1
Eukaryotic Gene Expression by RNA Polymerase II

Geetha Durairaj,[1,3,] Shivani Malik,[2,3,*] and Sukesh R. Bhaumik[3]*

[1]University of California, Department of Biological Chemistry, 825 Health Sciences Road, 240D Med Sci I, Irvine, CA 92697, USA
[2]University of California, School of Medicine, Department of Medicine, 1450 3rd Street, HD-340, San Francisco, CA 94158, USA
[3]Southern Illinois University School of Medicine, Department of Biochemistry and Molecular Biology, 1245 Lincoln Drive, Neckers Building, Carbondale, IL 62901, USA

1.1
Introduction

Expression of eukaryotic protein coding genes starts with transcription by RNA polymerase II in the nucleus (Figure 1.1). Transcription can be mechanistically divided into different phases, namely, transcriptional initiation, elongation and termination, and produces messenger RNAs (mRNAs). Following transcriptional initiation, nascent mRNA of about 20–24 nucleotides long associates with the cap-binding complex (CBC) at its 5′-end via the formation of the cap structure by capping enzymes. As mRNA synthesis progresses during transcriptional elongation, RNA polymerase II mediates the addition of many mRNA binding proteins that aid in preventing DNA:RNA hybrid formation and facilitating the 3′-end processing, splicing, surveillance, and export of mRNA to the cytoplasm for translation into proteins (Figure 1.1). All these steps are highly coordinated and precisely regulated. In this chapter, we provide broad overviews of the different steps of eukaryotic gene expression by RNA polymerase II.

1.2
Transcriptional Initiation of RNA Polymerase II Genes

Transcriptional initiation of RNA polymerase II genes typically involves an activator that binds to a specific sequence (known as upstream activating sequence

* Geetha Durairaj and Shivani Malik contributed equally to this chapter.

Gene Regulation, Epigenetics, and Hormone Signaling, First Edition. Edited by Subhrangsu S. Mandal.

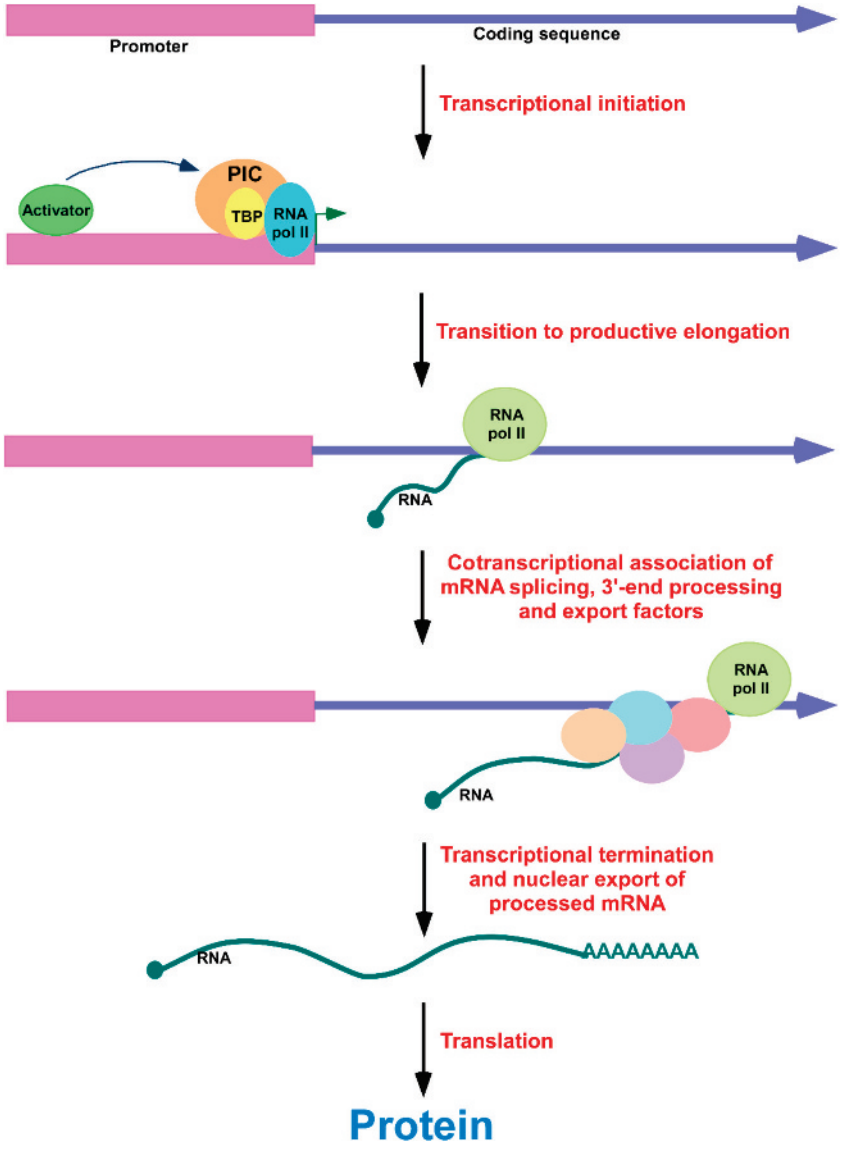

Figure 1.1 Schematic diagrams of different steps of eukaryotic gene expression by RNA polymerase II. PIC, preinitiation complex; and RNA pol II, RNA polymerase II.

or UAS) in the gene promoter via its DNA binding domain. Subsequently, the activation domain (AD) of the activator interacts with one or more transcription factors to facilitate the assembly of general transcription factors (GTFs) such as TFIIA, TFIIB, TFIID, TFIIE, TFIIF, TFIIH, and RNA polymerase II, resulting in the formation of preinitiation complex (PIC) for transcriptional initiation [1,2].

The ADs are generally acidic in nature (e.g., Gal4, Gcn4, VP16, and p53). However, other ADs such as glutamine-rich and proline-rich also exist in stimulation of transcriptional initiation.

A wealth of research in yeast has delineated two broad mechanisms of transcriptional initiation, namely, SAGA (Spt-Ada-Gnc5-Acetyltransferase)-dependent and TFIID (transcription factor IID)-dependent transcriptional initiation. SAGA is a large multisubunit protein complex, and is highly conserved among eukaryotes [2]. In yeast, SAGA comprises of 14 nonessential (Ada1, Ada2, Ada3, Gcn5, Spt3, Spt7, Spt8, Spt20, Sgf11, Sgf29, Sgf73, Ubp8, Sus1, and Chd1) and 6 essential (ATM/PI-3-kinase-related protein Tra1 and a subset of TBP-associated factors or TAFs such as TAF5, TAF6, TAF9, TAF10, and TAF12) subunits [2–4]. The Gcn5 subunit of SAGA has the histone acetyltransferase (HAT) activity, while the Ubp8 subunit possesses histone deubiquitylation or ubiquitin protease activity [2–4]. The Ada1, Spt7, and Spt20 subunits maintain overall structural integrity of SAGA [2,5,6]. The Tra1 subunit of SAGA interacts with activator, and subsequently, facilitates the formation of PIC for transcriptional initiation [2,7]. Biochemical and genetic experiments indicate that the Spt3 and Spt8 subunits of SAGA interact with TBP, and stimulate the PIC formation [2,6,8,9]. In addition to recruiting the PIC components, SAGA modulates chromatin structure through its HAT activity to regulate transcriptional initiation. SAGA acetylates nucleosomal histone H3 (K9, K14, K18, and K23 (K, lysine)) residues [10,11], promotes the formation of TFIIA–TBP–DNA complex, and hence transcriptional initiation. However, SAGA-mediated acetylation may also reduce TBP recruitment at certain genes (e.g., *ARG1*), thus adding to the complexity of gene regulation [12]. Further, SAGA regulates transcription through the histone H2B deubiquitylation activity of its Ubp8 subunit [13–15]. Ubp8 modulates the level of ubiquitylated-histone H2BK123 that is correlated with histone H3K4 methylation, and hence transcription activity of the SAGA-dependent genes [16]. SAGA plays similar functions in *Drosophila* and humans [2–4].

SAGA regulates transcription of about 10% genes in yeast [2]. However, transcription of ~90% genes in yeast is regulated by TFIID, but not SAGA [2]. TFIID is a large multisubunit protein complex that is composed of TBP and a set of TAFs. TAFs are evolutionarily conserved and ubiquitously expressed proteins with the exception of some tissue-specific TAFs in mammals. TAFs have a conserved structural motif, known as histone-fold that is found in the core histones. The histone folds of TAFs mediate subunit interactions within TFIID, and thus, act as building blocks of the TFIID complex [17,18]. Further, TAFs interact directly with activator, and facilitates the recruitment of TBP to the core promoter to enhance the PIC formation for transcriptional initiation [2,19–23].

Both the UAS and core promoter contribute to the TFIID dependency of a gene for transcriptional initiation. Sequence swapping experiments have revealed that the replacement of the UAS of a TFIID-dependent promoter with the UAS of a SAGA-regulated promoter impairs the recruitment of TFIID to the core promoter [2,22]. Consistently, fusion of the UAS of the TFIID-dependent promoter with the core promoter of a SAGA-regulated (or TFIID-independent)

promoter leads to the recruitment of TFIID to the TFIID-independent core promoter [2,22]. These results highlight the importance of UAS in determining the TFIID dependency of a gene, hence indicating the role of the UAS-bound activator in targeting TAFs to the core promoter. Consistently, certain TAFs of TFIID have been shown to interact with activator [19,20]. However, the core promoter may also contribute to this TFIID dependency [2]. It is proposed that a TATA-like element in the core promoter may participate in TFIID binding for transcriptional initiation [24]. TAFs have also been shown to interact with the initiator or downstream promoter element (DPE) in *Drosophila* and humans, thus supporting the role of the core promoter in TFIID dependency [25]. An additional layer of regulation by the core promoter is provided by histone H3K4 methylation [26]. TFIID interacts with histone H3K4 methylation at the promoter. Such interaction is mediated via the recognition of trimethylated-K4 of histone H3 by the plant homeodomain of TAF3 (absent in yeast, but present in higher eukaryotes) of TFIID [2]. Consistently, the loss of histone H3K4 methylation leads to decreased level of TFIID binding to the promoter [26]. Further, there is a cross talk between modifications where histone H3 Arginine 2 methylation inhibits the interaction of TFIID with histone H3K4 methylation, while acetylation of histone H3K9/14 promotes it [2,27]. Thus, there is a combination of interrelated regulatory mechanisms mediated by the initiation proteins, promoter structure, and chromatin structure/covalent modifications that contribute to the precise control of transcriptional initiation at the TFIID-dependent genes.

Both the TFIID- and SAGA-dependent transcriptional initiations are additionally modulated by the 26 S proteasome complex that is involved in targeted protein degradation via ubiquitylation [1,2,28–30]. The 26 S proteasome complex consists of the 19 S RP (regulatory particle) and 20 S CP (core particle). The 20 S CP is involved in proteolytic degradation, while the 19 S RP provides the specificity for ubiquitylated-proteins and ATP-dependence [1]. The 20 S CP is composed of two α-rings and two β-rings that are stacked in the order of α β β α forming the cylinder-like structure. The 19 S RP has at least 17 different proteins including six ATPases (Rpt1–Rpt6). Six AAA-ATPases (Rpt1–Rpt6) and three non-ATPase proteins (Rpn1, Rpn2 and Rpn10) form the base of the 19 S RP, while the lid of the 19 S RP consists of eight non-ATPase proteins (Rpn3, Rpn5–Rpn9, Rpn11, and Rpn12) [1]. Intriguingly, the 19 S subcomplex of the 26 S proteasome has been shown to promote the PIC formation at both the TFIID- and SAGA-dependent promoters independently of the 20 S CP [28–30]. At the TFIID-dependent ribosomal protein genes, the 19 S proteasome subcomplex enhances the targeting of NuA4-HAT (that acetylates histone H4) coactivator that facilitates TFIID recruitment to the promoter for transcriptional initiation [29]. At the SAGA-dependent *GAL* genes, the 19 S proteasome subcomplex promotes the targeting of SAGA to the activator [28,30]. SAGA, in turn, facilitates the PIC formation at the core promoter for transcriptional initiation [2,31]. Thus, the 19 S proteasome subcomplex positively regulates transcriptional initiation of both SAGA- and TFIID-dependent genes, independently of the proteolytic function of the proteasome. Such function of the 19 S proteasome subcomplex in stimulation

of transcriptional initiation is mediated via its ATPase activity [2,28–30]. In addition, the proteolytic function of the proteasome has also been shown to regulate transcriptional initiation of certain genes by controlling activator/coactivator/corepressor abundance, localization, or destruction [1,2].

Overall, we have described above two distinct mechanisms of eukaryotic transcriptional initiation by SAGA and TFIID that are dictated by the target specificity of the activator. These two distinct mechanisms are further differentially regulated by the mediator complex [2,21]. Mediator is a large multiprotein complex, and is composed of three modules [32,33]. These modules are known as Srb4 (Srb2, Srb4, Srb5, Srb6, Rox3, Med8, Med11, and Med6), Gal11/Sin4 (Gal11, Rgr1, Sin4, Pgd1, and Med2), and Med9/Med10 modules (Med1, Med4, Med7, Srb7, Med9, and Med10). The Srb4 module interacts with RNA polymerase II. In case of SAGA-dependent transcriptional initiation, mediator is required for TBP recruitment at the core promoter, while it is dispensable for TBP recruitment at the TFIID-dependent promoter [2,7,21]. Likewise, another GTF, namely TFIIB, is differentially required for TBP recruitment to the core promoters of SAGA- and TFIID-dependent genes [2,7,21]. TFIIB facilitates TBP recruitment to the SAGA-dependent promoter, while it is dispensable for recruitment of TBP to the TFIID-dependent promoter. Even though mediator and TFIIB are dispensable for TBP recruitment at the TFIID-dependent promoter, they are required for transcriptional initiation of the TFIID-dependent genes, subsequent to TBP recruitment [2,21]. These results imply the complex regulatory mechanisms of transcriptional initiation by SAGA, TFIID, and other GTFs. In addition, transcriptional initiation is further influenced by chromatin remodeling factors and environmental and intracellular signaling events [1,2,34–39].

1.3
Transcriptional Elongation of RNA Polymerase II Genes

Following transcriptional initiation, RNA polymerase II transitions to elongation. This transition consists of a series of steps, and is referred to as promoter clearance, where RNA polymerase II establishes stable contact with nascent RNA and weakens its interaction with GTFs. The first step in promoter clearance is the formation of an open complex structure where TFIIH unwinds the double-stranded DNA upstream of the TSS (transcription start site) via its DNA helicase activity [40,41]. This is followed by the synthesis of the first RNA bond, utilizing ribonucleotide triphosphate. There are a few cycles of abortive initiation, where the length of nascent RNA transcript varies from two to four nucleotides [40,42,43]. Studies on TATA-box-containing promoters have shown that the upstream edge of the open bubble complex remain fixed with respect to the TATA box and extends downstream of the TSS as the nascent transcript progresses till the bubble collapses and TFIIH is no longer required for keeping the strands of the bubble unannealed [40,42,44]. The nascent transcript becomes about 9–10 nucleotides long when the bubble collapses [40]. After the addition

of few more nucleotides, TFIIB loses contacts with RNA polymerase II [40,45]. There may be some backtracking of RNA polymerase II to an upstream location till the position +23 (relative to TSS) where the mRNA transcript becomes stable enough and the process of promoter clearance gets completed [40,46–48].

A key event in the promoter clearance is the phosphorylation of RNA polymerase II. RNA polymerase II is composed of 12 subunits, namely Rpb1-12. Rpb1 is the largest subunit of RNA polymerase II, and has a characteristic C-terminal domain (CTD). The CTD consists of multiple repeats of a conserved heptamer sequence, YSPTSPS, and is essential for cellular viability [49–51]. The serine (S) residues at positions 2, 5, and 7 are particularly important. These residues are phosphorylated and associated with different transcriptional states. Transcriptional initiating form of RNA polymerase II is not phosphorylated. During transition to elongation, RNA polymerase II is phosphorylated by the Kin28/Cdk7 subunit of TFIIH at S-5 [52,53]. Such phosphorylation does not only promote promoter clearance, but also acts as the recruitment platform for mRNA processing and histone modification factors. The level of S-5 phosphorylation is high close to the 5′-end of the coding region of active gene, and starts to drop as RNA polymerase II progresses more into elongation or toward the 3′-end of the coding sequence [52,53]. There is a concomitant increase in the S-2 phosphorylation as RNA polymerase II progresses through the coding region, which, in turn, modulates transcriptional elongation [52,53].

RNA polymerase II pauses at the promoter proximal regions in *Drosophila* and mammals [54,55]. Negative elongation factor (NELF) and DRB (5,6-dichloro-1-β-D ribofuranosylbenzimidazole) sensitivity-inducing factor (DSIF) play important roles in the pausing of RNA polymerase II [54,55]. DSIF alone does not pause RNA polymerase II [54,55]. DSIF targets NELF to associate with RNA polymerase II for pausing [54,55]. DSIF is present in all eukaryotes and archaea, and shares homology with a bacterial transcription factor, NusG [54,56]. However, NELF is conserved only in higher eukaryotes to cause promoter proximal pausing of RNA polymerase II [54,55]. The pausing of RNA polymerase II has emerged as an important step of transcriptional regulation [54,55]. The dissociation of NELF releases the paused-RNA polymerase II for productive transcriptional elongation. One of the important factors promoting the release of promoter proximal paused-RNA polymerase II to its productive elongating form is the positive transcription elongation factor b (P-TEFb). P-TEFb is a cyclin-dependent kinase consisting of a regulatory subunit, cyclin T, and the kinase subunit, Cdk9 (Ctk1 in yeast). P-TEFb phosphorylates NELF, and thereby triggers NELF's dissociation to release paused-RNA polymerase II into productive elongating form [54,55]. P-TEFb also phosphorylates DSIF and S-2 at the CTD of the Rpb1p subunit of RNA polymerase II [54,55]. The phosphorylations of DSIF and RNA polymerase II have stimulatory effects on transcriptional elongation [53,54]. Thus, P-TEFb plays an important role to release the paused-RNA polymerase II, and enhances transcriptional elongation. The recruitment of P-TEFb to the gene may occur through several ways, including association with DNA binding proteins such as c-Myc and NF-kB, mediator, or

Brd4 (that associates with acetylated tail of histone H4) [54,55,57–60]. Additionally, the amount and availability of active P-TEFb is regulated via sequestering of P-TEFb into an inactive complex with 7SK RNA and HEXIM protein [54,55,61,62]. Therefore, various factors, including signaling molecules and chromatin structure/modification, play crucial roles to control the function of P-TEFb, and hence the release of paused-RNA polymerase II to productive transcriptional elongation.

As mentioned above, P-TEFb phosphorylates the CTD of RNA polymerase II on S-2 as it progresses toward the 3′-end of active gene to ensure productive transcriptional elongation [53,54,62]. In addition, such phosphorylation of RNA polymerase II regulates mRNA processing and export. Consistently, the inhibition of P-TEFb reduces gene expression. Likewise, S-5 phosphorylation of the CTD of RNA polymerase II regulates 5′-end processing (or capping) of mRNA and recruitment of chromatin modification factors. Thus, phosphorylation status of the CTD of RNA polymerase II plays important roles in regulation of mRNA synthesis, processing, and export [53,54,62]. Ssu72 and Fcp1 dephosphorylate S-5 and S-2 residues, respectively, of the CTD as RNA polymerase II elongates the transcript [53,63]. Rtr1 (regulator of transcription) is another RNA polymerase II-associated phosphatase that regulates the transition from phospho-S5 to phospho-S2 of the CTD during transcription elongation [64]. Hence, phosphatases participate in regulation of the phosphorylation states of the CTD of RNA polymerase II, and hence transcriptional elongation and mRNA processing and export.

Productive transcriptional elongation is also ensured by the processivity factors such as TFIIS that facilitates transcription of stalled-RNA polymerase II and FACT (facilitates chromatin transcription), which promotes transcription through chromatin. The histones in the coding region are rapidly evicted during elongation and deposited back onto DNA in the wake of elongating RNA polymerase II [65]. This assembly/disassembly of histones is aided by histone chaperones, FACT and Asf1 (antisilencing function protein 1), histone H2B ubiquitylation, and ATP-dependent chromatin remodeling factors such as Swi/Snf and Rad26 [65–71]. Thus, these proteins/factors serve to facilitate transcriptional elongation on chromatin. Further, mutations in FACT and Asf1 lead to the generation of cryptic transcripts that are initiated from the internal start sites of the coding sequence [66,72]. Thus, in addition to promoting transcriptional elongation, FACT and Asf1 also suppress cryptic transcription. Likewise, another transcriptional elongation factor, Spt6, suppresses transcriptional initiation from the cryptic sites [72]. Additionally, an epigenetic mark at the coding regions, namely histone H3K36 methylation, also prevents spurious transcriptional initiation from the intragenic regions. In yeast, the Rpd3 S histone deacetylase complex recognizes the Set2-mediated histone H3K36 methylation at the coding regions of the active genes, and directs deacetylation of histones to suppress cryptic intragenic transcription [73,74]. Thus, chromatin structure/modifications and associated regulatory factors play important roles in controlling transcriptional elongation.

Overall, transcriptional elongation is a critical step in eukaryotic gene expression. It is regulated by a large number of proteins/factors for promoter clearance, releasing paused-RNA polymerase II at the promoter proximal site, backtracking, and productive elongation. In addition, chromatin structure/modifications and associated regulatory factors control transcriptional elongation. Further, transcriptional elongation is coupled to DNA repair, and maintains genome integrity [75–78]. In turn, DNA repair factors have also been shown to regulate transcription [79–82]. Thus, transcriptional elongation is a highly coordinated and precisely regulated process in eukaryotic gene expression. We have broadly described above this process. However, the detailed regulatory mechanisms of transcriptional elongation have been reviewed elsewhere [83–88].

1.4
Transcriptional Termination of RNA Polymerase II Gene

Termination is the last step of transcription, and marks the ends of various regulated events of gene expression. Transcriptional termination entails the dissociation of the elongation complex at the 3′-ends of the active genes, and is the least understood phase of transcription. Transcriptional termination does not take place at a specific site, but can occur from a few bases to several kilobases of the 3′-end of mRNA [89,90]. Most of the protein-coding mRNA precursors have a highly conserved polyadenylation or poly (A) signal, 5′-AAUAAA-3′, followed by a U/GU-rich sequence. Transcriptional termination closely depends on the poly (A) signal or site. The mRNA processing factors such as mRNA polyadenylation and cleavage proteins also participate in transcription termination. The detailed processing events are described later in this chapter.

There are two major models proposed for transcriptional termination of RNA polymerase II genes. The first model is an antiterminator or allosteric model. According to this model, transcriptional elongation through the poly (A) site leads to a conformational change in the elongation complex due to dissociation of elongation factors and/or association of termination factors [91,92]. Pcf11, an mRNA-processing protein, has been shown to promote such dismantling of the elongation complex. Pcf11 associates with the phosphorylated-CTD of the Rpb1 subunit of RNA polymerase II and mRNA transcript, thus coordinating transcriptional termination with mRNA processing [92–94]. Due to the dual association with RNA polymerase II CTD and mRNA transcript, Pcf11 may serve to relay the force generated by the conformational changes at the CTD to the nascent transcript disrupting the elongation complex.

The second model of transcriptional termination is the torpedo model. It is based on the entry of a 5′ → 3′ exonuclease (Rat1 in yeast/Xrn2 in humans) at the poly (A) site, which degrades nascent mRNA and promotes the release of RNA polymerase II [95,96]. Like the antiterminator model, S2-phosphorylated-CTD of RNA polymerase II and mRNA processing proteins are mediators in this model. The CTD of RNA polymerase II serves as the binding platform for

cleavage and polyadenylation specificity factor (CPSF) that aids in transcriptional termination. The CPSF complex also binds to the transcribed poly (A) site in the pre-mRNA, and contributes to RNA polymerase II pausing downstream of this site [97]. This eventually stimulates cleavage and release of mRNA from the 3′-ends of the genes. Upon cleavage of the nascent mRNA at the poly (A) site, the downstream RNA is rapidly degraded by Rat1/Xrn2, leading to transcriptional termination [95–97]. Rat1 is targeted to RNA polymerase II phosphorylated at S2 by Rtt103, an important accessory protein. In these two models, the poly (A) site or signal is essential for transcriptional termination, thus supporting the idea that transcriptional termination is the combination of both models. Even though considerable progress has been made, much work is needed to fully understand the complex regulatory mechanisms of transcriptional termination.

1.5 Capping of mRNA at the 5′-End

mRNA-processing proteins use the altering phosphorylation status of CTD of the Rpb1 subunit of RNA polymerase II to dynamically associate with transcription machinery, and process mRNA as it is synthesized. The foremost mRNA-processing event that occurs cotranscriptionally is the capping (or cap structure formation) of the 5′-end of nascent mRNA. Capping protects nascent mRNAs from degradation by exonucleases and increases their stabilities [98,99]. Further, the cap structure influences the splicing of the first intron of mRNA. In addition, the cap structure is crucial for translational initiation, and leads to circularization of mRNA in the cytoplasm to facilitate multiple rounds of translation. The mRNA cap structure consists of a guanosine residue harboring a methylation at the *N*-7 position, and binds to pre-mRNA via 5′–5′ triphosphate bridge. Once the nascent pre-mRNA is capped, it is bound by CBC that is involved in mediating a variety of cellular events, such as transcriptional initiation and elongation, mRNA export, histone H2B ubiquitylation, histone H3K36 methylation, and processing of micro RNAs, in addition to protecting nascent mRNA from exonuclease [98]. Hence, mRNA capping has crucial functions in regulation of gene expression.

The cap structure is formed shortly after transcriptional initiation, when nascent mRNA chains are about 20–23 nucleotides long. Capping is carried out by three enzymatic reactions. In the first step, RNA triphosphatase removes the γ-phosphate from the first nucleotide at the 5′-end of mRNA (pppRNA), thereby producing a diphosphate terminus (ppRNA). This is followed by the transfer of guanine monophosphate (GMP) to the diphosphate-end by RNA guanyltransferase to generate the cap (GpppRNA). Finally, the cap is methylated at the *N*-7 position of guanine by the action of guanine *N*7-methyltransferase, resulting in the formation of cap structure (m^7GpppRNA). Functions of mRNA capping enzymes and their associated mechanisms are conserved across eukaryotes. However, the enzymes that mediate mRNA capping are differentially organized in lower eukaryotes as compared to higher eukaryotes. For instance, RNA

triphosphatase and RNA guanyltransferase in budding yeast (*Saccharomyces cerevisiae*) are encoded by two different proteins (Cet1 and Ceg1), while mammals carry a bifunctional capping enzyme with both activities. Similar to budding yeast, fission yeast (*Schizosaccharomyces pombe*) encodes separate triphosphatase (Pct1) and guanylyltransferase (Pce1). However, unlike budding yeast, where Cet1 forms a functional heterodimer with Ceg1, Pct1 and Pce1 in fission yeast do not interact with each other (Ref. herein [100]). Phosphorylated CTD participates in the recruitment of the capping machinery, and hence capping of mRNA [101,102].

The mRNA cap structure at the 5′-end is recognized by CBC that is composed of two subunits, namely Cbp20 and Cbp80. Both Cbp20 and Cbp80 are evolutionarily conserved among eukaryotes. Neither Cbp20 nor Cbp80 binds to mRNA cap structure alone, suggesting that a complex of both Cbp20 and Cbp80 is required for recognition of mRNA cap structure to protect mRNA from exonuclease [98]. In addition to protecting mRNA, CBC also influences the splicing of mRNA by facilitating the recruitment of spliceosomal complex to the intron-containing genes. Further, CBC is involved in mediating the export of mRNA [103]. Moreover, CBC has been recently shown to regulate transcriptional initiation and elongation, histone H2B ubiquitylation, histone H3K36 methylation, and processing of microRNAs (Ref. herein [98]). Thus, CBC does not only maintain the stability of mRNA, but also regulates various transcriptional and posttranscriptional events. Following mRNA export to the cytoplasm, CBC participates in translational initiation for protein synthesis. In addition, CBC is also involved in nonsense-mediated mRNA decay. Collectively, mRNA capping at the 5′-end and CBC play important functions in the regulation of crucial steps of eukaryotic gene expression.

1.6 Processing of mRNA at the 3′-End

Another essential mRNA processing event that occurs after capping and subsequent CBC binding is the polyadenylation at the 3′-end. Eukaryotic mRNAs possess a poly (A) tail at their 3′-ends. Presence of poly (A) tail increases the stability and translational efficiency of mRNA. Only mRNAs that are properly polyadenylated are exported out of the nucleus. Indeed, the ratio of cytoplasmic to nuclear mRNAs is shown to go down in the absence of proper polyadenylation of mRNA [104]. Addition of poly (A) to the 3′-end of mRNA occurs in two steps. In the first step, pre-mRNA is cleaved at a specific site, and subsequently polyadenylated [105]. The efficiency of this process is dependent on the actions of multiprotein complexes that bind to specific sites at the 3′-end of pre-mRNA. Most cellular mRNAs possess some signature sites onto which the cleavage and polyadenylation machinery binds [106]. The most conserved one of these sites is the poly (A) signal, AAUAAAA, located 20–30 nucleotides upstream of cleavage site (CA) where the poly (A) tail is added. The other less-conserved U/GU-rich

sequence is located downstream of CA [105,106]. In higher eukaryotes, these sites are bound by CPSF and cleavage stimulation factor (Cstf) [105,106]. CPSF is composed of five subunits: CPSF-160, CPSF-100, CPSF-73, CPSF-30, and Fip1, and associates with the poly (A) site (or AAUAAAA), while Cstf (that consists of CstF-50, CstF-77, and CstF-64) binds to U/GU-rich region. Binding of Cstf increases the stability of CPSF–RNA complex. Apart from these sequences, there are some accessory elements in pre-mRNA (e.g., upstream sequence elements or USE) that provide additional stability for anchoring the 3′-end formation machinery [107]. Once these cleavage elements are bound, the CPSF-73 subunit of CPSF catalyses an endonucleolytic cleavage at the cleavage site, and the poly (A) tail is immediately added at the formed 3′-end by poly (A) polymerase. In yeast, this sequential processing of mRNA is brought about by the actions of cleavage and polyadenylation factor (CPF) and cleavage factor 1 (CF1) [106]. CF1 consists of two factors, namely CF1A and CF1B. CF1A is composed of Rna14, Rna15, Clp1, and Pcf11, while CF1B is composed of Hrp1. While the subunits of CPF are homologous to the CPSF subunits in higher eukaryotes, subunits of yeast CF1A are homologous to the Cstf subunits [105,107,108]. Thus, mRNA processing at the 3′-end appears to be conserved among eukaryotes.

1.7 Splicing of mRNA

Greater complexity in the gene structures of higher eukaryotes necessitates the involvement of additional mechanisms in processing of mRNA before it is exported out of the nucleus. While most of the lower eukaryotic genes lack introns, higher eukaryotic pre-mRNAs contain introns, and hence, are subject to splicing prior to be exported to the cytoplasm for translation. In the process of splicing, introns are removed from pre-mRNAs to form an export-competent mature mRNA. Pre-mRNA splicing is carried out by a large RNA–protein complex (60 S), known as spliceosome [109]. Spliceosome is composed of snRNAs (U1, U2, U4, U5, and U6) and their associated proteins. In yeast, apart from snRNAs and associated proteins, ~100 additional components make up the spliceosome. In humans, this complex has more than 300 components owing to the complexity of the genome. In fact, human genome contains more than 200 000 introns of varying lengths, and demands a complex spliceosomal apparatus [109]. Interestingly, evidence also suggests the existence of another splicing complex in humans, known as minor spliceosome that is involved in the splicing of a minor class introns with noncanonical consensus sequences [110,111]. Minor spliceosome contains U11, U12, U4atac, U6atac, and U5 snRNP (small nuclear ribonucleoprotein). U5 snRNP is shared with major spliceosome, while other components of minor spliceosome are functionally analogous to U1, U2, U4, and U6 snRNPs [112].

Pre-mRNA sequences define the boundaries of introns. Most introns have the GU sequence at the 5′ splice site and AG sequence at the 3′ splice site. Further, a

variable length of polypyrimidines (known as polypyrimidine tract) is present upstream of the 3′ splice site. Apart from these sequences, there is a branch point that includes adenosine and is located 18–40 bp upstream of the 3′ splice site. These sequences serve as the recognition sites of the spliceosomal components. Spliceosome carries out two transesterification reactions to excise introns. In the first step, phosphodiester bond at the 5′ splice site is attacked by 2′-hydroxyl of adenosine of the branch point, thereby generating a free 5′-exon and lariat-3′-exon intermediates. Following this, 3′-hydroxyl of the 5′ splice site attacks the phosphodiester bond at the 3′ splice site, leading to the excision of lariat and ligation of exons. These steps are mediated by the stepwise assembly of spliceosomal components as described below.

Splicing is initiated by the binding of U1 snRNP to the 5′ splice site of an intron. Simultaneously, the 3′ splice site is bound by an U2 auxiliary factor (U2AF). These factors along with some additional factors form the first complex known as commitment or E (early) complex. In the E complex, both exons are brought in close proximity. The E complex is followed by the recruitment of U2 snRNP by U2AF. U2 snRNP binds to the branch point adenosine to form pre-spliceosomal A complex. These steps are then followed by the recruitment of U4, U5, and U6 snRNPs, thereby forming the B complex. Subsequently, catalytically active spliceosomal C complex is formed by structural rearrangements of the B complex [113,114], and leads to the formation of lariat, excision of intron, and ligation of the exons. Intron-free mRNA thus generated becomes ready to be exported out of the nucleus [114]. While the sequential assembly of the spliceosomal components on the splice sites is well recognized, an alternative view of spliceosomal assembly has also been proposed by Stevens *et al.* [115]. According to their study, spliceosomal components are preassembled prior to the binding to pre-mRNA, thereby raising a possibility of exclusion of multistep assembly of spliceosomal components.

Pre-mRNA splicing requires ATP hydrolysis and RNA unwinding that is brought about by the action of eight splicing factors that belong to the family of DExH/D-box helicases (Prp5/DDX23, Sub2/UAP56, Prp28/DDX46, Brr2/U5200KD, Prp2/DHX16, Prp16/DHX38, Prp22/DHX8, and Prp43/DHX15) [113,114,116–118]. Once spliceosome brings about the removal of introns from pre-mRNA, a complex of proteins gets recruited to a position of 20–24 nucleotides upstream of the exon–exon junction [119,120]. This complex is termed as the exon junction complex (EJC), and travels with mRNA to the cytoplasm. EJC serves as the binding platform for many proteins, some of which are involved in mRNA export [120]. Thus, splicing of intron-containing genes is linked to nuclear mRNA export. Further, EJC is involved in nonsense-mediated mRNA decay that detects and degrades mRNAs containing premature stop codons [120]. EJC can also enhanced translation efficiency. Thus, splicing of intron-containing genes is coupled to the quality control and translation efficiency of mRNA via EJC.

While it is easy to envision the synthesis of single mRNA from a single gene, the complexity of higher eukaryotes requires the production of numerous

proteins (or functional mRNAs) involved in regulation of many processes as opposed to fewer proteins in lower eukaryotes. Hence, to accommodate such diversity of proteins, cells have come up with alternative splicing where a single gene encodes multiple functional mRNAs, and hence proteins. In fact, alternative splicing has been observed in ~25% of *Caenorhabditis elegans*, ~60% of *Drosophila melanogaster*, and ~95% of human genes [121]. While normal splicing of pre-mRNA involves four splicing signals, alternative splicing requires additional *trans*-acting factors that bind to *cis*-acting sequences in pre-mRNA to promote the synthesis of splice variants [121,122]. Alternative splicing can be classified into four main categories: (i) exon skipping, where an exon can be spliced out together with an intron; (ii) alternative 5′ and (iii) 3′ selections, where these splice sites may be recognized just at one end of an intron; and (iv) intron retention, where an intron can be retained along with exons. Such a mechanism is a smart move by the cell to promote proteomic diversity in higher eukaryotes. Overall, splicing of pre-mRNA is regulated by a large number of factors in a highly coordinated manner to precisely regulate eukaryotic gene expression.

1.8
Nuclear Export of mRNA for Translation

After pre-mRNA undergoes multiple processing events inside the nucleus, it needs to be exported to the cytoplasm for translation into proteins. Nuclear mRNA export occurs through nuclear pore complex (NPC), a complex of nucleoporins that is embedded within the nuclear envelope. NPCs are the only doorways for transport of molecules in and out of the nucleus. This evolutionarily conserved complex consists of a central transporter channel that is surrounded by a symmetrical spoke complex [123–129]. Spoke complex is flanked by nuclear and cytoplasmic rings on either sides of the nuclear membrane. Export of mRNAs through the NPC is mediated via transient interactions between the receptors that bind to mRNA cargos and the degenerate FG (phenylalanine–glycine) repeats of the nucleoporins [130–138]. Hence, mRNA receptors play crucial roles in transporting mRNAs out of the nucleus.

A vast majority of receptors involved in the transport process belong to a family called karyopherins (termed as importins or exportins based on the directionality of transport). The directionality of karyopherin family of receptors depends on a small GTPase Ran that plays a key role in the association and dissociation of substrates with transport receptor [133,139]. However, substantial evidence suggests that the export of mRNA from nucleus to cytoplasm is not mediated by the karyopherin-β family members, and is much more complicated as compared to karyopherin-mediated transport [138]. The export of mRNAs is carried out by the export receptor family of proteins termed as nuclear export factors (NXFs). While numerous NXFs have been discovered so far across eukaryotes (several NXFs in humans, NXF1–6; four in *Drosophila*, DmNXF1–4; two in *Caenorhabditis elegans*, CeNXF1–2; and one in *Saccharomyces cerevisiae*, Mex67), Mex67

and its human homolog, NXF1 (also known as TAP), are well characterized [140–147]. Mex67/TAP is an essential protein that forms a functional heterodimer with Mtr2/p15 to mediate the export of mRNA [145]. Being a shuttling protein, Mex67 travels with mRNA to the cytoplasm, and then gets dissociated from mRNA by the combined actions of Gle1, Dbp5 (a DEAD-box RNA helicase) and InsP6 [130,148]. The conditional inactivation of Mex67/TAP leads to an accumulation of poly (A) RNA in the nucleus, thereby supporting the roles of these factors in mRNA export.

Despite its function in mRNA export, Mex67/TAP displays a low affinity for binding to mRNAs, and hence requires adapter proteins to bridge this interaction. Genetic and biochemical studies in yeast have led to the identification of a Mex67 binding-adaptor protein, namely Yra1 (yeast RNA annealing protein) [149–152]. Yra1 belongs to the REF (RNA export factor) family of proteins. Mutation in Yra1 causes nuclear accumulation of poly (A) RNA, implicating its role in the export of mRNA [149,150]. Further, the association of Mex67 with newly synthesized mRNAs is also reduced in Yra1 mutant strain, thus indicating an essential role of Yra1 in enhancing the interaction of mRNA with Mex67 [149,150,153].

To understand the mechanism of recruitment of Yra1 to the nascent mRNA, a synthetic lethal screening was performed, which revealed the genetic interaction of Yra1 with Sub2, a DEAD-box helicase/ATPase that is involved in pre-mRNA splicing as well as mRNA export [154]. Further, Yra1 has been shown to interact with Sub2 [154]. Sub2 and Yra1 are also found to be loaded onto the active genes in a transcription-dependent manner via the THO (Suppressors of transcriptional defects of *hpr1Δ* by overexpression) complex [155–157]. THO is a tetrameric complex that consists of Tho2, Hpr1, Mft1, and Thp2. In addition, Tex1, a nonessential protein of unknown function, has been shown to be associated with THO [158]. The THO complex is evolutionarily conserved across eukaryotes. Human THO complex contains homologs of yeast Hpr1 (THOC1) and Tho2 (THOC2) with additional subunits fSAP79/THOC5, fSAP35/THOC6, and fSAP24/THOC7 [159–162]. Tex1 homolog, THOC3, is also present in human THO complex. Human THO complex resembles *Drosophila* THO complex. However, there is no apparent yeast Mft1 and Thp2 homologs in human and *Drosophilla* THO complexes. Moreover, human or *Drosophila* THOC5, THOC6, and THOC7 homologs are not apparently found in yeast THO complex [159,160,163]. THO plays an important role in transcriptional elongation. In yeast, null mutations of any of its four nonessential subunits (Hpr1, Thp2, Tho2, and Mft1) have been linked to increased hyperrecombination and defects in transcription elongation. This has been attributed to the formation of DNA: RNA hybrid (R-loop) between nascent transcript and template DNA, which impedes transcription and also leads to genomic instability associated with the accumulation of DNA breaks [155–157,159,164,165]. Recently, the depletion of human THO subunits has also been shown to generate similar effects on transcriptional elongation and genomic instability, thus indicating the functional conservations of the THO complex across species [166]. The THO complex has

been shown to be recruited to the active chromatin during transcriptional elongation in an RNA polymerase II-dependent fashion [159,167,168]. Once THO is recruited, it facilitates the association of export factor Sub2 that then loads mRNA export adapter, Yra1. In fact, the THO subunits have been shown to interact physically and genetically with Sub2 and Yra1. Particularly, Hpr1 interacts directly with Sub2, and helps in the recruitment of Sub2 and Yra1 to the actively transcribing genes [153,154,167,169,170]. Additionally, THO mutants are found to be defective in mRNA export. These observations provide evidence for the coupling of transcription elongation with mRNA export via THO. Once Yra1 is recruited, Sub2 dissociates from it, thereby allowing the interaction of Yra1 with the export receptor, Mex67. The interactions of Sub2 and Mex67 with Yra1 are mutually exclusive as both Sub2 and Mex67 bind competitively with Yra1 [154]. However, Sub2 and Mex67 can simultaneously bind with Yra1, since both Sub2 and Mex67 can independently interact with each of the two domains of Yra1 [154]. Sub2 and Yra1 together with the THO components form the TREX (transcription–export) complex. TREX is conserved among eukaryotes. The human homologs of yeast Sub2 and Yra1 are UAP56 and ALY, respectively. These components are essential for cell survival, and perform similar functions in yeast and humans [158].

The TREX complex gets recruited cotranscriptionally to the coding regions of active genes, but shows a bias toward the 3′-end. While the recruitment of yeast TREX complex to the active gene is coupled to ongoing transcription, the recruitment of metazoan TREX has been shown to be dependent on splicing [171,172]. Intriguingly, recent studies by Johnson *et al.* [173,174] showed that Yra1 can be loaded onto the active genes in a Sub2-independent manner. Their studies revealed that Yra1 is loaded onto the active gene via its interaction with the 3′-end processing factor, Pcf11. Pcf11 is recruited cotranscriptionally to the active genes through its interactions with the CTD of RNA Polymerase II. Yra1 that is recruited to the active genes in this manner is then transferred to the export factor, Sub2. Subsequently, Yra1 becomes available for interaction with the export receptor, Mex67, for nuclear export of mRNA. However, Hobeika et al [175,176] showed that Mex67 is recruited in an earlier step during transcription. Such recruitment is dependent on the Hpr1 component of the THO complex. The UBA (ubiquitin-associated) domain of Mex67 recognizes the ubiquitylated-form of Hpr1. This interaction contributes to the recruitment of Mex67 and protects ubiquitylated-Hpr1 from the 26 S proteasome complex. Consistently, the absence of the ubiquitylated-form of Hpr1 or the UBA domain of Mex67 is associated with the defects in mRNA export. Thus, cotranscriptional recruitment of Mex67 to the active genes is indispensable for mRNA export. As mentioned earlier, Mex67 has less affinity for binding to mRNA and Yra1 has been well demonstrated to bridge this interaction in yeast. While Yra1 is an essential protein in yeast, double-stranded RNA interference of REF/Aly in *Drosophila* and *Caenorhabditis elegans* concluded that Aly is dispensable for mRNA export [177,178]. This points out to the existence of additional export adaptors in bridging the interaction of mRNAs with export receptors in higher

eukaryotes [177,178]. Indeed, studies have indicated the roles of SR (serine–arginine)-rich proteins, SRp20 and 9G8, in mammals as export adaptors in bridging the interaction of mRNA with receptor TAP [179–182].

While a number of studies demonstrated the interconnection between transcription elongation and mRNA export, studies over the past several years have also shed light on tight coupling that exists between transcriptional activation and export of mRNA [183–185]. Such a link was established with the identification of a factor called Sus1 (sl gene upstream of Ysa1) [183–185]. Sus1 was initially identified by synthetic lethality screening with an Yra1 mutant allele, and was shown to interact genetically and physically with coactivator SAGA and mRNA export factors, Thp1 and Sac3 (components of the TREX-2 complex). Consistently, Sus1 has been found to be a common component of SAGA and TREX-2 [183]. Thus, Sus1 links SAGA with TREX-2, thereby playing a key role in coupling transcriptional activation with mRNA export.

The TREX-2 complex is constituted by Thp1, Sac3, Cdc31, and Sus1, and is located on the inner side of NPC. Sac3 forms the central scaffold of the TREX-2 complex [186,187]. Sus1 associates with the TREX-2 complex via Sac3, and functions in mRNA export. TREX-2 interacts with NPC via nucleoporins, such as Nup1 and Nup60, and also associates with Mex67–Mtr2 export receptor [183,184,186]. Recent studies have also identified Sem1 as a new component of this evolutionarily conserved complex [188–190]. The physical link between mRNA export complex (TREX-2) and transcriptional coactivator complex (SAGA) via Sus1 supported the "gene gating" hypothesis as originally proposed by Blobel [191]. The TREX-2 has been shown to play an important role in repositioning yeast genes such as *GAL1* and *INO1* in close proximity to NPC through its interactions with nucleoporins of NPC and SAGA [185,192–194]. Apart from yeast, gene-gating hypothesis has also been proved in *Drosophila* and humans [192,195–197].

Prior to nuclear export, mRNA undergoes a quality control check with the aid of nuclear surveillance complex, known as exosome [198]. Exosome complex contains several $3' \rightarrow 5'$ exonucleases, and degrades the improperly processed mRNAs prior to their export to the cytoplasm [199]. Only properly processed mRNAs or mRNPs are exported out of the nucleus. Once these processed mRNAs reach cytoplasm, most of them enter the translationally active pool that encodes proteins. During translation, mRNA is threaded through the space between the ribosomal subunits to undergo a pioneering round of translation [200–202]. This process removes any associated hnRNPs (heterogeneous nuclear ribonucleoproteins) that might hinder the translational activity. At this stage, nuclear 5′-cap and 3′-poly (A) binding nuclear protein are replaced by elF4E (cytoplasmic cap binding protein) and poly (A)-binding cytoplasmic protein, respectively. All these events result in functional circularization of mRNA due to the *cis* and *trans* interactions between its 5′- and 3′-ends, thereby promoting the efficient translation of the message [201,202]. Life of mRNAs ends after translation in the cytoplasm. The mRNA cap is removed by decapping enzymes, and mRNA is degraded by exonucleases [203–205]. The degradation

or at least the final step(s) of degradation of mRNA is believed to occur in certain cytoplasmic foci, known as P-bodies (processing bodies) [203–205]. P-bodies are the structures that are enriched with numerous proteins involved in mRNA degradation [203–205].

1.9 Conclusion

Eukaryotic gene expression is a highly coordinated and complex regulatory process. It is mechanistically composed of many different steps as described above. Misregulation of any of these steps leads to an altered pattern of gene expression, which can change cellular functions and threaten cellular viability. Thus, the factors involved in different steps of gene expression are correlated with a variety of diseases. Therefore, a large number of studies are focused on understanding the regulation of eukaryotic gene expression in a number of eukaryotic systems. In this chapter, we have broadly described eukaryotic gene expression from the birth to the end of mRNA life. Thus, this chapter would provide general views of gene expression, function, and regulation. However, the details of the different steps of eukaryotic gene expression by RNA polymerase II are described in a number of review articles cited in this chapter.

References

1 Bhaumik, S.R. and Mailk, S. (2008) Diverse regulatory mechanisms of eukaryotic transcriptional activation by the proteasome complex. *Crit. Rev. Biochem. Mol. Biol.*, **43**, 419–433.

2 Bhaumik, S.R. (2011) Distinct regulatory mechanisms of eukaryotic transcriptional activation by SAGA and TFIID. *Biochim. Biophys. Acta*, **1809**, 97–108.

3 Rodriguez-Navarro, S. (2009) Insights into SAGA function during gene expression. *EMBO Rep.*, **10**, 843–850.

4 Daniel, J.A. and Grant, P.A. (2007) Multi-tasking on chromatin with the SAGA coactivator complexes. *Mutat. Res.*, **618**, 135–148.

5 Bhaumik, S.R. and Green, M.R. (2002) Differential requirement of SAGA components for recruitment of TATA-box-binding protein to promoters *in vivo*. *Mol. Cell Biol.*, **22**, 7365–7371.

6 Sterner, D.E., Grant, P.A., Roberts, S.M., Duggan, L.J., Belotserkovskaya, R., Pacella, L.A., Winston, F., Workman, J.L., and Berger, S.L. (1999) Functional organization of the yeast SAGA complex: distinct components involved in structural integrity, nucleosome acetylation, and TATA-binding protein interaction. *Mol. Cell Biol.*, **19**, 86–98.

7 Bhaumik, S.R., Raha, T., Aiello, D.P., and Green, M.R. (2004) *In vivo* target of a transcriptional activator revealed by fluorescence resonance energy transfer. *Genes. Dev.*, **18**, 333–343.

8 Eisenmann, D.M., Arndt, K.M., Ricupero, S.L., Rooney, J.W., and Winston, F. (1992) SPT3 interacts with TFIID to allow normal transcription in *Saccharomyces cerevisiae*. *Genes Dev.*, **6**, 1319–1331.

9 Laprade, L., Rose, D., and Winston, F. (2007) Characterization of new Spt3 and TATA-binding protein mutants of *Saccharomyces cerevisiae*: Spt3 TBP allele-specific interactions and bypass of Spt8. *Genetics*, **177**, 2007–2017.

10 Kuo, M.H., Brownell, J.E., Sobel, R.E., Ranalli, T.A., Cook, R.G., Edmondson, D.G., Roth, S.Y., and Allis, C.D. (1996) Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. *Nature*, **383**, 269–272.

11 Grant, P.A., Eberharter, A., John, S., Cook, R.G., Turner, B.M., and Workman, J.L. (1999) Expanded lysine acetylation specificity of Gcn5 in native complexes. *J. Biol. Chem.*, **274**, 5895–5900.

12 Ricci, A.R., Genereaux, J., and Brandl, C.J. (2002) Components of the SAGA histone acetyltransferase complex are required for repressed transcription of ARG1 in rich medium. *Mol. Cell Biol.*, **22**, 4033–4042.

13 Daniel, J.A., Torok, M.S., Sun, Z.W., Schieltz, D., Allis, C.D., Yates, J.R. 3rd, and Grant, P.A. (2004) Deubiquitination of histone H2B by a yeast acetyltransferase complex regulates transcription. *J. Biol. Chem.*, **279**, 1867–1871.

14 Shukla, A., Stanojevic, N., Duan, Z., Sen, P., and Bhaumik, S.R. (2006) Ubp8p, a histone deubiquitinase whose association with SAGA is mediated by Sgf11p, differentially regulates lysine 4 methylation of histone H3 *in vivo*. *Mol. Cell Biol.*, **26**, 3339–3352.

15 Henry, K.W., Wyce, A., Lo, W.S., Duggan, L.J., Emre, N.C., Kao, C.F., Pillus, L., Shilatifard, A., Osley, M.A., and Berger, S.L. (2003) Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8. *Genes Dev.*, **17**, 2648–2663.

16 Shukla, A., Chaurasia, P., and Bhaumik, S.R. (2009) Histone methylation and ubiquitination with their cross-talk and roles in gene expression and stability. *Cell Mol. Life Sci.*, **66**, 1419–1433.

17 Gangloff, Y.G., Pointud, J.C., Thuault, S., Carre, L., Romier, C., Muratoglu, S., Brand, M., Tora, L., Couderc, J.L., and Davidson, I. (2001) The TFIID components human TAF(II)140 and *Drosophila* BIP2 (TAF(II)155) are novel metazoan homologues of yeast TAF(II)47 containing a histone fold and a PHD finger. *Mol. Cell Biol.*, **21**, 5109–5121.

18 Gangloff, Y.G., Romier, C., Thuault, S., Werten, S., and Davidson, I. (2001) The histone fold is a key structural motif of transcription factor TFIID. *Trends Biochem. Sci.*, **26**, 250–257.

19 Garbett, K.A., Tripathi, M.K., Cencki, B., Layer, J.H., and Weil, PA. (2007) Yeast TFIID serves as a coactivator for Rap1p by direct protein-protein interaction. *Mol. Cell Biol.*, **27**, 297–311.

20 Layer, J.H., Miller, S.G., and Weil, PA. (2010) Direct transactivator-transcription factor IID (TFIID) contacts drive yeast ribosomal protein gene transcription. *J. Biol. Chem.*, **285**, 15489–15499.

21 Li, X.Y., Bhaumik, S.R., and Green, M.R. (2000) Distinct classes of yeast promoters revealed by differential TAF recruitment. *Science*, **288**, 1242–1244.

22 Li, X.Y., Bhaumik, S.R., Zhu, X., Li, L., Shen, W.C., Dixit, B.L., and Green, M.R. (2002) Selective recruitment of TAFs by yeast upstream activating sequences. Implications for eukaryotic promoter structure. *Curr. Biol.*, **12**, 1240–1244.

23 Shen, W.C., Bhaumik, S.R., Causton, H.C., Simon, I., Zhu, X., Jennings, E.G., Wang, T.H., Young, R.A., and Green, M.R. (2003) Systematic analysis of essential yeast TAFs in genome-wide transcription and preinitiation complex assembly. *EMBO J.*, **22**, 3395–3402.

24 Rhee, H.S. and Pugh, B.F. (2012) Genome-wide structure and organization of eukaryotic pre-initiation complexes. *Nature*, **483**, 295–301.

25 Smale, S.T. (2001) Core promoters: active contributors to combinatorial gene regulation. *Genes. Dev.*, **15**, 2503–2508.

26 Vermeulen, M., Mulder, K.W., Denissov, S., Pijnappel, W.W., van Schaik, F.M., Varier, R.A., Baltissen, M.P., Stunnenberg, H.G., Mann, M., and Timmers, H.T. (2007) Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell*, **131**, 58–69.

27 Cler, E., Papai, G., Schultz, P., and Davidson, I. (2009) Recent advances in understanding the structure and function of general transcription factor TFIID. *Cell Mol. Life Sci.*, **66**, 2123–2134.

28 Malik, S., Shukla, A., Sen, P., and Bhaumik, S.R. (2009) The 19S proteasome subcomplex establishes a specific protein interaction network at the promoter for stimulated transcriptional initiation *in vivo*. *J. Biol. Chem.*, **284**, 35714–35724.

29 Uprety, B., Lahudkar, S., Malik, S., and Bhaumik, S.R. (2012) The 19S proteasome subcomplex promotes the targeting of NuA4 HAT to the promoters of ribosomal protein genes to facilitate the recruitment of TFIID for transcriptional initiation *in vivo*. *Nucleic Acids Res.*, **40**, 1969–1983.

30 Lee, D., Ezhkova, E., Li, B., Pattenden, S.G., Tansey, W.P., and Workman, J.L. (2005) The proteasome regulatory particle alters the SAGA coactivator to enhance its interactions with transcriptional activators. *Cell*, **123**, 423–436.

31 Bhaumik, S.R. and Green, M.R. (2001) SAGA is an essential *in vivo* target of the yeast acidic activator Gal4p. *Genes Dev.*, **15**, 1935–1945.

32 Kornberg, R.D. (2005) Mediator and the mechanism of transcriptional activation. *Trends Biochem. Sci.*, **30**, 235–239.

33 Myers, L.C. and Kornberg, R.D. (2000) Mediator of transcriptional regulation. *Annu. Rev. Biochem.*, **69**, 729–749.

34 Roeder, R.G. (2005) Transcriptional regulation and the role of diverse coactivators in animal cells. *FEBS Lett.*, **579**, 909–915.

35 Carrozza, M.J., Utley, R.T., Workman, J.L., and Cote, J. (2003) The diverse functions of histone acetyltransferase complexes. *Trends Genet.*, **19**, 321–329.

36 Peterson, C.L. and Workman, J.L. (2000) Promoter targeting and chromatin remodeling by the SWI/SNF complex. *Curr. Opin. Genet. Dev.*, **10**, 187–192.

37 Hassan, A.H., Neely, K.E., Vignali, M., Reese, J.C., and Workman, J.L. (2001) Promoter targeting of chromatin-modifying complexes. *Front. Biosci.*, **6**, D1054–D1064.

38 Smith, C.L. and Peterson, C.L. (2005) ATP-dependent chromatin remodeling. *Curr. Top. Dev. Biol.*, **65**, 115–148.

39 Workman, J.J., Chen, H., and Laribee, R.N. (2014) Environmental signaling through the mechanistic target of rapamycin complex 1: mTORC1 goes nuclear. *Cell Cycle*, **13** (5), 714–725.

40 Luse, D.S. and Samkurashvili, I. (1998) The transition from initiation to elongation by RNA polymerase II. *Cold Spring Harb. Symp. Quant. Biol.*, **63**, 289–300.

41 Lin, Y.C., Choi, W.S., and Gralla, J.D. (2005) TFIIH XPB mutants suggest a unified bacterial-like mechanism for promoter opening but not escape. *Nat. Struct. Mol. Biol.*, **12**, 603–607.

42 Holstege, F.C., Fiedler, U., and Timmers, H.T. (1997) Three transitions in the RNA polymerase II transcription complex during initiation. *EMBO J.*, **16**, 7468–7480.

43 Saunders, A., Core, L.J., and Lis, J.T. (2006) Breaking barriers to transcription elongation. *Nat. Rev. Mol. Cell Biol.*, **7**, 557–567.

44 Pal, M., Ponticelli, A.S., and Luse, D.S. (2005) The role of the transcription bubble and TFIIB in promoter clearance by RNA polymerase II. *Mol. Cell*, **19**, 101–110.

45 Kostrewa, D., Zeller, M.E., Armache, K.J., Seizl, M., Leike, K., Thomm, M., and Cramer, P. (2009) RNA polymerase II-TFIIB structure and mechanism of transcription initiation. *Nature*, **462**, 323–330.

46 Pal, M. and Luse, D.S. (2003) The initiation-elongation transition: lateral mobility of RNA in RNA polymerase II complexes is greatly reduced at +8/+9 and absent by +23. *Proc. Natl. Acad. Sci. USA*, **100**, 5700–5705.

47 Luse, D.S. (2013) Promoter clearance by RNA polymerase II. *Biochim. Biophys. Acta*, **1829**, 63–68.

48 Pal, M., McKean, D., and Luse, D.S. (2001) Promoter clearance by RNA polymerase II is an extended, multistep process strongly affected by sequence. *Mol. Cell Biol.*, **21**, 5815–5825.

49 Scafe, C., Martin, C., Nonet, M., Podos, S., Okamura, S., and Young, R.A. (1990) Conditional mutations occur predominantly in highly conserved residues of RNA polymerase II subunits. *Mol. Cell Biol.*, **10**, 1270–1275.

50 Bartolomei, M.S., Halden, N.F., Cullen, C.R., and Corden, J.L. (1988) Genetic analysis of the repetitive carboxyl-terminal domain of the largest subunit of mouse RNA polymerase II. *Mol. Cell Biol.*, **8**, 330–339.

51 Stiller, J.W. and Hall, B.D. (2002) Evolution of the RNA polymerase II C-terminal domain. *Proc. Natl. Acad. Sci. USA*, **99**, 6091–6096.

52 Komarnitsky, P., Cho, E.J., and Buratowski, S. (2000) Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev.*, **14**, 2452–2460.

53 Buratowski, S. (2009) Progression through the RNA polymerase II CTD cycle. *Mol. Cell*, **36**, 541–546.

54 Nechaev, S. and Adelman, K. (2011) Pol II waiting in the starting gates: regulating the transition from transcription initiation into productive elongation. *Biochim. Biophys. Acta*, **1809**, 34–45.

55 Li, J. and Gilmour, D.S. (2011) Promoter proximal pausing and the control of gene expression. *Curr. Opin. Genet. Dev.*, **21**, 231–235.

56 Hirtreiter, A., Damsma, G.E., Cheung, A.C., Klose, D., Grohmann, D., Vojnic, E., Martin, A.C., Cramer, P., and Werner, F. (2010) Spt4/5 stimulates transcription elongation through the RNA polymerase clamp coiled-coil motif. *Nucleic Acids Res.*, **38**, 4040–4051.

57 Rahl, P.B., Lin, C.Y., Seila, A.C., Flynn, R.A., McCuine, S., Burge, C.B., Sharp, P.A., and Young, R.A. (2010) c-Myc regulates transcriptional pause release. *Cell*, **141**, 432–445.

58 Barboric, M., Nissen, R.M., Kanazawa, S., Jabrane-Ferrat, N., and Peterlin, B.M. (2001) NF-kappaB binds P-TEFb to stimulate transcriptional elongation by RNA polymerase II. *Mol. Cell*, **8**, 327–337.

59 Hargreaves, D.C., Horng, T., and Medzhitov, R. (2009) Control of inducible gene expression by signal-dependent transcriptional elongation. *Cell*, **138**, 129–145.

60 Donner, A.J., Ebmeier, C.C., Taatjes, D.J., and Espinosa, J.M. (2010) CDK8 is a positive regulator of transcriptional elongation within the serum response network. *Nat. Struct. Mol. Biol.*, **17**, 194–201.

61 Diribarne, G. and Bensaude, O. (2009) 7SK RNA, a non-coding RNA regulating P-TEFb, a general transcription factor. *RNA Biol.*, **6**, 122–128.

62 Peterlin, B.M. and Price, D.H. (2006) Controlling the elongation phase of transcription with P-TEFb. *Mol. Cell*, **23**, 297–305.

63 Reyes-Reyes, M. and Hampsey, M. (2007) Role for the Ssu72 C-terminal domain phosphatase in RNA polymerase II transcription elongation. *Mol. Cell Biol.*, **27**, 926–936.

64 Mosley, A.L., Pattenden, S.G., Carey, M., Venkatesh, S., Gilmore, J.M., Florens, L., Workman, J.L., and Washburn, M.P. (2009) Rtr1 is a CTD phosphatase that regulates RNA polymerase II during the transition from serine 5 to serine 2 phosphorylation. *Mol. Cell*, **34**, 168–178.

65 Schwabish, M.A. and Struhl, K. (2004) Evidence for eviction and rapid deposition of histones upon transcriptional elongation by RNA polymerase II. *Mol. Cell Biol.*, **24**, 10111–10117.

66 Schwabish, M.A. and Struhl, K. (2006) Asf1 mediates histone eviction and deposition during elongation by RNA polymerase II. *Mol. Cell*, **22**, 415–422.

67 Belotserkovskaya, R., Oh, S., Bondarenko, V.A., Orphanides, G., Studitsky, V.M., and Reinberg, D. (2003) FACT facilitates transcription-dependent nucleosome alteration. *Science*, **301**, 1090–1093.

68 Malik, S. and Bhaumik, S.R. (2012) Rad26p, a transcription-coupled repair factor, promotes the eviction and prevents the reassociation of histone H2A-H2B dimer during transcriptional elongation *in vivo*. *Biochemistry*, **51**, 5873–5875

69 Malik, S., Chaurasia, P., Lahudkar, S., Uprety, B., and Bhaumik, S.R. (2012) Rad26p regulates the occupancy of histone H2A-H2B dimer at the active genes *in vivo*. *Nucleic Acids Res.*, **40**, 3348–3363.

70 Fleming, A.B., Kao, C.F., Hillyer, C., Pikaart, M., and Osley, M.A. (2008) H2B

ubiquitylation plays a role in nucleosome dynamics during transcription elongation. *Mol. Cell*, **31**, 57–66.
71 Schwabish, M.A. and Struhl, K. (2007) The Swi/Snf complex is important for histone eviction during transcriptional activation and RNA polymerase II elongation *in vivo*. *Mol. Cell Biol.*, **27**, 6987–6995.
72 Kaplan, C.D., Laprade, L., and Winston, F. (2003) Transcription elongation factors repress transcription initiation from cryptic sites. *Science*, **301**, 1096–1099.
73 Carrozza, M.J., Li, B., Florens, L., Suganuma, T., Swanson, S.K., Lee, K.K., Shia, W.J., Anderson, S., Yates, J., Washburn, M.P. *et al.* (2005) Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell*, **123**, 581–592.
74 Keogh, M.C., Kurdistani, S.K., Morris, S.A., Ahn, S.H., Podolny, V., Collins, S.R., Schuldiner, M., Chin, K., Punna, T., Thompson, N.J. *et al.* (2005) Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell*, **123**, 593–605.
75 Vermeulen, W. and Fousteri, M. (2013) Mammalian transcription-coupled excision repair. *Cold Spring Harb. Perspect. Biol.*, **5**, a012625.
76 Gaillard, H. and Aguilera, A. (2013) Transcription coupled repair at the interface between transcription elongation and mRNP biogenesis. *Biochim. Biophys. Acta*, **1829**, 141–150.
77 Tornaletti, S. (2009) DNA repair in mammalian cells: transcription-coupled DNA repair: directing your effort where it's most needed. *Cell Mol. Life Sci.*, **66**, 1010–1020.
78 Hanawalt, P.C. and Spivak, G. (2008) Transcription-coupled DNA repair: two decades of progress and surprises. *Nat. Rev. Mol. Cell Biol.*, **9**, 958–970.
79 Fong, Y.W., Cattoglio, C., and Tjian, R. (2013) The intertwined roles of transcription and repair proteins. *Mol. Cell*, **52**, 291–302.
80 Woudstra, E.C., Gilbert, C., Fellows, J., Jansen, L., Brouwer, J., Erdjument-Bromage, H., Tempst, P., and Svejstrup, J.Q. (2002) A Rad26-Def1 complex coordinates repair and RNA pol II proteolysis in response to DNA damage. *Nature*, **415**, 929–933.
81 Le May, N., Mota-Fernandes, D., Vélez-Cruz, R., Iltis, I., Biard, D., and Egly, J.M. (2010) NER factors are recruited to active promoters and facilitate chromatin modification for transcription in the absence of exogenous genotoxic attack. *Mol. Cell*, **38**, 54–66.
82 Chaurasia, P., Sen, R., and Bhaumik, S.R. (2013) Functional analysis of Rad14p, a DNA damage recognition factor in nucleotide excision repair, in regulation of transcription *in vivo*. *J. Biol. Chem.*, **288**, 793–806.
83 Selth, L.A., Sigurdsson, S., and Svejstrup, J.Q. (2010) Transcript elongation by RNA polymerase II. *Annu. Rev. Biochem.*, **79**, 271–293.
84 Hsin, J.P. and Manley, J.L. (2012) The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes Dev.*, **26**, 2119–2137.
85 Conaway, R.C. and Conaway, J.W. (2013) The Mediator complex and transcription elongation. *Biochim. Biophys. Acta*, **1829**, 69–75.
86 Dujardin, G., Lafaille, C., Petrillo, E., Buggiano, V., Gómez Acuña, L.I., Fiszbein, A., Godoy Herz, M.A., Nieto Moreno, N., Muñoz, M.J., Alló, M., Schor, I.E., and Kornblihtt, A.R. (2013) Transcriptional elongation and alternative splicing. *Biochim. Biophys. Acta*, **1829**, 134–140.
87 Zhou, J., Schweikhard, V., and Block, S.M. (2013) Single-molecule studies of RNAPII elongation. *Biochim. Biophys. Acta*, **1829**, 29–38.
88 Kulaeva, O.I., Hsieh, F.K., Chang, H.W., Luse, D.S., and Studitsky, V.M. (2013) Mechanism of transcription through a nucleosome by RNA polymerase II. *Biochim. Biophys. Acta*, **1829**, 76–83.
89 Proudfoot, N.J. (1989) How RNA polymerase II terminates transcription in higher eukaryotes. *Trends Biochem. Sci.*, **14**, 105–110.
90 Birse, C.E., Minvielle-Sebastia, L., Lee, B.A., Keller, W., and Proudfoot, N.J.

(1998) Coupling termination of transcription to messenger RNA maturation in yeast. *Science*, **280**, 298–301.

91 Richard, P. and Manley, J.L. (2009) Transcription termination by nuclear RNA polymerases. *Genes Dev.*, **23**, 1247–1269.

92 Barilla, D., Lee, B.A., and Proudfoot, N.J. (2001) Cleavage/polyadenylation factor IA associates with the carboxyl-terminal domain of RNA polymerase II in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA*, **98**, 445–450.

93 Licatalosi, D.D., Geiger, G., Minet, M., Schroeder, S., Cilli, K., McNeil, J.B., and Bentley, D.L. (2002) Functional interaction of yeast pre-mRNA 3′ end processing factors with RNA polymerase II. *Mol. Cell*, **9**, 1101–1111.

94 Zhang, Z., Fu, J., and Gilmour, D.S. (2005) CTD-dependent dismantling of the RNA polymerase II elongation complex by the pre-mRNA 3′-end processing factor, Pcf11. *Genes Dev.*, **19**, 1572–1580.

95 Kim, M., Krogan, N.J., Vasiljeva, L., Rando, O.J., Nedea, E., Greenblatt, J.F., and Buratowski, S. (2004) The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature*, **432**, 517–522.

96 West, S., Gromak, N., and Proudfoot, N.J. (2004) Human 5′ → 3′ exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. *Nature*, **432**, 522–525.

97 Hyman, L.E. and Moore, C.L. (1993) Termination and pausing of RNA polymerase II downstream of yeast polyadenylation sites. *Mol. Cell Biol.*, **13**, 5159–5167.

98 Lahudkar, S., Shukla, A., Bajwa, P., Durairaj, G., Stanojevic, N., and Bhaumik, S.R. (2011) The mRNA cap-binding complex stimulates the formation of pre-initiation complex at the promoter via its interaction with Mot1p *in vivo*. *Nucleic Acids Res.*, **39**, 2188–2209.

99 Stutz, F. and Izaurralde, E. (2003) The interplay of nuclear mRNP assembly, mRNA surveillance and export. *Trends Cell Biol.*, **13**, 319–327.

100 Lahudkar, S., Durairaj, G., Uprety, B., and Bhaumik, S.R. (2014) A Novel role for Cet1p mRNA 5′-triphosphatase in promoter proximal accumulation of RNA polymerase II in *Saccharomyces cerevisiase*. *Genetics*, **196**, 161–.

101 Komarnitsky, P., Cho, E.J., and Buratowski, S. (2000) Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev.*, **14**, 2452–2460.

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

103 Aguilera, A. (2005) Cotranscriptional mRNP assembly: from the DNA to the nuclear pore. *Curr. Opin. Cell Biol.*, **17**, 242–250.

104 Huang, Y.Q. and Carmichael, G.G. (1996) Role of polyadenylation in nucleocytoplasmic transport of mRNA. *Mol. Cell. Biol.*, **16**, 1534–1542.

105 Proudfoot, N. and O'Sullivan, J. (2002) Polyadenylation: a tail of two complexes. *Curr. Biol.*, **12**, R855–R857.

106 Mandel, C.R., Bai, Y., and Tong, L. (2008) Protein factors in pre-mRNA 3′-end processing. *Cell. Mol. Life Sci.*, **65**, 1099–1122.

107 Wahle, E. and Ruegsegger, U. (1999) 3′-End processing of pre-mRNA in eukaryotes. *FEMS Microbiol. Rev.*, **23**, 277–295.

108 Marzluff, W.F. (2010) More than one way to make a tail. *EMBO J.*, **29**, 4066–4067.

109 Hoskins, A.A. and Moore, M.J. (2012) The spliceosome: a flexible, reversible macromolecular machine. *Trends Biochem. Sci.*, **37**, 179–188.

110 Russell, A.G., Charette, J.M., Spencer, D.F., and Gray, M.W. (2006) An early evolutionary origin for the minor spliceosome. *Nature*, **443**, 863–866.

111 Steitz, J.A., Dreyfuss, G., Krainer, A.R., Lamond, A.I., Matera, A.G., and Padgett, R.A. (2008) Where in the cell is the minor spliceosome? *Proc. Natl. Acad. Sci. USA*, **105**, 8485–8486.

112 Patel, A.A. and Steitz, J.A. (2003) Splicing double: Insights from the second spliceosome. *Nat. Rev. Mol. Cell Biol.*, **4**, 960–970.

113 Jurica, M.S. and Moore, M.J. (2003) Pre-mRNA splicing: awash in a sea of proteins. *Mol. Cell*, **12**, 5–14.

114 Shen, H. (2009) UAP56-a key player with surprisingly diverse roles in pre-mRNA splicing and nuclear export. *BMB Rep.*, **42**, 185–188.

115 Stevens, S.W., Ryan, D.E., Ge, H.Y., Moore, R.E., Young, M.K., Lee, T.D., and Abelson, J. (2002) Composition and functional characterization of the yeast spliceosomal penta-snRNP. *Mol. Cell*, **9**, 31–44.

116 Bergkessel, M., Wilmes, G.M., and Guthrie, C. (2009) SnapShot: formation of mRNPs. *Cell*, **136**, 794–U791.

117 Cordin, O., Banroques, J., Tanner, N.K., and Linder, P. (2006) The DEAD-box protein family of RNA helicases. *Gene*, **367**, 17–37.

118 Fairman, M.E., Maroney, P.A., Wang, W., Bowers, H.A., Gollnick, P., Nilsen, T.W., and Jankowsky, E. (2004) Protein displacement by DExH/D "RNA helicases" without duplex unwinding. *Science*, **304**, 730–734.

119 Le Hir, H., Izaurralde, E., Maquat, L.E., and Moore, M.J. (2000) The spliceosome deposits multiple proteins 20–24 nucleotides upstream of mRNA exon-exon junctions. *EMBO J.*, **19**, 6860–6869.

120 Le Hir, H., Gatfield, D., Izaurralde, E., and Moore, M.J. (2001) The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *EMBO J.*, **20**, 4987–4997.

121 McManus, C.J. and Graveley, B.R. (2011) RNA structure and the mechanisms of alternative splicing. *Curr. Opin. Genet. Dev.*, **21**, 373–379.

122 Kim, E., Goren, A., and Ast, G. (2008) Alternative splicing: current perspectives. *Bioassays*, **30**, 38–47.

123 Pante, N. (2004) Nuclear pore complex structure: unplugged and dynamic pores. *Dev. Cell*, **7**, 780–781.

124 D'Angelo, M.A. and Hetzer, M.W. (2008) Structure, dynamics and function of nuclear pore complexes. *Trends Cell Biol.*, **18**, 456–466.

125 Tran, E.J. and Wente, S.R. (2006) Dynamic nuclear pore complexes: life on the edge. *Cell*, **125**, 1041–1053.

126 Lim, R.Y.H. and Fahrenkrog, B. (2006) The nuclear pore complex up close. *Curr. Opin. Cell Biol.*, **18**, 342–347.

127 Adam, S.A. (2001) The nuclear pore complex. *Genome. Biol.*, **2**, REVIEWS0007.

128 Lyman, S.K. and Gerace, L. (2001) Nuclear pore complexes: dynamics in unexpected places. *J. Cell Biol.*, **154**, 17–20.

129 Kiseleva, E., Goldberg, M.W., Cronshaw, J., and Allen, T.D. (2000) The nuclear pore complex: structure, function, and dynamics. *Crit. Rev. Eukaryot. Gene Expr.*, **10**, 101–112.

130 Stewart, M. (2007) Ratcheting mRNA out of the nucleus. *Mol. Cell*, **25**, 327–330.

131 Cole, C.N. and Scarcelli, J.J. (2006) Transport of messenger RNA from the nucleus to the cytoplasm. *Curr. Opin. Cell Biol.*, **18**, 299–306.

132 Weis, K. (2002) Nucleocytoplasmic transport: cargo trafficking across the border. *Curr. Opin. Cell Biol.*, **14**, 328–335.

133 Conti, E. and Izaurralde, E. (2001) Nucleocytoplasmic transport enters the atomic age. *Curr. Opin. Cell Biol.*, **13**, 310–319.

134 Mattaj, I.W. and Englmeier, L. (1998) Nucleocytoplasmic transport: the soluble phase. *Annu. Rev. Biochem.*, **67**, 265–306.

135 Nakielny, S. and Dreyfuss, G. (1999) Transport of proteins and RNAs in and out of the nucleus. *Cell*, **99**, 677–690.

136 Izaurralde, E. (2002) Nuclear export of messenger RNA. *Results Probl. Cell Differ.*, **35**, 133–150.

137 Cullen, B.R. (2000) Nuclear RNA export pathways. *Mol. Cell. Biol.*, **20**, 4181–4187.

138 Cullen, B.R. (2003) Nuclear RNA export. *J. Cell Sci.*, **116**, 587–597.

139 Pemberton, L.F. and Paschal, B.M. (2005) Mechanisms of receptor-mediated nuclear import and nuclear export. *Traffic*, **6**, 187–198.

140 Herold, A., Suyama, M., Rodrigues, J.P., Braun, I.C., Kutay, U., Carmo-Fonseca,

M., Bork, P., and Izaurralde, E. (2000) TAP (NXF1) belongs to a multigene family of putative RNA export factors with a conserved modular architecture. *Mol. Cell. Biol.*, **20**, 8996–9008.

141 Segref, A., Sharma, K., Doye, V., Hellwig, A., Huber, J., Luhrmann, R., and Hurt, E. (1997) Mex67p, a novel factor for nuclear mRNA export, binds to both poly(A)(+) RNA and nuclear pores. *EMBO J.*, **16**, 3256–3271.

142 Tan, W., Zolotukhin, A.S., Bear, J., Patenaude, D.J., and Felber, B.K. (2000) The mRNA export in *Caenorhabditis elegans* is mediated by Ce-NXF-1, an ortholog of human TAP/NXF and *Saccharomyces cerevisiae* Mex67p. *RNA*, **6**, 1762–1772.

143 Yang, J., Bogerd, H.P., Wang, P.J., Page, D.C., and Cullen, B.R. (2001) Two closely related human nuclear export factors utilize entirely distinct export pathways. *Mol. Cell*, **8**, 397–406.

144 Santos-Rosa, H., Moreno, H., Simos, G., Segref, A., Fahrenkrog, B., Pante, N., and Hurt, E. (1998) Nuclear mRNA export requires complex formation between Mex67p and Mtr2p at the nuclear pores. *Mol. Cell. Biol.*, **18**, 6826–6838.

145 Strasser, K., Bassler, J., and Hurt, E. (2000) Binding of the Mex67p/Mtr2p heterodimer to FXFG, GLFG, and FG repeat nucleoporins is essential for nuclear mRNA export. *J. Cell Biol.*, **150**, 695–706.

146 Gruter, P., Tabernero, C., von Kobbe, C., Schmitt, C., Saavedra, C., Bachi, A., Wilm, M., Felber, B.K., and Izaurralde, E. (1998) TAP, the human homolog of Mex67p, mediates CTE-dependent RNA export from the nucleus. *Mol. Cell*, **1**, 649–659.

147 Braun, I.C., Rohrbach, E., Schmitt, C., and Izaurralde, E. (1999) TAP binds to the constitutive transport element (CTE) through a novel RNA-binding motif that is sufficient to promote CTE-dependent RNA export from the nucleus. *EMBO J.*, **18**, 1953–1965.

148 Lund, M.K. and Guthrie, C. (2005) The DEAD-Box protein Dbp5p is required to dissociate Mex67p from exported mRNPs at the nuclear rim. *Mol. Cell*, **20**, 645–651.

149 Strasser, K. and Hurt, E. (2000) Yra1p, a conserved nuclear RNA-binding protein, interacts directly with Mex67p and is required for mRNA export. *EMBO J.*, **19**, 410–420.

150 Stutz, F., Bachi, A., Doerks, T., Braun, I.C., Seraphin, B., Wilm, M., Bork, P., and Izaurralde, E. (2000) REF, an evolutionarily conserved family of hnRNP-like proteins, interacts with TAP/Mex67p and participates in mRNA nuclear export. *RNA*, **6**, 638–650.

151 Kashyap, A.K., Schieltz, D., Yates, J., and Kellogg, D.R. (2005) Biochemical and genetic characterization of Yra1p in budding yeast. *Yeast*, **22**, 43–56.

152 Keene, J.D. (2003) Organizing mRNA export. *Nat. Genet.*, **33**, 111–112.

153 Tutucci, E. and Stutz, F. (2011) Keeping mRNPs in check during assembly and nuclear export. *Nat. Rev. Mol. Cell Biol.*, **12**, 376–383.

154 Strasser, K. and Hurt, E. (2001) Splicing factor Sub2p is required for nuclear mRNA export through its interaction with Yra1p. *Nature*, **413**, 648–652.

155 Chavez, S., Beilharz, T., Rondon, A.G., Erdjument-Bromage, H., Tempst, P., Svejstrup, J.Q., Lithgow, T., and Aguilera, A. (2000) A protein complex containing Tho2, Hpr1, Mft1 and a novel protein, Thp2, connects transcription elongation with mitotic recombination in *Saccharomyces cerevisiae*. *EMBO J.*, **19**, 5824–5834.

156 Rondon, A.G., Jimeno, S., Garcia-Rubio, M., and Aguilera, A. (2003) Molecular evidence that the eukaryotic THO/TREX complex is required for efficient transcription elongation. *J. Biol. Chem.*, **278**, 39037–39043.

157 Jimeno, S., Rondon, A.G., Luna, R., and Aguilera, A. (2002) The yeast THO complex and mRNA export factors link RNA metabolism with transcription and genome instability. *EMBO J.*, **21**, 3526–3535.

158 Strasser, K., Masuda, S., Mason, P., Pfannstiel, J., Oppizzi, M., Rodriguez-Navarro, S., Rondon, A.G., Aguilera, A., Struhl, K., Reed, R. *et al.* (2002) TREX is a conserved complex coupling

transcription with messenger RNA export. *Nature*, **417**, 304–308.

159 Luna, R., Rondon, A.G., and Aguilera, A. (2012) New clues to understand the role of THO and other functionally related factors in mRNP biogenesis. *Biochim. Biophys. Acta, Gene Regul. Mech.*, **1819**, 514–520.

160 Katahira, J. (2012) mRNA export and the TREX complex. *Biochim. Biophys. Acta, Gene Regul. Mech.*, **1819**, 507–513.

161 Kelly, S.M. and Corbett, A.H. (2009) Messenger RNA export from the nucleus: a series of molecular wardrobe changes. *Traffic*, **10**, 1199–1208.

162 Viphakone, N., Hautbergue, G.M., Walsh, M., Chang, C.-T., Holland, A., Folco, E.G., Reed, R., and Wilson, S.A. (2012) TREX exposes the RNA-binding domain of Nxf1 to enable mRNA export. *Nat. Commun.*, **3**, 1006.

163 Masuda, S., Das, R., Cheng, H., Hurt, E., Dorman, N., and Reed, R. (2005) Recruitment of the human TREX complex to mRNA during splicing. *Genes Dev.*, **19**, 1512–1517.

164 Castellano-Pozo, M., Garcia-Muse, T., and Aguilera, A. (2012) R-loops cause replication impairment and genome instability during meiosis. *EMBO Rep.*, **13**, 923–929.

165 Aguilera, A. and Garcia-Muse, T. (2012) R loops: from transcription byproducts to threats to genome stability. *Mol. Cell*, **46**, 115–124.

166 Dominguez-Sanchez, M.S., Barroso, S., Gomez-Gonzalez, B., Luna, R., and Aguilera, A. (2011) Genome instability and transcription elongation impairment in human cells depleted of THO/TREX. *PloS Genet.*, **7**, e1002386.

167 Jimeno, S. and Aguilera, A. (2010) The THO complex as a key mRNP biogenesis factor in development and cell differentiation. *J. Biol. (London)*, **9**, 6.

168 Chanarat, S., Seizl, M., and Straesser, K. (2011) The Prp19 complex is a novel transcription elongation factor required for TREX occupancy at transcribed genes. *Genes Dev.*, **25**, 1147–1158.

169 Reed, R. and Cheng, H. (2005) TREX, SR proteins and export of mRNA. *Curr. Opin. Cell Biol.*, **17**, 269–273.

170 Zenklusen, D., Vinciguerra, P., Wyss, J.C., and Stutz, F. (2002) Stable mRNP formation and export require cotranscriptional recruitment of the mRNA export factors Yra1p and Sub2p by Hpr1p. *Mol. Cell. Biol.*, **22**, 8241–8253.

171 Masuda, S., Das, R., Cheng, H., Hurt, E., Dorman, N., and Reed, R. (2005) Recruitment of the human TREX complex to mRNA during splicing. *Genes Dev.*, **19**, 1512–1517.

172 Chi, B., Wang, Q., Wu, G., Tan, M., Wang, L., Shi, M., Chang, X., and Cheng, H. (2013) Aly and THO are required for assembly of the human TREX complex and association of TREX components with the spliced mRNA. *Nucleic Acids Res.*, **41**, 1294–1306.

173 Johnson, S.A., Kim, H., Erickson, B., and Bentley, D.L. (2011) The export factor Yra1 modulates mRNA 3′ end processing. *Nat. Struct. Mol. Biol.*, **18**, 1164–U1104.

174 Johnson, S.A., Cubberley, G., and Bentley, D.L. (2009) Cotranscriptional Recruitment of the mRNA export factor Yra1 by direct interaction with the 3′ end processing factor Pcf11. *Mol. Cell*, **33**, 215–226.

175 Hobeika, M., Brockmann, C., Gruessing, F., Neuhaus, D., Divita, G., Stewart, M., and Dargemont, C. (2009) Structural Requirements for the ubiquitin-associated domain of the mRNA export factor Mex67 to bind its specific targets, the transcription elongation THO complex component Hpr1 and nucleoporin FXFG repeats. *J. Biol. Chem.*, **284**, 17575–17583.

176 Hobeika, M., Brockmann, C., Iglesias, N., Gwizdek, C., Neuhaus, D., Stutz, F., Stewart, M., Divita, G., and Dargemont, C. (2007) Coordination of Hpr1 and ubiquitin binding by the UBA domain of the mRNA export factor mex67. *Mol. Biol. Cell*, **18**, 2561–2568.

177 Longman, D., Johnstone, I.L., and Caceres, J.F. (2003) The Ref/Aly proteins are dispensable for mRNA export and development in *Caenorhabditis elegans*. *RNA*, **9**, 881–891.

178 Gatfield, D. and Izaurralde, E. (2002) REF1/Aly and the additional exon

junction complex proteins are dispensable for nuclear mRNA export. *J. Cell Biol.*, **159**, 579–588.

179 Hargous, Y., Hautbergue, G.M., Tintaru, A.M., Skrisovska, L., Golovanov, A.P., Stevenin, J., Lian, L.-Y., Wilson, S.A., and Allain, F.H.T. (2006) Molecular basis of RNA recognition and TAP binding by the SR proteins SRp20 and 9G8. *EMBO J.*, **25**, 5126–5137.

180 Huang, Y.Q. and Steitz, J.A. (2001) Splicing factors SRp20 and 9G8 promote the nucleocytoplasmic export of mRNA. *Mol. Cell*, **7**, 899–905.

181 Escudero-Paunetto, L., Li, L., Hernandez, F.P., and Sandri-Goldin, R.M. (2010) SR proteins SRp20 and 9G8 contribute to efficient export of herpes simplex virus 1 mRNAs. *Virology*, **401**, 155–164.

182 Caceres, J.F., Screaton, G.R., and Krainer, A.R. (1998) A specific subset of SR proteins shuttles continuously between the nucleus and the cytoplasm. *Genes Dev.*, **12**, 55–66.

183 Pascual-Garcia, P. and Rodriguez-Navarro, S. (2009) A tale of coupling, Sus1 function in transcription and mRNA export. *RNA Biol.*, **6**, 141–144.

184 Rodriguez-Navarro, S., Fischer, T., Luo, M.J., Antunez, O., Brettschneider, S., Lechner, J., Perez-Ortin, J.E., Reed, R., and Hurt, E. (2004) Sus1, a functional component of the SAGA histone acetylase complex and the nuclear pore-associated mRNA export machinery. *Cell*, **116**, 75–86.

185 Fischer, T., Strasser, K., Racz, A., Rodriguez-Navarro, S., Oppizzi, M., Ihrig, P., Lechner, J., and Hurt, E. (2002) The mRNA export machinery requires the novel Sac3p-Thp1p complex to dock at the nucleoplasmic entrance of the nuclear pores. *EMBO J.*, **21**, 5843–5852.

186 Wilmes, G.M. and Guthrie, C. (2009) Getting to the gate: crystallization of a Sac3(CID):Sus1:Cdc31 complex. *Mol. Cell*, **33**, 671–672.

187 Jani, D., Lutz, S., Marshall, N.J., Fischer, T., Koehler, A., Ellisdon, A.M., Hurt, E., and Stewart, M. (2009) Sus1, Cdc31, and the Sac3 CID region form a conserved interaction platform that promotes nuclear pore association and mRNA export. *Mol. Cell*, **33**, 727–737.

188 García-Oliver, E., García-Molinero, V., and Rodríguez-Navarro, S. (2012) mRNA export and gene expression: the SAGA-TREX-2 connection. *Biochim. Biophys. Acta*, **1819**, 555–565.

189 Faza, M.B., Kemmler, S., Jimeno, S., Gonzalez-Aguilera, C., Aguilera, A., Hurt, E., and Panse, V.G. (2009) Sem1 is a functional component of the nuclear pore complex-associated messenger RNA export machinery. *J. Cell Biol.*, **184**, 833–846.

190 Garcia-Oliver, E., Pascual-Garcia, P., Garcia-Molinero, V., Lenstra, T.L., Holstege, F.C.P., and Rodriguez-Navarro, S. (2013) A novel role for Sem1 and TREX-2 in transcription involves their impact on recruitment and H2B deubiquitylation activity of SAGA. *Nucleic Acids Res.*, **41**, 5655–5668.

191 Blobel, G. (1985) Gene gating – a hypothesis. *Proc. Natl. Acad. Sci. USA*, **82**, 8527–8529.

192 Taddei, A., Van Houwe, G., Hediger, F., Kalck, V., Cubizolles, F., and Schober, H., and Gasser, S.M. (2006) Nuclear pore association confers optimal expression levels for an inducible yeast gene. *Nature*, **441**, 774–778.

193 Taddei, A. (2007) Active genes at the nuclear pore complex. *Curr. Opin. Cell. Biol.*, **19**, 305–310.

194 Cabal, G.G., Genovesio, A., Rodriguez-Navarro, S., Zimmer, C., Gadal, O., Lesne, A., Buc, H., Feuerbach-Fournier, F., Olivo-Marin, J.-C., Hurt, E.C. *et al.* (2006) SAGA interacting factors confine sub-diffusion of transcribed genes to the nuclear envelope. *Nature*, **441**, 770–773.

195 Casolari, J.M., Brown, C.R., Komili, S., West, J., Hieronymus, H., and Silver, P.A. (2004) Genome-wide localization of the nuclear transport machinery couples transcriptional status and nuclear organization. *Cell*, **117**, 427–439.

196 Faria, A.M.C., Levay, A., Wang, Y., Kamphorst, A.O., Rosa, M.L.P., Nussenzveig, D.R., Balkan, W., Chook, Y.M., Levy, D.E., and Fontoura, B.M.A. (2006) The nucleoporin Nup96 is required for proper expression of

interferon-regulated proteins and functions. *Immunity*, **24**, 295–304.

197 Mendjan, S., Taipale, M., Kind, J., Holz, H., Gebhardt, P., Schelder, M., Vermeulen, M., Buscaino, A., Duncan, K., and Mueller, J. (2006) Nuclear pore components are involved in the transcriptional regulation of dosage compensation in *Drosophila*. *Mol. Cell*, **21**, 811–823.

198 Sommer, P. and Nehrbass, U. (2005) Quality control of messenger ribonucleoprotein particles in the nucleus and at the pore. *Curr. Opin. Cell Biol.*, **17**, 294–301.

199 Mitchell, P. and Tollervey, D. (2000) Musing on the structural organization of the exosome complex. *Nat. Struct. Biol.*, **7**, 843–846.

200 Gao, Q., Das, B., Sherman, F., and Maquat, L.E. (2005) The pioneer round of translation: features and functions, Cap-binding protein 1-mediated and eukaryotic translation initiation factor 4E-mediated pioneer rounds of translation in yeast. *Proc. Natl. Acad. Sci. USA*, **102**, 4258–4263.

201 Sachs, A.B. and Varani, G. (2000) Eukaryotic translation initiation: there are (at least) two sides to every story. *Nat. Struct. Biol.*, **7**, 356–361.

202 Mangus, D.A., Evans, M.C., and Jacobson, A. (2003) Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biol.*, **4**, 223.

203 Houseley, J. and Tollervey, D. (2009) The many pathways of RNA degradation. *Cell*, **136**, 763–776.

204 Fillman, C. and Lykke-Andersen, J. (2005) RNA decapping inside and outside of processing bodies. *Curr. Opin. Cell Biol.*, **17**, 326–331.

205 Eulalio, A., Behm-Ansmant, I., and Izaurralde, E. (2007) P bodies: at the crossroads of post-transcriptional pathways. *Nat. Rev. Mol. Cell Biol.*, **8**, 9–22.