analysis of actively transcribed versus untranscribed rDNA can also be performed using Miller spreading (see Figure 1(c)). However, quantifying the ratio of active versus inactive rRNA genes is intrinsically biased, as it underestimates the fraction of inactive rDNA. Transcriptionally inactive rRNA genes are not directly detectable, but can be indirectly visualized because they are flanked by active rRNA genes. This method allows rRNA genes to be characterized at the single-gene level. From such analyses, two important conclusions were reached: nucleosomes are not detectable on actively transcribed genes (data not shown) and nucleosome structures are detectable on some inactive rDNA genes but not all (Figure 3).

The presence of histones on open chromatin, as detected by ChIP, contrasts sharply with the absence or strong depletion of nucleosomes on open chromatin that is observed with psoralen crosslinking, ChEC, and Miller spreading. These observations are not incompatible if one considers that psoralen crosslinking indicates the presence of canonical nucleosomes, whereas ChIP analysis reveals presence of histone molecules. Histones might still be present on open rDNA copies, but a large body of evidence establishes that they are not arranged as canonical nucleosomes impermeable to psoralen. Alternative nucleosome structures have been described and occur specifically when DNA supercoiling is altered [96]. This observation agrees with older biochemical studies demonstrating that despite the absence of detectable beaded nucleosomes on active rRNA genes, the protein constituents of nucleosomes may still be present [8]. Moreover, by combining reagent accessibility analyses and electron microscopy of rDNA from *Physarum polycephalum*, the existence on active rRNA genes of an alternative nucleosome structure called the lexosome was suggested some time ago [97, 98]. The lexosome is an altered nucleosome specifically located on actively transcribed regions, which has properties that facilitate transcription. In a lexosome configuration, the histone-DNA interactions are different than those in intact nucleosomes and allow psoralen to access DNA. Therefore, even if this alternative structure was not confirmed when tested for *in vitro* transcription [99], the lexosome represents an attractive model that is consistent with the ChIP data, the psoralen-crosslinking results, and the electron microscopy images produced in our studies as well as those of other research teams. Similarly, the altered topology of actively transcribed rDNA might lead to other alternative histone configurations [96].

6. Regulator of Pol I Elongation

To date, most factors known to regulate Pol I elongation were characterized previously as Pol II elongation factors. Such dual functions make interpretation of phenotype difficult, since an indirect effect via Pol II is difficult to exclude. However, some Pol I elongation factors are now well characterized.

Spt5 copurifies with Pol I and has been shown to be associated with rDNA *in vivo* [100]. Spt5 is an evolutionarily conserved elongation factor with homologs found in eubacteria (NusG) and in *archaea* (RpoE) [101, 102]. Multisubunit RNA polymerases have the ability to stably interact with DNA through a structural feature called the DNA clamp. Prior to interacting with DNA, the DNA clamp must be in an open configuration and closed for processive elongation. Spt5/NusG can close the DNA clamp, making the polymerase able to carry out processive elongation [103]. Spt5 interacts with multiple Pol I subunits and Rrn3, and an Spt5 mutant (*spt5-C292R*) suppresses the growth defect of the *rpa49*Δ Pol I mutant [104]. Phenotypic analysis of rDNA transcription in the Spt5 mutant suggested that it positively and negatively regulates Pol I functions [105]. Other factors that interact with Spt5, such as Spt4, Paf1 and Spt6, have also been suggested to regulate Pol I. An *spt4* deletion mutant led to decreased rDNA copy numbers, and an rRNA processing defect [100]. The Paf1, complex interacts physically with Spt5 [106] and is involved in stimulating Pol I elongation [107]. Spt6 is a histone chaperone that might also be a good candidate to regulate Pol I *in vivo*. In addition, the FACT complex stimulates Pol I activity in human cells [108]. FACT consists of Spt16 and Pob3, interacts with Spt5 in yeast [106], and is genetically linked to *rpa49*Δ [109]. Another Pol II transcription factor has been proposed to regulate Pol I: the Elongator, a six-subunit complex, conserved between yeast and mammals. In African trypanosomes, mutation or down-regulation of the Elp3b subunit of the complex results in increased synthesis of rRNA by Pol I [110].

The growing list of Pol II elongation factors that also regulate Pol I activity is interesting, but some mechanistic insights are still lacking. The elongation mechanisms of Pol II and Pol I are very different. Pol II carries regulators via interactions with CTD [111]. With few Pol II enzymes acting simultaneously on transcribed genes, the stoichiometric association of elongation factors with Pol II results in a density of about one elongation factor per gene. In contrast, a large number of Pol I enzymes can transcribe a single rRNA gene. Thus, it is difficult to imagine that Pol II elongation factors would be stochiometrically bound to each elongating Pol I. Therefore, it seems unlikely that Pol II and Pol I complexes use similar mechanisms of action or factor recruitment strategies. It remains to be understood how the same elongation factors can act in two very different transcription systems.

One factor can now be defined as a *bona fide* Pol I elongation factor, the nucleolin, called Nsr1 in budding yeast and

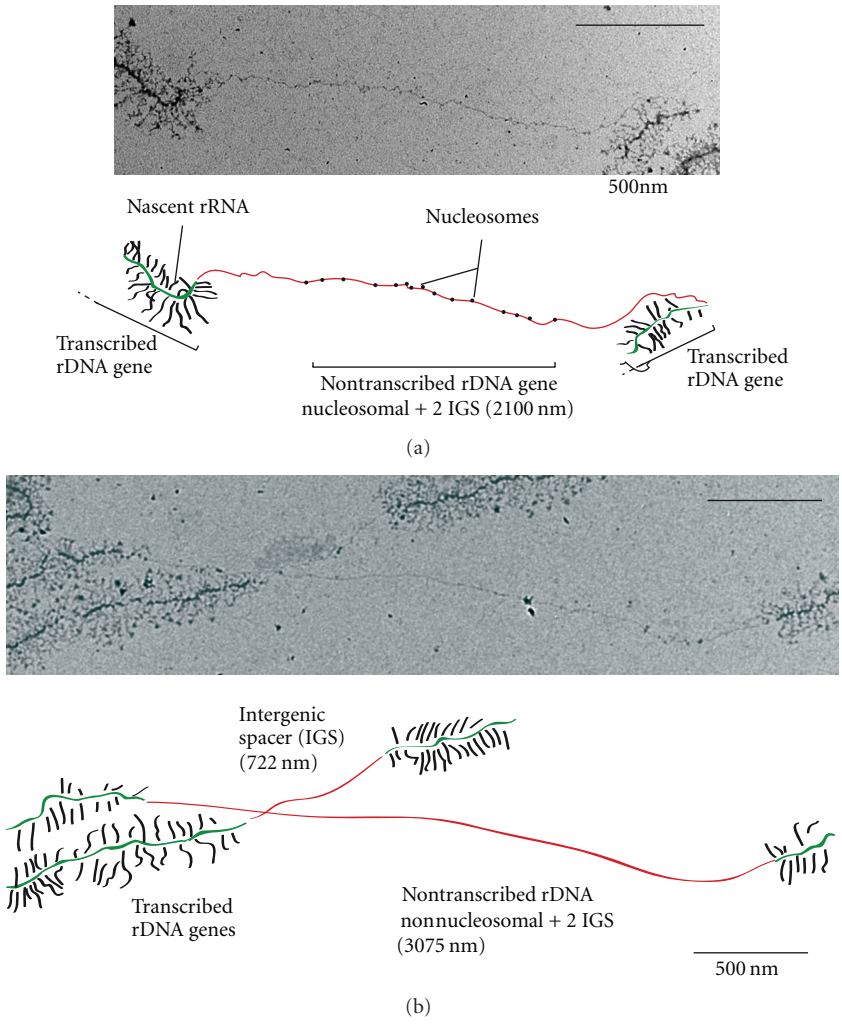


Figure 3: Miller spreading of nontranscribed rDNA. Single-gene analysis of nontranscribed rRNA genes reveals nucleosomal (a) and nonnucleosomal (b) organization. Chromatin spreading (upper panel) and a schematic representation of rDNA spreading (lower panel) are shown from strain NOY1071, bearing 25 rDNA copies. Transcribed rRNA genes (green) are identified by high Pol I density, and nascent rRNAs are shown in black. Non-transcribed regions are depicted in red. Intergenic spaces (IGSs) are short (600 nm) and can be easily distinguished from inactive rRNA genes. Inactive genes are flanked by two IGSs. Due to the DNA wrapping around nucleosomes, nontranscribed genes associated with nucleosomes appear shorter (2,100 nm − (2 × 600 nm) = 900 nm) than genes devoid of nucleosomes (3,100 nm − (2 × 600) = 1,900 nm). With 15 nucleosomes detected and each wrapped around 146 bp, we observed a length reduction of approximately 1,000 nm.

Gar2 in fission yeast. Following nucleolin's first identification as an abundant nucleolar protein, it has been implicated in numerous cellular processes [112, 113]. Among them, it is clearly involved in ribosome biogenesis. Nucleolin is required for early rRNA-processing events [114] and for Pol I activity through a nucleosomal template [115]. Nucleolin has a histone chaperone activity and stimulates transcription by a mechanism reminiscent of the activity of the FACT complex [116]. Nucleolin is clearly an important factor to understand the interplay between rDNA chromatin, Pol I transcription, and cotranscriptional rRNA processing.

7. Concluding Remarks

In this paper, we tried to focus on the unanswered questions of rRNA production, rather than make an exhaustive review of the large body of work addressing regulation of this complex multistep process. We still know little about how cells adjust the production of each ribosomal constituent in time and space to allow cotranscriptional assembly of preribosomal particles. The answer to this question probably lies in the existence of highly redundant pathways, which are all designed to coregulate rRNA production by Pol I, ribosomal protein production, and the availability of the ~200 transacting factors required to assemble eukaryotic ribosomes. The discovery of redundant pathways has clearly resulted from the extensive study of budding yeast. Most regulatory pathways affecting rRNA production are not essential for cell growth. Out of 14 Pol I subunits, four are not required for cell growth. However, when double inactivation is performed, their functions can be revealed and studied [74]. We have no doubt that important progress remains to

be made in understanding how Pol I is regulated. Because most of the complex interplay between rRNA production, assembly, cleavage, and folding occurs during elongation, we expect that most progress remaining to be made will uncover how rRNA elongation is coupled to rRNA assembly [81]. The central question of the exact structure and composition of open rDNA chromatin remains a major challenge. Pol I elongation is likely to be the most important step in controlling how nascent rRNA is folded and cleaved to yield pre-60S and pre-40S rRNA as they are being assembled into large preribosomes. We propose that elongation is the most regulatable step in rRNA production, making elongation the first target to regulate rRNA production.

Acknowledgments

The authors would like to thank all the members of the Gadal Lab for their contributions and critical evaluations of this paper. This work was supported by an ATIP-plus grant from CNRS, by the Agence Nationale de la Recherche (Nucleopol, Ribeuc and Jeune chercheur programme-ODynRib), and by Jeune Équipe from FRM. B. Albert is supported by a PhD fellowship from FRM. They thank Maxime Berthaud for preparing and quantifying Miller spreads obtained from mutant yeast cells. They also acknowledge Nacer Benmeradi and Stéphanie Balor for their help in EM image acquisition. This work also benefited from the assistance of the electron microscopy facility of the IFR 109 and from the imaging platform of Toulouse TRI.

References

[1] B. Albert, I. Leger-Silvestre, C. Normand et al., "RNA polymerase I-specific subunits promote polymerase clustering to enhance the rRNA gene transcription cycle," *The Journal of Cell Biology*, vol. 192, pp. 277–293, 2011.

[2] J. R. Warner, "The economics of ribosome biosynthesis in yeast," *Trends in Biochemical Sciences*, vol. 24, no. 11, pp. 437–440, 1999.

[3] O. L. Miller and B. R. Beatty, "Visualization of nucleolar genes," *Science*, vol. 164, no. 3882, pp. 955–957, 1969.

[4] T. D. Petes, "Yeast ribosomal DNA genes are located on chromosome XII," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 76, no. 1, pp. 410–414, 1979.

[5] E. Schweizer and H. O. Halvorson, "On the regulation of ribosomal RNA synthesis in yeast," *Experimental Cell Research*, vol. 56, no. 2-3, pp. 239–244, 1969.

[6] T. Moss, F. Langlois, T. Gagnon-Kugler, and V. Stefanovsky, "A housekeeper with power of attorney: the rRNA genes in ribosome biogenesis," *Cellular and Molecular Life Sciences*, vol. 64, no. 1, pp. 29–49, 2007.

[7] S. L. French, Y. N. Osheim, F. Cioci, M. Nomura, and A. L. Beyer, "In exponentially growing Saccharomyces cerevisiae cells, rRNA synthesis is determined by the summed RNA polymerase I loading rate rather than by the number of active genes," *Molecular and Cellular Biology*, vol. 23, no. 5, pp. 1558–1568, 2003.

[8] A. Conconi, R. M. Widmer, T. Koller, and J. M. Sogo, "Two different chromatin structures coexist in ribosomal RNA genes throughout the cell cycle," *Cell*, vol. 57, no. 5, pp. 753–761, 1989.

[9] S. P. Bell, R. M. Learned, H. M. Jantzen, and R. Tjian, "Functional cooperativity between transcription factors UBF1 and SL1 mediates human ribosomal RNA synthesis," *Science*, vol. 241, no. 4870, pp. 1192–1197, 1988.

[10] J. C. B. M. Zomerdijk, H. Beckmann, L. Comai, and R. Tjian, "Assembly of transcriptionally active RNA polymerase I initiation factor SL1 from recombinant subunits," *Science*, vol. 266, no. 5193, pp. 2015–2018, 1994.

[11] J. J. Gorski, S. Pathak, K. Panov et al., "A novel TBP-associated factor of SL1 functions in RNA polymerase I transcription," *EMBO Journal*, vol. 26, no. 6, pp. 1560–1568, 2007.

[12] D. Eberhard, L. Tora, J. M. Egly, and I. Grummt, "A TBP-containing multiprotein complex (TIF-IB) mediates transcription specificity of murine RNA polymerase I," *Nucleic Acids Research*, vol. 21, no. 18, pp. 4180–4186, 1993.

[13] C. A. Radebaugh, J. L. Matthews, G. K. Geiss et al., "TATA box-binding protein (TBP) is a constituent of the polymerase I- specific transcription initiation factor TIF-IB (SL1) bound to the rRNA promoter and shows differential sensitivity to TBP-directed reagents in polymerase I, II, and III transcription factors," *Molecular and Cellular Biology*, vol. 14, no. 1, pp. 597–605, 1994.

[14] J. Heix, J. C. B. M. Zomerdijk, A. Ravanpay, R. Tjian, and I. Grummt, "Cloning of murine RNA polymerase I-specific TAF factors: conserved interactions between the subunits of the species-specific transcription initiation factor TIF-IB/SL1," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 5, pp. 1733–1738, 1997.

[15] V. Y. Stefanovsky, G. Pelletier, D. P. Bazett-Jones, C. Crane-Robinson, and T. Moss, "DNA looping in the RNA polymerase I enhancesome is the result of non-cooperative in-phase bending by two UBF molecules," *Nucleic Acids Research*, vol. 29, no. 15, pp. 3241–3247, 2001.

[16] D. P. Bazett-Jones, B. Leblanc, M. Herfort, and T. Moss, "Short-range DNA looping by the Xenopus HMG-box transcription factor, xUBF," *Science*, vol. 264, no. 5162, pp. 1134–1137, 1994.

[17] K. I. Panov, J. K. Friedrich, J. Russell, and J. C. B. M. Zomerdijk, "UBF activates RNA polymerase I transcription by stimulating promoter escape," *EMBO Journal*, vol. 25, no. 14, pp. 3310–3322, 2006.

[18] A. C. O'Sullivan, G. J. Sullivan, and B. McStay, "UBF binding in vivo is not restricted to regulatory sequences within the vertebrate ribosomal DNA repeat," *Molecular and Cellular Biology*, vol. 22, no. 2, pp. 657–668, 2002.

[19] V. Stefanovsky, F. Langlois, T. Gagnon-Kugler, L. I. Rothblum, and T. Moss, "Growth factor signaling regulates elongation of RNA polymerase I transcription in mammals via UBF phosphorylation and r-chromatin remodeling," *Molecular Cell*, vol. 21, no. 5, pp. 629–639, 2006.

[20] I. Grummt, "Wisely chosen paths—regulation of rRNA synthesis: delivered on 30 June 2010 at the 35th FEBS Congress in Gothenburg, Sweden," *FEBS Journal*, vol. 277, no. 22, pp. 4626–4639, 2010.

[21] D. A. Keys, L. Vu, J. S. Steffan et al., "RRN6 and RRN7 encode subunits of a multiprotein complex essential for the initiation of rDNA transcription by RNA polymerase I in Saccharomyces cerevisiae," *Genes and Development*, vol. 8, no. 19, pp. 2349–2362, 1994.

[22] D. Lalo, J. S. Steffan, J. A. Dodd, and M. Nomura, "RRN11 encodes the third subunit of the complex containing Rrn6p and Rrn7p that is essential for the initiation of rDNA transcription by yeast RNA polymerase I," *Journal of Biological Chemistry*, vol. 271, no. 35, pp. 21062–21067, 1996.

[23] C. W. Lin, B. Moorefield, J. Payne, P. Aprikian, K. Mitomo, and R. H. Reeder, "A novel 66-kilodalton protein complexes with Rrn6, Rrn7, and TATA- binding protein to promote polymerase I transcription initiation in Saccharomyces cerevisiae," *Molecular and Cellular Biology*, vol. 16, no. 11, pp. 6436–6443, 1996.

[24] D. A. Keys, B. S. Lee, J. A. Dodd et al., "Multiprotein transcription factor UAF interacts with the upstream element of the yeast RNA polymerase I promoter and forms a stable preinitiation complex," *Genes and Development*, vol. 10, no. 7, pp. 887–903, 1996.

[25] M. Oakes, I. Siddiqi, L. Vu, J. Aris, and M. Nomura, "Transcription factor UAF, expansion and contraction of ribosomal DNA (rDNA) repeats, and RNA polymerase switch in transcription of yeast rDNA," *Molecular and Cellular Biology*, vol. 19, no. 12, pp. 8559–8569, 1999.

[26] L. Vu, I. Siddiqi, B. S. Lee, C. A. Josaitis, and M. Nomura, "RNA polymerase switch in transcription of yeast rDNA: Role of transcription factor UAF (upstream activation factor) in silencing rDNA transcription by RNA polymerase II," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 8, pp. 4390–4395, 1999.

[27] J. S. Steffan, D. A. Keys, J. A. Dodd, and M. Nomura, "The role of TBP in rDNA transcription by RNA polymerase I in Saccharomyces cerevisiae: TBP is required for upstream activation factor- dependent recruitment of core factor," *Genes and Development*, vol. 10, no. 20, pp. 2551–2563, 1996.

[28] O. Gadal, S. Labarre, C. Boschiero, and P. Thuriaux, "Hmo1, an HMG-box protein, belongs to the yeast ribosomal DNA transcription system," *EMBO Journal*, vol. 21, no. 20, pp. 5498–5507, 2002.

[29] A. H. Cavanaugh, A. Evans, and L. I. Rothblum, "Mammalian Rrn3 is required for the formation of a transcription competent preinitiation complex containing RNA polymerase I," *Gene Expression*, vol. 14, no. 3, pp. 131–147, 2007.

[30] R. T. Yamamoto, Y. Nogi, J. A. Dodd, and M. Nomura, "RRN3 gene of Saccharomyces cerevisiae encodes an essential RNA polymerase I transcription factor which interacts with the polymerase independently of DNA template," *EMBO Journal*, vol. 15, no. 15, pp. 3964–3973, 1996.

[31] P. Milkereit and H. Tschochner, "A specialized form of RNA polymerase I, essential for initiation and growth-dependent regulation of rRNA synthesis, is disrupted during transcription," *EMBO Journal*, vol. 17, no. 13, pp. 3692–3703, 1998.

[32] J. Bodem, G. Dobreva, U. Hoffmann-Rohrer et al., "TIF-IA, the factor mediating growth-dependent control of ribosomal RNA synthesis, is the mammalian homolog of yeast Rrn3p," *EMBO Reports*, vol. 1, no. 2, pp. 171–175, 2000.

[33] B. Moorefield, E. A. Greene, and R. H. Reeder, "RNA polymerase I transcription factor Rrn3 is functionally conserved between yeast and human," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 9, pp. 4724–4729, 2000.

[34] G. Peyroche, P. Milkereit, N. Bischler et al., "The recruitment of RNA polymerase I on rDNA is mediated by the interaction of the A43 subunit with Rrn3," *EMBO Journal*, vol. 19, no. 20, pp. 5473–5482, 2000.

[35] G. Miller, K. I. Panov, J. K. Friedrich, L. Trinkle-Mulcahy, A. I. Lamond, and J. C. B. M. Zomerdijk, "hRRN3 is essential in the SL1-mediated recruitment of RNA polymerase I to rRNA gene promoters," *EMBO Journal*, vol. 20, no. 6, pp. 1373–1382, 2001.

[36] I. Hirschler-Laszkiewicz, A. H. Cavanaugh, A. Mirza et al., "Rrn3 becomes inactivated in the process of ribosomal DNA transcription," *Journal of Biological Chemistry*, vol. 278, no. 21, pp. 18953–18959, 2003.

[37] R. Loewith, E. Jacinto, S. Wullschleger et al., "Two TOR complexes, only one of which is rapamycin sensitive, have distinct roles in cell growth control," *Molecular Cell*, vol. 10, no. 3, pp. 457–468, 2002.

[38] J. A. Claypool, S. L. French, K. Johzuka et al., "Tor pathway regulates Rrn3p-dependent recruitment of yeast RNA polymerase I to the promoter but does not participate in alteration of the number of active genes," *Molecular Biology of the Cell*, vol. 15, no. 2, pp. 946–956, 2004.

[39] A. Laferté, E. Favry, A. Sentenac, M. Riva, C. Carles, and S. Chédin, "The transcriptional activity of RNA polymerase I is a key determinant for the level of all ribosome components," *Genes and Development*, vol. 20, no. 15, pp. 2030–2040, 2006.

[40] F. Beckouet, S. Labarre-Mariotte, B. Albert et al., "Two RNA polymerase I subunits control the binding and release of Rrn3 during transcription," *Molecular and Cellular Biology*, vol. 28, no. 5, pp. 1596–1605, 2008.

[41] A. H. Cavanaugh, I. Hirschler-Laszkiewicz, Q. Hu et al., "Rrn3 phosphorylation is a regulatory checkpoint for ribosome biogenesis," *Journal of Biological Chemistry*, vol. 277, no. 30, pp. 27423–27432, 2002.

[42] J. Gerber, A. Reiter, R. Steinbauer et al., "Site specific phosphorylation of yeast RNA polymerase I," *Nucleic Acids Research*, vol. 36, no. 3, pp. 793–802, 2008.

[43] S. Rogers, R. Wells, and M. Rechsteiner, "Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis," *Science*, vol. 234, no. 4774, pp. 364–368, 1986.

[44] A. Philippi, R. Steinbauer, A. Reiter et al., "TOR-dependent reduction in the expression level of Rrn3p lowers the activity of the yeast RNA Pol I machinery, but does not account for the strong inhibition of rRNA production," *Nucleic Acids Research*, vol. 38, no. 16, pp. 5315–5326, 2010.

[45] A. Reiter, R. Steinbauer, A. Philippi et al., "Reduction in ribosomal protein synthesis is sufficient to explain major effects on ribosome production after short-term TOR inactivation in Saccharomyces cerevisiae," *Molecular and Cellular Biology*, vol. 31, pp. 803–817, 2010.

[46] A. Huber, S. L. French, H. Tekotte et al., "Sch9 regulates ribosome biogenesis via Stb3, Dot6 and Tod6 and the histone deacetylase complex RPD3L," *The EMBO Journal*, vol. 30, no. 15, pp. 3052–3064, 2011.

[47] A. Huber, B. Bodenmiller, A. Uotila et al., "Characterization of the rapamycin-sensitive phosphoproteome reveals that Sch9 is a central coordinator of protein synthesis," *Genes and Development*, vol. 23, no. 16, pp. 1929–1943, 2009.

[48] C. Cadwell, H. J. Yoon, Y. Zebarjadian, and J. Carbon, "The yeast nucleolar protein Cbf5p is involved in rRNA biosynthesis and interacts genetically with the RNA polymerase I transcription factor RRN3," *Molecular and Cellular Biology*, vol. 17, no. 10, pp. 6175–6183, 1997.

[49] D. A. Rojas, S. Moreira-Ramos, S. Zock-Emmenthal et al., "Rrn7 protein, an RNA polymerase I transcription factor, is required for RNA polymerase II-dependent transcription directed by core promoters with a HomolD Box Sequence," *The Journal of Biological Chemistry*, vol. 286, pp. 26480–26486, 2011.

[50] A. B. Berger, L. Decourty, G. Badis, U. Nehrbass, A. Jacquier, and O. Gadal, "Hmo1 is required for TOR-dependent regulation of ribosomal protein gene transcription," *Molecular and Cellular Biology*, vol. 27, no. 22, pp. 8015–8026, 2007.

[51] N. J. Proudfoot, A. Furger, and M. J. Dye, "Integrating mRNA processing with transcription," *Cell*, vol. 108, no. 4, pp. 501–512, 2002.

[52] S. Chanarat, M. Seizl, K. Strasser et al., "The Prp19 complex is a novel transcription elongation factor required for TREX occupancy at transcribed genes," *Genes & Development*, vol. 25, pp. 1147–1158, 2011.

[53] S. Rodriguez-Navarro and E. Hurt, "Linking gene regulation to mRNA production and export," *Current Opinion in Cell Biology*, vol. 23, pp. 302–309, 2011.

[54] 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.

[55] F. Dragon, J. E. G. Gallagher, P. A. Compagnone-Post et al., "A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis," *Nature*, vol. 417, no. 6892, pp. 967–970, 2002.

[56] J. E. G. Gallagher, D. A. Dunbar, S. Granneman et al., "RNA polymerase I transcription and pre-rRNA processing are linked by specific SSU processome components," *Genes and Development*, vol. 18, no. 20, pp. 2507–2517, 2004.

[57] N. J. Krogan, W. T. Peng, G. Cagney et al., "High-definition macromolecular composition of yeast RNA-processing complexes," *Molecular Cell*, vol. 13, no. 2, pp. 225–239, 2004.

[58] M. Wery, S. Ruidant, S. Schillewaert, N. Leporé, and D. L. J. Lafontaine, "The nuclear poly(A) polymerase and Exosome cofactor Trf5 is recruited cotranscriptionally to nucleolar surveillance," *RNA*, vol. 15, no. 3, pp. 406–419, 2009.

[59] J. L. Prieto and B. McStay, "Recruitment of factors linking transcription and processing of pre-rRNA to NOR chromatin is UBF-dependent and occurs independent of transcription in human cells," *Genes and Development*, vol. 21, no. 16, pp. 2041–2054, 2007.

[60] S. Hofmann and O. L. Miller Jr., "Visualization of ribosomal ribonucleic acid synthesis in a ribonuclease III deficient strain of Escherichia coli," *Journal of Bacteriology*, vol. 132, no. 2, pp. 718–722, 1977.

[61] R. M. Grainger and N. Maizels, "Dictyostelium ribosomal RNA is processed during transcription," *Cell*, vol. 20, no. 3, pp. 619–623, 1980.

[62] Y. N. Osheim, S. L. French, K. M. Keck et al., "Pre-18S ribosomal RNA is structurally compacted into the SSU processome prior to being cleaved from nascent transcripts in Saccharomyces cerevisiae," *Molecular Cell*, vol. 16, no. 6, pp. 943–954, 2004.

[63] M. Koš and D. Tollervey, "Yeast Pre-rRNA Processing and Modification Occur Cotranscriptionally," *Molecular Cell*, vol. 37, no. 6, pp. 809–820, 2010.

[64] J. Pérez-Fernández, A. Román, J. De Las Rivas, X. R. Bustelo, and M. Dosil, "The 90S preribosome is a multimodular structure that is assembled through a hierarchical mechanism," *Molecular and Cellular Biology*, vol. 27, no. 15, pp. 5414–5429, 2007.

[65] A. Segerstolpe, P. Lundkvist, Y. N. Osheim, A. L. Beyer, and L. Wieslander, "Mrd1p binds to pre-rRNA early during transcription independent of U3 snoRNA and is required for compaction of the pre-rRNA into small subunit processomes," *Nucleic Acids Research*, vol. 36, no. 13, pp. 4364–4380, 2008.

[66] Y. Nogi, R. Yano, and M. Nomura, "Synthesis of large rRNAs by RNA polymerase II in mutants of Saccharomyces cerevisiae defective in RNA polymerase I," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 88, no. 9, pp. 3962–3966, 1991.

[67] M. Oakes, Y. Nogi, M. W. Clark, and M. Nomura, "Structural alterations of the nucleolus in mutants of Saccharomyces cerevisiae defective in RNA polymerase I," *Molecular and Cellular Biology*, vol. 13, no. 4, pp. 2441–2455, 1993.

[68] S. Trumtel, I. Léger-Silvestre, P. E. Gleizes, F. Teulières, and N. Gas, "Assembly and functional organization of the nucleolus: ultrastructural analysis of Saccharomyces cerevisiae mutants," *Molecular Biology of the Cell*, vol. 11, no. 6, pp. 2175–2189, 2000.

[69] S. Fath, P. Milkereit, A. V. Podtelejnikov et al., "Association of yeast RNA polymerase I with a nucleolar substructure active in rRNA synthesis and processing," *Journal of Cell Biology*, vol. 149, no. 3, pp. 575–590, 2000.

[70] D. A. Schneider, A. Michel, M. L. Sikes et al., "Transcription elongation by RNA polymerase I is linked to efficient rRNA processing and ribosome assembly," *Molecular Cell*, vol. 26, no. 2, pp. 217–229, 2007.

[71] C. D. Kuhn, S. R. Geiger, S. Baumli et al., "Functional architecture of RNA polymerase I," *Cell*, vol. 131, no. 7, pp. 1260–1272, 2007.

[72] K. I. Panov, T. B. Panova, O. Gadal et al., "RNA polymerase I-specific subunit CAST/hPAF49 has a role in the activation of transcription by upstream binding factor," *Molecular and Cellular Biology*, vol. 26, no. 14, pp. 5436–5448, 2006.

[73] L. F. Liu and J. C. Wang, "Supercoiling of the DNA template during transcription," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 84, no. 20, pp. 7024–7027, 1987.

[74] O. Gadal, S. Mariotte-Labarre, S. Chedin et al., "A34.5, a nonessential component of yeast RNA polymerase I, cooperates with subunit A14 and DNA topoisomerase I to produce a functional rRNA synthesis machine," *Molecular and Cellular Biology*, vol. 17, no. 4, pp. 1787–1795, 1997.

[75] S. Gangloff, J. P. McDonald, C. Bendixen, L. Arthur, and R. Rothstein, "The yeast type I topoisomerase top3 interacts with Sgs1, a DNA helicase homolog: a potential eukaryotic reverse gyrase," *Molecular and Cellular Biology*, vol. 14, no. 12, pp. 8391–8398, 1994.

[76] M. F. Christman, F. S. Dietrich, N. A. Levin, B. U. Sadoff, and G. R. Fink, "The rRNA-encoding DNA array has an altered structure in topoisomerase I mutants of Saccharomyces cerevisiae," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 90, no. 16, pp. 7637–7641, 1993.

[77] R. P. Perry and D. E. Kelley, "Inhibition of RNA synthesis by actinomycin D: characteristic dose-response of different RNA species," *Journal of Cellular Physiology*, vol. 76, no. 2, pp. 127–139, 1970.

[78] B. Sollner-Webb and J. Tower, "Transcription of cloned eukaryotic ribosomal RNA genes," *Annual Review of Biochemistry*, vol. 55, pp. 801–830, 1986.

[79] K. V. Hadjiolova, A. A. Hadjiolov, and J. P. Bachellerie, "Actinomycin D stimulates the transcription of rRNA minigenes transfected into mouse cells. Implications for the in vivo hypersensitivity of rRNA gene transcription," *European Journal of Biochemistry*, vol. 228, no. 3, pp. 605–615, 1995.

[80] D. K. Trask and M. T. Muller, "Stabilization of type I topoisomerase-DNA covalent complexes by actinomycin D," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 85, no. 5, pp. 1417–1421, 1988.

[81] A. El Hage, S. L. French, A. L. Beyer, and D. Tollervey, "Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis," *Genes and Development*, vol. 24, no. 14, pp. 1546–1558, 2010.

[82] S. L. French, M. L. Sikes, R. D. Hontz et al., "Distinguishing the roles of Topoisomerases I and II in relief of transcription-induced torsional stress in yeast rRNA genes," *Molecular and Cellular Biology*, vol. 31, pp. 482–494, 2010.

[83] S. J. Brill, S. DiNardo, K. Voelkel-Meiman, and R. Sternglanz, "Need for DNA topoisomerase activity as a swivel for DNA replication for transcription of ribosomal RNA," *Nature*, vol. 326, no. 6111, pp. 414–416, 1987.

[84] C. Lavelle, "DNA torsional stress propagates through chromatin fiber and participates in transcriptional regulation," *Nature Structural and Molecular Biology*, vol. 15, no. 2, pp. 123–125, 2008.

[85] D. A. Koster, V. Croquette, C. Dekker, S. Shuman, and N. H. Dekker, "Friction and torque govern the relaxation of DNA supercoils by eukaryotic topoisomerase IB," *Nature*, vol. 434, no. 7033, pp. 671–674, 2005.

[86] J. Salceda, X. Fernández, and J. Roca, "Topoisomerase II, not topoisomerase I, is the proficient relaxase of nucleosomal DNA," *EMBO Journal*, vol. 25, no. 11, pp. 2575–2583, 2006.

[87] B. U. Sadoff, S. Heath-Pagliuso, I. B. Castano, Y. Zhu, F. S. Kieff, and M. F. Christman, "Isolation of mutants of Saccharomyces cerevisiae requiring DNA topoisomerase I," *Genetics*, vol. 141, no. 2, pp. 465–479, 1995.

[88] R. Dammann, R. Lucchini, T. Koller, and J. M. Sogo, "Chromatin structures and transcription of rDNA in yeast Saccharomyces cerevisiae," *Nucleic Acids Research*, vol. 21, no. 10, pp. 2331–2338, 1993.

[89] M. Schmid, T. Durussel, and U. K. Laemmli, "ChIC and ChEC: genomic mapping of chromatin proteins," *Molecular Cell*, vol. 16, no. 1, pp. 147–157, 2004.

[90] K. Merz, M. Hondele, H. Goetze, K. Gmelch, U. Stoeckl, and J. Griesenbeck, "Actively transcribed rRNA genes in S. cerevisiae are organized in a specialized chromatin associated with the high-mobility group protein Hmo1 and are largely devoid of histone molecules," *Genes and Development*, vol. 22, no. 9, pp. 1190–1204, 2008.

[91] A. Chindamporn, S. I. Iwaguch, Y. Nakagawa, M. Homma, and K. Tanaka, "Clonal size-variation of rDNA cluster region on chromosome XII of Saccharomyces cerevisiae," *Journal of General Microbiology*, vol. 139, no. 7, pp. 1409–1415, 1993.

[92] T. Kobayashi, D. J. Heck, M. Nomura, and T. Horiuchi, "Expansion and contraction of ribosomal DNA repeats in Saccharomyces cerevisiae: requirement of replication fork blocking (Fob1) protein and the role of RNA polymerase I," *Genes and Development*, vol. 12, no. 24, pp. 3821–3830, 1998.

[93] R. Lucchini and J. M. Sogo, "Replication of transcriptionally active chromatin," *Nature*, vol. 374, no. 6519, pp. 276–280, 1995.

[94] M. Wittner, S. Hamperl, U. Stockl et al., "Establishment and maintenance of alternative chromatin States at a multicopy gene locus," *Cell*, vol. 145, pp. 543–554, 2011.

[95] H. S. Jones, J. Kawauchi, P. Braglia, C. M. Alen, N. A. Kent, and N. J. Proudfoot, "RNA polymerase I in yeast transcribes dynamic nucleosomal rDNA," *Nature Structural and Molecular Biology*, vol. 14, no. 2, pp. 123–130, 2007.

[96] C. Lavelle and A. Prunell, "Chromatin polymorphism and the nucleosome superfamily: a genealogy," *Cell Cycle*, vol. 6, no. 17, pp. 2113–2119, 2007.

[97] C. P. Prior, C. R. Cantor, E. M. Johnson, and V. G. Allfrey, "Incorporation of exogenous pyrene-labeled histone into Physarum chromatin: a system for studying changes in nucleosomes assembled in vivo," *Cell*, vol. 20, no. 3, pp. 597–608, 1980.

[98] H. S. Judelson and V. M. Vogt, "Accessibility of ribosomal genes to trimethyl psoralen in nuclei of Physarum polycephalum," *Molecular and Cellular Biology*, vol. 2, no. 3, pp. 211–220, 1982.

[99] R. U. Protacio and J. Widom, "Nucleosome transcription studied in a real-time synchronous system: test of the lexosome model and direct measurement of effects due to histone octamer," *Journal of Molecular Biology*, vol. 256, no. 3, pp. 458–472, 1996.

[100] D. A. Schneider, S. L. French, Y. N. Osheim et al., "RNA polymerase II elongation factors Spt4p and Spt5p play roles in transcription elongation by RNA polymerase I and rRNA processing," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 34, pp. 12707–12712, 2006.

[101] F. W. Martinez-Rucobo, S. Sainsbury, A. C. Cheung, and P. Cramer, "Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity," *The EMBO Journal*, vol. 30, pp. 1302–1310, 2011.

[102] B. J. Klein, D. Bose, K. J. Baker et al., "RNA polymerase and transcription elongation factor Spt4/5 complex structure," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, pp. 546–550, 2011.

[103] G. A. Hartzog, T. Wada, H. Handa, and F. Winston, "Evidence that Spt4, Spt5, and Spt6 control transcription elongation by RNA polymerase II in Saccharomyces cerevisiae," *Genes and Development*, vol. 12, no. 3, pp. 357–369, 1998.

[104] O. V. Viktorovskaya, F. D. Appling, and D. A. Schneider, "Yeast transcription elongation factor Spt5 associates with RNA polymerase I and RNA polymerase II directly," *The Journal of Biological Chemistry*, vol. 286, pp. 18825–18833, 2011.

[105] S. J. Anderson, M. L. Sikes, Y. Zhang et al., "The transcription elongation factor spt5 influences transcription by RNA polymerase I positively and negatively," *The Journal of Biological Chemistry*, vol. 286, pp. 18816–18824, 2011.

[106] S. L. Squazzo, P. J. Costa, D. L. Lindstrom et al., "The Paf1 complex physically and functionally associates with transcription elongation factors in vivo," *EMBO Journal*, vol. 21, no. 7, pp. 1764–1774, 2002.

[107] Y. Zhang, M. L. Sikes, A. L. Beyer, and D. A. Schneider, "The Paf1 complex is required for efficient transcription elongation by RNA polymerase I," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 7, pp. 2153–2158, 2009.

[108] J. L. Birch, B. C. M. Tan, K. I. Panov et al., "FACT facilitates chromatin transcription by RNA polymerases i and III," *EMBO Journal*, vol. 28, no. 7, pp. 854–865, 2009.

[109] A. P. Davierwala, J. Haynes, Z. Li et al., "The synthetic genetic interaction spectrum of essential genes," *Nature Genetics*, vol. 37, no. 10, pp. 1147–1152, 2005.

[110] S. Alsford and D. Horn, "Elongator protein 3b negatively regulates ribosomal DNA transcription in african trypanosomes," *Molecular and Cellular Biology*, vol. 31, pp. 1822–1832, 2011.

[111] 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.

[112] L. R. Orrick, M. O. J. Olson, and H. Busch, "Comparison of nucleolar proteins of normal rat liver and Novikoff hepatoma ascites cells by two dimensional polyacrylamide gel electrophoresis," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 70, no. 5, pp. 1316–1320, 1973.

[113] F. Mongelard and P. Bouvet, "Nucleolin: a multiFACeTed protein," *Trends in Cell Biology*, vol. 17, no. 2, pp. 80–86, 2007.

[114] H. Ginisty, F. Amalric, and P. Bouvet, "Nucleolin functions in the first step of ribosomal RNA processing," *EMBO Journal*, vol. 17, no. 5, pp. 1476–1486, 1998.

[115] B. Rickards, S. J. Flint, M. D. Cole, and G. LeRoy, "Nucleolin is required for RNA polymerase I transcription in vivo," *Molecular and Cellular Biology*, vol. 27, no. 3, pp. 937–948, 2007.

[116] D. Angelov, V. A. Bondarenko, S. Almagro et al., "Nucleolin is a histone chaperone with FACT-like activity and assists remodeling of nucleosomes," *EMBO Journal*, vol. 25, no. 8, pp. 1669–1679, 2006.

SAGE-Hindawi Access to Research
Genetics Research International
Volume 2011, Article ID 626081, 10 pages
doi:10.4061/2011/626081

Review Article

Spatial Organization and Dynamics of Transcription Elongation and Pre-mRNA Processing in Live Cells

Miguel Sánchez-Álvarez,[1] Noemí Sánchez-Hernández,[2] and Carlos Suñé[2]

[1] *Dynamical Cell Systems Team, Section of Cellular and Molecular Biology, The Institute of Cancer Research, London SW3 6JB, UK*
[2] *Department of Molecular Biology, Instituto de Parasitología y Biomedicina "López Neyra" (IPBLN-CSIC), 18100 Armilla, Spain*

Correspondence should be addressed to Carlos Suñé, csune@ipb.csic.es

Received 22 July 2011; Accepted 5 September 2011

Academic Editor: Sebastián Chávez

During the last 30 years, systematic biochemical and functional studies have significantly expanded our knowledge of the transcriptional molecular components and the pre-mRNA processing machinery of the cell. However, our current understanding of how these functions take place spatiotemporally within the highly compartmentalized eukaryotic nucleus remains limited. Moreover, it is increasingly clear that "the whole is more than the sum of its parts" and that an understanding of the dynamic coregulation of genes is essential for fully characterizing complex biological phenomena and underlying diseases. Recent technological advances in light microscopy in addition to novel cell and molecular biology approaches have led to the development of new tools, which are being used to address these questions and may contribute to achieving an integrated and global understanding of how the genome works at a cellular level. Here, we review major hallmarks and novel insights in RNA polymerase II activity and pre-mRNA processing in the context of nuclear organization, as well as new concepts and challenges arising from our ability to gather extensive dynamic information at the single-cell resolution.

1. Introduction

In eukaryotic cells, the regulation, expression, and subsequent processing steps of genomic sequences tend to be localized to defined spaces within the nucleus [1]. In the interphase nucleus, uncondensed chromosomes do not expand randomly but occupy defined volumes termed "chromosome territories," whose relative positioning has recently been suggested to be determined by, or at least correlated with, differentiation stages and specific contexts [2–4]. This architecture facilitates the intermingling of specific subsets and combinations of genes that need to be coregulated in a given situation [4–7]. Indeed, active genes are most often positioned in the periphery of chromosome territories, while inactive genes remain located within more inaccessible areas of these regions. Although the molecular basis for this dynamic behavior of chromatin is not yet well understood, there are a significant number of studies supporting this concept, thus suggesting a novel layer of complexity in the regulation of gene expression. Genes are not inert entities waiting for the adequate subset of transcription factors to initiate the assembly of a processive RNA polymerase II (RNAPII) complex; instead, the dynamic positioning contributes to their activation state and correlation with other gene units and regulatory elements, such as enhancers and insulators [8–11].

Chromosome territories delimitate a region of the nucleus (usually termed the interchromatin space) that is relatively empty of dense chromatin and is hypothesized to be highly interconnected across the nucleus with a higher-order organization [12–15]. The delimited interchromatin volume contains not only most of the transcriptional activity at its boundaries but also contains several nonmembrane-bound dynamic structures—nuclear bodies—highly enriched with specific subsets of nuclear factors [16–18]. These nuclear bodies include Sam68 bodies, PML bodies, paraspeckles, Cajal bodies, and nuclear speckles.

2. The Transcription Factory: A Spatial and Functional Unit for RNAPII Transcription

In the early 1990s, two groups reported the use of novel techniques that allowed the visualization of transcriptionally

active sites within the nucleus through the incorporation of bromo-UTP in nascent transcripts [19, 20]. In these experiments, nascent transcripts remained immobilized at the site of their chromatin template, and they were studied in great detail using confocal and electron microscopy. Notably, the number of observed active sites appeared to be considerably lower than the estimated number of active molecules of RNAPII [21–23]. These discrete structures colocalize with hyperphosphorylated forms of RNAPII and are resistant to DNA digestion and extraction of soluble fractions [20, 21, 24–26]. These results suggest the existence of an immobile pool of hyperphosphorylated RNAPII within the eukaryotic nucleus. Subsequent fluorescence recovery after photobleaching- (FRAP-) based experiments performed on cells expressing a GFP-tagged construct of RNAPII support this interpretation [27]. Given that the number of observed transcription sites was significantly lower than the number of elongating RNAPII molecules as assessed by *in vitro* run on assays [22, 23], a model was proposed in which several active (mostly elongation-competent) RNAPII units assemble into higher-order structures termed "transcription factories" [21]. According to this model, chromatin loops are tethered to the factories through RNAPII or/and transcription factors for transcription to occur (recently reviewed in (Cook, [28])).

This model is consistent with the looped conformation model that several other independent approaches have suggested exists for an active eukaryotic gene [29–31]. In this model, upon recognition and activation by specific factors, the promoter sequence of the gene unit is tethered to the RNAPII subunit of the factories, and this attachment would be maintained during the transcription of the whole gene sequence, which is "reeled" on the RNAPII. This arrangement provides an additional layer of control and coordination over the different stages of transcription and positions the RNAPII units for subsequent rounds of transcription.

The existence of factories provides us with an elegant conceptual framework to explain the coregulation of functionally related groups of genes in specific contexts [30]. It has been observed that some of these active genes tend to be found in close proximity at a much higher frequency than would be expected by chance [32, 33]. Indeed, these genes tend to share a factory when they are positioned in close proximity, as observed by immunolabeling elongating RNAPII [33]. Although the structural resolution of transcribed genes is still technically limited, nevertheless, this reflects the potential crosstalk that can exist between the transcription factors recruited to each coregulated promoter. Some of the examples consistent with this model are the NF-κB/TNFα activation axis [34, 35] and the ERα module system [34]. Moreover, genes of different sizes and elongation timing coimmunoprecipitate with the elongating form of the RNAPII in a fashion consistent with the model in which they share the same factory and slide along the "polymerase reading heads" in sequential rounds of transcription, rather than just recruiting mobile polymerase complexes from proximal storage sites and undergoing independent read-throughs [35]. Transcription factories are also consistent with data suggesting that genes with shared features, such as promoter composition and the presence or absence of introns, tend to associate among each other [36]. Finally, transcription factories also provide an explanation for observations that indicate that promoter composition and associated events can influence subsequent stages of transcription elongation [37, 38].

Recent studies have reported on the stability of RNAPII foci upon disrupting transcription [39]. Interestingly, treatment of cells with 5,6-dichloro-1-β-d-ribofuranosylbenzimidazole (DRB; a highly specific inhibitor of the positive elongation transcription factor, P-TEFb, and thus an inhibitor of elongating polymerases) does not abolish the association of previously engaged genes with the RNAPII foci, at least for erythroid lineage-specific genes [39]. Observations in agreement with this model include independent genome-wide chromatin immunoprecipitation- (ChIP-) based studies that demonstrate that a significant number of genes is "primed" for transcription. These inactive genes have paused RNAPII complexes at their promoter regions and, upon gene activation, are released from the paused state, allowing elongation to proceed [40].

Initial studies using *in situ* spectroscopy have recently been carried out to determine the composition of the transcription factories [41]. In these studies, the authors demonstrated the existence of clearly defined ribonucleoprotein structures that coincide with sites of active transcription (the perichromatin fibrils), as assessed by BrU pulse incorporation and immunogold labeling. The size and estimated composition of carbon and nitrogen in these structures support the existence of the proposed model of assembled transcription factories, creating a more refined structural model in which the effector subunits of the RNAPII face outwards [41, 42].

Another important feature of transcription factories is the enhancement of the physical and functional coupling of transcription and downstream RNA processing steps. This is facilitated by the regulated recruitment of neighboring machinery for cotranscriptional mRNA maturation in an appropriate fashion and timing. This notion would expand the category of these structures to integrated "mRNA factories," providing an intuitive physical framework for the numerous observed interactions among transcriptional and mRNA processing factors [43–46]. Other essential processes in the regulation of pre-mRNA synthesis, such as chromatin remodeling and histone modification, would similarly benefit from such a design [30].

Although this model of transcription factories (Figure 1) explains many observations of gene expression and nuclear organization, there are many intriguing questions that need to be resolved. What are the molecular mechanisms underlying the appropriate targeting of activated gene sequences to a factory, and how are they integrated in a given regulatory context? What is the inner structure of factories in the cell at a resolution beyond the conventional light diffraction limits? How are these structures assembled and organized during the cell cycle according to the requirements of the cell? Does it require the existence of an underlying structural scaffold or "nucleoskeleton?" How are different regulatory hallmarks, such as the phosphorylation cycle proposed for the CTD of the RNAPII during its progression through the transcription

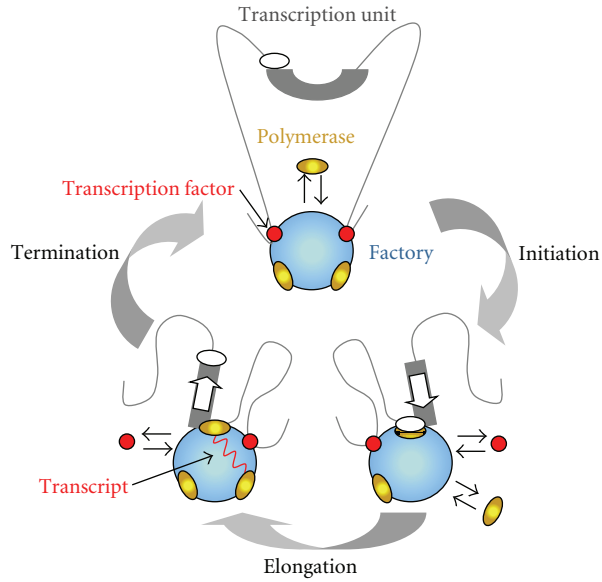


Figure 1: Model of the basic structure and function of RNAPII transcription factories. (a) A "gene-loop" is recruited upon activation to the transcription factory, which contains immobilized subunits of RNAPII. (b) The gene-loop is then "reeled" onto RNAPII for transcriptional elongation. (c) Upon termination, the anchoring of the locus allows for subsequent rounds of transcription. Adapted, with permission, from The Journal of Cell Science [47].

of a gene unit, integrated into the context of these structures? Finally, what is the functional relationship between transcription factories and other nuclear compartments related to the biogenesis of the mRNA?

3. Nuclear Speckles and the Regulation of Transcription and Pre-mRNA Processing

Many independent studies in the last two decades have led to a model in which the maturation of nascent transcripts take place simultaneously to their synthesis, that is, cotranscriptionally [48, 49]. This may be specific to a subset of genes or even to specific introns of a gene and is therefore considered not to be strictly required for the completion of pre-mRNA processing itself [50]. However, cotranscriptional processing allows for the functional coupling of the different steps of RNA biogenesis. The bidirectional interdependence among chromatin conformation and posttranslational modifications, in both transcription and different steps of pre-mRNA processing, constitutes an additional layer in gene expression regulation [51–56]. Moreover, it may play a pivotal role in complex processes, such as neuronal differentiation and activity, global integration of RNA processing signatures and DNA damage, and developmental programs [57–60].

If pre-mRNA processing is performed largely in a regulated cotranscriptional fashion, the dynamic distribution of pre-mRNA processing factors should be correlated with the organization of transcriptionally active sites in the nucleus. The distribution of pre-mRNA processing factors in the eukaryotic nucleus, as observed using immunofluorescence staining and light microscopy, is not homogeneous and shows a dynamic pattern of localized accumulation in 10–30 irregular domains termed speckles, "SC35 domains" or "splicing factor compartments" (SFCs). At the level of electron microscopy, they correspond to two distinct structures: (i) interchromatin granules clusters (IGCs), composed of particles measuring 20–25 nm in diameter and (ii) perichromatin fibrils, 3–5 nm fibrillar structures localized both at the periphery of IGCs and in other nucleoplasmic regions, which are the sites of nascent pre-mRNAs (for extensive reviews, see [61, 62]) These structures were first identified using immunostaining with specific antibodies against different small nuclear ribonucleoproteins (snRNPs) [63, 64]. This and other observations that show the presence of poly(A) RNA colocalizing with snRNPs and SC35-rich domains [65, 66] further support a link between nuclear speckles and pre-mRNA metabolism. Pioneering mass spectrometry studies [67, 68] and a still-growing list of publications using immunofluorescence or live-cell imaging labeled with tagged constructs corroborate a marked enrichment of these compartments with factors involved in pre-mRNA transcription and processing, especially pre-mRNA splicing.

Several models, which are not mutually exclusive, have been proposed to explain the role of these nuclear bodies in the regulation of gene expression: (1) they function as storage/assembly/modification compartments that can supply processing factors to the surrounding active transcription sites; (2) they function as sequestering sites for the dynamic control of transcription and processing factors; (3) they serve as functional "hubs" for coregulated genes and their products; (4) they play an active role in posttranscriptional pre-mRNA processing and surveillance and/or in the coupling of early steps of mRNA biogenesis (Figure 2).

The concept that nuclear speckles are transcriptionally inactive compartments that serve as storage or recycling sites of pre-mRNA metabolism complexes from which these complexes are recruited to nearby sites of active transcription according to demand is a widely held view supported by many experimental results [69–77]. Importantly, recruitment to active sites of transcription requires the integrity of the carboxyl terminal domain of the RNAPII [78], which indicates that transcriptional elongation plays a critical role in the recruitment of pre-mRNA processing factors.

This proposal is compatible with the view that speckles act as inhibitory sites where specific factors are actively sequestered when their functional repression is required. The essential splicing factor SRSF1 is sequestered into these regions upon the induction of stress through a mechanism dependent on the dynamic interaction of SRPK with stress chaperone complexes, including Hsp70 and Hsp90 [79]. Similarly, the *MALAT1* large noncoding (nc) RNA has been proposed to regulate the phosphorylation-dependent dynamics of splicing factors and their equilibrium between nucleoplasm availability and nuclear speckle sequestration [80]. Linking transcriptional elongation control to this model, the 7SK small ncRNA, which is a scaffold component of transcription elongation CDK9-CycT1 inactive complexes together with HEXIM proteins (see below), has been proposed to function, at least partially, by sequestering

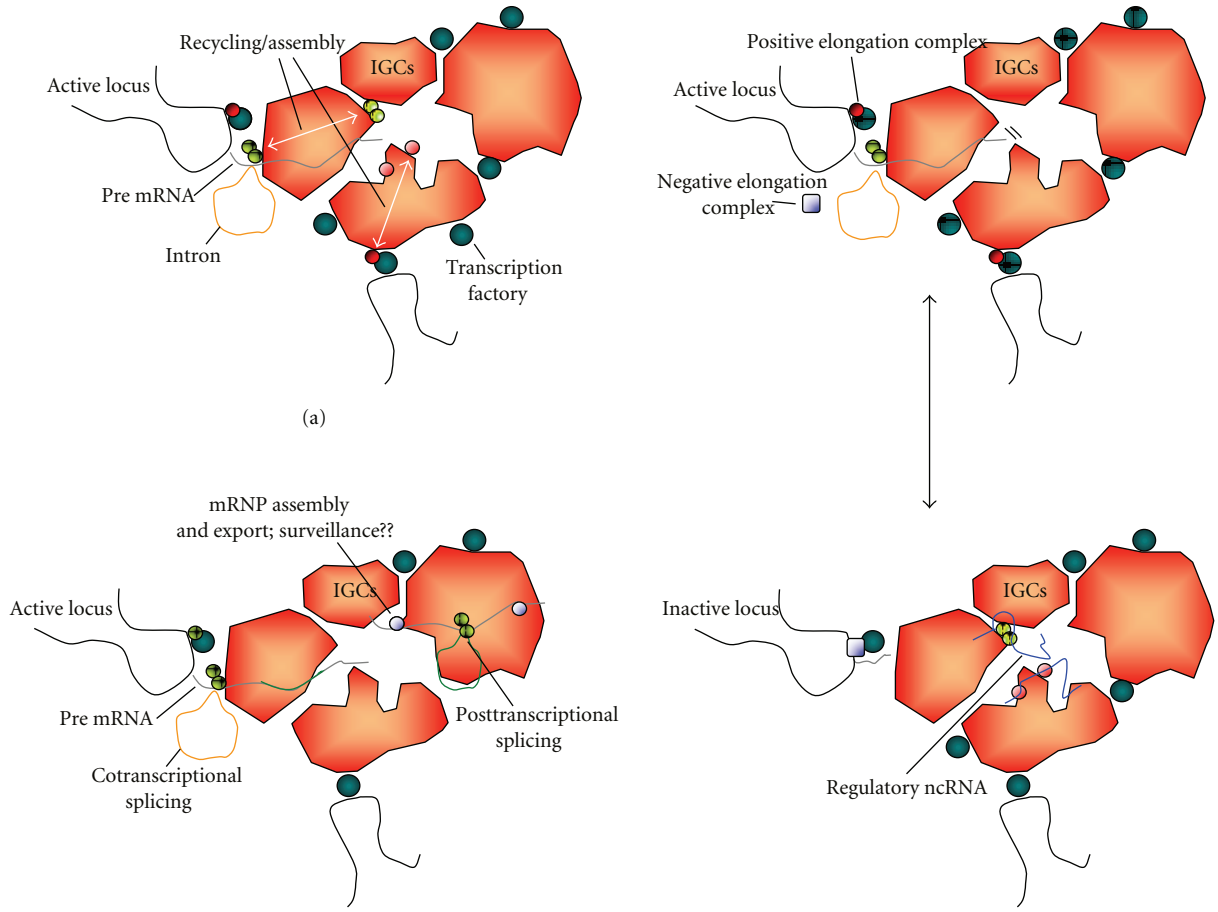


Figure 2: Three potential, nonexclusive models for the role of the interchromatin granule clusters (IGCs) in the regulation of transcription and pre-mRNA processing machinery. IGCs are depicted surrounded by transcriptionally competent sites or "factories" (grey beads). (a) IGCs may be specialized sites for the recycling and assembly of transcription (dark and light red beads) and pre-mRNA processing complexes (dark and light green beads) through regulated posttranslational modification cycles. Dark and light hue code denotes active and inactive pools of factors, respectively. (b) Posttranscriptional processing steps and potential surveillance of mRNA quality may be integrated in these structures, constituting a "checkpoint link" between mRNA transcription and mRNP assembly and export. A given transcript may include both introns that are spliced cotranscriptionally outside of the IGCs (orange lariat) and intron sequences that are processed posttranscriptionally (dark brown stretch and lariat). The later event may be also coupled in the IGCs to surveillance mechanisms, mRNP assembly (blue beads), and export. (c) Specific subsets of nuclear factors, such as ncRNAs (MALAT1, 7SK; see main text; depicted in the lower panel as thin blue threads), can function as active quenchers or sequesters of transcription and pre-mRNA processing factors (red and green beads, resp.), blocking the recruitment of these complexes from the IGCs to nearby active sites of transcription.

these inactive P-TEFb complexes at nuclear speckles [81]. However, there is currently little information known about the actual relevance of these dynamic interactions regarding the response of the cell to specific stimuli or its correlation with changes in transcriptomic profiles. For example, P-TEFb components colocalize at nuclear speckles with these negative regulators and also with the transcription activator adaptor Brd4 [82]. Rigorous quantitative approaches, as opposed to qualitative descriptions, especially those addressing the kinetics of these interactions, such as FRET-based studies, may complement these lines of research.

Nuclear speckles are most often adjacent to a relatively high density of transcriptionally active regions [21, 25, 46, 71, 83–85], and these active sites mostly represent elongation-competent complexes. Many of these active units correspond to specific, functionally interrelated protein-coding genes, and their juxtaposition to speckles may constitute an important part of their functional program, as has been suggested for genes involved in muscle differentiation [86]. A proposed role for nuclear speckles in these associations is the recruitment of splicing factors at specific active genes in early G1 phase, signaling for the subsequent recruitment of other functionally related gene units later on in the cell cycle [87, 88]. However, this model may be incomplete for explaining the highly dynamic behavior of these structures as observed by live-cell imaging. The inducible recruitment

of active genes compared to the dynamics of nuclear speckles has indeed been observed in live cells recently [89]. The authors proposed the following three different, nonmutually exclusive interpretations for this dynamic association such that taking the induced locus as a reference, the speckles could either be (i) assembling *de novo*, (ii) gathering by "coalescence" of smaller speckles, or (iii) recruiting the active gene to their surface. Interestingly, this inducible spatial correlation was dependent on the integrity of the inducible promoter driving the construct, *Hsp70*, which is known to be regulated by the activity of this model gene at the stage of elongation. Again, these observations might indirectly support a functional coordination between transcriptional elongation and the recruitment of pre-mRNA processing machinery. It remains to be fully resolved whether the nuclear speckles have an active role in the higher-order organization and functional coordination of the expression of specific genes or whether they rather arise as a consequence of the spatial concentration of required factors in areas proximal to active, coregulated genes. Importantly, it also remains to be elucidated if there is any posttranscriptional advantage for coregulated genes to converge at the same speckle. It would be interesting to unravel, for example, if subsets of genes that preferentially localize to the nuclear speckle periphery are enriched in genes that are mostly regulated at the level of elongation and whether both their synthesis and processing are enhanced upon appropriate recruitment to these compartments.

What is the behavior of the synthesized pre-mRNAs and mRNAs as related to nuclear speckles? A majority of introns are spliced, presumably in a cotranscriptional fashion, outside of the nuclear speckles [90–92]. This peripheral region of nuclear speckles can be therefore considered a potential interphase for cotranscriptional pre-mRNA processing. However, some introns undergo posttranscriptional processing, and their relative accumulation can be observed in these compartments, as in the case of intron 26 of the COL1A1 gene. Mutations that alter the splicing of this intron provoke increased accumulation of the transcript in nuclear speckles [91]. These data raise the possibility that nuclear speckles have a role in posttranscriptional or even postmaturation steps linked to mRNA surveillance and/or nuclear export pathways [90, 93]. Recently, a role for these structures has been proposed in the regulation of posttranscriptional "quality assessment" and the export of Herpesvirus mRNAs [94]. In this case, inducible recruitment of "viral transcription factories" to the peripheral areas of these structures was observed. Bimolecular fluorescence complementation (BiFC) experiments, showing the interaction between the exon-junction complex component Y14 and nuclear export factor 1 (NXF1) and their significant accumulation in nuclear speckles and peripheral areas, further indicate that export-competent spliced mRNAs localize at speckles and that this domain might play an active role in mRNA processing, including maturation and/or transport [95, 96]. This model is also compatible with the apparent subcompartmentalized organization of nuclear speckles [97, 98], which renders a "sponge-like" or "porous" structure that would easily allow for the transit of macromolecular complexes through its interior [99]. Coupling transcription elongation and mRNP assembly with export has been described in yeast and humans, and its potential relationship with the dynamics of mRNP transit through the speckles has been suggested recently [100–104]. Nuclear speckles may represent specialized compartments for the appropriate regulation and coordination of these functions.

Intriguingly, although transcription does not take place within nuclear speckles, a large subset of transcription factors accumulates in these regions, and transcription elongation factors are specifically enriched in these regions. RNAPII is also found to be associated with these compartments [44, 68, 105]. Quantitative laser confocal analysis of ultrathin cryosections has shown that nuclear speckles do not act as major storage sites for inactive complexes, but they instead contain a minor, stable pool of RNAPII molecules phosphorylated at the serine 2 residue of the carboxyl-terminal domain (CTD). Importantly, this subpopulation of RNAPII-2pSer is insensitive to DRB treatment [85, 105]. This fact and the absence of *de novo* synthesis of transcripts at nuclear speckles, as determined by UTP analogue incorporation, suggest that these hyperphosphorylated forms of RNAPII are not engaged in active transcription. An intriguing interpretation might be that these RNAPII molecules serve as a platform for posttranscriptional splicing of transcripts that are trafficking through the speckles.

4. Gene Expression Regulation at the Single-Cell Resolution: Studying the Kinetics of RNA Biogenesis

The study of spatial and dynamic properties is essential for understanding gene expression regulation. Techniques, such as FRAP and fluorescence loss in photobleaching (FLIP), are used to obtain very detailed information about diffusion rates, residency times, or proportion of immobile or stably tethered subpopulations of a given molecule in a delimited volume in the cell at a very high temporal resolution. Fluorescence Resonance Energy Transfer (FRET) and fluorescence lifetime imaging microscopy (FLIM), BiFluorescent complementation (BiFC), and specific applications based on fluorescence correlation spectroscopy (FCS) allow the monitoring and semiquantitation of close (mostly direct) interactions between molecules and the mapping of such interactions in relation to different structures of the cell. The use of these techniques in the study of transcription and pre-mRNA processing has led to interesting new concepts regarding their regulation and nuclear organization [106–112].

Not all nuclear factors freely diffuse through the nuclear space. The movement of some of these are compatible with a model in which factors "scan" unspecific genomic sequences or/and bind components of the RNA machinery through a weak and transient binding until they engage in a favorable, specific assembly on their target sites [113]. Nuclear functions and organization likely arise not from the static state of their components but from an extremely dynamic equilibrium between multiple functional interactions

[10, 114]. Rino and coworkers [115] found that nuclear speckles acquired a rounder and more quiescent morphology, as expected, upon transcription elongation inhibition using the P-TEFb inhibitor DRB. However, when studying the interchange rate of molecules bound to nuclear speckles within the nucleoplasm pool using FLIP, they found that the fluorescence was lost at a higher speed than it was in cells that had not been treated with DRB. Similar studies have been conducted for the transcription elongation complex P-TEFb in the context of Tat-mediated transactivation of the HIV-1 [53, 116]. Interestingly, these studies suggest that a potential mechanism by which Tat might contribute to P-TEFb-mediated transactivation is through the stabilization of CDK9 binding to the transcription site, increasing its residency time by almost tenfold. These observations suggest that cellular factors are constantly engaged in dynamic and highly transient interactions even within some apparently static structures.

Specific interactions among different spliceosomal components have also been studied in the context of nuclear organization and live-cell behavior using FRET and FLIM techniques. The dynamic model mentioned above is also compatible with the presence of preassembled subcomplexes, such as the spliceosomal components, which can be dynamically recruited to form even higher-order functional complexes [114]. Of note, these subcomplexes exist in even the absence of ongoing transcription and in nuclear compartments in which they do not actively function [75, 115, 117].

Recent adaptations of these techniques have been used to study fluorescent proteins that interact with high affinity and specificity to DNA and RNA sequences, such as the LacZ bacterial repressor and the MS2 and PP7 nucleocapsid coating proteins, respectively. The use of engineered constructs containing several tandem repeats of these target elements allows for the efficient recruitment of fluorescent molecules to the desired chromatin template or nascent transcript, amplification of the signal, and reliable detection by conventional light microscopy in living cells. FRAP measurements can then be used to estimate the rates of transcript elongation and release because these stages of mRNA biogenesis correlate with distinct kinetic steps, which can be mathematically dissected with appropriate modeling using the gathered data. Additionally, by combining this transcript-tagging system with the tagging of RNAPII with a distinct fluorophore, the dynamics of promoter binding and transcription initiation can also be inferred using the aforementioned data of transcript synthesis kinetics [118]. Interestingly, transcription appears to be a rather inefficient process, as only approximately 1% of recruited RNAPII molecules are engaged in processive transcription; however, these figures may vary considerably depending on the genetic model studied [119, 120]. These types of studies have led to an estimation of the rate of elongation for the human RNAPII of ~4.3 kb/min, although these figures can also vary widely depending on the experimental setting used [118, 121]. For example, a study based on the HIV-1 gene yielded unexpectedly long pausing times for RNAPII units located at the proximal promoter and 3′ terminal regions with an estimated elongation rate of ~1.7 kb/min [53, 119]. An important, unresolved issue is the unification of the mathematical models used to infer the kinetic properties of transcription elongation because considering or excluding certain phenomena in the reference model can drastically affect the entire interpretation [121].

These tools have been used not only for characterizing transcription dynamics, but also for facilitating the study of the *in vivo* functional coupling between transcription elongation and pre-mRNA processing, allowing us to obtain novel insights into the basis of pre-mRNA processing in the living cell. For example, it has been determined that early-spliceosomal components are actively recruited to transcribing genes lacking intronic sequences [122]. These observations are in accordance with previous biochemical and functional studies that describe a stable interaction between the initiating RNAPII complexes and the U1 snRNP [123]. In fact, stepwise cotranscriptional recruitment of the spliceosome has been reported recently [124]. Importantly, global splicing inhibition did not prevent recruitment of spliceosomal components to the active transcription site, further supporting that nuclear organization and coordination of pre-mRNA metabolism are significantly determined by transcription.

Finally, innovative microscopy and spectroscopy tools, coupled with powerful statistical analysis and modeling, have led to the first studies in estimating the dynamics of transcription at the single-molecule resolution [121, 125]. These novel approaches will allow us to gain further insights regarding single-cell behavior, and the aspects of noise, robustness, and cell-to-cell variability in pre-mRNA formation and processing, which may be important to globally understand the regulation of gene expression.

5. Concluding Remarks

The quantitative study of the spatial and dynamic aspects of transcription and pre-mRNA processing is revealing itself as an essential complement to well-established, classical biochemistry-based approaches to fully understand how the regulation of gene expression is exerted in the cell. As stated in this review, many recent insights that help to explain long-standing questions regarding mRNA biogenesis could not have been achieved otherwise. However, these studies also give rise to important new questions. What is the functional relevance of spatial organization and regulation of dynamics in the different stages of mRNA biogenesis for the cell in a given context? Can we obtain a genome-wide picture of these parameters for all protein-coding genes in a systematic manner? How is cell-to-cell variability regulated within a cellular population to be advantageous for the cellular population as a whole? Is the dynamic regulation of the spatial distribution of the involved factors an essential component for the fine tuning of functional coupling of transcription elongation and pre-mRNA processing? It is expected that there will be a remarkable increase in the usage and optimization of these approaches, combined with more conventional biochemical and functional approaches, in the study of all aspects of mRNA formation and function.

Acknowledgments

This work was supported by the Spanish Ministry of Science and Innovation Grant no. (BFU2008-01599), by Fundación para la Investigación y Prevención del SIDA en España Grant no. (36768), and by Junta de Andalucía (Proyecto de Excelencia 2009/CVI-4626). Support from the Fondo Europeo de Desarrollo Regional (FEDER) is also acknowledged.

References

[1] M. Carmo-Fonseca and C. Carvalho, "Nuclear organization and splicing control," *Advances in Experimental Medicine and Biology*, vol. 623, pp. 1–13, 2007.

[2] T. Cremer and C. Cremer, "Chromosome territories, nuclear architecture and gene regulation in mammalian cells," *Nature Reviews Genetics*, vol. 2, no. 4, pp. 292–301, 2001.

[3] L. Parada and T. Misteli, "Chromosome positioning in the interphase nucleus," *Trends in Cell Biology*, vol. 12, no. 9, pp. 425–432, 2002.

[4] R. Zhao, M. S. Bodnar, and D. L. Spector, "Nuclear neighborhoods and gene expression," *Current Opinion in Genetics and Development*, vol. 19, no. 2, pp. 172–179, 2009.

[5] M. R. Branco and A. Pombo, "Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations," *Plos Biology*, vol. 4, no. 5, article e138, 2006.

[6] G. Cavalli, "Chromosome kissing," *Current Opinion in Genetics and Development*, vol. 17, no. 5, pp. 443–450, 2007.

[7] J. Q. Ling and A. R. Hoffman, "Epigenetics of long-range chromatin interactions," *Pediatric Research*, vol. 61, no. 5, pp. 11R–16R, 2007.

[8] M. Carmo-Fonseca, "How genes find their way inside the cell nucleus," *Journal of Cell Biology*, vol. 179, no. 6, pp. 1093–1094, 2007.

[9] M. Dundr and T. Misteli, "Functional architecture in the cell nucleus," *Biochemical Journal*, vol. 356, no. 2, pp. 297–310, 2001.

[10] T. Misteli, "Concepts in nuclear architecture," *BioEssays*, vol. 27, no. 5, pp. 477–487, 2005.

[11] E. Soutoglou and T. Misteli, "Mobility and immobility of chromatin in transcription and genome stability," *Current Opinion in Genetics and Development*, vol. 17, no. 5, pp. 435–442, 2007.

[12] K. Ahmed, R. Li, and D. P. Bazett-Jones, "Electron spectroscopic imaging of the nuclear landscape," *Methods in Molecular Biology*, vol. 464, pp. 415–423, 2009.

[13] D. P. Bazett-Jones, R. Li, E. Fussner, R. Nisman, and H. Dehghani, "Elucidating chromatin and nuclear domain architecture with electron spectroscopic imaging," *Chromosome Research*, vol. 16, no. 3, pp. 397–412, 2008.

[14] H. Dehghani, G. Dellaire, and D. P. Bazett-Jones, "Organization of chromatin in the interphase mammalian cell," *Micron*, vol. 36, no. 2, pp. 95–108, 2005.

[15] G. Dellaire, R. Nisman, and D. P. Bazett-Jones, "Correlative light and electron spectroscopic imaging of chromatin in situ," *Methods in Enzymology*, vol. 375, pp. 456–478, 2004.

[16] L. de Jong, M. A. Grande, K. A. Mattern, W. Schul, and R. van Driel, "Nuclear domains involved in RNA synthesis, RNA processing, and replication," *Critical Reviews in Eukaryotic Gene Expression*, vol. 6, no. 2-3, pp. 215–246, 1996.

[17] A. I. Lamond and W. C. Earnshaw, "Structure and function in the nucleus," *Science*, vol. 280, pp. 547–553, 1998.

[18] Y. S. Mao, B. Zhang, and D. L. Spector, "Biogenesis and function of nuclear bodies," *Trends in Genetics*, vol. 27, no. 8, pp. 295–306, 2011.

[19] D. A. Jackson, A. B. Hassan, R. J. Errington, and P. R. Cook, "Visualization of focal sites of transcription within human nuclei," *The Embo Journal*, vol. 12, no. 3, pp. 1059–1065, 1993.

[20] D. G. Wansink, W. Schul, I. van der Kraan, B. van Steensel, R. van Driel, and L. de Jong, "Fluorescent labeling of nascent RNA reveals transcription by RNA polymerase II in domains scattered throughout the nucleus," *Journal of Cell Biology*, vol. 122, no. 2, pp. 283–293, 1993.

[21] F. J. Iborra, A. Pombo, D. A. Jackson, and P. R. Cook, "Active RNA polymerases are localized within discrete transcription "factories" in human nuclei," *Journal of Cell Science*, vol. 109, part 6, pp. 1427–1436, 1996.

[22] F. J. Iborra, A. Pombo, J. McManus, D. A. Jackson, and P. R. Cook, "The topology of transcription by immobilized polymerases," *Experimental Cell Research*, vol. 229, no. 2, pp. 167–173, 1996.

[23] D. A. Jackson, F. J. Iborra, E. M. Manders, and P. R. Cook, "Numbers and organization of RNA polymerases, nascent transcripts, and transcription units in HeLa nuclei," *Molecular Biology of the Cell*, vol. 9, no. 9, pp. 1523–1536, 1998.

[24] 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.

[25] X. Wei, S. Somanathan, J. Samarabandu, and R. Berezney, "Three-dimensional visualization of transcription sites and their association with splicing factor-rich nuclear speckles," *Journal of Cell Biology*, vol. 146, no. 3, pp. 543–558, 1999.

[26] C. Zeng, E. Kim, S. L. Warren, and S. M. Berget, "Dynamic relocation of transcription and splicing factors dependent upon transcriptional activity," *The Embo Journal*, vol. 16, no. 6, pp. 1401–1412, 1997.

[27] C. Lux, H. Albiez, R. D. Chapman et al., "Transition from initiation to promoter proximal pausing requires the CTD of RNA polymerase II," *Nucleic Acids Research*, vol. 33, no. 16, pp. 5139–5144, 2005.

[28] P. R. Cook, "A model for all genomes: the role of transcription factories," *Journal of Molecular Biology*, vol. 395, no. 1, pp. 1–10, 2010.

[29] P. Fraser, "Transcriptional control thrown for a loop," *Current Opinion in Genetics and Development*, vol. 16, no. 5, pp. 490–495, 2006.

[30] P. Fraser and W. Bickmore, "Nuclear organization of the genome and the potential for gene regulation," *Nature*, vol. 447, no. 7143, pp. 413–417, 2007.

[31] B. van Steensel and J. Dekker, "Genomics tools for unraveling chromosome architecture," *Nature Biotechnology*, vol. 28, no. 10, pp. 1089–1095, 2009.

[32] J. M. Brown, J. Leach, J. E. Reittie et al., "Coregulated human globin genes are frequently in spatial proximity when active," *Journal of Cell Biology*, vol. 172, no. 2, pp. 177–187, 2006.

[33] C. S. Osborne, L. Chakalova, K. E. Brown et al., "Active genes dynamically colocalize to shared sites of ongoing transcription," *Nature Genetics*, vol. 36, no. 10, pp. 1065–1071, 2004.

[34] Q. Hu, Y. S. Kwon, E. Nunez et al., "Enhancing nuclear receptor-induced transcription requires nuclear motor and LSD1-dependent gene networking in interchromatin granules," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 49, pp. 19199–19204, 2008.

[35] A. Papantonis and P. R. Cook, "Fixing the model for transcription: the DNA moves, not the polymerase," *Transcription*, vol. 2, no. 1, pp. 41–44, 2011.

[36] M. Xu and P. R. Cook, "Similar active genes cluster in specialized transcription factories," *Journal of Cell Biology*, vol. 181, no. 4, pp. 615–623, 2008.

[37] I. Montanuy, R. Torremocha, C. Hernandez-Munain, and C. Sune, "Promoter influences transcription elongation: TATA-box element mediates the assembly of processive transcription complexes responsive to cyclin-dependent kinase 9," *Journal of Biological Chemistry*, vol. 283, no. 12, pp. 7368–7378, 2008.

[38] H. Takahashi, T. J. Parmely, S. Sato et al., "Human mediator subunit MED26 functions as a docking site for transcription elongation factors," *Cell*, vol. 146, no. 1, pp. 92–104, 2011.

[39] J. A. Mitchell and P. Fraser, "Transcription factories are nuclear subcompartments that remain in the absence of transcription," *Genes and Development*, vol. 22, no. 1, pp. 20–25, 2008.

[40] J. Li and D. S. Gilmour, "Promoter proximal pausing and the control of gene expression," *Current Opinion in Genetics & Development*, vol. 21, pp. 231–235, 2011.

[41] C. H. Eskiw, A. Rapp, D. R. Carter, and P. R. Cook, "RNA polymerase II activity is located on the surface of protein-rich transcription factories," *Journal of Cell Science*, vol. 121, no. 12, pp. 1999–2007, 2008.

[42] D. R. Carter, C. Eskiw, and P. R. Cook, "Transcription factories," *Biochemical Society Transactions*, vol. 36, no. 4, pp. 585–589, 2008.

[43] 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.

[44] 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.

[45] S. Pandit, D. Wang, and X. D. Fu, "Functional integration of transcriptional and RNA processing machineries," *Current Opinion in Cell Biology*, vol. 20, no. 3, pp. 260–265, 2008.

[46] A. Pombo and P. R. Cook, "The localization of sites containing nascent RNA and splicing factors," *Experimental Cell Research*, vol. 229, no. 2, pp. 201–203, 1996.

[47] P. R. Cook, "Nongenic transcription, gene regulation and action at a distance," *Journal of Cell Science*, vol. 116, no. 22, pp. 4483–4491, 2003.

[48] T. Maniatis and R. Reed, "An extensive network of coupling among gene expression machines," *Nature*, vol. 416, pp. 499–506, 2002.

[49] G. Orphanides and D. Reinberg, "A unified theory of gene expression," *Cell*, vol. 108, no. 4, pp. 439–451, 2002.

[50] K. M. Neugebauer, "On the importance of being co-transcriptional," *Journal of Cell Science*, vol. 115, no. 20, pp. 3865–3871, 2002.

[51] P. Bjork and L. Wieslander, "Nucleocytoplasmic mRNP export is an integral part of mRNP biogenesis," *Chromosoma*, vol. 120, no. 1, pp. 23–38, 2011.

[52] R. F. Luco, M. Allo, I. E. Schor, A. R. Kornblihtt, and T. Misteli, "Epigenetics in alternative pre-mRNA splicing," *Cell*, vol. 144, no. 1, pp. 16–26, 2011.

[53] P. Maiuri, A. Knezevich, E. Bertrand, and A. Marcello, "Real-time imaging of the HIV-1 transcription cycle in single living cells," *Methods*, vol. 53, pp. 62–67, 2010.

[54] M. J. Moore and N. J. Proudfoot, "Pre-mRNA processing reaches back to transcription and ahead to translation," *Cell*, vol. 136, no. 4, pp. 688–700, 2009.

[55] 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.

[56] T. Trcek and R. H. Singer, "The cytoplasmic fate of an mRNP is determined cotranscriptionally: exception or rule?" *Genes and Development*, vol. 24, no. 17, pp. 1827–1831, 2010.

[57] M. Barboric, T. Lenasi, H. Chen, E. B. Johansen, S. Guo, and B. M. Peterlin, "7SK snRNP/P-TEFb couples transcription elongation with alternative splicing and is essential for vertebrate development," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 19, pp. 7798–7803, 2009.

[58] M. J. Munoz, M. S. P. Santangelo, M. P. Paronetto et al., "DNA damage regulates alternative splicing through inhibition of RNA polymerase II elongation," *Cell*, vol. 137, no. 4, pp. 708–720, 2009.

[59] S. E. Sanchez, E. Petrillo, E. J. Beckwith et al., "A methyl transferase links the circadian clock to the regulation of alternative splicing," *Nature*, vol. 468, no. 7320, pp. 112–116, 2010.

[60] I. E. Schor, N. Rascovan, F. Pelisch, M. Alio, 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.

[61] A. I. Lamond and D. L. Spector, "Nuclear speckles: a model for nuclear organelles," *Nature Reviews Molecular Cell Biology*, vol. 4, no. 8, pp. 605–612, 2003.

[62] D. L. Spector and A. I. Lamond, "Nuclear speckles," *Cold Spring Harbor Perspectives in Biology*, vol. 3, no. 2, 2011.

[63] M. R. Lerner, J. A. Boyle, J. A. Hardin, and J. A. Steitz, "Two novel classes of small ribonucleoproteins detected by antibodies associated with lupus erythematosus," *Science*, vol. 211, pp. 400–402, 1981.

[64] D. L. Spector, W. H. Schrier, and H. Busch, "Immunoelectron microscopic localization of SnRNPs," *Biology of the Cell*, vol. 49, no. 1, pp. 1–10, 1983.

[65] K. C. Carter, D. Bowman, W. Carrington et al., "A three-dimensional view of precursor messenger RNA metabolism within the mammalian nucleus," *Science*, vol. 259, pp. 1330–1335, 1993.

[66] K. C. Carter, K. L. Taneja, and J. B. Lawrence, "Discrete nuclear domains of poly(A) RNA and their relationship to the functional organization of the nucleus," *Journal of Cell Biology*, vol. 115, no. 5, pp. 1191–1202, 1991.

[67] P. J. Mintz, S. D. Patterson, A. F. Neuwald, C. S. Spahr, and D. L. Spector, "Purification and biochemical characterization of interchromatin granule clusters," *The Embo Journal*, vol. 18, no. 15, pp. 4308–4320, 1999.

[68] N. Saitoh, C. S. Spahr, S. D. Patterson, P. Bubulya, A. F. Neuwald, and D. L. Spector, "Proteomic analysis of interchromatin granule clusters," *Molecular Biology of the Cell*, vol. 15, no. 8, pp. 3876–3890, 2004.

[69] S. Huang and D. L. Spector, "Dynamic organization of pre-mRNA splicing factors," *Journal of Cellular Biochemistry*, vol. 62, no. 2, pp. 191–197, 1996.

[70] S. Huang and D. L. Spector, "Intron-dependent recruitment of pre-mRNA splicing factors to sites of transcription," *Journal of Cell Biology*, vol. 133, no. 4, pp. 719–732, 1996.

[71] D. Cmarko, P. J. Verschure, T. E. Martin et al., "Ultrastructural analysis of transcription and splicing in the cell nucleus after bromo-UTP microinjection," *Molecular Biology of the Cell*, vol. 10, no. 1, pp. 211–223, 1999.

[72] S. Fakan, "Perichromatin fibrils are in situ forms of nascent transcripts," *Trends in Cell Biology*, vol. 4, no. 3, pp. 86–90, 1994.

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

[74] Y. Xing, C. V. Johnson, P. T. Moen Jr., J. A. McNeil, and J. B. Lawrence, "Nonrandom gene organization: structural arrangements of specific pre-mRNA transcription and splicing with SC-35 domains," *Journal of Cell Biology*, vol. 131, no. 6, pp. 1635–1647, 1995.

[75] J. D. Ellis, D. Lleres, M. Denegri, A. I. Lamond, and J. F. Caceres, "Spatial mapping of splicing factor complexes involved in exon and intron definition," *Journal of Cell Biology*, vol. 181, no. 6, pp. 921–934, 2008.

[76] S. A. Mabon and T. Misteli, "Differential recruitment of pre-mRNA splicing factors to alternatively spliced transcripts in vivo," *Plos Biology*, vol. 3, no. 11, article e374, 1901 pages, 2005.

[77] A. Pombo, J. Ferreira, E. Bridge, and M. Carmo-Fonseca, "Adenovirus replication and transcription sites are spatially separated in the nucleus of infected cells," *The Embo Journal*, vol. 13, no. 21, pp. 5075–5085, 1994.

[78] T. Misteli and D. L. Spector, "RNA polymerase II targets pre-mRNA splicing factors to transcription sites in vivo," *Molecular Cell*, vol. 3, no. 6, pp. 697–705, 1999.

[79] X. Y. Zhong, J. H. Ding, J. A. Adams, G. Ghosh, and X. D. Fu, "Regulation of SR protein phosphorylation and alternative splicing by modulating kinetic interactions of SRPK1 with molecular chaperones," *Genes and Development*, vol. 23, no. 4, pp. 482–495, 2009.

[80] V. Tripathi, J. D. Ellis, Z. Shen et al., "The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation," *Molecular Cell*, vol. 39, no. 6, pp. 925–938, 2010.

[81] K. V. Prasanth, M. Camiolo, G. Chan et al., "Nuclear organization and dynamics of 7SK RNA in regulating gene expression," *Molecular Biology of the Cell*, vol. 21, no. 23, pp. 4184–4196, 2010.

[82] K. J. Moon, K. Mochizuki, M. Zhou, H. S. Jeong, J. N. Brady, and K. Ozato, "The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription," *Molecular Cell*, vol. 19, no. 4, pp. 523–534, 2005.

[83] S. Fakan and R. van Driel, "The perichromatin region: a functional compartment in the nucleus that determines large-scale chromatin folding," *Seminars in Cell and Developmental Biology*, vol. 18, no. 5, pp. 676–681, 2007.

[84] J. Niedojadlo, C. Perret-Vivancos, K. H. Kalland et al., "Transcribed DNA is preferentially located in the perichromatin region of mammalian cell nuclei," *Experimental Cell Research*, vol. 317, pp. 433–444, 2010.

[85] S. Q. Xie, S. Martin, P. V. Guillot, D. L. Bentley, and A. Pombo, "Splicing speckles are not reservoirs of RNA polymerase II, but contain an inactive form, phosphorylated on serine2 residues of the C-terminal domain," *Molecular Biology of the Cell*, vol. 17, no. 4, pp. 1723–1733, 2006.

[86] P. T. Moen Jr., C. V. Johnson, M. Byron et al., "Repositioning of muscle-specific genes relative to the periphery of SC-35 domains during skeletal myogenesis," *Molecular Biology of the Cell*, vol. 15, no. 1, pp. 197–206, 2004.

[87] J. M. Brown, J. Green, R. P. des Neves et al., "Association between active genes occurs at nuclear speckles and is modulated by chromatin environment," *Journal of Cell Biology*, vol. 182, no. 6, pp. 1083–1097, 2008.

[88] L. S. Shopland, C. V. Johnson, M. Byron, J. McNeil, and J. B. Lawrence, "Clustering of multiple specific genes and gene-rich R-bands around SC-35 domains: evidence for local euchromatic neighborhoods," *Journal of Cell Biology*, vol. 162, no. 6, pp. 981–990, 2003.

[89] Y. Hu, I. Kireev, M. Plutz, N. Ashourian, and A. S. Belmont, "Large-scale chromatin structure of inducible genes: transcription on a condensed, linear template," *Journal of Cell Biology*, vol. 185, no. 1, pp. 87–100, 2009.

[90] L. L. Hall, K. P. Smith, M. Byron, and J. B. Lawrence, "Molecular anatomy of a speckle," *Anatomical Record A Discoveries in Molecular, Cellular, and Evolutionary Biology*, vol. 288, no. 7, pp. 664–675, 2006.

[91] C. Johnson, D. Primorac, M. McKinstry, J. McNeil, D. Rowe, and J. B. Lawrence, "Tracking COL1A1 RNA in osteogenesis imperfecta: splice-defective transcripts initiate transport from the gene but are retained within the SC35 domain," *Journal of Cell Biology*, vol. 150, no. 3, pp. 417–431, 2000.

[92] L. S. Shopland, C. V. Johnson, and J. B. Lawrence, "Evidence that all SC-35 domains contain mRNAs and that transcripts can be structurally constrained within these domains," *Journal of Structural Biology*, vol. 140, no. 1–3, pp. 131–139, 2002.

[93] J. B. Lawrence and C. M. Clemson, "Gene associations: true romance or chance meeting in a nuclear neighborhood?" *Journal of Cell Biology*, vol. 182, no. 6, pp. 1035–1038, 2008.

[94] L. Chang, W. J. Godinez, I. H. Kim et al., "PNAS plus: herpesviral replication compartments move and coalesce at nuclear speckles to enhance export of viral late mRNA," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, pp. E136–E144, 2011.

[95] U. Schmidt, K. B. IM, C. Benzing et al., "Assembly and mobility of exon-exon junction complexes in living cells," *The Royal Naval Association*, vol. 15, no. 5, pp. 862–876, 2009.

[96] U. Schmidt, K. Richter, A. B. Berger, and P. Lichter, "In vivo BiFC analysis of Y14 and NXF1 mRNA export complexes: preferential localization within and around SC35 domains," *Journal of Cell Biology*, vol. 172, no. 3, pp. 373–381, 2006.

[97] G. Donnert, J. Keller, R. Medda et al., "Macromolecular-scale resolution in biological fluorescence microscopy," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 31, pp. 11440–11445, 2006.

[98] P. J. Mintz and D. L. Spector, "Compartmentalization of RNA processing factors within nuclear speckles," *Journal of Structural Biology*, vol. 129, no. 2-3, pp. 241–251, 2000.

[99] K. E. Handwerger, J. A. Cordero, and J. G. Gall, "Cajal bodies, nucleoli, and speckles in the Xenopus oocyte nucleus have a low-density, sponge-like structure," *Molecular Biology of the Cell*, vol. 16, no. 1, pp. 202–211, 2005.

[100] H. Cheng, K. Dufu, C. S. Lee, J. L. Hsu, A. Dias, and R. Reed, "Human mRNA export machinery recruited to the 5' end of mRNA," *Cell*, vol. 127, no. 7, pp. 1389–1400, 2006.

[101] A. P. Dias, K. Dufu, H. Lei, and R. Reed, "A role for TREX components in the release of spliced mRNA from nuclear speckle domains," *Nature Communications*, vol. 1, no. 7, 2010.

[102] R. Luna, H. Gaillard, C. Gonzalez-Aguilera, and A. Aguilera, "Biogenesis of mRNPs: integrating different processes in the eukaryotic nucleus," *Chromosoma*, vol. 117, no. 4, pp. 319–331, 2008.

[103] 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, no. 40, pp. 39037–39043, 2003.

[104] P. Valencia, A. P. Dias, and R. Reed, "Splicing promotes rapid and efficient mRNA export in mammalian cells," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 9, pp. 3386–3391, 2008.

[105] D. B. Bregman, L. Du, S. Van der Zee, and S. L. Warren, "Transcription-dependent redistribution of the large subunit of RNA polymerase II to discrete nuclear domains," *Journal of Cell Biology*, vol. 129, no. 2, pp. 287–298, 1995.

[106] A. B. Houtsmuller and W. Vermeulen, "Macromolecular dynamics in living cell nuclei revealed by fluorescence redistribution after photobleaching," *Histochemistry and Cell Biology*, vol. 115, no. 1, pp. 13–21, 2001.

[107] T. K. Kerppola, "Bimolecular fluorescence complementation (BiFC) analysis as a probe of protein interactions in living cells," *Annual Review of Biophysics*, vol. 37, pp. 465–487, 2008.

[108] M. Koster, T. Frahm, and H. Hauser, "Nucleocytoplasmic shuttling revealed by FRAP and FLIP technologies," *Current Opinion in Biotechnology*, vol. 16, no. 1, pp. 28–34, 2005.

[109] F. Mueller, D. Mazza, T. J. Stasevich, and J. G. McNally, "FRAP and kinetic modeling in the analysis of nuclear protein dynamics: what do we really know?" *Current Opinion in Cell Biology*, vol. 22, no. 3, pp. 403–411, 2010.

[110] J. Rino, J. Braga, R. Henriques, and M. Carmo-Fonseca, "Frontiers in fluorescence microscopy," *International Journal of Developmental Biology*, vol. 53, no. 8-10, pp. 1569–1579, 2009.

[111] T. C. Voss, I. A. Demarco, and R. N. Day, "Quantitative imaging of protein interactions in the cell nucleus," *BioTechniques*, vol. 38, no. 3, pp. 413–424, 2005.

[112] D. B. Zorov, E. Kobrinsky, M. Juhaszova, and S. J. Sollott, "Examining intracellular organelle function using fluorescent probes: from animalcules to quantum dots," *Circulation Research*, vol. 95, no. 3, pp. 239–252, 2004.

[113] R. D. Phair and T. Misteli, "High mobility of proteins in the mammalian cell nucleus," *Nature*, vol. 404, no. 6778, pp. 604–609, 2000.

[114] J. Rino and M. Carmo-Fonseca, "The spliceosome: a self-organized macromolecular machine in the nucleus?" *Trends in Cell Biology*, vol. 19, no. 8, pp. 375–384, 2009.

[115] J. Rino, J. M. Desterro, T. R. Pacheco, T. W. Gadella Jr., and M. Carmo-Fonseca, "Splicing factors SF1 and U2AF associate in extraspliceosomal complexes," *Molecular and Cellular Biology*, vol. 28, no. 9, pp. 3045–3057, 2008.

[116] D. Molle, P. Maiuri, S. Boireau et al., "A real-time view of the TAR:Tat:P-TEFb complex at HIV-1 transcription sites," *Retrovirology*, vol. 4, article 36, 2007.

[117] J. Chusainow, P. M. Ajuh, L. Trinkle-Mulcahy, J. E. Sleeman, J. Ellenberg, and A. I. Lamond, "FRET analyses of the U2AF complex localize the U2AF35/U2AF65 interaction in vivo and reveal a novel self-interaction of U2AF35," *The Royal Naval Association*, vol. 11, no. 8, pp. 1201–1214, 2005.

[118] X. Darzacq, Y. Shav-Tal, V. de Turris et al., "In vivo dynamics of RNA polymerase II transcription," *Nature Structural and Molecular Biology*, vol. 14, no. 9, pp. 796–806, 2007.

[119] S. Boireau, P. Maiuri, E. Basyuk et al., "The transcriptional cycle of HIV-1 in real-time and live cells," *Journal of Cell Biology*, vol. 179, no. 2, pp. 291–304, 2007.

[120] J. Yao, M. B. Ardehali, C. J. Fecko, W. W. Webb, and J. T. Lis, "Intranuclear distribution and local dynamics of RNA polymerase II during transcription activation," *Molecular Cell*, vol. 28, no. 6, pp. 978–990, 2007.

[121] T. Lionnet, B. Wu, D. Grunwald, R. H. Singer, and D. R. Larson, "Nuclear physics: quantitative single-cell approaches to nuclear organization and gene expression," *Cold Spring Harbor Symposia on Quantitative Biology*, vol. 75, pp. 113–126, 2010.

[122] B. Spiluttini, B. Gu, P. Belagal et al., "Splicing-independent recruitment of U1 snRNP to a transcription unit in living cells," *Journal of Cell Science*, vol. 123, no. 12, pp. 2085–2093, 2010.

[123] R. Das, J. Yu, Z. Zhang et al., "SR proteins function in coupling RNAP II transcription to pre-mRNA splicing," *Molecular Cell*, vol. 26, no. 6, pp. 867–881, 2007.

[124] U. Schmidt, E. Basyuk, M. C. Robert et al., "Real-time imaging of cotranscriptional splicing reveals a kinetic model that reduces noise: implications for alternative splicing regulation," *The Journal of Cell Biology*, vol. 193, pp. 819–829, 2011.

[125] D. R. Larson, D. Zenklusen, B. Wu, J. A. Chao, and R. H. Singer, "Real-time observation of transcription initiation and elongation on an endogenous yeast gene," *Science*, vol. 332, pp. 475–478, 2011.

SAGE-Hindawi Access to Research
Genetics Research International
Volume 2011, Article ID 653494, 10 pages
doi:10.4061/2011/653494

Review Article

Cryptic Transcription and Early Termination in the Control of Gene Expression

Jessie Colin,[1,2] Domenico Libri,[1,2] and Odil Porrua[1,2]

[1] *LEA Laboratory of Nuclear RNA Metabolism, Centre de Génétique Moléculaire (CNRS), UPR3404, 1 Avenue de la Terrasse, 91190 Gif sur Yvette, France*
[2] *Centre for mRNP Biogenesis and Metabolism, Aarhus University, C.F. Møllers Alle, Building 130, 8000 Aarhus C, Denmark*

Correspondence should be addressed to Odil Porrua, porrua-f@cgm.cnrs-gif.fr

Received 20 June 2011; Accepted 30 August 2011

Academic Editor: Damien Hermand

Recent studies on yeast transcriptome have revealed the presence of a large set of RNA polymerase II transcripts mapping to intergenic and antisense regions or overlapping canonical genes. Most of these ncRNAs (ncRNAs) are subject to termination by the Nrd1-dependent pathway and rapid degradation by the nuclear exosome and have been dubbed cryptic unstable transcripts (CUTs). CUTs are often considered as by-products of transcriptional noise, but in an increasing number of cases they play a central role in the control of gene expression. Regulatory mechanisms involving expression of a CUT are diverse and include attenuation, transcriptional interference, and alternative transcription start site choice. This review focuses on the impact of cryptic transcription on gene expression, describes the role of the Nrd1-complex as the main actor in preventing nonfunctional and potentially harmful transcription, and details a few systems where expression of a CUT has an essential regulatory function. We also summarize the most recent studies concerning other types of ncRNAs and their possible role in regulation.

1. Introduction

The development of new technologies in the field of transcriptome analysis has revealed an unexpected level of complexity in the eukaryotic transcription landscape. High-resolution techniques as tiling arrays and, more recently, RNA deep-sequencing have shown that a large proportion of transcripts are not associated to well-defined functional units as genes, rRNA, tRNA, and so forth, giving rise to the concepts of "pervasive" and "hidden" transcription [1, 2]. Those transcripts are often rapidly degraded, so that they remain invisible unless RNA degradation is prevented, for example, by inactivation of the degradation machinery.

Recent deep sequencing of nascent transcripts [3] has allowed a more direct analysis of RNA polymerase distribution in wild-type yeast cells obviating the need for working in mutants of the degradation pathway. These experiments have nicely confirmed the existence of hidden transcription.

In the yeast *Saccharomyces cerevisiae*, the main class of non-coding unstable RNAs transcribed by the RNA polymerase II is constituted by the Cryptic Unstable Transcripts (CUTs). CUTs are capped, relatively small, with an average length of 200 to 500 bp and with heterogeneous 3′ ends [4]. Their transcription is terminated by a pathway dependent upon the Nrd1 complex (see below), which targets them for polyadenylation and degradation by the TRAMP complex and the nuclear exosome, respectively [5, 6].

Work from two independent groups has provided a detailed picture of the genomic distribution of CUTs [7, 8], showing that the vast majority of these transcripts originate from nucleosome-free regions (NFRs) corresponding to promoter regions of *bona fide* genes. Importantly, most of the identified CUTs are transcribed divergently from the promoter region of annotated genes, suggesting that yeast promoters are intrinsically bidirectional. At least two different mechanisms control the intrinsic bidirectionality of promoters, which is a potential source of interference in gene expression. The first acts at the level of chromatin structure and involves different protein complexes implicated in modification of histones or chromatin remodelling that minimize spurious initiation [3, 9]. The second is the Nrd1-dependent termination pathway, mentioned above, which provokes

early termination and degradation of the transcripts (Figure 1). In addition to the more frequent antisense CUTs, a non-negligible number of CUTs overlap genes that are transcribed in the same sense. Some of these CUTs have been involved in the regulation of their cognate genes [7, 10–12]. Therefore, CUTs can be by-products of divergent transcription, but also functional units with an important role in the control of gene expression.

Another abundant class of non-coding RNAs has been named SUTs for Stable Unannotated Transcripts [8]. Their origin is the same as for the CUTs (5′ and 3′ end NFRs), but it has been proposed that they differ in their transcription termination mode since they are stable, and thus detectable in wild-type conditions, and often longer than CUTs [1]. However this aspect remains elusive because inactivation of a component of the canonical mRNA termination pathway that depends on the CPF-CF complex (Cleavage and Polyadenylation Factor- Cleavage Factor I) has only a minor effect on the termination of the SUTs tested [13]. A regulatory role for at least a subset of SUTs has also recently been described [14]. Finally, a new category of ncRNAs has been described very recently, which includes mainly antisense transcripts that are stabilized upon mutation of the major cytoplasmic 5′ to 3′ exoribonuclease Xrn1p [15]. These ncRNAs have been designed as XUTs for Xrn1-sensitive Unannotated Transcripts. It has to be noted that often the distinction between XUTs and other ncRNAs is blurry as considerable overlap exists between these three classes of transcripts [7, 8]. A role in repression of gene expression has been proposed for a subset of XUTs and their impact on transcription seems to be more prominent under stress conditions.

Because CUTs are by far the best-characterized class of ncRNAs in terms of origin, metabolism, and implication in regulation of gene expression, in this review we focus on the Nrd1-dependent termination pathway and its key role in limiting pervasive transcription and we describe the mechanisms of regulation that involve expression of a CUT. We also detail other cases of regulation mediated by ncRNAs belonging to other categories as SUTs or XUTs. The impact of cryptic transcription on global gene expression as well as the possible biological significance of this special way of regulation will be discussed.

2. Early Termination and Degradation in the Control of Cryptic Transcription

Pervasive transcription constitutes a risk for the cell that is controlled at different levels. Translation of aberrant or defective mRNAs that could result in toxic proteins is prevented by the nonsense-mediated decay (NMD), nonstop decay (NSD), and no-go decay (NGD) pathways in the cytoplasm [16]. However, those pathways act late in the expression process and cannot preclude possible interference of cryptic transcription with normal transcription of genes as, for example, by impeding binding of activator proteins to the promoter region of a downstream gene or by collision with elongating polymerases on a convergent gene [17–19]. Some of these deleterious effects are circumvented by the action of the Nrd1-complex that simultaneously elicits termination early in transcription and recruits the exosome-TRAMP complexes to their target RNAs, facilitating polyadenylation and consequential degradation of the transcripts [20]. The exosome is a conserved multisubunit complex with both endonuclease and 3′–5′ exonuclease activities that functions in degradation of defective transcripts as well as in processing of 3′ ends of stable ncRNAs (snRNAs, snoRNAs, and the 5.8S rRNA). The exosome has a nuclear and a cytoplasmic forms that share a core of ten proteins, being Rrp44p (also named Dis3p) the sole catalytic subunit of the core exosome. The nuclear form of the exosome possesses an additional exonuclease, Rrp6p, that also partakes in the degradation of CUTs (reviewed in [16]). The TRAMP complex has two alternative forms with a common structure, containing a poly-A polymerase (either Trf4p or Trf5p, Trf4p-containing complexes being more abundant), the DexH-box helicase Mtr4p and a zinc-knuckle RNA binding protein (either Air1p or Air2p). In contrast to the protective role of poly-A tails in mRNAs, polyadenylation by the TRAMP complex promotes exosome-mediated degradation, which is thought to be due to the lower processivity of Trf4/5, compared to the mRNA poly-A polymerase [21] and/or the shorter length of the poly-A tails added by the TRAMP complex [22]. The coupled action of the Nrd1-termination complex and the nuclear exosome allows controlling the production of aberrant transcripts at a stage prior to RNA export, possibly avoiding flooding the downstream RNA quality pathways mentioned above (NMD, NGD, *etc.*).

Even though the TRAMP complex and the nuclear exosome are important actors in the control of pervasive transcription, their role has been extensively reviewed elsewhere [23]. We will focus here on the properties of the Nrd1-complex and the data that contribute to elucidate the mechanisms of termination.

3. The Nrd1-Nab3-Sen1 Termination Complex

The Nrd1 complex was first identified for its role in termination and exosome-mediated maturation of sn- and snoRNAs [24]. Subsequently, it was shown to be responsible for termination of the novel class of ncRNAs designed as CUTs [5, 6].

The Nrd1 complex is composed by the RNA binding proteins Nrd1p and Nab3p and the RNA and RNA-DNA helicase Sen1p. Nrd1p is an essential 63 kDa protein that contains an Nab3p interacting region, an RNA recognition motif (RRM) and an N-terminal region that allows interaction with the C-terminal domain (CTD) of the large subunit of RNAP II (CID, CTD Interacting Domain [25, 26]). The CTD of RNAP II consists of tandem repeats of a hepta-peptide (YSPTSPS) that is subjected to different post-translational modifications throughout the transcription cycle and that serves as a landing pad for many proteins involved in key processes such as capping, elongation, termination, and splicing [27]. Phosphorylation at serines 2, 5, and 7 has been shown to be critical for the function and shape of the CTD. Nrd1p interacts preferentially with the Ser5-P form of RNAP II CTD *in vitro*

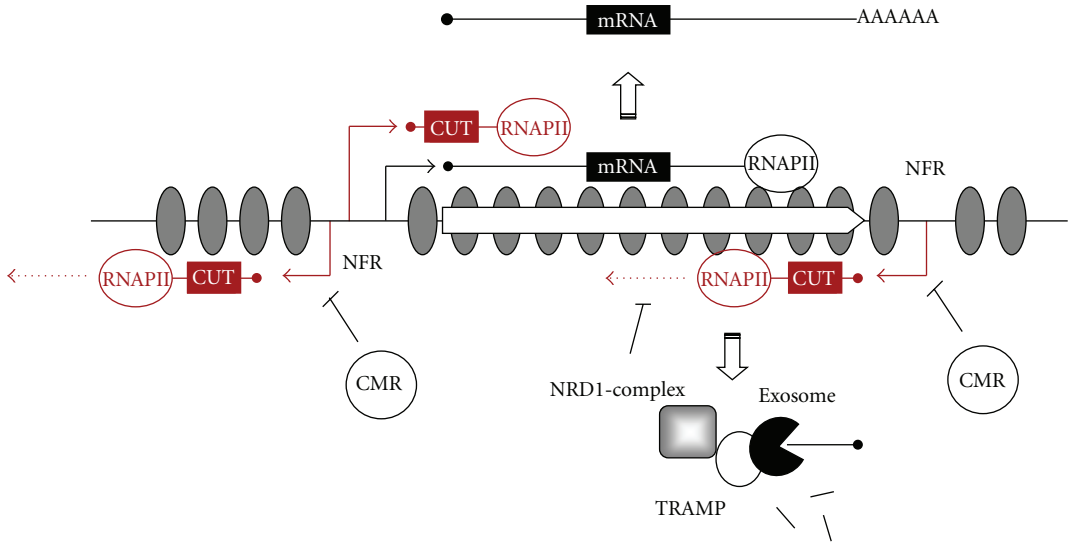


Figure 1: Complexity of the transcriptional landscape in yeast. Transcription of genes encoding stable RNAs by RNAP II is indicated by black lines and cryptic transcription originated from 5′ and 3′ nucleosome free regions (NFRs) is depicted in red. Initiation of transcription is represented by bent arrows and nucleosomes are depicted by grey ovals. Transcription can lead to the production of polyadenylated mRNAs that are competent for export to the cytoplasm and subsequent translation. Initiation of cryptic transcription is minimized by chromatin modifying and remodelling complexes (CMRs) that impose a repressive structure on the chromatin. When those mechanisms are insufficient, the Nrd1 complex terminates transcription and recruits the TRAMP and exosome complexes, which leads to polyadenylation and degradation of the generated CUT.

but colocalizes genomewide with the Ser7-P form *in vivo* [26, 28], suggesting an important role of the latter modification. Nab3p is an essential 90 kDa protein that contains an N-terminal domain rich in D and E, a central RRM, and an essential P/Q-rich C-terminal domain of unknown function [29]. It interacts directly with Nrd1p and with Sen1p [29, 30]. Mutational analysis of Nrd1-dependent termination substrates and RNA-binding assays performed with purified RRMs has led to identification of GUAA/G and UCUU as binding sites, respectively, for Nrd1p and Nab3p [31–33]. In addition, *in vivo* RNA-protein crosslinking experiments (CRAC) have recently shown that the preferred binding site for Nab3p *in vivo* is CTTG [22]. Sen1p is an essential 252 kDa protein with a role in termination of ncRNAs as well as several mRNAs [34] that also functions in DNA repair [35, 36]. The first 975 amino acids of Sen1p are dispensable for growth but are involved in the interaction with RNAPII, the RNAse III-like endonuclease Rnt1p, and the nucleotide excision repair endonuclease Rad2p [35], while the C-terminal half contains the essential helicase domain and a motif required for interaction with the Glc7p phosphatase, which dephosphorylates Sen1p *in vitro* [30].

Unlike the canonical mRNA termination pathway that depends on the CPF-CF complex and functions late in transcription, Nrd1-dependent termination is efficient only within a window of less than 1000 bp after transcription initiation, where the RNAP II CTD is phosphorylated mainly at Ser5 and Ser7 [28, 37–40].

Although it has been shown that Ser5-P (and possibly Ser7-P) favors Nrd1-dependent termination while Ser2-P antagonizes it [37], the correlation between CTD phosphorylation and termination remains not fully elucidated and the correct levels of each modification appear to be crucial. Consistent with this notion, mutation of proteins involved in the modification of the CTD, such as the Ser2 phosphatase Fcp1p [37], the Ser2 kinase Ctk1p [24], the Ser5 kinase Kin28p, the phosphatase Ssu72p [41, 42], and even the Ser-Pro isomerase Ess1p [43] affects negatively Nrd1-dependent termination. Interestingly, moreover, the role of the CTD and its modifications in termination might not pertain directly to its interaction with Nrd1p because a *nrd1*-ΔCID mutant does not exhibit any detectable termination defect [26].

In contrast to the interaction with the CTD, interaction with the RNA is strictly required for Nrd1-dependent termination [24, 26, 29, 37], although the abundance of the known Nrd1p and Nab3p recognition sites within the different termination substrates is highly variable, ranging from one to more than twelve sites [5, 10]. This variability suggests that additional termination signals yet unidentified might exist. Indeed, recent results obtained in our laboratory have revealed new motifs involved in Nrd1-dependent termination as well as specific arrangement of sites that are required for the termination signals to be functional (Porrua et al., unpublished). The heterogeneity among termination substrates concerns not only the *cis*-acting elements but also the *trans*-acting factors involved in Nrd1-dependent termination. For example, mutations in the catalytic site of the prolyl isomerase Ess1p provoke a defect in termination of a set of snoRNAs but termination of other snoRNAs and a large share of CUTs remain unaffected [43]. In addition, the poly A-binding protein Hrp1p seems to be implicated in termination of some CUTs tested but not others [10]. Furthermore, mutation or deletion of genes encoding proteins involved in histone modification as the histone methyltransferase *SET1*

and the histone deacetylase complex Rpd3L exacerbates the termination defects of *NRD1* mutants at most of the CUTs and snoRNAs analyzed but not all [44]. This apparent complexity is maybe the reason why the precise mechanism of Nrd1-dependent termination remains largely not understood. Further work, especially with *in vitro* systems, is needed to fully understand the exact role of each of the proteins involved as well as to dissect the different steps leading to transcription termination.

4. Cryptic Transcription in the Control of Gene Expression

Most of the CUTs that function in the regulation of gene expression identified thus far are located upstream or overlap the regulated gene and in the sense orientation. Production of the regulatory CUT normally has a negative impact on transcription of the downstream gene and this effect is exerted in *cis* CUTs can share the TATA box and/or the transcriptional start sites with the regulated gene or can use their owns, implying different mechanisms by which regulation occurs. In this section, we review the best-characterized examples of CUTs implicated in regulation and we briefly comment other types of ncRNAs that also control the expression of genes by different mechanisms.

4.1. Regulation by Attenuation. The term "attenuation" refers to negative control by a CUT that shares both the TATA box and the transcriptional start sites (TSS) with the regulated gene. After transcription initiation, a fraction of the elongating polymerases is subjected to early termination by the Nrd1-dependent pathway, which generates a non-functional CUT. Only the molecules that escape premature termination can proceed until the CPF-dependent terminator sequences, thus producing a functional mRNA molecule (Figure 2(a)).

The first and the best-characterized example of attenuation is autoregulation of *NRD1* itself. The *NRD1* transcript behaves both as an mRNA and as a CUT because it contains all the determinants for Nrd1-dependent termination and degradation, as well as the sequences required for normal termination and 3′ end formation by the canonical CPF-dependent pathway. The full-length *NRD1* mRNA is upregulated upon mutation or inactivation of *NRD1*, Nab3p, or *SEN1,* [24, 45]. Nrd1 has a relatively long 5′UTR of about 300 bp that contains Nrd1p and Nab3p binding sites. This sequence is sufficient for triggering Nrd1-dependent termination when inserted into an exogenous gene and provokes a 3- to-4-fold reduction in the levels of *NRD1* mRNA [24]. At least 13 additional Nrd1p- and Nab3p-binding sites are spread within the first 600 bp of *NRD1* coding sequence. Mutation of these motifs individually has a modest, if any, effect on *NRD1* autoregulation, but mutation of all of them in conjunction with modification of the sites in the 5′UTR provokes an additional 2-fold increase of mRNA levels, indicating that sequences within the coding sequence contribute to autorepression [6].

Another gene that shares this mode of regulation is *HRP1*. As in the case of *NRD1*, the *HRP1* 5′UTR is sufficient to trigger early termination when inserted in an exogenous gene. Termination is impaired upon mutation of the components of the Nrd1-complex, but, interestingly, it is also negatively affected by mutation of *HRP1* itself [10, 34]. Hrp1p is an RNA-binding protein that interacts with AU dinucleotide repeats and is implicated in pre-mRNA cleavage and polyadenylation [46] and mRNA export from the nucleus [47]. The work of Kuehner and coauthors suggests that Hrp1p acts in concert with the Nrd1-complex to regulate its own expression and possibly the expression of other genes that are subjected to attenuation.

4.2. Regulation by Transcriptional Interference. One of the best examples of transcriptional interference that implicates the production of a CUT relates to the control of the *SER3* gene, whose product catalyses a step in serine byosynthesis, in response to serine availability. The expression of *SER3* is activated in the absence of serine and is repressed under serine-replete conditions by the expression of an upstream CUT designed as *SER3* Regulatory Gene 1 (*SRG1*) [48]. *SRG1* is transcribed from its own promoter and TSS, which are different from those of *SER3,* and it extends up to the first nucleotides of the *SER3* coding region.

In the presence of serine, the sequence-specific activator Cha4p recruits the Swi/Snf chromatin remodeler and the SAGA complex to *SRG1* promoter region. These factors together activate transcription of *SRG1* that subsequently impedes binding of transcription activators and the TATA-binding protein to the *SER3* promoter region [11, 48]. Recent results have shown that the mechanism of transcriptional interference involves the assembly of nucleosomes over the *SER3* promoter region upon transcription of *SRG1*, a process that is mediated by the action of Spt6p and Spt16p, two histone chaperones that facilitate disassembly and reassembly of nucleosomes [49]. In the absence of serine, Cha4p does not interact with Swi/Snf and SAGA complexes, so transcription activation of *SRG1* does not take place and an NFR is formed at the *SER3* promoter region, allowing the binding of sequence-specific activators that are required for full expression of *SER3* (Figure 2(b)).

Termination of *SRG1* transcription occurs at two consecutive sites, being termination at the first site dependent on the Nrd1-complex and termination at the second site presumably dependent on the CPF-pathway. Given that expression of *SRG1* is strong and that each terminator alone might not be sufficiently efficient, the presence of a second terminator could constitute a fail-safe mechanism to prevent the production of a chimeric *SRG1-SER3* transcript. Indeed, depletion of the Nrd1 leads to the accumulation of *SRG1-SER3* RNAs that might eventually give rise to functional Ser3p protein under repressing conditions or to aberrant, potentially toxic translation products [5].

A similar mechanism of transcriptional interference might operate on regulation of *ADH1* and *ADH3* expression in response to zinc-limitation. *ADH1* and *ADH3* encode two different zinc-dependent alcohol dehydrogenases that are repressed during zinc deficiency by the small upstream transcripts *ZRR1* and *ZRR2*, respectively. Expression of

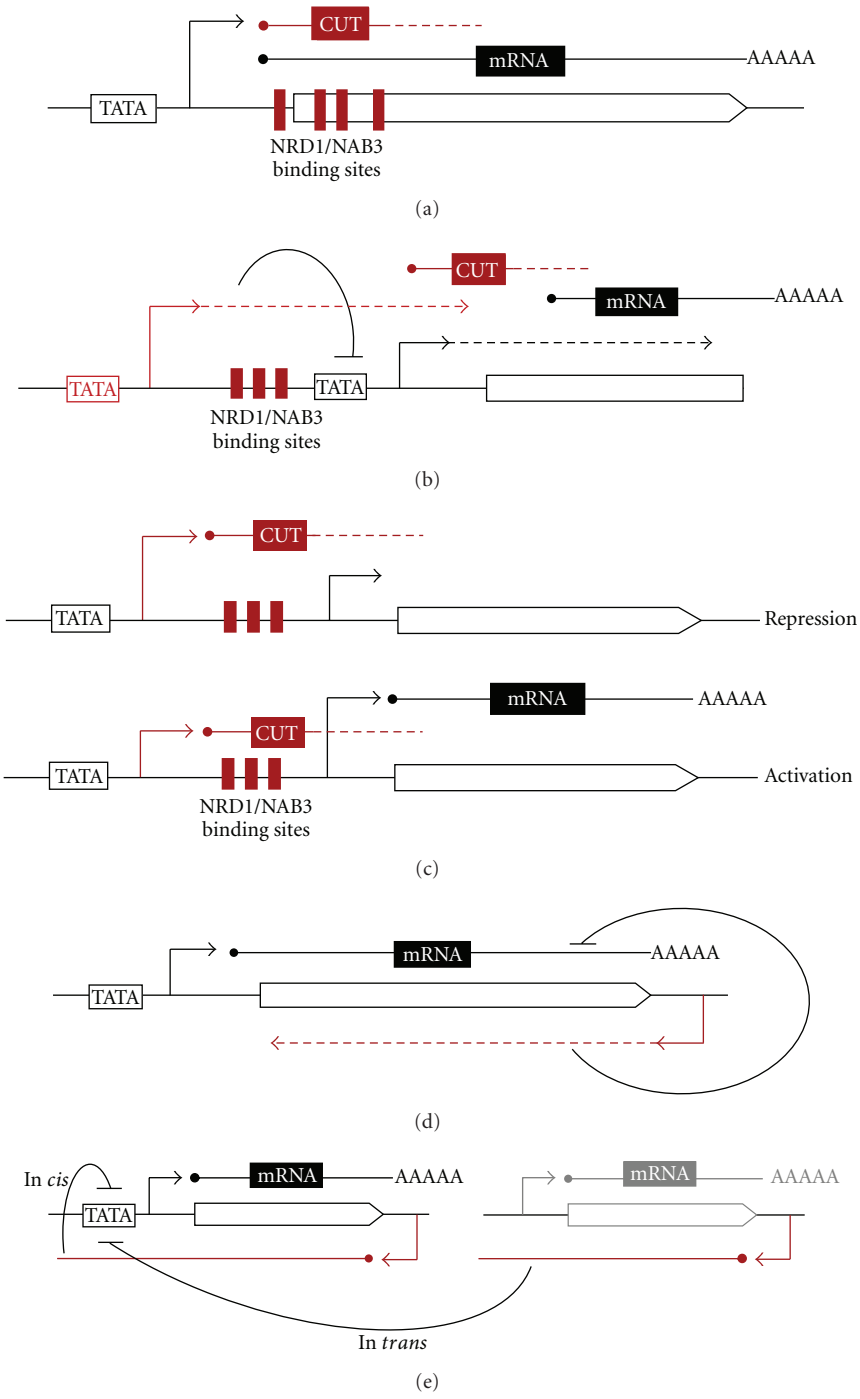


Figure 2: Summary of regulatory mechanisms involving production of ncRNAs Transcription is depicted by dashed arrows, ncRNAs by red lines, and mRNAs by black lines. Transcription start sites (TSSs) are indicated by bent arrows. Red boxes indicate Nrd1-dependent termination signals. (a) *Regulation by attenuation*: a given transcription initiation event can give rise to either an unstable transcript generated by premature termination by the Nrd1 complex or a stable mRNA if transcription is allowed to reach the CPF-dependent terminator. (b) *Regulation by transcriptional interference*: transcription of a CUT (or a stable RNA) occludes the promoter of a downstream mRNA gene thus impairing pre-initiation complex assembly and subsequent expression of the gene. (c) *Regulation by alternative TSS choice*: transcription initiation can occur either at an upstream or a downstream TSS. When the upstream TSS is selected, Nrd1-dependent termination signals are included in the transcript, leading to transcription termination and production of a CUT. When transcription starts at the downstream TSSs, these signals are skipped and a functional mRNA is produced. Regulation occurs at the level of TSS selection. (d) *Regulation by antisense transcription*: antisense transcription impairs the sense of mRNA production without affecting the initiation step. (e) *Regulation by antisense ncRNA*: long ncRNAs are able to act in *cis* and in *trans* to recruit chromatin-modifying enzymes and silence the sense gene. In this case, regulation is mediated by the ncRNA.

the upstream ncRNA is activated by the zinc-responsive regulator Zap1p, and transcription through the *ADH1* and *ADH3* promoter regions prevents binding of the required transcriptional activators [50]. Whether repression involves nucleosome deposition as for *SER3* regulation remains to be assessed. Both transcripts are relatively stable and detectable in a wild-type strain, however, at least part of the termination might be Nrd1-dependent since they are recovered in the CUT fraction in deep sequencing transcriptome analysis [7].

Finally, strong expression of upstream sense CUTs has been detected at some genes involved in glycolysis as *TPI1*, *GPM1,* and *FBA1*. These CUTs are expressed from their own promoter and TSS and are antiregulated relative to their associated mRNA, suggesting that they might be subjected to the same mechanism of transcriptional interference as *SER3* [7].

4.3. Alternative Transcriptional Start Site (TSS) Choice. Several genes of the nucleotides biosynthetic pathways *URA2*, *IMD2*, *URA8,* and *ADE12*, involved in synthesis of UTP, GTP, CTP, and ATP, respectively, are regulated by nucleotide availability and are activated when a given nucleotide is missing. All these loci share a similar organization with a 5′ overlapping CUT that is important for regulation (Figure 2(c)). The examples of *IMD2* and *URA2* are the best described [10, 12]. In both cases, expression of the CUT and the mRNA is driven by a common promoter but transcription can start either at an upstream or a downstream site. The region between the two consecutive TSSs contains all the sequence elements required for Nrd1-dependent termination and degradation, so that the use of the upstream TSS leads to the production of a CUT. On the contrary, when transcription starts at the downstream site, the Nrd1 complex termination signals are skipped and the whole coding sequence can be transcribed, leading to the production of a functional mRNA. Selection of the mRNA TSS only occurs under activating conditions (i.e., shortage of the specific nucleotide) while only the CUT TSS is used under nucleotide replete conditions. However, since the common promoter is always active, most of the regulation occurs after preinitiation complex assembly at the step of start site selection.

The presence of the upstream TSS has a clear inhibitory effect on expression of the mRNA under nonactivating con ditions, due to the fact that most polymerases never make it to the downstream TSS. Indeed, it has been shown that mutation or deletion of the *IMD2* CUT impairs full repression under guanine replete conditions [51]. Similarly, preventing production of the *URA2* CUT by mutation of its start site leads to expression of *URA2* mRNA even under nonactivating conditions [12].

What induces the selection of the downstream TSS upon activation remains, however, not fully clarified. It has been shown that under activating conditions the upstream CUT is still transcribed at the same levels as in repressing conditions in the case of *URA2* and to somewhat lower levels in the case of *IMD2.* Therefore it remains an open question whether the selection of the mRNA TSS occurs because of increased read-through at the upstream CUT TSS or by some independent mechanism. Consistent with the first notion, it has been proposed that under GTP shortage, the upstream TSS is skipped at the *IMD2* locus because it involves a G as the starting nucleotide thus allowing initiation to the downstream mRNA start site [10]. However, regulation at *URA2* has been shown to be different [12]. In this case, transcription of the CUT always starts at an A (not at U as the previous model would predict) and it is therefore difficult to imagine how UTP shortage could prevent initiation just based on nucleotide availability. Rather, it has been shown that activation of the mRNA sites requires a T-rich sequence comprised in the region between the CUT and the mRNA start sites and thus transcribed only at the CUT level [12]. This regulatory mechanism has revealed to be extremely complex and might be considerably different from one system to another. Additional work needs to be done on the different models to clarify the aspects that remain elusive. For example, it is unclear thus far whether, in addition to its repressive role under nonactivating conditions, expression of the upstream CUT plays any active role in selection of the downstream TSS upon nucleotide shortage [12]. However, it is clear that the Nrd1-dependent termination-degradation pathway plays an essential role in diverting the constitutive expression of a gene to a nonfunctional pathway, terminating and degrading transcripts initiated under conditions that do not require expression of the gene.

4.4. Other Ways of Regulation Involving an ncRNA. In this section, we include the most relevant cases of gene regulation that are mediated by other types of ncRNAs whose production is in principle independent of the Nrd1p pathway.

As we mentioned before, a second abundant class of ncRNAs includes stable transcripts that are originated by 5′ and 3′ NFRs as the CUTs and receive the name of SUTs. A recent genomewide study has shown that around 5% of genes overlap an SUT that is transcribed in the opposite strand and extends beyond the TSS of the gene. This set of ORFs is enriched in stress response and environment-specific genes and exhibits a larger expression dynamic range upon environmental changes, although identical maximal levels of expression, than genes with other configurations. The authors propose that in most cases, transcription of the antisense SUT represses expression of the sense gene under nonactivating conditions, which more efficiently locks the gene in an off-state under nonactivating conditions [14]. In addition to this possible general role of antisense transcription, condition-specific repression of a gene by an antisense ncRNA has been described in a number of cases. These sense-antisense modules can act through different mechanisms. They can act either in *cis* or in *trans* and several of them have been shown to be physiologically regulated as it is the case for the *PHO84*, *IME4*, *ZIP2,* or the *GAL1-GAL10* loci.

IME4 and *ZIP2* are two genes specifically expressed during meiosis and repressed in haploids by the cognate antisense ncRNAs *RME2* and *RME3,* respectively [18, 52]. These antisense transcripts provide a sophisticated way to activate genes as they are transcriptionally repressed in diploids by the a1p/α2p heterodimer, allowing expression of the meiotic genes. Control by *RME2* and *RME3* only works in *cis* and *via* a transcriptional interference-like mechanism. However, *RME2* and *RME3* ncRNAs do not need to extend until the

promoter region of *IME4* and *ZIP2* to exert their inhibitory function and TBP is always present in the 5′ of the sense genes *IME4* and *ZIP2*, even under repressive conditions. This suggests that the antisense RNA does not prevent formation of a preinitiation complex upstream of the sense gene but acts later, possibly by preventing elongation of the sense RNA (Figure 2(d)).

The *GAL1-GAL10* locus provides another example of *cis*-regulation by an ncRNA. At this locus, a transcript antisense to *GAL10* is generated only in glucose while *GAL1* and *GAL10* are silent whereas it becomes undetectable in galactose media when *GAL1* and *GAL10* are active. This antisense RNA allows full repression of the *GAL* locus in low glucose conditions presumably by controlling the recruitment of chromatin remodelling enzymes such as the histone methyltransferase Set1p and the histone deacetylase complex Rpd3S [53, 54].

The *PHO84* gene is also controlled by an antisense RNA. The ncRNA represses *PHO84* transcription through different mechanisms as it acts not only in *cis* but also in *trans* and in both cases, production of a long antisense RNA spanning the *PHO84* gene until its UAS sequence is necessary (Figure 2(e)). *PHO84* is regulated in ageing cells, when the antisense RNA is stabilized thus turning off *PHO84* expression. *cis*- antisense-mediated silencing requires recruitment of the histone deacetylase Hda1/2/3 which deacetylates histones at the *PHO84* promoter locus [55]. The endogenous copy of *PHO84* can also be silenced *in trans* by production of the antisense RNA from a plasmid, but in this case the silencing does not involve Hda1/2/3 [56].

The Ty1 transposon is also partially regulated by an antisense unstable transcript. This antisense RNA acts in *trans* to repress both mobility and expression of Ty1 [54]. The Ty1 antisense ncRNA is an XUT, and in a very recent work from van Dijk et al. [15], the authors propose that, as in the case of Ty1, many other genes are transcriptionally repressed by their antisense XUTs by a mechanism of silencing dependent on the action of the methyltransferase Set1p. Since repression is observed upon inactivation of the cytoplasmic RNAse Xrn1p, and subsequent stabilization of XUTs, an important mechanistic aspect that remains to be elucidated is how these cytoplasmic RNAs manage to exert their repressive role in the nucleus.

Additional data support a link between noncoding transcription and silencing [57, 58]. However, in contrast to the regulatory systems mentioned above where an ncRNA promotes silencing, here stabilization of an ncRNA derived from heterochromatic rDNA spacer region counteracts silencing through modifications in the structure of chromatin. Nonetheless, whether this effect on silencing is mediated by transcription itself or by the ncRNA is still unclear.

Finally, a different example of regulation is that of the *PHO5* locus, where transcription of an ncRNA antisense to *PHO5* is necessary for nucleosome eviction at the promoter, which is required for transcriptional activation of the gene [59]. Thus, in this case the ncRNA plays a positive role in transcription, in contrast to the previously mentioned ncRNAs, which impact negatively the expression of the cognate genes.

5. Concluding Remarks

Cryptic transcription is widespread in yeast. Recent genome-wide analyses of *Saccharomyces cerevisiae* transcriptome have revealed more than 1400 CUTs/SUTs generated by NFRs, mainly at promoter regions of *bona fide* genes, either sense or antisense to the associated gene [7, 8]. These studies have, however, been performed only in a few physiological conditions (exponential growth supported by a limited number of carbon sources) and upon mutation of only some of the components of the exosome and TRAMP complexes. Therefore, it is possible that many additional ncRNAs are produced under different conditions or in different growth phases. Consistently, a number of new unannotated transcripts have been detected upon deletion of the RNA exonuclease Xrn1p [15]. In addition, a recent analysis of the yeast transcriptome during sporulation has revealed the presence of new ncRNAs specific of the meiotic phase designed as MUTs for Meiotic Unannotated Transcripts [60]. Furthermore, several additional CUTs have been identified upon mutation of the catalytic subunit of the core exosome Dis3p (Gudipati et al., unpublished results). In addition to ncRNAs originated at 5′ and 3′ NFRs, cryptic transcription can occur at regions internal to genes upon mutation of the histone chaperones Spt6p and Spt16p [61], as well as the histone deacetylase complex Rpd3S and the histone methyltransferase Set2p [62], presumably because of the formation of transitory NFR in coding regions that can unveil cryptic promoters. This is in agreement with the notion that the chromatin structure plays an important role in the control of cryptic transcription.

The occurrence of such a high level of cryptic transcription might imply that transcription initiation by eukaryotic RNAPII can occur with relatively low specificity in the absence of active mechanisms to prevent it. This behavior is rather different from that of the prokaryotic RNA polymerase, which always binds motifs that are quite conserved and initiates transcription at a defined distance downstream from the promoter sequences [63]. In many ways, this intrinsic "promiscuity" of the eukaryotic RNAPII constitutes a disadvantage as it requires efficient pathways to (i) counteract the potentially deleterious effects of pervasive transcription on the expression of functional transcripts (e.g., the Nrd1-dependent termination pathway) and (ii) degrade a large amount of nonfunctional or aberrant RNAs that could impact negatively the physiology of the cell, for example, by sequestering the export and translation machineries or by producing toxic protein products upon translation (e.g., nuclear and cytoplasmic RNA quality control pathways). In spite of these drawbacks, the maintenance of such flexibility might imply some evolutionary and/or functional advantages. For instance, it is possible that the production of a plethora of different ncRNAs allows some of them to evolve towards functional molecules, which would then be stabilized and confer higher fitness to the cell. In addition, a high flexibility in transcription initiation provides multiple opportunities for regulation, enabling the development of sophisticated regulatory mechanisms as those reviewed here.

The expression of the ncRNA normally plays a repressive role on transcription of the cognate gene. In general, genes

that are subjected to both positive and negative control have a higher dynamic range of expression, which allows a fine-tuned response to environmental conditions [14]. In addition, active repression under nonactivating conditions minimizes "leakyness" in gene expression, avoiding production of proteins when they are not required by the cell and/or waste of energy. A remaining question is why the relatively complex negative control involving production of an ncRNA should be more advantageous than a classical regulation system based on protein repressors. The answer is not always obvious. In the case of regulation by attenuation, where the Nrd1-complex elicits partial premature termination, this mechanism provides an efficient way for the proteins involved in termination to control their own production by a feedback loop. The mechanism of transcriptional interference might allow transcriptional activators to be turned into repressors using the same molecular mechanism by which they both promote expression of the target genes and transcription of an ncRNA that prevents expression of the repressed genes. Finally, the regulation by alternative TSS selection seems to be associated to particular metabolic pathways (i.e., nucleotides biosynthesis) that require a fast response to environmental changes because they lead to the production of key molecules whose shortage would impact many cellular processes. A promoter that is already activated but not physiologically functional because it leads to the production of a CUT is more susceptible to be diverted to a functional state in the presence of the appropriate signal because it is endowed with the proper chromatin structure and a repertory of transcription factors and RNAPIIs. In that sense, production of the CUT would enable a "preactivated" state of the promoter, so that it would be ready for a fast response to changes in the intracellular environment.

Among the hundreds of ncRNAs that are associated to *bona fide* genes, only a few have been studied thus far and it is likely that expression of many more is modulated by the presence of an ncRNA and that additional regulatory mechanisms exist. Much more experimental work is needed to evaluate the global impact of cryptic transcription on gene expression and to unveil the multiple associated mechanisms of regulation.

In conclusion, this paper is focused on the main progresses in the study of ncRNAs done in the model organism *Saccharomyces cerevisiae* over the last years. However, the phenomenon of cryptic transcription and the regulation of gene expression by ncRNAs are conserved from yeast to animals. The most relevant works concerning the characterization of the noncoding transcriptome and the multiple categories of ncRNAs, as well as their associated regulatory functions, present in higher eukaryotes are nicely detailed in other recent reviews [1, 2, 64].

Acknowledgments

This work was supported by the CNRS, the ANR (Grant ANR-08-BLAN-0038-CSD 8), the Fondation pour la Recherche Médicale (FRM), and the Danish National Research Foundation. The research was carried out within the scope of the Associated European Laboratory (LEA) "Laboratory of Nuclear RNA Metabolism." O. Porrua is supported by a fellowship from the European Molecular Biology Organization.

References

[1] A. Jacquier, "The complex eukaryotic transcriptome: unexpected pervasive transcription and novel small RNAs," *Nature Reviews Genetics*, vol. 10, no. 12, pp. 833–844, 2009.

[2] J. Berretta and A. Morillon, "Pervasive transcription constitutes a new level of eukaryotic genome regulation," *EMBO Reports*, vol. 10, no. 9, pp. 973–982, 2009.

[3] L. S. Churchman and J. S. Weissman, "Nascent transcript sequencing visualizes transcription at nucleotide resolution," *Nature*, vol. 469, no. 7330, pp. 368–373, 2011.

[4] F. Wyers, M. Rougemaille, G. Badis et al., "Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase," *Cell*, vol. 121, no. 5, pp. 725–737, 2005.

[5] M. Thiebaut, E. Kisseleva-Romanova, M. Rougemaille, J. Boulay, and D. Libri, "Transcription termination and nuclear degradation of cryptic unstable transcripts: a role for the Nrd1-Nab3 pathway in genome surveillance," *Molecular Cell*, vol. 23, no. 6, pp. 853–864, 2006.

[6] J. T. Arigo, K. L. Carroll, J. M. Ames, and J. L. Corden, "Regulation of yeast NRD1 expression by premature transcription termination," *Molecular Cell*, vol. 21, no. 5, pp. 641–651, 2006.

[7] H. Neil, C. Malabat, Y. D'Aubenton-Carafa, Z. Xu, L. M. Steinmetz, and A. Jacquier, "Widespread bidirectional promoters are the major source of cryptic transcripts in yeast," *Nature*, vol. 457, no. 7232, pp. 1038–1042, 2009.

[8] Z. Xu, W. Wei, J. Gagneur et al., "Bidirectional promoters generate pervasive transcription in yeast," *Nature*, vol. 457, no. 7232, pp. 1033–1037, 2009.

[9] I. Whitehouse, O. J. Rando, J. Delrow, and T. Tsukiyama, "Chromatin remodelling at promoters suppresses antisense transcription," *Nature*, vol. 450, no. 7172, pp. 1031–1035, 2007.

[10] J. N. Kuehner and D. A. Brow, "Regulation of a eukaryotic gene by GTP-dependent start site selection and transcription attenuation," *Molecular Cell*, vol. 31, no. 2, pp. 201–211, 2008.

[11] J. A. Martens, P. Y. J. Wu, and F. Winston, "Regulation of an intergenic transcript controls adjacent gene transcription in Saccharomyces cerevisiae," *Genes and Development*, vol. 19, no. 22, pp. 2695–2704, 2005.

[12] M. Thiebaut, J. Colin, H. Neil et al., "Futile cycle of transcription initiation and termination modulates the response to nucleotide shortage in S. cerevisiae," *Molecular cell*, vol. 31, no. 5, pp. 671–682, 2008.

[13] S. Marquardt, D. Z. Hazelbaker, and S. Buratowski, "Distinct RNA degradation pathways and 3' extensions of yeast noncoding RNA species," *Transcription*, vol. 2, no. 3, pp. 145–154, 2011.

[14] Z. Xu, W. Wei, J. Gagneur et al., "Antisense expression increases gene expression variability and locus interdependency," *Molecular Systems Biology*, vol. 7, p. 468, 2011.

[15] E. L. van Dijk, C. L. Chen, Y. D'Aubenton-Carafa et al., "XUTs are a class of Xrn1-sensitive antisense regulatory non-coding RNA in yeast," *Nature*, vol. 475, no. 7354, pp. 114–117, 2011.

[16] J. Houseley, J. LaCava, and D. Tollervey, "RNA-quality control by the exosome," *Nature Reviews Molecular Cell Biology*, vol. 7, no. 7, pp. 529–539, 2006.

[17] I. H. Greger, A. Aranda, and N. Proudfoot, "Balancing transcriptional interference and initiation on the GAL7 promoter of Saccharomyces cerevisiae," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 15, pp. 8415–8420, 2000.

[18] C. F. Hongay, P. L. Grisafi, T. Galitski, and G. R. Fink, "Antisense transcription controls cell fate in Saccharomyces cerevisiae," *Cell*, vol. 127, no. 4, pp. 735–745, 2006.

[19] A. C. Palmer, J. B. Egan, and K. E. Shearwin, "Transcriptional interference by RNA polymerase pausing and dislodgement of transcription factors," *Transcription*, vol. 2, no. 1, pp. 9–14, 2011.

[20] L. Vasiljeva and S. Buratowski, "Nrd1 interacts with the nuclear exosome for 3′ processing of RNA polymerase II transcripts," *Molecular Cell*, vol. 21, no. 2, pp. 239–248, 2006.

[21] J. LaCava, J. Houseley, C. Saveanu et al., "RNA degradation by the exosome is promoted by a nuclear polyadenylation complex," *Cell*, vol. 121, no. 5, pp. 713–724, 2005.

[22] W. Wlotzka, G. Kudla, S. Granneman, and D. Tollervey, "The nuclear RNA polymerase II surveillance system targets polymerase III transcripts," *EMBO Journal*, vol. 30, no. 9, pp. 1790–1803, 2011.

[23] J. Houseley and D. Tollervey, "The many pathways of RNA degradation," *Cell*, vol. 136, no. 4, pp. 763–776, 2009.

[24] E. J. Steinmetz, N. K. Conrad, D. A. Brow, and J. L. Corden, "RNA-binding protein Nrd1 directs poly(A)-independent 3′-end formation of RNA polymerase II transcripts," *Nature*, vol. 413, no. 6853, pp. 327–331, 2001.

[25] N. K. Conrad, S. M. Wilson, E. J. Steinmetz et al., "A yeast heterogeneous nuclear ribonucleoprotein complex associated with RNA polymerase II," *Genetics*, vol. 154, no. 2, pp. 557–571, 2000.

[26] L. Vasiljeva, M. Kim, H. Mutschler, S. Buratowski, and A. Meinhart, "The Nrd1-Nab3-Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain," *Nature Structural and Molecular Biology*, vol. 15, no. 8, pp. 795–804, 2008.

[27] S. Buratowski, "Progression through the RNA polymerase II CTD cycle," *Molecular Cell*, vol. 36, no. 4, pp. 541–546, 2009.

[28] 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.

[29] K. L. Carroll, R. Ghirlando, J. M. Ames, and J. L. Corden, "Interaction of yeast RNA-binding proteins Nrd1 and Nab3 with RNA polymerase II terminator elements," *RNA*, vol. 13, no. 3, pp. 361–373, 2007.

[30] E. Nedea, D. Nalbant, D. Xia et al., "The Glc7 phosphatase subunit of the cleavage and polyadenylation factor is essential for transcription termination on snoRNA genes," *Molecular Cell*, vol. 29, no. 5, pp. 577–587, 2008.

[31] K. L. Carroll, D. A. Pradhan, J. A. Granek, N. D. Clarke, and J. L. Corden, "Identification of cis elements directing termination of yeast nonpolyadenylated snoRNA transcripts," *Molecular and Cellular Biology*, vol. 24, no. 14, pp. 6241–6252, 2004.

[32] F. Hobor, R. Pergoli, K. Kubicek et al., "Recognition of transcription termination signal by the nuclear polyadenylated RNA-binding (NAB) 3 protein," *Journal of Biological Chemistry*, vol. 286, no. 5, pp. 3645–3657, 2011.

[33] B. M. Lunde, M. Hörner, and A. Meinhart, "Structural insights into cis element recognition of non-polyadenylated RNAs by the Nab3-RRM," *Nucleic Acids Research*, vol. 39, no. 1, pp. 337–346, 2011.

[34] E. J. Steinmetz, C. L. Warren, J. N. Kuehner, B. Panbehi, A. Z. Ansari, and D. A. Brow, "Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase," *Molecular Cell*, vol. 24, no. 5, pp. 735–746, 2006.

[35] D. Ursic, K. Chinchilla, J. S. Finkel, and M. R. Culbertson, "Multiple protein/protein and protein/RNA interactions suggest roles for yeast DNA/RNA helicase Sen1p in transcription, transcription-coupled DNA repair and RNA processing," *Nucleic Acids Research*, vol. 32, no. 8, pp. 2441–2452, 2004.

[36] H. E. Mischo, B. Gómez-González, P. Grzechnik et al., "Yeast Sen1 helicase protects the genome from transcription-associated instability," *Molecular Cell*, vol. 41, no. 1, pp. 21–32, 2011.

[37] R. K. Gudipati, T. Villa, J. Boulay, and D. Libri, "Phosphorylation of the RNA polymerase II C-terminal domain dictates transcription termination choice," *Nature Structural and Molecular Biology*, vol. 15, no. 8, pp. 786–794, 2008.

[38] M. H. Jenks, T. W. O'Rourke, and D. Reines, "Properties of an intergenic terminator and start site switch that regulate IMD2 transcription in yeast," *Molecular and Cellular Biology*, vol. 28, no. 12, pp. 3883–3893, 2008.

[39] 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.

[40] J. R. Tietjen, D. W. Zhang, J. B. Rodríguez-Molina et al., "Chemical-genomic dissection of the CTD code," *Nature Structural and Molecular Biology*, vol. 17, no. 9, pp. 1154–1161, 2010.

[41] C. Ganem, F. Devaux, C. Torchet et al., "Ssu72 is a phosphatase essential for transcription termination of snoRNAs and specific mRNAs in yeast," *EMBO Journal*, vol. 22, no. 7, pp. 1588–1598, 2003.

[42] E. J. Steinmetz and D. A. Brow, "Ssu72 protein mediates both poly(A)-coupled and poly(A)-independent termination of RNA polymerase II transcription," *Molecular and Cellular Biology*, vol. 23, no. 18, pp. 6339–6349, 2003.

[43] N. Singh, Z. Ma, T. Gemmill et al., "The Ess1 prolyl isomerase is required for transcription termination of small noncoding RNAs via the Nrd1 pathway," *Molecular Cell*, vol. 36, no. 2, pp. 255–266, 2009.

[44] N. Terzi, L. S. Churchman, L. Vasiljeva, J. Weissman, and S. Buratowski, "H3K4 trimethylation by Set1 promotes efficient termination by the Nrd1-Nab3-Sen1 pathway," *Molecular and Cellular Biology*, vol. 31, pp. 3569–3583, 2011.

[45] E. J. Steinmetz, S. B. H. Ng, J. P. Cloute, and D. A. Brow, "Cis- and trans-acting determinants of transcription termination by yeast RNA polymerase II," *Molecular and Cellular Biology*, vol. 26, no. 7, pp. 2688–2696, 2006.

[46] L. Minvielle-Sebastia, K. Beyer, A. M. Krecic, R. E. Hector, M. S. Swanson, and W. Keller, "Control of cleavage site selection during mRNA 3' end formation by a yeast hnRNP," *EMBO Journal*, vol. 17, no. 24, pp. 7454–7468, 1998.

[47] M. M. Kessler, M. F. Henry, E. Shen et al., "Hrp1, a sequence-specific RNA-binding protein that shuttles between the nucleus and the cytoplasm, is required for mRNA 3'-end formation in yeast," *Genes and Development*, vol. 11, no. 19, pp. 2545–2556, 1997.

[48] J. A. Martens, L. Laprade, and F. Winston, "Intergenic transcription is required to repress the Saccharomyces cerevisiae SER3 gene," *Nature*, vol. 429, no. 6991, pp. 571–574, 2004.

[49] S. J. Hainer, J. A. Pruneski, R. D. Mitchell, R. M. Monteverde, and J. A. Martens, "Intergenic transcription causes repression

by directing nucleosome assembly," *Genes and Development*, vol. 25, no. 1, pp. 29–40, 2011.

[50] A. J. Bird, M. Gordon, D. J. Eide, and D. R. Winge, "Repression of ADH1 and ADH3 during zinc deficiency by Zap1-induced intergenic RNA transcripts," *EMBO Journal*, vol. 25, no. 24, pp. 5726–5734, 2006.

[51] K. A. Kopcewicz, T. W. O'Rourke, and D. Reines, "Metabolic regulation of IMD2 transcription and an unusual DNA element that generates short transcripts," *Molecular and Cellular Biology*, vol. 27, no. 8, pp. 2821–2829, 2007.

[52] B. Gelfand, J. Mead, A. Bruning et al., "Regulated antisense transcription controls expression of cell-type-specific genes in yeast," *Molecular and Cellular Biology*, vol. 31, no. 8, pp. 1701–1709, 2011.

[53] M. Pinskaya, S. Gourvennec, and A. Morillon, "H3 lysine 4 di- and tri-methylation deposited by cryptic transcription attenuates promoter activation," *EMBO Journal*, vol. 28, no. 12, pp. 1697–1707, 2009.

[54] J. Houseley, L. Rubbi, M. Grunstein, D. Tollervey, and M. Vogelauer, "A ncRNA modulates histone modification and mRNA induction in the yeast GAL gene cluster," *Molecular Cell*, vol. 32, no. 5, pp. 685–695, 2008.

[55] J. Camblong, N. Iglesias, C. Fickentscher, G. Dieppois, and F. Stutz, "Antisense RNA stabilization induces transcriptional gene silencing via histone deacetylation in S. cerevisiae," *Cell*, vol. 131, no. 4, pp. 706–717, 2007.

[56] J. Camblong, N. Beyrouthy, E. Guffanti, G. Schlaepfer, L. M. Steinmetz, and F. Stutz, "Trans-acting antisense RNAs mediate transcriptional gene cosuppression in S. cerevisiae," *Genes and Development*, vol. 23, no. 13, pp. 1534–1545, 2009.

[57] J. Houseley, K. Kotovic, A. El Hage, and D. Tollervey, "Trf4 targets ncRNAs from telomeric and rDNA spacer regions and functions in rDNA copy number control," *EMBO Journal*, vol. 26, no. 24, pp. 4996–5006, 2007.

[58] L. Vasiljeva, M. Kim, N. Terzi, L. M. Soares, and S. Buratowski, "Transcription termination and RNA degradation contribute to silencing of RNA polymerase II transcription within heterochromatin," *Molecular Cell*, vol. 29, no. 3, pp. 313–323, 2008.

[59] J. P. Uhler, C. Hertel, and J. Q. Svejstrup, "A role for noncoding transcription in activation of the yeast PHO5 gene," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 19, pp. 8011–8016, 2007.

[60] A. Lardenois, Y. Liu, T. Walther et al., "Execution of the meiotic noncoding RNA expression program and the onset of gametogenesis in yeast require the conserved exosome subunit Rrp6," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, no. 3, pp. 1058–1063, 2011.

[61] C. D. Kaplan, L. Laprade, and F. Winston, "Transcription elongation factors repress transcription initiation from cryptic sites," *Science*, vol. 301, no. 5636, pp. 1096–1099, 2003.

[62] M. J. Carrozza, B. Li, L. Florens et al., "Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription," *Cell*, vol. 123, no. 4, pp. 581–592, 2005.

[63] T. Ghosh, D. Bose, and X. Zhang, "Mechanisms for activating bacterial RNA polymerase," *FEMS Microbiology Reviews*, vol. 34, no. 5, pp. 611–627, 2010.

[64] M. Wery, M. Kwapisz, A. Morillon et al., "Noncoding RNAs in gene regulation," *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, vol. 3, no. 6, pp. 728–738, 2011.

SAGE-Hindawi Access to Research
Genetics Research International
Volume 2011, Article ID 989303, 12 pages
doi:10.4061/2011/989303

Review Article

Genomic Insights into the Different Layers of Gene Regulation in Yeast

José E. Pérez-Ortín, Daniel A. Medina, and Antonio Jordán-Pla

Departamento de Bioquímica y Biología Molecular, Facultad de Biológicas, Universitat de València, C/Dr. Moliner 50, 46100 Burjassot, Spain

Correspondence should be addressed to José E. Pérez-Ortín, jose.e.perez@uv.es

Received 26 June 2011; Accepted 26 August 2011

Academic Editor: Sebastián Chávez

The model organism *Saccharomyces cerevisiae* has allowed the development of new functional genomics techniques devoted to the study of transcription in all its stages. With these techniques, it has been possible to find interesting new mechanisms to control gene expression that act at different levels and for different gene sets apart from the known cis-trans regulation in the transcription initiation step. Here we discuss a method developed in our laboratory, Genomic Run-On, and other new methods that have recently appeared with distinct technical features. A comparison between the datasets generated by them provides interesting genomic insights into the different layers of gene regulation in eukaryotes.

1. Functional Genomics Techniques as a Driving Force for Biology

In the 1980s, Sidney Brenner stated that, "...progress (*in biology*) depends on the interplay of techniques, discoveries and new ideas, probably in that order of decreasing importance" [1]. It is absolutely true that most scientific revolutions have appeared after technological developments which are, directly or indirectly, the bases for obtaining new kinds of data which, in turn, have led to the emergence of new ideas among contemporary scientists. This statement also holds true for the commonest case of nonrevolutionary developments. All new tools created by scientists or technicians are always followed by novel data which, in most cases, have led to new proposals, hypotheses, or theories.

Since genome sequencing projects began in the early 1990s, a new biology concept has started. This concept was not new, but the development of new technologies to sequence and analyze whole genomes provided such an amount of new data, that a new kind of biological science was born, the so-called "omics" [2]. Genomics and other omics sciences have made a real revolution of molecular biology itself. This is especially true because the preceding molecular biology was an especially reductionist science; that is, genes, proteins, and pathways were mainly analyzed and screened individually in an attempt to decipher each one in the most in-depth way possible. Obviously, although the search for relationships among genes, proteins, and pathways was also underway, all integrative approaches lacked the most important component to be fully developed: data. At the same time, molecular systems biology came into being after Jacob and Monod's operon model [3]. Although, it was restricted to a few genes, proteins, and pathways, it never attempted to check if proposed mathematical models were more or less common in cells, and it certainly never dreamed of building comprehensive models to explain the whole behavior of a living cell.

The sequencing of the first eukaryote, the yeast *Saccharomyces cerevisiae*, made it possible to develop a totally new field in biology: functional genomics [4]. Until that time, genomics was a science devoted to obtaining genome sequences and "in silico" analyses of them. Given the availability of a whole genome sequence for a model organism, for which a huge amount of biological information existed and because of the awesome power of yeast genetics, it was possible to develop totally new tools and specialized mutant collections in a relatively short time (see [2], for a review). It also provided the data for establishing molecular systems

biology from an omics perspective [2]. The characteristic collaborative atmosphere of yeast genetics and a molecular biology community contributed to the rapid establishment of databanks (e.g., SGD or CYGD, [5, 6]), transnational projects [5], and strain repositories (e.g., Euroscarf), which are freely available for any interested scientist.

One of the most successful techniques in functional genomics has been microarray technology. Microarrays were fully developed by the mid-1990s, using mainly *S. cerevisiae* as a workhorse for many of the technological advances (reviewed in [7]). Different platforms have been created since 1997 for the whole yeast genome; for instance, glass cDNA microarrays [8] which have been the most widely used. Nylon macroarrays [9] were one of the first to be developed and are still a convenient alternative for specialized purposes [7]. Oligonucleotide arrays were developed by several laboratories and companies and are currently the most used alternative, especially the type known as tiling arrays, which cover the whole genome without leaving any gap in it. They have been used to discover totally unpredicted genes, noncanonical transcripts, either sense or antisense, as regards canonical genes [10] and to also locate the binding sites of many of the transcription-related proteins of this lower eukaryote [11].

Sidney Brenner's actual opinion does not correspond exactly to what people may think about his quoted sentence [1]. In fact, he has declared that this new emerging genomic approach is "low input, high throughput and no output biology" [12]. This opinion is widely extended among biologists because it seems that genomic techniques are just "fishing expeditions" in which there is no previous hypothesis to support them. Obviously this criticism is, at least in part, false. Each new genomic technique's own technical protocol is devoted to catching new kinds of fishes. Although the nature of these new fishes is not totally predicted in advance, it is rather obvious that there is a basal hypothesis in the technique's background: we are going to fish new, unknown specimens that will have special features that our new technique will be able to catch. A good example of this is the discovery of CUT (cryptic unstable transcripts) anywhere in the genome, but mainly in relation to canonical genes *loci* by means of tiling arrays and high-throughput sequencing (HTS) [13–15]. An additional corollary of this is the interpretation of the fishing expedition as not being a trivial question. Identifying the fishes and investigating the biological mechanisms that originate them are also ways of making science. In fact, the generation of new results is not only a natural consequence of the development of a new technique, but also a previous step to put forward a new hypothesis. For instance, the analysis of CUTs has brought about the discovery of new mechanisms of mRNA quality control and transcription termination [13].

In our lab, we study the whole gene transcription process using *S. cerevisiae* as a model system. RNAs, especially mRNAs, are unstable molecules. They are degraded by exo- and endonucleases, mostly in the cytoplasm [16]. The amount of mRNA (RA), therefore, is not just the result of transcription, but the equilibrium caused by transcription rates (TRs) and degradation rates (DRs). We realized that the genomic techniques available at the beginning of the 21st century were able to quantify RA, but that there were no techniques available to measure turnover rates. In 2002, Pat Brown's group developed a genomic technique to measure mRNA stabilities in yeast [17]. This technique was an upgrade of the well-established transcription shutoff protocols used for individual mRNA half-life studies [18]. These protocols cause stress in cells, which impedes the measurement of many mRNAs' half-lives involved in the stress response [19]. The cell's physiology is also affected: it is not necessarily true that the measured half-lives correspond to the real ones in nonperturbed cells [20]. For this reason, we developed a new protocol called Genomic Run-On (GRO), which is able to measure nascent transcription rates (TR) for all yeast genes and, at the same time, the mRNA amounts (RA) for them. In this way, mRNA stabilities can be determined from RA and TR in cases in which RA does not change, in steady-state conditions [21], and even during abrupt changes in RA after stress [20].

2. Genomic Run-On (GRO) for Yeast Cells. Features of the Nascent TR Dataset

The run-on method is a well-known procedure for detecting elongating RNA polymerases (RNA pols) in eukaryotic nuclei [22]. Figure 1 depicts the outline of the method. In most eukaryotes, it is necessary to isolate nuclei before doing the experiment [22]. In yeasts, however, whole cells can be directly used, which allows a physiological freezing of the actual transcription state in those organisms. The permeabilization of cells by means of a detergent (usually sarkosyl) provokes a sudden decrease of the NTP pools and stalls all elongating RNA pol. The detergent also disrupts the chromatin structure, thus avoiding any further initiation event. A subsequent pulse of externally added ribonucleotides, including labeled UTP (^{33}P-UTP or derivatized-UTP), induces a death rattle of these RNA pols, which actively elongate. This *postmortem* elongation labels the RNA molecules with a natural sequence of about 200–300 nucleotides. Those RNA pols that were backtracked or did not elongate do not incorporate nucleotides (red and yellow ovals in Figure 1). Most nascent RNA becomes labeled in this way. This RNA is probably less than 1% of the total RNA in the cell. Because it is labeled, it is possible to use it as a hybridizable molecule in a dot blot or a DNA array or, alternatively, to purify it based on any unique property conferred by the UTP analog used. The latter could be used for high-throughput sequencing or for DNA microarray hybridization. The signal associated with a particular sequence (the probe in the DNA array) is a reflection of the RNA pol density in it. If we assume a constant speed for the RNA pol, then RNA pol density is proportional to its transcription rate.

In 2004, we developed [23] a genomic upgrade of the yeast run-on technique which we called "genomic run-on" (GRO). It is based on a subsequent hybridization of RNA extracted onto whole yeast genome nylon macroarrays [24]. This acronym has been used for a similar technique in human cells [25]. The fact that this technique is user friendly

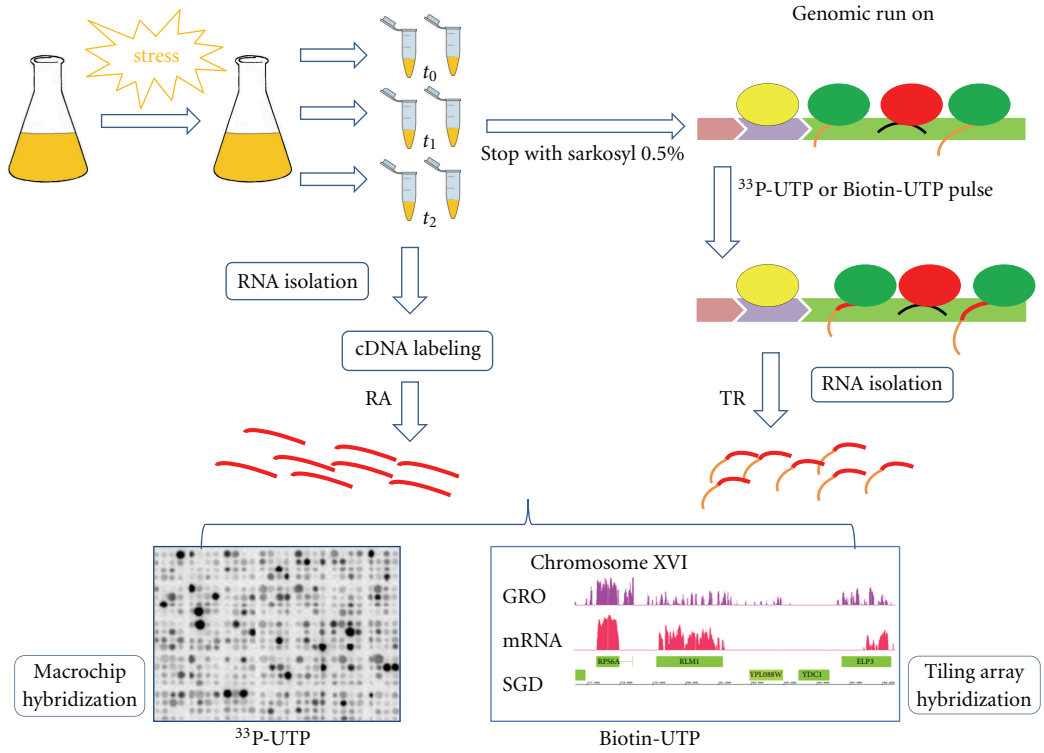


Figure 1: Outline of the GRO method for yeast cells. Cells grown to the desired condition are sampled in duplicate. One aliquot (arrow pointing rightwardly) is used for run-on labeling with either a radioactive precursor or 11-biotin-UTP. Total extracted RNA is used for macroarray or tiling microarray hybridization. The second cell's aliquot (arrow pointing downwardly) is frozen and used for RNA extraction and cDNA labeling, with either ^{33}P-dCTP or biotin allonamide triphosphate, and hybridized in a new tiling microarray (BioGRO) or in the same macroarray (GRO) previously used after stripping. This method can be used for single point determination in different strains or can be easily adapted to time-point series after a stress or drug treatment, as shown. RNA pol II molecules are shown in different colors, indicating molecules before the elongation step (yellow), elongating (green), or backtracked (red). Labeled parts of mRNAs or cDNA are drawn in red. RA = mRNA amounts, TR = transcription.

allows it to be used in many situations, such as the study of mutants, even when a metabolic fast change occurs. Indeed, 50 mL aliquots of yeast cells can be taken every two minutes if needed and processed very quickly in a few minutes. This allows to follow stress responses with a high resolution (see Figure 1; for a detailed experimental protocol; see [26]). In our protocol, another aliquot of cells is taken and frozen at the same time. This aliquot can be used for routine RNA purification (note that all RNAs are unlabeled in this case) and for subsequent standard mRNA amount determination after labeling as cDNA (Figure 1). Since both cell aliquots come from exactly the same time point of the culture, TR and RA values correspond to the average values of those parameters for a given cell population. The existence of experimentally determined TRs for all the genes of a given organism allowed us to compare the response profiles of each gene after an external change. We performed this kind of studies in the change from a glucose to a galactose medium [23] after oxidative stress [27], after osmotic stress [28] and after heat stress [29]. In all cases, genes cluster according to their response profiles by mostly following functional relatedness. This can be the direct result of the common regulation of those genes belonging to a same regulon by a transcription factor (TF). In fact, a meta-analysis done with our data by a different group showed that TR profiles were more suitable to predict functional relatedness than RA profiles [30]. The reason for this is quite obvious; nascent TR is the parameter directly affected by a TF. A change in RA can be the result of not only a change in TR, but also in mRNA stability (see Figure 2). Even the rate at which mRNA appeared in the cytoplasm (mature TR) could be less suitable for regulon finding because some posttranscriptional events, like mRNA export, can affect the mature TR profile. The effect of posttranscriptional changes blurs RA profiles because some mRNAs display different posttranscriptional behaviors. Therefore, the clustering of TR profiles seems to be the best tool to find transcriptional regulons. Although only formally demonstrated in *S. cerevisiae,* this statement seems to be reasonably extended to other organisms when methods to determine nascent TR become available. Nascent TR is also the best way to classify genes for active chromatin marks. It seems somewhat logical that the passage of RNA pol II molecules along the chromatin template is affected by particular nucleosomal organization. Because nascent TR, as

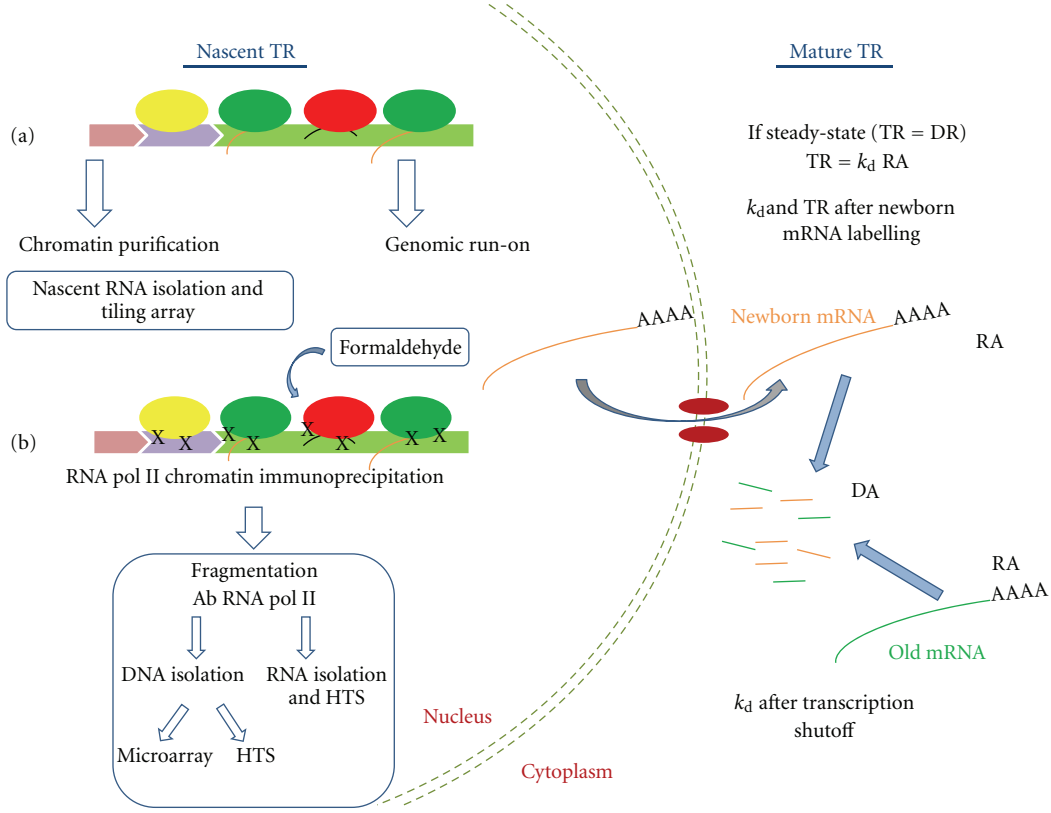


Figure 2: Outline of the different genomic techniques for TR and mRNA stability determination used in yeast. Determination of nascent TR in the nucleus is based on the detection of RNA polymerases or nascent RNAs. (a) Native chromatin can be purified and nascent RNA can be isolated. Then it is converted into cDNA and used for the tiling array analysis [32]. Alternatively, cells can be used for the GRO analysis (see Figure 1). (b) Cells can be fixed with formaldehyde and subjected to chromatin fragmentation and immunoprecipitation with RNA pol II antibodies. From this ternary complex, either DNA or RNA can be detached from polymerases and analyzed downstream. Recovered DNA is suitable for array hybridization (ChIP-Chip) [10, 33, 34] or HTS (ChIP-Seq) [35], whereas RNA can be converted into DNA and subjected to HTS [36]. The appearance of recently synthesized mRNA (newborn, orange) in the cytoplasm can be followed by thiouridine *in vivo* labeling, which is purified, quantified, and compared with the non labeled old mRNA (green). With this technique, it is possible to calculate mature TR and mRNA half-lives [37]. mRNA half-lives can also be calculated from a transcription shutoff experiment [38], from TR and RA data by assuming the steady state (k_d = TR/RA), or even for non-steady-state conditions [39]. DR: degradation rate. Other symbols are as in Figure 1.

determined by GRO, measures the actual elongation rate, what is actually affected by nucleosome positioning and readability should be a better predictor of the characteristic chromatin marks of active genes than mature TR which, as mentioned, considers other posttranscriptional steps. We have shown that this is precisely the case. When comparing the level of the different active chromatin features, such as H3-K_{36} trimethylation, or the presence of Esa1p or Gcn5p histone acetyltransferases, we found that the correlation with nascent TR (calculated by GRO) is better than when compared with the mature TR calculated from steady-state RA and mRNA stabilities [31].

The GRO protocol has allowed us to obtain a whole TR dataset for an organism for the first time [31]. Analyzing the dataset provided a number of surprises: 90% of yeast genes show TRs between 2 and 30 molecules/h, with a median of 7 molecules/h. This corresponds to 0.078 RNA pol II molecules/kb or 0.1 molecules/gene. As 25% of transcription corresponds to 5% of most transcribed genes, the distribution of RNA pol II molecules in a snapshot of an actively growing yeast cell is mostly like a desert: only 14% of genes have any actively transcribing RNA pol II molecule. Transcription onto canonical genes does not seem to be a common feature of the yeast genome in spite of the high number of RNA pol II molecules present in a cell (20000–30000) since only around 700–1400 would transcribe genes that encode proteins at a given time. One possibility is that part of these molecules is unable to transcribe, and that the amount of CTD-phosphorylated molecules (12000, according to [40]) suggests that they can be transcribing in other regions outside the canonical genes (see below), or perhaps that the mRNA molecules reaching the cytoplasm are merely a fraction of the nascent ones. The most transcribed genes are those that code for histones, which reach about 206 mRNAs/hour during the S phase. These results are rather similar to those obtained by a different technique (dynamic transcriptome analysis (DTA) in [37], see below) which measures mature TR. The

similarity between total nascent and mature TRs, however, does not mean genuine equality because the absolute units were obtained by normalization in both cases against the amount of mRNAs per cell. In any case, the similar medians and distributions, and the high correlation (r Pearson 0.63; see Figure 3) obtained, favor the quality of both datasets.

TR is not dependent on G+C content, but decreases for increasingly long genes [31]. This result was expected because of the probability of RNA pol II failure increasing with transcription unit length [41, 42]. The slope of this bias is greater for the GRO dataset than for the other TR evaluation method based on RNA pol II crosslinking to the gene (RNA pol II chip-on-ChIP: RPCC; see below). Moreover, the GRO curve shows a change in tendency for those genes whose ORFs are longer than 3 kb (see [31], Figure S1). This effect can be caused by the macroarrays used for both RPCC and GRO in which any gene whose ORF length is longer than 3 kb is represented by a probe covering only the last 1 kb downstream in the ORF, whereas the rest of the genes (<3 kb) are represented by the whole of the ORF. As the GRO method labels the elongating mRNA by extending it 200-300 nt downstream, it is expected to bias the label to the 3′ part of the genes. This would lead to an increased average label in the genes as their length shortens and the relative influence of their 3′ part increases. It also predicts a sudden increase for the genes represented by probes that cover only their last 1 kb of the ORF. With this in mind, we conclude that there is a general bias in the GRO which increases the calculated TR inversely with the ORF (i.e., probe) length. Additionally, there is an artifact for ORFs that are longer than 3 kb due to the use of 1 kb probes from their 3′ end. This gene-/probe-dependent effect was corrected by using the RPCC data for the lowest normalization of the GRO data. Our TR dataset has, therefore, been corrected for this artifact.

Another potential artifact of GRO (and other nascent TR methods) is the potential effect of cryptic transcription. Since GRO labels any elongating RNA polymerase, those RNA pol II molecules, that elongate anywhere inside the genome regions contained in probes (ORFs), are labeled regardless of making "canonical" transcripts or cryptic transcripts. In the original GRO method [23], the macroarrays used contain dsDNA probes. Because of this, both sense and antisense transcriptions will be summed. The existence of a vast number of cryptic transcripts has been demonstrated in many organisms, including yeast (see [42], for a recent review) in which two types have been defined: cryptic unstable transcripts (CUTs), which are only detectable in the absence of nuclear exosome activity and stable uncharacterized transcripts (SUTs) [14]. Some authors argue that cryptic transcription can be responsible for the differences observed in the genes' response to stress situations when comparing mRNA data and TR data [15, 43]. This contrasts with the recent unveiling of experimental evidence diminishing the possible quantitative contribution of antisense transcripts to the RNA pool when compared to their stable sense transcript counterparts in the bidirectional promoter's context [36, 42]. Moreover, we analyzed the different cryptic transcription datasets published and we observed that they are quite different, with very little overlap, and that the technique used to find them vastly affects the type (sense or antisense) of the cryptic transcripts found (García-Martínez et al., submitted). Thus, it seems that most yeast genes have cryptic transcripts, but mainly in a very low proportion (discussed in [42]). Thus, although we agree that cryptic transcription is a real contributor of nascent TR data, we believe that nascent TR reflects mainly "canonical transcription" for most genes.

3. Alternative Methods for Evaluating Transcription Rates and mRNA Degradation Rates

Thus the GRO technique allowed the comparison of the rate at which each mRNA is produced and its amount in the cell. This idea has been used quite recently by other authors, who have followed different TR evaluation methods (see below) both in yeast [37] and higher eukaryotic cells [44]. In all these studies, mRNA stabilities have also been calculated. The possibility of calculating mRNA half-lives indirectly from TR and RA is based on the aforementioned equilibrium between TR and DRs. If they are equal, then RA is constant over time. This is known as a steady-state situation, which mainly occurs for most mRNAs. For instance, we demonstrated that when yeast is grown in a flask (a batch culture), most of the exponential growth phase maintains the steady state for most mRNAs [21]. Although some mRNAs change slightly after several hours, the steady-state condition can be a good approximation to describe the mRNA pool. It is likely that the steady state is also true for the stationary phase or in chemostat cultures [45]. As DR follows a first-order kinetic law, it is proportional to RA and to a degradation constant (k_d). The GRO protocol calculates RA and TR for all the yeast genes. Because DR = TR, it is also possible to calculate k_d for them all. k_d has time^{-1} units and the reverse meaning of the half-life (k_d = ln2/half-life). When no steady-state situation occurs, again the chemical kinetic laws can be used to develop an equation in which the successive TR and RA time points are employed to infer k_d using the simplification, whose changes in both are lineal between the successive time points [20]. Although the mathematical computation increases the experimental associated error, this approach has allowed us to calculate k_d variations for those groups of genes with common RA and TR profiles during fast stress responses [27–29]. In all the individual genes tested, the k_d calculated at different times during the stress response qualitatively coincided with the experimentally determined one using Tet-off promoters.

In the last few years, other techniques apart from GRO have been developed to study TRs and mRNA stabilities in yeast and in other organisms at the genomic level. In all cases, they use recently developed, higher resolution methods, such as tiling arrays or high-throughput sequencing (HTS), which provide deeper insights into the transcription process than when using classical DNA arrays. For TR determination, most methods focus on nascent TR (Figure 2). The classical approach to unveil the dynamics of the transcriptional process at the TR level relies on the generation of RNA pol II density landscapes to precisely map where RNA pol sit in the whole genome, regardless of transcriptional

states (active, paused, backtracked, etc.; see Figure 2). This has been achieved with chromatin immunoprecipitation methods coupled with microarray analysis, ChIP-chip [33, 37, 43] or HTS, ChIP-Seq [35]. With these high-throughput methods, a plethora of different occupancy profiles for RNA polymerase II and its different phosphorylation forms are now publicly accessible. However, ChIP techniques cannot circumvent the fact that the presence of polymerase in a region should not be directly assumed as actual transcription because of there being nonelongating polymerases and because ChIP-associated techniques cannot discriminate the sense/antisense transcripts (see above). A recent variation of ChIP techniques has been able to partially skirt this drawback (Figure 2). By isolating and deep sequencing the nascent transcript associated with immunoprecipitated RNA polymerases (NET-Seq, [36]), both problems are avoided. Nascent mRNAs have been alternatively purified by chromatin fractionation approaches [32], thus providing a more direct measure of TR. The nonradioactive variants of GRO, coupled with tiling array analysis (BioGRO, Jordán-Pla et al., unpublished), or HTS (GRO-Seq), have been successfully applied to dissect the regulatory circuitry of yeast and human [25] cells. These high-resolution GRO techniques are beneficial because they can discriminate between active and nonactive transcription elongation states and can also detect any type of noncoding unstable transcript RNA polymerases generated in both the sense and antisense orientations, for which accumulating evidence shows that they play a crucial role in the regulation of gene expression [14], thus broadening our knowledge and understanding of gene regulation dynamics.

Mature TRs determine the rate of appearance of newborn mRNA in the cytoplasm (Figure 2, right). *In vivo* metabolic labeling of transcripts, with uracil or uridine analogs pulses, subsequent mRNA isolation, fractionation between labeled (newborn) and nonlabeled (old) mRNA and analysis, is an essentially nonperturbing system which provides us with a way of measuring mature TR directly [46]. The recently developed DTA methodology in yeast [37] and mammalian cells [44] focuses on newborn mature polyadenylated transcripts and uses a metabolic time lapse of variable extent (6 minutes in yeast). DTA technology is currently the only technology available that is able to measure mature TR in yeast experimentally. It can also determine mRNA half-lives at the same time. The technique is, however, time consuming and assumes that TR and DR are constant during the time lapse used. It also leaves an important bulk of non coding transcripts outside the frame [14]. Mature TR can also be calculated from RA and mRNA stability datasets indirectly (indirect TR, TRi = k_d RA) by assuming a steady state [47]. Alternatively, mRNA half-lives (k_d) can be calculated indirectly from experimental mature TR and RA datasets using the same equation or directly using the previously commented transcription shutoff methods [17, 19, 21].

The comparison of all existing yeast TR datasets [31, 33, 37, 47] (Jordán-Pla et al., unpublished) with each other and with a standardized RA dataset [31] has revealed that they all correlate quite well (Figure 3). Those corresponding to nascent TR correlate better to each other. This is also true for those corresponding to the mature TR (DTA, [37]; TRi, [47]). The last ones better correlate with the RA dataset. These results are logical because nascent mRNAs should be processed and exported to the cytoplasm where mature TR is measured. Moreover, nascent TR can be affected by cryptic transcription and mature TR cannot. The specific distribution of ribosomal protein (RP) genes (blue dots) is biased in all the comparisons shown. The meaning of this behavior is commented below.

4. Ribosomal Protein Genes: A Special Case of RNA Pol II Transcription

Ribosomal proteins in yeast are coded by a set of 137 genes. They are, perhaps, the most statistically significant group that clusters together in many of the genomic analyses done in yeast [23, 34]. This can be due to the fact that this group is more coordinated and/or more numerous than other regulons which are less coordinated (e.g., ribosome biogenesis regulon, RiBi, and ~200 genes), less abundant (Gal regulon, 7 genes), or both. Because translation is the most costly synthesis process for the cell [48], and as the ribosome is composed of stoichiometric amounts of RP, both the control and coordination of these genes is very strict. In yeast, they are mainly regulated at the transcriptional level [48], involving several TFs (see [49, 50], for a detailed updated review). RP mRNAs are among the most abundant in the cell. They have also been traditionally considered the most transcribed ones, representing as much as 50% of transcribing events [48]. Our results, however, quantify the overall TR of those genes as only 8.5% of total RNA pol II TR [31]. As previous estimations were based on indirect evaluations of RP transcription rates [48], we considered it merely a miscalculation. However, we now think that the main mistake seems to lie in the use of very few examples of RP mRNA half-lives at the time of the proposal. We recalculate now the total TR for RP genes (indirect TR or TRi) using all the available datasets of mRNA half-lives [17, 19] and RA datasets [31, 37]. In all cases, these genes represent less than 27% of total RNA pol II transcription in yeast. Therefore, the original 50% was clearly miscalculated. On the other hand, the plots of direct estimations of TR versus indirect ones (Figures 3(d) and 3(f)) show an overestimation of TRi values. The contribution of RP TR to the total can also be currently calculated from mature TR DTA data [37] to represent 16% of total TR. The differences with the previous 27% may be due to either a mistake in TRi or to a specific bias in DTA and GRO techniques for these genes. However, we think that the main reason is that transcription shutoff methods underestimate RP mRNA half-lives as they can provoke destabilization caused by stress. We have shown that this phenomenon does occur in all the stress responses we have analyzed to date (see below). In fact, RP mRNAs are 50% more stable than average in the DTA dataset [37] instead of merely coming close to it in transcription shutoff experiments [17, 19], or being less stable than average, as stated in Warner's review [48].

On the other hand, the bias observed in the RP nascent TRs measured by GRO, as mentioned above, may provide

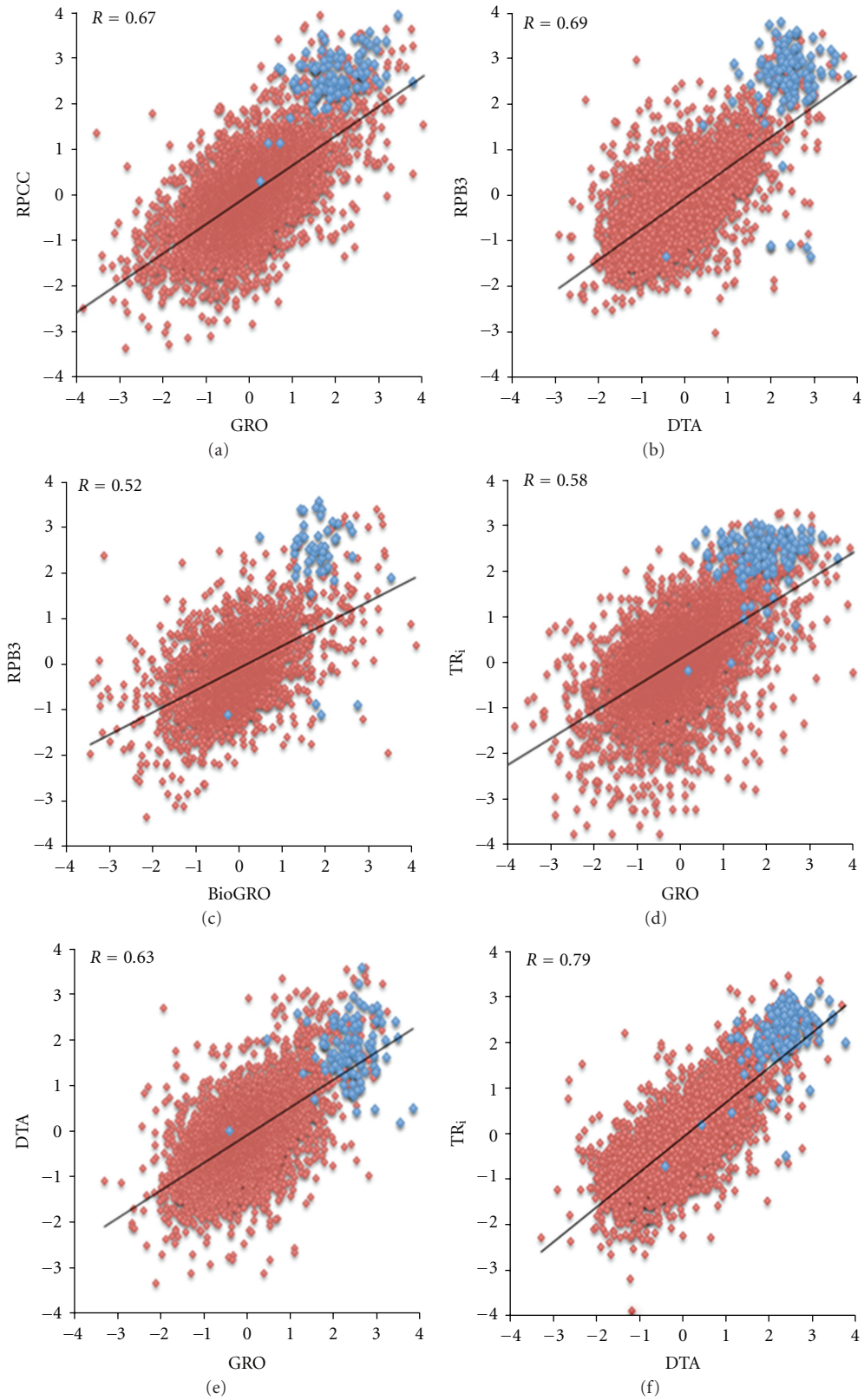


Figure 3: Comparison of the TR datasets obtained by different protocols. Available TR datasets were converted to $\log_2$ values and standardized by the Z-score. The tendency line for a linear correlation is shown. The *r* Pearson coefficient for the correlation is provided in the top left corner of each plot. In all cases, the RP genes (blue dots) distribution differs from the global gene distribution (red dots) according to the Wilcoxon text with a *P* value of < 0.0001. GRO is the radioactive nascent TR dataset. BioGRO is an unpublished nascent TR dataset obtained with Biotin-UTP and tiling arrays (Jordán-Pla et al, unpublished). RPCC is the RNA pol II chromatin immunoprecipitation performed by us [31]. RPB3 is the RNA pol II chromatin immunoprecipitation performed by Mayer et al. [33] using an anti-RPB3 antibody.

relevant insights into the elongation mechanism of these genes. We have seen that RP genes show a high percentage of nondetectable RNA pol II by GRO [34]. In exponential growth in glucose, they have about 36% more RNA pol II molecules present in the coding region of the RP, which are unable to do a run-on, more than the genome average (see Figures 3(a)–3(c)). We concluded that this reflects how the transcription of these genes involves a higher percentage of non elongating (probably backtracked) RNA pol II [31, 34]. This bias toward nonactive RNA pol II is also seen when comparing the results from mature TR (DTA technique, [37], Figure 3(e)) and RNA pol II immunoprecipitation [33], which reinforces our previous calculations. In this case, excess lies at about 31%, meaning that in RP genes, there is more RNA pol II present in those genes that do not produce mature mRNA in the cytoplasm than for the average gene. From the difference between the 36% and the 31% excess when using GRO or DTA in the comparison, or by directly comparing the DTA and GRO datasets (see Figure 3(e)), we conclude that between 15% and 40% of the RNA pol II molecules which are not labeled during the run-on are unable to resume elongation and to produce mature mRNA. This is perhaps because they are trapped in some step of the backtracking process [51]. The rest of the non labeled molecules (60–85%) would thus represent the efficiency of the backtracking process in recovering paused RNA pol II molecules, at least, for this group of genes.

What is the reason for the special abundance of non elongating RNA pol II molecules over RP genes? We hypothesized that a special chromatin structure of those genes could be responsible. In fact, TF Rap1p has a known chromatin-organizing activity (reviewed in [49]), which is predicted or demonstrated to activate 127 of the RP genes. However, this factor plays another, apparently opposed, role in telomeres: it organizes repressive subtelomeric chromatin [49]. Some years ago, it was shown that repressor activity resides in the C-terminal part of the protein, the *sil* domain, that a mutated version of Rap1 lacking it, *rap1-sil,* derepresses the genes within 50 kb from the telomere, and that it has increased levels of RP mRNAs [52]. We found that the excess of non elongating RNA pol II molecules disappeared in the *rap1-sil* mutant, which also occurs in a *tpk* single mutant [34]. Tpk1,2,3 are the alternative catalytic subunits of protein kinase A (PKA), which controls signaling from glucose [53]. Our model reveals that Rap1 not only recruits RNA pol II to the RP promoters, but it also organizes a partially repressed chromatin, which hinders RNA pol elongation, thus leading to pauses and arrest. This difficulty occurs at the beginning of the transcribed region [34] and could be an additional control mechanism for the regulation of these important genes. This mechanism would act only at the highest transcription rates (during exponential growth in glucose-rich media) as a way to accumulate truly elongating RNA pol II molecules which, due to this mechanism, slow down in the first part of the gene. After passing the control, these molecules increase their velocity and become sufficiently separated to avoid collisions. A similar elongation control has been proposed for ribosome translation [54] and has been mathematically demonstrated to increase the speed of RNA polymerases and to reduce noise (cell-to-cell variation, [55]), which is a typical, convenient feature of RP genes [56].

5. The Relative Importance of TR and DR in Controlling mRNA Amounts

At any time, RA is the result of the balance of TR and DR. When RA has to be changed, it is theoretically possible to act only on one side of the equilibrium, or on both sides. Traditionally, most studies on RA conducted at either the single gene level or the genomic level implicitly assumed that changes were due to only TR changes. A given gene is induced because a TF binds a promoter and attracts RNA pol II to transcribe it (by increasing TR) which, in turn, increases RA. Thus, RA profiles were considered to be a mere consequence of TR profiles. A few years ago, some authors started to pay attention to the potential effect of DR in the gene expression. Theoretically, the same effect on RA can be obtained by a change in DR rather than by an inverse change in TR. The mRNA half-lives determined by transcription shutoff methods were found to be very different between mRNAs and organisms [16]. More importantly, they were discovered to change for a given mRNA in different physiological situations [16]. When genomics strategies appeared [17], it was seen that the mRNAs belonging to the same pathways or functions tend to have similar half-lives, suggesting that regulons also exist at the mRNA stability level [57]. The trans factors acting in these posttranscriptional regulons were found to be mainly RNA-binding proteins (RBPs), which were relatively selective in the mRNA population because of their sequence specificity [58, 59].

As the GRO technique is able to indirectly determine mRNA stabilities in steady-state situations, as explained above, our first experiment [23] enabled us to reveal that those genes belonging to functionally related groups behave coordinately in DR. This was the first formal demonstration of the existence of post-transcriptional regulons. Similar studies in other organisms arrived at comparable conclusions, although not a whole genome-scale level (e.g., [38, 60]). In that first experiment, the times selected after changing cells from a glucose to a galactose medium were separated by hours, and the steady-state conditions can apply to each of them. Fast responses, typical of stress situations or sudden changes, did not meet the steady-state conditions. For such cases, we developed an algorithm, as previously described. With it, we have been able to determine approximate k_d profiles in response to different stresses for most genes and to verify the hypothesis of the influence of DR changes on RA profiles. Our studies [27–29] and those of others [61, 62] reveal that many genes have undergone changes in DR during stress responses. Many other genes, however, do not change their mRNA stability substantially. In line with this, interesting differences have been noted between various stresses [61], which probably depend on stress intensity. A good number of genes respond by slightly decreasing their RA level transiently to recover after several minutes. These genes tend to have flat k_d profiles, meaning they result from a mere transient decrease in TR. Other genes

that respond to stress by lowering their RA are the RP and RiBi genes. In all cases, we have seen that the mRNAs of all these genes showed a transient destabilization, which can be accompanied by different degrees of TR decrease. These results indicate that DR can be used to reduce mRNA levels and, as explained later, to also speed up this reduction. It is interesting to note that some genes, which do not exhibit coordinated behavior at the TR level in some instances, actually display coordinated behavior with mRNA stability. This is the case of mitochondrial RP genes which cluster in the mRNA stability analysis, but not in the TR analysis, during the shift from a glucose to a galactose medium [23]. These genes do not present obvious regulatory elements in their promoters, but a Puf3 element in their 3′UTR has been demonstrated to coordinate their stability after changes in respiratory behavior [63]. Thus, our study was able to show that, for some specific gene categories, coordination takes place at the posttranscriptional level and that DR is the main player in shaping a response.

Those genes that positively respond to stress by increasing TR also show interesting changes in DR. For instance after osmotic stress, many genes present increased RA by increasing their TR and decreasing their DR (increase in mRNA stability) for several minutes [28]. A similar observation was reported by other authors using a different method to determine mRNA stabilities [61, 62] in the oxidative stress response. Thus after some minutes of osmotic stress, these genes reverse the change in DR in parallel to transcription shut-off [28]. This effect has been interpreted as the DR change which precedes changes in RA [62]. Other authors postulate that the changes in DR in both yeast [61] and mammalian [44] cells contribute to sharp response peaks. In mammalian cells with substantially longer mRNA half-lives, the contribution of DR changes to speed up the increases and decreases in mRNA levels is probably more important than in yeast. Nevertheless, the short yeast generation time (~100 minutes) and the need for faster, more economically adjusted responses to environmental situations than for mammalian cells mean that it is also necessary to use changes in mRNA stability to sharpen RA peaks in order to restrict energy expenditure while they take place [20]. The width of the TR response peaks seen in all the stress-activated or repressed genes is about 15–30 minutes, which is somewhat narrower than the RA peaks in our experiments and in those performed by others [64]. It has to be considered that the results obtained in experiments using about 10^9cells are the average of all the possible individual states for a given cell. The difference in RA between individual cells is known as "transcriptional noise". It has been determined to considerably differ for two kinds of genes: the TATA-less genes, with nucleosome-depleted constitutive promoters, for example, RP genes [49], which are less noisy [56] than the inducible genes with a TATA box at their SAGA-dependent promoters, known to be transcribed in "bursts" of several consecutive mRNAs separated in time [65], and thus in a much noisier manner [56]. Stress-activated genes belong to this last group. Thus, it is conceivable that the relatively sharp TR peaks result from much sharper peaks in individual cells in which a single (or very few) transcription burst occurs at their single gene locus during a stress response. At the molecular level, this would correspond to a switch from off to on in the promoter chromatin structure. The consequence of a sharp TR peak in an RA peak would be sharp only if mRNA has very low stability (see [20]). Therefore for mRNAs with half-lives longer than 10–15 min, it is necessary to unstabilize them if a fast return to the original mRNA level is needed. At the molecular level, this also assumes a binary switch in a given cell represented by a change in the affinity of an RBP to its cognate 3′-UTR mRNA element [59]. In this case, however, the existence of multiple mRNA molecules per cell probably makes the "degradation burst" less acute.

All these results suggest the importance of DR in controlling mRNA levels during transcriptional responses. Reciprocally important quantitative analyses of TR and DR in shaping RA have shown that despite being theoretically equivalents, TR and DR do not seem to play the same role in determining the amount of mRNA in stress responses or at steady-state conditions. It is interesting to note that when comparing the different datasets for TR, RA, or mRNA half-lives, in all cases TR and RA show a positive significant correlation, which mRNA half-lives never do (e.g., stable mRNAs are not the most abundant, and unstable mRNAs are not the least abundant); indeed, they even show slightly negative correlations with RA and TR in yeast [37, 39] or mammalian [66] cells. When measuring the number of genes whose response profiles are significantly affected by DR changes, we [23, 27, 29] and other authors [44, 66], found a large majority of genes in which TR changes are the main determinant of RA profiles. Thus, it seems that DR is not used for the quantitative control of most mRNA levels in the majority of situations, but for classifying them into rapid or less-rapid response genes according to the stability of their mRNA [20, 44]. Many mRNAs have a relatively constant DR. Nevertheless, some special gene categories, such as stress-induced or RP genes in stress responses in yeast (see above), mitochondria-related genes during metabolic reprogramming from glucose to galactose in yeast [23], or inflammatory and immune signaling genes in dendritic cells [44], have a highly variable DR because of the cis elements in their 3′-UTRs targets of specialized RBPs [29, 58, 59]. In some instances in which growth stops, like the stationary phase [67], after a change from a glucose to a galactose medium in yeast or after strong stress, a general change in DR is observed [23, 68]. This situation probably relates more to a general change in DR machinery and/or p-body organization and can overlap with particular responses.

The corollary of this scheme is that each particular group of genes has a TR through their promoter organization, and a DR through their particular 3′-UTR sequences, which are subjected to transcriptional (regulons) and posttranscriptional (posttranscriptional regulons) regulation [58]. Both should evolve coordinately to achieve a common expression strategy (CES) for the group of genes. We analyzed the existence of CES not only for transcriptional regulation, but also for translational regulation [39]. We found that our hypothesis was true: each analyzed functional gene category had a statistically significant CES for both transcription and translation. Genes whose protein products belong to

large stable stoichiometric complexes, such as the cytosolic ribosome, the nucleosome, the proteasome, and many others, present characteristic profiles with relative unstable mRNAs and proteins, as well as relatively high transcription and translation rates. This probably reflects the need for fast changes in both mRNA and proteins in some instances to coordinate the amounts of subunits during a cell's life. Energy pathways genes, however, have more equilibrated profiles. A similar study in mammalian cells was done to find that CES also exists, but not quite the same as in yeast. For instance, in this case, RPs show very stable mRNAs, and proteins are very much like energy pathways genes [66]. These analyses reveal that, as expected, translation (translation rate and protein stability) is also a layer for gene regulation in all eukaryotes. From our study of yeast CES profiles, we conclude that the transcriptional layer is quantitatively more important for gene regulation than the translational one. However, Schwanhäusser et al. [66] conclude that the layer of translation is more important for determining the abundance of proteins in mouse fibroblasts. Hence, it seems that every life style needs particular gene expression strategies.

6. Future Trends

The similarities and differences observed between yeast and mammalian cells in organizing gene regulation indicate that the multiple layers used by eukaryotes provide a flexible network upon which every gene class can find its best strategy. Furthermore, this fitting can evolve according to the organism's requirements.

Variability in the single-cell gene expression in both microorganisms and tissue cells indicates that genomic transcription measurements should be complemented by techniques designed to quantify the gene expression at the single cell level [65, 69] and on the genomic scale.

Given the adjustment required between TR and DR (and also translation), it seems that cells have to contain mechanisms that allow cross-talk between mRNA transcription and degradation at both the single mRNA level (by coordinating transcription from a promoter with the fate of the mRNAs producing that) and the general level to coordinate the regulons and posttranscriptional regulons composed of several different gene species. The recent publications by M. Choder's group [70] and others [71] reveal that such mechanisms do exist.

Finally, it is worth mentioning that most of the results reviewed herein would not have been possible without comparing large datasets because the general tendencies and the differences between genes categories are only significant if we look at the whole genome at the same time. Therefore, we consider that although anyone is free to believe that high throughput genomic technologies may be of "low input", nowadays it comes over quite clearly that they actually provide some output to molecular biology.

Acknowledgments

The authors are grateful to the members of the GFL laboratory for discussion and support and to Drs. Vicent Pelechano and Paula Alepuz for critically reviewing the paper. D. Medina and A. Jordán-Pla are supported by Santiago Grisolía (Generalitat Valenciana) and FPI (Spanish MCINN) fellowships. José E. Pére-Ortín is supported by Grants from the Spanish MCINN (BFU2010-21975-C03-01) and from the Regional Valencian Government (Generalitat Valenciana-PROMETEO 2011/088).

References

[1] S. Brenner, "Detective Rummage investigates," *The Scientist*, vol. 16, no. 16, p. 15, 2002.

[2] M. Snyder and J. E. G. Gallagher, "Systems biology from a yeast omics perspective," *FEBS Letters*, vol. 583, no. 24, pp. 3895–3899, 2009.

[3] J. Beckwith, "The operon as paradigm: normal science and the beginning of biological complexity," *Journal Of Molecular Biology*, vol. 409, no. 1, pp. 7–13, 2011.

[4] A. Goffeau, B. G. Barrell, H. Bussey et al., "Life with 6000 genes," *Science*, vol. 274, no. 5287, pp. 546–567, 1996.

[5] S. R. Engel, R. Balakrishnan, G. Binkley et al., "Saccharomyces genome database provides mutant phenotype data," *Nucleic Acids Research*, vol. 38, supplement1, pp. D433–D436, 2009.

[6] U. Güldener, M. Münsterkötter, G. Kastenmüller et al., "CYGD: the comprehensive yeast genome database," *Nucleic Acids Research*, vol. 33, supplement 1, pp. D364–D368, 2005.

[7] J. E. Pérez-Ortín, J. García-Martínez, and T. M. Alberola, "DNA chips for yeast biotechnology. The case of wine yeasts," *Journal of Biotechnology*, vol. 98, no. 2-3, pp. 227–241, 2002.

[8] J. L. DeRisi, V. R. Iyer, and P. O. Brown, "Exploring the metabolic and genetic control of gene expression on a genomic scale," *Science*, vol. 278, no. 5338, pp. 680–686, 1997.

[9] N. C. Hauser, M. Vingron, M. Scheideler et al., "Transcriptional profiling on all open reading frames of Saccharomyces cerevisiae," *Yeast*, vol. 14, no. 13, pp. 1209–1221, 1998.

[10] L. David, W. Huber, M. Granovskaia et al., "A high-resolution map of transcription in the yeast genome," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 14, pp. 5320–5325, 2006.

[11] C. T. Harbison, D. B. Gordon, T. I. Lee et al., "Transcriptional regulatory code of a eukaryotic genome," *Nature*, vol. 430, no. 7004, pp. 99–104, 2004.

[12] S. Brenner, "Sequences and consequences," *Philosophical Transactions of the Royal Society B*, vol. 365, no. 1537, pp. 207–212, 2010.

[13] F. Wyers, M. Rougemaille, G. Badis et al., "Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase," *Cell*, vol. 121, no. 5, pp. 725–737, 2005.

[14] Z. Xu, W. Wei, J. Gagneur et al., "Bidirectional promoters generate pervasive transcription in yeast," *Nature*, vol. 457, no. 7232, pp. 1033–1037, 2009.

[15] M. Yassour, J. Pfiffner, J. Z. Levin et al., "Strand-specific RNA sequencing reveals extensive regulated long antisense transcripts that are conserved across yeast species," *Genome Biology*, 2010, article R87.

[16] R. Parker and H. Song, "The enzymes and control of eukaryotic mRNA turnover," *Nature Structural and Molecular Biology*, vol. 11, no. 2, pp. 121–127, 2004.

[17] Y. Wang, C. L. Liu, J. D. Storey, R. J. Tibshirani, D. Herschlag, and P. O. Brown, "Precision and functional specificity in mRNA decay," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 9, pp. 5860–5865, 2002.

[18] R. Parker, D. Herrick, S. W. Peltz, and A. Jacobson, "Measurement of mRNA decay rates in *Saccharomyces cerevisiae*," *Methods in Enzymology*, vol. 194, pp. 415–423, 1991.

[19] J. Grigull, S. Mnaimneh, J. Pootoolal, M. D. Robinson, and T. R. Hughes, "Genome-wide analysis of mRNA stability using transcription inhibitors and microarrays reveals posttranscriptional control of ribosome biogenesis factors," *Molecular and Cellular Biology*, vol. 24, no. 12, pp. 5534–5547, 2004.

[20] J. E. Pérez-Ortín, P. M. Alepuz, and J. Moreno, "Genomics and gene transcription kinetics in yeast," *Trends in Genetics*, vol. 23, no. 5, pp. 250–257, 2007.

[21] V. Pelechano and J. E. Pérez-Ortín, "There is a steady-state transcriptome in exponentially growing yeast cells," *Yeast*, vol. 27, no. 7, pp. 413–422, 2010.

[22] K. Hirayoshi and J. T. Lis, "Nuclear run-on assays: assessing transcription by measuring density of engaged RNA polymerases," *Methods in Enzymology*, vol. 304, pp. 351–362, 1999.

[23] J. García-Martínez, A. Aranda, and J. E. Pérez-Ortín, "Genomic run-on evaluates transcription rates for all yeast genes and identifies gene regulatory mechanisms," *Molecular Cell*, vol. 15, no. 2, pp. 303–313, 2004.

[24] T. M. Alberola, J. García-Martínez, O. Antúnez et al., "A new set of DNA macrochips for the yeast Saccharomyces cerevisiae: features and uses," *International Microbiology*, vol. 7, no. 3, pp. 199–206, 2004.

[25] L. J. Core, J. J. Waterfall, and J. T. Lis, "Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters," *Science*, vol. 322, no. 5909, pp. 1845–1848, 2008.

[26] J. García-Martínez, V. Pelechano, and J. E. Pérez-Ortín, "Genomic-wide methods to evaluate transcription rates in yeast," *Methods in Molecular Biology*, vol. 734, pp. 25–44, 2011.

[27] M. M. Molina-Navarro, L. Castells-Roca, G. Bellí et al., "Comprehensive transcriptional analysis of the oxidative response in yeast," *Journal of Biological Chemistry*, vol. 283, no. 26, pp. 17908–17918, 2008.

[28] L. Romero-Santacreu, J. Moreno, J. E. Pérez-Ortín, and P. Alepuz, "Specific and global regulation of mRNA stability during osmotic stress in Saccharomyces cerevisiae," *RNA*, vol. 15, no. 6, pp. 1110–1120, 2009.

[29] L. Castells-Roca, J. García-Martínez, J. Moreno, E. Herrero, G. Belli, and J. E. Perez-Ortin, "Heat shock response in yeast involves changes in both transcription rates and mRNA stabilities," *PLoS One*, vol. 6, Article ID e17272, 2011.

[30] B. Hayles, S. Yellaboina, and D. Wang, "Comparing transcription rate and mRNA abundance as parameters for biochemical pathway and network analysis," *PLoS One*, vol. 5, Article ID e9908, 2010.

[31] V. Pelechano, S. Chávez, and J. E. Pérez-Ortín, "A complete set of nascent transcription rates for yeast genes," *PLoS One*, vol. 5, no. 11, Article ID e15442, 2010.

[32] F. Carrillo-Oesterreich, S. Preibisch, and K. M. Neugebauer, "Global analysis of nascent rna reveals transcriptional pausing in terminal exons," *Molecular Cell*, vol. 40, no. 4, pp. 571–581, 2010.

[33] 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.

[34] V. Pelechano, S. Gimeno-González, A. Rodríguez-Gil, J. García-Martínez, J. E. Pérez-Ortín, and S. Chávez, "Regulon-specific control of transcription elongation across the yeast genome," *PLoS Genetics*, vol. 5, no. 8, Article ID e1000614, 2009.

[35] P. Lefrancois, G. M. Euskirchen, R. K. Auerbach et al., "Efficient yeast ChIP-Seq using multiplex short-read DNA sequencing," *BMC Genomics*, vol. 10, article 37, 2009.

[36] L. S. Churchman and J. S. Weissman, "Nascent transcript sequencing visualizes transcription at nucleotide resolution," *Nature*, vol. 469, pp. 368–373, 2011.

[37] C. Miller, B. Schwalb, K. Maier et al., "Dynamic transcriptome analysis measures rates of mRNA synthesis and decay in yeast," *Molecular Systems Biology*, vol. 7, article 458, 2011.

[38] J. Fan, X. Yang, W. Wang, W. H. Wood, K. G. Becker, and M. Gorospe, "Global analysis of stress-regulated mRNA turnover by using cDNA arrays," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 16, pp. 10611–10616, 2002.

[39] J. García-Martínez, F. González-Candelas, and J. E. Pérez-Ortín, "Common gene expression strategies revealed by genome-wide analysis in yeast," *Genome Biology*, vol. 8, no. 10, article R222, 2007.

[40] K. Struhl, "Transcriptional noise and the fidelity of initiation by RNA polymerase II," *Nature Structural and Molecular Biology*, vol. 14, no. 2, pp. 103–105, 2007.

[41] A. Marín, M. Gallardo, Y. Kato et al., "Relationship between G+C content, ORF-length and mRNA concentration in *Saccharomyces cerevisiae*," *Yeast*, vol. 20, no. 8, pp. 703–711, 2003.

[42] W. Wei, V. Pelechano, A. I. Järvelin, and L. M. Steinmetz, "Functional consequences of bidirectional promoters," *Trends in Genetics*, vol. 27, no. 7, pp. 267–276, 2011.

[43] T. S. Kim, C. L. Liu, M. Yassour et al., "RNA Polymerase mapping during stress responses reveals widespread nonproductive transcription in yeast," *Genome Biology*, 2010, article R75.

[44] M. Rabani, J. Z. Levin, L. Fan et al., "Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells," *Nature Biotechnology*, vol. 29, no. 5, pp. 436–442, 2011.

[45] E. M. Airoldi, C. Huttenhower, D. Gresham et al., "Predicting cellular growth from gene expression signatures," *PLoS Computational Biology*, vol. 5, no. 1, Article ID e1000257, 2009.

[46] M. Kenzelmann, S. Maertens, M. Hergenhahn et al., "Microarray analysis of newly synthesized RNA in cells and animals," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 15, pp. 6164–6169, 2007.

[47] F. C. P. Holstege, E. G. Jennings, J. J. Wyrick et al., "Dissecting the regulatory circuitry of a eukaryotic genome," *Cell*, vol. 95, no. 5, pp. 717–728, 1998.

[48] J. R. Warner, "The economics of ribosome biosynthesis in yeast," *Trends in Biochemical Sciences*, vol. 24, no. 11, pp. 437–440, 1999.

[49] M. C. Bosio, R. Negri, and G. Dieci, "Promoter architectures in the yeast ribosomal expression program," *Transcription*, vol. 2, no. 2, pp. 71–77, 2011.

[50] H. Lempiäinen and D. Shore, "Growth control and ribosome biogenesis," *Current Opinion in Cell Biology*, vol. 21, no. 6, pp. 855–863, 2009.

[51] D. Wang, D. A. Bushnell, X. Huang, K. D. Westover, M. Levitt, and R. D. Kornberg, "Structural basis of transcription: backtracked RNA Polymerase II at 3.4 angstrom Resolution," *Science*, vol. 324, no. 5931, pp. 1203–1206, 2009.

[52] I. R. Graham, R. A. Haw, K. G. Spink, K. A. Halden, and A. Chambers, "In vivo analysis of functional regions within yeast Rap1p," *Molecular and Cellular Biology*, vol. 19, no. 11, pp. 7481–7490, 1999.

[53] P. Jorgensen, I. Rupeš, J. R. Sharom, L. Schneper, J. R. Broach, and M. Tyers, "A dynamic transcriptional network

communicates growth potential to ribosome synthesis and critical cell size," *Genes and Development*, vol. 18, no. 20, pp. 2491–2505, 2004.

[54] K. Fredrick and M. Ibba, "How the sequence of a gene can tune its translation," *Cell*, vol. 141, no. 2, pp. 227–229, 2010.

[55] A. N. Boettiger, P. L. Ralph, and S. N. Evans, "Transcriptional regulation: effects of promoter proximal pausing on speed, synchrony and reliability," *PLoS Computational Biology*, vol. 7, no. 5, Article ID e1001136, 2011.

[56] J. R. S. Newman, S. Ghaemmaghami, J. Ihmels et al., "Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise," *Nature*, vol. 441, no. 7095, pp. 840–846, 2006.

[57] J. D. Keene, "RNA regulons: coordination of post-transcriptional events," *Nature Reviews Genetics*, vol. 8, no. 7, pp. 533–543, 2007.

[58] A. P. Gerber, D. Herschlag, and P. O. Brown, "Extensive association of functionally and cytotopically related mRNAs with Puf family RNA-binding proteins in yeast," *PLoS Biology*, vol. 2, no. 3, article e79, 2004.

[59] D. J. Hogan, D. P. Riordan, A. P. Gerber, D. Herschlag, and P. O. Brown, "Diverse RNA-binding proteins interact with functionally related sets of RNAs, suggesting an extensive regulatory system," *PLoS Biology*, vol. 6, no. 10, article e255, 2008.

[60] S. A. Tenenbaum, C. C. Carson, U. Atasoy, and J. D. Keene, "Genome-wide regulatory analysis using en masse nuclear run-ons and ribonomic profiling with autoimmune sera," *Gene*, vol. 317, no. 1-2, pp. 79–87, 2003.

[61] O. Shalem, O. Dahan, M. Levo et al., "Transient transcriptional responses to stress are generated by opposing effects of mRNA production and degradation," *Molecular Systems Biology*, vol. 4, article 4, 2008.

[62] C. Molin, A. Jauhiainen, J. Warringer, O. Nerman, and P. E. R. Sunnerhagen, "mRNA stability changes precede changes in steady-state mRNA amounts during hyperosmotic stress," *RNA*, vol. 15, no. 4, pp. 600–614, 2009.

[63] M. Chatenay-Laponte and G. S. Shadel, "Repression of mitochondrial translation, respiration and a metabolic cycle-regulated gene, SLF1, by the yeast pumilio-family protein Puf3p," *PLoS One*, vol. 6, no. 5, Article ID e20441, 2011.

[64] A. P. Gasch, P. T. Spellman, C. M. Kao et al., "Genomic expression programs in the response of yeast cells to environmental changes," *Molecular Biology of the Cell*, vol. 11, no. 12, pp. 4241–4257, 2000.

[65] D. Zenklusen, D. R. Larson, and R. H. Singer, "Single-RNA counting reveals alternative modes of gene expression in yeast," *Nature Structural and Molecular Biology*, vol. 15, no. 12, pp. 1263–1271, 2008.

[66] B. Schwanhäusser, D. Busse, N. Li et al., "Global quantification of mammalian gene expression control," *Nature*, vol. 473, pp. 337–342, 2011.

[67] A. D. Aragon, G. A. Quiñones, E. V. Thomas, S. Roy, and M. Werner-Washburne, "Release of extraction-resistant mRNA in stationary phase Saccharomyces cerevisiae produces a massive increase in transcript abundance in response to stress," *Genome Biology*, vol. 7, no. 2, article R9, 2006.

[68] G. Jona, M. Choder, and O. Gileadi, "Glucose starvation induces a drastic reduction in the rates of both transcription and degradation of mRNA in yeast," *Biochimica et Biophysica Acta*, vol. 1491, no. 1–3, pp. 37–48, 2000.

[69] S. Itzkovitz and and A. van Oudenaarden, "Validating transcripts with probes and imaging technology," *Nature Methods*, vol. 8, pp. S12–S19, 2011.

[70] L. Harel-Sharvit, N. Eldad, G. Haimovich, O. Barkai, L. Duek, and M. Choder, "RNA polymerase II subunits link transcription and mRNA decay to translation," *Cell*, vol. 143, no. 4, pp. 552–563, 2010.

[71] J. A. Kruk, A. Dutta, and J. Fu, "The multifunctional Ccr4-not complex directly promotes transcription elongation," *Genes and Development*, vol. 25, no. 6, pp. 581–593, 2011.

SAGE-Hindawi Access to Research
Genetics Research International
Volume 2011, Article ID 707641, 15 pages
doi:10.4061/2011/707641

Review Article

The Roles of the Paf1 Complex and Associated Histone Modifications in Regulating Gene Expression

Elia M. Crisucci and Karen M. Arndt

Department of Biological Sciences, University of Pittsburgh, Pittsburgh, PA 15260, USA

Correspondence should be addressed to Karen M. Arndt, arndt@pitt.edu

Received 14 June 2011; Accepted 31 August 2011

Academic Editor: David Gross

The conserved Paf1 complex (Paf1C) carries out multiple functions during transcription by RNA polymerase (pol) II, and these functions are required for the proper expression of numerous genes in yeast and metazoans. In the elongation stage of the transcription cycle, the Paf1C associates with RNA pol II, interacts with other transcription elongation factors, and facilitates modifications to the chromatin template. At the end of elongation, the Paf1C plays an important role in the termination of RNA pol II transcripts and the recruitment of proteins required for proper RNA 3′ end formation. Significantly, defects in the Paf1C are associated with several human diseases. In this paper, we summarize current knowledge on the roles of the Paf1C in RNA pol II transcription.

1. Introduction

The RNA pol II transcription cycle can be divided into three primary stages: initiation, elongation, and termination. During transcription initiation, the binding of the TATA-binding protein (TBP) subunit of TFIID to the promoter triggers the assembly of a preinitiation complex, which contains RNA pol II and the general transcription factors, TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH (reviewed in [1]). The general transcription factors position the polymerase at the transcription start site and unwind the DNA to expose the template strand for RNA synthesis (reviewed in [1]). During transcription elongation, multiple regulatory proteins associate with RNA pol II to facilitate its progression and modify the chromatin template (reviewed in [2]). Finally, during RNA 3′ end formation and transcription termination, the transcript is processed and released through the combined actions of multiple RNA processing factors (reviewed in [3]). Therefore, each stage of the transcription cycle is regulated by a plethora of proteins to ensure proper gene expression.

The regulation of transcription initiation is an important aspect of controlling gene expression and has thus been studied for many years. More recently, the regulation of post-initiation stages has been shown to be equally important for ensuring proper gene expression. Furthermore, eukaryotic cells have evolved many mechanisms to overcome the barriers imposed by chromatin on all three stages of the transcription cycle. One highly conserved protein complex that lies at the intersection between chromatin modification pathways and transcription is the Paf1C. This complex associates with RNA pol II [4–7] and influences multiple events during transcription elongation, including the posttranslational modification of histone proteins [8–12] and the recruitment of proteins required for RNA processing [13–16]. Not surprisingly given its key regulatory roles, the Paf1C and its functions are conserved throughout eukaryotes. Here, we review current knowledge on the Paf1C, emphasizing insights that have emerged from genetic and biochemical studies in budding yeast and also discussing more recent observations made in multicellular eukaryotes, where defects in the complex lead to developmental abnormalities and disease.

2. Initial Studies in Yeast Led to the Discovery of the Paf1C and Revealed Its Roles in Transcription Elongation

The search for accessory proteins that cooperate with general transcription factors and regulate transcription initiation prompted experiments that led to the identification of Paf1 (polymerase-associated factor 1) in *Saccharomyces cerevisiae* [7, 17]. As the name implies, Paf1 was found to associate with RNA pol II by affinity chromatography [17]. Subsequent studies demonstrated that Paf1 exists in a nuclear complex with Ctr9, Cdc73, Rtf1, and Leo1 [4, 18–20]. Prior to the discovery that they interacted with Paf1, other members of the Paf1C were identified through yeast genetics and recognized for their potential roles in transcription. For example, the *CTR9* gene (Cln three (*CLN3*) requiring 9) was originally identified by its genetic interaction with *CLN3* [21] and subsequently identified in a genetic screen for mutants with impaired transcription of G1 cyclin genes [22]. Additionally, *RTF1* (Restores TBP function 1) was originally discovered in a genetic selection for mutations that suppress transcriptional defects caused by TBP mutants with altered DNA binding specificity [23]. Cdc73 (Cell division cycle 73) [24] and Leo1 (Left open reading frame 1) [25] may not have been initially recognized for their roles in transcription, but were subsequently determined to be important transcriptional regulators in the context of the Paf1C.

Paf1C subunits have been implicated in transcription initiation by influencing the phenotypic effects of a TBP-altered specificity mutant [23] and in transcription termination and RNA 3′ end formation by mediating recruitment of 3′ end processing factors [13–16]. However, the Paf1C is currently best characterized for its critical roles during transcription elongation. Initial studies revealed that genetic disruption of the yeast Paf1C causes phenotypes associated with transcription elongation defects. For example, *S. cerevisiae* strains lacking Paf1C subunits exhibit sensitivity to 6-azauracil (6-AU) and mycophenolic acid (MPA) [20, 26]. These drugs reduce intracellular nucleotide pools, which is thought to increase polymerase pausing, making transcription more dependent on regulatory factors [27]. Consistent with these phenotypes, Paf1C members genetically and physically interact with elongation factors such as the Spt4-Spt5 (yDSIF) and Spt16-Pob3 (yFACT) complexes, suggesting that these complexes function cooperatively to modulate transcription elongation [4, 20, 26]. In agreement with the genetic data, a recently described transcription run-on assay revealed transcription elongation defects in the absence of Paf1C subunits *in vivo* [28]. Despite the strong evidence currently linking the Paf1C to the control of transcription elongation, a less direct role in regulating gene expression was also proposed in an earlier study [29]. The lack of an effect of *rtf1*Δ and *cdc73*Δ mutations on *in vivo* elongation rates or RNA pol II processivity, as measured on an inducible long gene, led to the conclusion that the Paf1C influenced cotranscriptional processes. Indeed, the Paf1C has been implicated in several cotranscriptional processes, including the phosphorylation of RNA pol II during elongation and the recruitment of a chromatin-remodeling enzyme, Chd1, to open reading frames [13, 30, 31]. Additionally, in its best-understood role, the Paf1C is important for the establishment of cotranscriptional histone modifications that influence gene expression [8–12]. Together, these observations suggest that the Paf1C influences gene expression through multiple functions during transcription (Figure 1). In this paper, we describe current information on these functions.

3. The Paf1C Associates with RNA Pol II and Influences the Phosphorylation State of the RNA Pol II CTD

The Paf1C accompanies the polymerase from the transcription start site to the poly(A) site [5, 32]. Rtf1 and Cdc73 are both required for the physical association of Paf1C with RNA pol II. In *rtf1*Δ or *cdc73*Δ cells, the remaining Paf1C subunits dissociate from the polymerase and chromatin, even though these subunits remain associated in a subcomplex [13, 31, 33]. Deletion analysis of *S. cerevisiae RTF1* defined a central region of the Rtf1 protein (amino acids 201 to 395), termed the ORF association region (OAR), that is required for the physical association between the Paf1C and active genes [34]. Although an NMR study has provided important structural information on the human Rtf1 OAR, also known as a Plus3 domain [35], the manner in which Rtf1 interacts with RNA pol II is unknown. Recombinant Cdc73 can interact with purified RNA pol II, suggesting that Cdc73 may directly contact RNA pol II *in vivo* [19]. Beyond the interactions of Rtf1 and Cdc73 with RNA pol II, Leo1 is also required for full association of the Paf1C with active genes [36]. In this case, evidence suggests that an interaction between Leo1 and the nascent mRNA stabilizes the association of the Paf1C with transcribed genes [36].

The interaction between Paf1C subunits and elongating RNA pol II is modulated by other transcription elongation factors. Several reports demonstrated that the Spt4-Spt5 complex promotes recruitment of the Paf1C to chromatin [33, 37–39]. Interestingly, recent studies suggest that the functional interactions between the Paf1C, Spt4-Spt5, and RNA polymerase are conserved beyond RNA pol II and are important for RNA pol I transcription as well [40–44]. Although their roles in Paf1C recruitment are less well characterized than that of Spt4-Spt5, the Spt6, FACT, and Ccr4-Not transcription factors have also been shown to modulate recruitment of the Paf1C to active genes [45–47].

The C-terminal domain (CTD) of the largest subunit of RNA pol II, Rpb1, consists of tandemly repeated copies of a heptapeptide sequence (YSPTSPS) that can be phosphorylated on the serines at positions 2, 5, and 7 of the repeat. Importantly, the phosphorylation state of the CTD changes throughout the transcription cycle and is important for recruiting the appropriate regulatory factors during each stage of transcription (reviewed in [48]). During initiation, the RNA pol II CTD is hypophosphorylated. Upon the transition from initiation to early elongation, the CTD becomes phosphorylated on serine 5 by CDK7

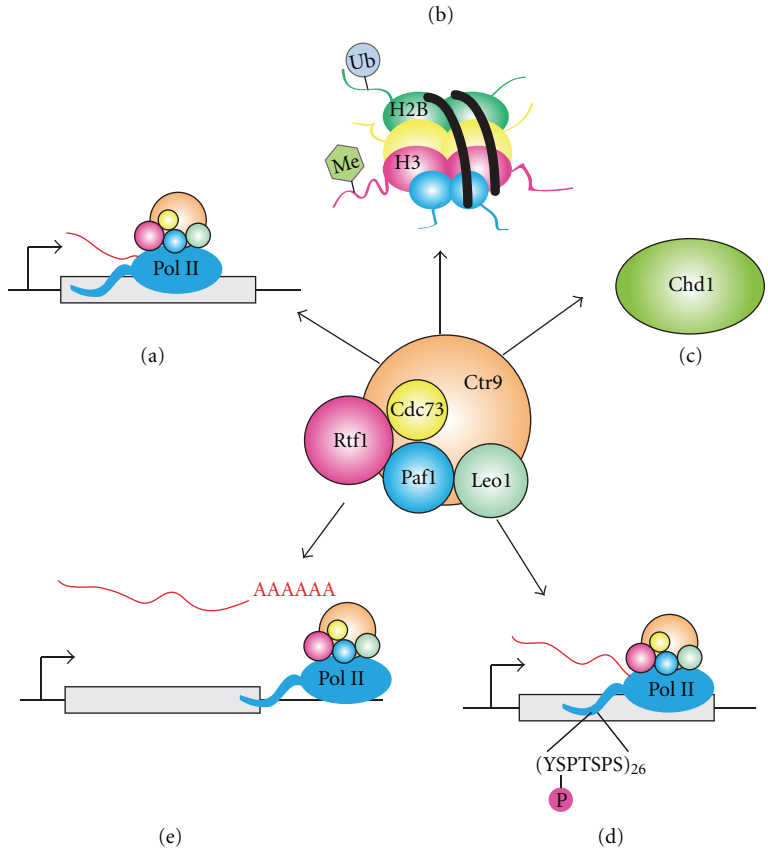


Figure 1: The multiple functions of the Paf1C. During transcription elongation, the Paf1C (a) associates with RNA pol II on coding regions [4, 6], (b) regulates histone modifications [8–12] (discussed in detail in Figure 2), and (c) recruits Chd1, an ATP-dependent chromatin remodeling enzyme [131, 132]. (d) During a later stage of transcription elongation, the Paf1C promotes phosphorylation of serine 2 of the RNA pol II CTD [13, 31]. (e) Additionally, the Paf1C is important for proper transcription termination and RNA 3′ end formation of both polyadenylated and nonpolyadenylated transcripts [13–16, 31].

(Kin28 in yeast) of the general transcription factor TFIIH [49]. Phosphorylated serine 5 is recognized by the mRNA capping machinery, thus coordinating mRNA 5′ end capping and early transcription elongation [50]. In yeast, phosphorylation of serine 5 can be reversed by Ssu72 and tends to decline as elongation proceeds [51]. Serine 7 of the CTD repeats is also phosphorylated by Kin28 [52–54]. Patterns of serine 5 and 7 phosphorylation overlap across genes; however, the elucidation of the functions of serine 7 phosphorylation is still at an early stage [52–55]. Later in elongation, serine 2 of the CTD becomes phosphorylated mainly by Ctk1 in yeast or P-TEFb in human cells [56, 57]. Serine 2 phosphorylation promotes the recruitment of cleavage and polyadenylation factors to RNA pol II, connecting the later stages of elongation to RNA 3′ end processing [58, 59]. Through mechanisms that are undefined, the Paf1C is required for normal levels of serine 2 phosphorylation [13, 31]. In addition to termination and RNA 3′ end formation factors, serine 2 phosphorylation recruits the histone H3 lysine (K) 36 methyltransferase, Set2 [60–63]. Therefore, the Paf1C most likely impacts these processes, in part, through influencing CTD phosphorylation.

4. The Paf1C Influences Gene Expression by Promoting Histone H2B K123 Ubiquitylation and Histone H3 K4 and K79 Methylation

During transcription elongation, RNA pol II encounters obstacles in the form of nucleosomes, the basic units of chromatin. Nucleosomes consist of two copies of each of the four histone proteins, H2A, H2B, H3, and H4, in a globular arrangement, wrapped by 147 base pairs of DNA [64, 65]. A large amount of evidence indicates that nucleosomes impede transcription elongation. For example, elongation efficiency is severely reduced during transcription of reconstituted chromatin templates compared to naked DNA *in vitro* [66, 67]. Furthermore, *in vivo*, transcription rates inversely correlate with nucleosome occupancy within open reading frames (ORFs) [68]. In a recent study that employed a deep-sequencing-based method to determine the positions of all active RNA pol II molecules, extensive pausing and backtracking of the polymerase were observed throughout the bodies of genes [69]. Paused polymerase was particularly noticeable at the positions of the first four nucleosomes, confirming that nucleosomes act as a barrier to transcription elongation *in vivo*.

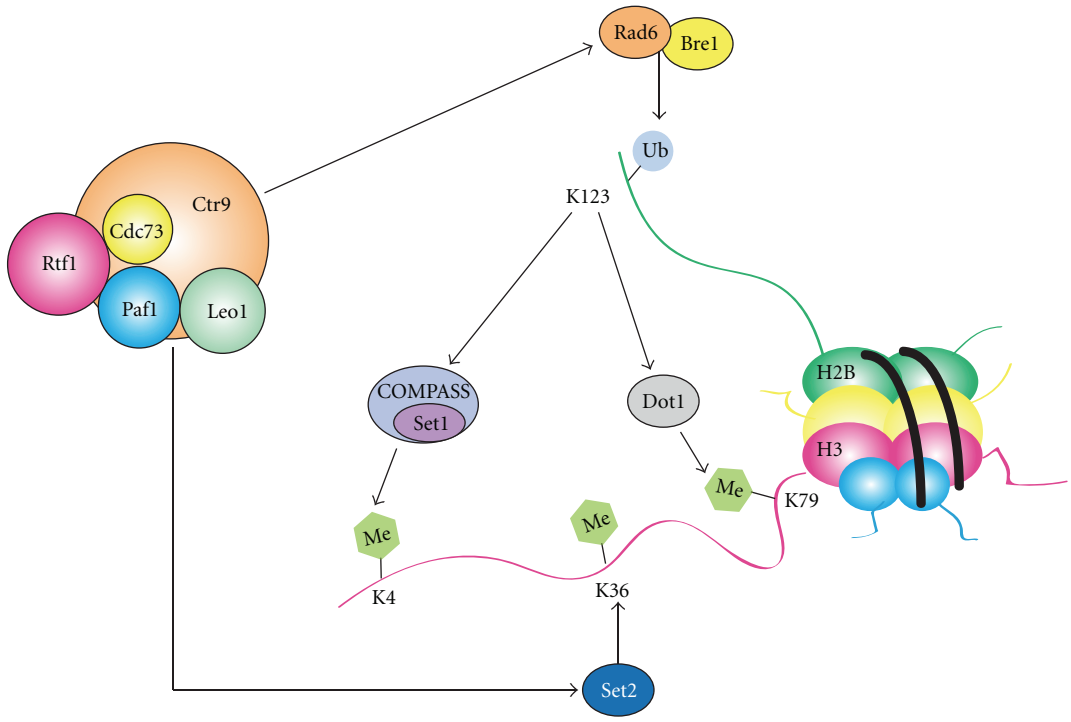


Figure 2: The Paf1C promotes histone H2B monoubiquitylation and histone H3 K4, K36, and K79 methylation. In yeast, the ubiquitin conjugating enzyme, Rad6, and the ubiquitin ligase, Bre1, monoubiquitylate H2B K123 [179–181]. H2B monoubiquitylation is a prerequisite for di- and trimethylation of H3 K4 and K79 by the histone methyltransferases Set1 and Dot1, respectively [71, 80, 182, 183]. Histone H3 is methylated on K36 by the methyltransferase Set2 [184]. Paf1 and Rtf1 subunits of the Paf1C are required for H2B K123 monoubiquitylation and the downstream di- and trimethylation of H3 K4 and K79 [9–12]. Paf1 and Ctr9 are required for trimethylation of K36 on histone H3 [8].

Histone proteins are subject to a wide variety of posttranslational modifications, including the acetylation, ubiquitylation, and methylation of lysine residues. These modifications regulate RNA pol II activity at all stages of the transcription cycle. Members of the Paf1C are required for several histone modifications that are associated with active transcription (Figure 2). Specifically, Paf1 and Rtf1 are required for the monoubiquitylation of H2B on K123 in yeast (K120 in humans) [12, 37, 70] and the subsequent di- and trimethylation of H3 K4 and K79 [9–11, 71]. Paf1 and Rtf1 promote H2B monoubiquitylation by facilitating the association of Rad6 with RNA pol II during transcription elongation [9–12, 70]. An *in vitro* assay using purified factors revealed a direct interaction between the Paf1C and Bre1 [72]; therefore, the Paf1C may tether Rad6 and Bre1 to the elongating polymerase. Since *paf1*Δ cells have greatly reduced levels of Rtf1 protein, Rtf1 is probably the primary subunit that promotes H2B monoubiquitylation and subsequent methylation of H3 on K4 and K79 [13]. In fact, mutational studies have shown that amino acids 62–152 of *S. cerevisiae* Rtf1 are required for these histone modifications, leading to the definition of a histone modification domain (HMD) in Rtf1 [34, 73].

H2B K123 ubiquitylation and H3 K4 and K79 methylation are enriched on the coding regions of active genes [70, 74, 75]. Consistent with a positive role in transcription, H2B monoubiquitylation has been shown to enhance the transcription elongation rate of a chromatin template *in vitro* [47]. *In vivo*, H2B K123 ubiquitylation facilitates transcription of galactose-inducible genes in yeast by promoting nucleosome-reassembly in the wake of RNA pol II in cooperation with the histone chaperone, Spt16 [76]. Additionally, a recent study using chemically defined nucleosome arrays demonstrated that H2B ubiquitylation interferes with chromatin compaction, which may facilitate transcription [77].

H2B K123 ubiquitylation and the downstream methylation of H3 on K4 and K79 regulate the silencing of reporter genes positioned near telomeres and other heterochromatic loci within the yeast genome [10, 78–83]. In *S. cerevisiae*, the silencing of genes near telomeres, and at the *HMR*, *HML*, and rDNA loci, is mediated by silent information regulator (Sir) proteins, which preferentially bind to hypomethylated histones (reviewed in [84]). The genome-wide loss of H3 K4 and K79 methylation has been proposed to cause a redistribution of Sir proteins from the silent loci, resulting in the loss of Sir-dependent transcriptional silencing (reviewed in [85]). Paf1 and Rtf1 are required for silencing of a telomere-proximal reporter gene in yeast [9, 10, 34, 73]. However, recent studies indicate that caution should be exercised in generalizing results obtained with the widely used *URA3*-based silencing reporter assays. In these studies, a *dot1*Δ mutation, which has been reported to cause a strong defect in telomeric silencing based on the reporter assays,

does not alleviate repression of natural genes near telomeres or lead to global changes in Sir protein occupancy [86, 87]. Based on these new observations, more work will be needed to clarify the roles of the Paf1C and its dependent histone modifications in heterochromatic gene silencing.

Microarray analysis of transcript levels in cells lacking the H2B ubiquitylation site (*htb1-K123R* substitution) has shown that H2B K123 ubiquitylation represses many genes throughout the yeast genome [88]. In fact, the majority of affected genes exhibited increased expression in *htb1-K123R* cells, indicating that this modification predominantly acts to repress transcription [88]. Consistent with repressive functions, reversal of H2B K123 ubiquitylation by the de-ubiquitylating enzyme, Ubp8, is required for full expression of certain inducible genes, including *GAL1*, *GAL10*, and *SUC2* [89–91]. Furthermore, the Paf1C mediates repression of a subset of genes, including the *ARG1* gene, by facilitating H2B K123 ubiquitylation [92]. These observations suggest that Paf1C-dependent H2B ubiquitylation has important functions for repression of global transcription. The mechanism by which H2B monoubiquitylation represses transcription is not completely understood. However, H2B monoubiquitylation has been shown to increase nucleosome stability at the promoters of repressed genes [93] and antagonize the recruitment of the positive transcription elongation factor TFIIS to genes in human cells [94].

Importantly, like its yeast counterpart, the human Paf1C controls gene expression through H2B monoubiquitylation and H3 K4 and K79 methylation [47, 95–98]. Furthermore, H2B monoubiquitylation in humans also has both positive and negative effects on transcription. For example, H2B monoubiquitylation is preferentially associated with highly expressed genes [97]. In addition, this modification has been shown to stimulate proper *HOX* gene expression in human cells [98] and the transcription of pluripotency genes in embryonic stem cells [95], thus promoting proper development and stem cell identity, respectively. However, de-ubiquitylation of H2B by Usp22, the human homolog of yeast Ubp8, inhibits heterochromatic silencing and promotes gene activation [99, 100]. Human Bre1/RNF20 acts as a tumor suppressor by promoting transcription of tumor suppressor genes and repressing proto-oncogenes, underscoring the importance of both positive and negative gene regulation by H2B monoubiquitylation [101]. Collectively, these observations indicate that H2B monoubiquitylation has important effects on gene expression in both yeast and humans.

5. The Paf1C Promotes Histone H3 K36 Trimethylation and Affects Histone Acetylation Levels on Genes

In addition to methylation of H3 K4 and K79, Paf1 and Ctr9 are required for trimethylation of H3 K36 by the histone methyltransferase, Set2 [8]. Set2 associates with the elongating form of RNA pol II in the body of actively transcribed genes in a Paf1C-dependent manner [8, 60, 63]. As stated above, the Paf1C may influence Set2 recruitment indirectly through its effects on CTD phosphorylation [13, 31]. Both H3 K4 and K36 methylation occur across most genes in a distinct pattern that is influenced by the phosphorylation state of the RNA pol II CTD (Figure 3). Serine 5 phosphorylation by Kin28 recruits Set1 to RNA pol II early in elongation, resulting in a peak of H3 K4 trimethylation near promoters [11]. Just downstream, K4 dimethylation peaks in 5′ coding regions, whereas K4 monomethylation occurs across the gene [102, 103]. Later in elongation, serine 2 phosphorylation of the RNA pol II CTD recruits Set2, resulting in H3 K36 methylation toward the 3′ end of the coding region [60–63].

Interestingly, H3 K4 and K36 methylation modulate histone acetylation by facilitating the recruitment or activity of histone acetyltransferase (HAT) and histone deacetylase complexes (HDACs). H3 K4 trimethylation recruits the NuA3 HAT complex, resulting in increased H3 K14 acetylation [104, 105]. Dimethylation of H3 K4 stimulates the activity of the Set3 HDAC [106, 107]. Consistent with this pathway of H3 K4 methylation-directed deacetylation, the loss of Paf1 results in increased acetylation at 5′ coding regions [8]. Histone H3 K36 dimethylation promotes the activity of the Rpd3S HDAC [108–111]. Through these pathways, H3 methylation restricts histone acetylation to promoters to prevent inappropriate transcription from cryptic start sites internal to coding regions and restore chromatin in the wake of the polymerase. Analysis of *paf1*Δ *set2*Δ double mutant strains suggests that Paf1 and Set2 function separately to reduce cryptic initiation and histone acetylation at 3′ coding regions [8]. These results may not be surprising since Paf1 is selectively required for H3 K36 trimethylation [8], yet dimethylation is sufficient for Rpd3S HDAC activity [111]. Therefore, at 5′ coding regions, the Paf1C reduces histone acetylation, possibly through H3 K4 methylation-mediated deacetylation by Set3. However, at 3′ coding regions, the Paf1C reduces acetylation through an undefined mechanism that is parallel to the established Set2-Rpd3S pathway.

Given its important roles in modulating several histone modifications, the Paf1C likely regulates gene expression by promoting histone modifications. However, while genome-wide analysis identified numerous genes that require the Paf1C for proper expression [14], only a subset of Paf1C-responsive genes exhibit altered expression in the absence of these same histone modifications [88]. Therefore, the Paf1C likely has roles aside from facilitating histone modifications that control gene expression. Consistent with this hypothesis, the repressive effect of the yeast Paf1C on *ARG1* transcription can be only partially explained by a loss of histone modifications [92]. Furthermore, *in vitro* transcription elongation assays have revealed a role for the Paf1C in stimulating transcription elongation of naked DNA templates by both RNA pol I and pol II [44, 112]. The histone modification-independent functions of the Paf1C may be conserved throughout eukaryotes, as the human Paf1C has recently been shown to stimulate *in vitro* transcription of a chromatin template independently of histone modifications [113]. Further investigation is required to elucidate critical histone modification-independent functions of the Paf1C.

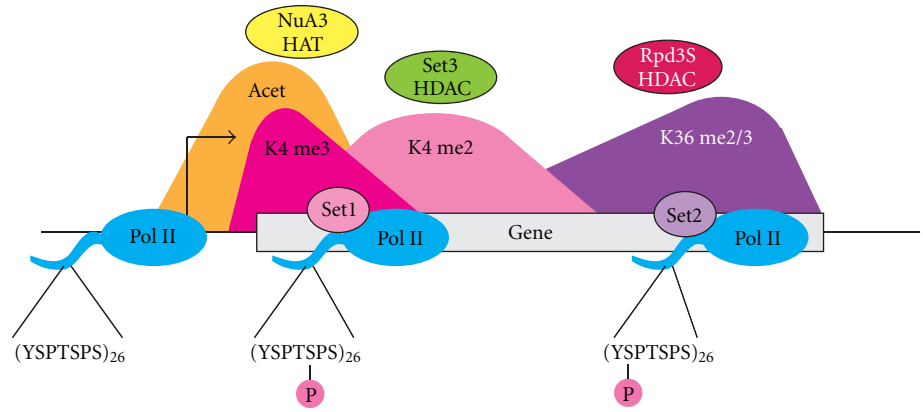


Figure 3: Typical distribution of histone modifications across a gene. Serine 5 phosphorylation on the RNA pol II CTD recruits Set1 to RNA pol II early in elongation, resulting in a peak of H3 K4 trimethylation near promoters [11]. Just downstream, K4 dimethylation peaks in 5′ coding regions [102, 103]. Later in elongation, serine 2 phosphorylation of the RNA pol II CTD recruits Set2, resulting in H3 K36 methylation toward the 3′ end of the coding region [60–63]. H3 K4 trimethylation recruits the NuA3 histone acetyltransferase (HAT) complex, resulting in increased H3 K14 acetylation near the promoter [104, 105]. Within the coding region, dimethylation of H3 K4 promotes the activity of the Set3 histone deacetylase complex (HDAC) [106, 107]. At the 3′ coding region, H3 K36 dimethylation promotes the activity of the Rpd3S HDAC [108–111]. Through these mechanisms, histone acetylation levels are maintained at low levels on coding regions.

6. The Paf1C Functions Cooperatively with Other Factors That Influence Chromatin Structure

Beyond influencing gene expression through establishing histone modifications, the Paf1C interacts physically and genetically with several factors that influence chromatin structure and transcription elongation, including the elongation complex Spt4-Spt5, the histone chaperone FACT, and the chromatin remodeler Chd1. Spt4-Spt5 is required for the recruitment of the Paf1C to the elongation complex and for H2B K123 ubiquitylation [33, 37–39]. Paf1C recruitment is regulated by phosphorylation of Spt5 by the Bur1 kinase [37–39]. Consistent with cooperative functions, the human Paf1C, Spt4-Spt5/DSIF, and Tat-SF1 cooperatively stimulate transcription elongation *in vitro* and promote transcription *in vivo* [114]. A recent study involving a novel transcription elongation reporter template demonstrated that Spt4 and the Paf1C facilitate elongation in yeast cells [28]. However, biochemical experiments in yeast and human cells have shown that Spt4-Spt5/DSIF interacts with RNA pol II during elongation and has both positive and negative effects on elongation [4, 115–118]. Although the functions of Spt4-Spt5/DSIF are not completely understood, genetic interactions with elongation and chromatin-related factors suggest that Spt4-Spt5/DSIF regulates transcription elongation through the modulation of chromatin structure. For example, in yeast Spt4-Spt5 genetically interacts with the ATP-dependent chromatin remodeler, Chd1 [30, 119], and kinases and phosphatases that modify the CTD of RNA pol II [120]. Spt4-Spt5 has been shown to recruit the Rpd3S HDAC to active genes in cooperation with Kin28 and Ctk1 [121]. Therefore, Spt4-Spt5 may affect chromatin in part by recruiting the Rpd3S HDAC. However, *in vitro* transcription using a naked DNA template has shown that the Paf1C and DSIF can also stimulate transcription elongation independently of chromatin [114].

The yeast Paf1C physically and genetically interacts with Spt16-Pob3/FACT [20, 26]. Consistent with this, the human Paf1C augments FACT-stimulated *in vitro* transcription of a chromatin template [47]. FACT was originally identified as a factor that facilitates transcription of a reconstituted chromatin template [67, 122]. Beyond this, many physical and genetic interactions suggest that Spt16-Pob3/FACT has important roles during transcription [20, 26, 30, 122, 123]. It is now known that FACT directly participates in the reorganization of nucleosomes within the ORFs of actively transcribed genes and reassembles chromatin in the wake of RNA pol II [123–127]. *In vitro* binding of an H2A-H2B dimer by FACT has led to the conclusion that FACT displaces a single H2A-H2B dimer to allow RNA pol II passage [122, 128, 129]. However, analysis of nucleosomes by *in vitro* hydroxyl radical accessibility and endonuclease cleavage experiments showed that FACT creates more accessibility than could be explained by the loss of an H2A-H2B dimer, yet partially protects nucleosomal DNA [130]. These results have led to a second model for FACT function in which FACT performs a more dramatic reorganization of the nucleosome without histone eviction.

In addition to its physical interactions with Spt4-Spt5 and FACT, the Paf1C interacts with Chd1, a conserved ATP-dependent chromatin remodeling enzyme [131, 132]. Chd1 associates with regions of active transcription [30, 133] and physically interacts with Spt4-Spt5/DSIF, FACT, and the Paf1C [4, 30, 134], pointing to an important role during transcription elongation. The mechanistic details of chromatin remodeling by Chd1 are not well understood. However, Chd1 has been shown to create a chromatin structure that inhibits cryptic transcription initiation [119]. Chd1 contains two N-terminal chromodomains, a central

ATPase domain and a C-terminal DNA-binding domain [135]. Interestingly, structural studies revealed that the two Chd1 chromodomains regulate the ATPase motor in a unique manner. Specifically, in the absence of nucleosome binding, the chromodomains physically block the ATPase domain, preventing association with naked DNA [136]. Chromodomains also bind methylated lysines [137]. It has been shown that the human homolog of Chd1 associates with chromatin by recognition of H3 K4 methylation [138, 139], but there are differing reports as to whether this occurs in yeast [138–140]. Instead, the Rtf1 subunit of the Paf1C in yeast has been shown to recruit Chd1 to chromatin through a region of Rtf1 distinct from its histone modification domain [30, 34].

7. The Paf1C Coordinates Transcription Elongation with Transcription Termination and RNA 3′ End Processing

In addition to its critical functions during transcription elongation, the Paf1C is important for proper transcription termination and RNA 3′ end formation [13–16, 31]. The loss of Paf1C members results in shorter poly(A) tail lengths [13]. Additionally, the Paf1C has been shown to modulate expression of a subset of genes, not by regulating elongation, but by controlling poly(A) site usage [14]. Specifically, the loss of Paf1 results in the read-through of poly(A) sites, producing 3′-extended transcripts that are subject to nonsense-mediated decay [14]. Termination and RNA 3′ end processing defects that occur in the absence of Paf1 can be attributed to the reduced recruitment of 3′ end processing factors to chromatin. In the absence of Paf1C members, altered poly(A) site usage is associated with reduced chromatin association of the cleavage and polyadenylation factor Pcf11 [13]. Additionally, Cft1, another 3′ end processing factor, associates with RNA pol II in a Paf1C-dependent manner [31]. The recruitment of cleavage and polyadenylation factors to RNA pol II and chromatin requires the serine 2-phosphorylated form of the RNA pol II CTD [58, 141]. Therefore, the Paf1C may regulate the recruitment of 3′ end processing factors indirectly through its effects on CTD phosphorylation. However, direct interactions between the Paf1C and 3′ end processing factors have been demonstrated in yeast and humans [31, 142]. Therefore, the Paf1C may recruit 3′ end processing factors through a mechanism that does not rely on RNA pol II CTD phosphorylation.

Together, these observations suggest that the Paf1C plays an important role in coordinating transcription with RNA 3′ end processing. Given that the Paf1C is required for the recruitment of 3′ end processing factors to chromatin [13, 31], yet it dissociates from RNA pol II shortly after the poly(A) site has been transcribed [5, 32], the Paf1C appears to participate in an exchange of elongation factors for 3′ end processing factors during transcription termination. Consistent with this hypothesis, when dissociated from chromatin, the Paf1C associates with RNA processing factors [31]. However, the exact mechanism by which the Paf1C regulates termination and 3′ end processing of polyadenylated transcripts remains unclear.

The Paf1C is also required for proper termination and 3′ end formation of nonpolyadenylated transcripts [15]. The loss of Paf1C members or Paf1C-dependent histone modifications results in the synthesis of small nucleolar RNAs (snoRNAs) extended at their 3′ ends [15, 73]. snoRNA termination defects in the absence of Paf1C members are associated with reduced recruitment of the 3′ end processing factors, Nrd1 and Nab3 [15]. Therefore, similar to its effects on the termination of polyadenylated transcripts, the Paf1C mediates snoRNA termination by promoting recruitment of 3′ end processing factors. Interestingly, it has recently been shown that the termination function of the Paf1C can be inhibited through an interaction with an activator [143]. Specifically, a physical interaction between Mpk1 MAPK and Paf1 prevents premature transcription termination by inhibiting recruitment of the Sen1-Nrd1-Nab3 complex [143]. However, the mechanism by which the Paf1C recruits 3′ end processing factors for termination remains to be revealed. Additionally, disruption of the Rtf1 HMD results in snoRNA termination defects, implicating H2B K123 ubiquitylation in the regulation of transcription termination [73]. Interestingly, nucleosome depletion in terminator regions has been shown to require RNA pol II transcription [144]. Therefore, in addition to facilitating recruitment of 3′ end processing factors, the Paf1C promotes proper transcription termination through H2B monoubiquitylation and its effects on chromatin structure.

The contribution of the Paf1C to transcription termination has yet to be assessed on a genome-wide scale. However, given the essential roles of transcription termination, which include regulating transcript stability and RNA pol II recycling (reviewed in [145–148]), Paf1C-dependent termination is likely to have wide-spread effects on gene expression. Importantly, the functions of the Paf1C in regulating termination and RNA 3′ end formation are conserved from yeast to humans [142, 149].

8. The Paf1C Has Critical Functions in Metazoans

As mentioned above, the known functions of the Paf1C, including RNA pol II-association [150] and roles in transcription elongation [113, 114], histone modification [95, 96, 98, 142], and RNA 3′ end formation [142, 149], are conserved between yeast and humans. However, there are some differences in complex composition in yeast and higher eukaryotes. In humans, the Paf1C is minimally composed of Paf1, Ctr9, Cdc73, Leo1, and the higher eukaryote-specific subunit, Ski8, which is involved in mRNA surveillance [150–152]. A few reports differ on whether human Rtf1 is absent from [98, 150, 151] or present in [113] the human complex. Therefore, human Rtf1 appears to be less stably associated with the Paf1C. Consistent with this, Rtf1 is not stably associated with the *Drosophila* Paf1C [153]. However, despite its less stable association with the Paf1C, human Rtf1 still influences gene expression and histone modification [95, 154].

The Paf1C has evolved critical functions in higher eukaryotes, including promoting proper development, maintaining pluripotency in stem cells, and preventing cancer. Consistent with an important role in development, the human Paf1C is required for proper transcription of Wnt target genes [155] and *HOX* genes [98]. Additionally, Rtf1 regulates the transcription of Notch target genes in *Drosophila* and zebrafish [153, 156, 157]. Given the regulation of important developmental genes, it is not surprising that, in zebrafish, the Paf1C is required for the development of the ears, neural crest, and heart [156, 158]. For proper heart development, the Paf1C is critical for the specification of cardiomyocytes and patterning of the primitive heart [159]. In addition to genes required for proper development, the human Paf1C regulates the expression of interleukin-6 responsive inflammatory genes [160] and those that maintain pluripotency in stem cells [95].

Members of the human Paf1C have also been implicated in cancer. Pancreatic differentiation factor 2/Paf1 is overexpressed in pancreatic cancer cell lines and overexpression in cell culture results in transformation [161]. Additionally, the gene encoding human Paf1 is amplified in many cancers, including breast and uterine cancers [162, 163]. Furthermore, parafibromin/Cdc73 is a tumor suppressor encoded by HRPT2, a gene that is mutated in hyperparathyroidism-jaw tumor syndrome [164–166]. The roles of the Paf1C in preventing cancer are not entirely understood. However, the Paf1C promotes leukemogenesis through interactions with MLL-rearranged oncoproteins, a topic which has been recently reviewed [167].

Like the yeast counterpart, the human Paf1C influences gene expression by facilitating histone modifications. For example, the Paf1C, H2B monoubiquitylation, and H3 K4 and K79 methylation promote *HOX* gene expression [98]. H2B monoubiquitylation also appears to play a role in maintaining pluripotency in stem cells, and the Paf1C promotes the transcription of genes required for pluripotency in both mouse and human embryonic stem cells [95, 168]. Cell differentiation is associated with reduced expression of Paf1C subunits [95, 168] and reduced levels of H2B ubiquitylation [169]. Interestingly, the silencing of pluripotency genes upon differentiation may be accomplished by the interaction between the Paf1C and DNA methyltransferases [170], which repress these genes [171].

In addition to its effects on histone modifications, the human Paf1C regulates gene expression through direct interactions with gene-specific activators. Human Ctr9 associates with Stat3 and recruits it to the promoters of interleukin-6 responsive genes [160]. Cdc73 in humans promotes the transcription of Wnt target genes through a direct interaction with β-catenin [155]. Additionally, the Paf1C is found in a complex with the transactivator Tat to promote transcription from the HIV-1 promoter [172].

9. Conclusions and Future Studies

The Paf1C performs multiple functions during RNA pol II transcription, and these functions are conserved from yeast to humans. Previous studies have provided a wealth of knowledge about the roles of the Paf1C in transcriptional regulation; however, many important questions remain. While it is known that the Paf1C associates with RNA polymerase during elongation and dissociates near the poly(A) site, the details of the Paf1C-RNA pol II interaction, and its regulation, remain undefined. Current information suggests that the association of the Paf1C with RNA pol II is facilitated by multiple interactions. Cdc73 and Rtf1 play nonredundant roles in tethering the Paf1C to RNA pol II, but it is unclear whether these proteins make direct or indirect contacts with the polymerase or whether, like Leo1 [36], interactions with the nascent transcript are involved. Furthermore, although phosphorylation of Spt5 stimulates Paf1C recruitment [37–39], the regulatory events that promote dissociation of the Paf1C near the poly(A) site have yet to be elucidated.

A requirement for the Paf1C in regulating transcription termination and RNA 3′ end formation has been observed at specific genes. However, determining the scope of this effect on RNA 3′ end formation or other steps in RNA maturation will require additional genome-wide studies on Paf1C-deficient cells. Existing data indicate that the Paf1C mediates transcription termination in several ways. Its physical association with RNA processing factors suggests that the Paf1C coordinates the exchange of transcription elongation factors for transcription termination and 3′ end processing factors [31], although the mechanisms remain to be characterized. Additionally, the transcription termination functions of the Paf1C correlate with its roles in promoting serine 2 phosphorylation of the RNA pol II CTD [13, 31] and H2B ubiquitylation [73]. It is not known whether the Paf1C promotes phosphorylation of serine 2 by affecting the recruitment and/or activity of the CTD kinase, Ctk1. Furthermore, the mechanisms by which the Paf1C promotes histone modifications have not been thoroughly investigated. While the Paf1C facilitates the recruitment of histone modifying enzymes to ORFs [9–12, 70], the molecular details of these interactions are uncharacterized, and it is uncertain whether the Paf1C plays a role in histone modification beyond simply recruiting the active players. Finally, while both positive and negative effects on gene expression have been described for the Paf1C [14, 17, 92, 173] and its downstream histone modifications [88, 92], the features of a gene that confer Paf1C-dependent expression are unknown. In the case of the H2B ubiquitylation mark, it would be especially interesting to know why some genes are repressed by this modification, while others are activated by it [88, 92].

Although it has been shown that the Paf1C is required for proper expression of numerous genes throughout the yeast genome [14], a role in regulating the expression of noncoding RNAs (ncRNAs) has not been determined. In addition to genome-wide analyses unexpectedly localizing RNA pol II to intergenic regions [174], genome-wide transcription analyses have revealed that up to 85% of the yeast genome is transcribed [175, 176]. Similar results were obtained in human cells, such that ncRNAs account for a large portion of the transcription observed [175–178]. Many ncRNAs arise from start sites within intergenic regions and overlap with

coding genes [175, 176]. Importantly, ncRNAs are becoming increasingly recognized as key regulators of gene expression. Therefore, to fully appreciate the mechanisms by which the Paf1C regulates gene expression, it will be important to know how its functions impact ncRNA synthesis. Upcoming investigations, which incorporate a multidisciplinary approach of structural, genetic, biochemical, and genomic experiments, will likely further establish the Paf1C as a critical regulator of gene expression, uncover new activities of the complex, and elucidate the molecular mechanisms of Paf1C-dependent functions that are crucial for the prevention of cancer and developmental defects.

Acknowledgments

The authors thank Joe Martens, Andrew VanDemark, Brett Tomson, Anthony Piro, Kristin Klucevsek, and Margaret Shirra for critical reading of this paper. Funding to K. M. Arndt is provided by the National Institutes of Health (R01 GM52593).

References

[1] T. W. Sikorski and S. Buratowski, "The basal initiation machinery: beyond the general transcription factors," *Current Opinion in Cell Biology*, vol. 21, no. 3, pp. 344–351, 2009.

[2] L. A. Selth, S. Sigurdsson, and J. Q. Svejstrup, "Transcript elongation by RNA polymerase II," *Annual Review of Biochemistry*, vol. 79, pp. 271–293, 2010.

[3] S. Lykke-Andersen and T. H. Jensen, "Overlapping pathways dictate termination of RNA polymerase II transcription," *Biochimie*, vol. 89, no. 10, pp. 1177–1182, 2007.

[4] N. J. Krogan, M. Kim, S. H. Ahn et al., "RNA polymerase II elongation factors of Saccharomyces cerevisiae: a targeted proteomics approach," *Molecular and Cellular Biology*, vol. 22, no. 20, pp. 6979–6992, 2002.

[5] 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.

[6] D. K. Pokholok, N. M. Hannett, and R. A. Young, "Exchange of RNA polymerase II initiation and elongation factors during gene expression in vivo," *Molecular Cell*, vol. 9, no. 4, pp. 799–809, 2002.

[7] P. A. Wade, W. Werel, R. C. Fentzke et al., "A novel collection of accessory factors associated with yeast RNA polymerase II," *Protein Expression and Purification*, vol. 8, no. 1, pp. 85–90, 1996.

[8] Y. Chu, R. Simic, M. H. Warner, K. M. Arndt, and G. Prelich, "Regulation of histone modification and cryptic transcription by the Bur1 and Paf1 complexes," *EMBO Journal*, vol. 26, no. 22, pp. 4646–4656, 2007.

[9] 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.

[10] H. H. Ng, S. Dole, and K. Struhl, "The Rtf1 Component of the Paf1 Transcriptional Elongation Complex Is Required for Ubiquitination of Histone H2B," *Journal of Biological Chemistry*, vol. 278, no. 36, pp. 33625–33628, 2003.

[11] 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.

[12] A. Wood, J. Schneider, J. Dover, M. Johnston, and A. Shilatifard, "The Paf1 complex is essential for histone monoubiquitination by the Rad6-Bre1 complex, which signals for histone methylation by COMPASS and Dot1p," *Journal of Biological Chemistry*, vol. 278, no. 37, pp. 34739–34742, 2003.

[13] C. L. Mueller, S. E. Porter, M. G. Hoffman, and J. A. Jaehning, "The Paf1 complex has functions independent of actively transcribing RNA polymerase II," *Molecular Cell*, vol. 14, no. 4, pp. 447–456, 2004.

[14] K. L. Penheiter, T. M. Washburn, S. E. Porter, M. G. Hoffman, and J. A. Jaehning, "A posttranscriptional role for the yeast Paf1-RNA polymerase II complex is revealed by identification of primary targets," *Molecular Cell*, vol. 20, no. 2, pp. 213–223, 2005.

[15] K. E. Sheldon, D. M. Mauger, and K. M. Arndt, "A requirement for the Saccharomyces cerevisiae Paf1 complex in snoRNA 3′ end formation," *Molecular Cell*, vol. 20, no. 2, pp. 225–236, 2005.

[16] B. N. Tomson, C. P. Davis, M. H. Warner, and K. M. Arndt, "Identification of a role for histone H2B ubiquitylation in noncoding RNA 3′-end formation through mutational analysis of Rtf1 in Saccharomyces cerevisiae," *Genetics*, vol. 188, no. 2, pp. 273–289, 2011.

[17] X. Shi, A. Finkelstein, A. J. Wolf, P. A. Wade, Z. F. Burton, and J. A. Jaehning, "Paf1p, an RNA polymerase II-associated factor in Saccharomyces cerevisiae, may have both positive and negative roles in transcription," *Molecular and Cellular Biology*, vol. 16, no. 2, pp. 669–676, 1996.

[18] C. L. Mueller and J. A. Jaehning, "Ctr9, Rtf1, and Leo1 are components of the Paf1/RNA polymerase II complex," *Molecular and Cellular Biology*, vol. 22, no. 7, pp. 1971–1980, 2002.

[19] X. Shi, M. Chang, A. J. Wolf et al., "Cdc73p and Paf1p are found in a novel RNA polymerase II-containing complex distinct from the Srbp-containing holoenzyme," *Molecular and Cellular Biology*, vol. 17, no. 3, pp. 1160–1169, 1997.

[20] S. L. Squazzo, P. J. Costa, D. L. Lindstrom et al., "The Paf1 complex physically and functionally associates with transcription elongation factors in vivo," *EMBO Journal*, vol. 21, no. 7, pp. 1764–1774, 2002.

[21] C. J. Di Como, H. Chang, and K. T. Arndt, "Activation of CLN1 and CLN2 G1 cyclin gene expression by BCK2," *Molecular and Cellular Biology*, vol. 15, no. 4, pp. 1835–1846, 1995.

[22] C. Koch, P. Wollmann, M. Dahl, and F. Lottspeich, "A role for Ctr9p and Paf1p in the regulation G1 cyclin expression in yeast," *Nucleic Acids Research*, vol. 27, no. 10, pp. 2126–2134, 1999.

[23] L. A. Stolinski, D. M. Eisenmann, and K. M. Arndt, "Identification of RTF1, a novel gene important for TATA site selection by TATA box-binding protein in Saccharomyces cerevisiae," *Molecular and Cellular Biology*, vol. 17, no. 8, pp. 4490–4500, 1997.

[24] S. I. Reed, J. Ferguson, and K. Y. Jahng, "Isolation and characterization of two genes encoding yeast mating pheromone signaling elements: CDC72 and CDC73," *Cold Spring Harbor Symposia on Quantitative Biology*, vol. 53, no. 2, pp. 621–627, 1988.

[25] V. Magdolen, P. Lang, G. Mages, H. Hermann, and W. Bandlow, "The gene LEO1 on yeast chromosome XV encodes a non-essential, extremely hydrophilic protein," *Biochimica et Biophysica Acta*, vol. 1218, no. 2, pp. 205–209, 1994.

[26] P. J. Costa and K. M. Arndt, "Synthetic lethal interactions suggest a role for the Saccharomyces cerevisiae Rtf1 protein in transcription elongation," *Genetics*, vol. 156, no. 2, pp. 535–547, 2000.

[27] F. Exinger and F. Lacroute, "6-Azauracil inhibition of GTP biosynthesis in Saccharomyces cerevisiae," *Current Genetics*, vol. 22, no. 1, pp. 9–11, 1992.

[28] C. Tous, A. G. Rondón, M. García-Rubio, C. González-Aguilera, R. Luna, and A. Aguilera, "A novel assay identifies transcript elongation roles for the Nup84 complex and RNA processing factors," *EMBO Journal*, vol. 30, no. 10, pp. 1953–1964, 2011.

[29] P. B. Mason and K. Struhl, "Distinction and relationship between elongation rate and processivity of RNA polymerase II in vivo," *Molecular Cell*, vol. 17, no. 6, pp. 831–840, 2005.

[30] R. Simic, D. L. Lindstrom, H. G. Tran et al., "Chromatin remodeling protein Chd1 interacts with transcription elongation factors and localizes to transcribed genes," *EMBO Journal*, vol. 22, no. 8, pp. 1846–1856, 2003.

[31] K. Nordick, M. G. Hoffman, J. L. Betz, and J. A. Jaehning, "Direct interactions between the Paf1 complex and a cleavage and polyadenylation factor are revealed by dissociation of Paf1 from RNA polymerase II," *Eukaryotic Cell*, vol. 7, no. 7, pp. 1158–1167, 2008.

[32] M. Kim, S. H. Ahn, N. J. Krogan, J. F. Greenblatt, and S. Buratowski, "Transitions in RNA polymerase II elongation complexes at the 3′ ends of genes," *EMBO Journal*, vol. 23, no. 2, pp. 354–364, 2004.

[33] H. Qiu, C. Hu, C. M. Wong, and A. G. Hinnebusch, "The Spt4p subunit of yeast DSIF stimulates association of the Paf1 complex with elongating RNA polymerase II," *Molecular and Cellular Biology*, vol. 26, no. 8, pp. 3135–3148, 2006.

[34] M. H. Warner, K. L. Roinick, and K. M. Arndt, "Rtf1 is a multifunctional component of the Paf1 complex that regulates gene expression by directing cotranscriptional histone modification," *Molecular and Cellular Biology*, vol. 27, no. 17, pp. 6103–6115, 2007.

[35] R. N. de Jong, V. Truffault, T. Diercks et al., "Structure and DNA Binding of the Human Rtf1 Plus3 Domain," *Structure*, vol. 16, no. 1, pp. 149–159, 2008.

[36] J. L. Dermody and S. Buratowski, "Leo1 subunit of the yeast Paf1 complex binds RNA and contributes to complex recruitment," *Journal of Biological Chemistry*, vol. 285, no. 44, pp. 33671–33679, 2010.

[37] R. N. Laribee, N. J. Krogan, T. Xiao et al., "BUR kinase selectively regulates H3 K4 trimethylation and H2B ubiquitylation through recruitment of the PAF elongation complex," *Current Biology*, vol. 15, no. 16, pp. 1487–1493, 2005.

[38] K. Zhou, W. H. W. Kuo, J. Fillingham, and J. F. Greenblatt, "Control of transcriptional elongation and cotranscriptional histone modification by the yeast BUR kinase substrate Spt5," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 17, pp. 6956–6961, 2009.

[39] 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.

[40] S. J. Anderson, M. L. Sikes, Y. Zhang et al., "The transcription elongation factor Spt5 influences transcription by RNA polymerase I positively and negatively," *Journal of Biological Chemistry*, vol. 286, no. 21, pp. 18816–18824, 2011.

[41] D. A. Schneider, S. L. French, Y. N. Osheim et al., "RNA polymerase II elongation factors Spt4p and Spt5p play roles in transcription elongation by RNA polymerase I and rRNA processing," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 34, pp. 12707–12712, 2006.

[42] O. V. Viktorovskaya, F. D. Appling, and D. A. Schneider, "Yeast transcription elongation factor Spt5 associates with RNA polymerase I and RNA polymerase II directly," *Journal of Biological Chemistry*, vol. 286, no. 21, pp. 18825–18833, 2011.

[43] Y. Zhang, M. L. Sikes, A. L. Beyer, and D. A. Schneider, "The Paf1 complex is required for efficient transcription elongation by RNA polymerase I," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 7, pp. 2153–2158, 2009.

[44] Y. Zhang, A. D. Smith, M. B. Renfrow, and D. A. Schneider, "The RNA polymerase-associated factor 1 complex (Paf1C) directly increases the elongation rate of RNA polymerase I and is required for efficient regulation of rRNA synthesis," *Journal of Biological Chemistry*, vol. 285, no. 19, pp. 14152–14159, 2010.

[45] C. D. Kaplan, M. J. Holland, and F. Winston, "Interaction between transcription elongation factors and mRNA 3′-end formation at the Saccharomyces cerevisiae GAL10-GAL7 locus," *Journal of Biological Chemistry*, vol. 280, no. 2, pp. 913–922, 2005.

[46] K. W. Mulder, A. B. Brenkman, A. Inagaki, N. J. F. van den Broek, and H. T. Marc Timmers, "Regulation of histone H3K4 tri-methylation and PAF complex recruitment by the Ccr4-Not complex," *Nucleic Acids Research*, vol. 35, no. 7, pp. 2428–2439, 2007.

[47] R. Pavri, B. Zhu, G. Li et al., "Histone H2B Monoubiquitination Functions Cooperatively with FACT to Regulate Elongation by RNA Polymerase II," *Cell*, vol. 125, no. 4, pp. 703–717, 2006.

[48] S. Buratowski, "Progression through the RNA Polymerase II CTD Cycle," *Molecular Cell*, vol. 36, no. 4, pp. 541–546, 2009.

[49] 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.

[50] 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.

[51] 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.

[52] 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.

[53] 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.

[54] 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.

[55] 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.

[56] 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.

[57] M. Patturajan, N. K. Conrad, D. B. Bregman, and J. L. Corden, "Yeast carboxyl-terminal domain kinase I positively and negatively regulates RNA polymerase II carboxyl-terminal domain phosphorylation," *Journal of Biological Chemistry*, vol. 274, no. 39, pp. 27823–27828, 1999.

[58] 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.

[59] Z. Ni, B. E. Schwartz, J. Werner, J. R. Suarez, and J. T. Lis, "Coordination of transcription, RNA processing, and surveillance by P-TEFb kinase on heat shock genes," *Molecular Cell*, vol. 13, no. 1, pp. 55–65, 2004.

[60] 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.

[61] B. Li, L. Howe, S. Anderson, J. R. Yates, and J. L. Workman, "The Set2 histone methyltransferase functions through the phosphorylated carboxyl-terminal domain of RNA polymerase II," *Journal of Biological Chemistry*, vol. 278, no. 11, pp. 8897–8903, 2003.

[62] D. Schaft, A. Roguev, K. M. Kotovic et al., "The histone 3 lysine methyltransferase, SET2, is involved in transcriptional elongation," *Nucleic Acids Research*, vol. 31, no. 10, pp. 2475–2482, 2003.

[63] 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.

[64] R. D. Kornberg, "Chromatin structure: a repeating unit of histones and DNA," *Science*, vol. 184, no. 4139, pp. 868–871, 1974.

[65] K. Luger, A. W. Mäder, R. K. Richmond, D. F. Sargent, and T. J. Richmond, "Crystal structure of the nucleosome core particle at 2.8 Å resolution," *Nature*, vol. 389, no. 6648, pp. 251–260, 1997.

[66] M. G. Izban and D. S. Luse, "Transcription on nucleosomal templates by RNA polymerase II in vitro: inhibition of elongation with enhancement of sequence-specific pausing," *Genes and Development*, vol. 5, no. 4, pp. 683–696, 1991.

[67] G. Orphanides, G. LeRoy, C. H. Chang, D. S. Luse, and D. Reinberg, "FACT, a factor that facilitates transcript elongation through nucleosomes," *Cell*, vol. 92, no. 1, pp. 105–116, 1998.

[68] C. K. Lee, Y. Shibata, B. Rao, B. D. Strahl, and J. D. Lieb, "Evidence for nucleosome depletion at active regulatory regions genome-wide," *Nature Genetics*, vol. 36, no. 8, pp. 900–905, 2004.

[69] L. S. Churchman and J. S. Weissman, "Nascent transcript sequencing visualizes transcription at nucleotide resolution," *Nature*, vol. 469, no. 7330, pp. 368–373, 2011.

[70] T. Xiao, C. F. Kao, N. J. Krogan et al., "Histone H2B ubiquitylation is associated with elongating RNA polymerase II," *Molecular and Cellular Biology*, vol. 25, no. 2, pp. 637–651, 2005.

[71] S. D. Briggs, T. Xiao, Z. W. Sun et al., "Trans-histone regulatory pathway in chromatin," *Nature*, vol. 418, no. 6897, p. 498, 2002.

[72] J. Kim and R. G. Roeder, "Direct Bre1-Paf1 complex interactions and RING finger-independent Bre1-Rad6 interactions mediate histone H2B ubiquitylation in yeast," *Journal of Biological Chemistry*, vol. 284, no. 31, pp. 20582–20592, 2009.

[73] B. N. Tomson, C. P. Davis, M. H. Warner, and K. M. Arndt, "Identification of a role for histone H2B ubiquitylation in noncoding RNA 3′-end formation through mutational analysis of Rtf1 in Saccharomyces cerevisiae," *Genetics*, vol. 188, no. 2, pp. 273–289, 2011.

[74] B. E. Bernstein, E. L. Humphrey, R. L. Erlich et al., "Methylation of histone H3 Lys 4 in coding regions of active genes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 13, pp. 8695–8700, 2002.

[75] H. Santos-Rosa, R. Schneider, A. J. Bannister et al., "Active genes are tri-methylated at K4 of histone H3," *Nature*, vol. 419, no. 6905, pp. 407–411, 2002.

[76] A. B. Fleming, C. F. Kao, C. Hillyer, M. Pikaart, and M. A. Osley, "H2B Ubiquitylation Plays a Role in Nucleosome Dynamics during Transcription Elongation," *Molecular Cell*, vol. 31, no. 1, pp. 57–66, 2008.

[77] B. Fierz, C. Chatterjee, R. K. McGinty, M. Bar-Dagan, D. P. Raleigh, and T. W. Muir, "Histone H2B ubiquitylation disrupts local and higher-order chromatin compaction," *Nature Chemical Biology*, vol. 7, pp. 113–119, 2011.

[78] N. J. Krogan, J. Dover, S. Khorrami et al., "COMPASS, a histone H3 (lysine 4) methyltransferase required for telomeric silencing of gene expression," *Journal of Biological Chemistry*, vol. 277, no. 13, pp. 10753–10755, 2002.

[79] C. Nislow, E. Ray, and L. Pillus, "SET1, a yeast member of the Trithorax family, functions in transcriptional silencing and diverse cellular processes," *Molecular Biology of the Cell*, vol. 8, no. 12, pp. 2421–2436, 1997.

[80] Z. W. Sun and C. D. Allis, "Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast," *Nature*, vol. 418, no. 6893, pp. 104–108, 2002.

[81] H. Huang, A. Kahana, D. E. Gottschling, L. Prakash, and S. W. Liebman, "The ubiquitin-conjugating enzyme Rad6 (Ubc2) is required for silencing in Saccharomyces cerevisiae," *Molecular and Cellular Biology*, vol. 17, no. 11, pp. 6693–6699, 1997.

[82] H. H. Ng, Q. Feng, H. Wang et al., "Lysine methylation within the globular domain of histone H3 by Dot1 is important for telomeric silencing and Sir protein association," *Genes and Development*, vol. 16, no. 12, pp. 1518–1527, 2002.

[83] F. van Leeuwen, P. R. Gafken, and D. E. Gottschling, "Dot1p modulates silencing in yeast by methylation of the nucleosome core," *Cell*, vol. 109, no. 6, pp. 745–756, 2002.

[84] M. D. Shahbazian and M. Grunstein, "Functions of Site-Specific histone acetylation and deacetylation," *Annual Review of Biochemistry*, vol. 76, pp. 75–100, 2007.

[85] L. N. Rusche, A. L. Kirchmaier, and J. Rine, "The establishment, inheritance, and function of silenced chromatin in Saccharomyces cerevisiae," *Annual Review of Biochemistry*, vol. 72, pp. 481–516, 2003.

[86] M. P. Rossmann, W. Luo, O. Tsaponina, A. Chabes, and B. Stillman, "A common telomeric gene silencing assay is affected by nucleotide metabolism," *Molecular Cell*, vol. 42, pp. 127–136, 2011.

[87] Y. H. Takahashi, J. Schulze, J. Jackson et al., "Dot1 and Histone H3K79 Methylation in Natural Telomeric and HM Silencing," *Molecular Cell*, vol. 42, no. 1, pp. 118–126, 2011.

[88] A. I. Mutiu, S. M. T. Hoke, J. Genereaux, G. Liang, and C. J. Brandl, "The role of histone ubiquitylation and deubiquitylation in gene expression as determined by the analysis of an HTB1K123R Saccharomyces cerevisiae strain," *Molecular Genetics and Genomics*, vol. 277, no. 5, pp. 491–506, 2007.

[89] J. A. Daniel, M. S. Torok, Z. W. Sun et al., "Deubiquitination of Histone H2B by a Yeast Acetyltransferase Complex Regulates Transcription," *Journal of Biological Chemistry*, vol. 279, no. 3, pp. 1867–1871, 2004.

[90] K. W. Henry, A. Wyce, W. S. Lo et al., "Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8," *Genes and Development*, vol. 17, no. 21, pp. 2648–2663, 2003.

[91] C. F. Kao, C. Hillyer, T. Tsukuda, K. Henry, S. Berger, and M. A. Osley, "Rad6 plays a role in transcriptional activation through ubiquitylation of histone H2B," *Genes and Development*, vol. 18, no. 2, pp. 184–195, 2004.

[92] E. M. Crisucci and K. M. Arndt, "The Paf1 complex represses ARG1 transcription in Saccharomyces cerevisiae by promoting histone modifications," *Eukaryotic Cell*, vol. 10, no. 6, pp. 712–723, 2011.

[93] M. B. Chandrasekharan, F. Huang, and Z. W. Sun, "Ubiquitination of histone H2B regulates chromatin dynamics by enhancing nucleosome stability," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 39, pp. 16686–16691, 2009.

[94] E. Shema, J. Kim, R. Roeder, and M. Oren, "RNF20 inhibits TFIIS-facilitated transcriptional elongation to suppress pro-oncogenic gene expression," *Molecular Cell*, vol. 42, no. 4, pp. 477–488, 2011.

[95] L. Ding, M. Paszkowski-Rogacz, A. Nitzsche et al., "A genome-scale RNAi screen for Oct4 modulators defines a role of the Paf1 complex for embryonic stem cell identity," *Cell Stem Cell*, vol. 4, no. 5, pp. 403–415, 2009.

[96] J. Kim, M. Guermah, R. K. McGinty et al., "RAD6-mediated transcription-coupled H2B ubiquitylation directly stimulates H3K4 methylation in human cells," *Cell*, vol. 137, no. 3, pp. 459–471, 2009.

[97] N. Minsky, E. Shema, Y. Field, M. Schuster, E. Segal, and M. Oren, "Monoubiquitinated H2B is associated with the transcribed region of highly expressed genes in human cells," *Nature Cell Biology*, vol. 10, no. 4, pp. 483–488, 2008.

[98] B. Zhu, Y. Zheng, A. D. Pham et al., "Monoubiquitination of human histone H2B: the factors involved and their roles in HOX gene regulation," *Molecular Cell*, vol. 20, no. 4, pp. 601–611, 2005.

[99] X. Y. Zhang, H. K. Pfeiffer, A. W. Thorne, and S. B. McMahon, "USP22, an hSAGA subunit and potential cancer stem cell marker, reverses the polycomb-catalyzed ubiquitylation of histone H2A," *Cell Cycle*, vol. 7, no. 11, pp. 1522–1524, 2008.

[100] Y. Zhao, G. Lang, S. Ito et al., "A TFTC/STAGA module mediates histone H2A and H2B deubiquitination, coactivates nuclear receptors, and counteracts heterochromatin silencing," *Molecular Cell*, vol. 29, no. 1, pp. 92–101, 2008.

[101] E. Shema, I. Tirosh, Y. Aylon et al., "The histone H2B-specific ubiquitin ligase RNF20/hBREl acts as a putative tumor suppressor through selective regulation of gene expression," *Genes and Development*, vol. 22, no. 19, pp. 2664–2676, 2008.

[102] L. L. Chih, T. Kaplan, M. Kim et al., "Single-nucleosome mapping of histone modifications in S. cerevisiae," *PLoS Biology*, vol. 3, no. 10, article no. e328, p. 17, 2005.

[103] D. K. Pokholok, C. T. Harbison, S. Levine et al., "Genome-wide map of nucleosome acetylation and methylation in yeast," *Cell*, vol. 122, no. 4, pp. 517–527, 2005.

[104] D. G. E. Martin, K. Baetz, X. Shi et al., "The Yng1p plant homeodomain finger is a methyl-histone binding module that recognizes lysine 4-methylated histone H3," *Molecular and Cellular Biology*, vol. 26, no. 21, pp. 7871–7879, 2006.

[105] S. D. Taverna, S. Ilin, R. S. Rogers et al., "Yng1 PHD finger binding to H3 Trimethylated at K4 promotes NuA3 HAT activity at K14 of H3 and transcription at a subset of targeted ORFs," *Molecular Cell*, vol. 24, no. 5, pp. 785–796, 2006.

[106] T. Kim and S. Buratowski, "Dimethylation of H3K4 by Set1 Recruits the Set3 Histone Deacetylase Complex to 5′ Transcribed Regions," *Cell*, vol. 137, no. 2, pp. 259–272, 2009.

[107] 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.

[108] M. J. Carrozza, B. Li, L. Florens et al., "Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription," *Cell*, vol. 123, no. 4, pp. 581–592, 2005.

[109] 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.

[110] B. Li, M. Gogol, M. Carey, D. Lee, C. Seidel, and J. L. Workman, "Combined action of PHD and chromo domains directs the Rpd3S HDAC to transcribed chromatin," *Science*, vol. 316, no. 5827, pp. 1050–1054, 2007.

[111] B. Li, J. Jackson, M. D. Simon et al., "Histone H3 lysine 36 dimethylation (H3K36me2) is sufficient to recruit the Rpd3s Histone deacetylase complex and to repress spurious transcription," *Journal of Biological Chemistry*, vol. 284, no. 12, pp. 7970–7976, 2009.

[112] A. G. Rondón, M. Gallardo, M. García-Rubio, and A. Aguilera, "Molecular evidence indicating that the yeast PAF complex is required for transcription elongation," *EMBO Reports*, vol. 5, no. 1, pp. 47–53, 2004.

[113] J. Kim, M. Guermah, and R. G. Roeder, "The human PAF1 complex acts in chromatin transcription elongation both independently and cooperatively with SII/TFIIS," *Cell*, vol. 140, no. 4, pp. 491–503, 2010.

[114] Y. Chen, Y. Yamaguchi, Y. Tsugeno et al., "DSIF, the Paf1 complex, and Tat-SF1 have nonredundant, cooperative roles in RNA polymerase II elongation," *Genes and Development*, vol. 23, no. 23, pp. 2765–2777, 2009.

[115] C. F. Bourgeois, Y. K. Kim, M. J. Churcher, M. J. West, and J. Karn, "Spt5 cooperates with human immunodeficiency virus type 1 Tat by preventing premature RNA release at terminator sequences," *Molecular and Cellular Biology*, vol. 22, no. 4, pp. 1079–1093, 2002.

[116] G. A. Hartzog, T. Wada, H. Handa, and F. Winston, "Evidence that Spt4, Spt5, and Spt6 control transcription elongation by RNA polymerase II in Saccharomyces cerevisiae," *Genes and Development*, vol. 12, no. 3, pp. 357–369, 1998.

[117] G. A. Hartzog, J. L. Speer, and D. L. Lindstrom, "Transcript elongation on a nucleoprotein template," *Biochimica et Biophysica Acta*, vol. 1577, no. 2, pp. 276–286, 2002.

[118] D. L. Lindstrom, S. L. Squazzo, N. Muster et al., "Dual roles for Spt5 in pre-mRNA processing and transcription elongation revealed by identification of Spt5-associated proteins," *Molecular and Cellular Biology*, vol. 23, no. 4, pp. 1368–1378, 2003.

[119] T. K. Quan and G. A. Hartzog, "Histone H3K4 and K36 methylation, Chd1 and Rpd3S oppose the functions of Saccharomyces cerevisiae Spt4-Spt5 in transcription," *Genetics*, vol. 184, no. 2, pp. 321–334, 2010.

[120] D. L. Lindstrom and G. A. Hartzog, "Genetic interactions of Spt4-Spt5 and TFIIS with the RNA polymerase II CTD and CTD modifying enzymes in Saccharomyces cerevisiae," *Genetics*, vol. 159, no. 2, pp. 487–497, 2001.

[121] S. Drouin, L. Laramée, 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 Genetics*, vol. 6, no. 10, article e1001173, pp. 1–12, 2010.

[122] G. Orphanides, W. H. Wu, W. S. Lane, M. Hampsey, and D. Reinberg, "The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins," *Nature*, vol. 400, no. 6741, pp. 284–288, 1999.

[123] T. Formosa, S. Ruone, M. D. Adams et al., "Defects in SPT16 or POB3 (yFACT) in Saccharomyces cerevisiae cause dependence on the Hir/Hpc pathway: polymerase passage may degrade chromatin structure," *Genetics*, vol. 162, no. 4, pp. 1557–1571, 2002.

[124] A. Jamai, A. Puglisi, and M. Strubin, "Histone chaperone Spt16 promotes redeposition of the original H3-H4 histones evicted by elongating RNA polymerase," *Molecular Cell*, vol. 35, no. 3, pp. 377–383, 2009.

[125] G. Orphanides and D. Reinberg, "RNA polymerase II elongation through chromatin," *Nature*, vol. 407, no. 6803, pp. 471–475, 2000.

[126] P. B. Mason and K. Struhl, "The FACT complex travels with elongating RNA polymerase II and is important for the fidelity of transcriptional initiation in vivo," *Molecular and Cellular Biology*, vol. 23, no. 22, pp. 8323–8333, 2003.

[127] C. D. Kaplan, L. Laprade, and F. Winston, "Transcription elongation factors repress transcription initiation from cryptic sites," *Science*, vol. 301, no. 5636, pp. 1096–1099, 2003.

[128] R. Belotserkovskaya, S. Oh, V. A. Bondarenko, G. Orphanides, V. M. Studitsky, and D. Reinberg, "FACT facilitates transcription-dependent nucleosome alteration," *Science*, vol. 301, no. 5636, pp. 1090–1093, 2003.

[129] D. Reinberg and R. J. Sims, "De FACTo nucleosome dynamics," *Journal of Biological Chemistry*, vol. 281, no. 33, pp. 23297–23301, 2006.

[130] H. Xin, S. Takahata, M. Blanksma, L. McCullough, D. J. Stillman, and T. Formosa, "yFACT induces global accessibility of nucleosomal DNA without H2A-H2B displacement," *Molecular Cell*, vol. 35, no. 3, pp. 365–376, 2009.

[131] C. Stockdale, A. Flaus, H. Ferreira, and T. Owen-Hughes, "Analysis of nucleosome repositioning by yeast ISWI and Chd1 chromatin remodeling complexes," *Journal of Biological Chemistry*, vol. 281, no. 24, pp. 16279–16288, 2006.

[132] H. G. Tran, D. J. Steger, V. R. Iyer, and A. D. Johnson, "The chrome domain protein Chd1p from budding yeast is an ATP-dependent chromatin-modifying factor," *EMBO Journal*, vol. 19, no. 10, pp. 2323–2331, 2000.

[133] D. G. Stokes, K. D. Tartof, and R. P. Perry, "CHD1 is concentrated in interbands and puffed regions of Drosophila polytene chromosomes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, no. 14, pp. 7137–7142, 1996.

[134] D. E. Kelley, D. G. Stokes, and R. P. Perry, "CHD1 interacts with SSRP1 and depends on both its chromodomain and its ATPase/helicase-like domain for proper association with chromatin," *Chromosoma*, vol. 108, no. 1, pp. 10–25, 1999.

[135] V. Delmas, D. G. Stokes, and R. P. Perry, "A mammalian DNA-binding protein that contains a chromodomain and an SNF2/SW12-like helicase domain," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 90, no. 6, pp. 2414–2418, 1993.

[136] G. Hauk, J. N. McKnight, I. M. Nodelman, and G. D. Bowman, "The Chromodomains of the Chd1 Chromatin Remodeler Regulate DNA Access to the ATPase Motor," *Molecular Cell*, vol. 39, no. 5, pp. 711–723, 2010.

[137] A. J. Bannister, P. Zegerman, J. F. Partridge et al., "Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain," *Nature*, vol. 410, no. 6824, pp. 120–124, 2001.

[138] J. F. Flanagan, L. Z. Mi, M. Chruszcz et al., "Double chromodomains cooperate to recognize the methylated histone H3 tail," *Nature*, vol. 438, no. 7071, pp. 1181–1185, 2005.

[139] R. J. Sims, C. F. Chen, H. Santos-Rosa, T. Kouzarides, S. S. Patel, and D. Reinberg, "Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains," *Journal of Biological Chemistry*, vol. 280, no. 51, pp. 41789–41792, 2005.

[140] M. G. Pray-Grant, J. A. Daniel, D. Schieltz, J. R. Yates, and P. A. Grant, "Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation," *Nature*, vol. 433, no. 7024, pp. 434–438, 2005.

[141] 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.

[142] O. Rozenblatt-Rosen, T. Nagaike, J. M. Francis et al., "The tumor suppressor Cdc73 functionally associates with CPSF and CstF 3′ mRNA processing factors," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 3, pp. 755–760, 2009.

[143] K. Y. Kim and D. E. Levin, "Mpk1 MAPK association with the paf1 complex blocks sen1-mediated premature transcription termination," *Cell*, vol. 144, no. 5, pp. 745–756, 2011.

[144] X. Fan, Z. Moqtaderi, Y. Jin, Y. Zhang, X. S. Liu, and K. Struhl, "Nucleosome depletion at yeast terminators is not intrinsic and can occur by a transcriptional mechanism linked to 3′-end formation," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 42, pp. 17945–17950, 2010.

[145] D. S. Gilmour and R. Fan, "Derailing the locomotive: transcription termination," *Journal of Biological Chemistry*, vol. 283, no. 2, pp. 661–664, 2008.

[146] 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, no. 5, pp. 283–294, 2011.

[147] P. Richard and J. L. Manley, "Transcription termination by nuclear RNA polymerases," *Genes and Development*, vol. 23, no. 11, pp. 1247–1269, 2009.

[148] E. Rosonina, S. Kaneko, and J. L. Manley, "Terminating the transcript: breaking up is hard to do," *Genes and Development*, vol. 20, no. 9, pp. 1050–1056, 2006.

[149] T. Nagaike, C. Logan, I. Hotta, O. Rozenblatt-Rosen, M. Meyerson, and J. L. Manley, "Transcriptional activators enhance polyadenylation of mRNA precursors," *Molecular Cell*, vol. 41, no. 4, pp. 409–418, 2011.

[150] O. Rozenblatt-Rosen, C. M. Hughes, S. J. Nannepaga et al., "The parafibromin tumor suppressor protein is part of

a human Paf1 complex," *Molecular and Cellular Biology*, vol. 25, no. 2, pp. 612–620, 2005.

[151] A. Yart, M. Gstaiger, C. Wirbelauer et al., "The HRPT2 tumor suppressor gene product parafibromin associates with human PAF1 and RNA polymerase II," *Molecular and Cellular Biology*, vol. 25, no. 12, pp. 5052–5060, 2005.

[152] B. Zhu, S. S. Mandal, A. D. Pham et al., "The human PAF complex coordinates transcription with events downstream of RNA synthesis," *Genes and Development*, vol. 19, no. 14, pp. 1668–1673, 2005.

[153] K. Adelman, W. Wei, M. B. Ardehali et al., "Drosophila Paf1 modulates chromatin structure at actively transcribed genes," *Molecular and Cellular Biology*, vol. 26, no. 1, pp. 250–260, 2006.

[154] A. G. Muntean, J. Tan, K. Sitwala et al., "The PAF Complex Synergizes with MLL Fusion Proteins at HOX Loci to Promote Leukemogenesis," *Cancer Cell*, vol. 17, no. 6, pp. 609–621, 2010.

[155] C. Mosimann, G. Hausmann, and K. Basler, "Parafibromin/Hyrax Activates Wnt/Wg Target Gene Transcription by Direct Association with β-catenin/Armadillo," *Cell*, vol. 125, no. 2, pp. 327–341, 2006.

[156] T. Akanuma, S. Koshida, A. Kawamura, Y. Kishimoto, and S. Takada, "Paf1 complex homologues are required for Notch-regulated transcription during somite segmentation," *EMBO Reports*, vol. 8, no. 9, pp. 858–863, 2007.

[157] K. Tenney, M. Gerber, A. Ilvarsonn et al., "Drosophila Rtf1 functions in histone methylation, gene expression, and Notch signaling," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 32, pp. 11970–11974, 2006.

[158] C. T. Nguyen, A. Langenbacher, M. Hsieh, and J. N. Chen, "The PAF1 complex component Leo1 is essential for cardiac and neural crest development in zebrafish," *Developmental Biology*, vol. 341, no. 1, pp. 167–175, 2010.

[159] A. D. Langenbacher, C. T. Nguyen, A. M. Cavanaugh, J. Huang, F. Lu, and J. -N. Chen, "The PAF1 complex differentially regulates cardiomyocyte specification," *Developmental Biology*, vol. 353, no. 1, pp. 19–28, 2011.

[160] M. Y. Youn, H. S. Yoo, M. J. Kim et al., "hCTR9, a component of Paf1 complex, participates in the transcription of interleukin 6-responsive genes through regulation of STAT3-DNA interactions," *Journal of Biological Chemistry*, vol. 282, no. 48, pp. 34727–34734, 2007.

[161] N. Moniaux, C. Nemos, B. M. Schmied et al., "The human homologue of the RNA polymerase II-associated factor 1 (hPaf1), localized on the 19q13 amplicon, is associated with tumorigenesis," *Oncogene*, vol. 25, no. 23, pp. 3247–3257, 2006.

[162] K. Heselmeyer, M. Macville, E. Schröck et al., "Advanced-stage cervical carcinomas are defined by a recurrent pattern of chromosomal aberrations revealing high genetic instability and a consistent gain of chromosome arm 3q," *Genes Chromosomes and Cancer*, vol. 19, no. 4, pp. 233–240, 1997.

[163] A. Kallioniemi, O. P. Kallioniemi, J. Piper et al., "Detection and mapping of amplified DNA sequences in breast cancer by comparative genomic hybridization," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 91, no. 6, pp. 2156–2160, 1994.

[164] K. J. Bradley, B. M. Cavaco, M. R. Bowl et al., "Parafibromin mutations in hereditary hyperparathyroidism syndromes and parathyroid tumours," *Clinical Endocrinology*, vol. 64, no. 3, pp. 299–306, 2006.

[165] J. D. Carpten, C. M. Robbins, A. Villablanca et al., "HRPT2, encoding parafibromin, is mutated in hyperparathyroidism-jaw tumor syndrome," *Nature Genetics*, vol. 32, no. 4, pp. 676–680, 2002.

[166] V. M. Howell, C. J. Haven, K. Kahnoski et al., "HRPT2 mutations are associated with malignancy in sporadic parathyroid tumours," *Journal of Medical Genetics*, vol. 40, no. 9, pp. 657–663, 2003.

[167] J. Tan, A. G. Muntean, and J. L. Hess, "PAFc, a key player in MLL-rearranged leukemogenesis," *Oncotarget*, vol. 1, no. 6, pp. 461–465, 2010.

[168] M. P. Ponnusamy, S. Deb, P. Dey et al., "RNA polymerase II associated factor 1/PD2 maintains self-renewal by its interaction with Oct3/4 in mouse embryonic stem cells," *Stem Cells*, vol. 27, no. 12, pp. 3001–3011, 2009.

[169] L. Xiong, A. Darwanto, S. Sharma et al., "Mass spectrometric studies on epigenetic interaction networks in cell differentiation," *Journal of Biological Chemistry*, vol. 286, no. 15, pp. 13657–13668, 2011.

[170] K. T.G. Rigbolt, T. A. Prokhorova, V. Akimov et al., "System-wide temporal characterization of the proteome and phosphoproteome of human embryonic stem cell differentiation," *Science Signaling*, vol. 4, no. 164, 2011.

[171] M. Imamura, K. Miura, K. Iwabuchi et al., "Transcriptional repression and DNA hypermethylation of a small set of ES cell marker genes in male germline stem cells," *BMC Developmental Biology*, vol. 6, article no. 34, 2006.

[172] B. Sobhian, N. Laguette, A. Yatim et al., "HIV-1 Tat Assembles a Multifunctional Transcription Elongation Complex and Stably Associates with the 7SK snRNP," *Molecular Cell*, vol. 38, no. 3, pp. 439–451, 2010.

[173] S. E. Porter, T. M. Washburn, M. Chang, and J. A. Jaehning, "The yeast Paf1-RNA polymerase II complex is required for full expression of a subset of cell cycle-regulated genes," *Eukaryotic Cell*, vol. 1, no. 5, pp. 830–842, 2002.

[174] E. J. Steinmetz, C. L. Warren, J. N. Kuehner, B. Panbehi, A. Z. Ansari, and D. A. Brow, "Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase," *Molecular Cell*, vol. 24, no. 5, pp. 735–746, 2006.

[175] L. David, W. Huber, M. Granovskaia et al., "A high-resolution map of transcription in the yeast genome," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 14, pp. 5320–5325, 2006.

[176] U. Nagalakshmi, Z. Wang, K. Waern et al., "The transcriptional landscape of the yeast genome defined by RNA sequencing," *Science*, vol. 320, no. 5881, pp. 1344–1349, 2008.

[177] T. H. Kim, L. O. Barrera, M. Zheng et al., "A high-resolution map of active promoters in the human genome," *Nature*, vol. 436, no. 7052, pp. 876–880, 2005.

[178] K. Fejes-Toth, V. Sotirova, R. Sachidanandam et al., "Post-transcriptional processing generates a diversity of 5′-modified long and short RNAs," *Nature*, vol. 457, no. 7232, pp. 1028–1032, 2009.

[179] W. W. Hwang, S. Venkatasubrahmanyam, A. G. Ianculescu, A. Tong, C. Boone, and H. D. Madhani, "A conserved RING finger protein required for histone H2B monoubiquitination and cell size control," *Molecular Cell*, vol. 11, no. 1, pp. 261–266, 2003.

[180] K. Robzyk, J. Recht, and M. A. Osley, "Rad6-dependent ubiquitination of histone H2B in yeast," *Science*, vol. 287, no. 5452, pp. 501–504, 2000.

[181] A. Wood, N. J. Krogan, J. Dover et al., "Bre1, an E3 ubiquitin ligase required for recruitment and substrate selection of

Rad6 at a promoter," *Molecular Cell*, vol. 11, no. 1, pp. 267–274, 2003.

[182] H. H. Ng, R. M. Xu, Y. Zhang, and K. Struhl, "Ubiquitination of histone H2B by Rad6 is required for efficient Dot1-mediated methylation of histone H3 lysine 79," *Journal of Biological Chemistry*, vol. 277, no. 38, pp. 34655–34657, 2002.

[183] J. Dover, J. Schneider, M. A. Tawiah-Boateng et al., "Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6," *Journal of Biological Chemistry*, vol. 277, no. 32, pp. 28368–28371, 2002.

[184] B. D. Strahl, P. A. Grant, S. D. Briggs et al., "Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression," *Molecular and Cellular Biology*, vol. 22, no. 5, pp. 1298–1306, 2002.

SAGE-Hindawi Access to Research
Genetics Research International
Volume 2011, Article ID 625210, 12 pages
doi:10.4061/2011/625210

Review Article

Histone Chaperones Spt6 and FACT: Similarities and Differences in Modes of Action at Transcribed Genes

Andrea A. Duina

Biology Department, Hendrix College, 1600 Washington Avenue, Conway, AR 72032, USA

Correspondence should be addressed to Andrea A. Duina, duina@hendrix.edu

Received 16 June 2011; Accepted 16 August 2011

Academic Editor: David Gross

The process of gene transcription requires the participation of a large number of factors that collectively promote the accurate and efficient expression of an organism's genetic information. In eukaryotic cells, a subset of these factors can control the chromatin environments across the regulatory and transcribed units of genes to modulate the transcription process and to ensure that the underlying genetic information is utilized properly. This article focuses on two such factors—the highly conserved histone chaperones Spt6 and FACT—that play critical roles in managing chromatin during the gene transcription process. These factors have related but distinct functions during transcription and several recent studies have provided exciting new insights into their mechanisms of action at transcribed genes. A discussion of their respective roles in regulating gene transcription, including their shared and unique contributions to this process, is presented.

1. Introduction

In eukaryotic cells, gene transcription takes place in the context of chromatin, a protein-DNA structure that includes the nucleosome—a particle composed of DNA and core histone proteins—as its fundamental unit [1]. The presence of nucleosomes over the regulatory and transcribed regions of genes poses unique problems not encountered by prokaryotic organisms and, as a result, eukaryotic cells have evolved sophisticated mechanisms that enable them to manipulate nucleosomes in a manner that allows for efficient transcriptional control. Different classes of protein factors that can modulate chromatin environments at transcribed loci have been identified and include chromatin remodeling complexes, which can alter DNA-histone interactions in an ATP-dependent manner, histone modifying enzymes, which can modulate the properties of nucleosomes by controlling the set of posttranslational modifications present on the histone proteins within nucleosomes, and histone chaperones, which can interact specifically with histones and can promote the assembly and/or disassembly of nucleosomes in an ATP-independent fashion [2–4]. Several histone chaperones with established roles in transcription have been identified and their contribution to this process is a current area of intense research [3, 5]. This review focuses on two of the better characterized histone chaperones—Spt6 and FACT—and summarizes our current understanding of their roles in the modulation of gene transcription. These highly conserved histone chaperones contribute to the transcription process in several ways and a discussion of the similarities and differences in their mechanisms of action is presented.

2. Identification and Initial Characterization of Spt6 and FACT

2.1. Spt6. *Spt6* was originally identified in the suppressor of Ty (Spt) screens and selection experiments in *Saccharomyces cerevisiae*, which were designed to identify genes that when mutated or expressed at abnormal levels could suppress the deleterious effects of insertions of Ty and solo $\boldsymbol{\delta}$ elements at certain biosynthetic genes [6, 7]. Soon after its initial discovery, it became clear that Spt6 plays essential roles in the control of transcription in yeast and that its role was likely to be genome-wide and not confined to regulatory aspects in the context of Ty and $\boldsymbol{\delta}$ elements [8–13]. Subsequent genetic and biochemical experiments further showed that Spt6 can interact directly with histones (with a preference for histone H3), that it possesses nucleosome assembly activity and that

it is an important player in the transcription elongation process [14–16]. These early experiments pointed to critical roles for Spt6 in the control of transcription through interactions with chromatin and provided the foundations for subsequent studies that have shed further light on the cellular processes impacted by Spt6 and the mechanistic aspects of its activities.

2.2. FACT. Similarly to Spt6, the components of the FACT (FAcilitates Chromatin Transcription) complex were first isolated through genetic and biochemical approaches in the *S. cerevisiae* model system. The gene encoding the first component of the yeast FACT (yFACT) complex, Spt16 (also known as Cdc68), was isolated as an *SPT* gene based on its ability to suppress the transcriptional defects of $\boldsymbol{\delta}$ element insertions at the *LYS2* and *HIS4* genes when expressed from a high copy number plasmid and was shown to encode an essential protein involved in transcription regulation of several genes [17–21]. The second component of yFACT, Pob3, was originally isolated as a protein that copurifies with DNA Polymerase $\boldsymbol{\alpha}$ in biochemical experiments [22] and later shown to also have roles in transcription [23]. Subsequent studies showed that Spt16 and Pob3 form a heterodimer involved in a variety of chromatin-based activities, likely through its ability to interact directly with nucleosomes with the assistance of Nhp6, a protein containing a DNA-binding region similar to the evolutionarily conserved high mobility group (HMG) motif found in several chromatin-interacting proteins [24, 25]. In the literature, the term yFACT has been used to refer to either the Spt16-Pob3 heterodimer or the Spt16-Pob3 dimer in association with Nhp6; for the purpose of this review, the term yFACT will refer to just the Spt16-Pob3 heterodimer, but it is important to keep in mind that the activities ascribed to yFACT are thought to require the participation of Nhp6 as well. The human FACT complex was identified in independent biochemical experiments as an activity required for productive transcription elongation on chromatin templates in *in vitro* reconstitution experiments [26, 27]. These landmark experiments provided critical initial insights into the biochemical properties for which the FACT complex is named. The human FACT complex comprises the homolog of Spt16 (hSpt16) and SSRP1, a protein that combines features of both the yeast Pob3 and Nhp6 proteins, suggesting that during evolution the functions conferred by Pob3 and Nhp6 in yeast have been condensed into a single polypeptide [28].

3. Histone Chaperoning and Transcription Regulation by Spt6 and FACT

3.1. Spt6. The initial discovery of the ability of Spt6 to interact with histones and to assemble nucleosomes *in vitro* [14] foreshadowed the now well-established role for Spt6 as a key histone chaperone during the transcription process. During transcription elongation, Spt6 is required for the maintenance of a chromatin structure that prevents improper usage of cryptic promoter elements, suggesting that the ability of Spt6 to reassemble nucleosomes in the wake of Pol II passage is critical for the prevention of spurious intragenic transcription initiation [29, 30]. The involvement of Spt6 in proper chromatin reconstitution during transcription elongation is also observed at certain activated stress genes [31, 32]. The requirement for Spt6 to reassemble nucleosomes during Pol II elongation, however, does not apply to all genes but instead appears to be associated predominantly with genes that are transcribed at high rates [33]. These findings are consistent with other studies that have proposed that the fate of nucleosomes during transcription elongation depends on the rate of transcription: in the context of high levels of transcription nucleosomes are completely dismantled in front of Pol II and reconstituted in its wake [34–36] whereas in the context of low levels of transcription hexamers devoid of an H2A-H2B dimer remain associated with DNA while still allowing for Pol II passage through a mechanism involving the formation of small DNA loops [35–37]. Thus, it is possible that the chaperoning activity of Spt6 is required in the former case but not in the latter [33], which would be consistent with the fact that Spt6 is not believed to be a histone H2A-H2B chaperone. Interestingly, however, even in the context of low transcription rates during which nucleosome loss is not detected in an *spt6* mutant, Spt6 function can still be required to prevent cryptic intragenic transcription, thus pointing to functions for Spt6 in preventing intragenic transcription that are, at least in some cases, independent from its classical role as a histone chaperone during the transcription elongation process [33].

The chaperoning activity of Spt6 is also required for proper control of transcription initiation. Recent studies in yeast have indicated that nucleosome reassembly activity driven by Spt6 occurs over certain gene promoters and that this activity is required for proper transcriptional repression [33, 38]. The observation that the expression of many genes is affected by an *spt6* mutation [30] but that these genes do not significantly overlap with a set of genes that suffer detectable loss of nucleosomes over their transcribed regions in the same *spt6* mutant background argues that Spt6-driven histone chaperoning activity during transcription elongation does not strongly impact the gene expression process [33]. Therefore, proper transcriptional output likely relies on Spt6-mediated functions that are related to its activity in transcription initiation and/or on Spt6-directed nucleosome reassembly-independent activities during transcription elongation. Finally, Spt6, as well as yFACT, has also been shown to regulate transcription initiation through their histone chaperoning activities during transcription elongation: in this case, transcription of intergenic noncoding DNA that overlaps the promoter of the yeast *SER3* gene is accompanied by Spt6- and yFACT-dependent reassembly of nucleosomes, which, in turn, are thought to prevent binding of transcription activators required for *SER3* expression, ultimately resulting in transcription repression of *SER3* [39].

3.2. FACT. The histone-chaperoning activity of the FACT complex during transcription elongation is required for two distinct, but potentially mechanistically related, processes: facilitation of histone removal in front of elongating Pol II and nucleosome reassembly in the wake of Pol II passage.

Evidence for the former process has come principally from *in vitro* experiments that showed that (i) FACT can interact with nucleosomes and its activity is required for efficient transcription elongation on nucleosomal templates, (ii) the two subunits of FACT can interact with H2A-H2B dimers and $(H3\text{-}H4)_2$ tetramers, and (iii) FACT activity can promote loss of H2A-H2B dimers from nucleosomes [26, 27, 40]. Additional evidence in support of this notion has come from recent experiments showing that mutations predicted to weaken interactions between H2A-H2B dimers and $(H3\text{-}H4)_2$ tetramers can suppress defects in yFACT function both *in vivo* and *in vitro* [41]—thus, histone mutations that favor nucleosome disassembly decrease the dependency on yFACT activity, a finding consistent with a role for FACT in promoting histone eviction. Interestingly, additional studies have shown that yFACT activity is required principally at genes that contain stable nucleosomes positioned over the 5′ end of their coding units, suggesting that FACT-mediated histone eviction at the early stages of transcription elongation is a particularly important event for ensuring proper Pol II progression throughout the length of a transcribed gene [42]. Whereas in some instances it has been speculated that the Spt6 histone chaperone may also possess nucleosome-disruption activity, no direct evidence for such an activity has been reported—therefore, facilitation of histone removal during the transcription elongation process may represent a major difference in the activities of FACT and Spt6.

Similarly to Spt6, the FACT complex has critical roles in the reassembly of nucleosomes following passage of Pol II over transcribed units. Initial evidence for a role for FACT in transcription-dependent nucleosome reassembly came from experiments in yeast showing synthetic lethal genetic interactions between *spt16* mutants and mutations in factors involved in deposition of histones onto DNA [43]. Whereas FACT was originally categorized as a histone H2A-H2B chaperone based on its ability to interact with H2A-H2B dimers but not with H3-H4 tetramers *in vitro* [27], subsequent biochemical experiments showed that human FACT can deposit all four core histones onto DNA *in vitro,* suggesting that part of FACT's *in vivo* function may include participation in nucleosome reassembly during transcription elongation through interactions with all core histones [40]. Recent experiments have provided support for the ability of FACT to interact with histones H3 and H4 and have highlighted the importance of these interactions in promoting histone deposition onto DNA during transcription elongation [34, 44–46]. Moreover, an elegant set of studies has shown that yFACT functions by incorporating the preexisting histones H3 and H4 back onto DNA following Pol II passage, a process with clear implications for the importance of maintenance of epigenetic marks on core histones over transcribed genes [46]. Recent work has implicated the Spt16-M domain, a structural domain originally identified through partial proteolysis experiments [47, 48], in directing histone deposition during transcription elongation [49]. Collectively, these findings establish FACT as a key chaperone for all four core histones during transcription elongation.

What are the consequences of defective FACT-mediated nucleosome reassembly during transcription elongation? Given the shared functions in transcription-dependent chromatin reassembly with Spt6, it is not surprising that mutations in FACT can also result in cryptic transcription initiation defects [29, 30, 49, 50]. However, unlike the case for Spt6 described earlier in which its histone deposition activity does not appear to be required at infrequently transcribed genes, loss of function of the Spt16 component of yFACT does result in nucleosome loss over certain infrequently transcribed genes and even genes expected to be in the "off" state [46]. Thus, it would appear that, at least in certain instances, infrequently and marginally transcribed genes can undergo nucleosome loss and that reassembly of proper nucleosome structure in these cases depends on FACT but not on Spt6, although one cannot exclude the possibility that the differential requirement observed for the two histone chaperones in this context could be due, at least in part, to the different experimental methodologies used in the two studies that addressed this issue [33, 46]. The requirement for FACT but not for Spt6 in nucleosome reassembly over the bodies of infrequently transcribed genes could be explained by a model in which for this class of genes loss of Spt6-mediated histone chaperoning can be compensated by FACT, which can chaperone all four core histones onto DNA whereas loss of FACT activity cannot be compensated by Spt6, which can only chaperone histones H3 and H4. An extension of this model would be that at highly transcribed genes, due to a demand for rapid and/or frequent nucleosome reassembly, the activities of both Spt6 and FACT become essential for maintenance of proper chromatin structure and loss of either one cannot be compensated by the other. Further genetic and biochemical experiments will need to be carried out to test the validity of this model.

An additional consequence of loss of FACT chaperoning activity has been recently described by Chávez and colleagues. In these studies, a failure of yFACT to properly deposit histones during transcription elongation was shown to lead to abnormally high intracellular levels of free histones, which, in turn, led to a delay in cell cycle progression at the G1 phase by repressing expression of a G1-cyclin gene [51]. Therefore, FACT chaperoning activity is critical both for events directly related to chromatin structure at sites where transcriptional elongation is occurring and, in a more indirect fashion, for proper progression through the cell cycle by controlling the proportion of nucleosomal versus nonnucleosomal histones in the cell. Interestingly, a mutation in *SPT6* was also shown to cause phenotypes consistent with excess accumulation of free histones in cells [51], thus raising the possibility that maintenance of proper levels of free histones in cells is a general property shared with other members of the histone chaperone family.

As is the case for Spt6, FACT activity is also required for proper regulation of transcription initiation. Whereas, as described earlier, Spt6 is required for repression of transcription initiation through its ability to directly promote nucleosome reassembly over gene promoter regions [33, 38], FACT-mediated chromatin alterations, including promotion of histone H2A-H2B removal from nucleosomes, have been implicated in activation of transcription initiation at a variety of genes in a number of different species [52–59]. Thus,

FACT and Spt6 can play opposite roles at gene promoters, but both functions are consistent with some of their known biochemical activities—facilitation of histone removal for the FACT complex and reassembly of nucleosomes for Spt6. However, at least in the context of the *SER3* gene as described earlier, yFACT and Spt6 can both repress transcription initiation through their transcription elongation-dependent nucleosome reassembly activities [39].

4. Mechanisms of Spt6 and FACT Histone Chaperoning Activity

4.1. Spt6. The mechanistic details for the interactions that occur between Spt6 and nucleosomes during the chaperoning process are still under investigation, but early and more recent studies have shown that Spt6 can bind double-stranded DNA *in vitro* [60] as well as free histones and nucleosomes and that the interaction between Spt6 and intact nucleosomes requires Nhp6, the same HMG protein described earlier that is also required for interactions between nucleosomes and yFACT [14, 61]. The interaction between Spt6 and nucleosomes is also regulated by the Spt6 binding partner Spn1/Iws1. Structural studies have identified a region located toward the N-terminus of Spt6 and a region located toward the C-terminus of Spn1/Iws1 containing two ARM repeats as being responsible for mediating the Spt6-Spn1/Iws1 interaction and functional studies have indicated that the integrity of this interface is critical for the proper function of the complex [61–63]. Interestingly, binding of Spn1/Iws1 to Spt6 interferes with the ability of Spt6 to interact with nucleosomes [61], thus suggesting that *in vivo* Spn1/Iws1 may assist Spt6 in releasing itself from nucleosomes following nucleosome reassembly. Together, the interactions observed between Spt6 and histones, nucleosomes and naked DNA are likely to represent snapshots of a series of events that normally occur during the Spt6-mediated nucleosome reassembly process in the context of gene transcription.

4.2. FACT. Similarly to Spt6, the yFACT complex requires the assistance of Nhp6 in order to bind to nucleosomes *in vitro* [25]. A series of elegant studies carried out by Formosa and colleagues has shown that several Nhp6 proteins are required to recruit yFACT to nucleosomes and that significant nucleosomal alterations occur upon Nhp6-mediated yFACT binding to nucleosomes [64, 65]. A major question that is still a subject of debate in the field is whether removal of H2A-H2B dimers from nucleosomes is a direct and necessary result of FACT activity or simply one of several potential outcomes [66]. Whereas the original model for FACT activity, which has been referred to as the "dimer eviction model," includes a direct role for the complex in dissociation of single histone H2A-H2B dimers from nucleosomes [66, 67], a more recently presented model, which has been referred to as the "global accessibility/noneviction model," proposes that interaction of FACT with nucleosomes results in the formation of reorganized nucleosomes in which all histone subunits are still tethered together but are in a dynamic structural state more prone to histone H2A-H2B loss [66, 68, 69]. In this latter model, histone H2A-H2B dimer loss from nucleosomes is not a necessary consequence of FACT activity but it is one that can be favored by extrinsic factors such as the force exerted by an oncoming Pol II complex.

Regardless of the exact mechanism, efficient FACT-facilitated eviction of histones during transcription elongation likely requires specific posttranslational histone modifications. In particular, monoubiquitination of histone H2B (H2BK123ub1 in yeast and H2BK120ub1 in mammals) has been shown to prime nucleosomes for FACT-mediated H2A-H2B dimer loss [70]. In addition, several histone acetyl transferase (HAT) complexes have been implicated as positive factors for transcription elongation (e.g., see [71–73]), with one of them, NuA3, having been shown to interact physically and genetically with the FACT complex [71]. Therefore, various histone modifications are likely to play important roles in regulating the efficacy of FACT in histone eviction during transcription elongation *in vivo* and future research will undoubtedly shed more light on the mechanistic details of these processes.

5. Functional Relationships between Spt6 and FACT and Histone Modifications

5.1. Spt6. Both Spt6 and FACT have the ability to influence the chromatin environment across transcribed genes by affecting histone posttranslational modifications. Spt6 activity has recently been linked to methylation of lysine 36 of histone H3 (H3K36me), a modification catalyzed by the Set2 histone methyltransferase associated with the reestablishment of proper nucleosome structure in the wake of Pol II passage through the recruitment of the Rpd3S histone deacetylase complex and subsequent histone deacetylation [74–76]. In the yeast system, a specific mutation in Spt6 leads to reduction in both dimethylation and trimethylation of H3K36 (H3K36me2 and H3K36me3, resp.); however, only the H3K36me3 modification appears to be directly promoted by Spt6 as the reduction in H3K36me2 in the *spt6* mutant appears to be due to an indirect effect resulting from decreased levels of the Set2 protein in the *spt6* mutant background [77, 78]. Interestingly, whereas H3K36me2 has been shown to be required for the prevention of cryptic intragenic transcription initiation through the Rpd3S pathway, H3K36me3 does not appear to be involved in this pathway, thus pointing to roles for Spt6 and Spt6-dependent H3K36me3 in transcription elongation independent from maintenance of proper chromatin structure [78]. A possible role for the Spt6-H3K36me3 pathway has come from studies in mammalian cells. These studies have indicated the existence of a complex bound to elongating Pol II containing Spt6, Iws1, and the HYPB/Set2 histone methyltransferase, which in mammalian cells catalyzes the H3K36me3 modification, and have shown that knockdown of HYPB/Set2 results in accumulation of bulk poly(A)$^+$ mRNA in the nucleus [79]. Thus, Spt6 and its partner Iws1 may promote HYPB/Set2-mediated H3K36me3 to facilitate mRNA nuclear export through a mechanism that has yet to be clearly defined.

5.2. FACT. As indicated earlier, monoubiquitination of histone H2B facilitates FACT-mediated histone eviction during transcription elongation. Interestingly, *in vitro* and *in vivo* experiments have also demonstrated a requirement for FACT activity in promoting H2B monoubiquitination, a role for H2B monoubiquitination in maintenance of FACT at transcribed regions and a cooperative relationship between H2B monoubiquitination and FACT activity in reassembling nucleosomes during transcription elongation [70, 80]. Thus, these studies establish a positive and dynamic relationship between H2B monoubiquitination and FACT activity during transcription elongation. As a testament to the versatility of histone modifications, recent work has shown that monoubiquitination of histone H2A, in stark contrast to histone H2B monoubiquitination, negatively impacts the process of transcription elongation by inhibiting FACT recruitment to chromatin [81]. These findings set the stage for additional studies exploring the potential interplay between additional histone modifications and FACT activity during transcription elongation.

6. Roles for Spt6 and FACT in mRNA Processing and Nuclear Export

In addition to being key contributors to the initiation and elongation phases of transcription, Spt6 and FACT are also involved in functions related to mRNA processing and nuclear export. In mammalian cells, the Pol II-associated complex discussed earlier composed of Spt6, Iws1, and HYPB/Set2 plays important roles in ensuring proper mRNA splicing and efficient mRNA export from the nucleus [79, 82]. The FACT complex has also been shown to participate in the process of mRNA nuclear export [83, 84], with recent experiments showing a direct interaction between the SSRP component of FACT and the mRNA export adaptor UIF [84]. Moreover, experiments performed in yeast have shown that Spt6 can regulate site selection for transcription termination and mRNA 3′ end formation [85, 86] and both Spt6 and Spt16 are required for efficient RNA splicing [87]. Taken together, these findings establish Spt6 and FACT as important players in the processes of mRNA processing and nuclear export and provide insights into the mechanisms that ensure the coordinated execution of the different phases that ultimately lead to the proper formation and localization of mRNA molecules.

7. Mechanisms of Spt6 and FACT Interactions with Transcribed Genes

7.1. Spt6. Pioneering immunofluorescence and biochemical experiments performed in *S. cerevisiae* and *D. melanogaster* provided compelling evidence that, as had been anticipated based on its characteristics as an elongation factor [15], Spt6 physically associates with Pol II and that it interacts with chromatin following patterns of interaction similar to those seen for transcribing Pol II [88–90]. More recent studies have provided additional insights into both the pattern of association of Spt6 across transcribed genes as well as the mechanisms that control its recruitment and association with chromatin.

Chromatin immunoprecipitation (ChIP) assays in yeast have shown that Spt6 interacts with several constitutively expressed genes in a manner similar to that seen for Pol II [91]. Similar results were obtained in genome-wide experiments in yeast that describe Spt6 as a component of a general transcription elongation complex acting at all transcribed genes [92]. In a recent study by the Lis laboratory in which the recruitment of Pol II and several transcription elongation factors were analyzed using a system that allows for a high degree of temporal resolution, Spt6 was shown to associate with the *Hsp70* loci in flies upon heat-shock treatment a few seconds after the recruitment of Pol II to the promoter [93]. Collectively, these experiments establish Spt6 as a general transcription elongation factor that is recruited to activated genes shortly after Pol II recruitment and that travels across transcribed units likely in association with the Pol II complex.

What are the mechanisms that control Spt6 association with transcribed genes? A critical function involved in directing Spt6 interaction with the elongating complex is carried out by a tandem SH2 (tSH2) domain located at the C-terminus of the protein. Analyses of recently solved crystal structures of the Spt6 tSH2 domain derived from different organisms have shown that the overall structure of this domain is evolutionarily conserved and a series of biochemical experiments has shown that this domain mediates interactions between Spt6 and the C-terminal domain (CTD) heptad repeats of Pol II [60, 82, 94, 95]. The interaction between the Spt6 tSH2 domain and the Pol II CTD is direct and requires phosphorylated serine residues on the CTD—more specifically, Ser2-phosphorylation on the CTD appears to be generally required for this interaction [60, 79, 82, 94, 95] whereas an involvement for Ser5-phosphorylation on the CTD, either by itself or in combination with Ser2-phosphorylation, in directing this interaction is less clear since conflicting results have been reported on this issue, likely as a consequence of differences in the assays and/or model systems used in the different studies [60, 79, 82, 94]. Interestingly, mammalian Spt6 is able to discriminate between different regions of the mammalian Pol II CTD and shows specific interactions with the N-terminal half of the CTD [79]. Additional properties of the tSH2 domain of Spt6 have been revealed through fluorescence anisotropy experiments, which have shown that the Spt6 tSH2 domain can also bind to CTD peptides that had been artificially phosphorylated on tyrosine residues present at the first position of the CTD heptad repeats, indicating that the tSH2 domain of Spt6 has phosphotyrosine-binding activity—which is the activity normally associated with SH2 motifs present in certain higher eukaryotic proteins—and raising the intriguing possibility that Spt6 may also specifically bind to target proteins through more canonical SH2-phosphotyrosine interactions [60].

Whereas the direct interaction between the tSH2 domain of Spt6 and Pol II clearly contributes to the association of Spt6 with transcribed genes, it is not the sole mechanism involved in recruitment of Spt6 to active genes. In support

of this notion, a mutation within the SH2 domain of murine Spt6 that lowers the affinity of Spt6 to Ser2-phoshporylated Pol II CTD did not affect transcription output levels in either *in vitro* or *in vivo* assays [82]. Furthermore, Mayer et al. have shown that in yeast the pattern of Pol II CTD Ser-2 phosphorylation across transcribed genes does not correlate with Spt6 occupancy and, more importantly, have reported that a mutant version of Spt6 that lacks the tSH2 domain can still be recruited to the 5′ ends of genes, albeit to a lesser degree than what is seen with wild-type Spt6 [92]. The fact that the Spt6 tSH2 domain is required for optimal recruitment of Spt6 to transcribed genes in these latter experiments suggests that the Spt6 tSH2-Pol II CTD interaction plays a role in the initial recruitment of Spt6 to chromatin: this recruitment may involve interactions between the Spt6 tSH2 and Ser5-phosphorylated versions of the Pol II CTD (which have been reported to occur *in vitro* [60]) or may be mediated through Ser-2 phosphorylated Pol II CTD, which, albeit present at low levels at 5′ ends of genes, could nevertheless recruit Spt6 since, at least in the context of the *Drosophila* Hsp70 genes, arrival of the Ser2 Pol II CTD kinase P-TEFb precedes Spt6 recruitment [93]. In addition to the role in Spt6 recruitment to transcribed genes, the interaction between Spt6 and Pol II mediated by the Spt6 tSH2 and the Ser2-phosphorylated Pol II CTD has been shown to play critical roles in regulating mRNA processing and nuclear export in mammals through Spt6-dependent recruitment of Iws1 and additional factors to nascent RNA molecules as discussed earlier [79, 82].

The observation that deletion of the Spt6 tSH2 domain does not abolish recruitment of Spt6 to transcribed genes indicates that additional mechanisms must exist to ensure proper Spt6 recruitment to 5′ ends of genes. Experiments performed in *Drosophila* and in *S. cerevisiae* provide some insights into the nature of these mechanisms. In *Drosophila*, impairment of the Paf1 complex—a multifunctional complex associated with Pol II that coordinates a variety of transcription-related processes, including recruitment of several transcription factors to genes and various posttranslational modifications of histones [96]—results in lower levels of Spt6 occupancy at the *Hsp70* gene [97]. Whereas the Paf1 complex has been shown to be required for full levels of Ser2 phosphorylation of the Pol II CTD in certain contexts [98, 99], the decrease in Spt6 association at the *Drosophila Hsp70* gene does not appear to be an indirect effect due to lower CTD phosphorylation since at this locus the levels of Ser2-phosphorylation of the Pol II CTD are not affected by the depletion of the Paf1 complex [97]. Thus, at least in certain cases, the Paf1 complex appears to be involved in recruitment of Spt6 to transcribed genes in a manner independent from its role in regulating Pol II modifications. An alternative mode of recruitment of Spt6 to a transcribing gene has been described for the yeast *CYC1* gene. In this case, Spt6 recruitment is dependent on Spn1/Iws1 [100], which has been shown to possess roles in transcription regulation downstream from initial recruitment of the TATA-box binding protein (TBP) to gene promoters [90, 101, 102]. Whether this latter recruitment mechanism is widespread in the yeast genome or is limited to those genes that are regulated at a post-TBP and post-Pol II recruitment step (as is the case for the *CYC1* gene) remains to be more fully elucidated (see Figure 1 for a cartoon depiction of the proposed mechanisms for Spt6 recruitment to active genes).

7.2. FACT. Several lines of evidence have demonstrated that the FACT complex, similarly to Spt6, physically associates with the bodies of transcribed genes *in vivo*. Numerous ChIP studies in yeast have shown specific interactions between yFACT and the transcribed regions of several actively transcribing genes [50, 91, 92, 103] and immunofluorescence and ChIP experiments in *Drosophila* have shown that FACT colocalizes with hyperphosphorylated Pol II at many transcriptionally active loci [104]. Interestingly, these latter experiments also showed that the patterns of association of FACT, Spt6, and Pol II at the activated heat shock gene *hsp70* are similar to one another [104], consistent with the notion that the two histone chaperones function in conjunction with each other to assist Pol II during the transcription elongation process. Comparative ChIP studies in yeast have provided further support for this notion since at certain active genes Spt6, yFACT, and Pol II associate with chromatin following similar patterns as one another [91, 105]. The hypothesis that FACT and Spt6 operate in conjunction with each other and with elongating Pol II during transcription elongation is also supported by several reports that have shown that the two chaperones can be found in the same physical complexes that also contain hyperphosphorylated Pol II [102, 104]. However, FACT and Spt6 are also likely to have roles independent from each other in transcription elongation since a recent genome-wide study has shown that at the global level, yFACT and Spt6 associate with an "average" gene in overlapping but distinct patterns, in which yFACT appears to be recruited at a location slightly more upstream from that used by Spt6 and is released earlier in the elongation process than Spt6 [92]. Thus, whereas it is likely that FACT and Spt6 can operate in conjunction during transcription elongation, their functions do not appear to be always coordinated with each other.

How is the FACT complex recruited to the transcribed regions of genes? Several studies have provided support for a role for the ATP-dependent chromatin remodeling factor Chd1 in recruitment of FACT to actively transcribing genes. Mammalian Chd1 and FACT interact physically with each other and in *Drosophila* they display similar patterns of association across polytene chromosomes [106]. Similarly, experiments carried out in yeast have shown that Chd1 associates with transcribed regions of active genes and that it can be found in complexes containing components of the yFACT complex [90, 107]. Since human Chd1 can associate with histone H3 methylated at lysine 4 (H3K4me3)—a histone modification associated with actively transcribed genes—through its two chromodomains [108–110], it is possible that at least in certain cases Chd1 can direct FACT recruitment to chromatin at sites enriched for the H3K4me3 modification [67, 111]. Strongly supporting this possibility, human FACT and Chd1 can be copurified with

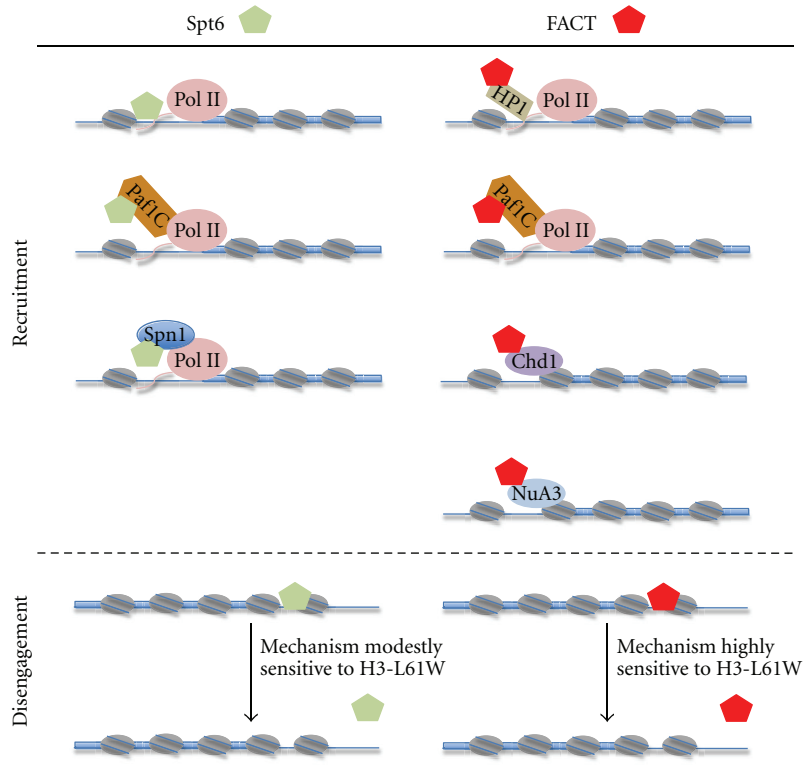


Figure 1: Proposed models for Spt6 and FACT recruitment to and disengagement from actively transcribing genes. *Recruitment (top panel):* Several mechanisms for Spt6 recruitment to active genes have been proposed and include direct interactions with the Pol II CTD (tail extending from Pol II in the figure) and indirect interactions with Pol II through either the Paf1 complex (Paf1C in the figure) or Spn1/Iws1 (Spn1 in the figure) serving as bridging factors. The FACT complex is thought to be able to associate indirectly with Pol II with either HP1 or the Paf1 complex bridging the interaction. The chromatin remodeling factor Chd1 and the histone acetyl transferase complex NuA3 can interact with histones and with FACT and likely promote FACT association with chromatin. *Disengagement (bottom panel):* at some genes, Spt6 and FACT dissociate at distinct locations during the transcription process and, therefore, in these cases each factor must utilize a unique dissociation mechanism (not addressed in this figure). At certain other genes, such as the yeast *PMA1* and *ADH1* genes, Spt6 and FACT depart chromatin at similar locations. However, the nature of the mechanisms used by the two histone chaperones at this class of genes is likely to be at least to some degree different, with Spt6 using a mechanism that is only modestly sensitive to the H3-L61W mutation and FACT using a mechanism that is very sensitive to the H3-L61W mutation. Green pentagons: Spt6; red pentagons: FACT; thin blue lines: DNA regions flanking the coding region of a transcribed gene; thick blue lines: coding region of a transcribed gene; gray ovals with two blue lines: nucleosomes.

H3K4me3-containing peptides with Chd1 being responsible for bridging the histone H3 and FACT interaction [112]. The generality of this model, however, is unclear since recent studies have shown that unlike its human counterpart, yeast Chd1 does not bind to histone H3 peptides methylated at lysine 4 [109, 110, 113]. The relationship between Chd1 and FACT is further complicated by the finding that, at least in certain cases, Chd1 and Spt16 can have opposing functions during transcription [56]. Thus, whereas it is likely that at least in higher eukaryotes Chd1 can directly recruit FACT to sites of active transcription through its ability to interact with H3K4me3-containing nucleosomes, Chd1 and FACT also display additional types of functional interactions that still remain to be more fully elucidated.

FACT recruitment to active genes is also controlled by the Paf1 complex. Evidence in support of this notion includes experiments carried out in yeast that have revealed physical and genetic interactions between components of the Paf1 complex and FACT [90, 114, 115], and studies in flies showing that depletion of Paf1 complex components decreases recruitment of FACT to the activated *Hsp70* gene [97]. As indicated earlier, these latter experiments also showed a requirement for the Paf1 complex in the association of Spt6 to chromatin, thus establishing a potential common route for Spt6 and FACT recruitment to actively transcribing genes. However, a physical relationship between Spt6 and the Paf1 complex has not been as clearly defined in the yeast model system; thus, whether the requirement for the Paf1 complex in efficient recruitment of Spt6 to the *Hsp70* gene in flies reflects a broader physical and functional connection between the two factors remains to be more thoroughly investigated. In addition to directly recruiting FACT to sites of active transcription, the Paf1 complex may also facilitate FACT recruitment through indirect mechanisms stemming from its ability to modulate histone modifications—for example, it could be envisioned that the Paf1 complex

localized to genes via direct interactions with Pol II leads to Paf1 complex-dependent histone modifications, which include H2B monoubiquitination and subsequent histone H3K4me3, which in turn recruits Chd1, ultimately leading to FACT association. The existence of such a pathway involving the Paf1 complex, Chd1, and FACT is consistent with experiments performed in yeast indicating that the three factors genetically interact with each other [114]. The H3K4me3 mark may also facilitate recruitment of the FACT complex through the NuA3 histone acetyl transferase complex since NuA3 can be bind directly to H3K4me3, as well as unmodified, histone H3 tails and, as indicated earlier, NuA3 physically associates with FACT [71, 116, 117]. Unlike the case with Spt6, the FACT complex does not appear to directly interact with Pol II; however, recent work in the fly system by Workman and colleagues has provided strong evidence that the FACT complex can interact with phosphorylated Pol II through a bridging protein, the heterochromatin protein 1 (HP1) [118]. In particular, the HP1c isoform is required for optimal recruitment of FACT to several heat shock loci following heat shock treatment and for normal levels of heat shock transcript levels [118]. Collectively, these studies reveal that different types of mechanisms can contribute to FACT recruitment to transcribed regions of genes and they pave the way for additional studies to assess whether all these mechanisms operate in a coordinated fashion at all loci to ensure optimal FACT recruitment or if different sets of recruiting mechanisms are utilized in a gene-specific fashion (see Figure 1 for a cartoon depiction of proposed mechanisms for FACT recruitment to active genes).

8. Mechanisms of Spt6 and FACT Disengagement from Transcribed Genes

The mechanisms that control the dissociation of transcription elongation factors, including those with histone chaperoning activity, still remain to be elucidated. It seems reasonable to speculate that at least in some instances the factors that are known to interact with Pol II and that disengage from transcribed units at the same locations as Pol II—that is, downstream from the polyadenylation (pA) sites—may simply dissociate from chromatin in conjunction with the Pol II complex. On the other hand, those factors shown to disengage from transcribed genes at or upstream from the pA sites, such as the Paf1 complex [91, 92], must use different mechanisms for dissociation, which may include alterations in Pol II CTD phosphorylation patterns, competition for Pol II binding with other transcription elongation factors—particularly those that are recruited towards the 3′ end of genes, such as Elf1 and termination factors [91, 92]—and conformational changes of Pol II and associated factors that may occur during the elongation process.

Whereas, as indicated earlier, global studies in yeast have indicated that at an "average" gene Spt6 and yFACT disengage from chromatin at different sites [92], suggesting that the two factors normally utilize different mechanisms of chromatin dissociation, at some loci, such as at the yeast *PMA1* and *ADH1* genes, Spt6 and yFACT disengage at similar locations past the pA sites [91, 103, 105] raising the possibility that at these genes Spt6 and yFACT may utilize similar dissociation mechanisms. An insight into this possibility was recently obtained from experiments in which a histone H3 mutant, H3-L61W, was shown to cause a marked accumulation of yFACT at the 3′ ends of transcribed genes in a transcription-dependent fashion [103, 105]. These studies led to a model in which normally yFACT requires a signal, possibly through posttranslational modification of one or more histone proteins, in order to properly dissociate from chromatin following the transcription process, and that the H3-L61W mutation interferes with this signal by preventing either the proper initiation or the propagation of the signal [103]. Interestingly, the same histone H3 mutation was shown not to significantly affect Pol II departure from *PMA1* and to only modestly affect the dissociation of Spt6 from the *PMA1* and *ADH1* genes [103, 105]. Therefore, it appears that even at genes in which Spt6 and yFACT normally disengage from chromatin at the same location, Spt6 and yFACT use distinct mechanisms for chromatin dissociation, with yFACT using an H3-L61W-sensitive mechanism and Spt6 using a mechanism that is significantly less sensitive to the H3-L61W mutation [105]. The Spt16-M domain appears to play a role in controlling yFACT dissociation from chromatin since specific mutations in this region have been shown to alleviate the yFACT 3′ accumulation defect seen in H3-L61W cells [49, 103]. Additional studies will be needed to obtain a more complete understanding of the mechanisms that govern dissociation of FACT and Spt6 from chromatin at the end of the transcription process (see Figure 1 for a cartoon depiction of the proposed mechanisms for Spt6 and FACT disengagement from active genes).

9. Conclusion

Whereas our understanding of the roles and the mechanisms of action of both Spt6 and FACT in the transcription process has increased dramatically since their original identification more than two decades ago, many questions remain to be addressed and future research will undoubtedly provide a more complete picture of the mechanistic details of the transcription process in general, and, more specifically, of the contributions of these two key histone chaperones to this process. In particular, it will be of interest to determine how the functions of Spt6 and FACT are coordinated with those of other histone chaperones known to operate during transcription, such as Asf1 and Nap1 [3, 5]. Traditional genetic and biochemical approaches coupled with more recent genome-wide strategies will continue to provide a powerful experimental platform with which to address these and other relevant questions.

Acknowledgments

The author would like to express his gratitude to Tim Formosa, Joseph Martens, Reine Protacio, and Fred Winston for providing helpful comments during the preparation of this paper and to the National Science Foundation (Grant no. 0919241) for funding.

References

[1] K. Luger, A. W. Mäder, R. K. Richmond, D. F. Sargent, and T. J. Richmond, "Crystal structure of the nucleosome core particle at 2.8 A resolution," *Nature*, vol. 389, no. 6648, pp. 251–260, 1997.

[2] B. R. Cairns, "Chromatin remodeling: insights and intrigue from single-molecule studies," *Nature Structural and Molecular Biology*, vol. 14, no. 11, pp. 989–996, 2007.

[3] M. Eitoku, L. Sato, T. Senda, and M. Horikoshi, "Histone chaperones: 30 years from isolation to elucidation of the mechanisms of nucleosome assembly and disassembly," *Cellular and Molecular Life Sciences*, vol. 65, no. 3, pp. 414–444, 2008.

[4] E. I. Campos and D. Reinberg, "Histones: annotating chromatin," *Annual Review of Genetics*, vol. 43, pp. 559–599, 2009.

[5] N. Avvakumov, A. Nourani, and J. Côté, "Histone chaperones: modulators of chromatin marks," *Molecular Cell*, vol. 41, no. 5, pp. 502–514, 2011.

[6] F. Winston, D. T. Chaleff, B. Valent, and G. R. Fink, "Mutations affecting Ty-mediated expression of the HIS4 gene of Saccharomyces cerevisiae," *Genetics*, vol. 107, no. 2, pp. 179–197, 1984.

[7] F. Winston, "Analysis of SPT genes: a genetic approach toward analysis of TFIID, histones, and other transcription factors in yeast," in *Transcription Regulation*, pp. 1271–1293, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA, 1992.

[8] C. D. Clark-Adams and F. Winston, "The SPT6 gene is essential for growth and is required for δ-mediated transcription in Saccharomyces cerevisiae," *Molecular and Cellular Biology*, vol. 7, no. 2, pp. 679–686, 1987.

[9] L. Neigeborn, J. L. Celenza, and M. Carlson, "SSN20 is an essential gene with mutant alleles that suppress defects in SUC2 transcription in Saccharomyces cerevisiae," *Molecular and Cellular Biology*, vol. 7, no. 2, pp. 672–678, 1987.

[10] C. L. Denis and T. Malvar, "The CCR4 gene from Saccharomyces cerevisiae is required for both nonfermentative and spt-mediated gene expression," *Genetics*, vol. 124, no. 2, pp. 283–291, 1990.

[11] M. S. Swanson and F. Winston, "SPT4, SPT5 and SPT6 interactions: effects on transcription and viability in Saccharomyces cerevisiae," *Genetics*, vol. 132, no. 2, pp. 325–336, 1992.

[12] C. Baniahmad, Z. Nawaz, A. Baniahmad, M. A. G. Gleeson, M. J. Tsai, and B. W. O'Malley, "Enhancement of human estrogen receptor activity by SPT6: a potential coactivator," *Molecular Endocrinology*, vol. 9, no. 1, pp. 34–43, 1995.

[13] P. A. Compagnone-Post and M. A. Osley, "Mutations in the SPT4, SPT5, and SPT6 genes alter transcription of a subset of histone genes in Saccharomyces cerevisiae," *Genetics*, vol. 143, no. 4, pp. 1543–1554, 1996.

[14] A. Bortvin and F. Winston, "Evidence that Spt6p controls chromatin structure by a direct interaction with histones," *Science*, vol. 272, no. 5267, pp. 1473–1476, 1996.

[15] G. A. Hartzog, T. Wada, H. Handa, and F. Winston, "Evidence that Spt4, Spt5, and Spt6 control transcription elongation by RNA polymerase II in Saccharomyces cerevisiae," *Genes and Development*, vol. 12, no. 3, pp. 357–369, 1998.

[16] M. Winkler, T. A. D. Siepen, and T. Stamminger, "Functional interaction between pleiotropic transactivator pUL69 of human cytomegalovirus and the human homolog of yeast chromatin regulatory protein SPT6," *Journal of Virology*, vol. 74, no. 17, pp. 8053–8064, 2000.

[17] C. D. Clark-Adams, D. Norris, M. A. Osley, J. S. Fassler, and F. Winston, "Changes in histone gene dosage alter transcription in yeast," *Genes & development*, vol. 2, no. 2, pp. 150–159, 1988.

[18] J. A. Prendergast, L. E. Murray, A. Rowley, D. R. Carruthers, R. A. Singer, and G. C. Johnston, "Size selection identifies new genes that regulate Saccharomyces cerevisiae cell proliferation," *Genetics*, vol. 124, no. 1, pp. 81–90, 1990.

[19] E. A. Malone, C. D. Clark, A. Chiang, and F. Winston, "Mutations of SPT16/CDC68 suppress cis- and trans-acting mutations that affect promoter function in Saccharomyces cerevisiae," *Molecular and Cellular Biology*, vol. 11, no. 11, pp. 5710–5717, 1991.

[20] A. Rowley, R. A. Singer, and G. C. Johnston, "CDC68, a yeast gene that effects regulation of cell proliferation and transcription, encodes a protein with a highly acidic carboxyl terminus," *Molecular and Cellular Biology*, vol. 11, no. 11, pp. 5718–5726, 1991.

[21] D. Lycan, G. Mikesell, M. Bunger, and L. Breeden, "Differential effects of Cdc68 on cell cycle-regulated promoters in Saccharomyces cerevisiae," *Molecular and Cellular Biology*, vol. 14, no. 11, pp. 7455–7465, 1994.

[22] J. Wittmeyer and T. Formosa, "The Saccharomyces cerevisiae DNA polymerase α catalytic subunit interacts with Cdc68/Spt16 and with Pob3, a protein similar to an HMG1-like protein," *Molecular and Cellular Biology*, vol. 17, no. 7, pp. 4178–4190, 1997.

[23] M. B. Schlesinger and T. Formosa, "POB3 is required for both transcription and replication in the yeast Saccharomyces cerevisiae," *Genetics*, vol. 155, no. 4, pp. 1593–1606, 2000.

[24] N. K. Brewster, G. C. Johnston, and R. A. Singer, "Characterization of the CP complex, an abundant dimer of Cdc68 and Pob3 proteins that regulates yeast transcriptional activation and chromatin repression," *Journal of Biological Chemistry*, vol. 273, no. 34, pp. 21972–21979, 1998.

[25] T. Formosa, P. Eriksson, J. Wittmeyer, J. Ginn, Y. Yu, and D. J. Stillman, "Spt16-Pob3 and the HMG protein Nhp6 combine to form the nucleosome-binding factor SPN," *EMBO Journal*, vol. 20, no. 13, pp. 3506–3517, 2001.

[26] G. Orphanides, G. LeRoy, C. H. Chang, D. S. Luse, and D. Reinberg, "FACT, a factor that facilitates transcript elongation through nucleosomes," *Cell*, vol. 92, no. 1, pp. 105–116, 1998.

[27] G. Orphanides, W. H. Wu, W. S. Lane, M. Hampsey, and D. Reinberg, "The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins," *Nature*, vol. 400, no. 6741, pp. 284–288, 1999.

[28] N. K. Brewster, G. C. Johnston, and R. A. Singer, "A bipartite yeast SSRP1 analog comprised of Pob3 and Nhp6 proteins modulates transcription," *Molecular and Cellular Biology*, vol. 21, no. 10, pp. 3491–3502, 2001.

[29] C. D. Kaplan, L. Laprade, and F. Winston, "Transcription elongation factors repress transcription initiation from cryptic sites," *Science*, vol. 301, no. 5636, pp. 1096–1099, 2003.

[30] V. Cheung, G. Chua, N. N. Batada et al., "Chromatin- and transcription-related factors repress transcription from within coding regions throughout the Saccharomyces cerevisiae genome.," *PLoS Biology*, vol. 6, no. 11, p. e277, 2008.

[31] M. M. Jensen, M. S. Christensen, B. Bonven, and T. H. Jensen, "Requirements for chromatin reassembly during transcriptional downregulation of a heat shock gene in Saccharomyces cerevisiae," *FEBS Journal*, vol. 275, no. 11, pp. 2956–2964, 2008.

[32] E. Klopf, L. Paskova, C. Solé et al., "Cooperation between the INO80 complex and histone chaperones determines adaptation of stress gene transcription in the yeast Saccharomyces cerevisiae," *Molecular and Cellular Biology*, vol. 29, no. 18, pp. 4994–5007, 2009.

[33] I. Ivanovska, P. E. Jacques, O. J. Rando, F. Robert, and F. Winston, "Control of chromatin structure by Spt6: different consequences in coding and regulatory regions," *Molecular and Cellular Biology*, vol. 31, no. 3, pp. 531–541, 2011.

[34] M. A. Schwabish and K. Struhl, "Evidence for eviction and rapid deposition of histones upon transcriptional elongation by RNA polymerase II," *Molecular and Cellular Biology*, vol. 24, no. 23, pp. 10111–10117, 2004.

[35] O. I. Kulaeva, D. A. Gaykalova, and V. M. Studitsky, "Transcription through chromatin by RNA polymerase II: histone displacement and exchange," *Mutation Research*, vol. 618, no. 1-2, pp. 116–129, 2007.

[36] O. I. Kulaeva, F. K. Hsieh, and V. M. Studitsky, "RNA polymerase complexes cooperate to relieve the nucleosomal barrier and evict histones," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 25, pp. 11325–11330, 2010.

[37] M. L. Kireeva, W. Walter, V. Tchernajenko, V. Bondarenko, M. Kashlev, and V. M. Studitsky, "Nucleosome remodeling induced by RNA polymerase II: loss of the H2A/H2B dimer during transcription," *Molecular Cell*, vol. 9, no. 3, pp. 541–552, 2002.

[38] M. W. Adkins and J. K. Tyler, "Transcriptional activators are dispensable for transcription in the absence of Spt6-mediated chromatin reassembly of promoter regions," *Molecular Cell*, vol. 21, no. 3, pp. 405–416, 2006.

[39] S. J. Hainer, J. A. Pruneski, R. D. Mitchell, R. M. Monteverde, and J. A. Martens, "Intergenic transcription causes repression by directing nucleosome assembly," *Genes and Development*, vol. 25, no. 1, pp. 29–40, 2011.

[40] R. Belotserkovskaya, S. Oh, V. A. Bondarenko, G. Orphanides, V. M. Studitsky, and D. Reinberg, "FACT facilitates transcription-dependent nucleosome alteration," *Science*, vol. 301, no. 5636, pp. 1090–1093, 2003.

[41] L. McCullough, R. Rawlins, A. E. Olsen, H. Xin, D. J. Stillman, and T. Formosa, "Insight into the mechanism of nucleosome reorganization from histone mutants that suppress defects in the FACT histone chaperone," *Genetics*, vol. 188, no. 4, pp. 835–646, 2011.

[42] S. Jimeno-González, F. Gómez-Herreros, P. M. Alepuz, and S. Chávez, "A gene-specific requirement for FACT during transcription is related to the chromatin organization of the transcribed region," *Molecular and Cellular Biology*, vol. 26, no. 23, pp. 8710–8721, 2006.

[43] T. Formosa, S. Ruone, M. D. Adams et al., "Defects in SPT16 or POB3 (yFACT) in Saccharomyces cerevisiae cause dependence on the Hir/Hpc pathway: polymerase passage may degrade chromatin structure," *Genetics*, vol. 162, no. 4, pp. 1557–1571, 2002.

[44] T. Stuwe, M. Hothorn, E. Lejeune et al., "The FACT Spt16 "peptidase" domain is a histone H3-H4 binding module," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 26, pp. 8884–8889, 2008.

[45] A. P. Vandemark, H. Xin, L. McCullough et al., "Structural and functional analysis of the Spt16p N-terminal domain reveals overlapping roles of yFACT subunits," *Journal of Biological Chemistry*, vol. 283, no. 8, pp. 5058–5068, 2008.

[46] A. Jamai, A. Puglisi, and M. Strubin, "Histone chaperone Spt16 promotes redeposition of the original H3-H4 histones evicted by elongating RNA polymerase," *Molecular Cell*, vol. 35, no. 3, pp. 377–383, 2009.

[47] A. F. O'Donnell, N. K. Brewster, J. Kurniawan, L. V. Minard, G. C. Johnston, and R. A. Singer, "Domain organization of the yeast histone chaperone FACT: the conserved N-terminal domain of FACT subunit Spt16 mediates recovery from replication stress," *Nucleic Acids Research*, vol. 32, no. 19, pp. 5894–5906, 2004.

[48] A. P. van Demark, M. Blanksma, E. Ferris, A. Heroux, C. P. Hill, and T. Formosa, "The structure of the yFACT Pob3-M domain, its interaction with the DNA replication factor RPA, and a potential role in nucleosome deposition," *Molecular Cell*, vol. 22, no. 3, pp. 363–374, 2006.

[49] C. N. Myers, G. B. Berner, J. H. Holthoff et al., "Mutant versions of the S. cerevisiae transcription elongation factor Spt16 define regions of Spt16 that functionally interact with histone H3," *PLoS ONE*, vol. 6, no. 6, article e20847, 2011.

[50] P. B. Mason and K. Struhl, "The FACT complex travels with elongating RNA polymerase II and is important for the fidelity of transcriptional initiation in vivo," *Molecular and Cellular Biology*, vol. 23, no. 22, pp. 8323–8333, 2003.

[51] M. Morillo-Huesca, D. Maya, M. C. Muñoz-Centeno et al., "FACT prevents the accumulation of free histones evicted from transcribed chromatin and a subsequent cell cycle delay in G1," *PLoS genetics*, vol. 6, no. 5, Article ID e1000964, p. e1000964, 2010.

[52] S. W. Kang, T. Kuzuhara, and M. Horikoshi, "Functional interaction of general transcription initiation factor TFIIE with general chromatin factor SPT16/CDC68," *Genes to Cells*, vol. 5, no. 4, pp. 251–263, 2000.

[53] T. Shimojima, M. Okada, T. Nakayama et al., "Drosophila FACT contributes to Hox gene expression through physical and functional interactions with GAGA factor," *Genes and Development*, vol. 17, no. 13, pp. 1605–1616, 2003.

[54] P. Eriksson, D. Biswas, Y. Yu, J. M. Stewart, and D. J. Stillman, "TATA-binding protein mutants that are lethal in the absence of the Nhp6 high-mobility-group protein," *Molecular and Cellular Biology*, vol. 24, no. 14, pp. 6419–6429, 2004.

[55] D. Biswas, Y. Yu, M. Prall, T. Formosa, and D. J. Stillman, "The yeast FACT complex has a role in transcriptional initiation," *Molecular and Cellular Biology*, vol. 25, no. 14, pp. 5812–5822, 2005.

[56] D. Biswas, R. Dutta-Biswas, and D. J. Stillman, "Chd1 and yFACT act in opposition in regulating transcription," *Molecular and Cellular Biology*, vol. 27, no. 18, pp. 6279–6287, 2007.

[57] T. Kihara, F. Kano, and M. Murata, "Modulation of SRF-dependent gene expression by association of SPT16 with MKL1," *Experimental Cell Research*, vol. 314, no. 3, pp. 629–637, 2008.

[58] M. Ransom, S. K. Williams, M. L. Dechassa et al., "FACT and the proteasome promote promoter chromatin disassembly and transcriptional initiation," *Journal of Biological Chemistry*, vol. 284, no. 35, pp. 23461–23471, 2009.

[59] S. Takahata, Y. Yu, and D. J. Stillman, "FACT and Asf1 regulate nucleosome dynamics and coactivator binding at the HO promoter," *Molecular Cell*, vol. 34, no. 4, pp. 405–415, 2009.

[60] D. Close, S. J. Johnson, M. A. Sdano et al., "Crystal structures of the S. cerevisiae Spt6 core and C-terminal tandem SH2 domain," *Journal of Molecular Biology*, vol. 408, no. 4, pp. 697–713, 2011.

[61] S. M. McDonald, D. Close, H. Xin, T. Formosa, and C. P. Hill, "Structure and biological importance of the Spn1-Spt6

interaction, and its regulatory role in nucleosome binding," *Molecular Cell*, vol. 40, no. 5, pp. 725–735, 2010.

[62] M. L. Diebold, M. Koch, E. Loeliger et al., "The structure of an Iws1/Spt6 complex reveals an interaction domain conserved in TFIIS, Elongin A and Med26," *EMBO Journal*, vol. 29, no. 23, pp. 3979–3991, 2010.

[63] V. Pujari, C. A. Radebaugh, J. V. Chodaparambil et al., "The transcription factor Spn1 regulates gene expression via a highly conserved novel structural motif," *Journal of Molecular Biology*, vol. 404, no. 1, pp. 1–15, 2010.

[64] S. Ruone, A. R. Rhoades, and T. Formosa, "Multiple Nhp6 molecules are required to recruit Spt16-Pob3 to form yFACT complexes and to reorganize nucleosomes," *Journal of Biological Chemistry*, vol. 278, no. 46, pp. 45288–45295, 2003.

[65] A. R. Rhoades, S. Ruone, and T. Formosa, "Structural Features of Nucleosomes Reorganized by Yeast FACT and Its HMG Box Component, Nhp6," *Molecular and Cellular Biology*, vol. 24, no. 9, pp. 3907–3917, 2004.

[66] D. D. Winkler and K. Luger, "The histone chaperone FACT: structural insights and mechanisms for nucleosome reorganization," *Journal of Biological Chemistry*, vol. 286, no. 21, pp. 18369–18374, 2011.

[67] D. Reinberg and R. J. Sims III, "De FACTo nucleosome dynamics," *Journal of Biological Chemistry*, vol. 281, no. 33, pp. 23297–23301, 2006.

[68] T. Formosa, "FACT and the reorganized nucleosome," *Molecular BioSystems*, vol. 4, no. 11, pp. 1085–1093, 2008.

[69] H. Xin, S. Takahata, M. Blanksma, L. McCullough, D. J. Stillman, and T. Formosa, "yFACT induces global accessibility of nucleosomal DNA without H2A-H2B displacement," *Molecular Cell*, vol. 35, no. 3, pp. 365–376, 2009.

[70] R. Pavri, B. Zhu, G. Li et al., "Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II," *Cell*, vol. 125, no. 4, pp. 703–717, 2006.

[71] S. John, L. Howe, S. T. Tafrov, P. A. Grant, R. Sternglanz, and J. L. Workman, "The something about silencing protein, Sas3, is the catalytic subunit of NuA3, a yTAF(II)30-containing HAT complex that interacts with the Spt16 subunit of the yeast CP (Cdc68/Pob3)-FACT complex," *Genes and Development*, vol. 14, no. 10, pp. 1196–1208, 2000.

[72] C. K. Govind, F. Zhang, H. Qiu, K. Hofmeyer, and A. G. Hinnebusch, "Gcn5 promotes acetylation, eviction, and methylation of nucleosomes in transcribed coding regions," *Molecular Cell*, vol. 25, no. 1, pp. 31–42, 2007.

[73] D. S. Ginsburg, C. K. Govind, and A. G. Hinnebusch, "NuA4 lysine acetyltransferase Esa1 is targeted to coding regions and stimulates transcription elongation with Gcn5," *Molecular and Cellular Biology*, vol. 29, no. 24, pp. 6473–6487, 2009.

[74] M. J. Carrozza, B. Li, L. Florens et al., "Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription," *Cell*, vol. 123, no. 4, pp. 581–592, 2005.

[75] 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.

[76] 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.

[77] Y. Chu, A. Sutton, R. Sternglanz, and G. Prelich, "The Bur1 cyclin-dependent protein kinase is required for the normal pattern of histone methylation by Set2," *Molecular and Cellular Biology*, vol. 26, no. 8, pp. 3029–3038, 2006.

[78] 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.

[79] 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.

[80] A. B. Fleming, C. F. Kao, C. Hillyer, M. Pikaart, and M. A. Osley, "H2B ubiquitylation plays a role in nucleosome dynamics during transcription elongation," *Molecular Cell*, vol. 31, no. 1, pp. 57–66, 2008.

[81] W. Zhou, P. Zhu, J. Wang et al., "Histone H2A monoubiquitination represses transcription by inhibiting RNA polymerase II transcriptional elongation," *Molecular Cell*, vol. 29, no. 1, pp. 69–80, 2008.

[82] 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.

[83] A. Herold, L. Teixeira, and E. Izaurralde, "Genome-wide analysis of nuclear mRNA export pathways in Drosophila," *EMBO Journal*, vol. 22, no. 10, pp. 2472–2483, 2003.

[84] G. M. Hautbergue, M. L. Hung, M. J. Walsh et al., "UIF, a new mRNA export adaptor that works together with REF/ALY, requires FACT for recruitment to mRNA," *Current Biology*, vol. 19, no. 22, pp. 1918–1924, 2009.

[85] C. D. Kaplan, M. J. Holland, and F. Winston, "Interaction between transcription elongation factors and mRNA 3′-end formation at the Saccharomyces cerevisiae GAL10-GAL7 locus," *Journal of Biological Chemistry*, vol. 280, no. 2, pp. 913–922, 2005.

[86] M. E. Bucheli and S. Buratowski, "Npl3 is an antagonist of mRNA 3′ end formation by RNA polymerase II," *EMBO Journal*, vol. 24, no. 12, pp. 2150–2160, 2005.

[87] T. Burckin, R. Nagel, Y. Mandel-Gutfreund et al., "Exploring functional relationships between components of the gene expression machinery," *Nature Structural and Molecular Biology*, vol. 12, no. 2, pp. 175–182, 2005.

[88] E. D. Andrulis, E. Guzmán, P. Döring, J. Werner, and J. T. Lis, "High-resolution localization of Drosophila Spt5 and Spt6 at heat shock genes in vivo: roles in promoter proximal pausing and transcription elongation," *Genes and Development*, vol. 14, no. 20, pp. 2635–2649, 2000.

[89] C. D. Kaplan, J. R. Morris, C. T. Wu, and F. Winston, "Spt5 and Spt6 are associated with active transcription and have characteristics of general elongation factors in D. melanogaster," *Genes and Development*, vol. 14, no. 20, pp. 2623–2634, 2000.

[90] N. J. Krogan, M. Kim, S. H. Ahn et al., "RNA polymerase II elongation factors of Saccharomyces cerevisiae: a targeted proteomics approach," *Molecular and Cellular Biology*, vol. 22, no. 20, pp. 6979–6992, 2002.

[91] M. Kim, S. H. Ahn, N. J. Krogan, J. F. Greenblatt, and S. Buratowski, "Transitions in RNA polymerase II elongation complexes at the 3' ends of genes," *EMBO Journal*, vol. 23, no. 2, pp. 354–364, 2004.

[92] 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.

[93] K. L. Zobeck, M. S. Buckley, W. R. Zipfel, and J. T. Lis, "Recruitment Timing and Dynamics of Transcription Factors

at the Hsp70 Loci in Living Cells," *Molecular Cell*, vol. 40, no. 6, pp. 965–975, 2010.

[94] M. L. Diebold, E. Loeliger, M. Koch, F. Winston, J. Cavarelli, and C. Romier, "Noncanonical tandem SH2 enables interaction of elongation factor Spt6 with RNA polymerase II," *Journal of Biological Chemistry*, vol. 285, no. 49, pp. 38389–38398, 2010.

[95] 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.

[96] J. A. Jaehning, "The Paf1 complex: platform or player in RNA polymerase II transcription?" *Biochimica et Biophysica Acta*, vol. 1799, no. 5-6, pp. 379–388, 2010.

[97] K. Adelman, W. Wei, M. B. Ardehali et al., "Drosophila Paf1 modulates chromatin structure at actively transcribed genes," *Molecular and Cellular Biology*, vol. 26, no. 1, pp. 250–260, 2006.

[98] C. L. Mueller, S. E. Porter, M. G. Hoffman, and J. A. Jaehning, "The Paf1 complex has functions independent of actively transcribing RNA polymerase II," *Molecular Cell*, vol. 14, no. 4, pp. 447–456, 2004.

[99] K. Nordick, M. G. Hoffman, J. L. Betz, and J. A. Jaehning, "Direct interactions between the Paf1 complex and a cleavage and polyadenylation factor are revealed by dissociation of Paf1 from RNA polymerase II," *Eukaryotic Cell*, vol. 7, no. 7, pp. 1158–1167, 2008.

[100] L. Zhang, A. G. L. Fletcher, V. Cheung, F. Winston, and L. A. Stargell, "Spn1 regulates the recruitment of Spt6 and the Swi/Snf complex during transcriptional activation by RNA polymerase II," *Molecular and Cellular Biology*, vol. 28, no. 4, pp. 1393–1403, 2008.

[101] J. A. Fischbeck, S. M. Kraemer, and L. A. Stargell, "SPN1, a conserved gene identified by suppression of a postrecruitment-defective yeast TATA-binding protein mutant," *Genetics*, vol. 162, no. 4, pp. 1605–1616, 2002.

[102] D. L. Lindstrom, S. L. Squazzo, N. Muster et al., "Dual roles for Spt5 in pre-mRNA processing and transcription elongation revealed by identification of Spt5-associated proteins," *Molecular and Cellular Biology*, vol. 23, no. 4, pp. 1368–1378, 2003.

[103] A. A. Duina, A. Rufiange, J. Bracey, J. Hall, A. Nourani, and F. Winston, "Evidence that the localization of the elongation factor Spt16 across transcribed genes is dependent upon histone H3 integrity in Saccharomyces cerevisiae," *Genetics*, vol. 177, no. 1, pp. 101–112, 2007.

[104] A. Saunders, J. Werner, E. D. Andrulis et al., "Tracking FACT and the RNA polymerase II elongation complex through chromatin in vivo," *Science*, vol. 301, no. 5636, pp. 1094–1096, 2003.

[105] A. Lloyd, K. Pratt, E. Siebrasse, M. D. Moran, and A. A. Duina, "Uncoupling of the patterns of chromatin association of different transcription elongation factors by a histone H3 mutant in Saccharomyces cerevisiae," *Eukaryotic Cell*, vol. 8, no. 2, pp. 257–260, 2009.

[106] D. E. Kelley, D. G. Stokes, and R. P. Perry, "CHD1 interacts with SSRP1 and depends on both its chromodomain and its ATPase/helicase-like domain for proper association with chromatin," *Chromosoma*, vol. 108, no. 1, pp. 10–25, 1999.

[107] R. Simic, D. L. Lindstrom, H. G. Tran et al., "Chromatin remodeling protein Chd1 interacts with transcription elongation factors and localizes to transcribed genes," *EMBO Journal*, vol. 22, no. 8, pp. 1846–1856, 2003.

[108] J. F. Flanagan, L. Z. Mi, M. Chruszcz et al., "Double chromodomains cooperate to recognize the methylated histone H3 tail," *Nature*, vol. 438, no. 7071, pp. 1181–1185, 2005.

[109] R. J. Sims, C. F. Chen, H. Santos-Rosa, T. Kouzarides, S. S. Patel, and D. Reinberg, "Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains," *Journal of Biological Chemistry*, vol. 280, no. 51, pp. 41789–41792, 2005.

[110] J. F. Flanagan, B. J. Blus, D. Kim, K. L. Clines, F. Rastinejad, and S. Khorasanizadeh, "Molecular implications of evolutionary differences in CHD double chromodomains," *Journal of Molecular Biology*, vol. 369, no. 2, pp. 334–342, 2007.

[111] S. K. Williams and J. K. Tyler, "Transcriptional regulation by chromatin disassembly and reassembly," *Current Opinion in Genetics and Development*, vol. 17, no. 2, pp. 88–93, 2007.

[112] R. J. Sims III, S. Millhouse, C. F. Chen et al., "Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing," *Molecular Cell*, vol. 28, no. 4, pp. 665–676, 2007.

[113] M. Okuda, M. Horikoshi, and Y. Nishimura, "Structural polymorphism of chromodomains in Chd1," *Journal of Molecular Biology*, vol. 365, no. 4, pp. 1047–1062, 2007.

[114] P. J. Costa and K. M. Arndt, "Synthetic lethal interactions suggest a role for the Saccharomyces cerevisiae Rtf1 protein in transcription elongation," *Genetics*, vol. 156, no. 2, pp. 535–547, 2000.

[115] S. L. Squazzo, P. J. Costa, D. L. Lindstrom et al., "The Paf1 complex physically and functionally associates with transcription elongation factors in vivo," *EMBO Journal*, vol. 21, no. 7, pp. 1764–1774, 2002.

[116] D. G. E. Martin, D. E. Grimes, K. Baetz, and L. Howe, "Methylation of histone H3 mediates the association of the NuA3 histone acetyltransferase with chromatin," *Molecular and Cellular Biology*, vol. 26, no. 8, pp. 3018–3028, 2006.

[117] A. Chruscicki, V. E. MacDonald, B. P. Young, C. J. R. Loewen, and L. J. Howe, "Critical determinants for chromatin binding by Saccharomyces cerevisiae Yng1 exist outside of the plant homeodomain finger," *Genetics*, vol. 185, no. 2, pp. 469–477, 2010.

[118] 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.

SAGE-Hindawi Access to Research
Genetics Research International
Volume 2011, Article ID 623718, 16 pages
doi:10.4061/2011/623718

Review Article

Updating the CTD Story: From Tail to Epic

Bartlomiej 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

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

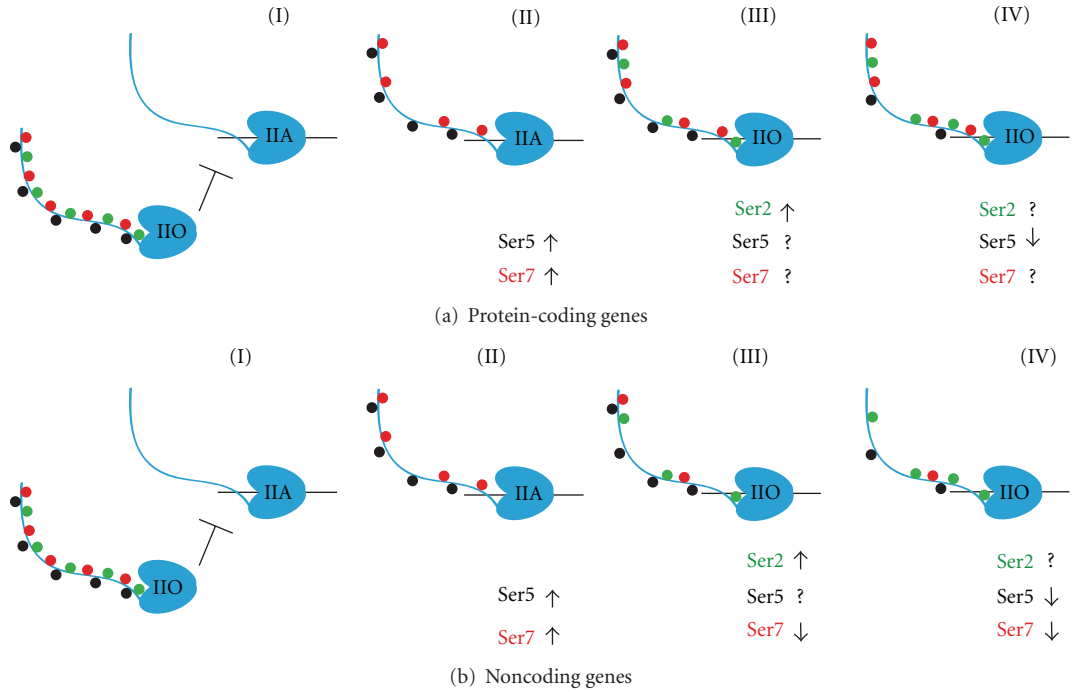


Figure 1: Revised "phospho-CTD cycle." (a) (I) For protein-coding genes, RNAPII is recruited to the promoter with an unphosphorylated CTD (IIA form); moreover, it appears that CTD phosphorylation prior to preinitiation complex (PIC) formation has an inhibitory effect on transcription. (II) Upon preinitiation 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 TFIIH: 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 splicesome 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 eoccupancy 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 Δ*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.

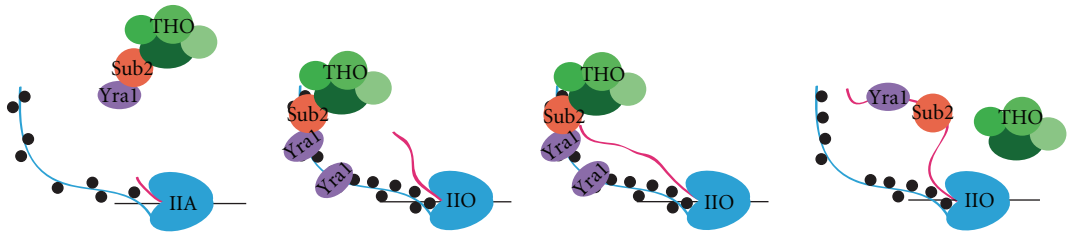


Figure 2: CTD-mediated recruitment of Yra1. Yra1 (possibly in complex with THO and Sub2) does not bind to Ser5P containing CTD and is not observed at high level at the 5′ end of genes. Yra1 is recruited to active genes during elongation when the CTD is doubly phosphorylated on Ser2 and Ser5. As it is 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* (*ctk1*Δ/*ctk1*Δ 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. Hoeppner, 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. Rodríguez-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. Rossettini, 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. Larochelle, 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. Søgaard, and J. Q. Svejstrup, "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. Chuikov, 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. Egloff, 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. Egloff, 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., "PITSLRE 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 RecQ5β 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.

SAGE-Hindawi Access to Research
Genetics Research International
Volume 2011, Article ID 726901, 9 pages
doi:10.4061/2011/726901

Review Article

The Role of RNA Polymerase II Elongation Control in HIV-1 Gene Expression, Replication, and Latency

Kyle A. Nilson[1] and David H. Price[1,2]

[1] *Molecular and Cellular Biology Program, The University of Iowa, Iowa City, IA 52242, USA*
[2] *Department of Biochemistry, The University of Iowa, Iowa City, IA 52242, USA*

Correspondence should be addressed to David H. Price, david-price@uiowa.edu

Received 5 July 2011; Accepted 22 July 2011

Academic Editor: Carles Sune

HIV-1 usurps the RNA polymerase II elongation control machinery to regulate the expression of its genome during lytic and latent viral stages. After integration into the host genome, the HIV promoter within the long terminal repeat (LTR) is subject to potent downregulation in a postinitiation step of transcription. Once produced, the viral protein Tat commandeers the positive transcription elongation factor, P-TEFb, and brings it to the engaged RNA polymerase II (Pol II), leading to the production of viral proteins and genomic RNA. HIV can also enter a latent phase during which factors that regulate Pol II elongation may play a role in keeping the virus silent. HIV, the causative agent of AIDS, is a worldwide health concern. It is hoped that knowledge of the mechanisms regulating the expression of the HIV genome will lead to treatments and ultimately a cure.

1. Introduction

According to the 2010 UNAIDS AIDS Epidemic Update, over 33 million people live with human immunodeficiency virus (HIV) type 1, a number that is increasing due to a combination of improved treatment and continued transmission. Upon crossing the mucosa, HIV docks with $CD4^+$ cells such as T-lymphocytes and macrophages, fuses with the host cell, and releases viral single-stranded RNA, reverse transcriptase, and integrase into the cytoplasm. Reverse transcriptase converts the HIV RNA into double-stranded DNA, at which point integrase chaperones the viral DNA into the nucleus for integration into the host genome. An initial round of host-induced gene expression by Pol II results in expression of Tat, the primary transactivator of HIV, which then recruits the positive transcription elongation factor P-TEFb containing Cdk9 and Cyclin T1 to the HIV LTR [1, 2]. This leads to increased viral gene expression and, eventually, replication of the HIV genome, assembly into new viral particles, and budding. HIV is capable of establishing life-long latent infection by suppressing its transcription, thus evading current antiretroviral therapies [3]. How HIV subverts Pol II elongation control during both active and latent infections has received a significant amount of attention, and it is hoped that these inquires will lead to the development of more effective treatments and an eventual cure.

Regulation of transcription of many human genes is accomplished by a process termed RNA polymerase II elongation control, and, after integration, the HIV LTR falls under this control. In fact, the HIV LTR has been used as a model to study the regulation of transcription at the level of elongation. In general, most human genes experience initiation, but the fraction of those initiation events that result in mRNAs is tightly regulated. After initiation, Pol II is directed by negative elongation factors that include DSIF and NELF to pause after synthesizing approximately 30–100 nucleotides of RNA [4]. These promoter proximally paused polymerases either prematurely terminate, or enter productive elongation under the influence of P-TEFb, thereby generating mRNAs or in the case of HIV, viral genomes [5]. Because of its important role in this process, the activity of P-TEFb is restricted by reversible association with 7SK snRNA-bound HEXIM1 or HEXIM2 proteins which inhibit the kinase activity of P-TEFb during its residence within the 7SK snRNP [6, 7]. A number of cellular activators

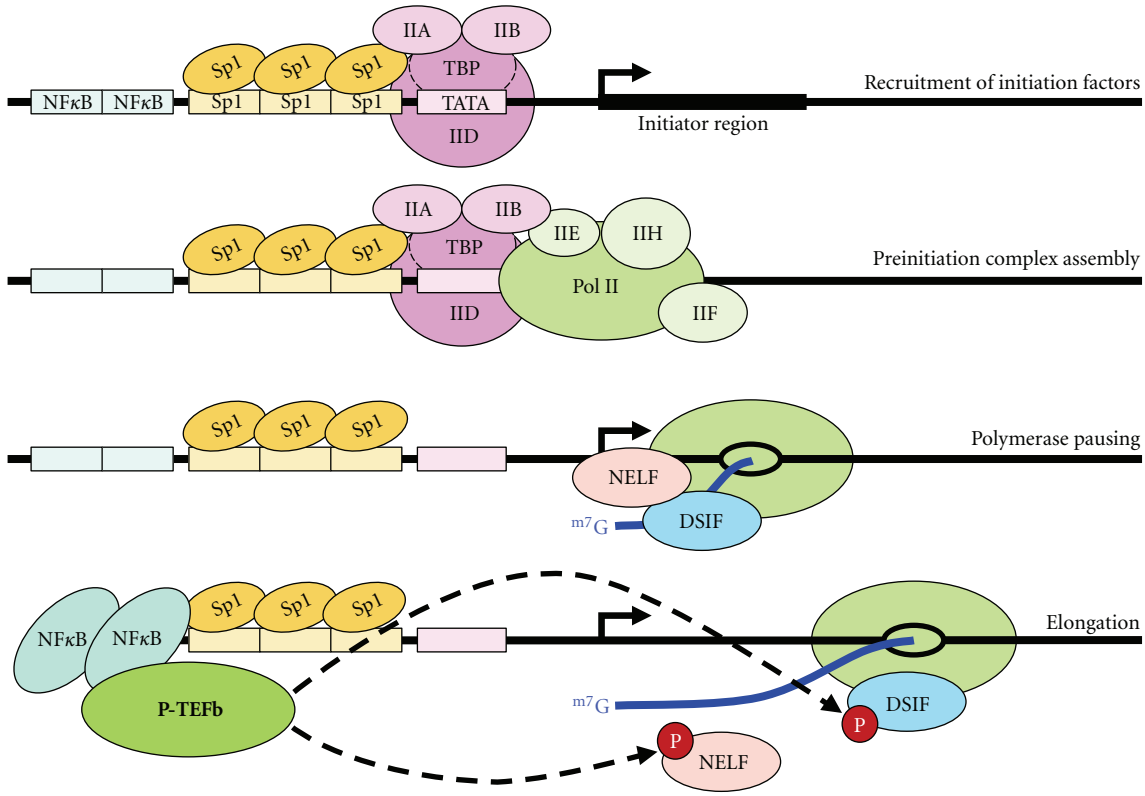


Figure 1: Early rounds of HIV transcription. At the HIV LTR, TATA-binding protein (TBP) is recruited to the TATA box with the aid of Sp1. This interaction is subsequently stabilized by TFIIA and TFIIB. Preinitiation complex assembly is completed by the sequential addition of Pol II•TFIIF, TFIIE, and TFIIH and is followed by promoter clearance. Pol II quickly falls under the negative influence of DSIF and NELF, pausing after the generation of a short, nascent transcript. The recruitment of P-TEFb through NFκB or other factors overcomes this inhibition by phosphorylating DSIF and NELF, allowing for productive elongation and the generation of HIV Tat.

including Brd4 [8, 9], c-Myc [10–12], NFκB [13], and others interact with and recruit P-TEFb to overcome this hurdle to transcription [14]. Recent ChIP-Seq experiments have revealed that promoter proximally paused polymerases are a prominent feature of chromatin [10] highlighting the relevance of elongation control to human transcription and disease. A growing body of evidence suggests that not only is HIV regulated by elongation control, but that the virus manipulates the machinery that regulates P-TEFb for the purposes of viral gene expression, replication, and latency.

2. The First Steps of HIV Gene Expression

The HIV LTR region is composed of several redundant elements that promote the swift and spontaneous assembly of the preinitiation complex (PIC) [15] (Figure 1). While most eukaryotic core promoters contain either a TATA box or a pyrimidine-rich initiator region, HIV plays host to both elements, encouraging the recruitment of transcription factors. Three tandem-repeat specificity protein 1 (Sp1) sites further promote PIC assembly and are indispensible in HIV transcription [16]. Sp1 stimulates the recruitment of TATA-binding protein (TBP), a subunit of TFIID, to the TATA box; this interaction is immediately stabilized by TFIIA and TFIIB (Figure 1, first panel). The subsequent assembly of Pol II•TFIIF, TFIIE, and TFIIH completes the formation of the PIC, allowing for initiation and promoter clearance [17] (Figure 1, second panel). In terms of initiation efficiency the HIV LTR is one of the strongest promoters known.

Immediately following initiation in the 5′ LTR, Pol II falls under the influence of negative elongation factors that include DSIF and NELF [18], pauses, and produces only short transcripts [19] (Figure 1, third panel). NELF has been shown to interact directly with the HIV nascent transcript TAR to inhibit elongation [20]. DSIF and NELF further encourage the formation of poised polymerases by inhibiting transcript cleavage factor TFIIS [21], but this factor may be needed ultimately to restart elongation [22]. In contrast to the highly efficient initiation that takes place on the LTR, the escape of promoter proximally paused polymerases into productive elongation is very inefficient. This is due to a combination of the strong interaction of NELF with the nascent HIV transcript [20] and potentially to the affinity of HEXIM proteins to the same RNA structure [23]. Due to this strong negative inhibition, HIV is incapable of productive elongation without elevated levels of P-TEFb. The reduction of available P-TEFb through inhibitors or expression of kinase dead Cdk9 mutant blocks HIV gene expression while leaving overall cellular transcription relatively unaffected [24–27]. HIV achieves its initial rounds of productive

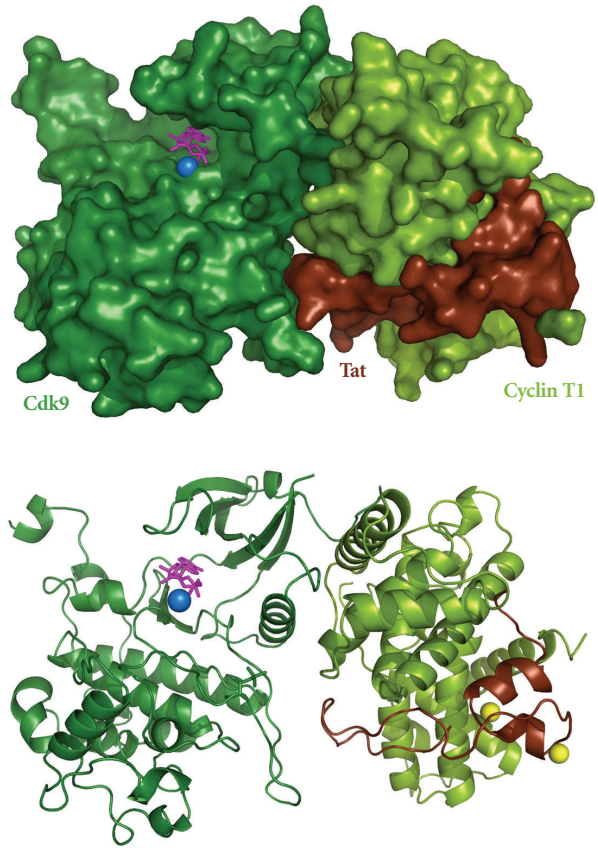


Figure 2: Structural detail of Tat•P-TEFb interaction. The two panels depict surface (upper) and cartoon (lower) representations of the crystal structure of the complex between HIV Tat and P-TEFb. Tat is full length 86 amino acid protein, but only residues 1–49 are visible. Cdk9 (dark green) is a 1-345 truncation of the 372 amino acid protein. Cyclin T1 (light green) is a 1–266 truncation of the 726 amino acid protein. The two zinc atoms coordinated by the interface of Tat and cyclin T1 are indicated as yellow spheres. ATP (magenta) and Mg (blue) are shown bound to the active site of Cdk9. The figure was created using Pymol from PDB entry 3MIA [34].

elongation through lymphocyte cellular activation, which increases Cyclin T1 expression and P-TEFb activity to levels sufficient for HIV gene expression [28].

T-cell activation triggers the activity of transcription factor NFκB, composed of subunits p50 and p65 (RelA), and NFAT. In response to a wide range of stimuli including TNF-α, NFκB inhibitor IκB becomes phosphorylated and subsequently degraded, allowing NFκB to translocate into the nucleus and localize to genomic binding sites [29], where the p65 subunit can be found in association with P-TEFb [13]. Two NFκB sites in close proximity to the Sp1 binding elements in the HIV LTR are shown to strongly enhance HIV transcription [30]; synergy between the p65 subunit and Sp1 further augments the shift from abortive to productive elongation in HIV [30] (Figure 1, fourth panel). Two additional sites downstream of the transcription start site also enhance HIV's transcriptional sensitivity to NFκB [31], though their role is less well understood. The nuclear factor of activated T cells, NFAT, likely plays a similar role [32] binding as a dimer to the same DNA binding elements [33]. These sites allow for the recruitment of enough P-TEFb to the HIV LTR to phosphorylate DSIF and NELF and trigger at least some poised polymerases to enter productive elongation, and this results in the production of Tat and the entry into the second phase of HIV transcription [13].

3. Maintenance of Highly Efficient HIV Gene Expression and Viral Replication

During the next stage of HIV infection an extremely high level of transcription of the viral genome is directed by Tat, the major transactivator of HIV transcription. Tat is an HIV protein designed for direct interaction with P-TEFb, mainly with Cyclin T1, but also with Cdk9 [1, 2, 24, 34]. Tat is required for efficient productive elongation of HIV genes [1, 19, 35, 36], and this stimulatory effect depends on P-TEFb [24–27]. Tat's ability to interact with P-TEFb (Figure 2) allows it to extract the kinase from the 7SK snRNP and bring it to poised polymerases on the LTR (Figure 3). This activates HIV transcription by exploiting the ability of P-TEFb to stimulate productive elongation of HIV-bound polymerases, allowing for effective viral gene expression and replication.

X-ray crystallography has proven useful in clarifying Tat's association with P-TEFb (Figure 2). Tat lacks a prominent secondary structure when free in solution, but upon interacting with P-TEFb, peptides 1–49 become highly organized

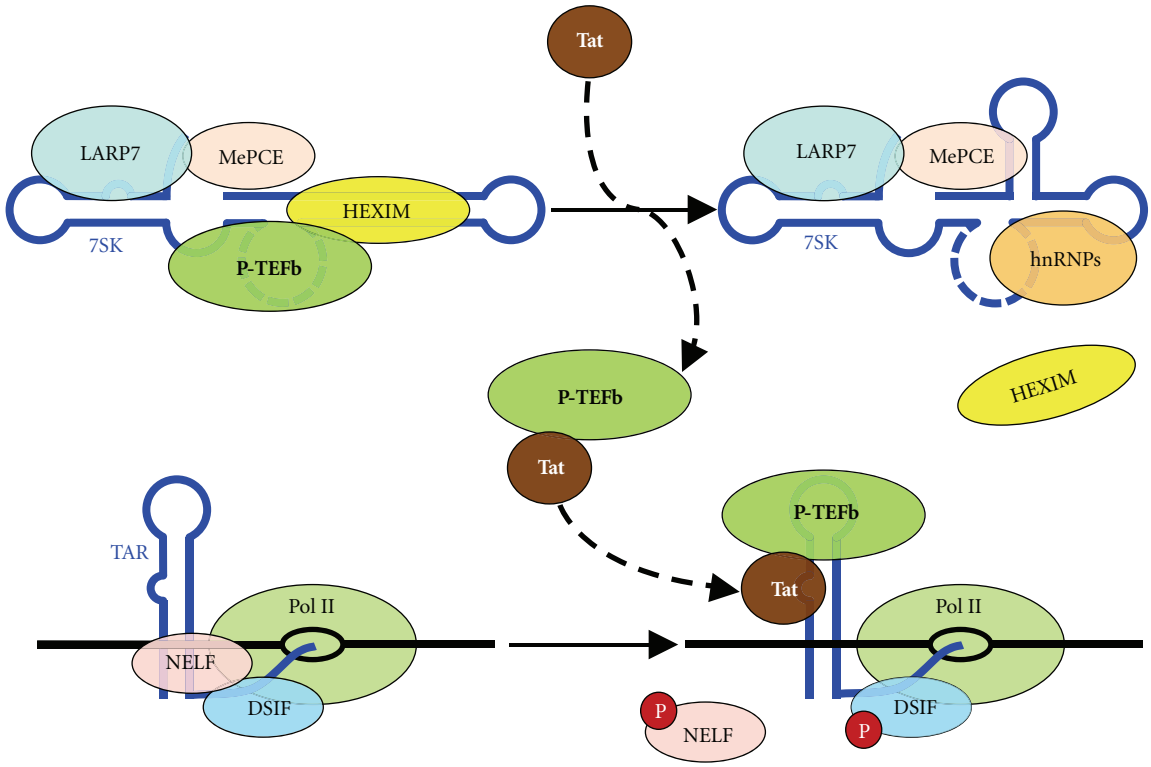


Figure 3: Tat overrides cellular elongation control. Tat interacts directly with P-TEFb sequestered in the 7SK snRNP. This results in a conformational change of 7SK, ejecting HEXIM proteins, and preventing P-TEFb•HEXIM1•7SK reassembly. Tat•P-TEFb migrates to the TAR element contained within HIV's nascent transcript, binds, and acts upon DSIF and NELF, efficiently overcoming inhibition of transcription.

and form a structure complimentary to the kinase [34]. Tat interacts primarily with Cyclin T1, using 88% of its surface area and a Zn-mediated bridge to stabilize the interface [34]. Tat also inserts into a groove between Cyclin T1 and Cdk9, resulting in a more stable and active P-TEFb kinase [34]. While comparisons of Tat•P-TEFb•ATP to a previous P-TEFb•ATP structure [37] suggested Tat altered the conformation of P-TEFb, it is possible that the differences in structures were due to 3 amino acid substitutions in the P-TEFb•ATP crystal, one of which lies in the critical surface between Cdk9 and Cyclin T1. What is clear is that Tat binds specifically to Cyclin T1 and forms a very stable complex. In fact, when Tat is overexpressed in HeLa cells, greater than 90%, of the Tat is found associated with P-TEFb [23]. Evidently excess Tat is degraded.

Tat's most striking behavior is its ability to recruit sequestered P-TEFb from the 7SK snRNP (Figure 3, top), subverting cellular elongation control, and guaranteeing a supply of P-TEFb for HIV replication. When P-TEFb inhibitors were titrated onto cells, a gradual decrease in the ratio of P-TEFb in the 7SK snRNP to free P-TEFb was observed and the IC_{50} for that change is identical to the IC_{50} for inhibition of HIV replication [27]. This and other experiments suggest that the 7SK snRNP is required for HIV replication. Importantly, HIV infection or expression of Tat in HeLa cells results in the release of P-TEFb from the 7SK snRNP [23, 38]. Under the later conditions the majority of the Tat is found in a Tat•P-TEFb complex that sediments in a glycerol gradient with significantly lower molecular weight than the 7SK snRNP [23]. Using a defined in vitro assay in which the 7SK snRNP was immunoprecipitated from HeLa cell lysates, recombinant Tat was able to extract P-TEFb directly and this release was completely dependent on the P-TEFb binding domain of Tat [8]. In the presence of Tat, cellular control of P-TEFb via the 7SK snRNP is no longer effective at limiting HIV transcription.

In addition to Tat being able to bind to and extract P-TEFb from the 7SK snRNP, it also can interact with 7SK directly. Electrophoretic mobility shift assays demonstrated that Tat could bind to 7SK RNA in a dose-dependent manner and that this interaction displaced HEXIM1 and prevented P-TEFb•HEXIM1•7SK reformation at increasing Tat concentrations [23]. The role for binding of Tat to 7SK snRNA is not clear. Although the P-TEFb binding domain of Tat was indispensible for extraction of P-TEFb from the 7SK snRNP, the RNA binding domain only had a slight stimulatory effect [8]. Another study provided evidence for Tat association with 7SK snRNP that lacked HEXIM1 [39]. Reconciliation of all of these results may be achieved by taking into consideration that 7SK snRNA undergoes a conformational change upon loss of P-TEFb [8] (Figure 3, top right). Chemical protection experiments provided strong evidence for this conformational change in 7SK RNA after loss of P-TEFb due to flavopiridol treatment of cells or

after treatment of the 7SK snRNP with recombinant Tat [8]. It was hypothesized that loss of HEXIM was caused by the conformational change in 7SK RNA. Binding of Tat to 7SK could also be negatively affected by the restructuring event. The study that detected Tat bound to the 7SK snRNP could be explained by the higher affinity of Tat compared to HEXIM1 for binding to 7SK RNA.

After Tat has extracted P-TEFb from the 7SK snRNP, the Tat•P-TEFb complex is recruited to the poised polymerase through an interaction with the transactivation response element, TAR (Figure 3, bottom). TAR, contained within HIV's nascent transcript, has a bulge and loop hairpin structure which, when combined with its RNA sequence, is used by Tat•P-TEFb for specific binding [40]. Due to the close proximity of P-TEFb kinase targets, Tat•P-TEFb•TAR efficiently phosphorylates DSIF and NELF and effects productive elongation. Recent studies suggest that interactions of Tat with other cellular factors may also be involved. Tagged Tat protein was found to form two distinct complexes, one containing P-TEFb, PAF1, AF9, ENL, AFF1, AFF4, ELL, and EAF1 (Tatcom1) and the other containing P-TEFb, 7SK, LARP7, and MePCE (Tatcom2) [39]. The superelongation complex [41] with Tat, Tatcom1 was shown to be more efficient at Pol II CTD phosphorylation than Tat•P-TEFb alone and may serve to overcome a diverse set of repressive cellular blocks to HIV replication [39]. While a role for Tatcom2 was not discovered, concurrent research demonstrated the presence of a repressive, Tatcom2-like Tat•P-TEFb•HEXIM1•7SK complex at poised polymerases before the synthesis of TAR [42]. The generation of TAR was necessary for Tat-mediated HEXIM1 displacement and the activation of P-TEFb, potentially explaining the origin of Tatcom2. It is still not known if P-TEFb•HEXIM1•7SK without Tat is capable of localizing to the HIV LTR, either as a part of the PIC or the poised polymerase [43]. Understanding that Tat may participate in both a highly stimulatory complex (Tatcom1) and a regulatory, inhibitive complex (Tat•P-TEFb•HEXIM1•7SK) provides a host of new questions regarding HIV's use of elongation control.

4. Elongation Control and Viral Latency

Current HIV therapies are extremely effective at reducing viremia, thereby improving patient health and reducing viral transmission; however, the persistence of latent viruses dictates that the antiviral treatments must be continued for life. Elimination of the latent viral reservoirs is required to cure a patient. The current idea is that this could be accomplished by forcing reactivation of latent viruses while blocking new infections with antivirals [44]. If the activation is thorough, all latently infected cells would be killed and the virus would be eliminated from the host. Most of the latent viruses are found in resting $CD4^+$ T cells, that were initially infected, but never began to lytically produce virus [45]. These cells can remain dormant for decades but, upon activation, can express virus leading to AIDS.

There are multiple mechanisms working to maintain stably integrated viruses in a latent state. Because most HIV integration events take place within active genes, transcriptional interference may play a role in latency [46, 47]. This occurs when elongating RNA polymerase II molecules travel through the HIV promoter, potentially hindering initiation and polymerase pausing. This mechanism has been demonstrated using a Jurkat cell model [48, 49], but the contribution of this mechanism in resting T cells that have lower levels of transcription in general is not as clear. One issue contributing to latency is that key factors needed for production of long HIV transcripts are found in very low concentrations in resting T cells. These include NFκB and/or NFAT as well as P-TEFb containing Cyclin T1, and the nuclear concentrations of these factors are substantially increased following T-cell activation [28, 29].

Because only 1 in a million resting T cells may harbor a latent virus, it is very difficult to analyze the state of the integrated genes. Chromatin immunoprecipitation (ChIP) can provide an indication of the occupancy by specific factors, and relative occupancy of a given factor at different sites can be fairly quantitatively assessed if ChIP-Seq is used. Without a method to enrich latently infected cells, however, ChIP lacks the sensitivity to obtain significant signals from the integrated viral genomes. Because of this, model systems using clonal populations of transformed cells from a single integration event or primary $CD4^+$ cells with heterogeneous integrations have been developed [50]. The choice of model system is important because the growth state of the cells influences the transcription factor environment.

The design of methods to activate latent viruses is hampered by the lack of certainty concerning the chromatin state of the HIV promoter. Two possibilities are a completely repressed chromatin structure (Figure 4 top) or an open promoter configuration with a poised polymerase (Figure 4 bottom), and latent cells could be a mixture of both. A repressive chromatin structure could be actively maintained by histone deacetylases (HDACs) recruited by factors such as Sp1 or by a default pathway that covers DNA in the absence of promoter use [51]. Supporting a role for HDACs in maintaining HIV latency, HDAC inhibitors have been demonstrated to activate latent viruses [52]. Evidence also suggests a significant role for open promoter regulation. In growing cells, ChIP-Seq experiments have revealed that most mammalian genes are occupied by promoter proximally paused polymerases including genes that have very low or undetectable expression [10]. The HIV LTR was one of the first promoters found to generate poised polymerases, and short, nonpolyadenylated transcripts containing TAR have been found in latently infected resting T cells [47, 53]. Therefore, it is very likely that the main block to expression in many latently infected cells may occur at the P-TEFb-dependent step in the transition into productive elongation.

If latent viral genomes are loaded with poised polymerases, what is blocking the function of P-TEFb? As described earlier, NFκB can recruit P-TEFb, but, in resting T cells, NFκB resides mainly in the cytoplasm. Without some transcription of the HIV genome, the primary recruiter of P-TEFb, Tat, will be absent. Another possible elongation repression mechanism was suggested by the finding that HEXIM1 could bind to TAR [23] (Figure 4, bottom).

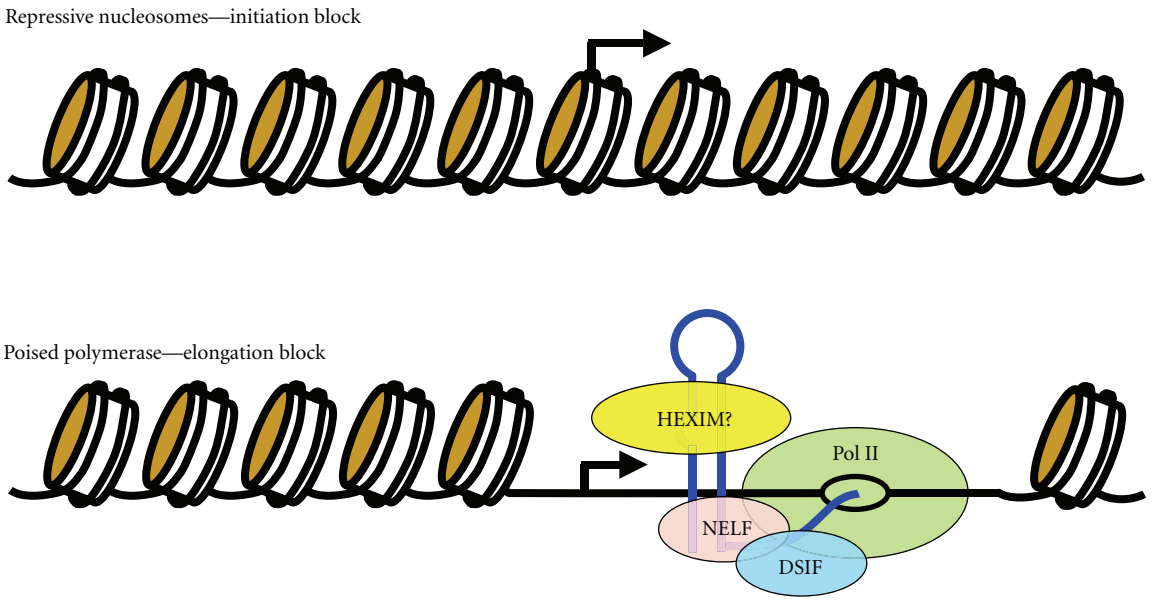


Figure 4: Possible mechanisms for maintaining latent viral genomes. The two panels represent two possible chromatin states over the HIV LTR in resting T cells in which the HIV genome is maintained in a silent state. In the upper panel, the entire LTR is covered in nucleosomes and initiation from the promoter is completely blocked. A less extreme but similar possibility (not pictured) is that accessibility of some factors is allowed, but initiation is still blocked. The lower panel depicts a state in which initiation is allowed, but all polymerases are left promoter proximally paused. HEXIM proteins may associate with TAR and act as a P-TEFb repressor.

The interaction of TAR with HEXIM1 would trigger the conformational change needed for P-TEFb binding [54] and, therefore, would act as a P-TEFb repressor [23]. The effects of reducing or increasing the level of HEXIM1 on expression from the HIV LTR has been interpreted as being mediated through the sequestration of P-TEFb by the 7SK snRNP [55–58]. However, direct binding of HEXIM1 to TAR can explain this influence on transcription [23]. It is currently not clear if HEXIM1 is bound to many nascent transcripts in human cells or if its interaction with TAR is specific.

5. Therapeutic Approaches Targeting HIV Transcription

The current cadre, of anti-HIV drugs target enzymes encoded by the virus and are quite effective until an HIV strain arises that is resistant. The error-prone nature of reverse transcriptase and recombination events lead to frequent mutations in the presence of drug-induced selective pressure, leading to resistance and proliferation despite therapy [59–61]. For this reason, most HIV treatments utilize at least three drugs with different targets. There are no currently available drugs that inhibit HIV transcription. Because P-TEFb is essential for HIV replication, the strong P-TEFb inhibitor flavopiridol blocks virus production [26]. Unfortunately, concentrations about 10-fold higher than those that block HIV replication also cause host cell death [27]. This is true for all P-TEFb inhibitors tested except for those that have other essential targets, in which case the therapeutic index is less than 10 [25, 27].

Because only HIV utilizes the Tat•P-TEFb complex, it is the logical target for a viral transcriptional inhibitor, and there are several possible avenues to develop compounds that block the function of this complex. A small molecule that blocked the interaction between Tat and P-TEFb would likely inhibit HIV transcription. It will be difficult to find such a compound, however, because Tat buries 3,499 Å^2 of surface area when bound to P-TEFb. An additional challenge in the search for such compounds is Tat's limited structure in the absence of P-TEFb. Another constraint to this approach is that compounds capable of preventing Tat from binding to Cyclin T1 or Cdk9 might also block binding of cellular factors such as HEXIM1, NFκB, Brd4, or CIITA, potentially causing unacceptable toxicity.

Another approach is to obstruct the recruitment of the Tat•P-TEFb complex to TAR. Compounds that block binding of Tat to TAR in vitro have been discovered [62–65], but have not been proven to work well in vivo for a variety of reasons including low affinity, poor uptake into cells, or other properties that make them inappropriate for delivery at efficacious concentrations. The Tat•P-TEFb complex has a higher affinity to TAR than Tat alone, and such a drug would have to overcome the additional contribution from Cyclin T1. Despite these setbacks, the potential for well-targeted drugs using this approach cannot be overlooked.

Finally, inhibition of the kinase activity of the Tat•P-TEFb complex specifically could block HIV transcription without affecting the cellular function of the important factor. Comparison of the structures of P-TEFb with [34] and without [37] Tat suggested that Tat induced a conformational change that could be exploited to develop a specific Tat•P-TEFb inhibitor. However, as mentioned above, change may have been due to mutations in the proteins used in the P-TEFb alone structure [34]. Solution of the structure of

wildtype P-TEFb is needed to resolve this issue. It may be possible to target the kinase activity of Tat•P-TEFb specifically by tethering a weak kinase inhibitor to a compound that binds tightly to the complex at the interface between Tat and P-TEFb. The anchor would provide specificity and increase the concentration of the inhibitor to effective levels. This method would take advantage of the tremendous stability of the Tat•P-TEFb complex but would require a significant amount of labor to design or discover the anchoring moiety of the compound.

Toward a cure for HIV, drugs that stimulate the reactivation of latently infected cells could allow for the subsequent eradication of the virus using existing antiretroviral treatments. Prostratin, a protein kinase c activator, induces the NFκB signaling pathway and stimulates HIV gene expression in latently infected cells without causing cell replication [66], but fails to uniformly activate latent cell populations [67]. Further, high-dose or prolonged treatments with prostratin were shown to induce growth arrest and cell death, hindering its use as a therapeutic [67]. If HEXIM proteins are found to be important in maintaining latency as described above, reactivation and subsequent elimination of the virus may also be possible by specifically relieving this block. For this approach to work, the mechanism must be unique to the HIV LTR or the HEXIM1•TAR interaction would have to be targeted specifically. Care would be needed to avoid disrupting the interaction of HEXIM with the 7SK snRNP. Structural information about the HEXIM•P-TEFb•TAR complex is needed to further define the function of HEXIM proteins on the LTR.

Acknowledgments

The authors would like to thank Tahir Tahirov (UNMC) for discussions concerning the Tat•P-TEFb complex as a potential drug target. This work was supported by NIH Grants GM35500 and AI074392 to D. H. Price, and K. A. Nilson is supported by a Presidential Graduate Fellowship at the University of Iowa.

References

[1] 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.

[2] P. Wei, M. E. Garber, S. M. Fang, W. H. Fischer, and K. A. Jones, "A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA," *Cell*, vol. 92, no. 4, pp. 451–462, 1998.

[3] D. Finzi, J. Blankson, J. D. Siliciano et al., "Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy," *Nature Medicine*, vol. 5, no. 5, pp. 512–517, 1999.

[4] D. B. Renner, Y. Yamaguchi, T. Wada, H. Handa, and D. H. Price, "A highly purified RNA polymerase II elongation control system," *Journal of Biological Chemistry*, vol. 276, no. 45, pp. 42601–42609, 2001.

[5] 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.

[6] V. T. Nguyen, T. Kiss, A. A. Michels, and O. Bensaude, "7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes," *Nature*, vol. 414, no. 6861, pp. 322–325, 2001.

[7] Z. Yang, Q. Zhu, K. Luo, and Q. Zhou, "The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription," *Nature*, vol. 414, no. 6861, pp. 317–322, 2001.

[8] B. J. Krueger, K. Varzavand, J. J. Cooper, and D. H. Price, "The mechanism of release of P-TEFb and HEXIM1 from the 7SK snRNP by viral and cellular activators includes a conformational change in 7SK," *PLoS ONE*, vol. 5, no. 8, Article ID e12335, 2010.

[9] Z. Yang, J. H. N. Yik, R. Chen et al., "Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4," *Molecular Cell*, vol. 19, no. 4, pp. 535–545, 2005.

[10] P. B. Rahl, C. Y. Lin, A. C. Seila et al., "C-Myc regulates transcriptional pause release," *Cell*, vol. 141, no. 3, pp. 432–445, 2010.

[11] S. Kanazawa, L. Soucek, G. Evan, T. Okamoto, and B. M. Peterlin, "c-Myc recruits P-TEFb for transcription, cellular proliferation and apoptosis," *Oncogene*, vol. 22, no. 36, pp. 5707–5711, 2003.

[12] S. R. Eberhardy and P. J. Farnham, "Myc recruits P-TEFb to mediate the final step in the transcriptional activation of the cad promoter," *Journal of Biological Chemistry*, vol. 277, no. 42, pp. 40156–40162, 2002.

[13] M. Barboric, R. M. Nissen, S. Kanazawa, N. Jabrane-Ferrat, and B. M. Peterlin, "NF-κB binds P-TEFb to stimulate transcriptional elongation by RNA polymerase II," *Molecular Cell*, vol. 8, no. 2, pp. 327–337, 2001.

[14] 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.

[15] M. Stevens, E. De Clercq, and J. Balzarini, "The regulation of HIV-1 transcription: molecular targets for chemotherapeutic intervention," *Medicinal Research Reviews*, vol. 26, no. 5, pp. 595–625, 2006.

[16] D. Harrich, J. Garcia, F. Wu, R. Mitsuyasu, J. Gonzalez, and R. Gaynor, "Role of SP1-binding domains in in vivo transcriptional regulation of the human immunodeficiency virus type 1 long terminal repeat," *Journal of Virology*, vol. 63, no. 6, pp. 2585–2591, 1989.

[17] M. C. Thomas and C. M. Chiang, "The general transcription machinery and general cofactors," *Critical reviews in biochemistry and molecular biology*, vol. 41, no. 3, pp. 105–178, 2006.

[18] Y. H. Ping and T. M. Rana, "DSIF and NELF interact with RNA polymerase II elongation complex and HIV-1 Tat stimulates P-TEFb-mediated phosphorylation of RNA polymerase II and DSIF during transcription elongation," *Journal of Biological Chemistry*, vol. 276, no. 16, pp. 12951–12958, 2001.

[19] S. Y. Kao, A. F. Calman, P. A. Luciw, and B. M. Peterlin, "Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product," *Nature*, vol. 330, no. 6147, pp. 489–493, 1987.

[20] K. Fujinaga, D. Irwin, Y. Huang, R. Taube, T. Kurosu, and B. M. Peterlin, "Dynamics of human immunodeficiency virus transcription: P-TEFb phosphorylates RD and dissociates negative effectors from the transactivation response element," *Molecular and Cellular Biology*, vol. 24, no. 2, pp. 787–795, 2004.

[21] M. Palangat, D. B. Renner, D. H. Price, and R. Landick, "A negative elongation factor for human RNA polymerase II inhibits the anti-arrest transcript-cleavage factor TFIIS," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 42, pp. 15036–15041, 2005.

[22] K. Adelman, M. T. Marr, J. Werner et al., "Efficient release from promoter-proximal stall sites requires transcript cleavage factor TFIIS," *Molecular Cell*, vol. 17, no. 1, pp. 103–112, 2005.

[23] S. C. Sedore, S. A. Byers, S. Biglione, J. P. Price, W. J. Maury, and D. H. Price, "Manipulation of P-TEFb control machinery by HIV: recruitment of P-TEFb from the large form by Tat and binding of HEXIM1 to TAR," *Nucleic Acids Research*, vol. 35, no. 13, pp. 4347–4358, 2007.

[24] 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.

[25] O. Flores, G. Lee, J. Kessler et al., "Host-cell positive transcription elongation factor b kinase activity is essential and limiting for HIV type 1 replication," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 13, pp. 7208–7213, 1999.

[26] S. H. Chao, K. Fujinaga, J. E. Marion et al., "Flavopiridol inhibits P-TEFb and blocks HIV-1 replication," *Journal of Biological Chemistry*, vol. 275, no. 37, pp. 28345–28348, 2000.

[27] S. Biglione, S. A. Byers, J. P. Price et al., "Inhibition of HIV-1 replication by P-TEFb inhibitors DRB, seliciclib and flavopiridol correlates with release of free P-TEFb from the large, inactive form of the complex," *Retrovirology*, vol. 4, article 47, 2007.

[28] J. Garriga, J. Peng, M. Parreño, D. H. Price, E. E. Henderson, and X. Graña, "Upregulation of cyclin T1/CDK9 complexes during T cell activation," *Oncogene*, vol. 17, no. 24, pp. 3093–3102, 1998.

[29] M. Karin and Y. Ben-Neriah, "Phosphorylation meets ubiquitination: the control of NF-κB activity," *Annual Review of Immunology*, vol. 18, pp. 621–663, 2000.

[30] M. J. Pazin, P. L. Sheridan, K. Cannon et al., "NF-κB-mediated chromatin reconfiguration and transcriptional activation of the HIV-1 enhancer in vitro," *Genes and Development*, vol. 10, no. 1, pp. 37–49, 1996.

[31] M. Mallardo, E. Dragonetti, F. Baldassarre, C. Ambrosino, G. Scala, and I. Quinto, "An NF-κB site in the 5'-untranslated leader region of the human immunodeficiency virus type 1 enhances the viral expression in response to NF-κB-activating stimuli," *Journal of Biological Chemistry*, vol. 271, no. 34, pp. 20820–20827, 1996.

[32] F. Pessler and R. Q. Cron, "Reciprocal regulation of the nuclear factor of activated T cells and HIV-1," *Genes and Immunity*, vol. 5, no. 3, pp. 158–167, 2004.

[33] D. L. Bates, K. K. B. Barthel, Y. Wu et al., "Crystal structure of NFAT bound to the HIV-1 LTR Tandem κB enhancer element," *Structure*, vol. 16, no. 5, pp. 684–694, 2008.

[34] T. H. Tahirov, N. D. Babayeva, K. Varzavand, J. J. Cooper, S. C. Sedore, and D. H. Price, "Crystal structure of HIV-1 Tat complexed with human P-TEFb," *Nature*, vol. 465, no. 7299, pp. 747–751, 2010.

[35] R. A. Marciniak, B. J. Calnan, A. D. Frankel, and P. A. Sharp, "HIV-1 Tat protein trans-activates transcription in vitro," *Cell*, vol. 63, no. 4, pp. 791–802, 1990.

[36] R. A. Marciniak and P. A. Sharp, "HIV-1 Tat protein promotes formation of more-processive elongation complexes," *EMBO Journal*, vol. 10, no. 13, pp. 4189–4196, 1991.

[37] S. Baumli, G. Lolli, E. D. Lowe et al., "The structure of P-TEFb (CDK9/cyclin T1), its complex with flavopiridol and regulation by phosphorylation," *EMBO Journal*, vol. 27, no. 13, pp. 1907–1918, 2008.

[38] M. Barboric, J. H. N. Yik, N. Czudnochowski et al., "Tat competes with HEXIM1 to increase the active pool of P-TEFb for HIV-1 transcription," *Nucleic Acids Research*, vol. 35, no. 6, pp. 2003–2012, 2007.

[39] B. Sobhian, N. Laguette, A. Yatim et al., "HIV-1 Tat assembles a multifunctional transcription elongation complex and stably associates with the 7SK snRNP," *Molecular Cell*, vol. 38, no. 3, pp. 439–451, 2010.

[40] K. A. Jones and B. M. Peterlin, "Control of RNA initiation and elongation at the HIV-1 promoter," *Annual Review of Biochemistry*, vol. 63, pp. 717–743, 1994.

[41] E. Smith, C. Lin, and A. Shilatifard, "The super elongation complex (SEC) and MLL in development and disease," *Genes and Development*, vol. 25, no. 7, pp. 661–672, 2011.

[42] I. D'Orso and A. D. Frankel, "RNA-mediated displacement of an inhibitory snRNP complex activates transcription elongation," *Nature Structural and Molecular Biology*, vol. 17, no. 7, pp. 815–821, 2010.

[43] M. Barboric and T. Lenasi, "Kick-sTARting HIV-1 transcription elongation by 7SK snRNP deporTATion," *Nature Structural and Molecular Biology*, vol. 17, no. 8, pp. 928–930, 2010.

[44] R. F. Siliciano, "What do we need to do to cure HIV infection," *Topics in HIV Medicine*, vol. 18, no. 3, pp. 104–108, 2010.

[45] T. W. Chun, D. Finzi, J. Margolick, K. Chadwick, D. Schwartz, and R. F. Siliciano, "In vivo fate of HIV-1-infected t cells: quantitative analysis of the transition to stable latency," *Nature Medicine*, vol. 1, no. 12, pp. 1284–1290, 1995.

[46] K. Lassen, Y. Han, Y. Zhou, J. Siliciano, and R. F. Siliciano, "The multifactorial nature of HIV-1 latency," *Trends in Molecular Medicine*, vol. 10, no. 11, pp. 525–531, 2004.

[47] Y. Han, K. Lassen, D. Monie et al., "Resting CD4+ T cells from human immunodeficiency virus type 1 (HIV-1)-infected individuals carry integrated HIV-1 genomes within actively transcribed host genes," *Journal of Virology*, vol. 78, no. 12, pp. 6122–6133, 2004.

[48] S. A. Williams, L. F. Chen, H. Kwon, C. M. Ruiz-Jarabo, E. Verdin, and W. C. Greene, "NF-κB p50 promotes HIV latency through HDAC recruitment and repression of transcriptional initiation," *EMBO Journal*, vol. 25, no. 1, pp. 139–149, 2006.

[49] T. Lenasi, X. Contreras, and B. M. Peterlin, "Transcriptional interference antagonizes proviral gene expression to promote HIV latency," *Cell Host and Microbe*, vol. 4, no. 2, pp. 123–133, 2008.

[50] H. C. Yang, "Primary cell models of HIV latency," *Current Opinion in HIV and AIDS*, vol. 6, no. 1, pp. 62–67, 2011.

[51] S. Hakre, L. Chavez, K. Shirakawa, and E. Verdin, "Epigenetic regulation of HIV latency," *Current Opinion in HIV and AIDS*, vol. 6, no. 1, pp. 19–24, 2011.

[52] D. M. Margolis, "Histone deacetylase inhibitors and HIV latency," *Current Opinion in HIV and AIDS*, vol. 6, no. 1, pp. 25–29, 2011.

[53] K. G. Lassen, J. R. Bailey, and R. F. Siliciano, "Analysis of human immunodeficiency virus type 1 transcriptional elongation in resting CD4+ T cells in vivo," *Journal of Virology*, vol. 78, no. 17, pp. 9105–9114, 2004.

[54] Q. Li, J. J. Cooper, G. H. Altwerger, M. D. Feldkamp, M. A. Shea, and D. H. Price, "HEXIM1 is a promiscuous double-stranded RNA-binding protein and interacts with RNAs in

addition to 7SK in cultured cells," *Nucleic Acids Research*, vol. 35, no. 8, pp. 2503–2512, 2007.

[55] J. H. N. Yik, R. Chen, A. C. Pezda, and Q. Zhou, "Compensatory contributions of HEXIM1 and HEXIM2 in maintaining the balance of active and inactive positive transcription elongation factor b complexes for control of transcription," *Journal of Biological Chemistry*, vol. 280, no. 16, pp. 16368–16376, 2005.

[56] J. H. N. Yik, R. Chen, A. C. Pezda, C. S. Samford, and Q. Zhou, "A human immunodeficiency virus type 1 Tat-like arginine-rich RNA-binding domain is essential for HEXIM1 to inhibit RNA polymerase II transcription through 7SK snRNA-mediated inactivation of P-TEFb," *Molecular and Cellular Biology*, vol. 24, no. 12, pp. 5094–5105, 2004.

[57] J. H. N. Yik, R. Chen, R. Nishimura, J. L. Jennings, A. J. Link, and Q. Zhou, "Inhibition of P-TEFb (CDK9/cyclin T) kinase and RNA polymerase II transcription by the coordinated actions of HEXIM1 and 7SK snRNA," *Molecular Cell*, vol. 12, no. 4, pp. 971–982, 2003.

[58] A. Fraldi, F. Varrone, G. Napolitano et al., "Inhibition of Tat activity by the HEXIM1 protein," *Retrovirology*, vol. 2, article 42, 2005.

[59] V. A. Johnson, F. Brun-Vézinet, B. Clotet et al., "Update of the drug resistance mutations in HIV-1: December 2010," *Topics in HIV Medicine*, vol. 18, no. 5, pp. 156–163, 2010.

[60] J. W. Tang and D. Pillay, "Transmission of HIV-1 drug resistance," *Journal of Clinical Virology*, vol. 30, no. 1, pp. 1–10, 2004.

[61] Z. Temesgen, D. Warnke, and M. J. Kasten, "Current status of antiretroviral therapy," *Expert Opinion on Pharmacotherapy*, vol. 7, no. 12, pp. 1541–1554, 2006.

[62] C. W. Lee, H. Cao, K. Ichiyama, and T. M. Rana, "Design and synthesis of a novel peptidomimetic inhibitor of HIV-1 Tat-TAR interactions: squaryldiamide as a new potential bioisostere of unsubstituted guanidine," *Bioorganic and Medicinal Chemistry Letters*, vol. 15, no. 19, pp. 4243–4246, 2005.

[63] M. A. Gelman, S. Richter, H. Cao, N. Umezawa, S. H. Gellman, and T. M. Rana, "Selective binding of TAR RNA by a Tat-derived β-peptide," *Organic Letters*, vol. 5, no. 20, pp. 3563–3565, 2003.

[64] V. Kesavan, N. Tamilarasu, H. Cao, and T. M. Rana, "A new class of RNA-binding oligomers: peptoid amide and ester analogues," *Bioconjugate Chemistry*, vol. 13, no. 6, pp. 1171–1174, 2002.

[65] N. Tamilarasu, J. Zhang, S. Hwang, and T. M. Rana, "A new strategy for site-specific protein modification: analysis of a TAT peptide-tar RNA interaction," *Bioconjugate Chemistry*, vol. 12, no. 2, pp. 135–138, 2001.

[66] Y. D. Korin, D. G. Brooks, S. Brown, A. Korotzer, and J. A. Zack, "Effects of prostratin on T-cell activation and human immunodeficiency virus latency," *Journal of Virology*, vol. 76, no. 16, pp. 8118–8123, 2002.

[67] S. A. Williams, L. F. Chen, H. Kwon et al., "Prostratin antagonizes HIV latency by activating NF-κB," *Journal of Biological Chemistry*, vol. 279, no. 40, pp. 42008–42017, 2004.

SAGE-Hindawi Access to Research
Genetics Research International
Volume 2011, Article ID 572689, 11 pages
doi:10.4061/2011/572689

Research Article

Comparative Study of Cyanobacterial and *E. coli* RNA Polymerases: Misincorporation, Abortive Transcription, and Dependence on Divalent Cations

Masahiko Imashimizu,[1,2] Kan Tanaka,[3] and Nobuo Shimamoto[1,4]

[1] *Structural Biology Center, National Institute of Genetics, and Department of Genetics, The Graduate University for Advanced Studies, Mishima, Shizuoka 411-8540, Japan*
[2] *Gene Regulation and Chromosome Biology Laboratory, National Cancer Institute at Frederick, MD 21702, USA*
[3] *Graduate School of Horticulture, Chiba University, 648 Matsudo, Matsudo-shi, Chiba 271-8518, Japan*
[4] *Faculty of Life Sciences, Kyoto Sangyo University, Kamigamo-Motoyama, Kita-Ku, Kyoto 603-8555, Japan*

Correspondence should be addressed to Nobuo Shimamoto, nshima@cc.kyoto-su.ac.jp

Received 7 June 2011; Accepted 31 July 2011

Academic Editor: Sebastián Chávez

If Mg^{2+} ion is replaced by Mn^{2+} ion, RNA polymerase tends to misincorporate noncognate nucleotide, which is thought to be one of the reasons for the toxicity of Mn^{2+} ion. Therefore, most cells have Mn^{2+} ion at low intracellular concentrations, but cyanobacteria need the ion at a millimolar concentration to maintain photosynthetic machinery. To analyse the mechanism for resistance against the abundant Mn^{2+} ion, we compared the properties of cyanobacterial and *E. coli* RNA polymerases. The cyanobacterial enzyme showed a lower level of abortive transcription and less misincorporation than the *E. coli* enzyme. Moreover, the cyanobacterial enzyme showed a slower rate of the whole elongation by an order of magnitude, paused more frequently, and cleaved its transcript faster in the absence of NTPs. In conclusion, cyanobacterial RNA polymerase maintains the fidelity of transcription against Mn^{2+} ion by deliberate incorporation of a nucleotide at the cost of the elongation rate. The cyanobacterial and the *E. coli* enzymes showed different sensitivities to Mg^{2+} ion, and the physiological role of the difference is also discussed.

1. Introduction

A DNA-dependent RNA polymerase (RNAP) has a catalytic center chelating Mg^{2+} ions to form phosphodiester bonds [1, 2]. Most other divalent cations inhibit the activity, but Mn^{2+}, which has a similar but only slightly larger ionic radius than Mg^{2+}, supports the polymerizations. In the case of *Escherichia coli* (*Esc*) RNAP, Mn^{2+} causes the enzyme to misincorporate deoxyribonucleoside triphosphate (dNTP) instead of ribonucleoside triphosphate (NTP) [3] as well as noncognate NTPs [4–6]. Therefore, the growth of *E. coli* is inhibited in the presence of Mn^{2+}, and the inhibition is mitigated by GreA and GreB which bind to RNAP [7].

In cyanobacteria, the intracellular concentration of Mn^{2+} is considered to be higher by two orders of magnitude than in *E. coli* [8], because Mn^{2+} is required at higher concentrations to assemble the photosynthetic reaction center carrying Mn^{2+} cluster, and/or to scavenge reactive oxygen species generated from photosynthesis [9, 10]. Since cyanobacteria have no genes orthologous to *greA* or *greB*, cyanobacteria should have the mechanism of detoxifying Mn^{2+}.

We thus made two purified reconstituted transcription systems from a mesophilic cyanobacterial species, *Synechococcus* sp. PCC 7942 (*Syn*), and from a moderately thermophilic species, *Thermosynechococcus elongatus* BP-1 (*The*). By comparing these systems with the system of *E. coli*, we here examine the possibility that the mechanism of detoxifying Mn^{2+} is installed in the poisoning target, RNAP, and ask whether or not the cyanobacterial RNAP has a sensitivity to Mg^{2+} different from that of *E. coli* RNAP.

2. Materials and Methods

2.1. Materials. All of the oligo DNAs and RNA, NTPs, and [γ-^{32}P]ATP were obtained from Hokkaido System Science, Yamasa and Perkin Elmer, respectively. Restriction enzymes were purchased from New England Biolabs and Takara. Primestar HS and Primestar Max DNA polymerases used for PCR were purchased from Takara. *Syn* cells used for RNAP purification were partially gifted from Dr. Mitsumasa Hanaoka. The culture condition of the gifted cells is the same as shown in the next section.

2.2. Protein Purification. Cells of *Syn* and *The* were cultivated and harvested as described [15, 16]. About 10 g of the wet-cell paste was suspended in 30 mL of TGED buffer [10 mM Tris-HCl (pH 8.0), 5% (v/v) glycerol, 0.1 mM EDTA, and 1 mM DTT] containing 1 mM phenylmethylsulfonyl fluoride and 0.2 M NaCl, and then, cells in suspension were lysed by sonication. A 35%- saturated solution of ammonium sulfate was added to a soluble fraction of the cell lysate which was obtained by centrifuging for 30 min at 20,000 $\times$ g. The obtained mixture was added to 10 mL of phenyl sepharose resin (GE Healthcare), pre-equilibrated with the TGED buffer containing 0.2 M NaCl and 35% saturated ammonium sulfate, and the mixture was washed by the same buffer, eluted with the TGED buffer. The eluate was applied to a column containing 10 mL of DEAE resin (TOSOH) pre-equilibrated with the TGED buffer, and the column was washed with TGED buffer containing 0.2 M NaCl, eluted with the TGED buffer containing 0.5 M NaCl. The eluted fraction was further chromatographically purified by using Hi-trap heparin affinity column (GE Healthcare) and MonoQ anion exchange column (GE Healthcare) as described [16]. The purified core enzyme was dialyzed against the TGED buffer containing 0.2 M NaCl and 50% glycerol and stored at -80°C. The core enzyme of *E. coli* was prepared according to [17], and σ^{70} of *E. coli* and σ^{A}s of *Syn* and *The* were obtained as described in [15, 16].

2.3. In Vitro Transcription Assays. The liner DNA template containing the T7A1 promoter from -147 to +87, when +1 is the transcription start site, was prepared by PCR using the plasmid pAR1435 [18] and the following digestion by *Hae*III. The template containing the *psbA2* promoter from -127 to +101 was prepared by PCR using a genomic DNA of a cyanobacterium *Synechocystis sp*. PCC 6803. These DNA templates were purified by PAGE. The holoenzyme was reconstituted by incubating core RNAP mixed with a 3-fold molar excess of the primary σ factor for 10 min at 37°C. We used twice more *Syn* and *The* RNAPs, because they have smaller affinities for the *psbA2* promoter. The reconstituted holoenzyme (50 nM for *E. coli* or 100 nM for *Syn* and *The*) and 20 nM DNA template was preincubated for 10 min at 37°C in 8 μL of T-buffer [50 mM Tris-HCl (pH 7.9), 100 mM KCl, 10 mM $MgCl_2$ (see figure legends and the main text in the cases that the concentration was changed or $MgCl_2$ was changed to $MnCl_2$), 1 mM DTT, and 150 μg mL^{-1} partially hydrolysed casein]. Reaction was started by adding 2 μL of prewarmed substrate mixture: 5 μM [γ-^{32}P]ATP (40 Ci mmol^{-1}) as well as 100 μM each of GTP, CTP, and UTP. In a single-round reaction, heparin (100 μg mL^{-1}) was added together with the substrates to eliminate enzyme turnover. After incubation for 20 min at 37°C, the reaction was stopped by phenol/chloroform/isoamyl alcohol (25 : 24 : 1). Transcripts were analyzed by PAGE using a 20% gel containing 7 M urea. The sequences of abortive transcripts from the T7A1 promoter were assigned according to [19], and those from the *psbA2* promoter were determined using 5′ end-labeled RNA marker 5′-AGUCAGUU-3′ and a chemical cleavage assay (Figure S2 which is available online at doi:10.4061/2012/572689).

All the experiments with presented results in this study were repeated two or more times and the represented ones are shown.

2.4. TEC Formation. In order to stop elongation at Position +18 by a lack of the cognate CTP for Position +19, the intrinsic cytosine residues (nontemplate strand) at Positions +4, +9, +10, and +14 of the *psbA2* promoter were replaced by thymine by PCR-based mutagenesis, following cloning of the DNA template from -127 to +101 into a plasmid pUC19 (see Figure 4(a)). For the TEC9 formation, the intrinsic cytosine residues at Positions +10 and +14 were retained. The binary complex was formed by mixing the holoenzyme (50 nM for *E. coli* or 100 nM for *Syn*) and 40 nM DNA template containing the *psbA2* promoter as described above. The ternary complex was formed by incubating 5 min at 37°C with 5 μM [γ-^{32}P]ATP (40 Ci mmol^{-1}) as well as 50 μM each of GTP and UTP in T-buffer containing 10 mM $MgCl_2$ or 1 mM $MnCl_2$. The TEC was isolated from the substrates including [γ-^{32}P]ATP as well as released abortive transcripts by triplicated passing through a MicroSpin G50 column (GE Healthcare) equilibrated with T2-buffer [50 mM Tris-HCl (pH 7.9), 100 mM KCl, 10 mM $MgCl_2$ or 1 mM $MnCl_2$, 1 mM DTT, and 5% glycerol] at room temperature. After this treatment, no elongations by the TEC9s and TEC18s of both *Syn* and *E. coli* in the T2 buffer were detected for 10 min at 37°C, without adding substrates.

2.5. Single-Step Elongation Assay. Reactions were started by adding cognate or noncognate NTP at the final concentration of 0.5 mM to 9 μL solution containing the TEC and the T2-buffer. The obtained mixture was incubated at 37°C for the indicated time and the reaction was stopped by adding an equal volume of gel-loading solution, containing 95% (v/v) deionized formamide, 20 mM EDTA (pH 8.0), bromophenol blue and xylene cyanol, 0.05% (w/v) each, and analyzed by PAGE using a 20% gel containing 7 M urea.

3. Results

3.1. Cyanobacterial Transcription System Reconstituted from Purified Components. For the comparative study, it is essential to purify cyanobacterial RNAPs. The tedious step of

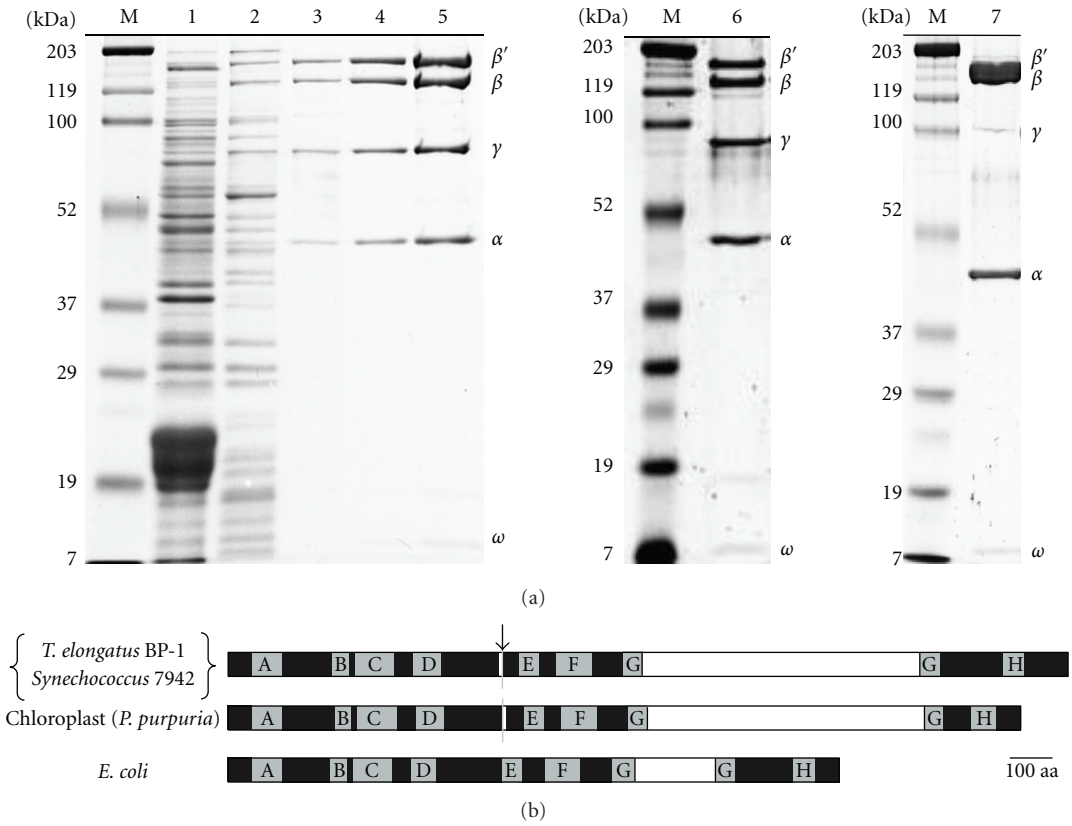


Figure 1: Cyanobacterial RNAPs. (a) Analysis of the fractions of *The* core enzyme by 10% SDS-PAGE after the treatment of hydrophobic resin (lane 1 : 26.9 μg), DEAE anion-exchange (lane 2 : 7.0 μg), heparin affinity (lane 3 : 0.8 μg), and MonoQ chromatography (lane 4 : 1.6 μg). Lanes 5–7 show 7.5 μg of the purified *The*, *Syn,* and *E. coli* enzymes used in transcription assays. Marker proteins (BIO-RAD) are also shown in lane M with the molecular weights indicated in the left margin. The gel was stained with Coomasie Brriliant Blue. (b) Schematic diagrams of β' subunits of *The*, *Syn*, a chloroplast of red alga *Porphyra purpurea,* and *E. coli*. The conserved regions A–H [11, 12] were also indicated. The white boxes represent the nonconserved domain inserted in Region G [13, 14]. The split sites are indicated by an arrow. The scale bar of 100 amino-acid residues (aa) is indicated at the bottom.

the purification is to remove thylakoid-membrane fragments and the associated proteins [20, 21]. We thus invented a batch-wise removal with a hydrophobic resin (see Materials and methods). The *Syn* and *The* core enzymes were purified by the improved procedure, resulting in a preparation which is more than 95% pure, judged by CBB stain, within a day (Figure 1(a)). The core enzyme of cyanobacteria is composed of $\alpha_2\beta\beta'\omega$ as *E. coli* enzyme, but its β' subunit is composed of two polypeptides [20]. A large nonconserved domain is inserted in the G region as chloroplast RNAP (Figure 1(b)) [21].

We examined the transcription by the holoenzymes of *Syn*, *The,* and *E. coli* retaining their primary σ factors (σ^s and σ^{70}) at the T7A1 promoter (Figure 2(a)), which is the strongest among the standard promoters for the *E. coli* enzyme [22]. The *Syn* enzyme was active on this transcription unit but much weaker than the *E. coli* enzyme: the runoff transcript was observed only at the multiround condition, and its amount was lower by an order of magnitude (Figure 2(b)). Moreover, the *The* enzyme did not synthesize any detectable transcripts in all conditions (data not shown).

We thus searched for a promoter driving transcription by both cyanobacteria and *E. coli* holoenzymes without adding specific activators. The cyanobacterial *psbA2* gene encoding D1 protein of the photosystem II is highly transcribed in the daytime [23] and the sequence of its promoter shares the -10 and -35 elements recognized by the *E. coli* σ^{70} holoenzyme (Figure 2(a)) [24]. As expected, the *E. coli* enzyme was active on this promoter and produced similar amounts of runoff transcripts from the *psbA2* promoter in the single-round condition (Figure 2(c)). We thus used this promoter in comparison between these enzymes.

3.2. Cyanobacterial RNAPs Are Less Abortive due to Their Core Parts. In contrast to the similar amounts of runoff transcript, *Syn* RNAP produced a much lower level of 2–11 nt long transcripts (Figure 2(c)), which is known as abortive transcription, an iterative synthesis and release of oligo-RNA in initiation [25, 26]. The ratio of abortive synthesis to runoff synthesis is smaller than that in *E. coli* by two orders of magnitude. The ratio for *The* RNAP could be determined only in a multiround transcription condition and was as low as that of *Syn* RNAP (Figure 2(c)).

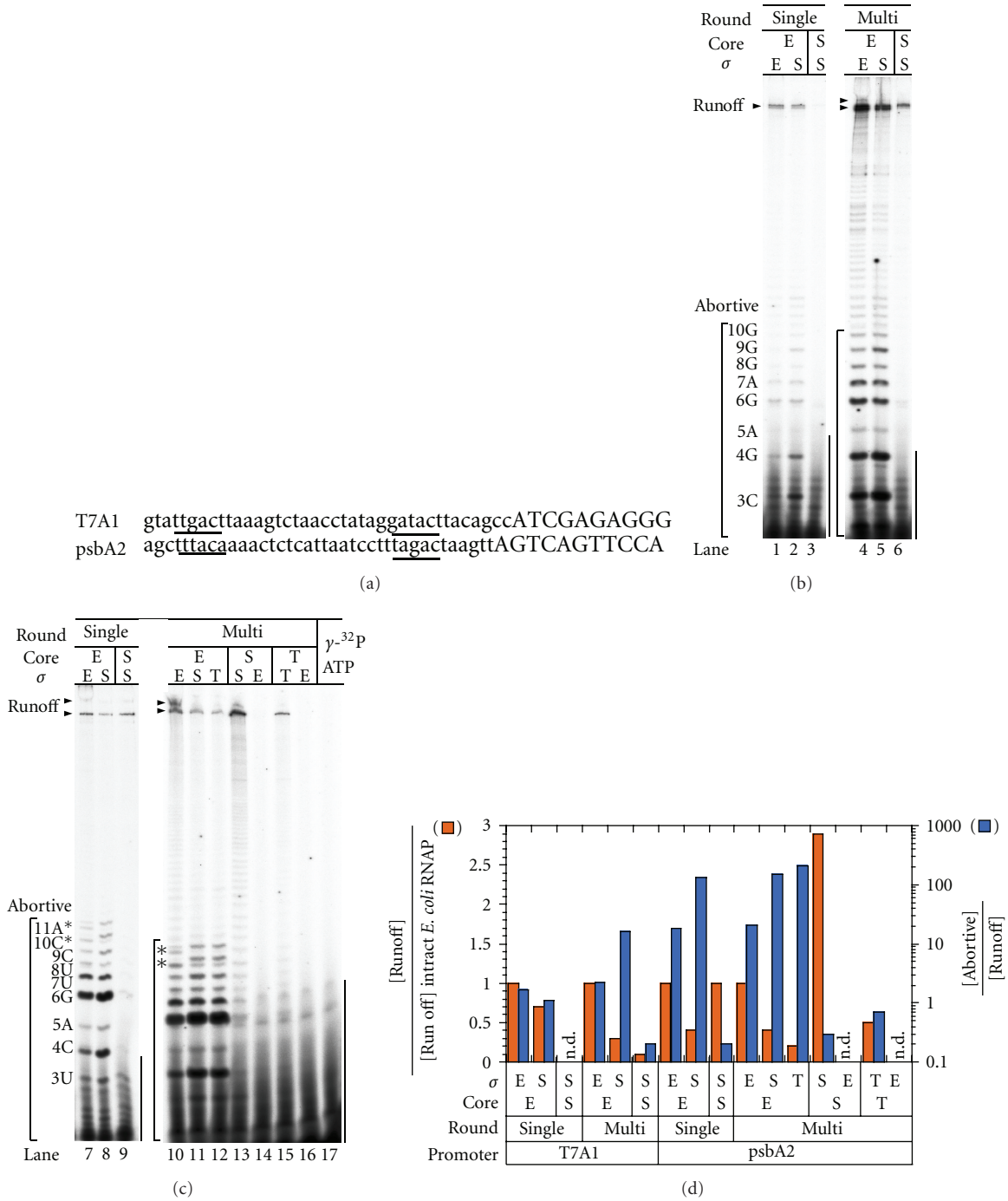


Figure 2: Transcription from the T7A1 and the *psbA2* promoters by cyanobacterial and *E. coli* RNAPs as well as their chimeric enzymes. (a) The DNA sequences of the T7A1 and the *psbA2* promoters. The putative −10 and −35 elements are underlined, and the transcribed sequences are indicated in uppercase letters. (b) The analysis of transcripts labeled with [γ-^{32}P]ATP from the T7A1 promoter with 20% polyacrylamide in the presence of 7M urea in TBE buffer. (c) The transcripts from the *psbA2* promoter. The round of transcription is indicated at the top of the gels. The runoff and abortive transcripts are indicated by arrowheads and parentheses, respectively. The length and the incorporated nucleotide at its 3′ end of an abortive transcript are indicated on the left margin. Asterisks indicate the abortive transcripts involving misincorporation. The radioactive contaminants that is contained in [γ-^{32}P]ATP are also indicated in lane 17. (d) The amounts of the run-off transcript normalized by that of the intact *E. coli* enzymes (red) and the ratio of the amount of abortive transcripts to that of the run-off transcript in logarithmic scale.

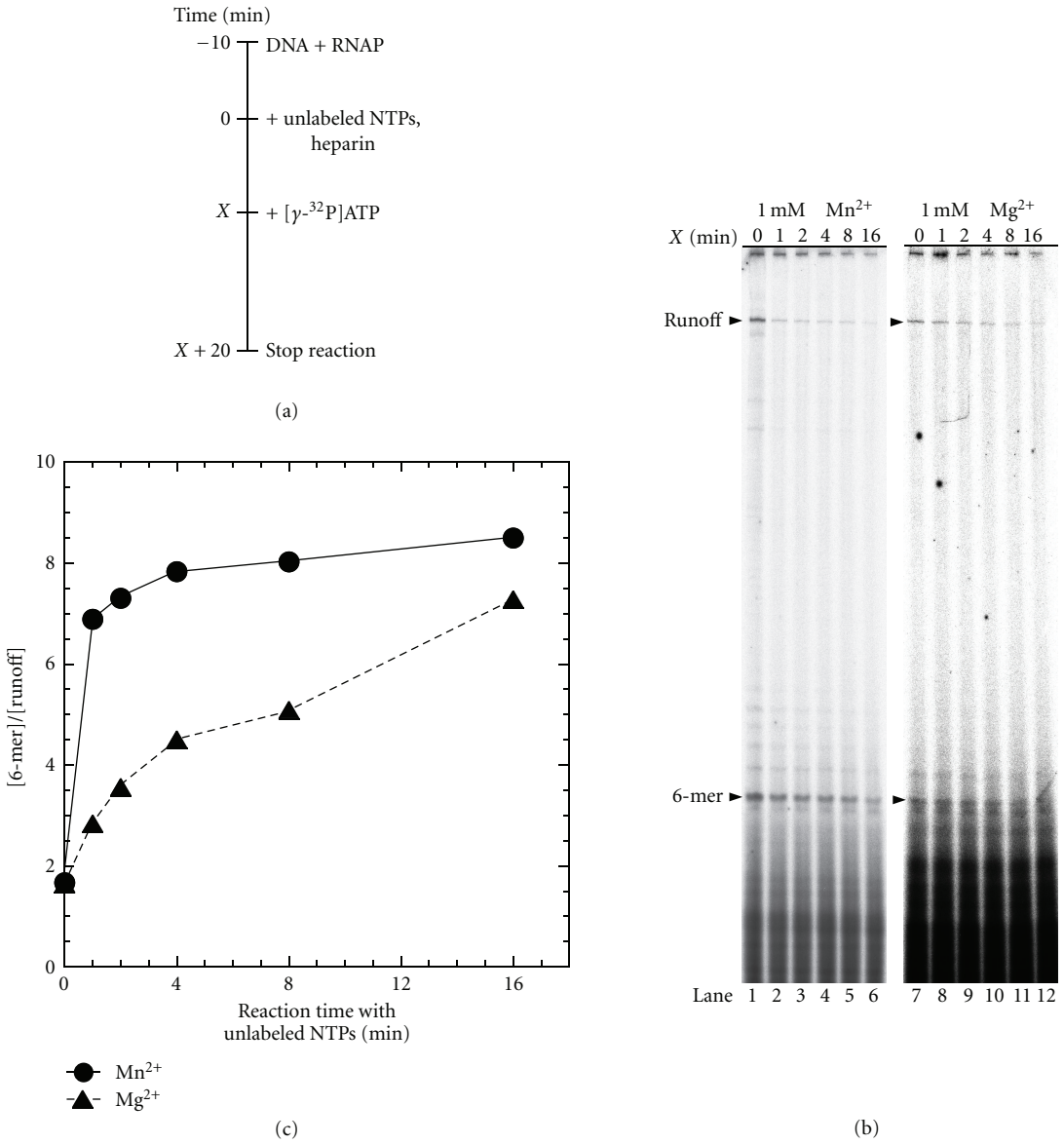


Figure 3: Inverse pulse-labeling assay for detecting persistent abortive synthesis at the *psbA2* promoter by *Syn* RNAP. (a) Schematic diagram of the assay. The reaction time X is varied. (b) The labeled products synthesized at 1 mM Mn^{2+} (left) and at 1 mM Mg^{2+} (right). The runoff transcript and the 6-mer abortive transcript are indicated by arrowheads. (c) The ratios of the amount of 6-mer abortive transcript to that of the runoff transcript were plotted against the reaction time X.

The observed small ratios were not specific to the *psbA2* promoter, because in multiround transcriptions from the T7A1 promoter, the ratio for *Syn* RNAP was also two orders of magnitude smaller than that for *E. coli* (lanes 4 and 6 in Figure 2). In this experiment, we used the holoenzymes reconstituted from the σ^As purified from the overproducing *E. coli* strains because of low yields of holoenzymes. The observed low activities of abortive transcription are not due to the artifacts of the reconstitution. We independently purified the histidine-tagged holoenzyme of a *Synechocystis* sp. PCC 6803 and it also produced abortive transcripts as low as the *Syn* and *The* RNAP (Figure S1A). In conclusion, cyanobacterial RNAPs are generally much less abortive than *E. coli* RNAP.

The level of abortive transcription is known to depend on mutations of both core enzyme and σ factor [27–29], indicating that both the components are responsible. We addressed the question of which component is more responsible by constructing chimeric holoenzymes. Taking into account the large difference between the levels of cyanobacterial and *E. coli* RNAPs, we found that the large difference was associated with the core part of RNAP but not σ (lanes 2, 5, 8, 11, and 12 in Figure 2). In spite of the absence of detectable transcripts by *The* enzyme in

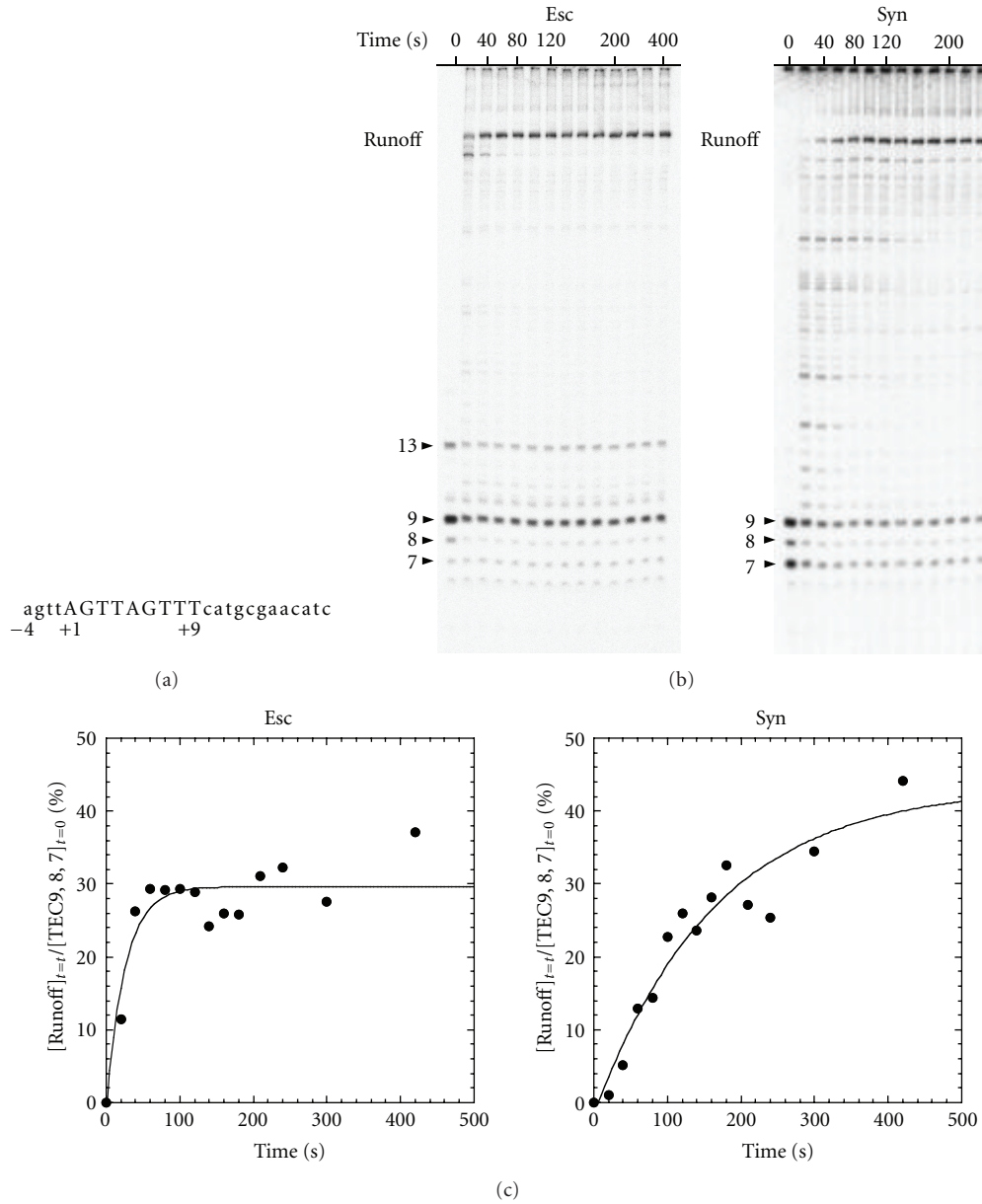


Figure 4: Elongation by the *Syn* and *E. coli* TEC9s formed at the *psbA2* promoter. (a) A sequence of the nontemplate strand near the initiation site (+1). The early transcribed region of 9 mer is shown in uppercase letters and the upstream sequence in lower case letters. (b) The transcripts obtained in elongation of the TEC9s which had been labeled with [γ-^{32}P]ATP. Four NTPs of 100 μM each were added to the TEC9 at time zero and incubated for the indicated times. Only this experiment was performed at 25°C because of distinct pausing of *Syn* enzyme. The transcripts existed at time zero as well as the runoff transcripts are indicated by arrowheads. (c) The amount of the runoff transcript at each time point is normalized with the initial amount of TECs and plotted against time. The curves were the best-fit single-exponential ones.

Figure 2, the above conclusion is also applicable to the enzyme because of the results of transcripts labeled with [α-^{32}P] UTP (Figure S1B).

Among the abortive transcripts at the *psbA2* promoter, the 10 nt- and 11 nt-transcripts of *E. coli* RNAP accompanied the bands with decreased migration (Figure 2(c)). These transcripts were shown to contain a purine nucleotide instead of the cognate cytosine at Position +10, a misincorporation, by a chemical cleavage assay (Figure S2). In contrast to abortive transcripts, the runoff transcript did not contain any detectable misincorporation. The chimeric enzymes containing *E. coli* core enzyme also showed the misincorporation (lanes 7, 8, 10–12 in Figure 2). Such a misincorporation by the *E. coli* enzyme had already been known for several promoters [19], and thus, this result suggests that the core enzyme is responsible for the misincorporation.

When the enzymes of two different species with different typical growing temperatures are compared, there are no absolute choice of the temperature. Since we focused to the

difference in the catalytic properties of RNAPs rather than the difference in its role in their growth, we selected the same 37°C for *in vitro* assays. In this way, we can avoid the effect of the different stabilities of DNA duplex in the promoter. The both cyanobacteria grow and survive at 37°C although the optimal temperature of *The* is 55°C. Another lines of circumstantial evidence for our choice is that the levels of the runoff RNA synthesis are similar for the *The* and *E. coli* enzymes for the *Syn rrnA* promoter (Panel B in Figure S1).

3.3. Syn RNAP Is Likely to Form Moribund Complex which Produces Only Abortive Transcripts. For *E. coli* RNAP, only a part of the promoter complex synthesizes the full-length RNA, while the rest produces the majority of abortive transcripts at the promoter (Figure S3), the latter being named moribund complex [7, 30, 31]. To test whether or not the *Syn* RNAP has a similar property, we carried out the most sensitive assay for detecting the moribund complex: inverse pulse-labeling which monitors the fates of promoter-RNAP complex [30]. In this assay, a single-round transcription from the *psbA2* promoter was started with unlabelled 4NTPs (ATP, CTP, GTP, and UTP) at time zero. The γ-^{32}P-labeled initiating ATP was then added at various time points, followed by incubating for a further 20 min to complete the round (Figure 3(a)). Since the [γ-^{32}P]-labeled runoff and abortive transcripts are both produced only by the RNAP that still survive at the promoter at the time point of adding [γ-^{32}P]ATP, the ratio of the two kinds of labeled transcripts should reflect the preference of producing abortive transcripts.

As shown in Figures 3(b) and 3(c), the preference becomes stronger at the later time points, showing that the RNAP bound at the promoter at time zero is not homogeneous. Although there is no evidence that cyanobacterial RNAP forms the same moribund complex as *E. coli* RNAP, which produces only abortive products, a fraction of the *Syn* enzyme produces abortive transcripts preferentially over others, and the fraction is enriched at the later time points.

3.4. Syn RNAP Pauses More Frequently and Elongates Transcript More Slowly Than the E. coli Enzyme. We next compared the elongation by *Syn* with that by *E. coli* RNAPs. To isolate elongation from initiation, we used a ternary elongation complex retaining 9-mer transcripts, TEC9 (Figure 4(a)). It was elongated with 4NTPs in the presence of 10 mM Mg^{2+}. The *E. coli* TEC9 smoothly elongated its transcript, and the elongation was completed at 40 s (Figure 4(b)). The 6- and 7-mer transcripts existed at time zero, and their amounts did not change, demonstrating that they had been contaminating TEC9. The 9-mer transcript, as well as a part of the 8-mer transcript, was elongated. Therefore, there is no significant pausing of the *E. coli* enzyme on this template.

In contrast, the elongation of the *Syn* TEC9 was paused at a number of lengths and was completed at later than 200 s (Figure 4(b)). When the time courses of syntheses of the run-off transcripts are approximated to single-exponential curves, the elongation of the *Syn* TEC9 has a time constant

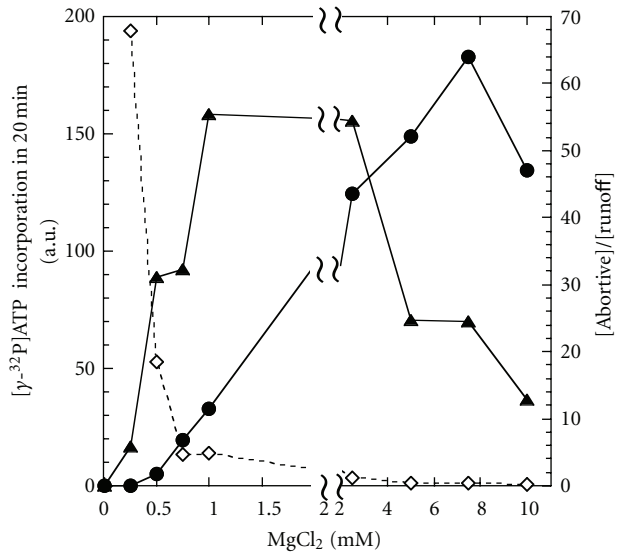


Figure 5: Abortive and runoff transcriptions by *Syn* RNAP at the *psbA2* promoter in the different concentration of Mg^{2+}. The levels of abortive synthesis of 3 to 11-mer (▲) and the runoff synthesis (•) in 20 min in a single-round are plotted against the concentrations of $MgCl_2$. The ratio of the levels (◊) is also plotted.

of 6 times longer than that of the *E. coli* TEC9, that is, time constants of 164 s compared with 26 s (Figure 4(c)).

The *Syn* TEC9 contains more amounts of 7-mer and 8-mer transcripts than the *E. coli* TEC9 at time zero. Since no such short transcripts were found prior to the step of removing NTPs in the preparation (data not shown), the shorter transcript must be generated in the absence of NTP by hydrolysis and/or pyrophosphorolysis of TEC9's, suggesting that the *Syn* TEC9 had a higher activity of shortening a transcript than *E. coli* had. In addition, the *E. coli* TEC9 is contaminated by a 13-mer transcript. This might be due to the slippage synthesis [32–34] on the *psbA2* promoter, where the segment of AGUU transcribed from +1 to +4 slips back to the agtt template sequence from −4 to −1, resulting in an addition of 4 base extra sequence at the 5′-end (Figure 4(a)). Again, the *Syn* enzyme does not catalyze such a misincorporation.

3.5. Sensitivity to Mg^{2+} Ion of Syn RNAP. The Mg^{2+} concentration in chloroplast is observed to be changed at dark and at light [35]. In cyanobacteria, the gene expressions are globally repressed at dark, including the *psbA2* expression, [36]. To test the possibility that transcription is directly controlled by Mg^{2+} concentration, we examined the sensitivity of *Syn* RNAP to the concentration (Figure 5). According to the increasing Mg^{2+} concentration from 0.5 to 2.5 mM, the level of runoff synthesis was increased by 25-fold, while the ratio of abortive synthesis to runoff synthesis was decreased to 1/15. In other words, it is possible to control RNAP to produce mainly abortive transcripts at a concentration and mainly mature transcripts in another concentration. However, such control would satisfy several quantitative conditions on the Mg^{2+} concentration.

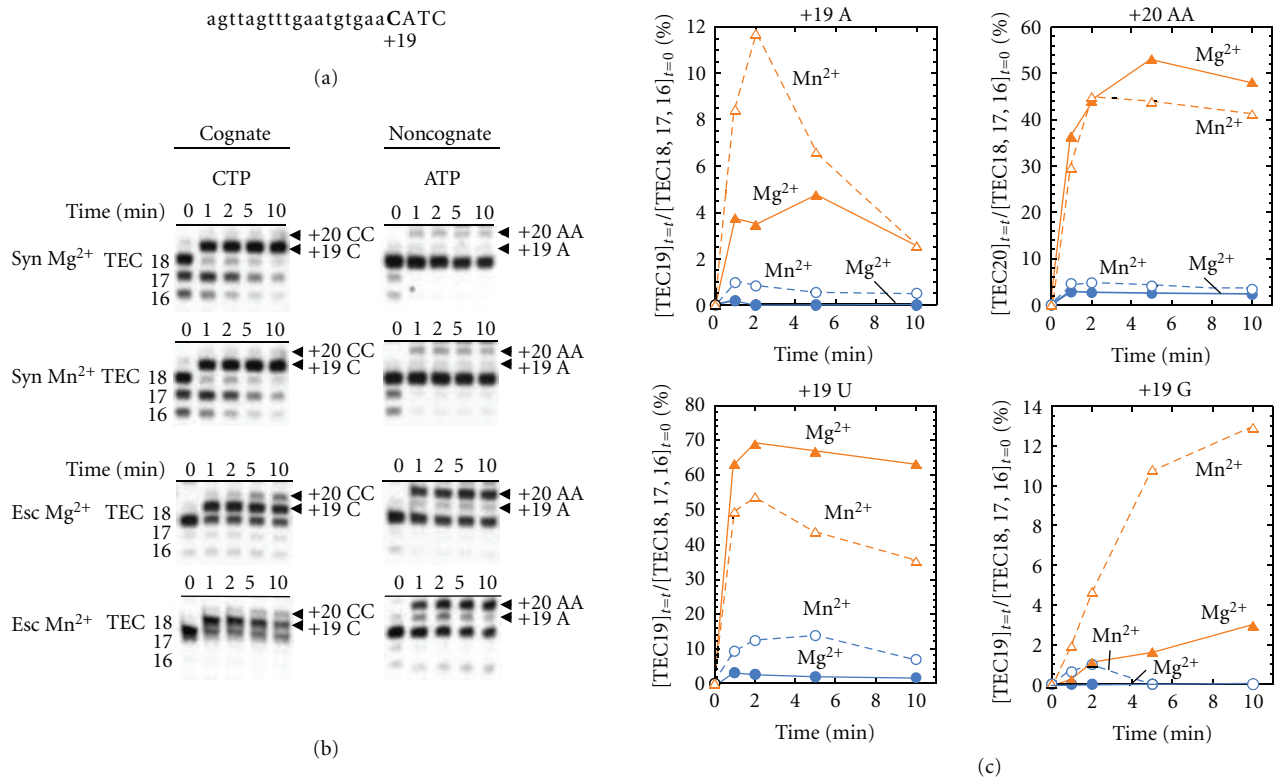


FIGURE 6: Incorporations of the cognate or noncognate NMP by the *Syn* and *E. coli* TEC18s in the presence of 10 mM Mg^{2+} or 1 mM Mn^{2+}. (a) The sequence of the nontemplate strand from +1 to +22. Position +19 and the 18-mer transcript are shown in the bold letter and lower case letters, respectively. (b) The products of TEC18 synthesized in the presence of the cognate CTP and the noncognate ATP (0.5 mM each). Reactions were performed in the presence of 10 mM $MgCl_2$ (upper) or 1 mM $MnCl_2$ (lower). The lengths of transcripts are indicated on the both margins. The incorporated nucleotides are also indicated by the arrowheads in the right margin. (c) The amount of transcript involving misincorporation was shown in the percent of the initial amount of TECs. The TECs of the *Syn* (blue circle) and *E. coli* (*Esc*, orange triangle) were incubated with 0.5 mM NTP for the indicated times in the presence of 10 mM $MgCl_2$ (filed symbols) or 1 mM $MnCl_2$ (open symbols).

3.6. Syn RNAP Shows Higher Fidelity in Elongation Than the E. coli Enzyme. The results above described show that the *Syn* RNAP is slower and more accurate in the whole elongation than the *E. coli* RNAP. We then examined the elongation at the resolution of one nucleotide and arbitrarily selected the elongation at +18. By using a *psbA2* DNA with cytosine replaced as shown in Figure 6(a), we have prepared TEC18s by the *Syn* and *E. coli* enzymes. The elongations of *Syn* and *E. coli* TEC18s were carried out for 10 min with CTP, the cognate substrate, or ATP, a noncognate substrate, in the presence of 10 mM Mg^{2+} or 1 mM Mn^{2+} (Figure 6(b)). Similarly to TEC9s, a part of the transcript of the *Syn* TEC18 was shortened during the preparation, as shown in the lanes at time zero in Figure 6(b), probably due to hydrolysis or pyrophosphorolysis in the absence of NTPs. However, in contrast to TEC9, the shortening was not distinct for the *E. coli* TEC18, demonstrating that the *Syn* enzyme tends to shorten the retaining transcript more than the *E. coli* enzyme at Position +18.

In the presence of the cognate CTP, the *Syn* and *E. coli* TEC18s incorporated CMP at Position +19 within 1 min and then slowly misincorporated CMP at Position +20 (Figure 6(b)). The level of the CMP misincorporation at 10 min was lower for the *Syn* TEC18 than for the *E. coli*, being consistent with the more accurate whole elongation by the *Syn* enzyme than by the *E. coli* enzyme. In the presence of the noncognate ATP, both the TEC18s incorporated two AMP molecules, the misincorporation at Position +19 and +20, and again the *Syn* enzyme was more accurate. The level of the successive two AMP incorporations was by an order of magnitude higher for the *E. coli* TEC18 than for the *Syn* one (Figure 6(c)). The *Syn* TEC18 showed higher fidelity than the *E. coli* TEC irrespective of the coexisting divalent cations, also for misincorporating UMP and GMP (Figure 5(c)). Because this misincorporation may be interpreted as the misincorporation at +19 with the substrate at +20, this misincorporation could be due to the misalignment mechanism [37, 38].

In conclusion, the *Syn* enzyme showed higher fidelity than the *E. coli* one, both in initiation and in elongation, higher shortening activity, and slower rates of elongation. For most sets of substrates and enzymes, misincorporations were enhanced by the replacement of Mg^{2+} with Mn^{2+} except for the misincorporation by the *E. coli* TEC of AMP at +20 (Figure 6(c)).

4. Discussion

As shown in Figure 1(b), cyanobacterial RNAPs have large inserted domains in the G region of the β' subunit [14]. In yeast RNAP II, the lack of Rpb9, which can be considered to be the counterpart of a part of the inserted domain of the bacterial β' subunit, makes elongation faster and fidelity lower [4]. Our results share a similar feature: *E. coli* RNAP bearing smaller insertion elongates faster with lower fidelity than the cyanobacterial enzyme bearing a larger insertion. Therefore, the size of the inserted domain could be a determinant of the rate of elongation and the fidelity of an RNAP.

The holoenzyme containing the primary σ factor occupies less than 0.2% of the soluble proteins in cyanobacteria, and the content is at least 3 to 10-fold smaller than that in *E. coli* [24, 39, 40]. This is also consistent with the slower rates of elongation by cyanobacterial RNAP. As proved in *E. coli*, the release of sigma is a time-dependent event costing several seconds [41, 42], but elongation of several kilo base takes several minutes at least. Therefore, slower elongation requires less σ factor. In addition, the less formation of moribund complex also decreases the requirement of σ factor, because the complex contains the factor.

In *E. coli*, moribund complex blocks a promoter at the cost of abortive transcripts which are elongated at a rate of orders of magnitude smaller than the productive transcripts [30]. Therefore, repression by forming moribund complex at a promoter is considered to be less wasteful than abortion of long transcripts by elongation pause or immature termination (Figure S3). The Gre factors enable the conversion of moribund complex into productive complex [31] to eliminate the repression by moribund complex [43]. The factors finely tune the levels of some proteins which are important to the growth in rich nutrient conditions [7, 44] although the factors are not essential. For example, the transcriptional level of *atp* operon encoding FoF1ATPase, the main generator of ATP, is reduced to one fourth in the disruptant of the *gre* genes [7]. Therefore, the Gre factors allow *E. coli* to make a quick uptake of nutrients at the moderate cost of abortive synthesis.

The two cyanobacteria examined in this study, as well as most cyanobacteria, grow in poorer nutrient conditions than *E. coli*. They may not be allowed to use the moderate cost as *E. coli* does with its Gre factors, and this might be the reason why they exceptionally lack the genes orthologous to *greA* which is widely conserved in eubacteria. Consistently, their RNAPs elongate transcripts slower and form less moribund complex, making the Gre factors insignificant.

If the discussion above is the case, why is the branched pathway preserved in the cyanobacterial RNAP? There could be two answers. The first is that the mechanism is inevitably accompanied with the Brownian ratchet mechanism of RNAP elongation [45]. The other is that the cyanobacterial RNAP may use moribund complex only in limited conditions such as low Mg^{2+} concentrations. The ratio of abortive to runoff transcripts was 50 or more at a Mg^{2+} concentration lower than 0.5 mM (Figure 5), which is large enough for the repression by moribund complex to be significant.

The concentration of Mg^{2+} in a stroma of chloroplast has been measured to change from 0.5 mM at dark to 2 mM at light [35]. This increase by illumination is due to the uptake of Mg^{2+} from the inner region of the thylakoid membrane, and electrically compensated by the antitransport of proton from a stroma due to photosynthesis [46, 47]. The cytosolic pH in cyanobacteria has been measured to change as in the stroma between light and dark [48, 49]. In contrast to Mg^{2+}, little is known about the variation of Mn^{2+} concentration in cyanobacterial cells. Although a measured value of the concentration of Mg^{2+} or Mn^{2+} *in vivo* cannot be directly correlated to its concentration effect on *in vitro* transcription (Figure S4), it is an important clue in comparative studies of RNAPs.

Acknowledgments

This work was supported by an NIG scholarship and Leaveanest Grant to M. Imashimizu. The authors thank Ms. Harriet Sallach for reading the paper.

References

[1] V. Sosunov, E. Sosunova, A. Mustaev, I. Bass, V. Nikiforov, and A. Goldfarb, "Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase," *EMBO Journal*, vol. 22, no. 9, pp. 2234–2244, 2003.

[2] E. Zaychikov, E. Martin, L. Denissova et al., "Mapping of catalytic residues in the RNA polymerase active center," *Science*, vol. 273, no. 5271, pp. 107–108, 1996.

[3] J. Hurwitz, L. Yarbrough, and S. Wickner, "Utilization of deoxynucleoside triphosphates by DNA-dependent RNA polymerase of *E. coli*," *Biochemical and Biophysical Research Communications*, vol. 48, no. 3, pp. 628–635, 1972.

[4] C. Walmacq, M. L. Kireeva, J. Irvin et al., "Rpb9 subunit controls transcription fidelity by delaying NTP sequestration in RNA polymerase II," *Journal of Biological Chemistry*, vol. 284, no. 29, pp. 19601–19612, 2009.

[5] S. K. Niyogi, R. P. Feldman, and D. J. Hoffman, "Selective effects of metal ions on RNA synthesis rates," *Toxicology*, vol. 22, no. 1, pp. 9–21, 1981.

[6] S. K. Niyogi and R. P. Feldman, "Effects of several metal ions on misincorporation during transcription," *Nucleic Acids Research*, vol. 9, no. 11, pp. 2615–2527, 1981.

[7] M. Susa, T. Kubori, and N. Shimamoto, "A pathway branching in transcription initiation in *Escherichia coli*," *Molecular Microbiology*, vol. 59, no. 6, pp. 1807–1817, 2006.

[8] N. Keren, M. J. Kidd, J. E. Penner-Hahn, and H. B. Pakrasi, "A light-dependent mechanism for massive accumulation of manganese in the photosynthetic bacterium *Synechocystis* sp. PCC 6803," *Biochemistry*, vol. 41, no. 50, pp. 15085–15092, 2002.

[9] S. Shcolnick and N. Keren, "Metal homeostasis in cyanobacteria and chloroplasts. Balancing benefits and risks to the photosynthetic apparatus," *Plant Physiology*, vol. 141, no. 3, pp. 805–810, 2006.

[10] M. J. Daly, E. K. Gaidamakova, V. Y. Matrosova et al., "Protein oxidation implicated as the primary determinant of bacterial radioresistance.," *PLoS biology*, vol. 5, no. 4, p. e92, 2007.

[11] L. A. Allison, M. Moyle, M. Shales, and C. James Ingles, "Extensive homology among the largest subunits of eukaryotic

and prokaryotic RNA polymerases," *Cell*, vol. 42, no. 2, pp. 599–610, 1985.

[12] G. Zhang, E. A. Campbell, L. Minakhin, C. Richter, K. Severinov, and S. A. Darst, "Crystal structure of Thermus aquaticus core RNA polymerase at 3.3 Å resolution," *Cell*, vol. 98, no. 6, pp. 811–824, 1999.

[13] I. Artsimovitch, V. Svetlov, K. S. Murakami, and R. Landick, "Co-overexpression of *Escherichia coli* RNA polymerase subunits allows isolation and analysis of mutant enzymes lacking lineage-specific sequence insertions," *Journal of Biological Chemistry*, vol. 278, no. 14, pp. 12344–12355, 2003.

[14] L. M. Iyer, E. V. Koonin, and L. Aravind, "Evolutionary connection between the catalytic subunits of DNA-dependent RNA polymerases and eukaryotic RNA-dependent RNA polymerases and the origin of RNA polymerases," *BMC Structural Biology*, vol. 3, pp. 1–23, 2003.

[15] A. Goto-Seki, M. Shirokane, S. Masuda, K. Tanaka, and H. Takahashi, "Specificity crosstalk among group 1 and group 2 sigma factors in the cyanobacterium *Synechococcus* sp. PCC7942: *in vitro* specificity and a phylogenetic analysis," *Molecular Microbiology*, vol. 34, no. 3, pp. 473–484, 1999.

[16] M. Imashimizu, M. Hanaoka, A. Seki, K. S. Murakami, and K. Tanaka, "The cyanobacterial principal σ factor region 1.1 is involved in DNA-binding in the free form and in transcription activity as holoenzyme," *FEBS Letters*, vol. 580, no. 14, pp. 3439–3444, 2006.

[17] D. A. Hager, D. J. Jin, and R. R. Burgess, "Use of mono Q high-resolution ion-exchange chromatography to obtain highly pure and active *Escherichia coli* RNA polymerase," *Biochemistry*, vol. 29, no. 34, pp. 7890–7894, 1990.

[18] J. J. Dunn and F. W. Studier, "Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements," *Journal of Molecular Biology*, vol. 166, no. 4, pp. 477–535, 1983.

[19] T. Kubori and N. Shimamoto, "Physical interference between *Escherichia coli* RNA polymerase molecules transcribing in tandem enhances abortive synthesis and misincorporation," *Nucleic Acids Research*, vol. 25, no. 13, pp. 2640–2647, 1997.

[20] G. J. Schneider and R. Hasekorn, "RNA polymerase subunit homology among cyanobacteria, other eubacteria and archaebacteria," *Journal of Bacteriology*, vol. 170, no. 9, pp. 4136–4140, 1988.

[21] G. Schyns, L. Jia, T. Coursin, N. Tandeau De Marsac, and J. Houmard, "Promoter recognition by a cyanobacterial RNA polymerase: *in vitro* studies with the *Calothrix* sp. PCC 7601 transcriptional factors RcaA and RcaD," *Plant Molecular Biology*, vol. 36, no. 5, pp. 649–659, 1998.

[22] W. Kammerer, U. Deuschle, R. Gentz, and H. Bujard, "Functional dissection of *Escherichia coli* promoters: information in the transcribed region is involved in late steps of the overall process," *EMBO Journal*, vol. 5, no. 11, pp. 2995–3000, 1986.

[23] M. R. Schaefer and S. S. Golden, "Differential expression of members of a cyanobacterial *psbA* gene family in response to light," *Journal of Bacteriology*, vol. 171, no. 7, pp. 3973–3981, 1989.

[24] S. Imamura, S. Yoshihara, S. Nakano et al., "Purification, characterization, and gene expression of all sigma factors of RNA polymerase in a cyanobacterium," *Journal of Molecular Biology*, vol. 325, no. 5, pp. 857–872, 2003.

[25] A. J. Carpousis and J. D. Gralla, "Cycling of ribonucleic acid polymerase to produce oligonucleotides during initiation in vitro at the lac UV5 promoter," *Biochemistry*, vol. 19, no. 14, pp. 3245–3253, 1980.

[26] D. E. Johnston and W. R. McClure, "Abortive initiation of *in vitro* RNA synthesis on bacteriophage DNA," in *RNA Polymerase*, R. Losick and M. Chamberlin, Eds., pp. 413–428, Cold Spring Harbor Laboratory Press, New York, NY, USA, 1976.

[27] L. M. Heisler, G. Feng, D. J. Jin, C. A. Gross, and R. Landick, "Amino acid substitutions in the two largest subunits of *Escherichia coli* RNA polymerase that suppress a defective rho termination factor affect different parts of the transcription complex," *Journal of Biological Chemistry*, vol. 271, no. 24, pp. 14572–14583, 1996.

[28] K. S. Murakami, S. Masuda, and S. A. Darst, "Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 A resolution," *Science*, vol. 296, no. 5571, pp. 1280–1284, 2002.

[29] R. Sen, H. Nagai, V. J. Hernandez, and N. Shimamoto, "Reduction in abortive transcription from the λP(R) promoter by mutations in region 3 of the σ70 subunit of *Escherichia coli* RNA polymerase," *Journal of Biological Chemistry*, vol. 273, no. 16, pp. 9872–9877, 1998.

[30] T. Kubori and N. Shimamoto, "A branched pathway in the early stage of transcription by *Escherichia coli* RNA polymerase," *Journal of Molecular Biology*, vol. 256, no. 3, pp. 449–457, 1996.

[31] R. Sen, H. Nagai, and N. Shimamoto, "Polymerase arrest at the λP(R) promoter during transcription initiation," *Journal of Biological Chemistry*, vol. 275, no. 15, pp. 10899–10904, 2000.

[32] S. Borukhov, V. Sagitov, C. A. Josaitis, R. L. Gourse, and A. Goldfarb, "Two modes of transcription initiation in vitro at the *rrnB* P1 promoter of *Escherichia coli*," *Journal of Biological Chemistry*, vol. 268, no. 31, pp. 23477–23482, 1993.

[33] J. P. Jacques and M. M. Susskind, "Pseudo-templated transcription by *Escherichia coli* RNA polymerase at a mutant promoter," *Genes and Development*, vol. 4, no. 10, pp. 1801–1810, 1990.

[34] R. C. Parker, "Synthesis of in vitro ColE1 transcripts with 5′-terminal ribonucleotides that exhibit noncomplementarity with the DNA template," *Biochemistry*, vol. 25, no. 21, pp. 6593–6598, 1986.

[35] S. Ishijima, A. Uchibori, H. Takagi, R. Maki, and M. Ohnishi, "Light-induced increase in free Mg^{2+} concentration in spinach chloroplasts: measurement of free Mg^{2+} by using a fluorescent probe and necessity of stromal alkalinization," *Archives of Biochemistry and Biophysics*, vol. 412, no. 1, pp. 126–132, 2003.

[36] H. Ito, M. Mutsuda, Y. Murayama et al., "Cyanobacterial daily life with Kai-based circadian and diurnal genome-wide transcriptional control in *Synechococcus elongatus*," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 33, pp. 14168–14173, 2009.

[37] E. Kashkina, M. Anikin, F. Brueckner et al., "Template Misalignment in Multisubunit RNA Polymerases and Transcription Fidelity," *Molecular Cell*, vol. 24, no. 2, pp. 257–266, 2006.

[38] R. T. Pomerantz, D. Temiakov, M. Anikin, D. G. Vassylyev, and W. T. McAllister, "A mechanism of nucleotide misincorporation during transcription due to template-strand misalignment," *Molecular Cell*, vol. 24, no. 2, pp. 245–255, 2006.

[39] M. Jishage and A. Ishihama, "Regulation of RNA polymerase sigma subunit synthesis in *Escherichia coli*: intracellular levels of σ70 and σ38," *Journal of Bacteriology*, vol. 177, no. 23, pp. 6832–6835, 1995.

[40] Y. Iwakura, K. Ito, and A. Ishihama, "Biosynthesis of RNA polymerase in *Escherichia coli*. I. Control of RNA polymerase content at various growth rates," *Molecular and General Genetics*, vol. 133, no. 1, pp. 1–23, 1974.

[41] N. Shimamoto, T. Kamigochi, and H. Utiyama, "Release of the σ subunit of *Escherichia coli* DNA-dependent RNA polymerase depends mainly on time elapsed after the start of initiation, not on length of product RNA," *Journal of Biological Chemistry*, vol. 261, no. 25, pp. 11859–11865, 1986.

[42] M. Raffaelle, E. I. Kanin, J. Vogt, R. R. Burgess, and A. Z. Ansari, "Holoenzyme switching and stochastic release of sigma factors from RNA polymerase *in vivo*," *Molecular Cell*, vol. 20, no. 3, pp. 357–366, 2005.

[43] R. Sen, H. Nagai, and N. Shimamoto, "Conformational switching of *Escherichia coli* RNA polymerase-promoter binary complex is facilitated by elongation factor GreA and GreB," *Genes to Cells*, vol. 6, no. 5, pp. 389–401, 2001.

[44] E. Stepanova, J. Lee, M. Ozerova et al., "Analysis of promoter targets for *Escherichia coli* transcription elongation factor GreA *in vivo* and *in vitro*," *Journal of Bacteriology*, vol. 189, no. 24, pp. 8772–8785, 2007.

[45] G. Bar-Nahum, V. Epshtein, A. E. Ruckenstein, R. Rafikov, A. Mustaev, and E. Nudler, "A ratchet mechanism of transcription elongation and its control," *Cell*, vol. 120, no. 2, pp. 183–193, 2005.

[46] G. H. Krause, "Light induced movement of magnesium ions in intact chloroplasts. Spectroscopic determination with Eriochrome Blue SE," *Biochimica et Biophysica Acta*, vol. 460, no. 3, pp. 500–510, 1977.

[47] G. Hind, H. Y. Nakatani, and S. Izawa, "Light dependent redistribution of ions in suspensions of chloroplast thylakoid membranes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 71, no. 4, pp. 1484–1488, 1974.

[48] H. W. Heldt, K. Werdan, M. Milovancev, and G. Geller, "Alkalization of the chloroplast stroma caused by light dependent proton flux into the thylakoid space," *Biochimica et Biophysica Acta*, vol. 314, no. 2, pp. 224–241, 1973.

[49] S. Belkin, R. J. Mehlhorn, and L. Packer, "Proton gradients in intact cyanobacteria," *Plant physiology*, vol. 84, pp. 25–30, 1987.

SAGE-Hindawi Access to Research
Genetics Research International
Volume 2011, Article ID 206290, 7 pages
doi:10.4061/2011/206290

Review Article

The Transition of Poised RNA Polymerase II to an Actively Elongating State Is a "Complex" Affair

Marie N. Yearling, Catherine A. Radebaugh, and Laurie A. Stargell

Biochemistry and Molecular Biology, Colorado State University, Fort Collins, CO 80523-1870, USA

Correspondence should be addressed to Laurie A. Stargell, laurie.stargell@colostate.edu

Received 24 June 2011; Accepted 31 July 2011

Academic Editor: Sebastián Chávez

The initial discovery of the occupancy of RNA polymerase II at certain genes *prior* to their transcriptional activation occurred a quarter century ago in *Drosophila*. The preloading of these poised complexes in this inactive state is now apparent in many different organisms across the evolutionary spectrum and occurs at a broad and diverse set of genes. In this paper, we discuss the genetic and biochemical efforts in *S. cerevisiae* to describe the conversion of these poised transcription complexes to the active state for productive elongation. The accumulated evidence demonstrates that a multitude of coactivators and chromatin remodeling complexes are essential for this transition.

1. Introduction

RNA Polymerase II (RNAPII) is a 12-subunit enzyme that binds promoter DNA and catalyzes the synthesis of messenger RNA in eukaryotes. Although the recruitment of RNAPII to a promoter is necessary for productive gene expression, it is not sufficient in many cases. Early studies in *Drosophila* [1] and more recent genome-wide analyses in both flies and humans have revealed that thousands of genes contain poised RNAPII at their promoters [2–5]. These poised promoters allow for rapid and synchronous activation, thereby providing the precise timing of gene expression critical for developmental processes [6, 7]. Indeed, postrecruitment events necessary to convert RNAPII into a productively elongating form are increasingly considered general regulatory features of transcription in higher eukaryotes [8–10].

In yeast cultured to stationary phase, approximately 40% of the genes in the genome show association of RNAPII in their inactive state [11]. These polymerases are thought to be poised for rapid and concerted activation upon transition to more opportunistic growth conditions. In actively growing yeast cultures, genome-wide studies indicate that partial but inactive PIC complexes are a widespread phenomenon across the genome [12] and a majority of bound RNAPII may be in an inactive state [11, 13]. Gene regulation at postrecruitment steps in *S. cerevisiae* is also supported by differences in 5′ to 3′ RNAPII occupancy and the frequent pausing of RNAPII within coding regions of genes [14, 15]. In addition, accumulation of inactive RNAPII within ribosomal protein genes [16] and at the promoter of the uninduced *CYC1* gene [17] provides further support for postrecruitment transcriptional regulation in yeast. Due to the genetic and biochemical amenability of the yeast system, studies of the transition of poised RNAPII to the active form have provided key insights into the sophisticated molecular requirements involved in this postrecruitment process.

2. The Yeast *CYC1* Gene: A Model for Postrecruitment Regulation via Poised RNAPII

The yeast *CYC1* gene encodes iso-1-cytochrome c, a nuclear-encoded protein involved in the electron transport chain in the mitochondria [18]. In the presence of a fermentable carbon source (such as dextrose), *CYC1* gene expression is extremely low [19, 20]. When cells are grown on a nonfermentable carbon source (such as lactate or ethanol), *CYC1* is activated and transcript levels increase 10-fold. In contrast to the dramatic changes in transcriptional output,

the occupancy of RNAPII [17, 21], as well as a number of other factors [22, 23], is maintained during the carbon source change (Figure 1). The *CYC1* promoter contains preloaded RNAPII, the general transcription factors TATA-binding protein (TBP) and TFIIH, the SAGA (Spt-Ada-Gcn5 acetyltransferase) complex, and Spn1, a highly conserved chromatin-associated transcription factor [22, 23]. Intriguingly, RNAPII is serine 5 phosphorylated on the C-terminal domain (CTD) of Rpb1 prior to activation [23]. The CTD is hypophosphorylated prior to initiation and typically becomes serine 5 hyperphosphorylated during the transition from initiation to elongation [24, 25]. The phosphorylation of the CTD at *CYC1* prior to activation is consistent with TFIIH occupancy, since TFIIH has CTD kinase as well as DNA unwinding activities [26–30]. Under inducing conditions for *CYC1,* a number of new factors are recruited to the promoter including the Mediator complex, and the chromatin regulatory factors the Swi/Snf complex and Spt6 [22]. This poised promoter could be advantageous in the native environment, allowing for rapid induction due to changing nutritional needs [31].

3. The Role of SAGA in the Inactive-to-Active Transition

The SAGA complex is a large multisubunit coactivator that facilitates gene expression at multiple steps within the transcription cycle [32, 33], including initiation [34–42] and more recently identified activities in the stimulation of elongation [33]. SAGA localization within gene coding regions [43–47] and elongation defects in SAGA deficient strains [46, 48, 49] demonstrate that the function of SAGA in transcription activation extends beyond the well-characterized activities of TBP delivery and posttranslational modifications of histones. It is unclear how the emerging functions in elongation pertain to the traditional roles of SAGA except at *CYC1,* where studies indicate that they appear to be functionally distinct.

The poised *CYC1* promoter requires SAGA for the transition from a preloaded complex to an actively transcribing unit since deletion of SAGA-integrity subunits blocks activated transcription [22]. Several well-characterized functions of SAGA are not relevant to this transition. For example, SAGA contains a TBP interaction module essential for delivering TBP to certain promoters [34–38]. Since the preloaded promoter has both TBP and SAGA present under noninducing conditions, a functional connection between the two seemed likely. Surprisingly, although abolishing the SAGA complex results in loss of activated transcription, it does not alter TBP occupancy [22]. SAGA also has two known histone modifying enzymatic capabilities, a histone acetyltransferase (HAT) module responsible for acetylation events involved in facilitating active transcription [39–42] and a histone deubiquitinase (DUB) module known to aid in elongation [50, 51]. Yet, strains deficient for HAT activity or the DUB module are competent for activation [22].

In summary, the preloaded promoter is not dependent on the traditional well-characterized roles of SAGA, and yet SAGA integrity is required for the transition to an actively elongating complex after the recruitment of the PIC. These elongation activities may also be important at other genes, but difficult to observe because those genes require SAGA for recruitment of the general transcription machinery. The functions of numerous components within the SAGA complex remain to be elucidated and the preloaded promoter provides an excellent archetype for further investigations. Despite the necessity for SAGA function, SAGA is not sufficient for activation and another coactivator is critical for induction of the poised promoter.

4. Mediator-RNAPII Connections at *CYC1*

The Mediator complex is a large coactivator that is conserved from yeast to humans [52] and acts as an integrator of the transcription process, traditionally linking upstream signals from the activator with the general transcription machinery [53–57]. Mediator is essential for *CYC1* activation and is recruited after the transfer to inducing conditions [22]. Mediator is well characterized for its ability to recruit RNAPII to promoters [57, 58], although this function is unnecessary for *CYC1* since RNAPII is present at the poised promoter prior to activation. Mediator has also been shown to stimulate TFIIH-dependent phosphorylation of the CTD [57, 59]. However, as previously noted, serine 5 phosphorylation of the CTD at *CYC1* is observed prior to activation when Mediator is absent from the promoter, although subsequent rounds of transcription may be impacted. Mediator has also been shown to be involved in the isomerization of the PIC into a transcriptionally competent conformation [60]. This function fits well with the Mediator requirement for activation of the poised promoter as the subunits of Mediator essential for activating the poised promoter [22] are involved in interactions with RNAPII and the general transcription factors [54, 57, 61].

5. Chromatin Components with Critical Roles in the Transition to Active RNAPII

The transition of the poised promoter to its actively elongating form is highly dependent on a number of chromatin regulatory factors, including the Spn1/Spt6 [62, 63] complex and the Swi/Snf complex [23]. Notably, RNAPII and Spn1 occupy the poised promoter in the uninduced state, whereas Spt6 and Swi/Snf are recruited upon activation (Figure 1). Spn1 interacts with both RNAPII and Spt6 [23, 64–69], thereby linking the regulation of the poised promoter to the chromatin architecture. Spt6 is a histone chaperone that promotes reassembly of nucleosomes following passage of RNAPII [70–74], and Spn1 is an important regulator of the Spt6-nucleosome interaction [75]. In addition to a nucleosome maintenance role during elongation, Spt6 also has other chromatin-dependent [76] and chromatin-independent [77] roles in transcription. Importantly, the loss of Spn1 at *CYC1* under noninducing conditions leads to a failure to recruit Spt6 under inducing conditions [23], consistent with their direct interaction [75, 78]. Mutations

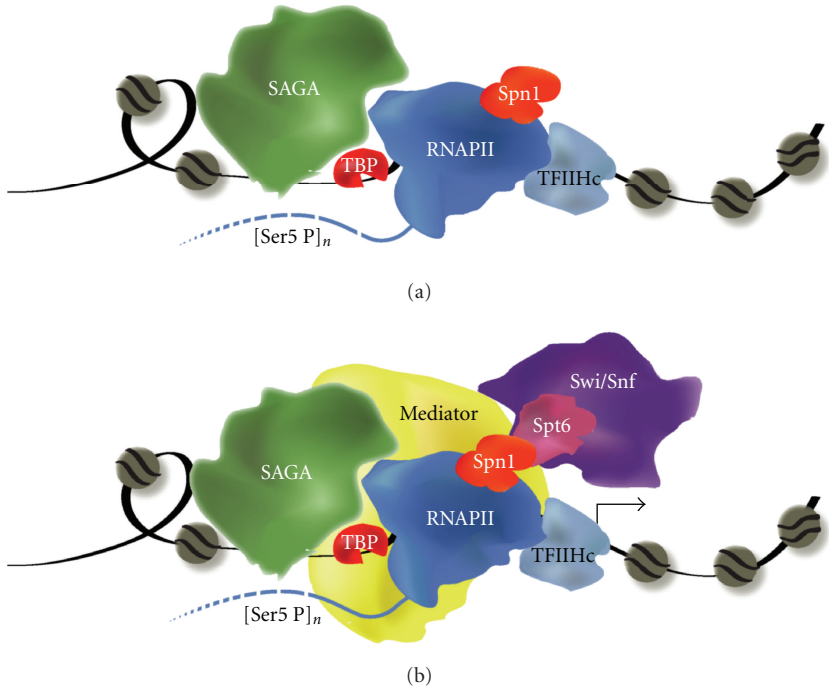


FIGURE 1: The poised *CYC1* promoter contains preloaded transcription components. (a) Prior to activation, the preloaded *CYC1* promoter contains TATA binding protein (TBP), RNA polymerase II (RNAPII), the core TFIIH complex (TFIIHc), Spt-Ada-GCN5 acetyltransferase (SAGA), and the transcription factor Spn1. The CTD, shown by the hashed line trailing RNAPII exhibits serine 5 phosphorylation potentially on multiple repeats (denoted by "*n*"). These components occupy the promoter prior to high levels of transcriptional output. (b) The occupancy of the preloaded factors is maintained under induced conditions, and Mediator, Spt6, and Swi/Snf are recruited, leading to an increase in transcriptional output (indicated by the arrow).

in either Spt6 [70, 79] or Spn1 [23] suppress mutant phenotypes associated with the loss of the Swi/Snf complex.

Swi/Snf is an ATP-dependent chromatin remodeler that disassembles nucleosomes resulting in a loss of histones from DNA [80]. Swi/Snf is involved in remodeling at several recruitment-regulated promoters [81–84] as well as in coding regions [85, 86]. Suppressing mutations as observed for Spn1, Spt6, and Swi/Snf are typical indicators of factors that function in the same pathway and are physically connected [87, 88]. Specifically at *CYC1*, the absence of promoter binding by Spn1 results in constitutive recruitment of the Swi/Snf complex [23]. Thus, the binding of Spn1 blocks the recruitment of the Swi/Snf complex in the uninduced state and also serves as a platform for recruiting Spt6 during the activated state. An attractive model for the functions of these factors in proper *CYC1* expression is that, in order for the transition to an actively elongating state to occur, the Swi/Snf complex evicts nucleosomes and the Spn1/Spt6 complex reassembles them. Precisely how these activities are related to the poised RNAPII in the uninduced state remains to be investigated, but it is tempting to speculate that the interplay between RNAPII and the nucleosomal architecture contributes to the inactive state. The involvement of the chromatin context and inactive RNAPII complexes has also been observed at particular silent loci in the yeast genome [89].

6. Poising as a "Complex" Affair

Several important questions remain. For example, what creates the poised polymerase in the first place? We have found that the occupancy of RNAPII at *CYC1* is an incredibly robust phenomenon: single deletion of dozens of different transcription factors and coactivator complex subunits has not resulted in RNAPII occupancy defects (data not shown and [22]). It could be that RNAPII preloading is an intrinsic property of the *CYC1* promoter and/or its nuclear and chromosomal context. In contrast to the resiliency of RNAPII occupancy, the transition to an actively elongating form is a highly demanding phenomenon, requiring the efforts of several prodigious and powerful transcription complexes: SAGA, Mediator, and Swi/Snf. Intriguingly, these complexes and their functions appear to work autonomously at *CYC1*. As shown previously [22], Mediator and SAGA occupancy are not dependent on each other, and mutations that result in constitutive occupancy of Swi/Snf do not bypass the need for SAGA or Mediator for activation of the preloaded promoter (Figure 2). As such, three distinct pathways are required to shift the polymerase into its active form. Further studies are needed to elucidate how each complex directly contributes to the transition from the poised to the active form. However, it is clear that in accordance with Newton's first law (a body at rest tends to stay at rest), these large macromolecular

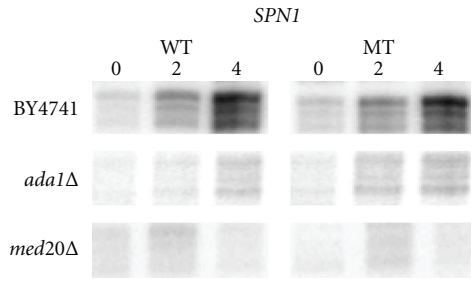


Figure 2: Mutating *SPN1*, which results in loss of Spn1 and constitutive recruitment of Swi/Snf to the promoter, does not bypass the requirement for SAGA or Mediator. *CYC1* transcript levels were analyzed before and two or four hours after induction in ethanol using an S1 nuclease protection assay [23] with RNA isolated from wild-type, *ada1*Δ or *med20*Δ strains. Each strain harbors either a wild-type (WT) or mutant (MT) form of *SPN1*. Similar results were obtained for other SAGA and Mediator deletion strains, including *gcn5*Δ, *spt7*Δ, *spt8*Δ, *spt20*Δ, *med5*Δ, *med15*Δ, and *med18*Δ (data not shown).

assemblies must provide the essential outside forces to initiate the process.

7. Perspectives

How related is the RNAPII poising observed in yeast to that in metazoans? Studies in flies and human cells have clearly established that in many cases the polymerase has started transcribing and is paused just downstream of the start site. In contrast, there is no evidence for initiated transcripts that are stalled in yeast [13, 90]. Whether the poised RNAPII in yeast is an evolutionary precursor to the more sophisticated version of paused RNAPII in metazoans is an open question. Nevertheless, it is intriguing that occupancy of SAGA [47, 91], a requirement for Mediator [53, 60, 92], a dependency on Spt6 [74], the involvement of Spn1 [73], and the chromatin architecture [93] play critical roles in pausing and/or postrecruitment transcriptional events in metazoan cells. Taken together, these results suggest that there are universal requirements for the activities of multiple complexes in the transition of RNAPII from a poised to an actively elongating state.

Acknowledgments

The authors would like to thank Adam Almeida, Sarah Lee, and Tingting Yao for critical reading and feedback. The work in the Stargell laboratory is supported by grants to L.A.S. from the National Science Foundation (MCB-0843073) and National Institutes of Health (GM088409).

References

[1] D. S. Gilmour and J. T. Lis, "RNA polymerase II interacts with the promoter region of the noninduced hsp70 gene in *Drosophila melanogaster* cells," *Molecular and Cellular Biology*, vol. 6, no. 11, pp. 3984–3989, 1986.

[2] T. H. Kim, L. O. Barrera, M. Zheng et al., "A high-resolution map of active promoters in the human genome," *Nature*, vol. 436, no. 7052, pp. 876–880, 2005.

[3] M. G. Guenther, S. S. Levine, L. A. Boyer, R. Jaenisch, and R. A. Young, "A chromatin landmark and transcription initiation at most promoters in human cells," *Cell*, vol. 130, no. 1, pp. 77–88, 2007.

[4] G. W. Muse, D. A. Gilchrist, S. Nechaev et al., "RNA polymerase is poised for activation across the genome," *Nature Genetics*, vol. 39, no. 12, pp. 1507–1511, 2007.

[5] I. M. Min, J. J. Waterfall, L. J. Core, R. J. Munroe, J. Schimenti, and J. T. Lis, "Regulating RNA polymerase pausing and transcription elongation in embryonic stem cells," *Genes and Development*, vol. 25, no. 7, pp. 742–754, 2011.

[6] J. Zeitlinger, A. Stark, M. Kellis et al., "RNA polymerase stalling at developmental control genes in the *Drosophila melanogaster* embryo," *Nature Genetics*, vol. 39, no. 12, pp. 1512–1516, 2007.

[7] R. N. Saha, E. M. Wissink, E. R. Bailey et al., "Rapid activity-induced transcription of Arc and other IEGs relies on poised RNA polymerase II," *Nature Neuroscience*, vol. 14, no. 7, pp. 848–856, 2011.

[8] L. A. Selth, S. Sigurdsson, and J. Q. Svejstrup, "Transcript elongation by RNA polymerase II," *Annual Review of Biochemistry*, vol. 79, pp. 271–293, 2010.

[9] M. Levine, "Paused RNA polymerase II as a developmental checkpoint," *Cell*, vol. 145, no. 4, pp. 502–511, 2011.

[10] 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, no. 1, pp. 34–45, 2011.

[11] M. Radonjic, J. C. Andrau, P. Lijnzaad et al., "Genome-wide analyses reveal RNA polymerase II located upstream of genes poised for rapid response upon *S. cerevisiae* stationary phase exit," *Molecular Cell*, vol. 18, no. 2, pp. 171–183, 2005.

[12] S. J. Zanton and B. F. Pugh, "Full and partial genome-wide assembly and disassembly of the yeast transcription machinery in response to heat shock," *Genes and Development*, vol. 20, no. 16, pp. 2250–2265, 2006.

[13] V. Pelechano, S. Chávez, and J. E. Pérez-Ortín, "A complete set of nascent transcription rates for yeast genes," *PLoS ONE*, vol. 5, no. 11, Article ID e15442, 2010.

[14] B. J. Venters and B. F. Pugh, "A canonical promoter organization of the transcription machinery and its regulators in the *Saccharomyces* genome," *Genome Research*, vol. 19, no. 3, pp. 360–371, 2009.

[15] L. S. Churchman and J. S. Weissman, "Nascent transcript sequencing visualizes transcription at nucleotide resolution," *Nature*, vol. 469, no. 7330, pp. 368–373, 2011.

[16] V. Pelechano, S. Jimeno-González, A. Rodríguez-Gil, J. García-Martínez, J. E. Pérez-Ortín, and S. Chávez, "Regulon-specific control of transcription elongation across the yeast genome," *PLoS Genetics*, vol. 5, no. 8, Article ID e1000614, 2009.

[17] C. Martens, B. Krett, and P. J. Laybourn, "RNA polymerase II and TBP occupy the repressed *CYC1* promoter," *Molecular Microbiology*, vol. 40, no. 4, pp. 1009–1019, 2001.

[18] F. Sherman, J. W. Stewart, E. Margoliash, J. Parker, and W. Campbell, "The structural gene for yeast cytochrome C," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 55, no. 6, pp. 1498–1504, 1966.

[19] L. Guarente and T. Mason, "Heme regulates transcription of the *CYC1* gene of *S. cerevisiae* via an upstream activation site," *Cell*, vol. 32, no. 4, pp. 1279–1286, 1983.

[20] L. Guarente, B. Lalonde, P. Gifford, and E. Alani, "Distinctly regulated tandem upstream activation sites mediate catabolite

repression of the *CYC1* gene of *S. cerevisiae*," *Cell*, vol. 36, no. 2, pp. 503–511, 1984.

[21] L. Kuras and K. Struhl, "Binding of TBP to promoters in vivo is stimulated by activators and requires Pol II holoenzyme," *Nature*, vol. 399, no. 6736, pp. 609–613, 1999.

[22] S. K. Lee, A. G. L. Fletcher, L. Zhang, X. Chen, J. A. Fischbeck, and L. A. Stargell, "Activation of a poised RNAPII-dependent promoter requires both SAGA and mediator," *Genetics*, vol. 184, no. 3, pp. 659–672, 2010.

[23] L. Zhang, A. G. L. Fletcher, V. Cheung, F. Winston, and L. A. Stargell, "Spn1 regulates the recruitment of Spt6 and the Swi/Snf complex during transcriptional activation by RNA polymerase II," *Molecular and Cellular Biology*, vol. 28, no. 4, pp. 1393–1403, 2008.

[24] 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.

[25] C. Cheng and P. A. Sharp, "RNA polymerase II accumulation in the promoter-proximal region of the dihydrofolate reductase and γ-actin genes," *Molecular and Cellular Biology*, vol. 23, no. 6, pp. 1961–1967, 2003.

[26] W. J. Feaver, O. Gileadi, Y. Li, and R. D. Kornberg, "CTD kinase associated with yeast RNA polymerase II initiation factor b," *Cell*, vol. 67, no. 6, pp. 1223–1230, 1991.

[27] H. Lu, L. Zawel, L. Fisher, J. M. Egly, and D. Reinberg, "Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II," *Nature*, vol. 358, no. 6388, pp. 641–645, 1992.

[28] H. Serizawa, R. C. Conaway, and J. W. Conaway, "A carboxyl-terminal-domain kinase associated with RNA polymerase II transcription factor d from rat liver," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 89, no. 16, pp. 7476–7480, 1992.

[29] H. Serizawa, J. W. Conaway, and R. C. Conaway, "Phosphorylation of C-terminal domain of RNA polymerase II is not required in basal transcription," *Nature*, vol. 363, no. 6427, pp. 371–374, 1993.

[30] L. Schaeffer, R. Roy, S. Humbert et al., "DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor," *Science*, vol. 259, no. 5104, pp. 58–63, 1993.

[31] M. Johnston, "Feasting, fasting and fermenting: glucose sensing in yeast and other cells," *Trends in Genetics*, vol. 15, no. 1, pp. 29–33, 1999.

[32] E. Koutelou, C. L. Hirsch, and S. Y. R. Dent, "Multiple faces of the SAGA complex," *Current Opinion in Cell Biology*, vol. 22, no. 3, pp. 374–382, 2010.

[33] S. Rodriguez-Navarro, "Insights into SAGA function during gene expression," *EMBO Reports*, vol. 10, no. 8, pp. 843–850, 2009.

[34] A. M. Dudley, C. Rougeulle, and F. Winston, "The Spt components of SAGA facilitate TBP binding to a promoter at a post-activator-binding step in vivo," *Genes and Development*, vol. 13, no. 22, pp. 2940–2945, 1999.

[35] R. Belotserkovskaya, D. E. Sterner, M. Deng, M. H. Sayre, P. M. Lieberman, and S. L. Berger, "Inhibition of TATA-binding protein function by SAGA subunits Spt3 and Spt8 at Gcn4-activated promoters," *Molecular and Cellular Biology*, vol. 20, no. 2, pp. 634–647, 2000.

[36] S. R. Bhaumik and M. R. Green, "SAGA is an essential in vivo target of the yeast acidic activator Gal4p," *Genes and Development*, vol. 15, no. 15, pp. 1935–1945, 2001.

[37] E. Larschan and F. Winston, "The *S. cerevisiae* SAGA complex functions in vivo as a coactivator for transcriptional activation by Gal4," *Genes and Development*, vol. 15, no. 15, pp. 1946–1956, 2001.

[38] H. Qiu, C. Hu, F. Zhang et al., "Interdependent recruitment of SAGA and Srb mediator by transcriptional activator Gcn4p," *Molecular and Cellular Biology*, vol. 25, no. 9, pp. 3461–3474, 2005.

[39] J. E. Brownell, J. Zhou, T. Ranalli et al., "*Tetrahymena* histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation," *Cell*, vol. 84, no. 6, pp. 843–851, 1996.

[40] P. A. Grant, L. Duggan, J. Cote et al., "Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an ada complex and the saga (spt/ada) complex," *Genes and Development*, vol. 11, no. 13, pp. 1640–1650, 1997.

[41] P. A. Grant, D. Schieltz, M. G. Pray-Grant et al., "A subset of TAF(II)s are integral components of the SAGA complex required for nucleosome acetylation and transcriptional stimulation," *Cell*, vol. 94, no. 1, pp. 45–53, 1998.

[42] R. T. Utley, K. Ikeda, P. A. Grant et al., "Transcriptional activators direct histone acetyltransferase complexes to nucleosomes," *Nature*, vol. 394, no. 6692, pp. 498–502, 1998.

[43] C. Desmoucelles, B. Pinson, C. Saint-Marc, and B. Daignan-Fornier, "Screening the yeast "Disruptome" for mutants affecting resistance to the immunosuppressive drug, mycophenolic acid," *Journal of Biological Chemistry*, vol. 277, no. 30, pp. 27036–27044, 2002.

[44] C. K. Govind, F. Zhang, H. Qiu, K. Hofmeyer, and A. G. Hinnebusch, "Gcn5 promotes acetylation, eviction, and methylation of nucleosomes in transcribed coding regions," *Molecular Cell*, vol. 25, no. 1, pp. 31–42, 2007.

[45] A. Wyce, K. W. Henry, S. L. Berger, Verdin, Gu, and Yao, "H2B ubiquitylation and de-ubiquitylation in gene activation," *Novartis Foundation Symposium*, vol. 259, pp. 63–73, 2004.

[46] P. Pascual-García, C. K. Govind, E. Queralt et al., "Sus1 is recruited to coding regions and functions during transcription elongation in association with SAGA and TREX2," *Genes and Development*, vol. 22, no. 20, pp. 2811–2822, 2008.

[47] V. M. Weake, J. O. Dyer, C. Seidel et al., "Post-transcription initiation function of the ubiquitous SAGA complex in tissue-specific gene activation," *Genes & Development*, vol. 25, pp. 1499–1509, 2011.

[48] M. Morillo-Huesca, M. Vanti, and S. Chávez, "A simple in vivo assay for measuring the efficiency of gene length-dependent processes in yeast mRNA biogenesis," *FEBS Journal*, vol. 273, no. 4, pp. 756–769, 2006.

[49] M. Sanso, I. Vargas-Perez, L. Quintales, F. Antequera, J. Ayte, and E. Hidalgo, "Gcn5 facilitates Pol II progression, rather than recruitment to nucleosome-depleted stress promoters, in *Schizosaccharomyces pombe*," *Nucleic Acids Research*, vol. 39, no. 15, pp. 6369–6379.

[50] K. W. Henry, A. Wyce, W. S. Lo et al., "Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8," *Genes and Development*, vol. 17, no. 21, pp. 2648–2663, 2003.

[51] J. A. Daniel, M. S. Torok, Z. W. Sun et al., "Deubiquitination of histone H2B by a yeast acetyltransferase complex regulates transcription," *Journal of Biological Chemistry*, vol. 279, no. 3, pp. 1867–1871, 2004.

[52] H. M. Bourbon, "Comparative genomics supports a deep evolutionary origin for the large, four-module transcriptional mediator complex," *Nucleic Acids Research*, vol. 36, no. 12, pp. 3993–4008, 2008.

[53] K. D. Meyer, S. C. Lin, C. Bernecky, Y. Gao, and D. J. Taatjes, "P53 activates transcription by directing structural shifts in Mediator," *Nature Structural and Molecular Biology*, vol. 17, no. 6, pp. 753–760, 2010.

[54] J. A. Davis, Y. Takagi, R. D. Kornberg, and F. J. Asturias, "Structure of the yeast RNA polymerase II holoenzyme: mediator conformation and polymerase interaction," *Molecular Cell*, vol. 10, no. 2, pp. 409–415, 2002.

[55] G. Cai, T. Imasaki, K. Yamada, F. Cardelli, Y. Takagi, and F. J. Asturias, "Mediator Head module structure and functional interactions," *Nature Structural and Molecular Biology*, vol. 17, no. 3, pp. 273–279, 2010.

[56] H. Sakurai and T. Fukasawa, "Functional connections between mediator components and general transcription factors of *Saccharomyces cerevisiae*," *Journal of Biological Chemistry*, vol. 275, no. 47, pp. 37251–37256, 2000.

[57] C. Esnault, Y. Ghavi-Helm, S. Brun et al., "Mediator-dependent recruitment of TFIIH modules in preinitiation complex," *Molecular Cell*, vol. 31, no. 3, pp. 337–346, 2008.

[58] R. K. Biddick, G. L. Law, K. K. B. Chin, and E. T. Young, "The transcriptional coactivators SAGA, SWI/SNF, and mediator make distinct contributions to activation of glucose-repressed genes," *Journal of Biological Chemistry*, vol. 283, no. 48, pp. 33101–33109, 2008.

[59] Y. J. Kim, S. Bjorklund, Y. Li, M. H. Sayre, and R. D. Kornberg, "A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II," *Cell*, vol. 77, no. 4, pp. 599–608, 1994.

[60] G. Wang, M. A. Balamotis, J. L. Stevens, Y. Yamaguchi, H. Handa, and A. J. Berk, "Mediator requirement for both recruitment and postrecruitment steps in transcription initiation," *Molecular Cell*, vol. 17, no. 5, pp. 683–694, 2005.

[61] J. Soutourina, S. Wydau, Y. Ambroise, C. Boschiero, and M. Werner, "Direct interaction of RNA polymerase II and mediator required for transcription in vivo," *Science*, vol. 331, no. 6023, pp. 1451–1454, 2011.

[62] J. A. Fischbeck, S. M. Kraemer, and L. A. Stargell, "SPN1, a conserved gene identified by suppression of a postrecruitment-defective yeast TATA-binding protein mutant," *Genetics*, vol. 162, no. 4, pp. 1605–1616, 2002.

[63] V. Pujari, C. A. Radebaugh, J. V. Chodaparambil et al., "The transcription factor Spn1 regulates gene expression via a highly conserved novel structural motif," *Journal of Molecular Biology*, vol. 404, no. 1, pp. 1–15, 2010.

[64] D. L. Lindstrom, S. L. Squazzo, N. Muster et al., "Dual roles for Spt5 in pre-mRNA processing and transcription elongation revealed by identification of Spt5-associated proteins," *Molecular and Cellular Biology*, vol. 23, no. 4, pp. 1368–1378, 2003.

[65] N. J. Krogan, M. Kim, S. H. Ahn et al., "RNA polymerase II elongation factors of *Saccharomyces cerevisiae*: a targeted proteomics approach," *Molecular and Cellular Biology*, vol. 22, no. 20, pp. 6979–6992, 2002.

[66] A. C. Gavin, M. Bösche, R. Krause et al., "Functional organization of the yeast proteome by systematic analysis of protein complexes," *Nature*, vol. 415, no. 6868, pp. 141–147, 2002.

[67] 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.

[68] K. Tarassov, V. Messier, C. R. Landry et al., "An in vivo map of the yeast protein interactome," *Science*, vol. 320, no. 5882, pp. 1465–1470, 2008.

[69] L. Li, H. Ye, H. Guo, and Y. Yin, "*Arabidopsis* IWS1 interacts with transcription factor BES1 and is involved in plant steroid hormone brassinosteroid regulated gene expression," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 8, pp. 3918–3923, 2010.

[70] A. Bortvin and F. Winston, "Evidence that Spt6p controls chromatin structure by a direct interaction with histones," *Science*, vol. 272, no. 5267, pp. 1473–1476, 1996.

[71] C. D. Kaplan, L. Laprade, and F. Winston, "Transcription elongation factors repress transcription initiation from cryptic sites," *Science*, vol. 301, no. 5636, pp. 1096–1099, 2003.

[72] M. W. Adkins and J. K. Tyler, "Transcriptional activators are dispensable for transcription in the absence of Spt6-mediated chromatin reassembly of promoter regions," *Molecular Cell*, vol. 21, no. 3, pp. 405–416, 2006.

[73] 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.

[74] M. B. Ardehali, J. Yao, K. Adelman et al., "Spt6 enhances the elongation rate of RNA polymerase II in vivo," *EMBO Journal*, vol. 28, no. 8, pp. 1067–1077, 2009.

[75] S. M. McDonald, D. Close, H. Xin, T. Formosa, and C. P. Hill, "Structure and biological importance of the Spn1-Spt6 interaction, and its regulatory role in nucleosome binding," *Molecular Cell*, vol. 40, no. 5, pp. 725–735, 2010.

[76] I. Ivanovska, P. -É. Jacques, O. J. Rando, F. Robert, and F. Winston, "Control of chromatin structure by Spt6: different consequences in coding and regulatory regions," *Molecular and Cellular Biology*, vol. 31, no. 3, pp. 531–541, 2011.

[77] M. Endoh, W. Zhu, J. Hasegawa et al., "Human Spt6 stimulates transcription elongation by RNA polymerase II in vitro," *Molecular and Cellular Biology*, vol. 24, no. 8, pp. 3324–3336, 2004.

[78] M. L. Diebold, M. Koch, E. Loeliger et al., "The structure of an Iws1/Spt6 complex reveals an interaction domain conserved in TFIIS, Elongin A and Med26," *EMBO Journal*, vol. 29, no. 23, pp. 3979–3991, 2010.

[79] L. Neigeborn, K. Rubin, and M. Carlson, "Suppressors of SNF2 mutations restore invertase derepression and cause temperature-sensitive lethality in yeast," *Genetics*, vol. 112, no. 4, pp. 741–753, 1986.

[80] Y. Lorch, B. Maier-Davis, and R. D. Kornberg, "Chromatin remodeling by nucleosome disassembly in vitro," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 9, pp. 3090–3093, 2006.

[81] M. W. Adkins, S. K. Williams, J. Linger, and J. K. Tyler, "Chromatin disassembly from the PHO5 promoter is essential for the recruitment of the general transcription machinery and coactivators," *Molecular and Cellular Biology*, vol. 27, no. 18, pp. 6372–6382, 2007.

[82] T. Gkikopoulos, K. M. Havas, H. Dewar, and T. Owen-Hughes, "SWI/SNF and Asf1p cooperate to displace histones during induction of the *Saccharomyces cerevisiae* HO promoter," *Molecular and Cellular Biology*, vol. 29, no. 15, pp. 4057–4066, 2009.

[83] P. Korber, S. Barbaric, T. Luckenbach et al., "The histone chaperone Asf1 increases the rate of histone eviction at the yeast PHO5 and PHO8 promoters," *Journal of Biological Chemistry*, vol. 281, no. 9, pp. 5539–5545, 2006.

[84] S. Takahata, Y. Yu, and D. J. Stillman, "FACT and Asf1 regulate nucleosome dynamics and coactivator binding at the HO promoter," *Molecular Cell*, vol. 34, no. 4, pp. 405–415, 2009.

[85] M. A. Schwabish and K. Struhl, "The Swi/Snf complex is important for histone eviction during transcriptional activation and RNA polymerase II elongation in vivo," *Molecular and Cellular Biology*, vol. 27, no. 20, pp. 6987–6995, 2007.

[86] S. Shivaswamy and V. R. Iyer, "Stress-dependent dynamics of global chromatin remodeling in yeast: dual role for SWI/SNF in the heat shock stress response," *Molecular and Cellular Biology*, vol. 28, no. 7, pp. 2221–2234, 2008.

[87] S. R. Collins, K. M. Miller, N. L. Maas et al., "Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map," *Nature*, vol. 446, no. 7137, pp. 806–810, 2007.

[88] A. Roguev, S. Bandyopadhyay, M. Zofall et al., "Conservation and rewiring of functional modules revealed by an epistasis map in fission yeast," *Science*, vol. 322, no. 5900, pp. 405–410, 2008.

[89] L. Gao and D. S. Gross, "Sir2 silences gene transcription by targeting the transition between RNA polymerase II initiation and elongation," *Molecular and Cellular Biology*, vol. 28, no. 12, pp. 3979–3994, 2008.

[90] 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.

[91] L. A. Lebedeva, E. N. Nabirochkina, M. M. Kurshakova et al., "Occupancy of the Drosophila hsp70 promoter by a subset of basal transcription factors diminishes upon transcriptional activation," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 50, pp. 18087–18092, 2005.

[92] J. M. Park, J. Werner, J. M. Kim, J. T. Lis, and Y. J. Kim, "Mediator, not holoenzyme, is directly recruited to the heat shock promoter by HSF upon heat shock," *Molecular Cell*, vol. 8, no. 1, pp. 9–19, 2001.

[93] D. A. Gilchrist, G. Dos Santos, D. C. Fargo et al., "Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation," *Cell*, vol. 143, no. 4, pp. 540–551, 2010.

SAGE-Hindawi Access to Research
Genetics Research International
Volume 2011, Article ID 652461, 10 pages
doi:10.4061/2011/652461

Review Article

Transcriptional Elongation and mRNA Export Are Coregulated Processes

Maria Micaela Molina-Navarro, Celia Pilar Martinez-Jimenez, and Susana Rodriguez-Navarro

Centro de Investigación Príncipe Felipe (CIPF), Avenida Saler, 16. 46012, Valencia, Spain

Correspondence should be addressed to Susana Rodriguez-Navarro, srodriguez@cipf.es

Received 23 June 2011; Accepted 10 August 2011

Academic Editor: Carles Sune

Chromatin structure complexity requires the interaction and coordinated work of a multiplicity of factors at different transcriptional regulation stages. Transcription control comprises a set of processes that ensures proper balance in the gene expression under different conditions, such as signals, metabolic states, or development. We could frame those steps from epigenetic marks to mRNA stability to support the holistic view of a fine-tune balance of final mRNA levels through mRNA transcription, export, stability, translation, and degradation. Transport of mRNA from the nucleus to the cytoplasm is a key process in regulated gene expression. Transcriptional elongation and mRNA export are coregulated steps that determine the mature mRNA levels in the cytoplasm. In this paper, recent insights into the coordination of these processes in eukaryotes will be summarised.

1. Introduction

Gene expression is a coordinated multistep process that begins with transcription and RNA processing in the nucleus, followed by mRNA export to the cytoplasm for translation. In eukaryotes, it depends on the orchestrated action of several multiprotein complexes which regulate the gene expression by RNA polymerase II (RNAP-II) at multiple levels.

Given the association of DNA with nucleosomes, they become a physical barrier for the transcription elongation by RNAP-II, where histone chaperones play an essential role in facilitating transcription *in vitro* and *in vivo,* and by establishing a bridge between RNAP-II and elongation factors (reviewed in [1]).

Interestingly, recent genome-wide profiling has provided a partial picture of the chromatin landscape, including a wide variety of epigenetic information about posttranslational modifications (PTMs), histone variants, DNA methylation patterns, and nucleosome occupancy (reviewed in [2]). In this scenario, histone turnover determines continuous access to sequence-specific DNA-binding proteins; therefore, nucleosome positioning is actively involved in gene-expression regulation.

Mechanisms to export mRNA are integrated into the biogenesis of messenger ribonucleoparticles (mRNPs). In this context, the nuclear pore complex (NPC) provides a molecular environment that regulates gene tethering and mRNA export. Moreover, a functional connection between transcription and RNA export has been reinforced by human genetic disorders caused by mutation in mRNA export factors and adaptors [3, 4]. mRNA biogenesis requires the transport of competent mRNPs to the cytoplasm, coupling nuclear export, and formation of mature RNA. This process is complex and requires coordination between gene transcription and RNA capping, splicing, 3′-end formation, termination, and polyadenylation (reviewed in [5]).

Emerging studies have provided strong evidence for a role of the NPC in gene expression control. For instance, gene positioning is a dynamic process which determines gene regulation [6]. Actually, from the gene gating hypothesis proposed by Blobel in 1985 arises the resulting link between mRNA transcription and the NPC. Therefore, mRNA transcription and export processes are orchestrated in time and space.

As we have seen, the complexity of chromatin structure requires the interaction and coordinated work of

a multiplicity of factors at different transcriptional regulation levels. Despite the fact that proteins involved in gene expression are highly conserved from yeast to higher eukaryotes, several processes such as mRNA biogenesis, which are coupled to nuclear export, are more divergent. This paper will pay attention to the factors implicated in gene transcriptional elongation coupled to mRNA export processes, more specifically to complexes which coordinate mRNP biogenesis and whose misregulation causes important diseases.

2. TREX Couples Transcription Elongation and mRNA Export

Due to the compartmentalisation of the eukaryotic cell, the transcribed mRNAs need to be transported from the nucleus towards the cytoplasm to be translated into proteins. After mRNAs have been packaged into mRNPs, they can be exported. One of these proteins involved in the formation of mature mRNPs is Npl3 (9G8 and STp20 in mammals), which is an mRNA-binding protein that shuttles between the nucleus and cytoplasm for rounds of mRNPs export. Npl3 is directly involved in the mRNA export process, since mutations in this gene cause nuclear accumulation of mRNA [18, 19]. Interestingly, Npl3 was one of the first pieces of evidence in yeast to show that an export factor was required cotranscriptionally for proper mRNA export [20]. Using chromatin immunoprecipitation experiments, the Silver lab demonstrated that Npl3 binds to chromatin in a transcription-dependent manner [21]. Later, work done in different laboratories has extended this cotranscriptional regulation concept of mRNA biogenesis (reviewed in [22]). A paradigm of a complex involved in transcription elongation coupled to mRNA export is the identification of the TREX complex, which contains both factors involved in transcription elongation (THO complex) and mRNA export (Yra1 and Sub2) [5, 23, 24]. These findings allowed to put forward a first model in which TREX functions in cotranscriptional mRNP formation by promoting efficient transcription elongation through RNAP-II and by guiding the newly formed mRNP to the downstream mRNA export machinery [25]. Supporting this model, it was also found that the general mRNA export receptor can be cotranscriptionally recruited to genes and, hence, contributes to gene positioning at the NPC in an RNA-independent manner [26]. The evolutionary conservation of the TREX complex from yeast to human reveals the biological importance of the association of mRNA export and transcription components in the same complex.

The TREX complex in yeast comprises Yra1, Sub2, Gbp2, Hrb1, and the THO complex (suppressor of the transcriptional defect of Hpr1 by overexpression). The THO complex integrates the core proteins Tho2, Hpr1, Mft1, and Thp2 [23]. Among them, Yra1 (Aly/REF) is required for the export of many mRNAs [27], and it plays a key role as an mRNA adaptor factor involved in splicing-coupled mRNA export in metazoans [28, 29]. The *YRA1* gene encodes an mRNA-binding protein which acts as an adaptor for the ATP-dependent RNA-helicase Sub2 (UAP56) and is involved in splicing and mRNA export [30, 31]. Besides Yra1 and Sub2, THO interacts with other nuclear mRNA export factors, such as the SR-proteins Gbp2 and Hrb1 and Tex1 (hTex1) [24, 32, 33].

The THO complex and the aforementioned RNA export factors provide evidence for a functional link between transcription and mRNA metabolism [34]. Yra1 and Sub2 are recruited by a direct transcription-coupled mechanism as part of the TREX complex to be then transferred to nascent mRNA. In contrast, Yra1 and Sub2 accumulate poorly in a subset of intron-containing genes even though Sub2 is related to splicing, probably because the cotranscriptional spliceosome assembly inhibits their recruitment in yeast [25]. During transcription elongation, Sub2 is recruited through a direct interaction with the THO component Hpr1 [26]. Recent experiments have shown that Yra1 recruitment is Sub2 independent, but it is dependent on a direct interaction with the 3′end-processing factor Pcf11 (hPcf11), which is associated with RNAP-II early on in transcription [35]**.** Furthermore, the mRNA-binding protein Yra1 recruits the essential export factor Mex67 (mammalian Tap/NXF1) [36] and its partner Mtr2 (p15/NXT1) to nascent mRNA [37, 38]. The Mex67/Mtr2 heterodimer, together with Yra1, escorts the mRNP to the NPC [39]. It has been shown that Mex67 binds to the Nup84 complex, which is crucial for nuclear mRNA export [40]. Nup84 has been recently described to have a transcriptional elongation role as part of a functional NPC linked to its function in mRNA export [41]. These findings extend the role of those factors involved in mRNA transport during transcription elongation, for instance, nucleoporins. Along with Yra1 and Sub2, another Mex67 adaptor is the poly(A) binding protein Nab2, which can be transferred from the transcription machinery onto mRNA, thus ensuring that only mature mRNPs access the NPC. Nab2 is required for poly(A) tail length control and mRNA export, linking 3′end processing and export-like Npl3 [42, 43].

Recently, novel links between transcription elongation and mRNA export have been described. For instance, the Prp19 splicing complex (homologous to human XAB2) also has a novel function in transcription elongation, since it ensures the stabilisation of the recruitment of the TREX complex at the transcribed genes [44]. This recruitment is mediated by a component of Prp19, named Syf1, which genetically interacts with THO and is involved in both splicing and transcription-coupled DNA repair. The C-terminus of Syf1 is necessary for the interaction between the Prp19 complex and RNAP-II at the transcribed genes. Furthermore, Jimeno et al. identified two suppressors of an *hpr1*Δ mutant, Thp3 and Csn12, which form a complex that is recruited to transcribed genes, establishing a further link between transcription elongation and mRNA processing [45].

It is interesting to note how histone chaperones seem to play a key role in transcription elongation by working together with the TREX complex. Nap1 is a histone chaperone recruited to specific transcribing ORFs by the TREX component Yra1, and it is important for chromatin remodelling during transcription elongation. In addition, Nap1 shows a genetic interaction with Mex67. These data

suggest a new connection between transcription elongation and mRNPs formation [46]. In human cells, a histone chaperone, Spt6, is also involved in mRNA export through its binding to the transcription elongating protein Iws1 which, in turn, triggers recruitment of REF/Yra1 to Spt6-responsive genes [47, 48]. Furthermore, UIF has been described as a novel mRNA adaptor that interacts with NXF1/Mex67 and is required for delivering mRNA to the NPC. The histone chaperone FACT, involved in transcription elongation, specifically binds UIF and is required for loading UIF onto mRNA, which ensures efficient mRNA export [49]. Accordingly, this reinforces the tight link between transcription elongation and mRNA export.

3. Role of TREX-2 and NPC-Regulating Gene Expression

Interaction of genes with nuclear pores also contributes to the coupling of mRNP biogenesis and export. In accordance with the "gene gating" hypothesis [50], the link between mRNA transcription and NPC export sites helps preferential processing and export of transcripts [51].

After nuclear quality control mechanisms (reviewed in [52]), mRNPs are tethered and exported to the cytoplasm. Transport of mRNA occurs through NPCs. Therefore, the NPC provides a molecular channel for the trafficking of export factors associated with their cargoes from the nucleus to the cytoplasm, and *vice versa* [53]. Nucleoporins bind to regulated genes and induce their transcription. This suggests a role of NPC components in regulating the gene expression programs in multicellular organisms [6, 54]. An example that depicts the importance of the NPC components in mRNA export is the interaction between Mlp1 and the nucleoporin Nup60, which are required for mRNA export to the cytoplasm [55].

An early discovery, which suggested a physical connection between the transcription machinery and the NPC, was the identification of Sus1 in yeast as a factor that is part of two complexes involved in transcription and mRNA export [56]. Sus1 was identified as an mRNA export factor in a synthetic lethal screening by using a *yra1* mutant allele [56]. It is a nuclear protein localised around the nuclear periphery through its binding to the nuclear pore-associated Sac3-Thp1-Cdc31 complex (TREX-2) involved in mRNA export [56, 57]. Sus1 is also a component of the SAGA histone acetylase complex implicated in transcription initiation. Associated with the deubiquitination module (DUBm) of the transcriptional coactivator SAGA, Sus1 is involved in chromatin modification and transcriptional activation [58].

Initially, the role of TREX-2 docking mRNPs to specific nucleoporins at the nuclear entrance of the NPC was reported. Sac3 interacts with Mex67p-Mtr2p for proper mRNA export. Furthermore, Sac3p-Thp1p TREX-2 components connect with the NPC environment through Nup1p [59]. *THP1* was originally identified as a eukaryotic-conserved gene whose null mutations conferred genetic instability, transcription defects and hyperrecombination phenotypes like those on THO mutants [60]. Recently, a new integrant of the TREX-2 complex has been found. Sem1 (Dss1 in humans) interacts physically and functionally with Thp1 and Sac3, indicating that Sem1 could be a bona-fide candidate for the TREX-2 complex [61, 62].

Most of the mutants in TREX-2 components compromise transcription elongation to different extents. Transcription elongation has been traditionally studied by genetic and biochemical approaches both *in vivo* and *in vitro*. *In vitro*, purified RNAP-II engaged directly on an oligonucleotide with a dC-tail [63], or analyses of elongation in naked DNA using whole cell extracts and plasmids with two G-less cassettes [64], have been widely used. *In vivo*, transcriptional run-on assays using a pulse of radioactive UTP and further array hybridisation [65] and RNA Pol-ChIP [66, 67] have also been extensively employed. Other assays, for instance, the GLAM ratio method, have also been successfully applied to identify factors whose mutation affects transcription elongation [68–70]. However, controversial results have been obtained depending on the assays performed, which reveal that multiple parameters might affect the specificity and yield of the factors analysed by these techniques. For instance, the role of Sus1 during transcription elongation illustrates this discordance. Several authors, including ourselves, have proposed a role for Sus1 during transcription elongation. Gonzalez-Aguilera and coworkers showed that transcription elongation is impaired in *sus1*Δ mutants *in vivo*, but only slightly *in vitro*, and they concluded that *sus1*Δ leads to similar gene expression defects as those of mutants *thp1*Δ and *sac3*Δ [71]. We demonstrated elsewhere by ChIP assays that Sus1 is recruited to coding regions during transcription elongation in association with SAGA and TREX-2 [68] and that Sus1 physically associates with export factors and with the RNAP-II. We observed that its absence elicits a decrease in total RNAP-II recruitment in pGAL-YLR454w constructs. This fits in nicely with the results obtained by the GLAM assay, which measures the efficiency of gene-length-dependent mRNA accumulation [69]. All these approaches suggest a functional role of Sus1 during transcription elongation. However, these results differ from a recent study which concluded that Sus1 does not significantly affect transcription elongation [41]. These discrepancies could be explained by antibody specificities [72, 73]. For instance, slight differences in the ChIP protocols could bring about different results due to the crosslinking effects between elongating RNAP-II and nascent mRNP [74]. It has been shown that in some studies the 8WG16 antibody specifically recognises the hypophosphorylated and, therefore, the initiating form of RNAP-II [75, 76]. Other studies, especially those in yeast and *Drosophila*, have clearly detected the elongating form of RNAP-II in chromatin immunoprecipitations [77]. However, Gilchrist et al. have demonstrated that the commonly used 8WG16 antibody has a higher affinity for initiating polymerase than for elongation-competent polymerase (hyperphosphorylated). Thus, the ChIP material derived from immunoprecipitation with 8WG16 will be inherently and substantially biased towards enrichment in the promoter-proximal RNAP-II signal, rendering this material not the most suitable for analyses of transcription elongation or RNAP-II stalling [73].

Gilchrist et al. used an antibody against the Rpb3 subunit in genome-wide analyses using ChIP-chip and ChIP-Seq to study widespread regulation of transcription elongation [73]. An antibody against Rpb3 recognises RNAP-II regardless of the phosphorylation state of the CTD of the Rpb1 subunit. Since transcription is coupled to mRNA export and the NPC plays a prominent role during this coupling, it is possible that many physiological factors account for these experimental divergences.

Once mRNP has been properly formed and assembled, it is thought that the export is facilitated by a close location to the NPC. TREX-2 mediates the location of the active genes to the NPC through a binding to both the NPC nuclear face and the SAGA complex. The crystal structure of Sus1 and Cdc31 [16], bound to a central region of Sac3, forms a conserved interaction platform that promotes NPC association and mRNA export and provides a scaffold which integrates the interaction between protein complexes and facilitates the coupling of transcription and mRNA export [17]. Deciphering the exact TREX-2 complex structure will shed light on the mechanistic coordination of the process by the complex in response to physiological or environmental changes.

4. SAGA-TREX-2 Mediates Initiation, Elongation, and Export

The SAGA complex's function in transcription activation has been largely characterised, mainly its role as a histone-modifying complex (reviewed in [78–80]). SAGA contains two enzymatic activities involved in posttranslational histone modifications: histone acetylation and deubiquitinylation [81, 82]. The SAGA chromatin-modifying complex regulates accessibility to promoter DNA in part through modification of histone amino terminal tails. Histone acetylation is one of the best-studied posttranslational modifications. SAGA contains a Gcn5-related acetyltransferase (GNAT) as a catalytic subunit. The well-established members of the complex were identified by biochemical and genetic studies first in yeast and later in *Drosophila* and humans. For simplicity reasons, the mammalian (human) GCN5-containing STAGA, TFTC, and PCAF complexes have been renamed as human SAGA by Pijnappel and Timmers in 2008 [79, 83, 84].

In fact, SAGA-promoted histone acetylation appears to enhance the processivity of RNAP-II during transcription elongation, indicating its direct contribution to this process [82, 85]. Furthermore, several pieces of evidence suggest a link between SAGA and mRNA export that is not only restricted to the presence of Sus1 in SAGA and TREX-2. Firstly, Sgf11 deletion enhances the mRNA export defects observed in *sus1*Δ cells [58]. Secondly, Ubp8 and Sgf73 deletions show defects in mRNA export. Hence, the interaction of Sac3 and Thp1 TREX-2 components with SAGA is crucial for mRNA export. Sgf73 appears to be a molecular scaffold which integrates regulation of H2B ubiquitination in *GAL1* mRNA export by tethering the gene to the NPC [86]. Thirdly, Sgf73 is necessary for the association of Sus1 not only with SAGA, but also with TREX-2, which suggests that both complexes are coordinated in their role of coupling transcription to mRNA export [68]. Finally Mlp1, a protein involved in gene anchoring at the nuclear periphery, has been described to interact with SAGA and to tether the actively transcribed *GAL1* gene to the NPC [87] by expanding the links between SAGA and mRNA export.

Several nuclear steps in mRNPs formation occur in close temporal and physical proximity to the NPC. In this scenario, SAGA and TREX-2 could coordinate the expression of a subset of genes in the vicinity of the NPC. The precise sequence of events is poorly understood, but a consensus model for the mechanism of this process is framed by the Blobel gene-gating hypothesis [50]. According to this model (Figure 1), the SAGA chromatin remodelling complex is recruited to the promoter of a subset of inducible genes and is needed for their transcription. As mentioned above, components of the TREX-2 complex interact with SAGA. The Sus1 protein, a member of the SAGA DUBm, facilitates the interaction between SAGA and TREX-2 components, which conducts a relocation of transcriptionally active genes to the nuclear periphery. Hence, the gene-gating process triggers the loading of mRNA export factors on the nascent transcript. In a subsequent step, the THO complex might interact with the new generated mRNPs throughout Sub2 and Yra1 by transferring mRNA to Mex67 and Mtr2, resulting in the formation of an export-competent mRNP. It is currently accepted that other adaptors, such as Nab2 and Npl3, are players in the process and are determinants in the association of mature mRNPs with the NPC. Moreover, the presence of Sus1 in coding regions and its physical interaction with RNAP-II imply that Mex67 and Yra1 are suggestive of a further role of Sus1 in downstream events during mRNA biogenesis [68]. Subsequently, the THO complex is released from the transcription site before transcription termination occurs [88], and Sub2 association with Sac3p establishes a physical link between THO and TREX-2 [24], whence TREX-2 is involved in mRNA export with a role on the nuclear side of the NPC.

Besides Gcn5-mediated histone acetylation, several studies have demonstrated that K123 at the C-terminus of H2B is ubiquitylated (ubH2B). ubH2B is essential for the trans-tail methylation of histone H3 and is also required for optimal gene activation [81]. Thus, histone modification marks could determine the regulatory mechanisms in transcription elongation coupled to export. Strikingly, deletions in both the *UBP8* and *GCN5* genes responsible for SAGA enzymatic activities cause synergistic transcription elongation defects, suggesting that the roles of histone acetyl transferase (HAT) and DUB modules in elongation are functionally distinct [89]. Among the epigenetic marks, ubiquitination of histone H2B, therefore, regulates chromatin dynamics by enhancing nucleosome stability. It is interesting to speculate that histone H2B deubiquitination might be of special relevance in the promoters recruited to the NPC via the SAGA-TREX-2 connection. Further work is required to determine this intriguing possibility.

All-encompassing, these models have been established from the different experimental and structural data obtained in *S. cerevisiae*. Nevertheless, several pieces of evidence open up the possibility that gene tethering to the nuclear periphery

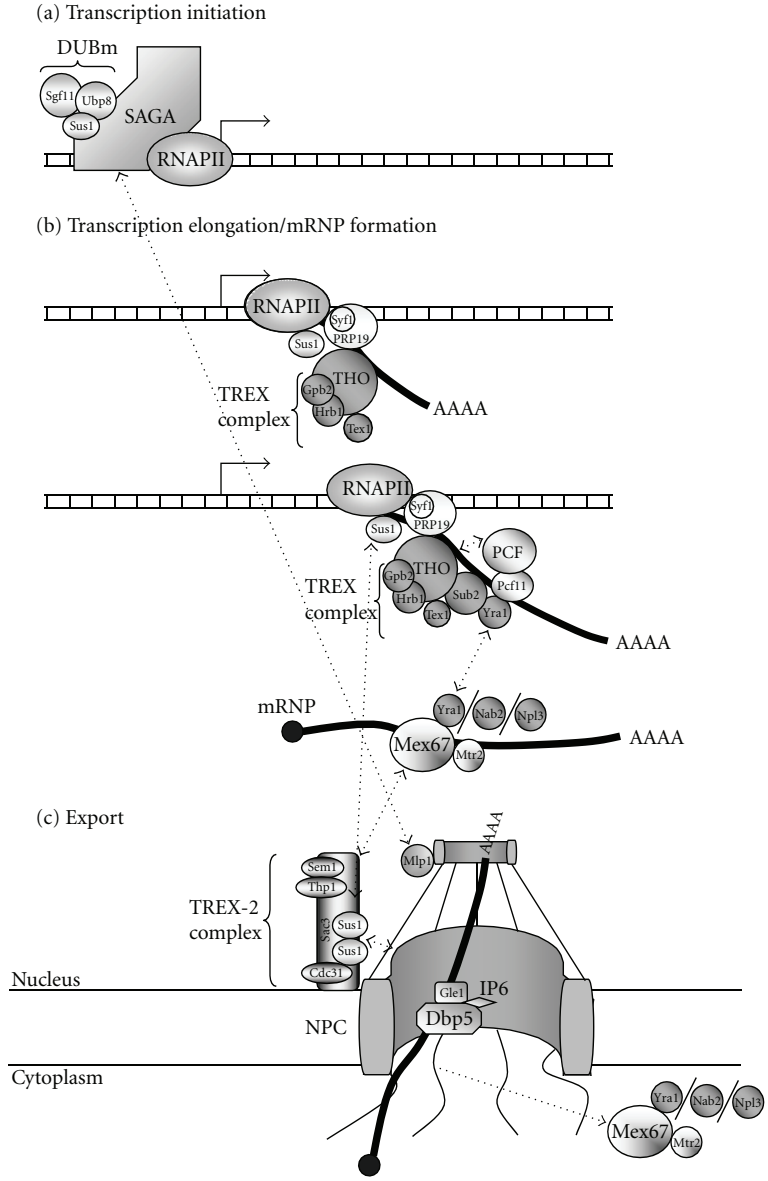


Figure 1: Coordination of different steps in transcription initiation, mRNP biogenesis, and export facilitates gene expression. (a) Active genes are recruited to the nuclear periphery through the factors involved in both transcription and mRNPs export. In the transcription activation process, there is an interaction via SAGA and Sus1 between the promoter and the NPC tethering the genes to the nuclear periphery. Mlp1 (myosin-like protein 1) is also involved in the recruitment of the *GAL1* gene to the NPC upon transcription activation. (b) The TREX complex is recruited to nascent mRNA in the early steps of transcription elongation although there are differences between yeast and metazoan when considering genome complexity. Whereas TREX is recruited cotranscriptionally in yeast, it is recruited by spliceosome components in metazoan likely due to the highest percentage of intron-containing genes [14, 15]. Adaptor proteins (Yra1, Nab2, and Npl3) recognise RNA when the transcript is competent for export to the cytoplasm, and they recruit it to export factors such as Mex67-Mtr2. This recruitment to export factors is crucial for generating mature mRNP. (c) Once the mRNP is properly formed and assembled, it is assumed that the export is facilitated by a close location to the NPC. TREX-2 (also known as THSC) mediates the location of active genes to the NPC through a binding to both the NPC nuclear face and the SAGA complex. The crystal structure of Sus1 and Cdc31 [16], bound to a central region of Sac3, forms a conserved interaction platform that promotes NPC association and mRNA export to provide a scaffold that integrates the interaction between protein complexes and facilitates the coupling of transcription and mRNA export [17]. Thp1 and Sac3 mediate the docking of mRNP at the NPC through its interaction with Mex67-Mtr2 and nucleoporins. mRNPs share the 5′-to-3′ polarity of movement through the NPC. Although the exact manner of how the transport of mRNPs takes place has not been precisely described, the most widely accepted hypothesis is that mRNPs are pulled through the NPC via ATP hydrolysis by the shuttling ATPase Dbp5 (hDbp5). Dbp5/Rat8 binds to the cytoplasmic filaments of the NPC by interacting with two nucleoporins (Nup159 and Nup 42) and Gle1. Gle1, together with its cofactor IP6, stimulates the ATPase activity of Dbp5. Once inside the cytoplasm, mRNA is released, and mRNP proteins are removed by entering a new export cycle.

Table 1: Human disorders associated with mRNA biogenesis and mRNA export defects.

Gene	Pathologies/disorders	Comments	References
GLE1	Lethal congenital contracture syndrome 1 (LCCS1)	Encodes a protein required for the export of mRNAs from the nucleus to the cytoplasm and is critical in motoneuron development and maturation.	[3]
X-linked mental retardation (XLMR)	Fragile X syndrome (FXS)	Inactivation of the X-linked FMR1 gene leads to the loss of its encoded protein FMRP and RNA export factor NXF2, causing defects in neuronal development and function as well as in male germ cells.	[4]
NUP155	Atrial fibrillation (AF)	Mutations in the gene are an inherited form of clinical arrhythmia that can lead to sudden cardiac death.	[7]
NUP98 and NUP214	Acute myelogenous leukemia (AML).	Translocations in the gene have been characterised as mutations leading to several types of leukaemia.	[8–10]
ATXN7	Neurodegenerative disease spinocerebellar ataxia type 7 (SCA7)	The expansion of an unstable CAG repeat in the first exon of the SCA7 gene causes this neurodegenerative disease	[11]
USP22	Associated with poor prognosis of diverse cancer types	Catalyses the deubiquitylation of histone H2B and is required for appropriate cell-cycle progression. Component of the 11-gene polycomb/cancer stem-cell signature	[12]
TREX84	Breast cancer	Its expression is strongly associated with an aggressive phenotype of human breast tumour	[13]

might also occur in some cases in other eukaryotes (reviewed in [90]). Interestingly, upon the inactivation of two NPC components, the transcription of the X-linked genes reduces, suggesting that the connection between the NPC and gene activation might be a conserved mechanism in eukaryotes [91].

5. mRNA Biogenesis and mRNP Export Defects Leads to Diverse Human Disorders

Nuclear factor export proteins, along with NPC components, play a critical role in the selective transport of proteins, RNA and ribonucleoproteins across the nuclear envelope. Defects in mRNA export, as well as Nups mutations, have been linked to several human diseases Table 1. Interestingly, the observed phenotypes are often manifested in specific cell types and in particular molecular pathways [3, 4, 6–8, 92]. For instance, a dynamic and stable transcription of Nups takes place in dividing *versus* terminally differentiated cells. D'Angelo et al. showed in 2009 that lack of replacement of NPC scaffold components in the somatic cells from *C. elegans* and rat brain neurons comes with an age-dependent deterioration of the NPC, leading to an aberrant nuclear accumulation of cytoplasmic tubulin [93]. In humans, there is a strong tissue-specific requirement for Nups, which are associated with specific pathologies (reviewed [6]). One example that depicts Nups requirements is atrial fibrillation (AF) caused by a mutation in the human *NUP155* which, in turn, can lead to sudden cardiac death [7]. Experiments with heterozygous $Nup155^{-/+}$ mice have shown an AF phenotype, suggesting that reduction in the level of *NUP155,* or its mistargeting, results in a tissue-specific disorder [12]. Other examples are the translocations in *NUP98* and *NUP214* genes, which have been characterised as mutations leading to several types of leukaemia [8–10].

Furthermore, the human TREX complex has been identified as a culprit of aggressive human breast cancer [13]. hTREX84, a subunit of the hTREX complex, is highly expressed in this kind of cancer, and its expression is strongly associated with an aggressive phenotype of human breast tumours [13]. Hence, Guo et al. identified not only hTREX84 as a prognosticator of breast cancer, but also the delineated human TREX complex as a target for therapeutic drugs against breast cancer [13].

Several lines of experiments in yeast have proposed Sus1 (ENY2 orthologue) to be a mediator between the NPC, Nups, and active genes [56] and also a link between transcription and mRNA export [68]. Sus1 plays a critical role in the modularity of the DUBm, which is conserved in *Drosophila* and humans. Thus, we could also expect that defects in DUBm composition would affect mRNA export. Therefore, mRNA processing defects could trigger diverse genetic disorders Table 1. Indeed, the interaction between DUBm and the SAGA complex has been shown to be mediated by TAF5L and ATXN7 [94]. Interestingly a decade ago, ATXN7 was correlated with neurodegenerative disease spinocerebellar ataxia type 7 (SCA7) [11]. ATXN7 is an integral component of SAGA [95], and several studies in yeast have demonstrated that the yeast orthologue ATXN7 (Sgf73) anchors DUBm to the SAGA complex [86]. Mutations in Sgf73 result in a release of DUBm from SAGA [56, 86]. These data open up several possibilities, for instance, (i) SAGA-independent DUBm could regulate a subset of target genes, (ii) DUBm could determine the recruitment of SAGA in particular gene promoters, (iii) or even a combination of both.

Further studies are needed to shed light on the molecular mechanism underlying the specificity and the biological role of mRNA export factors, TREX components, and NPC coupling transcriptional initiation, elongation, and mRNA export of *in vivo* targets. Likewise, understanding the time-specific interaction of these components during development will unravel the tight-coupled mechanism that regulates the cytoplasmic fate or their target mRNAs.

6. Concluding Remarks

The latest advances in proteomics have improved knowledge about the composition of protein complexes, and the newly established protein-protein networks show additional levels of plasticity to coordinate mRNA transcription, elongation, and export. However, further studies are needed to comprehend the particular contribution of each component orchestrating the gene expression in eukaryotes.

Deciphering the molecular mechanisms that coordinate the transcriptional output and mRNA level in response to cellular signals will be a future determinant to discover new therapeutic targets and new cellular pathways involved in different processes, for example, development and differentiation in response to stress conditions or even in complex diseases such cancer and human genetic disorders.

Acknowledgments

This work has been supported by grants to S. Rodriguez-Navarro from the Spanish Ministry of Science and Innovation (BFU2008-04082-C02-02) and the GV (Valencian Regional Government; ACOMP2011/057). S. Rodriguez-Navarro and M. M. Molina-Navarro have been supported by Ramon y Cajal/IP3 and Juan de la Cierva Contracts (MEC), respectively. C. P. Martinez-Jimenez is holder of an FEBS postdoctoral fellowship. Maria Micaela Molina-Navarro and Celia Pilar Martinez-Jimenez are equaly contributed in this paper.

References

[1] A. Saunders, L. J. Core, and J. T. Lis, "Breaking barriers to transcription elongation," *Nature Reviews Molecular Cell Biology*, vol. 7, no. 8, pp. 557–567, 2006.

[2] R. Margueron and D. Reinberg, "Chromatin structure and the inheritance of epigenetic information," *Nature Reviews Genetics*, vol. 11, no. 4, pp. 285–296, 2010.

[3] H. O. Nousiainen, M. Kestilä, N. Pakkasjärvi et al., "Mutations in mRNA export mediator GLE1 result in a fetal motoneuron disease," *Nature Genetics*, vol. 40, no. 2, pp. 155–157, 2008.

[4] M. Zhang, Q. Wang, and Y. Huang, "Fragile X mental retardation protein FMRP and the RNA export factor NXF2 associate with and destabilize Nxf1 mRNA in neuronal cells," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 24, pp. 10057–10062, 2007.

[5] S. Rodríguez-Navarro and E. Hurt, "Linking gene regulation to mRNA production and export," *Current Opinion in Cell Biology*, vol. 23, no. 3, pp. 302–309, 2011.

[6] M. Capelson and M. W. Hetzer, "The role of nuclear pores in gene regulation, development and disease," *EMBO Reports*, vol. 10, no. 7, pp. 697–705, 2009.

[7] X. Zhang, S. Chen, S. Yoo et al., "Mutation in nuclear pore component NUP155 leads to atrial fibrillation and early sudden cardiac death," *Cell*, vol. 135, no. 6, pp. 1017–1027, 2008.

[8] S. Saito, M. Miyaji-Yamaguchi, and K. Nagata, "Aberrant intracellular localization of set-can fusion protein, associated with a leukemia, disorganizes nuclear export," *International Journal of Cancer*, vol. 111, no. 4, pp. 501–507, 2004.

[9] T. Nakamura, D. A. Largaespada, M. P. Lee et al., "Fusion of the nucleoporin gene NUP98 to HOXA9 by the chromosome translocation t(7;11)(p15;p15) in human myeloid leukaemia," *Nature Genetics*, vol. 12, no. 2, pp. 154–158, 1996.

[10] C. Slape and P. D. Aplan, "The role of NUP98 gene fusions in hematologic malignancy," *Leukemia and Lymphoma*, vol. 45, no. 7, pp. 1341–1350, 2004.

[11] G. David, A. Dürr, G. Stevanin et al., "Molecular and clinical correlations in autosomal dominant cerebellar ataxia with progressive macular dystrophy (SCA7)," *Human Molecular Genetics*, vol. 7, no. 2, pp. 165–170, 1998.

[12] X. Y. Zhang, M. Varthi, S. M. Sykes et al., "The putative cancer stem cell marker USP22 is a subunit of the human SAGA complex required for activated transcription and cell-cycle progression," *Molecular Cell*, vol. 29, no. 1, pp. 102–111, 2008.

[13] S. Guo, M. A. Hakimi, D. Baillat et al., "Linking transcriptional elongation and messenger RNA export to metastatic breast cancers," *Cancer Research*, vol. 65, no. 8, pp. 3011–3016, 2005.

[14] N. Custódio, C. Carvalho, I. Condado, M. Antoniou, B. J. Blencowe, and M. Carmo-Fonseca, "In vivo recruitment of exon junction complex proteins to transcription sites in mammalian cell nuclei," *RNA*, vol. 10, no. 4, pp. 622–633, 2004.

[15] M. J. Wolyniak and C. N. Cole, "Harnessing genomics to explore the processes and evolution of mRNA export," *RNA Biology*, vol. 5, no. 2, pp. 68–72, 2008.

[16] D. Jani, S. Lutz, N. J. Marshall et al., "Sus1, Cdc31, and the Sac3 CID region form a conserved interaction platform that promotes nuclear pore association and mRNA export," *Molecular Cell*, vol. 33, no. 6, pp. 727–737, 2009.

[17] D. Egecioglu and J. H. Brickner, "Gene positioning and expression," *Current Opinion in Cell Biology*, vol. 23, no. 3, pp. 338–345, 2011.

[18] D. R. Singleton, S. Chen, M. Hitomi, C. Kumagai, and A. M. Tartakoff, "A yeast protein that bidirectionally affects nucleocytoplasmic transport," *Journal of Cell Science*, vol. 108, part 1, pp. 265–272, 1995.

[19] M. S. Lee, M. Henry, and P. A. Silver, "A protein that shuttles between the nucleus and the cytoplasm is an important mediator of RNA export," *Genes and Development*, vol. 10, no. 10, pp. 1233–1246, 1996.

[20] E. P. Lei, H. Krebber, and P. A. Silver, "Messenger RNAs are recruited for nuclear export during transcription," *Genes and Development*, vol. 15, no. 14, pp. 1771–1782, 2001.

[21] E. P. Lei and P. A. Silver, "Intron status and 3′-end formation control cotranscriptional export of mRNA," *Genes and Development*, vol. 16, no. 21, pp. 2761–2766, 2002.

[22] D. L. Bentley, "Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors," *Current Opinion in Cell Biology*, vol. 17, no. 3, pp. 251–256, 2005.

[23] S. Chávez, T. Beilharz, A. G. Rondón et al., "A protein complex containing Tho2, Hpr1, Mft1 and a novel protein, Thp2, connects transcription elongation with mitotic recombination

in Saccharomyces cerevisiae," *EMBO Journal*, vol. 19, no. 21, pp. 5824–5834, 2000.

[24] K. Stäßer, S. Masuda, P. Mason et al., "TREX is a conserved complex coupling transcription with messenger RNA export," *Nature*, vol. 417, no. 6886, pp. 304–308, 2002.

[25] K. C. Abruzzi, S. Lacadie, and M. Rosbash, "Biochemical analysis of TREX complex recruitment to intronless and intron-containing yeast genes," *EMBO Journal*, vol. 23, no. 13, pp. 2620–2631, 2004.

[26] D. Zenklusen, P. Vinciguerra, J. C. Wyss, and F. Stutz, "Stable mRNP formation and export require cotranscriptional recruitment of the mRNA export factors Yra1p and Sub2p by Hpr1p," *Molecular and Cellular Biology*, vol. 22, no. 23, pp. 8241–8253, 2002.

[27] H. Hieronymus and P. A. Silver, "Genome-wide analysis of RNA-protein interactions illustrates specificity of the mRNA export machinery," *Nature Genetics*, vol. 33, no. 2, pp. 155–161, 2003.

[28] H. Le Hir, D. Gatfield, E. Izaurralde, and M. J. Moore, "The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay," *EMBO Journal*, vol. 20, no. 17, pp. 4987–4997, 2001.

[29] 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.

[30] K. Sträßer 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.

[31] M. Zhang and M. R. Green, "Identification and characterization of yUAP/Sub2p, a yeast homolog of the essential human pre-mRNA splicing factor hUAP56," *Genes and Development*, vol. 15, no. 1, pp. 30–35, 2001.

[32] E. Kiesler, F. Miralles, and N. Visa, "HEL/UAP56 binds cotranscriptionally to the Balbiani ring pre-mRNA in an intron-independent manner and accompanies the BR mRNP to the nuclear pore," *Current Biology*, vol. 12, no. 10, pp. 859–862, 2002.

[33] E. Hurt, M. J. Luo, S. Röther, R. Reed, and K. Sträßer, "Cotranscriptional recruitment of the serine-arginine-rich (SR)-like proteins Gbp2 and Hrb1 to nascent mRNA via the TREX complex," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 7, pp. 1858–1862, 2004.

[34] 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.

[35] 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.

[36] 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.

[37] K. Sträßer and E. Hurt, "Yra1p, a conserved nuclear RNA-binding protein, interacts directly with Mex67p and is required for mRNA export," *EMBO Journal*, vol. 19, no. 3, pp. 410–420, 2000.

[38] F. Stutz, A. Bachi, T. Doerks et al., "REF, an evolutionarily conserved family of hnRNP-like proteins, interacts with TAP/Mex67p and participates in mRNA nuclear export," *RNA*, vol. 6, no. 4, pp. 638–650, 2000.

[39] W. Yao, D. Roser, A. Köhler, B. Bradatsch, J. Baßler, and E. Hurt, "Nuclear export of ribosomal 60S subunits by the general mRNA export receptor Mex67-Mtr2," *Molecular Cell*, vol. 26, no. 1, pp. 51–62, 2007.

[40] W. Yao, M. Lutzmann, and E. Hurt, "A versatile interaction platform on the Mex67-Mtr2 receptor creates an overlap between mRNA and ribosome export," *EMBO Journal*, vol. 27, no. 1, pp. 6–16, 2008.

[41] C. Tous, A. G. Rondón, M. García-Rubio, C. González-Aguilera, R. Luna, and A. Aguilera, "A novel assay identifies transcript elongation roles for the Nup84 complex and RNA processing factors," *EMBO Journal*, vol. 30, no. 10, pp. 1953–1964, 2011.

[42] S. R. Carmody, E. J. Tran, L. H. Apponi, A. H. Corbett, and S. R. Wente, "The mitogen-activated protein kinase Slt2 regulates nuclear retention of non-heat shock mRNAs during heat shock-induced stress," *Molecular and Cellular Biology*, vol. 30, no. 21, pp. 5168–5179, 2010.

[43] N. Iglesias, E. Tutucci, C. Gwizdek et al., "Ubiquitin-mediated mRNP dynamics and surveillance prior to budding yeast mRNA export," *Genes and Development*, vol. 24, no. 17, pp. 1927–1938, 2010.

[44] S. Chanarat, M. Seizl, and K. Sträßer, "The prp19 complex is a novel transcription elongation factor required for TREX occupancy at transcribed genes," *Genes and Development*, vol. 25, no. 11, pp. 1147–1158, 2011.

[45] S. Jimeno, C. Tous, M. L. García-Rubio et al., "New suppressors of THO mutations identify Thp3 (Ypr045c)-Csn12 as a protein complex involved in transcription elongation," *Molecular and Cellular Biology*, vol. 31, no. 4, pp. 674–685, 2011.

[46] B. C. Del Rosario and L. F. Pemberton, "Nap1 links transcription elongation, chromatin assembly, and messenger RNP complex biogenesis," *Molecular and Cellular Biology*, vol. 28, no. 7, pp. 2113–2124, 2008.

[47] 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.

[48] 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.

[49] G. M. Hautbergue, M. L. Hung, M. J. Walsh et al., "UIF, a new mRNA export adaptor that works together with REF/ALY, requires FACT for recruitment to mRNA," *Current Biology*, vol. 19, no. 22, pp. 1918–1924, 2009.

[50] G. Blobel, "Gene gating: a hypothesis," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 82, no. 24, pp. 8527–8529, 1985.

[51] G. G. Cabal, A. Genovesio, S. Rodriguez-Navarro et al., "SAGA interacting factors confine sub-diffusion of transcribed genes to the nuclear envelope," *Nature*, vol. 441, no. 7094, pp. 770–773, 2006.

[52] M. B. Fasken and A. H. Corbett, "Mechanisms of nuclear mRNA quality control," *RNA Biology*, vol. 6, no. 3, pp. 237–241, 2009.

[53] E. Kiseleva, M. W. Goldberg, J. Cronshaw, and T. D. Allen, "The nuclear pore complex: structure, function, and dynamics," *Critical Reviews in Eukaryotic Gene Expression*, vol. 10, no. 1, pp. 101–112, 2000.

[54] B. Kalverda, H. Pickersgill, V. V. Shloma, and M. Fornerod, "Nucleoporins directly stimulate expression of developmental and cell-cycle genes inside the nucleoplasm," *Cell*, vol. 140, no. 3, pp. 360–371, 2010.

[55] E. A. Powrie, D. Zenklusen, and R. H. Singer, "A nucleoporin, Nup60p, affects the nuclear and cytoplasmic localization of ASH1 mRNA in S. cerevisiae," *RNA*, vol. 17, no. 1, pp. 134–144, 2011.

[56] S. Rodríguez-Navarro, T. Fischer, M. J. Luo et al., "Sus1, a functional component of the SAGA histone acetylase complex and the nuclear pore-associated mRNA export machinery," *Cell*, vol. 116, no. 1, pp. 75–86, 2004.

[57] T. Fischer, S. Rodríguez-Navarro, G. Pereira, A. Rácz, E. Schiebel, and E. Hurt, "Yeast centrin Cdc31 is linked to the nuclear mRNA export machinery," *Nature Cell Biology*, vol. 6, no. 9, pp. 840–848, 2004.

[58] A. Köhler, P. Pascual-García, A. Llopis et al., "The mRNA export factor Sus1 is involved in Spt/Ada/Gcn5 acetyltransferase-mediated H2B deubiquitinylation through its interaction with Ubp8 and Sgf11," *Molecular Biology of the Cell*, vol. 17, no. 10, pp. 4228–4236, 2006.

[59] T. Fischer, K. Sträßer, A. Rácz et al., "The mRNA export machinery requires the novel Sac3p-Thp1p complex to dock at the nucleoplasmic entrance of the nuclear pores," *EMBO Journal*, vol. 21, no. 21, pp. 5843–5852, 2002.

[60] M. Gallardo and A. Aguilera, "A new hyperrecombination mutation identifies a novel yeast gene, THP1, connecting transcription elongation with mitotic recombination," *Genetics*, vol. 157, no. 1, pp. 79–89, 2001.

[61] G. M. Wilmes, M. Bergkessel, S. Bandyopadhyay et al., "A genetic interaction map of RNA-processing factors reveals links between Sem1/Dss1-containing complexes and mRNA export and splicing," *Molecular Cell*, vol. 32, no. 5, pp. 735–746, 2008.

[62] M. B. Faza, S. Kemmler, S. Jimeno et al., "Sem1 is a functional component of the nuclear pore complex-associated messenger RNA export machinery," *Journal of Cell Biology*, vol. 184, no. 6, pp. 833–846, 2009.

[63] T. R. Kadesch and M. J. Chamberlin, "Studies of in vitro transcription by calf thymus RNA polymerase II using a novel duplex DNA template," *Journal of Biological Chemistry*, vol. 257, no. 9, pp. 5286–5295, 1982.

[64] A. G. Rondón, M. Gallardo, M. García-Rubio, and A. Aguilera, "Molecular evidence indicating that the yeast PAF complex is required for transcription elongation," *EMBO Reports*, vol. 5, no. 1, pp. 47–53, 2004.

[65] J. R. Warner, "Labeling of RNA and phosphoproteins in Saccharomyces cerevisiae," *Methods in Enzymology*, vol. 194, pp. 423–428, 1991.

[66] P. B. Mason and K. Struhl, "Distinction and relationship between elongation rate and processivity of RNA polymerase II in vivo," *Molecular Cell*, vol. 17, no. 6, pp. 831–840, 2005.

[67] J. Sandoval, J. L. Rodríguez, G. Tur et al., "RNAPol-ChIP: a novel application of chromatin immunoprecipitation to the analysis of real-time gene transcription," *Nucleic Acids Research*, vol. 32, no. 11, p. e88, 2004.

[68] P. Pascual-García, C. K. Govind, E. Queralt et al., "Sus1 is recruited to coding regions and functions during transcription elongation in association with SAGA and TREX2," *Genes and Development*, vol. 22, no. 20, pp. 2811–2822, 2008.

[69] M. Morillo-Huesca, M. Vanti, and S. Chávez, "A simple in vivo assay for measuring the efficiency of gene length-dependent processes in yeast mRNA biogenesis," *FEBS Journal*, vol. 273, no. 4, pp. 756–769, 2006.

[70] H. Gaillard, C. Tous, J. Botet et al., "Genome-wide analysis of factors affecting transcription elongation and DNA repair: a new role for PAF and Ccr4-not in transcription-coupled repair," *PLoS Genetics*, vol. 5, no. 2, Article ID e1000364, 2009.

[71] C. González-Aguilera, C. Tous, B. Gómez-González, 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, no. 10, pp. 4310–4318, 2008.

[72] J. Zeitlinger, A. Stark, M. Kellis et al., "RNA polymerase stalling at developmental control genes in the Drosophila melanogaster embryo," *Nature Genetics*, vol. 39, no. 12, pp. 1512–1516, 2007.

[73] D. A. Gilchrist, D. C. Fargo, and K. Adelman, "Using ChIP-chip and ChIP-seq to study the regulation of gene expression: genome-wide localization studies reveal widespread regulation of transcription elongation," *Methods*, vol. 48, no. 4, pp. 398–408, 2009.

[74] D. F. Tardiff, K. C. Abruzzi, and M. Rosbash, "Protein characterization of Saccharomyces cerevisiae RNA polymerase II after in vivo cross-linking," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 50, pp. 19948–19953, 2007.

[75] T. H. Kim, L. O. Barrera, M. Zheng et al., "A high-resolution map of active promoters in the human genome," *Nature*, vol. 436, no. 7052, pp. 876–880, 2005.

[76] T. I. Lee, R. G. Jenner, L. A. Boyer et al., "Control of developmental regulators by polycomb in human embryonic stem cells," *Cell*, vol. 125, no. 2, pp. 301–313, 2006.

[77] A. K. Boehm, A. Saunders, J. Werner, and J. T. Lis, "Transcription factor and polymerase recruitment, modification, and movement on dhsp70 in vivo in the minutes following heat shock," *Molecular and Cellular Biology*, vol. 23, no. 21, pp. 7628–7637, 2003.

[78] S. P. Baker and P. A. Grant, "The SAGA continues: expanding the cellular role of a transcriptional co-activator complex," *Oncogene*, vol. 26, no. 37, pp. 5329–5340, 2007.

[79] Z. Nagy and L. Tora, "Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription factor and global histone acetylation," *Oncogene*, vol. 26, no. 37, pp. 5341–5357, 2007.

[80] E. Koutelou, C. L. Hirsch, and S. Y. R. Dent, "Multiple faces of the SAGA complex," *Current Opinion in Cell Biology*, vol. 22, no. 3, pp. 374–382, 2010.

[81] K. W. Henry, A. Wyce, W. S. Lo et al., "Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8," *Genes and Development*, vol. 17, no. 21, pp. 2648–2663, 2003.

[82] C. K. Govind, F. Zhang, H. Qiu, K. Hofmeyer, and A. G. Hinnebusch, "Gcn5 promotes acetylation, eviction, and methylation of nucleosomes in transcribed coding regions," *Molecular Cell*, vol. 25, no. 1, pp. 31–42, 2007.

[83] W. W. M. P. Pijnappel and H. T. M. Timmers, "Dubbing SAGA unveils new epigenetic crosstalk," *Molecular Cell*, vol. 29, no. 2, pp. 152–154, 2008.

[84] M. S. Torok and P. A. Grant, "Histone acetyltransferase proteins contribute to transcriptional processes at multiple levels," *Advances in Protein Chemistry*, vol. 67, pp. 181–199, 2004.

[85] A. Johnsson, M. Durand-Dubief, Y. Xue-Franzén, M. Rönnerblad, K. Ekwall, and A. Wright, "HAT-HDAC interplay modulates global histone H3K14 acetylation in gene-coding regions during stress," *EMBO Reports*, vol. 10, no. 9, pp. 1009–1014, 2009.

[86] A. Köhler, M. Schneider, G. G. Cabal, U. Nehrbass, and E. Hurt, "Yeast Ataxin-7 links histone deubiquitination with gene gating and mRNA export," *Nature Cell Biology*, vol. 10, no. 6, pp. 707–715, 2008.

[87] R. Luthra, S. C. Kerr, M. T. Harreman et al., "Actively transcribed GAL genes can be physically linked to the nuclear pore by the SAGA chromatin modifying complex," *Journal of Biological Chemistry*, vol. 282, no. 5, pp. 3042–3049, 2007.

[88] M. Kim, S. H. Ahn, N. J. Krogan, J. F. Greenblatt, and S. Buratowski, "Transitions in RNA polymerase II elongation complexes at the 3′ ends of genes," *EMBO Journal*, vol. 23, no. 2, pp. 354–364, 2004.

[89] A. Wyce, T. Xiao, K. A. Whelan et al., "H2B ubiquitylation acts as a barrier to Ctk1 nucleosomal recruitment prior to removal by Ubp8 within a SAGA-related complex," *Molecular Cell*, vol. 27, no. 2, pp. 275–288, 2007.

[90] R. Luna, H. Gaillard, C. González-Aguilera, and A. Aguilera, "Biogenesis of mRNPs: integrating different processes in the eukaryotic nucleus," *Chromosoma*, vol. 117, no. 4, pp. 319–331, 2008.

[91] S. Mendjan, M. Taipale, J. Kind et al., "Nuclear pore components are involved in the transcriptional regulation of dosage compensation in Drosophila," *Molecular Cell*, vol. 21, no. 6, pp. 811–823, 2006.

[92] J. Wesierska-Gadek, A. Klima, O. Komina, C. Ranftler, P. Invernizzi, and E. Penner, "Characterization of autoantibodies against components of the nuclear pore complexes: high frequency of anti-p62 nucleoporin antibodies," *Annals of the New York Academy of Sciences*, vol. 1109, pp. 519–530, 2007.

[93] M. A. D'Angelo, M. Raices, S. H. Panowski, and M. W. Hetzer, "Age-dependent deterioration of nuclear pore complexes causes a loss of nuclear integrity in postmitotic cells," *Cell*, vol. 136, no. 2, pp. 284–295, 2009.

[94] Y. Zhao, G. Lang, S. Ito et al., "A TFTC/STAGA module mediates histone H2A and H2B deubiquitination, coactivates nuclear receptors, and counteracts heterochromatin silencing," *Molecular Cell*, vol. 29, no. 1, pp. 92–101, 2008.

[95] D. Helmlinger, S. Hardy, S. Sasorith et al., "Ataxin-7 is a subunit of GCN5 histone acetyltransferase-containing complexes," *Human Molecular Genetics*, vol. 13, no. 12, pp. 1257–1265, 2004.

SAGE-Hindawi Access to Research
Genetics Research International
Volume 2011, Article ID 970968, 10 pages
doi:10.4061/2011/970968

Review Article

Control of V(D)J Recombination through Transcriptional Elongation and Changes in Locus Chromatin Structure and Nuclear Organization

Beatriz del Blanco, Vanina García, Alberto García-Mariscal, and Cristina Hernández-Munain

Instituto de Parasitología y Biomedicina López-Neyra (IPBLN-CSIC), Consejo Superior de Investigaciones Cientťficas, Parque Tecnológico de Ciencias de la Salud, Avenida del Conocimiento s/n. 18100 Armilla, Spain

Correspondence should be addressed to Cristina Hernández-Munain, chmunain@ipb.csic.es

Received 7 July 2011; Accepted 29 July 2011

Academic Editor: Carles Sune

V(D)J recombination is the assembly of gene segments at the antigen receptor loci to generate antigen receptor diversity in T and B lymphocytes. This process is regulated, according to defined developmental programs, by the action of a single specific recombinase complex formed by the recombination antigen gene (RAG-1/2) proteins that are expressed in immature lymphocytes. V(D)J recombination is strictly controlled by RAG-1/2 accessibility to specific recombination signal sequences in chromatin at several levels: cellular lineage, temporal regulation, gene segment order, and allelic exclusion. DNA cleavage by RAG-1/2 is regulated by the chromatin structure, transcriptional elongation, and three-dimensional architecture and position of the antigen receptor loci in the nucleus. Cis-elements specifically direct transcription and V(D)J recombination at these loci through interactions with transacting factors that form molecular machines that mediate a sequence of structural events. These events open chromatin to activate transcriptional elongation and to permit the access of RAG-1/2 to their recombination signal sequences to drive the juxtaposition of the V, D, and J segments and the recombination reaction itself. This chapter summarizes the advances in this area and the important role of the structure and position of antigen receptor loci within the nucleus to control this process.

1. Introduction

The immune system is considered one of the best models to study the molecular mechanisms of epigenetic control of cellular differentiation *in vivo*. Development of B and T lymphocytes occurs through a series of well-defined differentiation stages initiating from a common hematopoietic stem cell. Each of these differentiation stages involves the activation and repression of antigen receptor loci that influence cellular identity. Expression of these loci in immature lymphocytes requires the activation and silencing of genomic DNA rearrangements that are regulated through the accessibility of chromatin, transcriptional elongation, three-dimensional structure, and nuclear positioning during cell development [1, 2]. In fact, lymphocytes are the only vertebrate cells that use genomic DNA rearrangements as an integral component of their developmental program. This DNA rearrangement process is known as V(D)J recombination and consists of the assembly of the genomically dispersed gene segments V (variable), D (diversity), and J (joining) to generate the functional variable region of antigen receptors: immunoglobulins (Igs) in B lymphocytes and T-cell receptors (TCRs) in T lymphocytes (Figure 1). This process results in the expression of a unique antigen receptor in each developed lymphocyte and is therefore responsible for the generation of antigen receptor diversity in T and B lymphocytes that defines the vertebrate adaptive immune responses.

V(D)J recombination is initiated through the action of the protein products of the recombination activating gene (RAG) 1 and 2; together, RAG-1 and RAG-2 form a specific endonuclease in immature lymphoid cells. The RAG-1/2

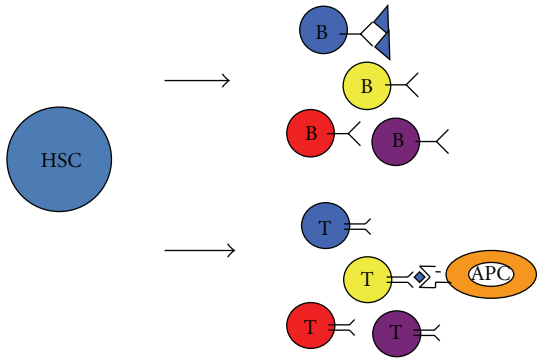


Figure 1: Lymphocyte maturation and expression of clonotypic antigen receptors. Scheme represents lymphocyte maturation from hematopoietic stem cells (HSCs). Each lymphocyte expresses a specific antigen receptor during cell development. B lymphocyte antigen receptors and immunoglobulins (Ig) recognize soluble antigens, whereas T lymphocyte antigen receptors (TCRs) recognize antigenic peptides presented by antigen presenting cells (APCs).

complex is responsible for the double-strand DNA cleavage between the segments that allows recombination through the recognition of specific recombination signal sequences (RSSs) that flank them (Figure 2); these RSSs consist of conserved heptamer and nonamer elements that are separated by a less conserved spacer region of 12 or 23 base pairs. The 12–23 rule limits recombination to segments between a 12-bp RSS and a 23-bp RSS (here, we will refer to RSSs with 12-bp and 23-bp spacers as compatible RSSs) [3]. Proteins responsible for the joining of the nonhomologous ends then process the double-strand broken ends to generate the coding DNA and the extrachromosomal circles containing the RSSs and the deleted internal gene regions.

2. V(D)J Recombination Control by Chromatin Structure

The 5′ region of the *Ig* and *Tcr* loci (which encode the variable regions responsible for antigen recognition) contains the V, D (only in some loci), and J gene segments (Figure 3) that are assembled through the action of RAG-1/2 proteins in lymphocyte precursors. The restricted expression of RAG-1 and RAG-2 in immature lymphocytes explains the specificity of the V(D)J recombination process in these cells. However, antigen receptor loci (*Igh, Igk, Igl, Tcra, Tcrb, Tcrg, or Tcrd*) and lineage (B or T) specific regulation is defined by controlled RAG-1/2 accessibility to the specific locus chromatin in B and T lymphocytes precursors [6]. This control establishes that the immunoglobulin loci (*Igh, Igk,* and *Igl*) only rearrange in B-lymphocyte precursors, and the T-cell receptor loci (*Tcra, Tcrb, Tcrd,* and *Tcrg*) rearrange exclusively in T lymphocyte precursors. Additionally, there is temporal control of V(D)J recombination that ensures that this process occurs in a developmental stage-specific manner during lymphocyte development. During T-lymphocyte development, *Tcrb, Tcrd,* and *Tcrg* loci rearrange earlier than *Tcra* locus. Similarly, during B lymphocyte development, *Igh* rearranges earlier than the *Igk* and *Igl* loci. Furthermore, there is an additional developmental control imposed on sets of gene segments within each antigen receptor locus. For example, D-to-J rearrangements precede V-to-DJ rearrangements at the *Tcrb* and *Igh* loci. This locus-, lineage-, temporal-, and gene segment order-specific regulation of V(D)J recombination is mediated through the control of RSS accessibility to the RAG-1/2 proteins. Hence, the chromatin imposes a barrier to RAG-1/2 accessibility that is controlled through strict epigenetic control, which is dependent on the specific antigen receptor locus, gene segment, cellular lineage, and developmental stage. This is the basis for the accessibility model proposed 25 years ago by Yancopoulos and Alt [7]. These investigators observed that the developmental activation of V_H gene segment recombination at the *Igh* locus coincided with V_H germline transcription (the process of transcription of sterile transcripts at an un-rearranged locus originating from V-associated promoters) during B lymphocyte development [7]. Based on these results, they proposed that the transcription of the V_H gene segments reflects an increase in the accessibility of the V_H gene segments to both the transcriptional and recombinational machineries (RNA polymerase II (RNAPII) and RAG-1/2 proteins, resp.). Since then, germline transcripts initiating at V, D, and J gene segments have been found to developmentally coincide with the activation of V(D)J recombination at each antigen receptor locus [7–9]. In addition to reports of sense transcription, developmentally regulated antisense intergenic transcription across the V_H gene segments that correlates with V_H to DJ_H recombination has also been reported [10]. In agreement with this model, it has been proven that the barrier that the chromatin imposes on RAG-1/2 accessibility is eliminated through the activation of cis-transcriptional elements present at these loci during lymphocyte development [1]. Each *Ig* or *Tcr* locus is equipped with at least one transcriptional enhancer in the vicinity of the constant region and numerous promoters associated with V, D, and J gene segments (Figure 3). The essential role of each of these cis-elements in controlling the accessibility to the RAG-1/2 proteins was demonstrated in numerous studies using transgenic mini-loci as recombination reporters and directed mutagenesis at the endogenous loci [1]. These studies clearly established that the enhancers are the elements that are responsible for specific lineage determination and temporal control of V(D)J recombination through the general regulation of locus chromatin structure; thus, enhancers control the accessibility of the RAG-1/2 proteins to multiple gene segments separated by large distances, whereas promoters are the elements that mediate the accessibility of the RAG-1/2 proteins to regions located at the proximal regions of the specific gene segments [1]. The accessibility model was reinforced by observations demonstrating a direct correlation between V(D)J recombination and activating epigenetic modifications such as histone H3 and H4 acetylation (H3ac and H4ac), methylation of lysine 4 of histone H3 (H3K4me), nuclease accessibility and DNA hypomethylation [1, 11–14], and changes in nucleosomal structure [15]. Furthermore, establishment of inactive chromatin suppresses V(D)J recombination [16].

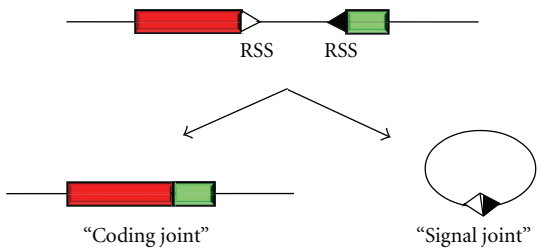


Figure 2: V(D)J recombination process. Gene segments are represented as red and green rectangles, and a pair of compatible RSSs is represented as white and black triangles.

Additional *in vitro* studies have demonstrated that assembly of RSSs into nucleosomes inhibits V(D)J recombination [17–20], supporting the notion that nucleosomes impede RAG1/2 binding or function. The barrier for V(D)J recombination imposed by nucleosomes can be surmounted by ATP-dependent chromatin remodeling complexes, such as SWI/SNF [17, 18, 21–23]. Recently, it has been directly demonstrated that chromatin accessibility to RAG-1/2 is indeed mediated by enhancers and promoters [24]. In this study, it was proven that the enhancers control global RAG-1 binding, whereas promoters direct local RAG-1 binding at the antigen receptor loci. RAG-1 binding to accessible RSSs can be targeted in the absence of RAG-2, which is recruited directly to trimethylated H3K4 (H3K4me3), a mark of open and active chromatin [25–27]. Hence, both the enhancer and promoter elements are necessary to confer RAG-1/2 accessibility to specific RSSs within a given locus and to facilitate the recombination synapse between the RSSs (Figure 4). To date, the precise molecular mechanisms by which distal enhancers control the transcriptional activation from promoters separated by large distances and V(D)J recombination are not known. Chromatin immunoprecipitation and chromosomal conformation capture experiments have demonstrated that V(D)J recombination involves physical interactions between distal transcriptional regulatory elements such as promoters and enhancers to permit RAG-1/2 accessibility to their target RSSs [28–30].

It is interesting that the antigen receptor loci undergo a process of gene contraction (juxtaposition of V and D-J regions), which is strictly regulated during development. As demonstrated by three-dimensional fluorescence *in situ* hybridization (3D-FISH) experiments using distal DNA probes, this contraction correlates with transcription and V(D)J recombination [29, 31–33]. In fact, gene contraction occurs at the same moment that a particular locus is transcribed and is ready to recombine even in the absence of recombinase activity in $Rag^{-/-}$ mice. These contractions could be mediated by interactions between regulatory regions and/or by specific nuclear structures. Enhancer-promoter interactions are thought to direct the long-distance communications (Figure 4). Comparative analysis of *Tcra/Tcrd* locus contraction in wild-type, *Tcra* locus enhancer $(E\alpha)^{-/-}$ and *Tcrd* locus enhancer $(E\delta)^{-/-}$ thymocytes revealed no significant differences between the cell types [33]. Hence, enhancer-promoter interaction is not necessarily sufficient to mediate *Tcra* locus contraction and RSSs synapse, but locus contraction might facilitate both enhancer-promoter interactions and RSSs synapse that are required for transcription and V(D)J recombination. These data suggest that a preexisting conformation of the locus that is mediated by an enhancer-independent mechanism may promote the enhancer-promoter interactions necessary to activate transcription and V(D)J recombination of the distant gene segments [32, 33]. Hence, enhancer-promoter interactions could establish molecular bridges over long distances but only after they have been brought into proximity by locus contraction mediated by chromatin-organizing proteins. The mechanism for locus contraction itself is not known. Deficiencies in specific transcription factors such as Pax5, Ikaros, and YY1 disrupt the contracted *Igh* locus configuration, but it is not known how they do this [34–38]. Additionally, locus contraction could be mediated by chromatin-organizing proteins such as SATB1, CCCTC-binding factor (CTCF), and CTCF-associated cohesin that have been shown to promote long-distance looping interactions at other loci [39–44]. In fact, CTCF and cohesin have been shown to colocalize at multiple sites within the *Igh* and *Igk* loci in immature B lymphocytes [45, 46]. Consistent with this, very recent experiments have functionally demonstrated that CTCF and cohesin influence the genomic structure of the *Igh* locus in developing B lymphocytes [47]. The precise molecular mechanisms involved in how locus-specific conformational changes can regulate the enhancer-promoter interactions to subsequently direct the different V(D)J programs during lymphocyte development is an issue of intense research in the field.

3. V(D)J Recombination Control by Transcriptional Elongation

The finding that germline transcription at a given antigen receptor locus occurs concomitantly with its recombination [7] suggests a linkage between transcription and V(D)J recombination. Based on this evidence, it is accepted that both transcription and V(D)J recombination are consequences of locus accessibility that is mediated through the activation of promoters and enhancers during lymphocyte development [1]. These regulatory elements serve as docking elements to recruit transcription factors that initiate and help to propagate changes in chromatin structure that are essential for the accessibility of the RNAPII and RAG-1/2 proteins (Figure 4). Consistent with this, several transcription factors have also been shown to coordinately regulate both transcription and V(D)J recombination. For example, overexpression of E2A in nonlymphoid cells that express RAG-1 and RAG-2 proteins induced germline transcription and V(D)J recombination at the *Igk*, *Tcrg,* and *Tcrd* loci [48, 49]; $OcaB^{-/-}$ mice displayed defective transcription and recombination of a subset of Vκ gene segments [50]; Stat5 is required for the transcription and VγJγ recombination at the *Tcrg* locus in response to IL-7 [51, 52]; and deletion of the enhancers and promoters at *Tcr* and *Ig* loci or inclusion of mutations at motifs for required transcription factors

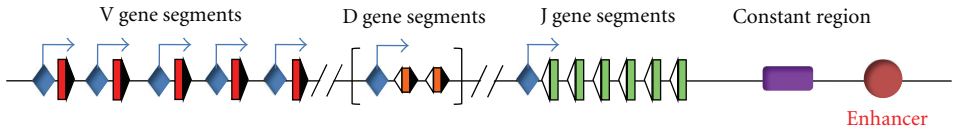


Figure 3: Representation of an antigen receptor locus. V, D, and J gene segments are represented as red, orange, and green rectangles, respectively. RSSs are represented as white and black triangles. The constant region is represented as a purple rectangle. Promoters are represented as blue diamonds, and arrows indicate transcription. The enhancer is represented as a red circle.

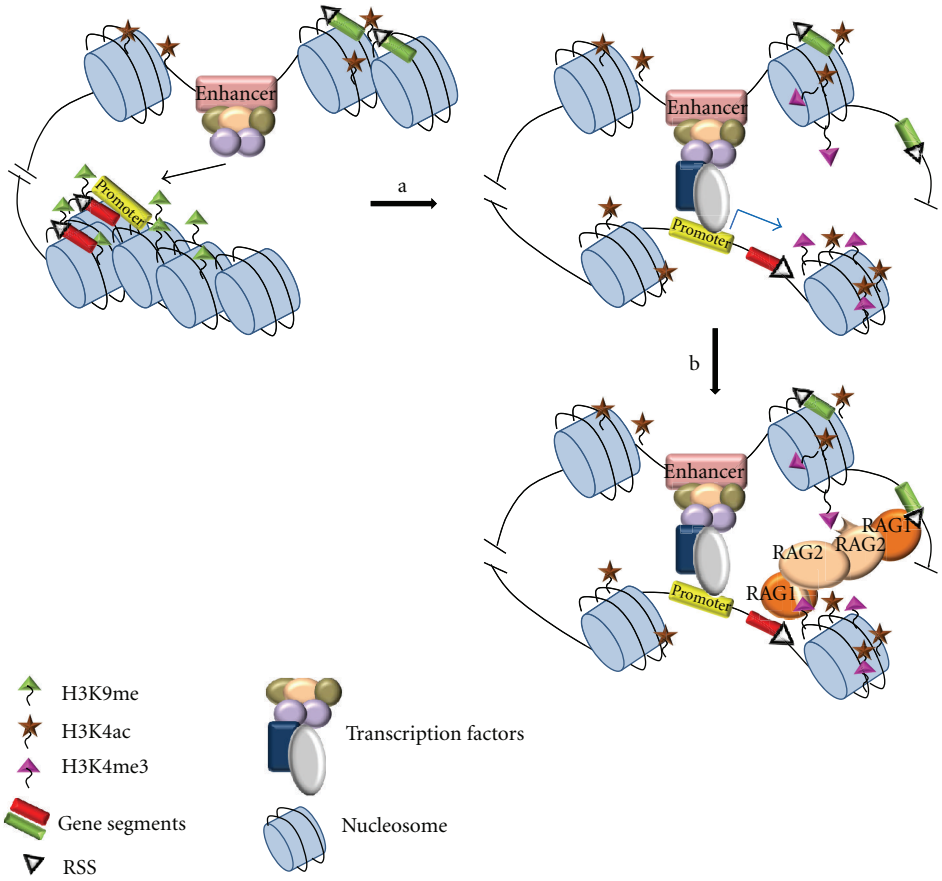


Figure 4: Representation of the molecular mechanism for activation of V(D)J recombination by enhancers, promoters, and transcription. Shown is a depiction of the physical interaction between the enhancer and the promoter within an antigen receptor locus. This interaction is mediated through protein-protein interactions among transcription factors and triggers the processes of transcription and V(D)J recombination that is derived from chromatin opening and subsequent accessibility of the RSSs to the RNAPII and RAG-1/2 proteins. Gene segments are represented as red and green rectangles and RSSs as white triangles. (a) Enhancer activation by the assembly of a functional multiprotein complex on the enhancer mediates the recruitment of the RNAPII to the promoter. This activates germline transcription (blue arrow) and opens the chromatin structure by repositioning nucleosomes, evicting nucleosomes, and/or changing the covalent modifications of histones (e.g., changing H3K9me, which is indicative of repressive chromatin, to H3K4me and H3/H4ac, which are indicative of activated chromatin). (b) The new chromatin configuration allows recruitment of the RAG-2 protein through H3K4me3 and recruitment of the RAG-1 protein to accessible RSSs. Recruitment of a RAG-1/2 complex to two compatible RSSs allows initiation of V(D)J recombination.

within these cis-elements inhibit both transcription and V(D)J recombination at each locus [1].

For many years, it was not clear whether germline transcription was merely a side effect of chromatin accessibility generated by the activation of enhancers and promoters or whether it was causal in the V(D)J recombination process itself. In an elegant and definitive study, Abarrategui and Krangel have proven that germline transcription is a key developmental regulator of accessibility for Vα-to-Jα recombination at the *Tcra* locus [4]. This locus spans 1.5 Mb and contains around 100 Vα/δ gene segments in a 1 Mb region at the 5′ end of the locus and 61 Jα gene segments in a 65 kb region at the 3′ end of the locus [53] (Figure 5). Vα-to-Jα recombination events depend on both Eα, located at the 3′ end of the locus [54], and the promoters associated with the Jα gene segments [55]. These

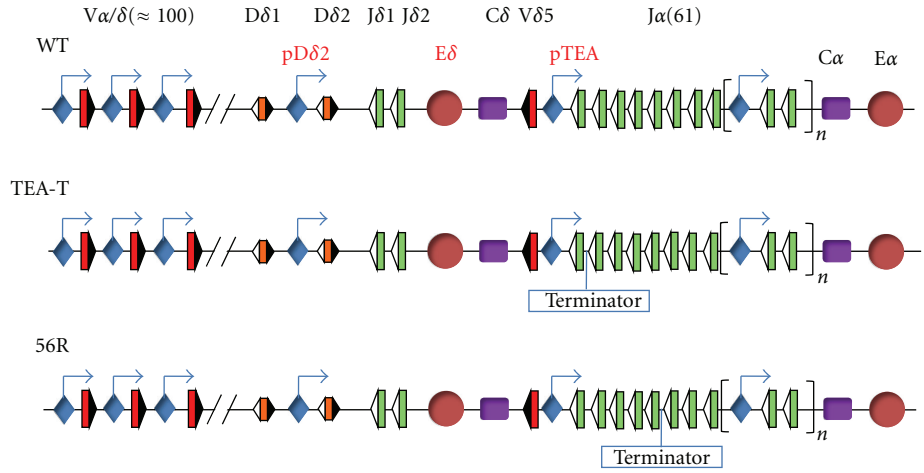


Figure 5: Representation of WT *Tcra* locus and mutant versions containing a terminator sequence [4, 5]. V, D, and J gene segments are represented as red, orange, and green rectangles, respectively. RSSs are represented as white and black triangles. Constant regions are represented as purple rectangles. Promoters, including the TEA promoter, are represented as blue diamonds, and arrows indicate transcription. Eδ and Eα are represented as red circles. The position of the terminator sequence in mutant TEA-T and 56R *Tcra* is indicated.

investigators introduced a strong transcription terminator downstream of either the TEA promoter (TEA-T) or the Jα56 gene segment (56R) in the endogenous mouse locus to block transcription originating at the upstream TEA promoters [4, 5] (Figure 5). The terminator sequence they used consists of four polyadenylation sites followed by an array of twelve bacterial lac operons that are thought to function as strong pause sites for RNAPII. The introduced terminator was able to impose an effective block to RNAPII passage. Interestingly, the transcriptional block in both TEA-T and 56R mice caused a strong reduction in recombination at the Jα gene segments located immediately downstream of the terminator sequence. Thus, these experiments clearly demonstrated that transcriptional elongation by RNAPII is necessary for creating accessible chromatin for the RAG-1/2 proteins to initiate *Tcra* recombination.

Although instances of recombination in the apparent absence of transcription have also been reported, credible explanations for each of them can be found. For example, isolated nuclei of lymphocytes from *Rag*$^{-/-}$ mice can rearrange their receptor antigen loci *in vitro* by addition of RAG-1 and RAG-2 proteins in the absence of ongoing transcription [6]; however, the chromatin structure of the receptor antigen loci could remain accessible during nuclei isolation. Additionally, it has been demonstrated that the V_H gene segments rearrange in pro-B lymphocytes with no detectable transcription of these segments [56]; however, the state of transcription of these segments at the time of recombination could not be analyzed. Furthermore, the requirement for transcription itself might not be necessary for RSSs that are located near a cis-element; this is the case for RSSs that are positioned adjacent to a promoter. Several examples of this phenomenon include the inducible mouse mammary tumor virus long terminal repeat that can confer accessibility to a tightly associated RSS [57], the endogenous *Tcrb* locus that has a *Tcrb* enhancer-(Eβ-) dependent promoter tightly associated with the Dβ1 segment to confer accessibility to this segment [58, 59], and the physical interaction of the Dβ1 promoter with Eβ to deliver SWI/SNF chromatin remodeling complexes that results in a decrease of nucleosome occupancy at the Dβ1 gene segment [28, 30, 60]. These results argue that transcription is not required to confer accessibility to RAG-1/2 when the RSS is tightly associated with a promoter. However, this is not the situation for many RSSs within the antigen receptor loci *in vivo* because they are positioned far away from the promoters.

More recent experiments with the TEA-T and 56R mouse models have definitively demonstrated that chromatin accessibility to the RAG-1 protein is mediated by transcriptional elongation itself [24]. In this study, it was proven that the transcriptional terminator introduced within the Jα cluster in TEA-T and 56R mice (Figure 5) inhibited the recruitment of the RAG-1 protein to downstream chromatin. Hence, transcriptional elongation itself confers accessibility to the RAG-1/2 proteins to specific RSSs. Although the pattern of RAG-2 binding was not assessed in this study, it is expected that it would closely resemble that of RAG-1 binding because the pattern of H3K4me3 (which accurately predicts RAG-2 binding) was similar to that of RAG-1 [4, 5, 24, 27]. Hence, H3K4me3 recruits the RAG-2 protein [25–27] and directly contributes to the opening of the chromatin by either repositioning or evicting nucleosomes, which allows free access of the RAG-1 protein to RSSs. Although it has not been directly demonstrated, it is expected that this is likely to be the case for the RSSs that are distantly located from the promoters in the other *Ig* and *Tcr* loci. However, a direct role for transcripts in directing V(D)J recombination, as has been shown for class switching at the *Igh* locus, is not expected because the RAG-1/2 substrate is double strand DNA.

Thus, Abarrategui and Krangel's studies [4, 5] have demonstrated that V(D)J recombination at the *Tcra* locus

requires that the elongation machinery travels or has traveled through the RSS's DNA to allow the RAG-1/2 proteins to access to the RSS's chromatin. However, the precise molecular mechanism by which transcriptional elongation directs the process of VαJα recombination is not known. Chromatin structure imposes significant obstacles on the passage of the RNAPII through the DNA. Because the elongation of transcription is associated with the transient disruption of nucleosome structure [61] and because of the transcription-dependent reduction of nucleosome density at the coding region [62], it is possible that RAG-1 is recruited to the RSSs due to either a transient disruption of the chromatin structure or to less compacted chromatin derived from RNAPII transit. Consistent with the role of transcriptional elongation in transient disrupting of nucleosomal structure, it has been shown using the TEA-T mouse model (Figure 5) that germline transcription originating from the TEA promoter at the *Tcra* locus causes covalent histone modifications related to opening of chromatin [5], as well as repositioning and loss of the nucleosomes at a 600 base-pair region including the TEA promoter itself and its closest 3′ Jα gene segment, Jα61 [15].

Gene transcription is initiated by the binding of transcription factors to promoters and enhancers (Figure 4). The binding of transcription factors to the enhancer recruits histone acetyltransferases (HATs) that acetylate the N-terminal tails of histones H3 and H4. These acetylated histones provide binding sites for the bromodomains present in other chromatin remodeling complexes and histone-modifying enzymes. Binding of these complexes and enzymes results in nucleosome displacement or disassembly and thus frees promoters for binding by RNAPII thereby allowing transcriptional initiation [63–65]. The RNAPII complexes assembled on the promoter subsequently transit through the chromatin to mediate transcriptional elongation. Clearance of the RNAPII from the promoter requires the phosphorylation of its carboxy-terminal domain (CTD) [66], which is a molecular platform that can recruit a variety of histone modifier complexes, chromatin remodeling complexes, histone chaperones, and elongation factors that are associated and travel with the elongating form of RNAPII; these cofactors are required for efficient transcription through the chromatin [65, 67–72]. Among the most important chromatin modifications associated with transcriptional elongation are the methylation of lysines 4, 36, and 79 of histone H3 and the monoubiquitination of histone H2A.

Abarrategui and Krangel observed that H3ac, H3K4me3, H3K4me2, and H3K36me3 were significantly reduced at the 3′ end of the terminator introduced immediately downstream of the TEA promoter at the *Tcra* locus [5] (Figure 5); these results suggest that the histone methyl transferases and HATs responsible for these histone covalent modifications might be involved in activating VαJα recombination through RAG-1/2 recruitment mediated by transcriptional elongation [70, 73]. In addition, H3K4me2/3 that recruits RAG-2 [25–27] also recruits other chromatin remodeling complexes such as the ATPase RSC, ISWI, or SWI/SNF that can reposition or evict nucleosomes to facilitate the passage of the RNAPII that can be involved in recruitment of RAG-1 [23, 74–76]. Currently, the identity of the specific chromatin modifying activities that are involved in facilitating the recruitment of RAG-1/2 proteins to RSSs is unknown. Additionally, chaperones that travel with the elongating RNAPII and the chromatin remodeling complexes that facilitate RNAPII's transit could also favor RAG-1 binding [61, 77]. It is also possible that RAG-1 recruitment to the RSSs might be facilitated by direct interaction with the elongating RNAPII through the RNAPII CTD. This would require a total coupling between transcription and V(D)J recombination; the coupling of transcription and RNA splicing is already accepted to occur at the RNAPII CTD [78]. Understanding how precisely transcription activates V(D)J recombination is an important goal for future research in this field.

4. Control of V(D)J Recombination by the Nuclear Position of the Antigen Receptor Loci

In spite of the above cited data, there is evidence that shows that transcription and RSS accessibility are not necessarily sufficient to activate V(D)J recombination *in vivo*. For example, in pre-T lymphocytes, a deletion within the *Tcrb* locus that placed the Vβ gene segments under the influence of the Eβ promoted high levels of Vβ transcription but not Vβ to DJβ rearrangement [79]. Additionally, in pre-T lymphocytes, ectopic introduction of Eα within the cluster of Vβ gene segments enhanced transcription of these segments but did not induce Vβ-to-DJβ rearrangement [80]. Furthermore, germline Vβ transcripts are detected from both alleles of the *Tcrb* locus even in the presence of allelic exclusion during Vβ-to DβJβ recombination [81, 82]. Finally, germline transcription similarly occurs at the *Igk* locus, which undergoes biallelic germline transcription in pre-B lymphocytes during allelic exclusion of Vκ-to-Jκ recombination [83–86]. These findings indicate that additional regulatory constraints on V(D)J recombination exist that operate beyond transcription and chromatin accessibility.

Over the last ten years, it has become clear that the position of *Ig* and *Tcr* loci in the nucleus has an essential role in directing V(D)J recombination between distantly located gene segments. It is now accepted that, in addition to transcriptional competence, a particular locus or allele must move away from repressive chromatin to allow distant RSSs to form a recombinational synapse through RAG-1/2 binding that initiates V(D)J recombination [87]. It is now evident that the *Ig* and *Tcr* loci move away from repressive compartments such as the nuclear periphery or pericentric heterochromatin when they undergo recombination; the loci then reassociate with them following recombination [85, 87]. Furthermore, associations of *Igh*, *Igk,* and *Tcrb* loci with repressive nuclear compartments seem to be responsible for the establishment of allelic exclusion [81, 86, 88]. Hence, there is a clear connection between the association with repressive nuclear compartments and the inhibition of V(D)J recombination, but germline transcription does not always correlate with this phenomenon.

The mechanism by which locus association with repressive compartments inhibits V(D)J recombination without

inhibiting transcription remains unknown. RAG-1/2 binding to D and J gene segments is robust at the *Igh* and *Tcrb* loci in pro-B/pre-B and pro-T/pre-T lymphocytes, respectively, even when V-to-DJ recombination is inhibited and one or both alleles are associated with repressive compartments in both stages [24, 27, 31, 81]. It is known that the *Igh* locus is tethered to the nuclear membrane through the distal V_H region cluster, whereas the D_HJ_H region is located away [85]. Thus, persistent RAG-1/2 binding to *Igh* and *Tcrb* D and J gene segments in pro-B/T and pre-B/T lymphocytes might be consequence of RSS accessibility due to the spatial orientation of these loci within the nucleus [27, 85]. It is known that V_H and Vβ gene segment transcription and accessibility are reduced in the transition from pro-B/T to pre-B/T lymphocytes, respectively, and hence both parameters do seem to correlate with allelic exclusion of *Igh* and *Tcrb* loci at pre-B/T lymphocytes [89–91]. These results support the current model that feedback inhibition of *Igh* and *Tcrb* loci in pre-B/T lymphocytes, but not in pro-B/T lymphocytes, operates primarily on V_H and Vβ gene segment accessibility, respectively. The different chromatin structure at the V_H and Vβ gene segments in the two stages indicates clear differences about what might be happening in pro-B/T versus pre-B/T lymphocytes [27, 81, 89–91]. Identification of cis-elements and transfactors that are involved in controlling the association of specific antigen receptor alleles with repressive nuclear compartments is required to elucidate the function of nuclear positioning in regulating V(D)J recombination and allelic exclusion. In the case of the *Igk* locus, which is also relocated during B lymphocyte development [85], a cis-element that binds the Ikaros transcriptional repressor targets *Igk* transgenes to centromeric chromatin and inhibits Vκ-to-Jκ recombination [92]; another candidate is IRF-4, which directs the *Igk* allele away from the pericentromeric heterochromatin [93]. In the case of the *Tcrb* locus, the helix-loop-helix protein, E47, is a good candidate to direct the interaction of this locus with pericentromeric heterochromatin since its dosage is rate-limiting with regard to V(D)J recombination and forced E47 expression interferes with pre-TCR-mediated feedback inhibition [94]. Additionally, it was also proposed that V(D)J recombination events occurring on one allele could activate signals that inhibit rearrangements on the second allele [95]. Consistent with this idea, it has been shown recently that homologous pairing of *Ig* alleles occurs during recombination and is mediated by RAG-1/2 binding [88, 96]. Furthermore, it has been demonstrated that RAG-mediated cleavage on one allele induces the other allele to relocate to pericentromeric heterochromatin by a mechanism related to the recognition of the cleaved allele by the DNA damage sensor ataxia telangiectasia mutated (ATM) protein [88]. Hence, activation of ATM by the cleaved allele acts in trans on the uncleaved allele to prevent recombination. Interallelic pairing has been proposed as a general mechanism for establishing the monoallelic gene expression that contributes to the maintenance of genomic integrity and suppresses oncogenic translocations during V(D)J recombination of antigen receptor loci.

5. Conclusions

V(D)J recombination is essential for the development of adaptive immune responses in vertebrates. In developing lymphocytes, V(D)J recombination is subjected to very tight spatial and temporal regulation. The regulation of this process is very complex and involves nuclear dynamics and changes in higher-order chromatin architecture to create gene segment accessibility to RAG-1/2 proteins. Active chromatin is bound by RAG-2 through interactions with specific H3K4me3, whereas RAG-1 binds to accessible RSSs derived from transcriptional elongation in large loci. The correlation between transcription and V(D)J recombination in both the recombined and allelic excluded antigen receptor loci led to the studies that established that the differential positioning of such loci at transcriptionally repressive nuclear regions might be responsible for allelic exclusion. Allelic association with the repressive nuclear compartments can inhibit V(D)J recombination by a mechanism other than transcription or RAG-1/2 accessibility to chromatin. Future experiments should be focused on identifying the cis-elements and trans-factors that regulate V(D)J recombination *in vivo*.

Acknowledgments

The authors are grateful for support from the Spanish Ministry of Science and Innovation Grant BFU2009-08796, Junta de Andalucía Grants CTS-6587 and CVI-4526, and CSIC Grants 200620I040 and 201020E060. The research in our lab is partly supported by European Region Development Funds (ERDF/FEDER). They are also grateful to an anonymous reviewer for comments on the paper.

References

[1] R. M. Cobb, K. J. Oestreich, O. A. Osipovich, and E. M. Oltz, "Accessibility control of V(D)J recombination," *Advances in Immunology*, vol. 91, pp. 45–109, 2006.

[2] D. G. Schatz and Y. Ji, "Recombination centres and the orchestration of V(D)J recombination," *Nature Reviews Immunology*, vol. 11, pp. 252–263, 2011.

[3] H. Sakano, K. Huppi, G. Heinrich, and S. Tonegawa, "Sequences at the somatic recombination sites of immunoglobulin light-chain genes," *Nature*, vol. 280, no. 5720, pp. 288–294, 1979.

[4] I. Abarrategui and M. S. Krangel, "Regulation of T cell receptor-alpha gene recombination by transcription," *Nature Immunology*, vol. 7, no. 10, pp. 1109–1115, 2006.

[5] I. Abarrategui and M. S. Krangel, "Noncoding transcription controls downstream promoters to regulate T-cell receptor α recombination," *EMBO Journal*, vol. 26, no. 20, pp. 4380–4390, 2007.

[6] P. Stanhope-Baker, K. M. Hudson, A. L. Shaffer, A. Constantinescu, and M. S. Schlissel, "Cell type-specific chromatin structure determines the targeting of V(D)J recombinase activity in vitro," *Cell*, vol. 85, no. 6, pp. 887–897, 1996.

[7] G. D. Yancopoulos and F. W. Alt, "Developmentally controlled and tissue-specific expression of unrearranged V(H) gene segments," *Cell*, vol. 40, no. 2, pp. 271–281, 1985.

[8] J. D. Fondell and K. B. Marcu, "Transcription of germ line Vα segments correlates with ongoing T-cell receptor α-chain

rearrangement," *Molecular and Cellular Biology*, vol. 12, no. 4, pp. 1480–1489, 1992.

[9] M. S. Schlissel, L. M. Corcoran, and D. Baltimore, "Virus-transformed preB-cells show ordered activation but not inactivation of immunoglobulin gene rearrangement and transcription," *Journal of Experimental Medicine*, vol. 173, no. 3, pp. 711–720, 1991.

[10] D. J. Bolland, A. L. Wood, C. M. Johnston et al., "Antisense intergenic transcription in V(D)J recombination," *Nature Immunology*, vol. 5, no. 6, pp. 630–637, 2004.

[11] D. Jung, C. Giallourakis, R. Mostoslavsky, and F. W. Alt, "Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus," *Annual Review of Immunology*, vol. 24, pp. 541–570, 2006.

[12] M. Goldmit and Y. Bergman, "Monoallelic gene expression: a repertoire of recurrent themes," *Immunological Reviews*, vol. 200, pp. 197–214, 2004.

[13] M. S. Krangel, "T cell development: better living through chromatin," *Nature Immunology*, vol. 8, no. 7, pp. 687–694, 2007.

[14] M. S. Schlissel, "Regulating antigen-receptor gene assembly," *Nature Reviews Immunology*, vol. 3, no. 11, pp. 890–899, 2003.

[15] H. D. Kondilis-Mangum, R. M. Cobb, O. Osipovich, S. Srivatsan, E. M. Oltz, and M. S. Krangel, "Transcription-dependent mobilization of nucleosomes at accessible TCR gene segments in vivo," *Journal of Immunology*, vol. 184, no. 12, pp. 6970–6977, 2010.

[16] O. Osipovich, R. Milley, A. Meade et al., "Targeted inhibition of V(D)J recombination by a histone methyltransferase," *Nature Immunology*, vol. 5, no. 3, pp. 309–316, 2004.

[17] J. Kwon, K. B. Morshead, J. R. Guyon, R. E. Kingston, and M. A. Oettinger, "Histone acetylation and hSWI/SNF remodeling act in concert to stimulate V(D)J cleavage of nucleosomal DNA," *Molecular Cell*, vol. 6, no. 5, pp. 1037–1048, 2000.

[18] J. Kwon, A. N. Imbalzano, A. Matthews, and M. A. Oettinger, "Accessibility of nucleosomal DNA to V(D)J cleavage is modulated by RSS positioning and HMG1," *Molecular Cell*, vol. 2, no. 6, pp. 829–839, 1998.

[19] A. Golding, S. Chandler, E. Ballestar, A. P. Wolffe, and M. S. Schlissel, "Nucleosome structure completely inhibits in vitro cleavage by the V(D)J recombinase," *EMBO Journal*, vol. 18, no. 13, pp. 3712–3723, 1999.

[20] F. McBlane and J. Boyes, "Stimulation of V(D)J recombination by histone acetylation," *Current Biology*, vol. 10, no. 8, pp. 483–486, 2000.

[21] O. Osipovich, R. M. Cobb, K. J. Oestreich, S. Pierce, P. Ferrier, and E. M. Oltz, "Essential function for SWI-SNF chromatin-remodeling complexes in the promoter-directed assembly of *Tcrb* genes," *Nature Immunology*, vol. 8, no. 8, pp. 809–816, 2007.

[22] N. Patenge, S. K. Elkin, and M. A. Oettinger, "ATP-dependent remodeling by SWI/SNF and ISWI proteins stimulates V(D)J cleavage of 5 S arrays," *Journal of Biological Chemistry*, vol. 279, no. 34, pp. 35360–35367, 2004.

[23] H. Du, H. Ishii, M. J. Pazin, and R. Sen, "Activation of 12/23-RSS-dependent RAG cleavage by hSWI/SNF complex in the absence of transcription," *Molecular Cell*, vol. 31, no. 5, pp. 641–649, 2008.

[24] Y. Ji, A. J. Little, J. K. Banerjee et al., "Promoters, enhancers, and transcription target RAG1 binding during V(D)J recombination," *Journal of Experimental Medicine*, vol. 207, no. 13, pp. 2809–2816, 2010.

[25] Y. Liu, R. Subrahmanyam, T. Chakraborty, R. Sen, and S. Desiderio, "A plant homodomain in RAG-2 that binds hypermethylated lysine 4 of histone H3 is necessary for efficient antigen-receptor-gene rearrangement," *Immunity*, vol. 27, no. 4, pp. 561–571, 2007.

[26] S. Ramón-Maiques, A. J. Kuo, D. Carney et al., "The plant homeodomain finger of RAG2 recognizes histone H3 methylated at both lysine-4 and arginine-2," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 48, pp. 18993–18998, 2007.

[27] Y. Ji, W. Resch, E. Corbett, A. Yamane, R. Casellas, and D. G. Schatz, "The in vivo pattern of binding of RAG1 and RAG2 to antigen receptor loci," *Cell*, vol. 141, no. 3, pp. 419–431, 2010.

[28] K. J. Oestreich, R. M. Cobb, S. Pierce, J. Chen, P. Ferrier, and E. M. Oltz, "Regulation of TCRβ gene assembly by a promoter/enhancer holocomplex," *Immunity*, vol. 24, no. 4, pp. 381–391, 2006.

[29] J. A. Skok, R. Gisler, M. Novatchkova, D. Farmer, W. de Laat, and M. Busslinger, "Reversible contraction by looping of the *Tcra* and *Tcrb* loci in rearranging thymocytes," *Nature Immunology*, vol. 8, no. 4, pp. 378–387, 2007.

[30] S. Spicuglia, S. Kumar, J.-H. Yeh et al., "Promoter activation by enhancer-dependent and -independent loading of activator and coactivator complexes," *Molecular Cell*, vol. 10, no. 6, pp. 1479–1487, 2002.

[31] E. Roldan, M. Fuxa, W. Chong et al., "Locus "decontraction" and centromeric recruitment contribute to allelic exclusion of the immunoglobulin heavy-chain gene," *Nature Immunology*, vol. 6, no. 1, pp. 31–41, 2005.

[32] H. Y. Shih, B. Hao, and M. S. Krangel, "Orchestrating T-cell receptor α gene assembly through changes in chromatin structure and organization," *Immunologic Research*, vol. 49, no. 1–3, pp. 192–201, 2011.

[33] H. Y. Shih and M. S. Krangel, "Distinct contracted conformations of the *Tcra/Tcrd* locus during *Tcra* and *Tcrd* recombination," *Journal of Experimental Medicine*, vol. 207, no. 9, pp. 1835–1841, 2010.

[34] M. Fuxa, J. Skok, A. Souabni, G. Salvagiotto, E. Roldán, and M. Busslinger, "Pax5 induces V-to-DJ rearrangements and locus contraction of the immunoglobulin heavy-chain gene," *Genes and Development*, vol. 18, no. 4, pp. 411–422, 2004.

[35] D. Reynaud, I. A. Demarco, K. L. Reddy et al., "Regulation of B cell fate commitment and immunoglobulin heavy-chain gene rearrangements by Ikaros," *Nature Immunology*, vol. 9, no. 8, pp. 927–936, 2008.

[36] H. Liu, M. Schmidt-Supprian, Y. Shi et al., "Yin Yang 1 is a critical regulator of B-cell development," *Genes and Development*, vol. 21, no. 10, pp. 1179–1189, 2007.

[37] D. G. T. Hesslein, D. L. Pflugh, D. Chowdhury, A. L. M. Bothwell, R. Sen, and D. G. Schatz, "Pax5 is required for recombination of transcribed, acetylated, 5′*IgH* V gene segments," *Genes and Development*, vol. 17, no. 1, pp. 37–42, 2003.

[38] Z. Zhang, C. R. Espinoza, Z. Yu et al., "Transcription factor Pax5 (BSAP) transactivates the RAG-mediated V_H-to-DJV_H rearrangement of immunoglobulin genes," *Nature Immunology*, vol. 7, no. 6, pp. 616–624, 2006.

[39] E. Splinter, H. Heath, J. Kooren et al., "CTCF mediates long-range chromatin looping and local histone modification in the β-globin locus," *Genes and Development*, vol. 20, no. 17, pp. 2349–2354, 2006.

[40] L. Han, D. H. Lee, and P. E. Szabo, "CTCF is the master organizer of domain-wide allele-specific chromatin at the

H19/Igf2 imprinted region," *Molecular and Cellular Biology*, vol. 28, no. 3, pp. 1124–1135, 2008.

[41] S. Hadjur, L. M. Williams, N. K. Ryan et al., "Cohesins form chromosomal cis-interactions at the developmentally regulated *IFNG* locus," *Nature*, vol. 460, no. 7253, pp. 410–413, 2009.

[42] S. Cai, C. C. Lee, and T. Kohwi-Shigematsu, "SATB1 packages densely looped, transcriptionally active chromatin for coordinated expression of cytokine genes," *Nature Genetics*, vol. 38, no. 11, pp. 1278–1288, 2006.

[43] R. Nativio, K. S. Wendt, Y. Ito et al., "Cohesin is required for higher-order chromatin conformation at the imprinted IGF_2-H_{19} locus," *PLoS Genetics*, vol. 5, no. 11, Article ID e1000739, 2009.

[44] S. Galande, P. K. Purbey, D. Notani, and P. P. Kumar, "The third dimension of gene regulation: organization of dynamic chromatin loopscape by SATB1," *Current Opinion in Genetics and Development*, vol. 17, no. 5, pp. 408–414, 2007.

[45] S. C. Degner, T. P. Wong, G. Jankevicius, and A. J. Feeney, "Developmental stage-specific recruitment of cohesin to CTCF sites throughout immunoglobulin loci during B lymphocyte development," *Journal of Immunology*, vol. 182, no. 1, pp. 44–48, 2009.

[46] S. C. Degner-Leisso and A. J. Feeney, "Epigenetic and 3-dimensional regulation of V(D)J rearrangement of immunoglobulin genes," *Seminars in Immunology*, vol. 22, no. 6, pp. 346–352, 2010.

[47] S. C. Degner, J. Verma-Gaur, T. P. Wong et al., "CCCTC-binding factor (CTCF) and cohesin influence the genomic architecture of the *Igh* locus and antisense transcription in pro-B cells," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, no. 23, pp. 9566–9571, 2011.

[48] W. J. Romanow, A. W. Langerak, P. Goebel et al., "E2A and EBF act in synergy with the V(D)J recombinase to generate a diverse immunoglobulin repertoire in nonlymphoid cells," *Molecular Cell*, vol. 5, no. 2, pp. 343–353, 2000.

[49] J. K. Ghosh, W. J. Romanow, and C. Murre, "Induction of a diverse T cell receptor γ/δ repertoire by the helix-loop-helix proteins E2A and HEB in nonlymphoid cells," *Journal of Experimental Medicine*, vol. 193, no. 6, pp. 769–775, 2001.

[50] R. Casellas, M. Jankovic, G. Meyer et al., "OcaB is required for normal transcription and V(D)J recombination of a subset of immunoglobulin κ genes," *Cell*, vol. 110, no. 5, pp. 575–585, 2002.

[51] E. Bertolino, K. Reddy, K. L. Medina, E. Parganas, J. Ihle, and H. Singh, "Regulation of interleukin 7-dependent immunoglobulin heavy-chain variable gene rearrangements by transcription factor STAT5," *Nature Immunology*, vol. 6, no. 8, pp. 836–843, 2005.

[52] S. K. Ye, Y. Agata, H. C. Lee et al., "The IL-7 receptor controls the accessibility of the TCRγ locus by Stat5 and histone acetylation," *Immunity*, vol. 15, no. 5, pp. 813–823, 2001.

[53] M. S. Krangel, J. Carabana, I. Abbarategui, R. Schlimgen, and A. Hawwari, "Enforcing order within a complex locus: Current perspectives on the control of V(D)J recombination at the murine T-cell receptor α/δ locus," *Immunological Reviews*, vol. 200, pp. 224–232, 2004.

[54] B. P. Sleckman, C. G. Bardon, R. Ferrini, L. Davidson, and F. W. Alt, "Function of the TCR α enhancer in $\alpha\beta$ and $\gamma\delta$ T cells," *Immunity*, vol. 7, no. 4, pp. 505–515, 1997.

[55] A. Hawwari, C. Bock, and M. S. Krangel, "Regulation of T cell receptor α gene assembly by a complex hierarchy of germline Jα promoters," *Nature Immunology*, vol. 6, no. 5, pp. 481–489, 2005.

[56] C. Angelin-Duclos and K. Calame, "Evidence that immunoglobulin V_H-DJ recombination does not require germ line transcription of the recombining variable gene segment," *Molecular and Cellular Biology*, vol. 18, no. 11, pp. 6253–6264, 1998.

[57] S. R. Cherry and D. Baltimore, "Chromatin remodeling directly activates V(D)J recombination," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 19, pp. 10788–10793, 1999.

[58] C. E. Whitehurst, S. Chattopadhyay, and J. Chen, "Control of V(D)J recombinational accessibility of the Dβ1 gene segment at the TCRβ locus by a germline promoter," *Immunity*, vol. 10, no. 3, pp. 313–322, 1999.

[59] N. Mathieu, W. M. Hempel, S. Spicuglia, C. Verthuy, and P. Ferrier, "Chromatin remodeling by the T cell receptor (TCR)-β gene enhancer during early T cell development: implications for the control of TCR-β locus recombination," *Journal of Experimental Medicine*, vol. 192, no. 5, pp. 625–636, 2000.

[60] M. L. Sikes, A. Meade, R. Tripathi, M. S. Krangel, and E. M. Oltz, "Regulation of V(D)J recombination: a dominant role for promoter positioning in gene segment accessibility," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 19, pp. 12309–12314, 2002.

[61] M. A. Schwabish and K. Struhl, "Asf1 mediates histone eviction and deposition during elongation by RNA polymerase II," *Molecular Cell*, vol. 22, no. 3, pp. 415–422, 2006.

[62] A. Kristjuhan and J. Q. Svejstrup, "Evidence for distinct mechanisms facilitating transcript elongation through chromatin in vivo," *EMBO Journal*, vol. 23, no. 21, pp. 4243–4252, 2004.

[63] A. Saha, J. Wittmeyer, and B. R. Cairns, "Chromatin remodelling: the industrial revolution of DNA around histones," *Nature Reviews Molecular Cell Biology*, vol. 7, no. 6, pp. 437–447, 2006.

[64] C. L. Smith and C. L. Peterson, "ATP-dependent chromatin remodeling," *Current Topics in Developmental Biology*, vol. 65, pp. 115–148, 2005.

[65] J. L. Workman, "Nucleosome displacement in transcription," *Genes and Development*, vol. 20, no. 15, pp. 2009–2017, 2006.

[66] R. J. Sims, R. Belotserkovskaya, and D. Reinberg, "Elongation by RNA polymerase II: the short and long of it," *Genes and Development*, vol. 18, no. 20, pp. 2437–2468, 2004.

[67] B. Li, M. Carey, and J. L. Workman, "The role of chromatin during transcription," *Cell*, vol. 128, no. 4, pp. 707–719, 2007.

[68] B. O. Wittschieben, G. Otero, T. de Bizemont et al., "A novel histone acetyltransferase is an integral subunit of elongating RNA polymerase II holoenzyme," *Molecular Cell*, vol. 4, no. 1, pp. 123–128, 1999.

[69] 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.

[70] 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.

[71] 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.

[72] T. Xiao, C. F. Kao, N. J. Krogan et al., "Histone H2B ubiquitylation is associated with elongating RNA polymerase

II," *Molecular and Cellular Biology*, vol. 25, no. 2, pp. 637–651, 2005.

[73] M. Albert and K. Helin, "Histone methyltransferases in cancer," *Seminars in Cell and Developmental Biology*, vol. 21, no. 2, pp. 209–220, 2010.

[74] H. Santos-Rosa, R. Schneider, B. E. Bernstein et al., "Methylation of histone H3 K4 mediates association of the Isw1p ATPase with chromatin," *Molecular Cell*, vol. 12, no. 5, pp. 1325–1332, 2003.

[75] M. Carey, B. Li, and J. L. Workman, "RSC exploits histone acetylation to abrogate the nucleosmal block to RNA polymerase II elongation," *Molecular Cell*, vol. 24, no. 3, pp. 481–487, 2006.

[76] L. L. Corey, C. S. Weirich, I. J. Benjamin, and R. E. Kingston, "Localized recruitment of a chromatin-remodeling activity by an activator in vivo drives transcriptional elongation," *Genes and Development*, vol. 17, no. 11, pp. 1392–1401, 2003.

[77] R. Belotserkovskaya, S. Oh, V. A. Bondarenko, G. Orphanides, V. M. Studitsky, and D. Reinberg, "FACT facilitates transcription-dependent nucleosome alteration," *Science*, vol. 301, no. 5636, pp. 1090–1093, 2003.

[78] S. Pandit, D. Wang, and X. D. Fu, "Functional integration of transcriptional and RNA processing machineries," *Current Opinion in Cell Biology*, vol. 20, no. 3, pp. 260–265, 2008.

[79] M. Senoo, L. Wang, D. Suzuki, N. Takeda, Y. Shinkai, and S. Habu, "Increase of TCR Vβ accessibility within Eβ regulatory region influences its recombination frequency but not allelic exclusion," *Journal of Immunology*, vol. 171, no. 2, pp. 829–835, 2003.

[80] A. Jackson, H. D. Kondilis, B. Khor, B. P. Sleckman, and M. S. Krangel, "Regulation of T cell receptor β allelic exclusion at a level beyond accessibility," *Nature Immunology*, vol. 6, no. 2, pp. 189–197, 2005.

[81] R. J. Schlimgen, K. L. Reddy, H. Singh, and M. S. Krangel, "Initiation of allelic exclusion by stochastic interaction of *Tcrb* alleles with repressive nuclear compartments," *Nature Immunology*, vol. 9, no. 7, pp. 802–809, 2008.

[82] J. Jia, M. Kondo, and Y. Zhuang, "Germline transcription from T-cell receptor Vβ gene is uncoupled from allelic exclusion," *EMBO Journal*, vol. 26, no. 9, pp. 2387–2399, 2007.

[83] N. Singh, Y. Bergman, H. Cedar, and A. Chess, "Biallelic germline transcription at the κ immunoglobulin locus," *Journal of Experimental Medicine*, vol. 197, no. 6, pp. 743–750, 2003.

[84] R. H. Amin, D. Cado, H. Nolla et al., "Biallelic, ubiquitous transcription from the distal germline *Igk* locus promoter during B cell development," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 2, pp. 522–527, 2009.

[85] S. T. Kosak, J. A. Skok, K. L. Medina et al., "Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development," *Science*, vol. 296, no. 5565, pp. 158–162, 2002.

[86] M. Goldmit, Y. Ji, J. Skok et al., "Epigenetic ontogeny of the Igk locus during B cell development," *Nature Immunology*, vol. 6, no. 2, pp. 198–203, 2005.

[87] S. Jhunjhunwala, M. C. van Zelm, M. M. Peak, and C. Murre, "Chromatin architecture and the generation of antigen receptor diversity," *Cell*, vol. 138, no. 3, pp. 435–448, 2009.

[88] S. L. Hewitt, B. Yin, J. Chaumeil et al., "RAG-1 and ATM coordinate monoallelic recombination and nuclear positioning of immunoglobulin loci," *Nature Immunology*, vol. 10, pp. 655–664, 2009.

[89] J. Maes, L. P. O'Neill, P. Cavelier, B. M. Turner, F. Rougeon, and M. Goodhardt, "Chromatin remodeling at the Ig loci prior to V(D)J recombination," *Journal of Immunology*, vol. 167, no. 2, pp. 866–874, 2001.

[90] R. Tripathi, A. Jackson, and M. S. Krangel, "A change in the structure of Vβ chromatin associated with TCR β allelic exclusion," *Journal of Immunology*, vol. 168, no. 5, pp. 2316–2324, 2002.

[91] D. Chowdhury and R. Sen, "Transient IL-7/IL-7R signaling provides a mechanism for feedback inhibition of immunoglobulin heavy chain gene rearrangements," *Immunity*, vol. 18, no. 2, pp. 229–241, 2003.

[92] Z. Liu, P. Widlak, Y. Zou et al., "A recombination silencer specifies heterochromatin positioning and Ikaros association in the immunoglobulin κ locus," *Immunity*, vol. 24, no. 4, pp. 405–415, 2006.

[93] K. Johnson, T. Hashimshony, C. M. Sawai et al., "Regulation of immunoglobulin light chain recombination by the transcription factor IRF-4 and the attenuation of interleukin-7 signaling," *Immunity*, vol. 28, no. 3, pp. 335–345, 2008.

[94] Y. Agata, N. Tamaki, S. Sakamoto et al., "Regulation of T cell receptor β gene rearrangement and allelic exclusion by the helix-loop-helix protein, E47," *Immunity*, vol. 27, no. 6, pp. 871–884, 2007.

[95] F. W. Alt, V. Enea, A. L. M. Bothwell, and D. Baltimore, "Activity of multiple light chain genes in murine myeloma cells producing a single, functional light chain," *Cell*, vol. 21, no. 1, pp. 1–12, 1980.

[96] V. L. Brandt, S. L. Hewitt, and J. A. Skok, "It takes two. Communication between homologous alleles preserves genomic stability during V(D)J recombination," *Nucleus*, vol. 1, no. 1, pp. 23–29, 2010.

SAGE-Hindawi Access to Research
Genetics Research International
Volume 2011, Article ID 309865, 9 pages
doi:10.4061/2011/309865

Review Article

RNA Polymerase II Elongation at the Crossroads of Transcription and Alternative Splicing

Manuel de la Mata,[1,2] Manuel J. Muñoz,[1] Mariano Alló,[1] Juan Pablo Fededa,[1,3] Ignacio E. Schor,[1] and Alberto R. Kornblihtt[1]

[1] *Laboratorio de Fisiología y Biología Molecular, Departamento de Fisiología, Biología Molecular, y Celular, IFIBYNE-CONICET, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, C1428EHA Buenos Aires, Argentina*
[2] *Friedrich Miescher Institute for Biomedical Research, 4002 Basel, Switzerland*
[3] *Department of Biology, Institute of Biochemistry, ETH Zurich, 8093 Zurich, Switzerland*

Correspondence should be addressed to Alberto R. Kornblihtt, ark@fbmc.fcen.uba.ar

Received 15 June 2011; Accepted 23 June 2011

Academic Editor: Carles Sune

The elongation phase of transcription lies at the core of several simultaneous and coupled events leading to alternative splicing regulation. Although underestimated in the past, it is at this phase of the transcription cycle where complexes affecting the transcription machinery itself, chromatin structure, posttranscriptional gene regulation and pre-mRNA processing converge to regulate each other or simply to consolidate higher-order complexes and functions. This paper focuses on the multiple processes that take place during transcription elongation which ultimately regulate the outcome of alternative splicing decisions.

1. Introduction

Regulation of gene expression was originally conceived as a hierarchy of steps linked together on a time scale and physically separated in different cell compartments in accordance with the central dogma of biology. This concept has long been abandoned, with a significant accumulation of evidence describing an extensive network of events, encompassing transcription, mRNA processing, chromatin regulation, and posttranscriptional gene regulation, which take place simultaneously and in a mutually regulated or coupled manner [1, 2]. Distinctions between complexes and processes governing gene expression have been blurred to a large extent, adding complexity to the ever-increasing fraction of genes subjected to alternative promoter usage, alternative splicing (AS) (>90% genes), alternative polyadenylation, editing, and posttranscriptional gene silencing by small RNAs [3, 4]. Additionally, this complexity takes a new dimension when studied in the context of chromatin and its regulation upon gene expression (for reviews, see [5–7]). This paper will focus on the main features of coupling between transcriptions elongation and splicing, and its implications on AS regulation.

2. The Benefits of Coupling

Initial visualization of Drosophila-embryo nascent transcripts by electron microscopy, showed that splicing can occur cotranscriptionally [8]. This was later directly demonstrated for the human dystrophin gene [9], which spans 2400 kb and can take 16 hr to complete transcription. More recently, a quantitative study of the c-Src and fibronectin mRNAs compared chromatin-bound and nucleoplasmic RNA fractions. There, it was shown that most introns are excised efficiently in the chromatin-bound fractions, with a gradient of cotranscriptional splicing efficiency decreasing from promoter-proximal to promoter-distal introns, that is, the direction of transcription [10].

One implication of the cotranscriptional nature of splicing is that the two processes can be coupled. In a broad sense, coupling implies that the involved processes can happen efficiently only as the result of their combined action, even

for processes that are constitutive and non-regulated. For instance, whereas both transcription and splicing can take place independently at low efficiency as *in vitro* reactions, it is only *in vivo* or in coupled *in vitro* systems where maximal efficiency can be achieved [11–16]. Cotranscriptional processing is necessary to allow for coupling between transcription and splicing, although it does not necessarily guarantee it. There are examples of both cotranscriptional splicing that seems to be uncoupled, as well as purely posttranscriptional splicing [17–20]. Noteworthy, the consequences of the different types of splicing can be considerable; whereas cotranscriptional splicing can be regulated by mechanisms dependent on transcription, postranscriptional splicing can be subjected to additional regulatory mechanisms linked to events downstream of transcription (e.g., RNA export) [20]. However, cotranscriptional splicing seems to predominate for most introns in mammalian genes [10, 17, 20–23] pointing at an evolutionary conserved role in allowing for coupling of transcription and splicing. Cotranscriptional splicing is more efficient than posttranscriptional splicing by driving nascent pre-mRNAs to the association with spliceosome components [24, 25] and splicing regulatory factors, such as serine/arginine-rich (SR) proteins [22]. This allows for different levels of regulation of AS and prevents backhybridization of the nascent pre-mRNA to the DNA template strand, which can cause genome instability [26, 27]. Even in the case of posttranscriptional splicing, coupling with transcription can be determinant for AS regulation. Since pre-mRNA splicing is a multistep reaction, it is possible that commitment to splicing takes place cotranscriptionally during early splice-site recognition, while completion of the splicing reaction occurs posttranscriptionally [20, 28, 29], consistent with the fact that introns are not necessarily removed in the exact order that they are transcribed [17, 30–32]. This mechanism, which can be viewed as a cotranscriptional commitment rather than a cotranscriptional catalysis, tends to apply largely to splicing and not to other RNA-processing events like capping and cleavage/polyadenylation [33–37]. Intermediate scenarios are also possible, with both splicing commitment and catalysis taking place cotranscriptionally but not following a strict 5′ to 3′ direction of intron removal [10, 38].

Another implication of cotranscriptional splicing is that it allows for a bidirectional coupling of the two processes [1, 20, 39, 40]. For instance, the splicing machinery can reciprocally affect transcription in different ways by either stimulating transcriptional elongation [41, 42], transcriptional initiation [43, 44] and, as recently shown in yeast, by imposing a transient pausing checkpoint around the 3′ end of introns and on terminal exons [45, 46]. This bidirectional feedback might in turn reinforce splicing efficiency, conferring important advantages for gene expression. Nevertheless, reciprocal coupling might not be a widespread general phenomenon considering that elongation kinetics seems to be independent of splicing in some model genes [47]. This highlights the possibility that specific exon-intron architectures and/or cis-acting sequences might be required for reciprocal coupling to occur.

Despite all the seemingly clear advantages of cotranscriptional over posttranscriptional splicing, particularly in allowing for coupling, the true proportions of these two modes of splicing in mammals still await to be determined on a genome-wide scale. It will be interesting to use global approaches to answer this question, which might have profound implications in reorienting our current research and on our understanding of the regulation of AS.

3. Early Discoveries

In our current view, the regulation of AS is the result of the combined action of splicing factors acting on splicing enhancers and silencers, regulatory secondary structure motifs of mRNAs, and the coupling with RNA polymerase II (Pol II) transcription [22, 48–51]. An early indication for coupling was the finding that promoter identity affects splicing decisions independently of the strength of the promoter, opposing the classical view whereby promoters are limited to affect transcription levels ([52–56], for review, see [57]). Different promoters, such as those of the α-globin and fibronectin (FN) genes, were shown to induce a 10-fold difference in inclusion of the human FN alternative exon 33 (E33, also referred to as EDI or EDA) when driving its expression from reporter minigenes in transiently transfected mammalian cells [52, 53] (Figure 1).

One implication of the promoter effect on AS is that splicing factors could regulate AS through promoters. Cell-specific AS events could then arise from cell-specific promoter occupation rather than from the differential abundance of ubiquitous splicing factors. Under physiological conditions, promoter architecture could then control AS through the differential occupation by different transcription factors. Supporting this hypothesis, evidence shows that transcriptional activators and coactivators, with different effects on Pol II, indeed affect AS differentially [58, 59].

4. The Kinetics of Coupling

Transcription appears to influence pre-mRNA splicing through at least two independent modes: by kinetically coupling the processing reactions, where the rate of Pol II elongation influences the outcome of the alternative events [60, 61], and by physically and functionally recruiting mRNA processing factors to the transcription machinery, in particular to Pol II's carboxyterminal domain (CTD) [13, 62–65]. In fact, recruitment of splicing factors to sites of transcription is dependent on RNA Pol II CTD [66] and deletion of the CTD impairs capping, cleavage/polyadenylation, and splicing of the b-globin transcript [34]. Here we provide a view of the kinetic mode of coupling, although accumulated evidence supports that both modes can operate simultaneously in a nonmutually exclusive manner (for reviews on the recruitment mode of coupling, see [2, 65, 67]).

It should be pointed out that transcription itself is a complex multistep and regulated process, organized as a transcription cycle, with each step subjected to extensive regulation [68]. However, it is mainly at the elongation step of transcription where most of the connections between

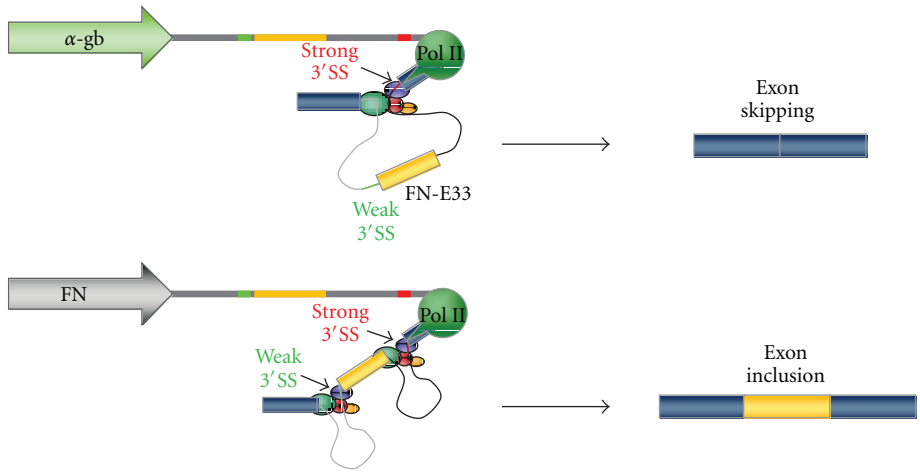


Figure 1: Promoters affect alternative splicing. α-globin/FN hybrid minigenes under the control of different promoters, used in transient transfections of mammalian cells in culture to assess inclusion levels of the alternatively spliced E33 (EDI or EDA) cassette exon (dark yellow). Inclusion level with the FN promoter is >10-fold higher when compared to the α-globin promoter.

the transcription and splicing machineries actually occur. A role for Pol II elongation on AS had been suggested before the finding of the promoter effect [69] and was later supported by several lines of evidence. Eperon et al. [69] showed that, in contrast to *in vitro* conditions, ongoing RNA synthesis *in vivo* affects the potential secondary structure of long—but not that of short—RNA substrates, which in turn affects splicing, pointing at a kinetic link between transcription and splicing. Additional evidence came from experiments in which RNA Pol II's local pausing caused by elements inserted into the tropomyosin gene, promoted higher inclusion of tropomyosin exon 3 [60]. However, more conclusive evidence for a role of elongation on AS regulation was shown in a series of reports demonstrating that several factors globally impacting Pol II elongation also affect AS. (i) Replication of AS reporter minigenes greatly stimulates FN E33 inclusion, and is counteracted by trichostatin A (TSA), a potent inhibitor of histone deacetylation considered to drive chromatin into an "open" state. This suggested that replication conveys a more compact chromatin structure to the template, thus slowing elongation and leading to higher E33 inclusion [70]. (ii) Drugs known to inhibit elongation, such as DRB [58, 70], flavopiridol, or camptothecin [38], favor E33 inclusion. (iii) Transcriptional activation by VP16, a factor that promotes both initiation and elongation, decreases E33 inclusion while Sp1, acting only on initiation, has no effect on E33 inclusion [58]. (iv) The presence of the SV40 transcriptional enhancer near a promoter stimulates Pol II elongation and provokes a 3–10-fold reduction in FN E33 inclusion independently of the promoter used [71]. (v) Slow mutants of RNA Pol II increase FN E33 inclusion in human cells, affect AS of the endogenous gene ultrabithorax (Ubx) in Drosophila, and modulate the inclusion of an artificially created alternative exon in yeast [61, 72]. (vi) DNA-damage signaling triggered by UV irradiation affects the AS of fibronectin, caspase 9, Bcl-x, and other human genes by inducing hyperphosphorylation of Pol II CTD and blocking Pol II elongation [73]. These data are in agreement with a "first come, first served" model for regulation of AS [74] (Figure 2). One interpretation of this model is that slow elongation favors the removal of the intron upstream of an alternative cassette exon before removal of the downstream intron. Alternatively, slow elongation would favor recruitment of splicing factors to the upstream intron before the downstream intron is synthesized. Once commitment to include the exon is achieved, the order of intron removal becomes irrelevant (Figure 2). The latter interpretation is supported by recent evidence showing a preferential removal of the intron downstream of the FN cassette exon 33 before the upstream intron has been removed [38]. Most importantly, whereas cis-acting mutations and trans-acting factors that upregulate E33 inclusion alter the relative order of intron removal, elongation slowdown also induces higher E33 inclusion without affecting the order of intron removal, suggesting that slow elongation favors commitment to exon inclusion during spliceosome assembly [38]. In light of these findings, "first served" would not be equivalent to "first excised" but to "first committed," in agreement with the observed preferential cotranscriptionality of spliceosome recruitment rather than catalysis. The control of elongation on AS is not restricted to a few cases. Recent data confirms and extends the findings of elongation control on AS using a global approach based on AS microarray profiling [48]. This study demonstrates that a variety of conditions that impact Pol II elongation, including drug treatments (DRB and camptothecin), Pol II mutants that inhibit elongation (slow and CTD phosphorylation Pol II mutants), and cellular stress (UV irradiation, know to impact Pol II elongation), globally affect both the mRNA levels and AS of a significant number of the genes. Moreover, the largest statistically significant fraction of affected genes showed a concomitant decrease in steady-state mRNA levels with increase in exon inclusion levels, coincident with many of the examples previously reported [60, 61, 70, 72, 73, 75, 76]. Nevertheless, there was also a smaller but significant number of genes displaying decreased mRNA levels with

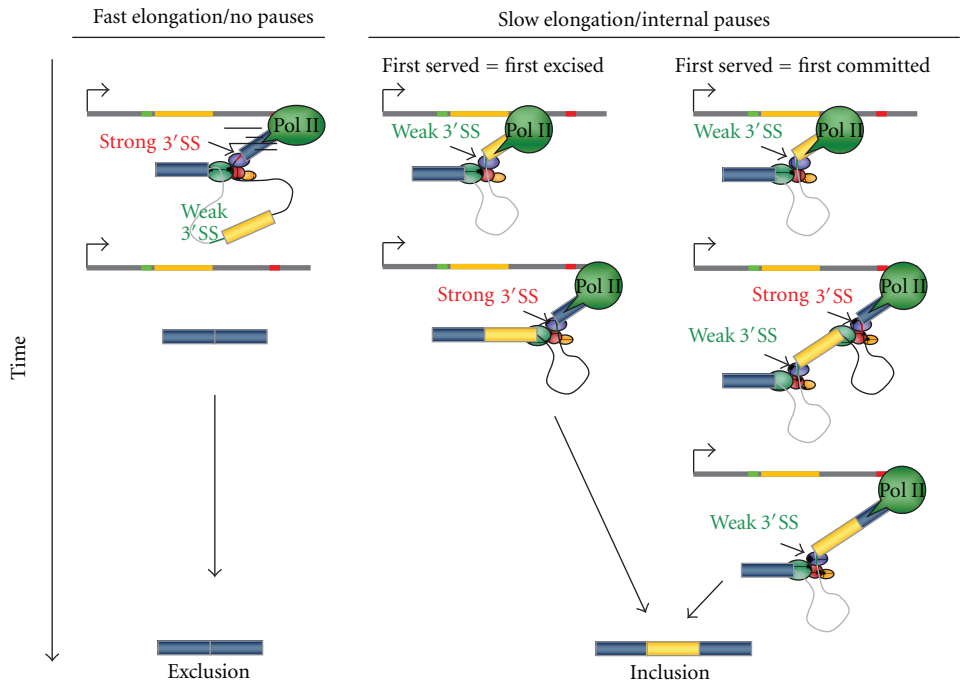


Figure 2: Alternative models for the "first come, first served" mechanism of splice site selection. (a) Fast elongation promotes usage of the stronger downstream 3′ splice site. (b) Slow elongation causes preferential excision of the upstream intron (first served = first excised). (c) Slow elongation causes commitment to E33 inclusion via recruitment of splicing factors (first served = first committed). Both introns are excised individually and in an order that is independent of elongation. (Based on [38].)

decreased exon inclusion levels [48], consistent with the idea that inhibition of Pol II elongation can lead to increased exon inclusion as well as increased exon skipping depending on the underlying splicing regulatory mechanism involved in each case [55, 61]. Pol II's elongation-dependent changes in AS regulation also display a high preference to modulate the expression levels of genes involved in RNA metabolism, including pre-mRNA splicing factors and other RNA binding proteins. Interestingly, about one-third of those genes contain AS events that introduce a premature termination codon (PTC) when spliced into the mature mRNA, subsequently leading to nonsense-mediated mRNA decay (NMD) [48]. This represents another example of evolutionarily conserved elongation-coupled events, acting together to coordinate the levels or RNA binding proteins with their steady-state mRNA levels.

5. Elongation Links to Chromatin

The promoter effect, together with the kinetic, physical, and functional coupling modes of transcription and AS, immediately shifted the attention to other factors thought to be restricted to transcriptional regulation, such as the chromatin structure. It soon became clear that the unanticipated complexity of splicing regulation could not be explained solely based on the current models, which lacked a clear connection to *in vivo* situations. Chromatin is the natural substrate upon which transcriptional regulation acts *in vivo*, and major discoveries have recently pointed at chromatin structure and post-translational histone modifications as key regulators of AS.

The chromatin role on AS regulation is broad, involving both direct (elongation-independent) interactions with the splicing machinery and effects on AS through changes in transcription elongation. Here, we concentrate on the later, although recent findings on the direct roles of chromatin on splicing have significantly changed our view of how exon-intron architecture is achieved by intimately linking nucleosome to exon structure ([77–84], for reviews, see [5–7]. In fact, genome-wide mapping of nucleosome positioning on exons at the DNA level may shed light on one of the most striking puzzles in the field of splicing, that is, how does the splicing machinery recognizes, with high fidelity, short exons (on average 150 bp) "floating" in a "sea" of introns (on average 5.4 kbp), an exon-intron architecture typical of vertebrate genes [85, 86]. Notably, the average size of a mammalian exon is similar to the length of DNA wrapped around a nucleosome, suggesting a conserved function for the nucleosome in exon definition [79, 80, 82]. According to the exon definition model, originally postulated by S. M. Berget [87], the spliceosome and auxiliary factors achieve this recognition by preferentially assembling on 3′ and 5′ sites paired across exons and not across introns (i.e., not following an intron definition mode of recognition, typical in lower eukaryotes like yeast). This favors exon recognition and acts as a selective force for short exon size. As described below, nucleosome positioning on exons may help in exon definition by creating roadblocks for Pol II elongation that

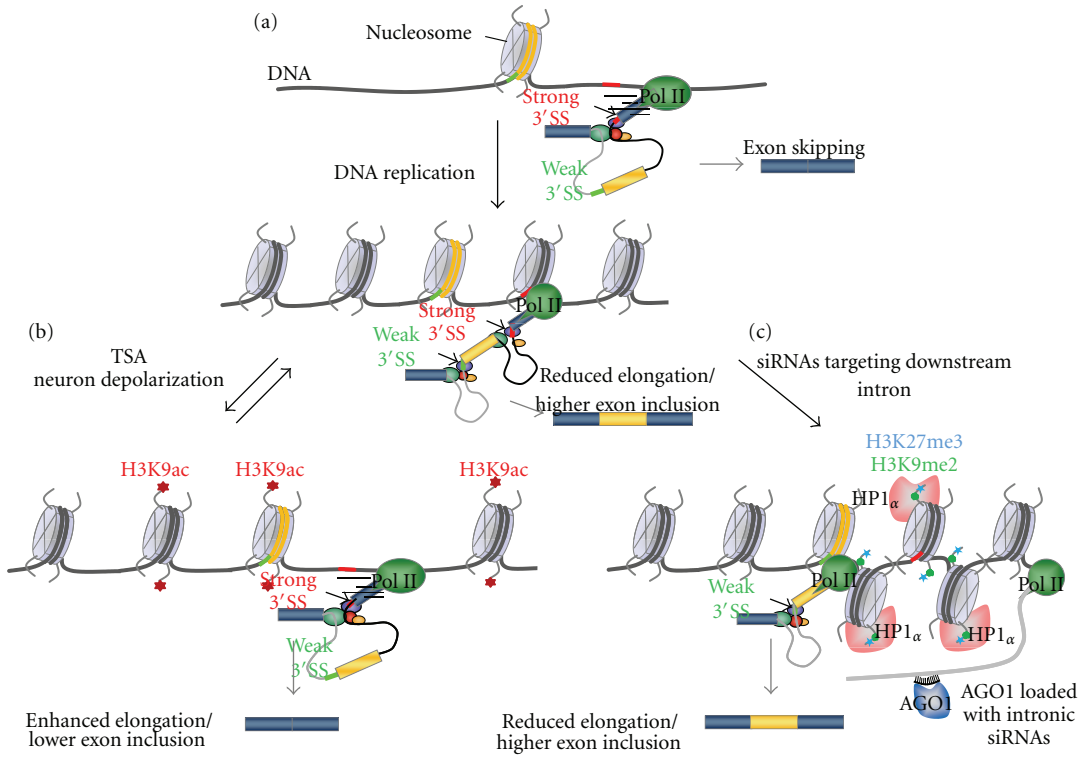


Figure 3: Chromatin couples elongation to alternative splicing. (a) Early evidence for a role of chromatin on splicing: replication affects alternative splicing. Loose nucleosome assembly (as in transiently transfected reporter minigenes) gives rise to low inclusion of the alternative exon (yellow) into the mature mRNA. After replication, nucleosome organization becomes more compact, promoting much higher inclusion of the alternative exon. (b) Depolarization of neuronal cells or treatment with TSA triggers intragenic histone acetylation and looser nucleosome compaction which in turn causes skipping of the alternative exon (yellow). (c) Model for TGS-AS. Transfection with siRNAs targeting the intron downstream from the alternative exon (yellow) promotes dimethylation and trimethylation of H3K9 and H3K27 (green and blue marks, resp.), triggered by siRNA's guide strand entering a silencing complex containing AGO1. HP1α is recruited and the resulting condensed chromatin structure generates roadblocks to Pol II elongation, causing higher inclusion of the alternative exon according to the kinetic coupling model.

provide longer time for cotranscriptional recognition of splice sites in the nascent pre-mRNA.

6. Chromatin Structure in Coupling

A role for chromatin in the coupling of elongation with AS was first suggested with the finding that two copies of the same adenovirus genome, either unreplicated or replicated in the same nucleus, gave rise to different alternatively spliced RNAs [57]. It was then speculated that the chromatin organization acquired after replication was more compact, with a subsequent reduction in Pol II elongation rates and more time to assemble splicing complexes at a suboptimal upstream 5′ splice site, thus favoring its use compared to the downstream 5′ splice site. This was confirmed by the finding that FN E33 inclusion is sensitive to the replication-mediated chromatinization status of the reporter plasmid [58, 70] (Figure 3(a). See above). Both in plants and humans genes, RNA Pol II distribution correlates with nucleosome deposition with preferential accumulation at exons relative to introns [79, 80, 83, 88], consistent with nucleosomes acting as barriers that locally modulate RNA Pol II density by inducing its pausing [89, 90]. This could in turn modulate splicing efficiency, in agreement with RNA Pol II being more highly enriched at alternatively spliced exons than at constitutive ones [88]. Nucleosome density also varies according to splice site strength, with stronger positioning of nucleosomes in included alternatively spliced exons than excluded ones [80] and stronger positioning at exons defined by weaker splice sites [81, 82]. Consistent with these results, overexpression of the ATPase-dependent chromatin-remodeling complex SWI/SNF subunit Brm in human cells, induces accumulation of Pol II with a modified CTD phosphorylation pattern on regions encoding variant exons of the CD44 gene and causes increased inclusion of these exons into mature mRNA [91].

7. Histone Modifications in Coupling

Histone modifications have emerged as major regulators of AS by either impacting the coupling of elongation with AS or by direct means such as recruiting splicing factors to the nascent pre-mRNA [7]. An indication for a role of histone modifications in the coupling of elongation with AS was the observation that treatment of cells with the histone deacetylase inhibitor TSA induces skipping of the alternatively

spliced fibronectin E33 and the neural cell adhesion molecule (NCAM) exon 18 [58, 75, 76]. In a more physiological context, depolarization of neuronal cells increases H3K9 acetylation and H3K36 methylation locally around the alternatively spliced exon 18 of NCAM which correlates with increased exon skipping [76] (Figure 3(b)). Interestingly, no changes in histone acetylation were observed at the NCAM promoter, suggesting that this reversible effect may be due to an intragenic and local modulation of the RNA Pol II elongation rate [76]. For instance, acetylated and open chromatin would induce fast Pol II elongation and skipping of the alternative exon. In line with these observations, targeting of an intronic sequence downstream of the alternatively spliced E33 of fibronectin with small interfering RNAs induces local heterochromatinization and increased E33 inclusion without changes in general transcription levels [75] (Figure 3(c)). In addition, inhibition of histone deacetylation, DNA methylation, H3K9 methylation, and downregulation of heterochromatin protein 1α (HP1α) abolish the siRNA-mediated effect on exon E33 splicing [75], suggesting a role of these modifications in AS regulation. Considering the multiple evidence for the involvement of chromatin in siRNA effects on splicing, this mechanism has been referred to as TGS-AS for transcriptional gene-silencing-regulated AS [92] (Figure 3(c)).

Several other epigenetic modifications, including DNA methylation, may also directly or indirectly act via histone modifications to affect splice site decisions [83, 93, 94]. This adds to the additional mechanism involving direct physical crosstalk between chromatin and the splicing machinery via an adaptor complex [7, 95, 96].

8. Conclusions and Outlook

The fact that transcriptional elongation is largely connected to pre-mRNA processing and particularly to AS regulation has boosted extensive efforts to understand its mechanisms and physiological implications. The link to chromatin seems to be the natural way in which elongation and splicing are truly engaged and might indeed explain the poorly understood mechanisms by which tissue- and cell-type-specific AS patterns are established, propagated, and maintained. Although some cell- and tissue-specific differences in expression of constitutive splicing factors have been reported [97], it is tempting to speculate that, in analogy to the histone code mechanisms used to specify gene expression, a histone-based system may also encode information that specifies cell- and tissue-specific AS patterns. As recently proposed, this would provide an epigenetic memory for splicing decisions likely to be heritable during proliferation and susceptible to modification along differentiation, without the need for major changes in the AS rules at each step of differentiation [7]. Nevertheless, it still remains to be determined whether the effects of histone modification on RNA processing are heritable and therefore authentic epigenetic modifications or whether they are just transient modulators.

Given the bidirectional nature of coupling, it is conceivable that splicing might reciprocally feed back on the chromatin structure via affecting the transcription machinery. In fact, besides recruiting splicing regulators such as SR proteins or U2 snRNP subunits [25, 64], Pol II can interact with histone modifiers such as the histone 3 lysine 36 (H3K36) methyltransferase Set2 [98], known, for example, to regulate histone deacetylation and prevent inappropriate initiation within the body of transcribed genes [99]. This interesting possibility would build up on the notion that transcription, chromatin, and splicing are intimately dependent on each other.

Future directions of the field will most likely aim at a more comprehensive view on the histone modifications role in AS. Histone modifications must be comprehensively mapped on a genome-wide scale in multiple cell types and tissues and compared to global AS patterns. New candidate players will be most likely be studied, such as noncoding RNAs recently shown to be involved in heterochromatin structure and associated with AS regulation [100, 101]. Ultimately, a better understanding of the multiple links between transcription and splicing will head the way to decipher the complexity of gene expression both in physiological and pathological conditions.

Acknowledgments

The authors thank Valeria Buggiano and the rest of the members of the Kornblihtt and Srebrow laboratories for their invaluable help. Work in their laboratory is supported by grants to A. R. Kornblihtt from the Agencia Nacional de Promoción de Ciencia y Tecnología of Argentina, the Universidad de Buenos Aires, and the European Alternative Splicing Network (EURASNET). A. R. Kornblihtt is a career investigator at the Consejo Nacional de Investigaciones Científicas y Técnicas of Argentina (CONICET) and an international research scholar of the Howard Hughes Medical Institute.

References

[1] A. R. Kornblihtt, M. de la Mata, J. P. Fededa, M. J. Muñoz, and G. Nogués, "Multiple links between transcription and splicing," *RNA*, vol. 10, no. 10, pp. 1489–1498, 2004.

[2] S. Pandit, D. Wang, and X. D. Fu, "Functional integration of transcriptional and RNA processing machineries," *Current Opinion in Cell Biology*, vol. 20, no. 3, pp. 260–265, 2008.

[3] Q. Pan, O. Shai, L. J. Lee, B. J. Frey, and B. J. Blencowe, "Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing," *Nature Genetics*, vol. 40, no. 12, pp. 1413–1415, 2008.

[4] E. T. Wang, R. Sandberg, S. Luo et al., "Alternative isoform regulation in human tissue transcriptomes," *Nature*, vol. 456, no. 7221, pp. 470–476, 2008.

[5] I. E. Schor, M. Alló, and A. R. Kornblihtt, "Intragenic chromatin modifications: a new layer in alternative splicing regulation," *Epigenetics*, vol. 5, no. 3, pp. 174–179, 2010.

[6] M. Alló, I. E. Schor, M. J. Muñoz et al., "Chromatin and alternative splicing," *Cold Spring Harbor Laboratory of Quantitative Biology*, vol. 75, pp. 103–111, 2010.

[7] R. F. Luco, M. Allo, I. E. Schor, A. R. Kornblihtt, and T. Misteli, "Epigenetics in alternative pre-mRNA splicing," *Cell*, vol. 144, no. 1, pp. 16–26, 2011.

[8] A. L. Beyer and Y. N. Osheim, "Splice site selection, rate of splicing, and alternative splicing on nascent transcripts," *Genes & Development*, vol. 2, no. 6, pp. 754–765, 1988.

[9] C. N. Tennyson, H. J. Klamut, and R. G. Worton, "The human dystrophin gene requires 16 hours to be transcribed and is cotranscriptionally spliced," *Nature Genetics*, vol. 9, no. 2, pp. 184–190, 1995.

[10] A. Pandya-Jones and D. L. Black, "Co-transcriptional splicing of constitutive and alternative exons," *RNA*, vol. 15, no. 10, pp. 1896–1908, 2009.

[11] S. Ghosh and M. A. Garcia-Blanco, "Coupled in vitro synthesis and splicing of RNA polymerase II transcripts," *RNA*, vol. 6, no. 9, pp. 1325–1334, 2000.

[12] I. Wetterberg, J. Zhao, S. Masich, L. Wieslander, and U. Skoglund, "In situ transcription and splicing in the Balbiani ring 3 gene," *The EMBO Journal*, vol. 20, no. 10, pp. 2564–2574, 2001.

[13] 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.

[14] B. J. Natalizio and M. A. Garcia-Blanco, "In vitro coupled transcription splicing," *Methods*, vol. 37, no. 4, pp. 314–322, 2005.

[15] R. Das, K. Dufu, B. Romney, M. Feldt, M. Elenko, and R. Reed, "Functional coupling of RNAP II transcription to spliceosome assembly," *Genes and Development*, vol. 20, no. 9, pp. 1100–1109, 2006.

[16] M. J. Hicks, C. R. Yang, M. V. Kotlajich, and K. J. Hertel, "Linking splicing to Pol II transcription stabilizes pre-mRNAs and influences splicing patterns.," *PLoS Biology*, vol. 4, no. 6, article e147, 2006.

[17] G. Bauren and L. Wieslander, "Splicing of Balbiani ring 1 gene pre-mRNA occurs simultaneously with transcription," *Cell*, vol. 76, no. 1, pp. 183–192, 1994.

[18] J. Laurencikiene, A. M. Källman, N. Fong, D. L. Bentley, and M. Öhman, "RNA editing and alternative splicing: the importance of co-transcriptional coordination," *EMBO Reports*, vol. 7, no. 3, pp. 303–307, 2006.

[19] D. Lazarev and J. L. Manley, "Concurrent splicing and transcription are not sufficient to enhance splicing efficiency," *RNA*, vol. 13, no. 9, pp. 1546–1557, 2007.

[20] J. Han, J. Xiong, D. Wang, and X.-D. Fu, "Pre-mRNA splicing: Where and when in the nucleus," *Trends in Cell Biology*, vol. 21, no. 6, pp. 336–343, 2011.

[21] G. Zhang, K. L. Taneja, R. H. Singer, and M. R. Green, "Localization of pre-mRNA splicing in mammalian nuclei," *Nature*, vol. 372, no. 6508, pp. 809–812, 1994.

[22] R. Das, J. Yu, Z. Zhang et al., "SR proteins function in coupling RNAP II transcription to pre-mRNA splicing," *Molecular Cell*, vol. 26, no. 6, pp. 867–881, 2007.

[23] J. Singh and R. A. Padgett, "Rates of in situ transcription and splicing in large human genes," *Nature Structural and Molecular Biology*, vol. 16, no. 11, pp. 1128–1133, 2009.

[24] S. A. Lacadie, D. F. Tardiff, S. Kadener, and M. Rosbash, "In vivo commitment to yeast cotranscriptional splicing is sensitive to transcription elongation mutants," *Genes and Development*, vol. 20, no. 15, pp. 2055–2066, 2006.

[25] I. Listerman, A. K. Sapra, and K. M. Neugebauer, "Cotranscriptional coupling of splicing factor recruitment and precursor messenger RNA splicing in mammalian cells," *Nature Structural and Molecular Biology*, vol. 13, no. 9, pp. 815–822, 2006.

[26] 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.

[27] A. Aguilera and B. Gómez-González, "Genome instability: a mechanistic view of its causes and consequences," *Nature Reviews Genetics*, vol. 9, no. 3, pp. 204–217, 2008.

[28] M. J. Moore, E. M. Schwartzfarb, P. Silver, and M. C. Yu, "Differential recruitment of the splicing machinery during transcription predicts genome-wide patterns of mRNA splicing," *Molecular Cell*, vol. 24, no. 6, pp. 903–915, 2006.

[29] D. F. Tardiff, S. A. Lacadie, and M. Rosbash, "A genome-wide analysis indicates that yeast pre-mRNA splicing is predominantly posttranscriptional," *Molecular Cell*, vol. 24, no. 6, pp. 917–929, 2006.

[30] M. F. LeMaire and C. S. Thummel, "Splicing precedes polyadenylation during Drosophila E74A transcription," *Molecular and Cellular Biology*, vol. 10, no. 11, pp. 6059–6063, 1990.

[31] O. Kessler, Y. Jiang, and L. A. Chasin, "Order of intron removal during splicing of endogenous adenine phosphoribosyltransferase and dihydrofolate reductase pre-mRNA," *Molecular and Cellular Biology*, vol. 13, no. 10, pp. 6211–6222, 1993.

[32] C. Attanasio, A. David, and M. Neerman-Arbez, "Outcome of donor splice site mutations accounting for congenital afibrinogenemia reflects order of intron removal in the fibrinogen alpha gene (FGA)," *Blood*, vol. 101, no. 5, pp. 1851–1856, 2003.

[33] 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.

[34] 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. 357–360, 1997.

[35] Y. Hirose and J. L. Manley, "RNA polymerase II is an essential mRNA polyadenylation factor," *Nature*, vol. 395, no. 6697, pp. 93–96, 1998.

[36] T. Maniatis and R. Reed, "An extensive network of coupling among gene expression machines," *Nature*, vol. 416, no. 6880, pp. 499–506, 2002.

[37] M. J. Moore and N. J. Proudfoot, "Pre-mRNA processing reaches back totranscription and ahead to translation," *Cell*, vol. 136, no. 4, pp. 688–700, 2009.

[38] M. de la Matta, C. Lafaille, and A. R. Kornblihtt, "First come, first served revisited: factors affecting the same alternative splicing event have different effects on the relative rates of intron removal," *RNA*, vol. 16, no. 5, pp. 904–912, 2010.

[39] A. Furger, J. M. O'Sullivan, A. Binnie, B. A. Lee, and N. J. Proudfoot, "Promoter proximal splice sites enhance transcription," *Genes and Development*, vol. 16, no. 21, pp. 2792–2799, 2002.

[40] J. P. Fededa and A. R. Kornblihtt, "A splicing regulator promotes transcriptional elongation," *Nature Structural and Molecular Biology*, vol. 15, no. 8, pp. 779–781, 2008.

[41] Y. W. Fong and Q. Zhou, "Stimulatory effect of splicing factors on transcriptional elongation," *Nature*, vol. 414, no. 6866, pp. 929–933, 2001.

[42] S. Lin, G. Coutinho-Mansfield, D. Wang, S. Pandit, and X. D. Fu, "The splicing factor SC35 has an active role in transcriptional elongation," *Nature Structural and Molecular Biology*, vol. 15, no. 8, pp. 819–826, 2008.

[43] K. Y. Kwek, S. Murphy, A. Furger et al., "U1 snRNA associates with TFIIH and regulates transcriptional initiation," *Nature Structural Biology*, vol. 9, no. 11, pp. 800–805, 2002.

[44] C. K. Damgaard, S. Kahns, S. Lykke-Andersen, A. L. Nielsen, T. H. Jensen, and J. Kjems, "A 5′ splice site enhances the recruitment of basal transcription initiation factors in vivo," *Molecular Cell*, vol. 29, no. 2, pp. 271–278, 2008.

[45] R. D. Alexander, S. A. Innocente, J. D. Barrass, and J. D. Beggs, "Splicing-dependent RNA polymerase pausing in yeast," *Molecular Cell*, vol. 40, no. 4, pp. 582–593, 2010.

[46] F. Carrillo Oesterreich, S. Preibisch, and K. M. Neugebauer, "Global analysis of nascent rna reveals transcriptional pausing in terminal exons," *Molecular Cell*, vol. 40, no. 4, pp. 571–581, 2010.

[47] Y. Brody, N. Neufeld, N. Bieberstein et al., "The in vivo kinetics of RNA polymerase II elongation during co-transcriptional splicing," *PLoS Biology*, vol. 9, no. 1, Article ID e1000573, 2011.

[48] J. Y. Ip, D. Schmidt, Q. Pan et al., "Global impact of RNA polymerase II elongation inhibition on alternative splicing regulation," *Genome Research*, vol. 21, no. 3, pp. 390–401, 2011.

[49] G. E. May, O. Sara, C. J. McManus, and B. R. Graveley, "Competing RNA secondary structures are required for mutually exclusive splicing of the Dscam exon 6 cluster," *RNA*, vol. 17, no. 2, pp. 222–229, 2011.

[50] J. T. Witten and J. Ule, "Understanding splicing regulation through RNA splicing maps," *Trends in Genetics*, vol. 27, no. 3, pp. 89–97, 2011.

[51] Y. Yang, L. Zhan, W. Zhang et al., "RNA secondary structure in mutually exclusive splicing," *Nature Structural and Molecular Biology*, vol. 18, pp. 159–168, 2011.

[52] P. Cramer, C. G. Pesce, F. E. Baralle, and A. R. Kornblihtt, "Functional association between promoter structure and transcript alternative splicing," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 21, pp. 11456–11460, 1997.

[53] P. Cramer, J. F. Cáceres, D. Cazalla et al., "Coupling of transcription with alternative splicing: RNA pol II promoters modulate SF2/ASF and 9G8 effects on an exonic splicing enhancer," *Molecular Cell*, vol. 4, no. 2, pp. 251–258, 1999.

[54] D. Auboeuf, A. Hönig, S. M. Berget, and B. W. O'Malley, "Coordinate regulation of transcription and splicing by steroid receptor coregulators," *Science*, vol. 298, no. 5592, pp. 416–419, 2002.

[55] F. Pagani, C. Stuani, E. Zuccato, A. R. Kornblihtt, and F. E. Baralle, "Promoter architecture modulates CFTR exon 9 skipping," *Journal of Biological Chemistry*, vol. 278, no. 3, pp. 1511–1517, 2003.

[56] N. D. Robson-Dixon and M. A. Garcia-Blanco, "MAZ elements alter transcription elongation and silencing of the fibroblast growth factor receptor 2 exon IIIb," *Journal of Biological Chemistry*, vol. 279, no. 28, pp. 29075–29084, 2004.

[57] G. Adami and L. E. Babiss, "DNA template effect on RNA splicing: two copies of the same gene in the same nucleus are processed differently," *The EMBO Journal*, vol. 10, no. 11, pp. 3457–3465, 1991.

[58] G. Nogués, S. Kadener, P. Cramer, D. Bentley, and A. R. Kornblihtt, "Transcriptional activators differ in their abilities to control alternative splicing," *Journal of Biological Chemistry*, vol. 277, no. 45, pp. 43110–43114, 2002.

[59] D. Auboeuf, D. H. Dowhan, X. Li et al., "CoAA, a nuclear receptor coactivator protein at the interface of transcriptional coactivation and RNA splicing," *Molecular and Cellular Biology*, vol. 24, no. 1, pp. 442–453, 2004.

[60] G. C. Roberts, C. Gooding, H. Y. Mak, N. J. Proudfoot, and C. W. J. Smith, "Co-transcriptional commitment to alternative splice site selection," *Nucleic Acids Research*, vol. 26, no. 24, pp. 5568–5572, 1998.

[61] 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.

[62] 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.

[63] D. L. Bentley, "Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors," *Current Opinion in Cell Biology*, vol. 17, no. 3, pp. 251–256, 2005.

[64] 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.

[65] M. J. Muñoz, 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.

[66] T. Misteli and D. L. Spector, "RNA polymerase II targets pre-mRNA splicing factors to transcription sites in vivo," *Molecular Cell*, vol. 3, no. 6, pp. 697–705, 1999.

[67] A. R. Kornblihtt, "Coupling transcription and alternative splicing," *Advances in Experimental Medicine and Biology*, vol. 623, pp. 175–189, 2007.

[68] N. J. Fuda, M. B. Ardehali, and J. T. Lis, "Defining mechanisms that regulate RNA polymerase II transcription in vivo," *Nature*, vol. 461, no. 7261, pp. 186–192, 2009.

[69] L. P. Eperon, I. R. Graham, A. D. Griffiths, and I. C. Eperon, "Effects of RNA secondary structure on alternative splicing of Pre-mRNA: is folding limited to a region behind the transcribing RNA polymerase?" *Cell*, vol. 54, no. 3, pp. 393–401, 1988.

[70] 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," *The EMBO Journal*, vol. 20, no. 20, pp. 5759–5768, 2001.

[71] S. Kadener, J. P. Fededa, M. Rosbash, and A. R. Kornblihtt, "Regulation of alternative splicing by a transcriptional enhancer through RNA pol II elongation," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 12, pp. 8185–8190, 2002.

[72] K. J. Howe, C. M. Kane, and M. Ares, "Perturbation of transcription elongation influences the fidelity of internal exon inclusion in Saccharomyces cerevisiae," *RNA*, vol. 9, no. 8, pp. 993–1006, 2003.

[73] M. J. Muñoz, M. S. P. Santangelo, M. P. Paronetto et al., "DNA damage regulates alternative splicing through inhibition of RNA polymerase II elongation," *Cell*, vol. 137, no. 4, pp. 708–720, 2009.

[74] M. Aebi and C. Weissman, "Precision and orderliness in splicing," *Trends in Genetics*, vol. 3, no. 4, pp. 102–107, 1987.

[75] 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.

[76] 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.

[77] R. Andersson, S. Enroth, A. Rada-Iglesias, C. Wadelius, and J. Komorowski, "Nucleosomes are well positioned in exons and carry characteristic histone modifications," *Genome Research*, vol. 19, no. 10, pp. 1732–1741, 2009.

[78] P. Kolasinska-Zwierz, T. Down, I. Latorre, T. Liu, X. S. Liu, and J. Ahringer, "Differential chromatin marking of introns and expressed exons by H3K36me3," *Nature Genetics*, vol. 41, no. 3, pp. 376–381, 2009.

[79] S. Nahkuri, R. J. Taft, and J. S. Mattick, "Nucleosomes are preferentially positioned at exons in somatic and sperm cells," *Cell Cycle*, vol. 8, no. 20, pp. 3420–3424, 2009.

[80] S. Schwartz, E. Meshorer, and G. Ast, "Chromatin organization marks exon-intron structure," *Nature Structural and Molecular Biology*, vol. 16, no. 9, pp. 990–995, 2009.

[81] N. Spies, C. B. Nielsen, R. A. Padgett, and C. B. Burge, "Biased chromatin signatures around polyadenylation sites and exons," *Molecular Cell*, vol. 36, no. 2, pp. 245–254, 2009.

[82] H. Tilgner, C. Nikolaou, S. Althammer et al., "Nucleosome positioning as a determinant of exon recognition," *Nature Structural and Molecular Biology*, vol. 16, no. 9, pp. 996–1001, 2009.

[83] R. K. Chodavarapu, S. Feng, Y. V. Bernatavichute et al., "Relationship between nucleosome positioning and DNA methylation," *Nature*, vol. 466, no. 7304, pp. 388–392, 2010.

[84] P. Dhami, P. Saffrey, A. W. Bruce et al., "Complex exon-intron marking by histone modifications is not determined solely by nucleosome distribution," *PLoS One*, vol. 5, no. 8, Article ID e12339, 2010.

[85] M. K. Sakharkar, V. T. K. Chow, and P. Kangueane, "Distributions of exons and introns in the human genome," *In Silico Biology*, vol. 4, no. 4, pp. 387–393, 2004.

[86] K. L. Fox-Walsh, Y. Dou, B. J. Lam, S. P. Hung, P. F. Baldi, and K. J. Hertel, "The architecture of pre-mRNAs affects mechanisms of splice-site pairing," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 45, pp. 16176–16181, 2005.

[87] S. M. Berget, "Exon recognition in vertebrate splicing," *Journal of Biological Chemistry*, vol. 270, no. 6, pp. 2411–2414, 1995.

[88] A. S. Brodsky, C. A. Meyer, I. A. Swinburne et al., "Genomic mapping of RNA polymerase II reveals sites of co-transcriptional regulation in human cells," *Genome Biology*, vol. 6, no. 8, p. R64, 2005.

[89] C. Hodges, L. Bintu, L. Lubkowska, M. Kashlev, and C. Bustamante, "Nucleosomal fluctuations govern the transcription dynamics of RNA polymerase II," *Science*, vol. 325, no. 5940, pp. 626–628, 2009.

[90] L. S. Churchman and J. S. Weissman, "Nascent transcript sequencing visualizes transcription at nucleotide resolution," *Nature*, vol. 469, no. 7330, pp. 368–373, 2011.

[91] E. Batsché, M. Yaniv, and C. Muchardt, "The human SWI/SNF subunit Brm is a regulator of alternative splicing," *Nature Structural and Molecular Biology*, vol. 13, no. 1, pp. 22–29, 2006.

[92] M. Alló and A. R. Kornblihtt, "Gene silencing: small RNAs control RNA polymerase II elongation," *Current Biology*, vol. 20, no. 17, pp. R704–R707, 2010.

[93] E. Hodges, A. D. Smith, J. Kendall et al., "High definition profiling of mammalian DNA methylation by array capture and single molecule bisulfite sequencing," *Genome Research*, vol. 19, no. 9, pp. 1593–1605, 2009.

[94] F. C. Oesterreich, N. Bieberstein, and K. M. Neugebauer, "Pause locally, splice globally," *Trends in Cell Biology*, vol. 21, no. 6, pp. 328–335, 2011.

[95] R. J. Sims III, S. Millhouse, C. F. Chen et al., "Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing," *Molecular Cell*, vol. 28, no. 4, pp. 665–676, 2007.

[96] 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.

[97] A. Hanamura, J. F. Cáceres, A. Mayeda, B. R. Franza, and A. R. Krainer, "Regulated tissue-specific expression of antagonistic pre-mRNA splicing factors," *RNA*, vol. 4, no. 4, pp. 430–444, 1998.

[98] 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.

[99] J. S. Lee and A. Shilatifard, "A site to remember: H3K36 methylation a mark for histone deacetylation," *Mutation Research*, vol. 618, no. 1-2, pp. 130–134, 2007.

[100] S. Kishore and S. Stamm, "The snoRNA HBII-52 regulates alternative splicing of the serotonin receptor 2C," *Science*, vol. 311, no. 5758, pp. 230–232, 2006.

[101] A. Khanna and S. Stamm, "Regulation of alternative splicing by short non-coding nuclear RNAs," *RNA Biology*, vol. 7, no. 4, pp. 480–485, 2010.

SAGE-Hindawi Access to Research
Genetics Research International
Volume 2011, Article ID 326286, 7 pages
doi:10.4061/2011/326286

Review Article

Elongating under Stress

Eulàlia de Nadal and Francesc Posas

Cell Signaling Unit, Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra (UPF), C/ Doctor Aiguader 88, 08003 Barcelona, Spain

Correspondence should be addressed to Eulàlia de Nadal, eulalia.nadal@upf.edu and Francesc Posas, francesc.posas@upf.edu

Received 13 June 2011; Accepted 6 July 2011

Academic Editor: Sebastián Chávez

In response to extracellular stimuli, mitogen-activated protein kinases (MAPKs) modulate gene expression to maximize cell survival. Exposure of yeast to high osmolarity results in activation of the p38-related MAPK Hog1, which plays a key role in reprogramming the gene expression pattern required for cell survival upon osmostress. Hog1 not only regulates initiation but also modulates other steps of the transcription process. Recent work indicates that other yeast signalling MAPKs such as Mpk1 modulate transcriptional elongation in response to cell wall stress. Similarly, mammalian MAPKs have also been found associated to coding regions of stress-responsive genes. In this paper, significant progress in MAPK-regulated events that occur during the transcriptional elongation step is summarized, and future directions are discussed. We expect that the principles learned from these studies will provide a new understanding of the regulation of gene expression by signalling kinases.

1. Introduction

Signal transduction pathways allow cells to sense and respond to extracellular stimuli. MAPK modules are conserved signalling elements utilized in many intracellular signal transduction pathways in eukaryotic organisms [1]. Each MAPK module is activated by specific types of stimuli to induce specific adaptive response. The central core of MAPK systems consists of a tier of kinases, a MAPK kinase kinase (MAPKKK) that phosphorylates and activates a MAPK kinase (MAPKK) on serine and threonine that in turn phosphorylates the MAPK on a threonine (sometimes serine) and tyrosine residue. In yeast *Saccharomyces cerevisiae,* there are several MAPK cascades (for reviews of function and regulation in yeast MAPK signalling, see [2–8]).

The HOG (high osmolarity glycerol) pathway is activated in response to osmostress by two upstream independent branches that converge on the MAPKK Pbs2, which controls Hog1 MAPK activity (see Figure 1) [2, 9–11]. Upon stress, Hog1 translocated into the nucleus [12, 13]. The CWI (cell wall integrity) pathway, which is comprised of Bck1, Mkk1/2, and the Mpk1/Slt2 (and its pseudokinase paralogue Mlp1) MAPK, becomes activated under a number of different conditions that compromise the structure and function of the yeast cell wall, including elevated growth temperature, pheromone-induced morphogenesis, and chemical cell wall antagonist (see Figure 1) [5, 14]. MAPK pathways are known to be conserved during the evolution of the entire eukaryotic kingdom. The Hog1 functional ortholog in mammalian cells is the p38 family of stress-activated MAPKs (SAPKs) and responds to several stresses [15, 16]. Mpk1 is a functional homolog of human Erk5 (for extracellular signal-regulated kinase), which is activated in response to growth factors as well as various stresses [17].

In yeast, these two MAPKs play an important role in controlling gene expression, and they both modulate transcription initiation and elongation steps of the transcription cycle. We will focus here on the role of Hog1 and Mpk1 in regulation of transcription elongation.

2. An Overview of Regulation of Transcriptional Initiation by Hog1 and Mpk1

Among other functions, Hog1 is a master regulator of reprogramming gene expression in response to osmostress. Upon stress, the yeast genome alters its expression pattern up to 20% depending on the strength and duration of the

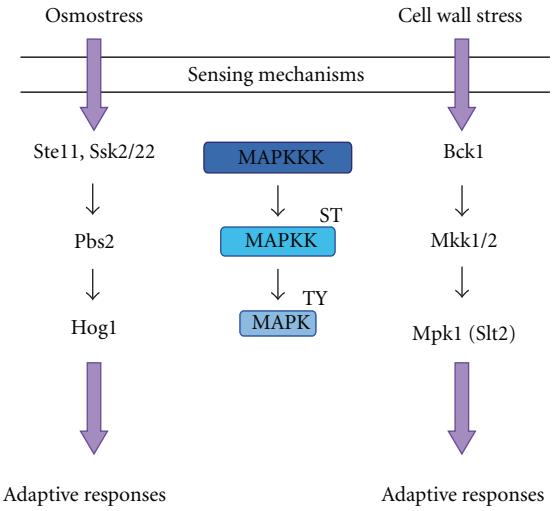


Figure 1: The HOG and CWI pathways. CWI signalling pathway is initiated at the plasma membrane through different sensing mechanisms and is activated by Pkc1 upon different cell wall stresses. The linear cascade consists in the Bck1 MAPKKK, which activates a pair of redundant MAPKK (Mkk1 and Mkk2) that in turn activates the Mpk1/Stl2 MAPK. In the HOG pathway, two independent upstream osmosensing mechanisms lead to the activation of the MAPKKK Ste11 and Ssk2/22. The Pbs2 MAPKK integrates both signals and activates the Hog1 MAPK. Both the HOG and CWI pathways are involved in the regulation of transcriptional elongation by specific types of stimuli to induce specific adaptive response.

stress [18–24]. A major part of these changes are regulated by Hog1 through several unrelated transcription factors such as Msn2/4, Hot1, Smp1, or Sko1, which work in combination at the specific stress-dependent promoters [19, 21, 25–28]. Recently, it has been reported the dynamics of binding of these transcription factors to their target genes in response to osmostress [19, 25]. The integration of this analysis with gene expression patterns reveals a complex dynamic and hierarchical network in which specific combinations of transcription factors activate distinct sets of genes at discrete times to coordinate a rapid and transient stress-adaptive response [25]. It is well known that, when transcription is initiated in response to osmostress, the MAPK is recruited to the osmoresponsive genes by specific transcription factors [26, 27, 29–31] and directly phosphorylates some of them [32–34]. Once bound to chromatin, Hog1 serves as a platform to recruit the RNA Pol II [32] and associated general transcription factors such as the Mediator or SAGA [30, 35] as well as histone-modifying factors [36]. It is worth noting that Hog1 is not the unique kinase that binds to chromatin. Actually, most MAPKs in yeast associate with genes that are their targets of transcriptional control [27, 37]. Further details in the regulation of transcriptional initiation by MAPKs and their implications for understanding control of gene expression are described in [38–40].

In response to cell wall stress, two known transcription factors, Rlm1 and SBF (Swi4 and Swi6), are controlled by Mpk1 by different mechanisms. Whereas Rlm1 is activated through direct phosphorylation [41–44], SBF is activated by a noncatalytic mechanism [37, 45–47]. Swi4 forms a complex with Mpk1 upon stress, and it associates with SBF-binding sites in the promoters of cell wall stress target genes [48]. Moreover, Mpk1 regulates Swi6 nucleocytoplasmic shuttling in a biphasic manner: first, formation of the Mpk1-Swi4 complex recruits Swi6 to the nucleus for transcriptional activation and, second, Mpk1 negatively regulates Swi6 by phosphorylation, which inhibits nuclear entry [49, 50].

3. The Hog1 MAPK Regulates Transcriptional Elongation

The transcription cycle consists of several steps, and elongation is a critical phase of transcription susceptible of strong regulation [51–53]. In addition to its association to promoters, Hog1 is also present at coding regions of stress-responsive genes, suggesting to have a more general role as chromatin-associated enzyme than previously expected. Actually, genome-wide chromatin binding of the MAPK has revealed that Hog1 is recruited to most of the transcribed regions of osmoinducible genes [26, 27]. Moreover, it is recruited to transcribed regions independently of the promoter bound-specific transcription factors dedicated to osmostress adaptation. By uncoupling Hog1-dependent transcription initiation from transcription elongation, it has been demonstrated that binding of Hog1 to stress-responsive coding regions depends on the 3′UTR regions. However, how SAPK is recruited to these specific 3′regions in response to stress remains so far unknown [31].

Which are the tasks of Hog1 at coding regions is still an open question. Recruitment of the kinase to the open reading frames (ORFs) is essential for an increased association of RNA Pol II and proper mRNA production in response to osmostress [31]. Moreover, Hog1 interacts with elongating RNA Pol II (phosphorylated at serine 2 and 5 of the C-terminal domain) as well as with general components of the transcription elongation complex. It is worth noting that the catalytic activity of Hog1 is required both for its binding to chromatin and to stimulate mRNA production during the elongation process. However, the identification of phosphorylation events mediated by the MAPK during transcription elongation remains open.

Binding of Hog1 to the stress-responsive ORFs is restricted temporally. Although the initial recruitment of the MAPK and RNA Pol II is similar, Pol II association is observed for a longer period upon osmostress, whereas binding of Hog1 is restricted at the initial phase of elongation [31]. This suggests a role for the MAPK at early stages of the elongation process. Chromatin structure is tightly regulated through multiple mechanisms, including chromatin remodelling, histone variant incorporation, histone eviction, and histone modification [54–56]. Actually, several genome-wide studies found a significant loss of histones from the promoter and coding regions of heavily transcribed genes throughout the genome [57].

Nucleosome positioning of stress-responsive loci is altered dramatically in a Hog1 MAPK-dependent manner during osmostress [58]. Hog1 physically interacts with the RSC chromatin remodelling complex to direct its association

with the coding regions of osmoresponsive genes and allow for nucleosome rearrangements during transcriptional elongation upon stress. In the RSC mutants, RNA Pol II accumulates on stress-dependent promoters but not in coding regions. Moreover, the RSC mutants display reduced stress gene expression and enhanced sensitivity to osmostress [58]. Other chromatin remodelling complexes such as INO80 associated with the ORFs of stress genes in a stress-specific way [59, 60]. Mutants defective in subunits of the INO80 complex, as well as in several histone chaperone systems, lead to globally increased transcript levels upon osmostress and delayed repositioning of histones in ORF regions of stress genes. Thus, it seems that INO80 is relevant for the efficient downregulation of stress genes under acute stress conditions.

Single-cell experiments have shown that Hog1 nuclear accumulation increases linearly with stimulus. However, at low stress levels, the transcriptional output shows two distinct subpopulations, one responding and the other one not. Of note, this bimodality is reflected in chromatin remodelling and depends on both the retention time and concentration of Hog1 in the nucleus [61]. Thus, chromatin dynamics, together with transient MAPK activation, determines a transcriptional threshold in response to linear increase in signalling upon stress.

4. The Mpk1 MAPK Serves in Transcription Elongation

The role of MAPK in the modulation of transcriptional elongation is not restricted to Hog1. Other yeast signalling kinases, such as Fus3, PKA, or Mpk1, have been reported to associate to coding regions of activated genes [62]. Mpk1 associates with the coding region of the *FKS2* gene in response to cell wall stress although such binding does not require MAPK kinase activity [49]. This diverges from Hog1, where catalytic activity is essential for kinase recruitment.

How does Mpk1 associate with the *FKS2* coding region during transcription? Mpk1 is tethered to the elongation complex through its interaction with RNA Pol II-associated complex Paf1C complex [49, 63, 64]. Paf1 subunit interacts directly with Mpk1 through its docking motif (D-motif) in a cell wall stress-dependent manner. Interaction between both proteins requires the Swi4 and Swi6 but not Rlm1 transcription initiation factor. The *paf1-4A* mutant, which is unable to interact with Mpk1, is deficient for transcription elongation of the *FKS2* gene and renders cells sensitive to cell wall stress. Actually, in this mutant, Mpk1 is recruited to the promoter but does not progress into the coding region of *FKS2* in response to stress [49]. This is again a different scenario when compared to Hog1 and osmostress, since Mpk1 is moving from the initiation to the elongation complexes using the Paf1 complex as a scaffold. Of note, Paf1 and Hog1 are able to coimmunoprecipitate in response to osmostress [31] if such interaction is through the D-motif in Paf1 remains to be elucidated. A large number of short sense transcripts across the *FKS2* promoter-proximal region terminate under noninducing conditions [65, 66]. The critical function of the Mpk1-Paf1 association in transcription elongation is to prevent such premature termination under inducing conditions. Indeed, Mpk1-Paf1 interaction blocks recruitment of the Sen1-Nrd1-Nab3 termination complex to allow effective elongation of cell-wall stress genes [49]. Therefore, it is becoming increasingly apparent that yeast MAPKs play a key role in regulation of transcriptional elongation in response to cellular stress although the molecular mechanisms involved differ among kinases.

5. MAPK Signalling and Transcriptional Elongation in Higher Eukaryotes

Binding of signalling kinases to chromatin has been now shown in organisms other than yeast. Several reports support an essential role of p38 MAPK in the regulation of transcription upon inflammation and stress responses [67, 68] as well as during cell growth and differentiation [69–71]. In response to stress, p38 associates to chromatin as Hog1 does in yeast and allows for recruitment of RNA Pol II and transcriptional initiation. Similarly, anchoring of active p38 to target stress-dependent promoters is mediated by specific transcription factors [72]. Moreover, p38 is also present at coding regions depending on its activity upon stress, clearly suggesting that it might be travelling along with the nascent mRNA elongating machinery in a similar way as described in yeast [72]. Actually, it is described that the MAPK interacts with the RNA Pol II in mammalian cells [32].

p38 controls skeletal muscle differentiation by regulating the sequential activation of myogenic regulatory factors and their transcriptional coactivators, including chromatin remodelling enzymes (reviewed in [73]). However, whether the MAPK has a specific role during transcriptional elongation in muscle differentiation remains to be determined.

The extracellular signal-regulated kinase (ERK) pathway also regulates gene expression. Erk1 is activated by progesterone and phosphorylates the progesterone receptor. Then, a complex of activated progesterone receptor, Erk1, and its target kinase Msk1 is recruited to the target promoters, where Msk1 phosphorylates histone H3 at serine 10 promoting chromatin remodelling and gene regulation (reviewed in [74, 75]). Once more, the specific role of the Erk1 kinase in transcriptional elongation in response to progestins remains to be elucidated.

When expressed in yeast, the human Erk5 MAPK is activated in response to cell wall stress and suppresses the phenotypic defects of a *mpk1* mutant. Moreover, Erk5 is able to activate gene expression through Rlm1 as well as SBF transcription factors [17, 37]. There are evidences suggesting that MAPK regulation of Paf1C function is conserved in humans. Actually, Erk5 can interact with a predicted D-motif in the human PD2/PAF1 within the same region as that found in yeast Paf1 to drive transcription elongation [47].

Therefore, it is likely that MAPK regulation of transcription in higher eukaryotes is not restricted to initiation but also to elongation upon stress and that MAPK-driven mechanisms are conserved among eukaryotic cells.

6. Conclusions

Proper adaptation to stress is critical for cell survival. There has been a number of signalling networks involved in stress signal transduction, and, amongst them, MAPK signalling networks stand out. One important piece of the different adaptive strategies consists of a massive reorganization of the gene expression capacity. Protein kinases not only phosphorylate target proteins, but also themselves become part of the transcriptional complex. This suggests a new scenario in which signalling kinases, rather than simply relaying the signal to the transcriptional effectors, may function as integral components of the transcriptional complexes that activate expression of distinct target genes (see Figure 2). Actually, several recent reports have shown that the role of MAPK in the regulation of the transcription cycle is not limited to transcription initiation but rather extends to the process of transcriptional elongation. For instance, the Hog1 MAPK fits with the description of a bona fide elongation factor, with the feature that its role in elongation is restricted to osmoresponsive genes, and it has a pivotal role in nucleosome remodelling upon stress. Moreover, other MAPKs such as Mpk1 block premature transcription termination of stress-induced genes. Further studies are required to establish whether the mechanisms by which Hog1 and Mpk1 regulate transcriptional elongation are conserved in both MAPKs.

We are just beginning to understand the overlapping link between signal transduction pathways and modulation of transcriptional elongation. The molecular mechanisms of such relationship are far from fully being understood. For example, one unresolved question is how the MAPK Hog1 is recruited specifically to the osmo-responsive coding regions. Is there some special feature in the 3′UTR region of stress-dependent genes targeting the kinase? or is there some transcription elongation factor that specifically recruits the kinase? Another key question remaining in the field is which are the phosphorylation events mediated by MAPK on the transcriptional elongation machinery. Although it is clear that there are noncatalytic functions in transcription regulation by MAPK, activity of Hog1 is needed for the recruitment of the kinase onto the promoters and coding regions of osmostress-responsive genes. Recent findings shed light on the link between upstream signalling kinases and direct phosphorylation of histones and/or histone modifiers (reviewed in [76]). Moreover, transcriptional stress responses and the chromatin structure are tightly linked [77, 78]. Why regulation of Paf1 by Mpk1 only affects those cell wall stress-activated genes that are under the control of the noncatalytic Mpk1 pathway is another question that remains unsolved. And in these lines, are the Rlm1-dependent genes regulated by the MAPK during transcriptional elongation?

Last but not least, to what extent regulation of transcriptional elongation will apply to other yeast and mammalian signalling modules will be interesting avenues for future investigation. In any case, it is becoming clear that mechanisms, which regulate transcription, have been preserved from yeast to mammals.

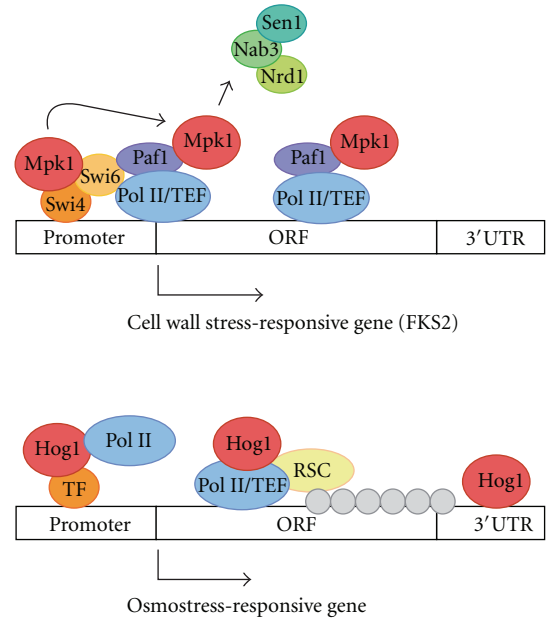


Figure 2: Regulation of transcriptional elongation by MAPK signalling pathways. Mpk1 and Hog1 MAPKs are key regulators of transcription in response to specific stresses. Whereas activated Mpk1 is moving from initiation to elongation complexes through its interaction with Paf1C complex, binding of Hog1 to the ORFs seems to be dependent on the 3′UTR region of the osmoresponsive gene. Association of Mpk1 with Paf1 serves as an antitermination factor by blocking recruitment of the Sen1-Nrd1-Nab3 termination complex to the cell wall stress-responsive gene. Hog1 acts as a stress-specific transcription elongation factor (TEF) and targets selectively the RSC complex to osmoinducible ORFs, which displaces nucleosomes contributing to the efficient activation of transcription.

Acknowledgments

The authors thank all the members of the Cell Signaling Unit. They apologize to colleagues whose work could not be cited due to size limitation. The laboratory of F. Posas and E. de Nadal is being supported by Fundación Marcelino Botín (FMB), Ministerio de Ciéncia y Innovación, BFU program (BFU2008-00530 to E. de Nadal and BIO2009-07762 to F. Posas), Consolider Ingenio 2010 programme (Grant no. CSD2007-0015), and FP7 (*UNICELLSYS*) framework programs. F. Posas is recipient of an ICREA Acadèmia (Generalitat de Catalunya).

References

[1] C. Widmann, S. Gibson, M. B. Jarpe, and G. L. Johnson, "Mitogen-activated protein kinase: conservation of a three-kinase module from yeast to human," *Physiological Reviews*, vol. 79, no. 1, pp. 143–180, 1999.

[2] R. E. Chen and J. Thorner, "Function and regulation in MAPK signaling pathways: lessons learned from the yeast Saccharomyces cerevisiae," *Biochimica et Biophysica Acta*, vol. 1773, no. 8, pp. 1311–1340, 2007.

[3] E. de Nadal, P. M. Alepuz, and F. Posas, "Dealing with osmostress through MAP kinase activation," *The EMBO Reports*, vol. 3, no. 8, pp. 735–740, 2002.

[4] S. Hohmann, "Osmotic stress signaling and osmoadaptation in yeasts," *Microbiology and Molecular Biology Reviews*, vol. 66, no. 2, pp. 300–372, 2002.

[5] D. E. Levin, "Cell wall integrity signaling in Saccharomyces cerevisiae," *Microbiology and Molecular Biology Reviews*, vol. 69, no. 2, pp. 262–291, 2005.

[6] S. M. O'Rourke, I. Herskowitz, and E. K. O'Shea, "Yeast go the whole HOG for the hyperosmotic response," *Trends in Genetics*, vol. 18, no. 8, pp. 405–412, 2002.

[7] H. Saito and K. Tatebayashi, "Regulation of the osmoregulatory HOG MAPK cascade in yeast," *Journal of Biochemistry*, vol. 136, no. 3, pp. 267–272, 2004.

[8] P. J. Westfall, D. R. Ballon, and J. Thorner, "When the stress of your environment makes you go HOG wild," *Science*, vol. 306, no. 5701, pp. 1511–1512, 2004.

[9] S. Hohmann, M. Krantz, and B. Nordlander, "Yeast osmoregulation," *Methods in Enzymology*, vol. 428, pp. 29–45, 2007.

[10] E. de Nadal, F. X. Real, and F. Posas, "Mucins, osmosensors in eukaryotic cells?" *Trends in Cell Biology*, vol. 17, no. 12, pp. 571–574, 2007.

[11] H. Saito, "Histidine phosphorylation and two-component signaling in eukaryotic cells," *Chemical Reviews*, vol. 101, no. 8, pp. 2497–2509, 2001.

[12] P. Ferrigno, F. Posas, D. Koepp, H. Saito, and P. A. Silver, "Regulated nucleo/cytoplasmic exchange of HOG1 MAPK requires the importin beta homologs NMD5 and XPO1," *The EMBO Journal*, vol. 17, no. 19, pp. 5606–5614, 1998.

[13] V. Reiser, H. Ruis, and G. Ammerer, "Kinase activity-dependent nuclear export opposes stress-induced nuclear accumulation and retention of Hog1 mitogen-activated protein kinase in the budding yeast Saccharomyces cerevisiae," *Molecular Biology of the Cell*, vol. 10, no. 4, pp. 1147–1161, 1999.

[14] J. C. Harrison, T. R. Zyla, E. S. Bardes, and D. J. Lew, "Stress-specific activation mechanisms for the "cell integrity" MAPK pathway," *Journal of Biological Chemistry*, vol. 279, no. 4, pp. 2616–2622, 2004.

[15] J. Han, J. D. Lee, L. Bibbs, and R. J. Ulevitch, "A MAP kinase targeted by endotoxin and hyperosmolarity in mammalian cells," *Science*, vol. 265, no. 5173, pp. 808–811, 1994.

[16] D. Sheikh-Hamad and M. C. Gustin, "MAP kinases and the adaptive response to hypertonicity: functional preservation from yeast to mammals," *The American Journal of Physiology—Renal Physiology*, vol. 287, no. 6, pp. F1102–F1110, 2004.

[17] A. W. Truman, S. H. Millson, J. M. Nuttall et al., "Expressed in the yeast Saccharomyces cerevisiae, human ERK5 is a client of the Hsp90 chaperone that complements loss of the Slt2p (Mpk1p) cell integrity stress-activated protein kinase," *Eukaryotic Cell*, vol. 5, no. 11, pp. 1914–1924, 2006.

[18] A. P. Gasch, P. T. Spellman, C. M. Kao et al., "Genomic expression programs in the response of yeast cells to environmental changes," *Molecular Biology of the Cell*, vol. 11, no. 12, pp. 4241–4257, 2000.

[19] A. P. Capaldi, T. Kaplan, Y. Liu et al., "Structure and function of a transcriptional network activated by the MAPK Hog1," *Nature Genetics*, vol. 40, no. 11, pp. 1300–1306, 2008.

[20] H. C. Causton, B. Ren, S. S. Koh et al., "Remodeling of yeast genome expression in response to environmental changes," *Molecular Biology of the Cell*, vol. 12, no. 2, pp. 323–337, 2001.

[21] C. Miller, B. Schwalb, K. Maier et al., "Dynamic transcriptome analysis measures rates of mRNA synthesis and decay in yeast," *Molecular Systems Biology*, vol. 7, p. 458, 2011.

[22] F. Posas, J. R. Chamber, J. A. Heyman, J. P. Hoeffler, E. de Nadal, and J. Arino, "The transcriptional response of yeast to saline stress," *Journal of Biological Chemistry*, vol. 275, no. 23, pp. 17249–17255, 2000.

[23] M. Rep, M. Krantz, J. M. Thevelein, and S. Hohmann, "The transcriptional response of Saccharomyces cerevisiae to osmotic shock. Hot1p and Msn2p/Msn4p are required for the induction of subsets of high osmolarity glycerol pathway-dependent genes," *Journal of Biological Chemistry*, vol. 275, no. 12, pp. 8290–8300, 2000.

[24] J. Yale and H. J. Bohnert, "Transcript expression in Saccharomyces cerevisiae at high salinity," *Journal of Biological Chemistry*, vol. 276, no. 19, pp. 15996–16007, 2001.

[25] L. Ni, C. Bruce, C. Hart et al., "Dynamic and complex transcription factor binding during an inducible response in yeast," *Genes and Development*, vol. 23, no. 11, pp. 1351–1363, 2009.

[26] A. Pascual-Ahuir, K. Struhl, and M. Proft, "Genome-wide location analysis of the stress-activated MAP kinase Hog1 in yeast," *Methods*, vol. 40, no. 3, pp. 272–278, 2006.

[27] D. K. Pokholok, J. Zeitlinger, N. M. Hannett, D. B. Reynolds, and R. A. Young, "Activated signal transduction kinases frequently occupy target genes," *Science*, vol. 313, no. 5786, pp. 533–536, 2006.

[28] M. Proft, F. D. Gibbons, M. Copeland, F. P. Roth, and K. Struhl, "Genomewide identification of Sko1 target promoters reveals a regulatory network that operates in response to osmotic stress in Saccharomyces cerevisiae," *Eukaryotic Cell*, vol. 4, no. 8, pp. 1343–1352, 2005.

[29] P. M. Alepuz, A. Jovanovic, V. Reiser, and G. Ammerer, "Stress-induced MAP kinase Hog1 is part of transcription activation complexes," *Molecular Cell*, vol. 7, no. 4, pp. 767–777, 2001.

[30] M. Proft and K. Struhl, "Hog1 kinase converts the Sko1-Cyc8-Tup1 repressor complex into an activator that recruits SAGA and SWI/SNF in response to osmotic stress," *Molecular Cell*, vol. 9, no. 6, pp. 1307–1317, 2002.

[31] M. Proft, G. Mas, E. de Nadal et al., "The stress-activated Hog1 kinase is a selective transcriptional elongation factor for genes responding to osmotic stress," *Molecular Cell*, vol. 23, no. 2, pp. 241–250, 2006.

[32] P. M. Alepuz, E. de Nadal, M. Zapater, G. Ammerer, and F. Posas, "Osmostress-induced transcription by Hot1 depends on a Hog1-mediated recruitment of the RNA Pol II," *The EMBO Journal*, vol. 22, no. 10, pp. 2433–2442, 2003.

[33] E. de Nadal, L. Casadome, and F. Posas, "Targeting the MEF2-like transcription factor Smp1 by the stress-activated Hog1 mitogen-activated protein kinase," *Molecular and Cellular Biology*, vol. 23, no. 1, pp. 229–237, 2003.

[34] M. Proft, A. Pascual-Ahuir, E. de Nadal, J. Arino, R. Serrano, and F. Posas, "Regulation of the Sko1 transcriptional repressor by the Hog1 MAP kinase in response to osmotic stress," *The EMBO Journal*, vol. 20, no. 5, pp. 1123–1133, 2001.

[35] M. Zapater, M. Sohrmann, M. Peter, F. Posas, and E. de Nadal, "Selective requirement for SAGA in Hog1-mediated gene expression depending on the severity of the external osmostress conditions," *Molecular and Cellular Biology*, vol. 27, no. 11, pp. 3900–3910, 2007.

[36] E. de Nadal, M. Zapater, P. M. Alepuz, L. Sumoy, G. Mas, and F. Posas, "The MAPK Hog1 recruits Rpd3 histone deacetylase to activate osmoresponsive genes," *Nature*, vol. 427, no. 6972, pp. 370–374, 2004.

[37] K. Y. Kim, A. W. Truman, and D. E. Levin, "Yeast Mpk1 mitogen-activated protein kinase activates transcription through Swi4/Swi6 by a noncatalytic mechanism that requires upstream signal," *Molecular and Cellular Biology*, vol. 28, no. 8, pp. 2579–2589, 2008.

[38] E. de Nadal and F. Posas, "Multilayered control of gene expression by stress-activated protein kinases," *The EMBO Journal*, vol. 29, no. 1, pp. 4–13, 2010.

[39] F. Martínez-Montañés, A. Pascual-Ahuir, and M. Proft, "Toward a genomic view of the gene expression program regulated by osmostress in yeast," *OMICS*, vol. 14, no. 6, pp. 619–627, 2010.

[40] V. M. Weake and J. L. Workman, "Inducible gene expression: diverse regulatory mechanisms," *Nature Reviews Genetics*, vol. 11, no. 6, pp. 426–437, 2010.

[41] E. Dodou and R. Treisman, "The Saccharomyces cerevisiae MADS-box transcription factor Rlm1 is a target for the Mpk1 mitogen-activated protein kinase pathway," *Molecular and Cellular Biology*, vol. 17, no. 4, pp. 1848–1859, 1997.

[42] U. S. Jung and D. E. Levin, "Genome-wide analysis of gene expression regulated by the yeast cell wall integrity signalling pathway," *Molecular Microbiology*, vol. 34, no. 5, pp. 1049–1057, 1999.

[43] U. S. Jung, A. K. Sobering, M. J. Romeo, and D. E. Levin, "Regulation of the yeast Rlm1 transcription factor by the Mpk1 cell wall integrity MAP kinase," *Molecular Microbiology*, vol. 46, no. 3, pp. 781–789, 2002.

[44] Y. Watanabe, G. Takaesu, M. Hagiwara, K. Irie, and K. Matsumoto, "Characterization of a serum response factor-like protein in Saccharomyces cerevisiae, Rlm1, which has transcriptional activity regulated by the Mpk1 (Slt2) mitogen-activated protein kinase pathway," *Molecular and Cellular Biology*, vol. 17, no. 5, pp. 2615–2623, 1997.

[45] K. Baetz, J. Moffat, J. Haynes, M. Chang, and B. Andrews, "Transcriptional coregulation by the cell integrity mitogen-activated protein kinase Slt2 and the cell cycle regulator Swi4," *Molecular and Cellular Biology*, vol. 21, no. 19, pp. 6515–6528, 2001.

[46] J. C. Igual, A. L. Johnson, and L. H. Johnston, "Coordinated regulation of gene expression by the cell cycle transcription factor Swi4 and the protein kinase C MAP kinase pathway for yeast cell integrity," *The EMBO Journal*, vol. 15, no. 18, pp. 5001–5013, 1996.

[47] K. Y. Kim, A. W. Truman, S. Caesar, G. Schlenstedt, and D. E. Levin, "Yeast Mpk1 cell wall integrity mitogen-activated protein kinase regulates nucleocytoplasmic shuttling of the Swi6 transcriptional regulator," *Molecular Biology of the Cell*, vol. 21, no. 9, pp. 1609–1619, 2010.

[48] A. W. Truman, K. Y. Kim, and D. E. Levin, "Mechanism of Mpk1 mitogen-activated protein kinase binding to the Swi4 transcription factor and its regulation by a novel caffeine-induced phosphorylation," *Molecular and Cellular Biology*, vol. 29, no. 24, pp. 6449–6461, 2009.

[49] K. Y. Kim and D. E. Levin, "Mpk1 MAPK association with the Paf1 complex blocks Sen1-mediated premature transcription termination," *Cell*, vol. 144, pp. 745–756, 2011.

[50] K. Madden, Y. J. Sheu, K. Baetz, B. Andrews, and M. Snyder, "SBF cell cycle regulator as a target of the yeast PKC-MAP kinase pathway," *Science*, vol. 275, no. 5307, pp. 1781–1784, 1997.

[51] N. J. Fuda, M. B. Ardehali, and J. T. Lis, "Defining mechanisms that regulate RNA polymerase II transcription in vivo," *Nature*, vol. 461, no. 7261, pp. 186–192, 2009.

[52] A. Saunders, L. J. Core, and J. T. Lis, "Breaking barriers to transcription elongation," *Nature Reviews Molecular Cell Biology*, vol. 7, no. 8, pp. 557–567, 2006.

[53] S. Egloff and S. Murphy, "Cracking the RNA polymerase II CTD code," *Trends in Genetics*, vol. 24, no. 6, pp. 280–288, 2008.

[54] B. R. Cairns, "The logic of chromatin architecture and remodelling at promoters," *Nature*, vol. 461, no. 7261, pp. 193–198, 2009.

[55] C. Jiang and B. F. Pugh, "Nucleosome positioning and gene regulation: advances through genomics," *Nature Reviews Genetics*, vol. 10, no. 3, pp. 161–172, 2009.

[56] J. L. Workman, "Nucleosome displacement in transcription," *Genes and Development*, vol. 20, no. 15, pp. 2009–2017, 2006.

[57] W. Lee, D. Tillo, N. Bray et al., "A high-resolution atlas of nucleosome occupancy in yeast," *Nature Genetics*, vol. 39, no. 10, pp. 1235–1244, 2007.

[58] G. Mas, E. de Nadal, R. Dechant et al., "Recruitment of a chromatin remodelling complex by the Hog1 MAP kinase to stress genes," *The EMBO Journal*, vol. 28, pp. 326–336, 2009.

[59] I. Gorzer, C. Schuller, E. Heidenreich, L. Krupanska, K. Kuchler, and U. Wintersberger, "The nuclear actin-related protein Act3p/Arp4p of Saccharomyces cerevisiae is involved in transcription regulation of stress genes," *Molecular Microbiology*, vol. 50, pp. 1155–1171, 2003.

[60] E. Klopf, L. Paskova, C. Solé et al., "Cooperation between the INO80 complex and histone chaperones determines adaptation of stress gene transcription in the yeast Saccharomyces cerevisiae," *Molecular and Cellular Biology*, vol. 29, no. 18, pp. 4994–5007, 2009.

[61] S. Pelet, F. Rudolf, M. Nadal-Ribelles, E. de Nadal, F. Posas, and M. Peter, "Transient activation of the HOG MAPK pathway regulates bimodal gene expression," *Science*, vol. 332, no. 6030, pp. 732–735, 2011.

[62] C. W. Chow and R. J. Davis, "Proteins kinases: chromatin-associated enzymes?" *Cell*, vol. 127, no. 5, pp. 887–890, 2006.

[63] M. Chang, D. French-Cornay, H. Y. Fan, H. Klein, C. L. Denis, and J. A. Jaehning, "A complex containing RNA polymerase II, Paf1p, Cdc73p, Hpr1p, and Ccr4p plays a role in protein kinase C signaling," *Molecular and Cellular Biology*, vol. 19, no. 2, pp. 1056–1067, 1999.

[64] J. A. Jaehning, "The Paf1 complex: platform or player in RNA polymerase II transcription?" *Biochimica et Biophysica Acta*, vol. 1799, no. 5-6, pp. 379–388, 2010.

[65] F. Miura, N. Kawaguchi, J. Sese et al., "A large-scale full-length cDNA analysis to explore the budding yeast transcriptome," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 47, pp. 17846–17851, 2006.

[66] H. Neil, C. Malabat, Y. D'Aubenton-Carafa, Z. Xu, L. M. Steinmetz, and A. Jacquier, "Widespread bidirectional promoters are the major source of cryptic transcripts in yeast," *Nature*, vol. 457, no. 7232, pp. 1038–1042, 2009.

[67] I. Ferreiro, M. Joaquin, A. Islam et al., "Whole genome analysis of p38 SAPK-mediated gene expression upon stress," *BMC Genomics*, vol. 11, no. 1, article 144, 2010.

[68] D. Viemann, M. Goebeler, S. Schmid et al., "Transcriptional profiling of IKK2/NF-kappa B- and p38 MAP kinase-dependent gene expression in TNF-alpha-stimulated primary human endothelial cells," *Blood*, vol. 103, no. 9, pp. 3365–3373, 2004.

[69] A. Gazel, R. I. Nijhawan, R. Walsh, and M. Blumenberg, "Transcriptional profiling defines the roles of ERK and p38 kinases in epidermal keratinocytes," *Journal of Cellular Physiology*, vol. 215, no. 2, pp. 292–308, 2008.

[70] C. Serra, D. Palacios, C. Mozzetta et al., "Functional interdependence at the chromatin level between the MKK6/p38 and IGF1/PI3K/AKT pathways during muscle differentiation," *Molecular Cell*, vol. 28, no. 2, pp. 200–213, 2007.

[71] C. Simone, S. V. Forcales, D. A. Hill, A. N. Imbalzano, L. Latella, and P. L. Puri, "p38 pathway targets SWI-SNF chromatin-remodeling complex to muscle-specific loci," *Nature Genetics*, vol. 36, no. 7, pp. 738–743, 2004.

[72] I. Ferreiro, M. Barragan, A. Gubern, E. Ballestar, M. Joaquin, and F. Posas, "The p38 SAPK is recruited to chromatin via its interaction with transcription factors," *Journal of Biological Chemistry*, vol. 285, no. 41, pp. 31819–31828, 2010.

[73] F. Lluís, E. Perdiguero, A. R. Nebreda, and P. Muñoz-Cánoves, "Regulation of skeletal muscle gene expression by p38 MAP kinases," *Trends in Cell Biology*, vol. 16, no. 1, pp. 36–44, 2006.

[74] G. P. Vicent, R. Zaurin, C. Ballaré, A. S. Nacht, and M. Beato, "Erk signaling and chromatin remodeling in MMTV promoter activation by progestins," *Nuclear Receptor Signaling*, vol. 7, article e008, 2009.

[75] G. P. Vicent, A. S. Nacht, R. Zaurin, C. Ballare, J. Clausell, and M. Beato, "Minireview: role of kinases and chromatin remodeling in progesterone signaling to chromatin," *Molecular Endocrinology*, vol. 24, pp. 2088–2098, 2010.

[76] S. H. Baek, "When signaling kinases meet histones and histone modifiers in the nucleus," *Molecular Cell*, vol. 42, pp. 274–284, 2011.

[77] S. Shivaswamy, A. Bhinge, Y. Zhao, S. Jones, M. Hirst, and V. R. Iyer, "Dynamic remodeling of individual nucleosomes across a eukaryotic genome in response to transcriptional perturbation.," *PLoS Biology*, vol. 6, no. 3, article e65, 2008.

[78] J. Dai, E. M. Hyland, D. S. Yuan, H. Huang, J. S. Bader, and J. D. Boeke, "Probing nucleosome function: a highly versatile library of synthetic histone H3 and H4 mutants," *Cell*, vol. 134, no. 6, pp. 1066–1078, 2008.