



Published in final edited form as:

Gene. 2007 July 15; 396(2): 373–390. doi:10.1016/j.gene.2007.04.021.

Formation of the 3' end of histone mRNA: Getting closer to the end

Zbigniew Dominski* and William F. Marzluff

Department of Biochemistry and Biophysics and Program in Molecular Biology and Biotechnology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

Abstract

Nearly all eukaryotic mRNAs end with a poly (A) tail that is added to their 3' end by the ubiquitous cleavage/polyadenylation machinery. The only known exception to this rule are metazoan replication dependent histone mRNAs, which end with a highly conserved stem-loop structure. This distinct 3' end is generated by specialized 3' end processing machinery that cleaves histone pre-mRNAs 4–5 nucleotides downstream of the stem-loop and consists of the U7 small nuclear RNP (snRNP) and number of protein factors. Recently, the U7 snRNP has been shown to contain a unique Sm core that differs from that of the spliceosomal snRNPs, and an essential heat labile processing factor has been identified as symplekin. In addition, cross-linking studies have pinpointed CPSF-73 as the endonuclease, which catalyzes the cleavage reaction. Thus, many of the critical components of the 3' end processing machinery are now identified. Strikingly, this machinery is not as unique as initially thought but contains a number of factors involved in cleavage/polyadenylation, suggesting that the two mechanisms have a common evolutionary origin. The greatest challenge that lies ahead is to determine how all these factors interact with each other to form a catalytically competent processing complex capable of cleaving histone pre-mRNAs.

1. Introduction

The presence of a poly(A) tail at the 3' end is one of the most recognizable features of eukaryotic mRNAs. Formation of this 3' end occurs in a two step reaction directed by the cleavage/polyadenylation machinery and depends on the presence of two sequence elements in mRNA precursors (pre-mRNAs); a highly conserved hexanucleotide AAUAAA and a downstream G/U-rich sequence (Colgan and Manley, 1997; Wahle and Ruegsegger, 1999; Zhao et al., 1999a). In the first step, pre-mRNAs are cleaved between these two elements. In the second tightly coupled step the newly formed 3' end is extended by addition of a poly(A) tail consisting of 200–250 adenosines. Completion of both steps *in vitro* depends on at least 12 protein factors that engage in a network of interactions that serve to properly juxtapose the cleavage site with the catalytic components of the processing complex.

The poly(A) tail affects virtually all aspects of mRNA metabolism, including mRNA export, stability and translation (Kuhn and Wahle, 2004; Mangus et al., 2003). Yet, one class of eukaryotic mRNAs, the metazoan replication-dependent histone mRNAs encoding all five histone proteins, do not contain the poly(A) tail and instead end with a highly conserved stem-loop. This unique 3' end is formed by specialized processing machinery that recognizes two

© 2007 Elsevier B.V. All rights reserved.

*To whom correspondence should be addressed: Program in Molecular Biology and Biotechnology, CB #3280, University of North Carolina, Chapel Hill, NC 27599, Phone: (919) 962-2141; FAX: (919) 962-1274, dominski@med.unc.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

characteristic features present in histone pre-mRNAs, the stem-loop and a purine-rich histone downstream element (HDE) (Dominski and Marzluff, 1999). Cleavage occurs between these two elements and is not followed by addition of the poly(A) tail. Since histone genes do not contain introns, this single cleavage reaction is sufficient to form mature histone mRNAs that are subsequently exported to the cytoplasm and translated to generate histone proteins.

Synthesis of all histone proteins increases dramatically during S phase reflecting the necessity to rapidly package the newly replicated DNA into chromatin (Marzluff, 2005; Osley, 1991). In mammalian cells, the increased synthesis of histone proteins during S phase is achieved by a combination of both transcriptional activation of histone genes (Stein et al., 1996) and posttranscriptional mechanisms among which the unusual mechanism of 3' end processing of histone pre-mRNAs and the resultant 3' end play the critical role (Harris et al., 1991). During 3' end processing the HDE interacts with the U7 snRNP and the stem-loop is recognized by a protein termed the stem-loop binding protein (SLBP) (Dominski and Marzluff, 1999). After completion of processing, SLBP remains associated with the stem-loop at the end of mature histone mRNAs and stimulates their translation into histone proteins in the cytoplasm (Gorgoni et al., 2005; Sanchez and Marzluff, 2002). The intracellular concentration of SLBP reaches the highest level during S phase and rapidly declines during the S/G2 transition due to proteasome degradation (Whitfield et al., 2000). The disappearance of SLBP from the cell is concurrent with degradation of the existing pool of histone mRNAs, suggesting that these two events are functionally related. Thus, SLBP likely functions as a master regulator that turns on 3' end processing of histone pre-mRNAs, activates translation of the mature histone mRNA during S phase and finally triggers their degradation during the S/G2 transition. This post-transcriptional regulation results in coordinated synthesis of all classes of histones at the time of DNA replication and effectively eliminates expression of histone genes outside S phase. All mRNAs that do not end with the stem-loop are excluded from this global cell cycle regulation. These mRNAs are instead subject to more individual regulatory mechanisms mediated by various sequence elements primarily located within their 3' untranslated region (UTR) (Guhaniyogi and Brewer, 2001; Mazumder et al., 2003; Ross, 1995; Wickens et al., 2002).

The differences between sequence elements and factors involved in formation of polyadenylated mRNAs and histone mRNAs for many years justified the view that these two processing machineries have nothing in common besides dealing with the same end (Gilmartin, 2005). Perhaps, the most obvious reason for this distinction was the involvement of U7 snRNA in formation of histone pre-mRNAs and the lack of any RNA component in cleavage/polyadenylation. Surprisingly, recent studies demonstrated that these two processing machineries share at least two factors, and the number of common components may be even higher (Dominski et al., 2005a; Gilmartin, 2005; Kolev and Steitz, 2005; Weiner, 2005). This review summarizes what is currently known about 3' end processing of histone pre-mRNAs and emphasizes important breakthroughs that have been made in recent years in studying factors and mechanisms involved in this unique processing event. Since mechanisms involved in formation of histone pre-mRNAs and polyadenylated mRNAs are not as different as initially thought, the review will begin with a description of the cleavage/polyadenylation machinery. For more comprehensive information on other aspects of histone gene expression the reader is referred to a number of previously published reviews (Dominski and Marzluff, 1999; Jaeger et al., 2005; Marzluff, 2005; Marzluff and Duronio, 2002; Muller and Schumperli, 1997; Osley, 1991; Schumperli and Pillai, 2004).

2. 3' end processing by cleavage and polyadenylation

The vast majority of eukaryotic pre-mRNAs are processed at the 3' end by cleavage coupled to polyadenylation (Colgan and Manley, 1997; Wahle and Rieger, 1999; Zhao et al., 1999a). Although these two steps are intimately associated *in vivo*, they can be uncoupled in

vitro by a number of reagents, including EDTA, which at low concentrations inhibits polyadenylation but has no major effect on the cleavage reaction (Hirose and Manley, 1997). In mammalian cell, cleavage and polyadenylation depend on the presence of a highly conserved polyadenylation signal, AAUAAA, and a downstream G/U-rich sequence (Fig. 1). The two elements are often located several thousand nucleotides downstream of the stop codon (Mazumder et al., 2003; Mignone et al., 2002; Pesole et al., 2002). Pre-mRNA is cleaved between the two sequence elements, usually after a CA located 10–30 nucleotides downstream of the AAUAAA, and the newly generated 3' end is extended by addition of the poly(A) tail.

2.1. Factors involved in cleavage and polyadenylation

The cleavage reaction in mammalian nuclear extracts requires at least 12 protein components (Fig. 1). The AAUAAA is recognized by cleavage/polyadenylation specificity factor (CPSF), which consists of 5 subunits: CPSF-160, CPSF-100, CPSF-73, CPSF-30 and Fip1 (Kaufmann et al., 2004). CPSF-160 directly binds the AAUAAA element (Murthy and Manley, 1995), whereas CPSF-30 and Fip1 bind other regions in pre-mRNA, thus improving the recognition of the AAUAAA by CPSF-160 (Kaufmann et al., 2004). The downstream G/U-rich sequence is recognized by cleavage stimulation factor (CstF), which consists of three subunits, CstF-77, CstF-64 and CstF-50, with the CstF-64 subunit directly binding this RNA element (MacDonald et al., 1994). Under in vivo conditions CPSF and CstF may form a larger complex also containing symplekin, which is similar to the yeast protein PTA1 shown to be an essential component of the yeast cleavage/polyadenylation machinery (Zhao et al., 1999b). In mammalian cells symplekin may function as a platform that brings together various processing factors into a single entity (Takagaki and Manley, 2000). In *Xenopus* oocytes, symplekin is highly concentrated in the nuclear structures called Cajal bodies and is also present in the cytoplasm (Hofmann et al., 2002) where it is required for polyadenylation and thus translational activation of dormant mRNAs (Barnard et al., 2004). It is uncertain whether symplekin is directly required for cleavage/polyadenylation in mammalian nuclear extracts (Takagaki and Manley, 2000).

The in vitro cleavage reaction also depends on cleavage factor I (CF I_m) consisting of at least two subunits (Ruegsegger et al., 1996; Ruegsegger et al., 1998), and the poorly characterized mammalian cleavage factor II (CF II_m) (De Vries et al., 2000). CF I_m binds to the UGUA sequence often present in multiple copies upstream of the cleavage site and stimulates binding of CPSF to pre-mRNA through a functional interaction with Fip1 (Brown and Gilmartin, 2003; Venkataraman et al., 2005). This interaction is critically important in processing of pre-mRNAs lacking the canonical AAUAAA that constitute nearly 50% of all mammalian pre-mRNAs (Tian et al., 2005). Finally, the cleavage step depends on the C-terminal domain (CTD) of the large subunit of RNA polymerase II (Hirose and Manley, 1998), and poly(A) polymerase (PAP) (Christofori and Keller, 1989; Takagaki et al., 1988; Terns and Jacob, 1989). The requirement for the CTD may be circumvented by addition of creatine phosphate, which is routinely used in processing reactions reconstituted from purified components (Hirose and Manley, 1997). Following cleavage, the newly formed 3' end is extended by addition of a poly(A) tail. The polyadenylation reaction is catalyzed by PAP. This step in 3' end processing depends on CPSF, and on poly(A)-binding protein II, which associates with the nascent poly(A) sequence (Wahle, 1995).

A number of interactions have been identified between individual components of the processing complex (Colgan and Manley, 1997; Zhao et al., 1999a). CPSF-160 interacts with CstF-77, thus significantly improving binding of each protein to its RNA target through the co-operative mechanism (Murthy and Manley, 1995). CPSF-160 and Fip1 interact with PAP and these interactions are important for the polyadenylation step by stabilizing PAP on the upstream cleavage product (Kaufmann et al., 2004; Murthy and Manley, 1995). Within CstF,

CstF-77 makes contacts with each of the two remaining subunits of the complex, and CstF-64 interacts with symplekin (Takagaki and Manley, 2000). CstF-50 on the other hand interacts with the CTD, reflecting the tight coupling between transcription and 3' end processing (McCracken et al., 1997). Within CPSF, CPSF-73 through its C-terminal portion interacts with CPSF-100 (Dominski et al., 2005c). Nothing is known about other interactions that occur within this factor.

2.2. Identification of the endonuclease in cleavage coupled to polyadenylation

The identity of the endonuclease in cleavage coupled to polyadenylation remained unknown until recently and has been the subject of an intensive search driven by the presumption that it was a yet unrecognized component of the processing machinery (Wickens and Gonzalez, 2004). However, as early as 1999 it was suggested based on sequence analysis that the cleavage reaction might be catalyzed by CPSF-73 (Aravind, 1999). CPSF-73 contains a metallo- β -lactamase domain that is shared by a large family of enzymes active on a broad spectrum of substrates, including nucleic acids (Aravind, 1999). Proteins of the metallo- β -lactamase family are typically zinc-dependent hydrolases, as exemplified by the founding members of the family, class B β -lactamases that utilize zinc ions to hydrolyze and thus inactivate β -lactam antibiotics (Heinz and Adolph, 2004; Wang et al., 1999a). The family also includes tRNase Z, the best known metallo- β -lactamase nuclease, which is involved in 3' end processing of tRNA precursors in both prokaryotes and eukaryotes (Vogel et al., 2005). CPSF-73 contains all essential residues of the metallo- β -lactamase domain required for zinc binding and catalysis, including the histidine motif HxHxDH, which is the hallmark of the entire metallo- β -lactamase family (Aravind, 1999; Wang et al., 1999a). Sequence analysis and recent structural studies revealed that CPSF-73 and most nucleases of the metallo- β -lactamase family differ from the canonical members of the family as they have an additional domain called β -CASP that is inserted as a cassette into the C-terminal portion of the metallo- β -lactamase domain (Callebaut et al., 2002; Mandel et al., 2006).

The ultimate proof that CPSF-73 is the endonuclease in the 3' end processing has been recently provided by demonstrating that a bacterially expressed N-terminal portion of human CPSF-73 (amino acids 1–460) encompassing the metallo- β -lactamase and the β -CASP domains can cleave single stranded RNA substrates in vitro (Mandel et al., 2006). A CPSF-73 mutant containing alanine substitutions within the histidine motif was enzymatically inactive. Interestingly, CPSF-73 is closely related to another component of CPSF, CPSF-100 (Jenny et al., 1994), suggesting that their genes have evolved through duplication of a common ancestral gene. CPSF-100, despite adopting the same fold as CPSF-73, contains alterations within the histidine motif and other conserved residues of the catalytic domain (Aravind, 1999; Callebaut et al., 2002; Mandel et al., 2006). Consistently, the bacterially expressed yeast CPSF-100 does not bind zinc ions and has no nuclease activity in vitro (Mandel et al., 2006).

Cleavage of the AAUAAA-containing pre-mRNAs in vitro is not inhibited by 5 mM EDTA (Hirose and Manley, 1997; Ryan et al., 2004). This intriguing feature initially suggested that the cleavage reaction is catalyzed by a metal-independent mechanism. Recent data indicate that this reaction in fact requires zinc ions and can be abolished by zinc-specific chelators, consistent with CPSF-73 being the endonuclease (Ryan et al., 2004). Based on structural studies, the active site of CPSF-73 appears to be hidden deep in the interface between the β -CASP and the metallo- β -lactamase domains and contains two zinc ions that are bound with extremely high affinity by CPSF-73 explaining the inability of EDTA to inhibit the cleavage reaction in vitro (Mandel et al., 2006).

3. Sequence elements in 3' end processing of histone pre-mRNAs

Formation of the 3' end of histone mRNAs was first studied a quarter of a century ago by injecting sea urchin histone genes into *Xenopus* oocytes (Birchmeier et al., 1982; Birchmeier et al., 1983) and was initially believed to result from accurate termination of transcription. Later studies using the same in vivo system proved that histone mRNAs are in fact formed from longer precursors by a post-transcriptional endonucleolytic cleavage at the 3' end (Birchmeier et al., 1984; Krieg and Melton, 1984). These studies resulted in the identification of a 60-nucleotide RNA designated U7 snRNA as an essential component of this 3' end processing reaction (Galli et al., 1983; Strub et al., 1984). Important insights into the mechanism of 3' end processing of histone pre-mRNAs have been subsequently provided using in vitro systems based on nuclear extracts from mammalian cells capable of accurate and efficient cleavage of pre-synthesized histone pre-mRNAs (Dominski and Marzluff, 1999; Furger et al., 1998; Gick et al., 1986; Mowry and Steitz, 1987a).

Two sequence elements required for processing that usually lie within 100 nucleotides downstream of the stop codon have been identified in both sea urchin (Birchmeier et al., 1982; Birchmeier et al., 1983; Birchmeier et al., 1984) and mammalian histone pre-mRNAs (Birchmeier et al., 1984; Cotten et al., 1988; Mowry et al., 1989; Mowry and Steitz, 1987a; Vasserot et al., 1989). The first element is a stem-loop, and the second, a purine-rich histone downstream element (HDE) that begins 15–20 nucleotides 3' of the stem-loop (Fig. 2). The stem-loop sequence is highly conserved in all organisms and consists of a 6-base pair stem and a 4-nucleotide loop. In sea urchins, the HDE has a highly conserved sequence CAAGAAAGA (Birchmeier et al., 1983). In vertebrate histone pre-mRNAs, the HDE is more variable although it contains a purine-rich core that resembles the HDE sequence of sea urchins (Birnstiel et al., 1985; Cotten et al., 1988; Mowry et al., 1989; Mowry and Steitz, 1987b). The endonucleolytic cleavage occurs between the two elements, after the fourth (sea urchins) or the fifth nucleotide (mammals) downstream of the stem-loop, which is typically an adenosine, yielding the mature histone mRNA that ends with the stem-loop followed by a short single stranded tail (Fig. 2). The precision of cleavage is not absolute and nearby sites can be selected if they are preceded by an adenosine, indicating that the processing machinery has some flexibility (Furger et al., 1998; Scharl and Steitz, 1994). Cleavage after other nucleotides occurs with reduced efficiency (Furger et al., 1998). Consistently, the vicinity of natural cleavage site in histone pre-mRNAs is rich in adenosines and contains relatively few guanosines, which create a particularly poor environment for the cleavage reaction (Dominski et al., 1999; Furger et al., 1998; Scharl and Steitz, 1994).

In vitro studies using mammalian nuclear extracts demonstrated that the two sequence elements in histone pre-mRNAs are not equivalent in their importance for 3' end processing. In some pre-mRNAs the stem-loop can be mutated without significantly affecting the efficiency of 3' end processing, whereas mutations within the HDE invariably result in complete inhibition of the cleavage reaction (Mowry et al., 1989; Vasserot et al., 1989). Moving the HDE by as few as 5–6 nucleotides away from the stem-loop results in either complete (Georgiev and Birnstiel, 1985) or partial inhibition of 3' end processing (Scharl and Steitz, 1996; Scharl and Steitz, 1994). Thus, the spatial arrangement of the two processing signals in histone pre-mRNAs is important for processing and suggested the existence of a communication between factors interacting with each of the two elements. These results justify the view that the two processing signals form in fact a single bipartite processing element in which two conserved sequences are separated by nonconserved nucleotides.

4. Factors involved in 3' end processing of mammalian histone pre-mRNAs

A number of components of 3' end processing machinery that cleaves histone pre-mRNAs have been identified in the past 25 years using both genetic and biochemical approaches. Two of these factors directly interact with the two sequence elements that flank the site of cleavage and are the defining feature of all histone pre-mRNAs. The remaining factors are required to form a complex competent to recruit and activate the endonuclease. While some of these factors are exclusively devoted to processing of histone pre-mRNAs, at least two of them are also utilized during cleavage/polyadenylation.

4.1. U7 snRNP

U7 snRNP was the first trans-acting factor identified in 3' end processing of histone pre-mRNAs and it interacts with the HDE located downstream of the cleavage site (Birnstiel and Schaufele, 1988). An average mammalian cell contains only about 10^3 particles of U7 snRNP. Thus, this 3' end processing factor is 2–3 orders of magnitude less abundant than major spliceosomal snRNPs. The U7 snRNP contains a small RNA component, U7 snRNA, which is the shortest known transcript generated by RNA polymerase II (Fig. 2). The exact length of the U7 snRNA varies slightly in different species between 57 (sea urchins) and 71 nucleotides (*Drosophila*) (Dominski et al., 2003b).

The 5' terminal part of the U7 snRNA is unstructured (Fig. 2) and in nuclear extracts is readily accessible for base-pairing with complementary oligonucleotides or digestion with micrococcal nuclease, indicating that it is not tightly associated with proteins (Cotten et al., 1988; Gilmartin et al., 1988; Mowry and Steitz, 1987b; Soldati and Schümperli, 1988). This part of the U7 snRNA is involved in formation of a duplex with the HDE (Fig. 2). Blocking the U7 snRNA by a complementary oligonucleotide (Cotten et al., 1991) or mutations within the HDE that reduce the base pair potential to U7 snRNA abolish 3' end processing (Bond et al., 1991; Scharl and Steitz, 1996; Schaufele et al., 1986). The negative effect of the HDE mutations can be restored by compensatory mutations within the U7 snRNA, demonstrating that the formation of an RNA hybrid rather than the exact sequence of each RNA element involved in base pairing is essential for processing (Bond et al., 1991; Scharl and Steitz, 1996; Schaufele et al., 1986).

The 3' part of U7 snRNA folds into an extended stem-loop, which varies between organisms in the size of the stem and the loop. The terminal stem-loop can be significantly shortened (Kolev and Steitz, 2006) or changed at the sequence level (Gilmartin et al., 1988) without affecting the function of the U7 snRNP in 3' end processing. The central part of the U7 snRNA contains an Sm site with the consensus AAUUUGUCUAG, in which the invariant nucleotides are underlined (Fig. 3) (Dominski et al., 2003b; Kolev and Steitz, 2006). This sequence differs in few positions from the AAUUUUUGG consensus determined for the spliceosomal snRNAs (Fig. 3). The U7-specific Sm site must be separated from the 3' terminal stem-loop by a minimum of two nucleotides to allow assembly of a functional U7 snRNP (Kolev and Steitz, 2006).

In all major spliceosomal snRNPs with the exception of U6 snRNP, the Sm site serves as a nucleation center for the assembly of the Sm protein core. The Sm core forms a ring around the Sm site and consists of seven common Sm proteins: B, D1, D2, D3, E, F and G (Kambach et al., 1999; Raker et al., 1999) (Fig. 3). The Sm proteins vary in molecular weight between 9 and 29 kDa and are characterized by the presence of two short conserved sequences termed Sm motifs 1 and 2 (Hermann et al., 1995). The U7 snRNP, like the spliceosomal snRNPs, reacts with the Y12 anti-Sm antibodies (Gick et al., 1986; Strub and Birnstiel, 1986), which primarily recognize Sm B/B' and to a lesser extent Sm D1 and D3 proteins (Pillai et al., 2001). Based on this observation, U7 snRNP was believed for many years to contain the

spliceosomal type Sm core. Initial purification procedures based on precipitating U7 snRNP by an oligonucleotide complementary to the 5' end of U7 snRNA revealed the presence of two U7-specific proteins of 14 and 50 kDa but did not determine their identity, function or sites of binding within the particle (Smith et al., 1991). Subsequent studies resulted in the important finding that these two proteins are in fact subunits of an unusual U7 Sm core and replace Sm D1 and D2 proteins (Fig. 3). Both the smaller and the larger protein contain two Sm motifs and were designated Lsm10 (Pillai et al., 2001) and Lsm11 (Pillai et al., 2003), respectively. Lsm11 is larger than any known Sm protein and is characterized by an unusually long N-terminal domain and a relatively long spacer between the two Sm motifs. Human Lsm10 resembles in size and structure other proteins of the family and shares the strongest sequence similarity with the Sm D1 that occupies the equivalent position in the spliceosomal Sm ring (Pillai et al., 2001).

UV cross-linking studies demonstrated that three proteins make contacts with the U7-specific Sm site: G, B and Lsm11 (Mital et al., 1993; Pillai et al., 2003; Stefanovic et al., 1995). The common Sm proteins G and B also UV-cross link to the AAUUUUUGG spliceosomal Sm site by making direct contacts with the first and third uridine of this site, respectively (Fig. 3) (Urlaub et al., 2001). Mutations that convert the U7-specific Sm site into the spliceosomal consensus result in the incorporation of Sm D1 and D2 proteins instead of Lsm10 and Lsm11. This mutant U7 snRNP is inactive in 3' end processing of histone pre-mRNAs, strongly arguing that the unique Sm-core of U7 snRNP is functionally essential (Grimm et al., 1993; Pillai et al., 2001; Pillai et al., 2003; Stefanovic et al., 1995). Variants of U7 snRNP that contain Lsm11 without the N-terminal domain or are mutated within this domain are also inactive in processing (Azzouz et al., 2005b; Pillai et al., 2003). Thus, in addition to playing structural functions in forming the Sm ring, Lsm11 directly participates in recruiting an essential processing factor (Azzouz et al., 2005a).

A detailed mutational analysis identified residues of the U7-specific Sm site that are required for the assembly of the unusual Sm core of the U7 snRNP and 3' end processing (Kolev and Steitz, 2006). Interestingly, among the 11 nucleotides of the AAUUUGUCUAG sequence found in the mouse U7 snRNA only four (underlined) are functionally important and cannot be substituted with any other nucleotide. The importance of these nucleotides is most likely related to their role in binding Lsm11, which was shown by UV-cross-linking studies to interact with this part of the Sm site (Stefanovic et al., 1995). The remaining nucleotides, including the first three uridines which are considered the hallmark of the Sm site, can be changed individually to other nucleotides without causing any adverse effect on the assembly of the Sm core and 3' end processing (Kolev and Steitz, 2006). Surprisingly, a substitution of the central uridine of the three consecutive uridines with a nucleotide analog, pseudouridine, is detrimental to the assembly of a functional U7 snRNP. Since pseudouridine is known to confer structural rigidity on the sugar-phosphate backbone, this result indicates that formation of the U7-specific Sm core containing Lsm10 and Lsm11 requires backbone flexibility at this particular nucleotide position. This restriction seems to be unique to the U7 Sm site and does not apply to the spliceosomal Sm site consensus, which assembles into the canonical Sm ring in the presence of the pseudouridine substitution (Kolev and Steitz, 2006).

4.2. Stem-loop binding protein (SLBP)

The conserved stem-loop structure in histone pre-mRNA is a binding site for the stem-loop binding protein (SLBP) (Wang et al., 1996), also referred to as the hairpin binding factor (HBF) (Martin et al., 1997). The cDNA for SLBP was cloned using the RNA three-hybrid system (Martin et al., 1997; Wang et al., 1996), a modification of the yeast two-hybrid system specifically designed for cloning RNA-binding proteins (SenGupta et al., 1996; Zhang et al., 1999). Human SLBP is a protein of 269 amino acids and molecular weight of 31 kDa, although

it has an unusual electrophoretic mobility and migrates at about 45 kDa. All known vertebrate orthologues of SLBP contain a centrally located RNA binding domain (RBD) of approximately 70 amino acids that has no similarity to any other known domain involved in RNA binding (Wang et al., 1996). Interestingly, two SLBPs have been identified in *Xenopus*. xSLBP1 shares strong similarity with the canonical SLBP over the entire protein length, whereas xSLBP2 displays similarity only within the highly conserved RBD (Wang et al., 1999b). xSLBP1 can substitute for human and mouse SLBP in 3' end processing, demonstrating that these proteins are orthologous (Ingledue et al., 2000; Wang et al., 1999b). xSLBP2 does not support 3' end processing and instead performs a specialized function in *Xenopus* oocytes by binding to the terminal stem-loop in stored histone mRNAs (Wang et al., 1999b) and in silencing their translational activity (Sanchez and Marzluff, 2004).

A number of in vitro studies have demonstrated that the role of SLBP in processing is to stabilize binding of U7 snRNP to the HDE (Dominski et al., 1999; Melin et al., 1992; Spycher et al., 1994; Streit et al., 1993). The importance of SLBP in 3' end processing depends on the type of histone pre-mRNA used in the assay. Processing of histone pre-mRNAs whose HDE has a relatively weak complementarity to U7 snRNA strongly depends on SLBP. On the other hand, substrates that form strong duplexes with the U7 snRNA can be processed in vitro with a significant efficiency in the absence of SLBP (Dominski et al., 1999; Melin et al., 1992; Spycher et al., 1994; Streit et al., 1993). Under in vivo conditions, SLBP is likely to be an essential factor in processing of all histone pre-mRNAs regardless of the HDE sequence and the extent of base pairing to the U7 snRNA (Pandey et al., 1994). Following processing, SLBP remains associated with the stem-loop at the 3' end and accompanies the mature histone mRNA to the cytoplasm (Dominski et al., 1995; Erkmann et al., 2005; Whitfield et al., 2004), where it stimulates its translation into histone proteins (Gorgoni et al., 2005; Sanchez and Marzluff, 2002).

In human SLBP, only the RBD and the following 20 amino acids are required for pre-mRNA processing (Dominski et al., 1999). Mutational analysis of the RBD identified a number of positively charged and hydrophobic residues that play an important role in binding of this domain to the stem-loop (Dominski et al., 2001; Martin et al., 2000). Interestingly, the amino acids required for RNA binding are interspersed with amino acids that are involved in processing while being dispensable for RNA binding (Dominski et al., 2001). Regions of the stem-loop RNA involved in the interaction with SLBP were mapped by analyzing effects of various mutations within the RNA target (Battle and Doudna, 2001; Dominski et al., 2003a; Martin et al., 2000; Williams and Marzluff, 1995). Important clues were also provided by the identification of a protein, 3'hExo, which can bind the stem-loop RNA both alone and simultaneously with SLBP (Dominski et al., 2003a). 3'hExo contains a 3'-5' exonuclease domain capable of degrading single stranded RNA substrates and a SAP domain that mediates binding to the stem-loop RNA (Cheng and Patel, 2004; Yang et al., 2006). SLBP binds to the 5' side of the stem-loop structure, whereas the 3' side of the structure is occupied by 3'hExo (Fig. 4). The binding of SLBP critically depends on two highly conserved adenosines located 2 and 3 nucleotides upstream of the stem, the second and the sixth (top) base pairs of the stem, and uridines 1 and 3 in the loop (Battle and Doudna, 2001; Dominski et al., 2003a; Williams and Marzluff, 1995). On the other hand, 3'hExo has a strong preference for the terminal ACCCA and binding also depends on the overall sequence of the stem and the loop (Dominski et al., 2003a; Yang et al., 2006). It is likely that binding of SLBP and 3'hExo is co-operative, suggesting that the two proteins may directly interact with each other in the ternary complex (Dominski et al., 2003a). Undoubtedly, binding of the two proteins to this 26-nucleotide stem-loop RNA placed a strong evolutionary constraint on this short RNA target, contributing to its unusual sequence conservation in all vertebrates. A possible role of 3'hExo is to directly participate in the degradation of histone mRNAs at the end of S phase, although the evidence for this role is still lacking.

SLBP levels are regulated during the mammalian cell cycle (Whitfield et al., 2000). The intracellular concentration of SLBP increases rapidly at the end of G1 phase and remains high during S phase. At the end of S phase, SLBP is degraded by the proteasome pathway. The degradation is triggered by phosphorylation of two adjacent threonines within SLBP, with phosphorylation of one residue being catalyzed by a cyclin-dependent kinase (Zheng et al., 2003). The degradation of SLBP is temporally associated with the disappearance of histone mRNAs from the cell. The absence of SLBP outside the S phase additionally blocks 3' end processing thus preventing formation of the mature histone mRNAs (Zheng et al., 2003). Thus, SLBP functions as a versatile regulator that ensures high levels of histone mRNAs and proteins only during S phase concomitant with DNA replication and prevents harmful accumulation of free histone proteins in the cell outside S phase. Not surprisingly, a deficit of SLBP results in defects in chromosome condensation due to the lack of histones and impair early developmental stages in *Drosophila* (Sullivan et al., 2001) and *C. elegans* (Kodama et al., 2002; Pettitt et al., 2002).

4.4. 100 kDa zinc finger protein (ZFP100)

ZFP100 is a 100 kDa protein that contains 18 C2H2 zinc fingers and was cloned using a modified version of the yeast two-hybrid system designed to detect proteins interacting with SLBP bound to the SL RNA (Dominski et al., 2002a; Dominski and Marzluff, 2001). ZFP100 does not interact with each component of the complex separately, suggesting that it may recognize binding determinates present on both SLBP and the stem-loop RNA. Alternatively, binding of SLBP to the RNA may result in appropriate structural rearrangements of SLBP and/or RNA allowing their interaction with ZFP100. The interaction with ZFP100 involves two regions of SLBP that are required for 3' end processing: a 9-amino acid region within the RBD that is not involved in RNA binding, and the 20-amino acid region located immediately to the C-terminus of the RBD. Over-expression of ZFP100 strongly stimulates 3' end processing of exogenous histone pre-mRNAs in *Xenopus* oocytes (Dominski et al., 2002a) and increases expression of a reporter gene containing the stem-loop and HDE in human cells (Wagner and Marzluff, 2006). These results are consistent with ZFP100 being a limiting factor in 3' end processing of histone pre-mRNAs. In contrast, over-expression of Lsm10 and Lsm11, despite resulting in increased levels of the U7 snRNP has no effect on the efficiency of 3' end processing in mammalian cells (Wagner and Marzluff, 2006).

Immunoprecipitation experiments demonstrated that ZFP100 is associated with a fraction of the U7 snRNP (Azzouz et al., 2005a; Dominski et al., 2002a). Both ZFP100 and U7 snRNP are concentrated in Cajal bodies (Frey and Matera, 1995; Pillai et al., 2001; Wagner et al., 2006; Wu and Gall, 1993), some of which are localized near histone genes and are likely sites of 3' end processing of histone pre-mRNAs and/or recycling of the U7 snRNP. ZFP100 does not co-purify with Lsm10 and Lsm11 during various chromatographic procedures (Pillai et al., 2001; Smith et al., 1991), suggesting that it may be only loosely bound to the U7 snRNP. Alternatively, ZFP100 may be present in substoichiometric amounts in the U7 snRNP, possibly as a component of its active fraction (Wagner and Marzluff, 2006), or may associate with the U7 snRNP only in the presence of histone pre-mRNA substrate and/or other processing factors (Azzouz et al., 2005a). Further studies are required to distinguish between these three possibilities. In vitro binding experiments revealed that ZFP100 associates with the U7 snRNP by making contacts with two regions of Lsm11; the N-terminal portion and Sm motifs 1 and 2 (Azzouz et al., 2005a; Wagner et al., 2006). The interaction involves a number of zinc fingers of ZFP100 and its N-terminal KRAB domain. The fact that ZFP100 interacts with both SLBP bound to the stem-loop and Lsm11 makes it a perfect candidate for a putative factor that stabilizes the U7 snRNP on the pre-mRNA (Dominski et al., 2002a). ZFP100 as a bridging factor spanning through the cleavage site may also play a role in recruiting other factors to the

processing machinery (Fig. 5). However, direct evidence supporting these possibilities is missing.

4.5. CPSF-73 as the endonuclease and 5' exonuclease

Cleavage of mammalian histone pre-mRNA by an endonuclease generates the upstream product terminated with the stem-loop followed by the ACCCA, which corresponds to the mature histone mRNA (Fig. 2). Cleavage occurs in the presence of 20 mM EDTA supporting the initial notion that U7-dependent processing, in common with the cleavage that precedes polyadenylation, does not require metal ions (Dominski and Marzluff, 1999; Gick et al., 1986; Mowry and Steitz, 1987a). The downstream cleavage product containing the HDE is degraded *in vitro* by a 5'-3' exonuclease that shares two features with the endonuclease; it is also dependent on the U7 snRNP and is active in the presence of EDTA (Walther et al., 1998). The major, if not the only, role of this 5' exonucleolytic activity may be to liberate the U7 snRNP from the HDE for another round of processing. The U7-dependent degradation of the downstream cleavage product requires a 5' phosphate and can be uncoupled from the cleavage reaction (Walther et al., 1998).

Until recently, the endonuclease cleaving histone pre-mRNAs remained unknown although a number of observations indicated that it might be related to the endonuclease involved in cleavage/polyadenylation (Dominski and Marzluff, 1999; Scharl and Steitz, 1994). The chemistry of the two cleavage reactions displays a number of common features, including resistance to EDTA, generation of a 3' hydroxyl at the end of the upstream product, and the preference for an adenosine preceding the cleavage site (Dominski et al., 2005c). One of the proteins initially considered to function as the endonuclease for histone pre-mRNAs was a homologue of CPSF-73 designated RC-68 (Dominski et al., 2005c). RC-68 interacts with a closely related metallo- β -lactamase protein, RC-74, which has lost a number of residues required for zinc binding, hence resembling CPSF-100. Surprisingly, UV-cross-linking experiments revealed that the cleavage site in histone pre-mRNA is contacted by CPSF-73, implicating this protein rather than RC-68 as an endonuclease in 3' end processing of histone pre-mRNAs (Dominski et al., 2005a). RC-68 was later shown to be involved 3' end processing of precursors to snRNAs, another group of transcripts also generated by RNA polymerase II (Baillat et al., 2005). UV-cross-linking of CPSF-73 to histone pre-mRNA depends on binding of the U7 snRNP to the HDE and correlates with the processing efficiency. Most importantly, detection of the CPSF-73 cross-link requires the presence of a phosphorothioate modification at the cleavage site, which likely slows down the cleavage reaction and increases the time window when the protein is present in the direct proximity to the RNA (Dominski et al., 2005a).

CPSF-73 can be also UV-cross-linked to the downstream cleavage product during its 5' exonucleolytic degradation, which is dependent on the U7 snRNP. This suggests that the same protein functions as both the endonuclease and 5'-3' exonuclease in 3' end processing of histone pre-mRNAs (Dominski et al., 2005a). The 5'-3' exonuclease activity seems to be an intrinsic feature of a number of nucleases of the metallo- β -lactamase family. *In vitro* assays identified this activity for the three DNA-specific nucleases involved in various aspects of DNA metabolism: Artemis (Ma et al., 2002), Apollo (Lenain et al., 2006) and the yeast Snm1 (Li et al., 2005). Artemis can also function as an endonuclease on single stranded DNA substrates but this activity requires its association with the catalytic subunit of DNA protein kinase (DNA-PK_{cs}) (Ma et al., 2002). Artemis is involved in V(D)J recombination and the endonuclease activity is essential for opening DNA hairpins formed during joining V, D and J segments. Thus far there is no evidence that the bacterially expressed N-terminal portion of human CPSF-73 displays the 5'-3' exonuclease activity (Mandel et al., 2006). However, is possible

that CPSF-73 requires other factors that normally function in 3' end processing to demonstrate the full spectrum of nucleolytic activities.

In cleavage/polyadenylation, CPSF-73 and CPSF-100 co-exist in CPSF (Kaufmann et al., 2004; Murthy and Manley, 1992) and likely directly interact with each other (Dominski et al., 2005c). In 3' end processing of pre-snRNAs a similar heterodimer is formed between RC-68 and RC-74 (Baillat et al., 2005; Dominski et al., 2005c). It is therefore very likely that CPSF-73 interacts with CPSF-100 also during 3' end processing of histone pre-mRNAs. The biological sense of a structural arrangement in which a catalytically active metallo- β -lactamase protein forms a heterodimer with an inactive homologue is unknown. One possibility is that the inactive component (CPSF-100 or RC-74) serves as an adaptor involved in a number of important interactions that are required to bring the catalytically active metallo- β -lactamase (CPSF-73 or RC-68) to the cleavage site. An alternative although not mutually exclusive possibility is that CPSF-100 and RC-74 normally act to inhibit the nuclease activity of their active partners and the repression is relieved only when the processing complex is assembled. Of note, CPSF-100 does not UV-cross-link to the vicinity of the cleavage site, suggesting that CPSF-73 may be the only protein of the heterodimer that directly contacts the RNA substrate (Dominski et al., 2005a).

4.6. The heat labile factor (HLF)

Twenty years ago it was shown that the incubation of a HeLa nuclear extract at 50° C for 15 min abolishes its activity in 3' end processing of histone pre-mRNAs (Gick et al., 1987). The processing activity could be restored by adding chromatographic fraction of a HeLa nuclear extract containing proteins with molecular weight of approximately 40 kDa. The factor capable of restoring the processing activity was not sensitive to treatment with micrococcal nuclease and did not react with anti-Sm antibodies, indicating that it was not the U7 snRNP. Subsequent experiments showed that this factor was also different from SLBP and therefore represented a new component of the processing machinery (Vasserot et al., 1989). This factor was designated the heat labile factor or HLF (Gick et al., 1987). Gel filtration studies demonstrated that the HLF, in addition to being present in fractions containing proteins of 40 kDa, is also detected in high molecular weight fractions, which retain normal processing activity. It is likely that the HLF was present in these fractions as part of a larger complex that co-fractionated with other complexes of the same size containing all other essential processing factors. In independent studies, HLF was shown to oscillate during the cell cycle reaching the highest levels in S phase cells when 3' end processing of histone pre-mRNAs is the most efficient (Lüscher and Schümperli, 1987; Stauber and Schümperli, 1988).

Recently, a set of experiments based on fractionation of a nuclear extract and complementation of a heat inactivated extract identified a heat sensitive component of the processing machinery for histone pre-mRNAs as symplekin, a protein previously implicated in cleavage/polyadenylation (Kolev and Steitz, 2005). Symplekin is a protein of approximately 150 kDa, significantly larger than 40 kDa determined by gel filtration as the activity complementing heat inactivated extracts. However, this ambiguity could be due to proteolysis of symplekin into a functionally active fragment of 40 kDa during the original gel filtration experiments (Gick et al., 1987). Symplekin synthesized *in vitro* and added to heat treated extracts restored its activity in processing of histone pre-mRNAs (Kolev and Steitz, 2005). This result demonstrates that symplekin is an essential processing factor and the major, if not the only, heat sensitive component of the processing machinery. Interestingly, the active chromatographic fractions in addition to symplekin contained other components of the cleavage/polyadenylation machinery, including all five CPSF subunits (CPSF-160, 100, 73, 30 and Fip1) and two CstF subunits; CstF-77 and CstF-64, with the notable absence of CstF-50 (Kolev and Steitz, 2005). In an untreated extract, all these proteins appear to be part of the same multi-subunit complex, which

losses its integrity upon heat treatment. This multi-subunit complex is likely identical to the heat labile factor (HLF) initially identified in high molecular weight fractions by gel filtration (Gick et al., 1987). The restoration of the processing activity by symplekin is accompanied by the reappearance of the multi-subunit complex, suggesting that symplekin may function as a platform in assembling a number of other processing factors into the HLF. The same role of a scaffold has been first suggested for symplekin in cleavage/polyadenylation (Takagaki and Manley, 2000).

It is unknown whether all cleavage/polyadenylation factors present together with symplekin in the active fractions are required for 3' end processing of histone pre-mRNAs. The main role of CPSF-160 in cleavage/polyadenylation is to bind the AAUAAA. Fip1 and CPSF-30 on the other hand preferentially interact with U-rich sequences upstream of the cleavage site, whereas CstF-64 binds to the downstream G/U-rich sequence. These interactions are essential for recruiting CPSF-73 to the cleavage site. The sequence elements recognized by CPSF-160, Fip1 and CPSF-30 are absent from histone pre-mRNA and the task of bringing the CPSF-73 endonuclease to this substrate is conducted primarily by U7 snRNP, with the auxiliary role being played by SLBP. Indeed, competition experiments with a molar excess of RNA containing the AAUAAA hexamer had no effect on 3' end processing of histone pre-mRNAs, indicating that CPSF-160 may have exclusive function in cleavage/polyadenylation (Dominski et al., 2005a). CPSF-160 and Fip1 play also important role in proper positioning of poly(A) polymerase in the cleavage/polyadenylation complex for the subsequent polyadenylation step (Kaufmann et al., 2004). This step is not part of the maturation of histone mRNAs and it is unclear how the function of these two proteins can be accommodated to the U7-dependent processing mechanism. However, one possibility is that CPSF-160 and the remaining components of the HLF are only required to form a larger complex containing CPSF-73 as the essential endonuclease although, in contrast to cleavage/polyadenylation, they do not perform any direct functions in processing of histone pre-mRNAs. According to this model, the HLF fits the definition of the cleavage factor, i.e. a multi component complex containing the catalytic subunit.

5. Known and predicted interactions within the processing complex

The generation of the mature histone mRNA during 3' end processing in mammalian nuclear extracts is not preceded by a pronounced lag phase. The final cleavage product begins to accumulate 3–5 min after the start of the *in vitro* reaction (Dominski et al., 1995; Gick et al., 1986; Mowry and Steitz, 1987a). This indicates that the formation of the catalytically competent processing complex is rapid and, in contrast to the formation of the spliceosome, likely requires a relatively small number of components. Formation of the processing complex involves a number of RNA-RNA, RNA-protein and protein-protein interactions that are required to recruit and activate the endonuclease at the cleavage site. The model presented below is a summary of known and predicted interactions within the processing complex (Fig. 5). This model will certainly continue to evolve as new components are identified within the processing machinery and the role of known components is further extended.

5.1. The role of SLBP and ZFP100 in recruiting U7 snRNP

The stem-loop is bound by SLBP, whereas the U7 snRNA base pairs with the HDE, thus anchoring the U7 snRNP downstream of the cleavage site (Fig. 5). In at least some histone pre-mRNAs, binding of SLBP to the stem-loop may ensure a favorable structure of the downstream sequences that facilitates the formation of a duplex between the U7 snRNA and the HDE (Jaeger et al., 2006). SLBP and the U7 snRNP are the major if not the only components of a stable processing complex that survives both electrophoresis in native gels (Melin et al., 1992) and immunoprecipitation with anti-SLBP (Dominski et al., 1999). These two factors

also associate with histone pre-mRNA in the presence of mild detergents, which prevent recruitment of CPSF-73 (Dominski et al., 2005a).

The SLBP bound to the stem-loop interacts with ZFP100 (Dominski et al., 2002a), which in turn interacts with Lsm11 (Azzouz et al., 2005a; Wagner et al., 2006). Since SLBP binds to its target with very high affinity, this bridging interaction is predicted to stabilize the U7 snRNP on the pre-mRNA, and is of particular importance for pre-mRNAs that only weakly base pair with the U7 snRNA. ZFP100 is a limiting factor in 3' end processing *in vivo* and is predicted to associate with only a small percentage of stable processing complexes containing SLBP and U7 snRNP (Wagner and Marzluff, 2006). It is not yet clear whether ZFP100 is an auxiliary factor that functions to stabilize the U7 snRNP on pre-mRNA or itself has an essential role in 3' end processing. *In vivo*, the stem-loop appears in the nascent transcript before the HDE located several nucleotides downstream. It is therefore possible that SLBP first binds to the stem-loop and subsequently recruits ZFP-100 and the U7 snRNP to the complex prior to the formation of the duplex between the U7 snRNA and the HDE. This initial interaction may be important for enforcing the proper alignment of these two RNAs and thus preventing nonproductive base-pairing schemes or those that might result in imprecise cleavage sites. This sequence of events is supported by *in vitro* studies, which revealed contacts between an anti Sm-reactive factor, most likely the U7 snRNP, and the stem-loop region in the initial phase of processing (Mowry and Steitz, 1987a).

5.2. The role of proper spatial configuration of the bipartite processing signal

Insertion of as few as 5–6 nucleotides between the stem-loop and the HDE in mammalian histone mRNA reduces processing (Scharl and Steitz, 1994). This negative effect is likely a result of disrupting the communication between SLBP and U7 snRNP hence the binding of the U7 snRNP to the pre-mRNA is not stabilized by SLBP and relies solely on the base pairing interaction. Indeed, processing of mammalian pre-mRNAs containing a 12-nucleotide insertion between the stem-loop and the HDE is not further inhibited by the presence of the stem-loop RNA competitor, i.e. does not depend on SLBP (Dominski et al., 2005b).

In addition to reducing cleavage efficiency, insertions of extra nucleotides between the stem-loop and HDE in mammalian histone pre-mRNAs result in a shift of the cleavage site that moves in concert with the HDE rather than staying in the proximity of the stem-loop (Scharl and Steitz, 1994). The fact that histone pre-mRNAs are cleaved a fixed distance from the HDE strongly argues that the U7 snRNP bound to this element functions as a molecular ruler that specifies the site of cleavage (Scharl and Steitz, 1994). As noticed soon after identifying the sequence of the U7 snRNA from sea urchin and mammalian species (Cotten et al., 1988; Soldati and Schümperli, 1988), base pairing between the 5' end of U7 snRNA and the HDE potentially places the UCU sequence of the Sm site opposite the cleavage site in the pre-mRNA (Fig. 2). This portion of the U7 Sm site departs from the spliceosomal consensus and directly contacts Lsm11 (Fig. 3) (Stefanovic et al., 1995). Thus, the mechanism of the molecular ruler may be executed by bringing this unique component of the U7 Sm ring to the site of cleavage, which is precisely determined by anchoring the U7 snRNP to the HDE through the base pair interaction.

The reduced processing efficiency and the shift of the cleavage site resulting from the separation of the bipartite processing element in mammalian histone pre-mRNA can be fully suppressed by using modified U7 snRNAs containing comparable insertions between the Sm site and the region complementary to the HDE (Scharl and Steitz, 1996). The extended U7 snRNA restores the original configuration of the processing complex by bringing the Sm site and the attached Sm core, including Lsm11, to the proximity of the stem-loop and normal cleavage site. Interestingly, sequences inserted into the U7 snRNA must be at least partially complementary to histone pre-mRNA to be effective in suppression (Scharl and Steitz,

1996). The presence of too long single stranded regions in the U7 snRNA and histone pre-mRNA between the duplex and the cleavage site may provide too much structural flexibility, precluding formation of the catalytically competent processing complex and/or precise specification of the cleavage site. The requirement for the proper distance between the region in U7 snRNA involved in base pairing and the Sm site may also indicate that the duplex between the U7 snRNA and the HDE is stabilized by a novel factor, which additionally rigidifies the region of both RNAs that spans to the site of cleavage. That this region is rigidified during 3' end processing is supported by the fact that nucleotides inserted between the stem-loop and the HDE do not loop out and instead disrupt the apparently relatively weak interaction that holds together SLBP and the U7 snRNP (Scharl and Steitz, 1994).

5.3. Recruitment of the endonuclease

The next step in 3' end processing of histone pre-mRNAs is the recruitment of the cleavage factor containing the endonuclease, recently identified as CPSF-73 (Dominski et al., 2005a). Since the same protein also functions as the endonuclease in cleavage/polyadenylation (Mandel et al., 2006), the critical question is how the same catalytic subunit is recruited to the two vastly different processing machineries. Given the predicted juxtaposition of Lsm11 and the cleavage site, the central role in properly positioning CPSF-73 must be played by this unique component of the U7 snRNP. Lsm11 is the largest protein of the Sm/Lsm family and contains an extended N-terminal region that is involved in at least two distinct interactions. One interaction, which is dispensable for processing involves ZFP100, whereas the other interaction is absolutely essential for processing and may link CPSF-73 to the U7 snRNP (Azzouz et al., 2005a). CPSF-73 is remarkably conserved among evolutionarily distant organisms, whereas Lsm11 has changed over its entire length during evolution and retains only a weak similarity between vertebrates and invertebrates (Azzouz and Schumperli, 2003). Therefore, the recruitment of CPSF-73 by Lsm11 is most likely indirect and may involve a yet unidentified component of the processing machinery. The recruitment of CPSF-73 probably also involves CPSF-100, which forms a heterodimer with CPSF-73 (Dominski et al., 2005c) and may function as a mediator in bringing this endonuclease to the cleavage site.

In a model proposed by Kolev and Steitz, CPSF-73 functions in 3' end processing of histone pre-mRNAs in a larger heat labile factor (HLF) that also includes symplekin and all components of CPSF and CstF with the notable exception of CstF-50 (Kolev and Steitz, 2005). Interestingly, CstF-50 consists of seven WD-40 repeats, which form a ring structurally similar to that of the Sm core proteins. Therefore, an attractive possibility is that a component of the HLF interacts with the Sm ring of the U7 snRNP by making contacts with Lsm11 rather than with a specific region of CstF-50.

5.4. Cleavage and disruption of the complex

Formation of the processing complex containing all necessary components, possibly accompanied by catalytic activation of CPSF-73, results in cleavage of histone pre-mRNA and rapid dissociation of the complex. The upstream cleavage product ending with the stem-loop and the ACCCA tail corresponds to the mature histone mRNA and remains associated with SLBP. The downstream product with the bound U7 snRNP is rapidly degraded in the 5'-3' direction by the exonucleolytic activity of CPSF-73, resulting in liberation of the U7 snRNP from the HDE, which can be recycled for another round of processing (Walther et al., 1998). Termination of transcription on histone genes is dependent on 3' end processing (Chodchoy et al., 1991), indicating that the two events are tightly coupled and resemble in this respect coupling between transcription termination and cleavage/polyadenylation (Bentley, 2005; Buratowski, 2005; Proudfoot et al., 2002). It is possible that the 5'-3' exonuclease activity of CPSF-73 plays an active role in coupling the U7-dependent processing to transcription termination on histone genes by a "torpedo" mechanism, which in generation of polyadenylated

transcripts was shown to require another 5'-3' exonuclease, Rat1/Xrn2 (Luo et al., 2006; Rosonina et al., 2006; Tollervey, 2004).

6. 3' end processing of histone pre-mRNAs in *Drosophila*

In recent years, an efficient and accurate in vitro 3' end processing system has been developed for *Drosophila* and studies using this system significantly extended our knowledge of how histone pre-mRNAs are cleaved in this organism (Adamson and Price, 2003; Dominski et al., 2002b). Most but not all features of mammalian processing are conserved in *Drosophila* (Dominski et al., 2002b; Dominski et al., 2005b). The stem-loop in *Drosophila* histone pre-mRNAs contains all highly conserved nucleotides, whereas the HDE is relatively weakly defined but contains a short purine core that is essential for 3' end processing (Dominski et al., 2002b). The two sequence elements are spaced by a similar number of nucleotides as those present in mammalian histone pre-mRNAs. Thus, all critical features of the bipartite organization of the processing signal in histone pre-mRNAs are conserved in *Drosophila*. *Drosophila* extracts cleave histone pre-mRNAs four nucleotides after the stem-loop, one nucleotide closer to the stem-loop compared to the major cleavage site in the mammalian processing (Dominski et al., 2002b). All *Drosophila* histone pre-mRNAs contain an adenosine in this position, suggesting that the *Drosophila* and mammalian processing machineries prefer the same nucleotide in front of the cleavage site. Interestingly, unlike processing in mammalian nuclear extracts, cleavage in *Drosophila* pre-mRNAs does not shift to more distal sites upon moving the HDE away from the stem-loop (Dominski et al., 2002b).

Drosophila U7 snRNA with 71 nucleotides is the longest known U7 snRNA, contains the non-canonical Sm site (Dominski et al., 2003b) and associates with Lsm10 and Lsm11 (Azzouz and Schumperli, 2003), demonstrating that the structural uniqueness of the Sm complex of the U7 snRNP is also preserved in invertebrates. Because of its length, the 5' end of the U7 snRNA has a larger capacity to form a number of base pairs with each *Drosophila* histone pre-mRNA (Dominski et al., 2005b). However, the potential duplexes are rich in AU base pairs and hence may be relatively unstable.

Drosophila SLBP contains a readily recognizable RNA binding domain (RBD) but otherwise has no similarity to SLBP from vertebrate species (Sullivan et al., 2001). The RBD in *Drosophila* is located close to the C-terminus and is followed by a short stretch of amino acids rich in aspartates and serines. The serines are heavily phosphorylated in vivo, thus creating an extremely acidic domain at the end of SLBP (Dominski et al., 2002b). In contrast to mammalian processing, SLBP is indispensable for in vitro processing of all *Drosophila* histone pre-mRNAs (Dominski et al., 2002b). Both the essential role of SLBP in processing and the inability of large insertions placed between the stem-loop and HDE to change to position of the cleavage site suggest that in *Drosophila* the U7 snRNP and SLBP together function in specifying the site of cleavage and the two factors interact with each other, either directly or indirectly, with a high affinity (Dominski et al., 2005b). This putative interaction may occur prior to formation of the duplex and serve to properly align the U7 snRNA on HDE, hence precluding illegitimate and nonproductive base pair schemes facilitated by the A/U-richness of *Drosophila* HDEs and the 5' end of U7 snRNA (Dominski et al., 2005b). The acidic C-terminus of SLBP is a primary candidate for a domain that functions in anchoring U7 snRNP to the vicinity of the stem-loop. This domain is absolutely required for processing in vitro and its dephosphorylation significantly reduces the ability of SLBP to complement SLBP-depleted nuclear extracts (Dominski et al., 2002b).

No orthologue of ZFP100 or its functional counterpart has been identified in *Drosophila*. It is possible that this protein, like many other zinc finger proteins, has evolved rapidly and can not be recognized in the *Drosophila* protein database (Huntley et al., 2006). Alternatively, 3' end

processing of histone pre-mRNAs in *Drosophila* differs in this respect from processing in mammalian cells. The identity of the *Drosophila* endonuclease has not been yet determined. However, a number of observations, including the resistance of *Drosophila* in vitro processing to EDTA, a preference for an adenosine directly preceding the cleavage site and the degradation of the downstream cleavage product in the 5' to 3' direction strongly suggest that *Drosophila* histone pre-mRNAs are also cleaved by CPSF-73 (Dominski et al., 2005b).

7. Maturation of the U7 snRNP

Spliceosomal snRNPs, U1, U2, U4 and U5, undergo initial maturation steps in the cytoplasm, where they acquire the Sm core complex, consisting of the common Sm proteins B, D1, D2, D3, E, F and G (Fig. 3) (Kiss, 2004; Will and Luhrmann, 2001). The seven Sm proteins form a ring around the canonical Sm site in the spliceosomal snRNAs (Kambach et al., 1999; Raker et al., 1999). Formation of the Sm ring is a highly complex process controlled by the so called SMN complex, which consists of the SMN protein and five additional proteins termed Gemin2–7 (Gubitz et al., 2004). Components of the SMN complex associate with common Sm proteins and transfer them on the spliceosomal RNA. The assembly process is facilitated by the fact that the Sm proteins are pre-assembled in the cytoplasm, forming three intermediates: D1/D2, B/D3 and F/E/G (Raker et al., 1996). The spliceosomal snRNAs are distinguished from other cellular RNAs by Gemin5, which recognizes the 3' terminal stem-loop and the sequence of the Sm binding site, AAUUUUUGG, with the most critical role being played by the second adenosine and the first and third uridines (Battle et al., 2006; Golembe et al., 2005). The spliceosomal snRNPs containing the fully assembled Sm core are imported to Cajal bodies for final modifications (Stanek and Neugebauer, 2006).

The U7-specific Sm core appears to be assembled by a specialized SMN complex that associates with Lsm10 and Lsm11 instead of SmD1 and D2, and the five remaining common Sm proteins; B, D3, E, F and G (Azzouz et al., 2005b; Pillai et al., 2003). It is possible that like Sm proteins D1 and D2, Lsm10 and Lsm11 also interact with each other in the cytoplasm and are incorporated into the U7-specific Sm core as a heterodimer. The U7 snRNA does not interact with Gemin5 despite containing all necessary features recognized by this protein, i.e. the 3' terminal stem-loop, and the three critical nucleotides of its AAUUUGUCUAG Sm site: the second adenosine, and first and third uridines (Fig. 3) (Battle et al., 2006). The remaining part of the unusual sequence of the Sm binding site may somehow interfere with recognition of the U7 snRNA by Gemin5 and prevents this RNA from receiving the spliceosome type Sm core. One possibility is that the key role in the formation of the U7-specific Sm core is played by Lsm11 itself, which interacts with the Sm site in U7 snRNA but not with the canonical Sm site in spliceosomal snRNAs (Pillai et al., 2003). According to this scenario Lsm11, likely as a heterodimer with Lsm10, bound to the unusual Sm site in U7 snRNA creates a nucleation center recognized by the SMN complex, which subsequently completes the assembly process by adding the missing common Sm proteins. The inefficient formation of the unique U7-specific Sm core contributes significantly to the low abundance of complete U7 snRNP particles in mammalian cells (Grimm et al., 1993). Conversion of the U7-specific Sm site to that existing in the spliceosomal snRNAs results in a several-fold increase in the cellular concentration of the hybrid U7 snRNP. This increase can almost certainly be attributed to the much higher cellular abundance of the D1 and D2 proteins in comparison with Lsm10 and Lsm11 and/or more efficient co-operation of the SMN complex with the common Sm proteins in the Sm core assembly (Pillai et al., 2003).

8. The role of Cajal bodies in histone gene expression

Cajal bodies, previously known as coiled bodies, are small nuclear structures that vary in a number of copies from 1 to several hundreds per single cell, depending on the developmental

and metabolic status of the cell (Gall, 2000; Gall, 2003). One of the features of Cajal bodies in vertebrates is the presence of an 80 kDa protein called coilin. A number of observations indicate that Cajal bodies are involved in modification/remodeling of the snRNPs and snoRNPs before their final distribution to other cellular locations (Stanek and Neugebauer, 2006).

In vertebrate cells, Cajal bodies are highly enriched in the U7 snRNA and a subset of these structures are located close to the histone gene loci (Frey and Matera, 1995; Wu and Gall, 1993). Cajal bodies in *Xenopus* oocytes were also shown to contain small amounts of SLBP (Abbott et al., 1999), whereas human Cajal bodies are the exclusive site for ZFP100 (Wagner et al., 2006). Moreover, three factors only recently implicated in 3' end processing of histone pre-mRNAs as components of the heat labile factor; symplekin (Hofmann et al., 2002; Schul et al., 1996), CstF-64 and CPSF-100 (Hofmann et al., 2002; Schul et al., 1996) are also concentrated in Cajal bodies. In mammalian cells, Cajal bodies that are located close to the histone gene loci contain a protein termed NPAT that undergoes phosphorylation by cyclin E-Cdk2 during the G1/S transition and activates transcription of histone genes (Ma et al., 2000; Miele et al., 2005; Zhao, 2004; Zhao et al., 2000). Collectively, these data suggest that vertebrate Cajal bodies associated with the histone gene loci are actively engaged in transcription of histone genes and 3' end processing of the nascent histone transcripts (Gall, 2000). A clear advantage of such a structural organization might be to enhance the efficiency and accuracy of formation of histone mRNAs by concentrating all necessary transcription and processing factors in a single compartment located close to histone genes. Importantly, replication-dependent histone genes are clustered in all eukaryotes, strongly arguing that direct proximity on chromosomes of these coordinately regulated class of genes provides a clear evolutionary advantage (Marzluff et al., 2002). In *Drosophila*, the U7 snRNP resides in a distinct nuclear structure, the histone locus body (HLB), which is separated from Cajal body but invariably co-localizes with the histone gene locus (Liu et al., 2006).

It has been reported that in mammalian cells, CstF-64 and CPSF-100 have an unusual distribution. During G1 phase the two processing factors in addition to being evenly distributed throughout the nucleoplasm are concentrated in Cajal bodies (Schul et al., 1999). Interestingly, in S phase CstF-64 and CPSF-100 redistribute to separate structures called the cleavage bodies, which in contrast to Cajal bodies precisely co-localize with the histone gene loci and contain nascent histone transcripts. It is possible that the cleavage bodies, although not always easily distinguishable from closely associated Cajal bodies, also contain U7 snRNP and other essential processing factors, and constitute the true sites for 3' end processing of histone pre-mRNAs. Notably, the identification of CstF-64 and CPSF-100 in the close proximity to the histone gene loci for the first time suggested that these two known cleavage/polyadenylation processing factors may also be involved in formation of the nonpolyadenylated histone mRNAs (Schul et al., 1999).

9. Integration of 3' end processing and transcription

Although in vitro cleavage/polyadenylation of pre-mRNA can be separated from transcription, the two events are tightly coupled in vivo (Bentley, 2005; Hirose and Manley, 2000; Proudfoot, 2004; Proudfoot et al., 2002). The central role in coupling transcription and cleavage/polyadenylation is played by the C-terminal domain (CTD) of the large RNA Pol II subunit, which functions as a platform for binding numerous 3' end processing factors during transcription. A number of components of the cleavage/polyadenylation machinery may interact with transcription factors already at the promoter during the initiation phase of transcription and are later transferred on the CTD and other subunits of the elongating RNA polymerase complex (Calvo and Manley, 2003). The human CTD consists of 52 highly similar heptad repeats with the consensus YSPTSPS. Within the repeat, the serines 2 and 5 undergo reversible phosphorylation thus regulating the function of the CTD during different phases of

the transcription cycle. Phosphorylation of serine 2 by cyclin-dependent kinase 9 (Cdk9) is required for the elongation phase of transcription and thus formation of the full length pre-mRNAs and stimulates binding of cleavage/polyadenylation factors to the CTD (Ahn et al., 2004; Bird et al., 2004; Peterlin and Price, 2006; Phatnani and Greenleaf, 2006). The coupling between transcription and cleavage/polyadenylation is believed to improve the efficiency and accuracy of 3' end processing by increasing the local concentration of processing factors in the vicinity of the nascent transcript. Indeed, a significantly higher rate of cleavage coupled to transcription compared to cleavage of a pre-synthesized RNA substrate was recently recapitulated in vitro (Adamson et al., 2005). The CTD is also required for the cleavage reaction in an uncoupled in vitro system, suggesting that it may play a more general function in 3' end processing (Hirose and Manley, 1998).

RNA Pol II also generates pre-snRNAs and histone pre-mRNAs. These transcripts are much shorter, intron-less, processed at the 3' end by separate machineries and controlled by distinct gene promoters. Accurate 3' end processing of pre-snRNAs in mammalian cells requires that transcription of snRNA genes initiates from an snRNA gene promoter (Hernandez and Weiner, 1986). This suggests that the 3' end processing machinery for pre-snRNAs assembles only at a compatible promoter region and travels with the elongating polymerase until it reaches the appropriate 3' end processing signals in the nascent transcript. Pre-snRNAs are processed by a 12-subunit complex called Integrator that contains a putative endonuclease RC-68/Int11 and directly associates with the CTD (Baillat et al., 2005). Inhibition of Cdk9 activity does not affect generation of pre-snRNAs although it abolishes their 3' end processing (Jacobs et al., 2004; Medlin et al., 2005; Medlin et al., 2003), demonstrating that the CTD phosphorylated on serine 2 is selectively required for 3' end processing of pre-snRNAs but is dispensable for transcription elongation. In agreement with these results, addition of the phosphorylated CTD stimulates in vitro 3' end processing of snRNA precursors (Uguen and Murphy, 2003).

In contrast to snRNAs, mature histone mRNAs accumulate to normal levels in the presence of Cdk9 inhibitors (Medlin et al., 2005). Thus, neither transcription of histone genes nor 3' end processing of histone pre-mRNAs depends on phosphorylation of serine 2. Accordingly, chromatin immunoprecipitation assays revealed that only relatively low levels of Cdk9 and phosphorylated serine 2 are associated with RNA Pol II transcribing histone genes (Medlin et al., 2005). In addition, contrary to the in vitro functional coupling of transcription with cleavage/polyadenylation, no stimulation of 3' end processing was observed when histone pre-mRNA was tethered to DNA template in a transcription elongation complex (Adamson and Price, 2003). Taken together, these results suggest that the CTD may play only a minor role in coupling transcription to 3' end processing in expression of histone genes, in sharp contrast to its role in expression of other protein-encoding genes.

So far no interaction between transcription factors and components of the 3' end processing machinery has been documented in formation of histone mRNAs. Among known processing factors, the best candidate for a factor that links transcription with cleavage is ZFP100. This protein has a number of features of a transcription factor, including the N-terminal KRAB domain (Huntley et al., 2006) and the ability to strongly activate expression of reporter genes in the yeast two-hybrid system (Dominski et al., 2002a). It is possible that NPAT, the known transcriptional activator of histone genes, may also assist 3' end processing since the region of NPAT not required for histone gene transcription is essential for progression through S phase (Ye et al., 2003). During S phase, both transcription of histone genes and 3' end processing of the resultant pre-mRNAs are highly stimulated, coordinately contributing to the great increase in the intracellular concentration of mature histone mRNAs concomitant with DNA replication (Marzluff, 2005; Stein et al., 1996). Undoubtedly, there is an extensive network of interactions, possibly mediated by factors residing in Cajal (or cleavage) bodies that integrate transcription and 3' end processing for production of histone mRNAs in S phase. Future studies should result

in identification of new components of this network and shed some light on how this integration is achieved.

10. Conclusions

Since our last review in 1999, five new components of 3' end processing of histone pre-mRNAs have been identified. The list begins with Lsm10 and Lsm11, which are part of the U7-specific Sm core and are absent from the spliceosomal snRNPs. Lsm11 is significantly larger than any other known Sm or Lsm protein and, in addition to playing a structural role in forming the U7-specific Sm core, is believed to direct the assembly of the catalytically active complex due to its perfect juxtaposition with the cleavage site. The third recently identified factor, ZFP100, which interacts with Lsm11 links SLBP bound to the stem-loop with the U7 snRNP. ZFP100 may also help recruit additional processing factors to histone pre-mRNA and couple 3' end processing to transcription. The two factors most recently added to the list, CPSF-73 and symplekin, are the endonuclease and the heat sensitive component of the processing machinery, respectively. CPSF-73 and symplekin are known components of the cleavage/polyadenylation machinery and in 3' end processing of histone pre-mRNAs they may function as a larger heat labile factor (HLF) that contains other subunits of CPSF and two subunits of CstF; CstF-64 and CstF-77 (Kolev and Steitz, 2005). However, it is unknown whether these additional components are indeed required for the U7-dependent processing and whether the activity of any of these components is regulated during the cell cycle, as previously reported for HLF (Gick et al., 1987). Finding answers to these questions will be helpful in understanding the most important part of the puzzle: how the same endonuclease, CPSF-73, is recruited to the two vastly different pre-mRNAs that assemble into two specific processing complexes.

The U7-dependent mechanism for 3' end processing of replication-dependent histone pre-mRNAs is present only in metazoans and two flagellate green algae, *Volvox* and *Chlamydomonas*. In plants, fungi and other lower eukaryotes, histone mRNAs are polyadenylated and generated by the common 3' end processing machinery that also acts on all other mRNA precursors (Osley, 1991). Since the formation of histone mRNAs and polyadenylated mRNAs requires a set of the same factors including the endonuclease, the two relevant pre-mRNA processing machineries likely have a common evolutionary origin. An intriguing question to which we may never find an answer is whether, as suggested (Gilmartin, 2005), the U7-dependent mechanism for 3' end formation of histone mRNAs and its links to cells cycle regulatory pathways are a recent development that was invented after the divergence of plants and animals. If so, this specialized processing pathway could have evolved in early metazoan ancestors from the cleavage/polyadenylation pathway by adopting one of the spliceosomal snRNPs to the new role in formation of histone mRNAs. Or perhaps the U7-dependent processing is a relic of the primordial 3' end processing machinery that has been lost during evolution by plants and fungi and replaced in these organisms by the universal cleavage/polyadenylation machinery composed of exclusively protein components (Mowry and Steitz, 1988).

Acknowledgments

We thank our colleagues from UNC at Chapel Hill and many collaborators for their contribution to our work on 3' end processing of histone pre-mRNAs. We also thank J. Spychala (VisiScience) for help with the figures. This work is supported by the NIH grants GM 29832 and GM58921.

References

Abbott J, Marzluff WF, Gall JG. The stem loop binding protein (SLBP1) is present in coiled bodies of the *Xenopus* germinal vesicle. *Mol. Biol. Cell* 1999;10:487–499. [PubMed: 9950690]

- Adamson TE, Price DH. Cotranscriptional processing of *Drosophila* histone mRNAs. *Mol. Cell Biol* 2003;23:4046–4055. [PubMed: 12773550]
- Adamson TE, Shutt DC, Price DH. Functional coupling of cleavage and polyadenylation with transcription of mRNA. *J. Biol. Chem.* 2005
- Ahn SH, Kim M, Buratowski S. Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. *Mol. Cell* 2004;13:67–76. [PubMed: 14731395]
- Aravind L. An evolutionary classification of the metallo-beta-lactamase fold proteins. *In Silico. Biol* 1999;1:69–91. [PubMed: 11471246]
- Azzouz TN, Gruber A, Schumperli D. U7 snRNP-specific Lsm11 protein: dual binding contacts with the 100 kDa zinc finger processing factor (ZFP100) and a ZFP100-independent function in histone RNA 3' end processing. *Nucleic Acids Res* 2005a;33:2106–2117. [PubMed: 15824063]
- Azzouz TN, Pillai RS, Dapp C, Chari A, Meister G, Kambach C, Fischer U, Schumperli D. Toward an assembly line for U7 snRNPs: interactions of U7-specific Lsm proteins with PRMT5 and SMN complexes. *J. Biol. Chem* 2005b;280:34435–34440. [PubMed: 16087681]
- Azzouz TN, Schumperli D. Evolutionary conservation of the U7 small nuclear ribonucleoprotein in *Drosophila melanogaster*. *RNA* 2003;9:1532–1541. [PubMed: 14624008]
- Baillat D, Hakimi MA, Naar AM, Shilatifard A, Cooch N, Shiekhattar R. Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. *Cell* 2005;123:265–276. [PubMed: 16239144]
- Barnard DC, Ryan K, Manley JL, Richter JD. Symplekin and xGLD-2 are required for CPEB-mediated cytoplasmic polyadenylation. *Cell* 2004;119:641–651. [PubMed: 15550246]
- Battle DJ, Doudna JA. The stem-loop binding protein forms a highly stable and specific complex with the 3' stem-loop of histone mRNAs. *RNA* 2001;7:123–132. [PubMed: 11214174]
- Battle DJ, Lau CK, Wan L, Deng H, Lotti F, Dreyfuss G. The Gemin5 protein of the SMN complex identifies snRNAs. *Mol. Cell* 2006;23:273–279. [PubMed: 16857593]
- Bentley DL. Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. *Curr. Opin. Cell Biol* 2005;17:251–256. [PubMed: 15901493]
- Birchmeier C, Folk W, Birnstiel ML. The terminal RNA stem-loop structure and 80 bp of spacer DNA are required for the formation of 3' termini of sea urchin H2A mRNA. *Cell* 1983;35:433–440. [PubMed: 6317188]
- Birchmeier C, Grosschedl R, Birnstiel ML. Generation of authentic 3' termini of an H2A mRNA in vivo is dependent on a short inverted DNA repeat and on spacer sequences. *Cell* 1982;28:739–745. [PubMed: 6284372]
- Birchmeier C, Schümperli D, Sconzo G, Birnstiel ML. 3' editing of mRNAs: sequence requirements and involvement of a 60-nucleotide RNA in maturation of histone mRNA precursors. *Proc. Natl. Acad. Sci. USA* 1984;81:1057–1061. [PubMed: 6583695]
- Bird G, Zorio DA, Bentley DL. RNA polymerase II carboxy-terminal domain phosphorylation is required for cotranscriptional pre-mRNA splicing and 3'-end formation. *Mol. Cell Biol* 2004;24:8963–8969. [PubMed: 15456870]
- Birnstiel ML, Busslinger M, Strub K. Transcription termination and 3' processing: the end is in site! *Cell* 1985;41:349–359. [PubMed: 2580642]
- Birnstiel, ML.; Schaufele, FJ. Structure and function of minor snRNPs. In: Birnstiel, ML., editor. *Structure and Function of Major and Minor Small Ribonucleoprotein Particles*. Berlin: Springer-Verlag; 1988. p. 155-182.
- Bond UM, Yario TA, Steitz JA. Multiple processing-defective mutations in a mammalian histone premessenger RNA are suppressed by compensatory changes in U7 RNA both in vivo and in vitro. *Genes Dev* 1991;5:1709–1722. [PubMed: 1885007]
- Brown KM, Gilmartin GM. A mechanism for the regulation of pre-mRNA 3' processing by human cleavage factor Im. *Mol. Cell* 2003;12:1467–1476. [PubMed: 14690600]
- Buratowski S. Connections between mRNA 3' end processing and transcription termination. *Curr. Opin. Cell Biol* 2005;17:257–261. [PubMed: 15901494]
- Callebaut I, Moshous D, Mornon JP, De Villartay JP. Metallo-beta-lactamase fold within nucleic acids processing enzymes: the beta-CASP family. *Nucleic Acids Res* 2002;30:3592–3601. [PubMed: 12177301]

- Calvo O, Manley JL. Strange bedfellows: polyadenylation factors at the promoter. *Genes Dev* 2003;17:1321–1327. [PubMed: 12782649]
- Cheng Y, Patel DJ. Crystallographic structure of the nuclease domain of 3'hExo, a DEDDh family member, bound to rAMP. *J. Mol. Biol* 2004;343:305–312. [PubMed: 15451662]
- Chodchoy N, Pandey NB, Marzluff WF. An intact histone 3' processing site is required for transcription termination in a mouse histone H2a gene. *Mol. Cell. Biol* 1991;11:497–509. [PubMed: 1986243]
- Christofori G, Keller W. Poly(A) polymerase purified from HeLa cell nuclear extract is required for both cleavage and polyadenylation of pre- mRNA in vitro. *Mol. Cell. Biol* 1989;9:193–203. [PubMed: 2538718]
- Colgan DF, Manley JL. Mechanism and regulation of mRNA polyadenylation. *Genes and Development* 1997;11:2755–2766. [PubMed: 9353246]
- Cotten M, Gick O, Vasserot A, Schaffner G, Birnstiel ML. Specific contacts between mammalian U7 snRNA and histone precursor RNA are indispensable for the in vitro RNA processing reaction. *EMBO J* 1988;7:801–808. [PubMed: 3396543]
- Cotten M, Oberhauser B, Brunar H, Holzner A, Issakides G, Noe CR, Schaffner G, Wagner E, Birnstiel ML. 2'-O-methyl, 2'-O-ethyl oligoribonucleotides and phosphorothioate oligodeoxyribonucleotides as inhibitors of the *in vitro* U7 snRNP-dependent mRNA processing event. *Nucleic Acids Res* 1991;19:2629–2635. [PubMed: 1710340]
- De Vries H, Rügsegger U, Hübner W, Friedlein A, Langen H, Keller W. Human pre-mRNA cleavage factor II_m contains homologs of yeast proteins and bridges two other cleavage factors. *EMBO J* 2000;19:5895–5904. [PubMed: 11060040]
- Dominski Z, Erkmann JA, Greenland JA, Marzluff WF. Mutations in the RNA binding domain of stem-loop binding protein define separable requirements for RNA binding and histone pre-mRNA processing. *Mol. Cell. Biol* 2001;21:2008–2017. [PubMed: 11238936]
- Dominski Z, Erkmann JA, Yang X, Sanchez R, Marzluff WF. A novel zinc finger protein is associated with U7 snRNP and interacts with the stem-loop binding protein in the histone pre-mRNP to stimulate 3'-end processing. *Genes Dev* 2002a;16:58–71. [PubMed: 11782445]
- Dominski Z, Marzluff WF. Formation of the 3' end of histone mRNA. *Gene* 1999;239:1–14. [PubMed: 10571029]
- Dominski Z, Marzluff WF. Three-hybrid screens for RNA-binding proteins: Proteins binding the 3' end of histone mRNA. *Methods in Molecular Biology* 2001;177:291–318. [PubMed: 11530613]
- Dominski Z, Sumerel J, Hanson RJ, Marzluff WF. The polyribosomal protein bound to the 3' end of histone mRNA can function in histone pre-mRNA processing. *RNA* 1995;1:915–923. [PubMed: 8548656]
- Dominski Z, Yang X, Kaygun H, Marzluff WF. A 3' exonuclease that specifically interacts with the 3' end of histone mRNA. *Molecular Cell* 2003a;12:295–305. [PubMed: 14536070]
- Dominski Z, Yang X, Raska CS, Santiago CS, Borchers CH, Duronio RJ, Marzluff WF. 3' end processing of *Drosophila* histone pre-mRNAs: Requirement for phosphorylated dSLBP and co-evolution of the histone pre-mRNA processing system. *Mol. Cell. Biol* 2002b;22:6648–6660. [PubMed: 12192062]
- Dominski Z, Yang XC, Marzluff WF. The Polyadenylation Factor CPSF-73 Is Involved in Histone-Pre-mRNA Processing. *Cell* 2005a;123:37–48. [PubMed: 16213211]
- Dominski Z, Yang XC, Purdy M, Marzluff WF. Cloning and characterization of the *Drosophila* U7 small nuclear RNA. *Proc. Natl. Acad. Sci. U. S. A* 2003b;100:9422–9427. [PubMed: 12872004]
- Dominski Z, Yang XC, Purdy M, Marzluff WF. Differences and similarities between *Drosophila* and mammalian 3' end processing of histone pre-mRNAs. *RNA* 2005b;11:1–13.
- Dominski Z, Yang XC, Purdy M, Wagner EJ, Marzluff WF. A CPSF-73 homologue is required for cell cycle progression but not cell growth and interacts with a protein having features of CPSF-100. *Mol. Cell Biol* 2005c;25:1489–1500. [PubMed: 15684398]
- Dominski Z, Zheng LX, Sanchez R, Marzluff WF. Stem-loop binding protein facilitates 3'-end formation by stabilizing U7 snRNP binding to histone pre-mRNA. *Mol. Cell Biol* 1999;19:3561–3570. [PubMed: 10207079]
- Erkmann JA, Wagner EJ, Dong J, Zhang Y, Kutay U, Marzluff WF. Nuclear import of the stem-loop binding protein and localization during the cell cycle. *Mol. Biol. Cell* 2005;16:2960–2971. [PubMed: 15829567]

- Frey MR, Matera AG. Coiled bodies contain U7 small nuclear RNA and associate with specific DNA sequences in interphase human cells. *Proc. Natl. Acad. Sci. USA* 1995;92:5915–5919. [PubMed: 7597053]
- Furger A, Schaller A, Schümperli D. Functional importance of conserved nucleotides at the histone RNA 3' processing site. *RNA* 1998;4:246–256. [PubMed: 9510327]
- Gall JG. Cajal bodies: The first 100 years. *Annu. Rev. Cell Dev. Biol* 2000;16 273–NIL25.
- Gall JG. The centennial of the Cajal body. *Nat. Rev. Mol. Cell Biol* 2003;4:975–980. [PubMed: 14685175]
- Galli G, Hofstetter H, Stunnenberg HG, Birnstiel ML. Biochemical complementation with RNA in the *Xenopus* oocyte: a small RNA is required for the generation of 3' histone mRNA termini. *Cell* 1983;34:823–828. [PubMed: 6194891]
- Georgiev O, Birnstiel ML. The conserved CAAGAAAGA spacer sequence is an essential element for the formation of 3' termini of the sea urchin H3 histone mRNA by RNA processing. *EMBO J* 1985;4:481–489. [PubMed: 2410259]
- Gick O, Kramer A, Keller W, Birnstiel ML. Generation of histone mRNA 3' ends by endonucleolytic cleavage of the pre-mRNA in a snRNP-dependent *in vitro* reaction. *EMBO J* 1986;5:1319–1326. [PubMed: 3015597]
- Gick O, Kramer A, Vasserot A, Birnstiel ML. Heat-labile regulatory factor is required for 3' processing of histone precursor mRNAs. *Proc. Natl. Acad. Sci. U. S. A* 1987;84:8937–8940. [PubMed: 2962194]
- Gilmartin GM. Eukaryotic mRNA 3' processing: a common means to different ends. *Genes Dev* 2005;19:2517–2521. [PubMed: 16264187]
- Gilmartin GM, Schaufele F, Schaffner G, Birnstiel ML. Functional analysis of the sea urchin U7 small nuclear RNA. *Mol. Cell. Biol* 1988;8:1076–1084. [PubMed: 2835659]
- Golembe TJ, Yong J, Dreyfuss G. Specific sequence features, recognized by the SMN complex, identify snRNAs and determine their fate as snRNPs. *Mol. Cell Biol* 2005;25:10989–11004. [PubMed: 16314521]
- Gorgoni B, Andrews S, Schaller A, Schümperli D, Gray NK, Muller B. The stem-loop binding protein stimulates histone translation at an early step in the initiation pathway. *RNA* 2005;11:1030–1042. [PubMed: 15987814]
- Grimm C, Stefanovic B, Schümperli D. The low abundance of U7 snRNA is partly determined by its Sm binding site. *EMBO J* 1993;12:1229–1238. [PubMed: 8458335]
- Gubitz AK, Feng W, Dreyfuss G. The SMN complex. *Exp. Cell Res* 2004;296:51–56. [PubMed: 15120993]
- Guhaniyogi J, Brewer G. Regulation of mRNA stability in mammalian cells. *Gene* 2001;265:11–23. [PubMed: 11255003]
- Harris ME, Böhni R, Schneiderman MH, Ramamurthy L, Schümperli D, Marzluff WF. Regulation of histone mRNA in the unperturbed cell cycle: Evidence suggesting control at two posttranscriptional steps. *Mol. Cell. Biol* 1991;11:2416–2424. [PubMed: 2017161]
- Heinz U, Adolph HW. Metallo-beta-lactamases: two binding sites for one catalytic metal ion? *Cell Mol. Life Sci* 2004;61:2827–2839. [PubMed: 15558212]
- Hermann H, Fabrizio P, Raker VA, Foulaki K, Hornig H, Brahms H, Lührmann R. snRNP Sm proteins share two evolutionarily conserved sequence motifs which are involved in Sm protein-protein interactions. *EMBO J* 1995;14:2076–2088. [PubMed: 7744013]
- Hernandez N, Weiner AM. Formation of the 3' end of U1 snRNA requires compatible snRNA promoter elements. *Cell* 1986;47:249–258. [PubMed: 3768956]
- Hirose Y, Manley JL. Creatine phosphate, not ATP, is required for 3' end cleavage of mammalian pre-mRNA *in vitro*. *J. Biol. Chem* 1997;272:29636–29642. [PubMed: 9368030]
- Hirose Y, Manley JL. RNA polymerase II is an essential mRNA polyadenylation factor. *Nature* 1998;395:93–96. [PubMed: 9738505]
- Hirose Y, Manley JL. RNA polymerase II and the integration of nuclear events. *Genes Dev* 2000;14:1415–1429. [PubMed: 10859161]

- Hofmann I, Schnolzer M, Kaufmann I, Franke WW. Symplekin, a constitutive protein of karyo- and cytoplasmic particles involved in mRNA biogenesis in *Xenopus laevis* oocytes. *Mol. Biol. Cell* 2002;13:1665–1676. [PubMed: 12006661]
- Huntley S, Baggott DM, Hamilton AT, Tran-Gyamfi M, Yang S, Kim J, Gordon L, Branscomb E, Stubbs L. A comprehensive catalog of human KRAB-associated zinc finger genes: insights into the evolutionary history of a large family of transcriptional repressors. *Genome Res* 2006;16:669–677. [PubMed: 16606702]
- Ingledue TC, Dominski Z, Sanchez R, Erkmann JA, Marzluff WF. Dual role for the RNA binding domain of *Xenopus laevis* SLBP1 in histone pre-mRNA processing. *RNA* 2000;6:1635–1648. [PubMed: 11105762]
- Jacobs EY, Ogiwara I, Weiner AM. Role of the C-terminal domain of RNA polymerase II in U2 snRNA transcription and 3' processing. *Mol. Cell Biol* 2004;24:846–855. [PubMed: 14701755]
- Jaeger S, Barends S, Giege R, Eriani G, Martin F. Expression of metazoan replication-dependent histone genes. *Biochimie* 2005;87:827–834. [PubMed: 16164992]
- Jaeger S, Martin F, Rudinger-Thirion J, Giege R, Eriani G. Binding of human SLBP on the 3'-UTR of histone precursor H4–12 mRNA induces structural rearrangements that enable U7 snRNA anchoring. *Nucleic Acids Res.* 2006
- Jenny A, Hauri H-P, Keller W. Characterization of cleavage and polyadenylation specificity factor and cloning of its 100-kilodalton subunit. *Mol. Cell. Biol* 1994;14:8183–8190. [PubMed: 7969155]
- Kambach C, Walke S, Young R, Avis JM, De la Fortelle E, Raker VA, Lührmann R, Li J, Nagai K. Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs. *Cell* 1999;96:375–387. [PubMed: 10025403]
- Kaufmann I, Martin G, Friedlein A, Langen H, Keller W. Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase. *EMBO J* 2004;23:616–626. [PubMed: 14749727]
- Kiss T. Biogenesis of small nuclear RNPs. *J. Cell Sci* 2004;117:5949–5951. [PubMed: 15564372]
- Kodama Y, Rothman JH, Sugimoto A, Yamamoto M. The stem-loop binding protein CDL-1 is required for chromosome condensation, progression of cell death and morphogenesis in *Caenorhabditis elegans*. *Development* 2002;129:187–196. [PubMed: 11782412]
- Kolev NG, Steitz JA. Symplekin and multiple other polyadenylation factors participate in 3'-end maturation of histone mRNAs. *Genes & Development* 2005;19
- Kolev NG, Steitz JA. In vivo assembly of functional U7 snRNP requires RNA backbone flexibility within the Sm-binding site. *Nat. Struct. Mol. Biol* 2006;13:347–353. [PubMed: 16547514]
- Krieg PA, Melton DA. Formation of the 3' end of histone mRNA by post-transcriptional processing. *Nature* 1984;308:203–206. [PubMed: 6700722]
- Kuhn U, Wahle E. Structure and function of poly(A) binding proteins. *Biochim. Biophys. Acta* 2004;1678:67–84. [PubMed: 15157733]
- Lenain C, Bauwens S, Amiard S, Brunori M, Giraud-Panis MJ, Gilson E. The Apollo 5' exonuclease functions together with TRF2 to protect telomeres from DNA repair. *Curr. Biol* 2006;16:1303–1310. [PubMed: 16730175]
- Li X, Hejna J, Moses RE. The yeast Snm1 protein is a DNA 5'-exonuclease. *DNA Repair (Amst)* 2005;4:163–170. [PubMed: 15590324]
- Liu JL, Murphy C, Buszczak M, Clatterbuck S, Goodman R, Gall JG. The *Drosophila melanogaster* Cajal body. *J. Cell Biol* 2006;172:875–884. [PubMed: 16533947]
- Luo W, Johnson AW, Bentley DL. The role of Rat1 in coupling mRNA 3'-end processing to transcription termination: implications for a unified allosteric-torpedo model. *Genes Dev* 2006;20:954–965. [PubMed: 16598041]
- Lüscher B, Schümperli D. RNA 3' processing regulates histone mRNA levels in a mammalian cell mutant. A processing factor becomes limiting in G1-arrested cells. *EMBO J* 1987;6:1721–1726. [PubMed: 3608992]
- Ma TL, Van Tine BA, Wei Y, Garrett MD, Nelson D, Adams PD, Wang J, Qin J, Chow LT, Harper JW. Cell cycle-regulated phosphorylation of p220^{NPAT} by cyclin E/Cdk2 in Cajal bodies promotes histone gene transcription. *Genes and Development* 2000;14:2298–2313. [PubMed: 10995387]

- Ma Y, Pannicke U, Schwarz K, Lieber MR. Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. *Cell* 2002;108:781–794. [PubMed: 11955432]
- MacDonald CC, Wilusz J, Shenk T. The 64-kilodalton subunit of the CstF polyadenylation factor binds to pre-mRNAs downstream of the cleavage site and influences cleavage site location. *Mol. Cell. Biol* 1994;14:6647–6654. [PubMed: 7935383]
- Mandel CR, Kaneko S, Zhang H, Gebauer D, Vethantham V, Manley JL, Tong L. Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease. *Nature*. 2006
- Mangus DA, Evans MC, Jacobson A. Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biol* 2003;4:223. [PubMed: 12844354]
- Martin F, Michel F, Zenklusen D, Müller B, Schümperli D. Positive and negative mutant selection in the human histone hairpin-binding protein using the yeast three-hybrid system. *Nucleic Acids Res* 2000;28:1594–1603. [PubMed: 10710426]
- Martin F, Schaller A, Eglite S, Schümperli D, Müller B. The gene for histone RNA hairpin binding protein is located on human chromosome 4 and encodes a novel type of RNA binding protein. *EMBO J* 1997;16:769–778. [PubMed: 9049306]
- Marzluff WF. Metazoan replication-dependent histone mRNAs: a distinct set of RNA polymerase II transcripts. *Curr. Opin. Cell Biol* 2005;17:274–280. [PubMed: 15901497]
- Marzluff WF, Duronio RJ. Histone mRNA expression: Multiple levels of cell cycle regulation and important developmental consequences. *Curr. Opin. Cell Biol* 2002;14:692–699. [PubMed: 12473341]
- Marzluff WF, Gongidi P, Woods KR, Jin JP, Maltais L. The human and mouse replication-dependent histone genes. *Genomics* 2002;80:487–498. [PubMed: 12408966]
- Mazumder B, Seshadri V, Fox PL. Translational control by the 3'-UTR: the ends specify the means. *Trends Biochem. Sci* 2003;28:91–98. [PubMed: 12575997]
- McCracken S, Fong N, Yankulov K, Ballantyne S, Pan GH, Greenblatt J, Patterson SD, Wickens M, Bentley DL. The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature* 1997;385:357–361. [PubMed: 9002523]
- Medlin J, Scurry A, Taylor A, Zhang F, Peterlin BM, Murphy S. P-TEFb is not an essential elongation factor for the intronless human U2 snRNA and histone H2b genes. *EMBO J* 2005;24:4154–4165. [PubMed: 16308568]
- Medlin JE, Uguen P, Taylor A, Bentley DL, Murphy S. The C-terminal domain of pol II and a DRB-sensitive kinase are required for 3' processing of U2 snRNA. *EMBO J* 2003;22:925–934. [PubMed: 12574128]
- Melin L, Soldati D, Mital R, Streit A, Schümperli D. Biochemical demonstration of complex formation of histone pre-mRNA with U7 small nuclear ribonucleoprotein and hairpin binding factors. *EMBO J* 1992;11:691–697. [PubMed: 1531633]
- Miele A, Braastad CD, Holmes WF, Mitra P, Medina R, Xie R, Zaidi SK, Ye X, Wei Y, Harper JW, Van Wijnen AJ, Stein JL, Stein GS. HiNF-P directly links the cyclin E/CDK2/p220NPAT pathway to histone H4 gene regulation at the G1/S phase cell cycle transition. *Mol. Cell Biol* 2005;25:6140–6153. [PubMed: 15988025]
- Mignone F, Gissi C, Liuni S, Pesole G. Untranslated regions of mRNAs. *Genome Biol* 2002;3 REVIEWS0004.
- Mital R, Albrecht U, Schümperli D. Detection of UV-induced RNA:protein crosslinks in snRNPs by oligonucleotides complementary to the snRNA. *Nucleic Acids Res* 1993;21:1049–1050. [PubMed: 8451179]
- Mowry KL, Oh R, Steitz JA. Each of the conserved sequence elements flanking the cleavage site of mammalian histone pre-mRNAs has a distinct role in the 3'-end processing reaction. *Mol. Cell. Biol* 1989;9:3105–3108. [PubMed: 2779556]
- Mowry KL, Steitz JA. Both conserved signals on mammalian histone pre-mRNAs associate with small nuclear ribonucleoproteins during 3' end formation in vitro. *Mol. Cell. Biol* 1987a;7:1663–1672. [PubMed: 2955216]

- Mowry KL, Steitz JA. Identification of the human U7 snRNP as one of several factors involved in the 3' end maturation of histone pre-messenger RNA's. *Science* 1987b;238:1682–1687. [PubMed: 2825355]
- Mowry KL, Steitz JA. snRNP Mediators of 3' end processing: Functional fossils. *Trends Biochem. Sci* 1988;13:447–451. [PubMed: 2908086]
- Muller B, Schumperli D. The U7 snRNP and the hairpin binding protein: Key players in histone mRNA metabolism. *Semin. Cell Dev. Biol* 1997;8:567–576. [PubMed: 9642171]
- Murthy KG, Manley JL. Characterization of the multisubunit cleavage-polyadenylation specificity factor from calf thymus. *J. Biol. Chem* 1992;267:14804–14811. [PubMed: 1634525]
- Murthy KGK, Manley JL. The 160-kD subunit of human cleavage polyadenylation specificity factor coordinates pre-mRNA 3'-end formation. *Genes and Development* 1995;9:2672–2683. [PubMed: 7590244]
- Osley MA. The regulation of histone synthesis in the cell cycle. *Annu. Rev. Biochem* 1991;60:827–861. [PubMed: 1883210]
- Pandey NB, Williams AS, Sun J-H, Brown VD, Bond U, Marzluff WF. Point mutations in the stem-loop at the 3' end of mouse histone mRNA reduce expression by reducing the efficiency of 3' end formation. *Mol. Cell. Biol* 1994;14:1709–1720. [PubMed: 8114706]
- Pesole G, Liuni S, Grillo G, Licciulli F, Mignone F, Gissi C, Saccone C. UTRdb and UTRsite: specialized databases of sequences and functional elements of 5' and 3' untranslated regions of eukaryotic mRNAs. Update 2002. *Nucleic Acids Res* 2002;30:335–340. [PubMed: 11752330]
- Peterlin BM, Price DH. Controlling the elongation phase of transcription with P-TEFb. *Mol. Cell* 2006;23:297–305. [PubMed: 16885020]
- Pettitt J, Crombie C, Schumperli D, Müller B. The *Caenorhabditis elegans* histone hairpin-binding protein is required for core histone expression and is essential for embryonic and postembryonic cell division. *J. Cell Science* 2002;115:857–866. [PubMed: 11865041]
- Phatnani HP, Greenleaf AL. Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev* 2006;20:2922–2936. [PubMed: 17079683]
- Pillai RS, Grimmmer M, Meister G, Will CL, Luhrmann R, Fischer U, Schumperli D. Unique Sm core structure of U7 snRNPs: assembly by a specialized SMN complex and the role of a new component, Lsm11, in histone RNA processing. *Genes Dev* 2003;17:2321–2333. [PubMed: 12975319]
- Pillai RS, Will CL, Luhrmann R, Schumperli D, Müller B. Purified U7 snRNPs lack the Sm proteins D1 and D2 but contain Lsm10, a new 14 kDa Sm D1-like protein. *EMBO J* 2001;20:5470–5479. [PubMed: 11574479]
- Proudfoot N. New perspectives on connecting messenger RNA 3' end formation to transcription. *Curr. Opin. Cell Biol* 2004;16:272–278. [PubMed: 15145351]
- Proudfoot NJ, Furger A, Dye MJ. Integrating mRNA processing with transcription. *Cell* 2002;108:501–512. [PubMed: 11909521]
- Raker VA, Hartmuth K, Kastner B, Luhrmann R. Spliceosomal U snRNP core assembly: Sm proteins assemble onto an Sm site RNA nonnucleotide in a specific and thermodynamically stable manner. *Mol. Cell. Biol* 1999;19:6554–6565. [PubMed: 10490595]
- Raker VA, Plessel G, Luhrmann R. The snRNP core assembly pathway: Identification of stable core protein heteromeric complexes and an snRNP subcore particle *in vitro*. *EMBO J* 1996;15:2256–2269. [PubMed: 8641291]
- Rosonina E, Kaneko S, Manley JL. Terminating the transcript: breaking up is hard to do. *Genes Dev* 2006;20:1050–1056. [PubMed: 16651651]
- Ross J. mRNA stability in mammalian cells. *Microbiol. Rev* 1995;59:423–450. [PubMed: 7565413]
- Rueggsegger U, Beyer K, Keller W. Purification and characterization of human cleavage factor Im involved in the 3' end processing of messenger RNA precursors. *J. Biol. Chem* 1996;271:6107–6113. [PubMed: 8626397]
- Rueggsegger U, Blank D, Keller W. Human pre-mRNA cleavage factor Im is related to spliceosomal SR proteins and can be reconstituted *in vitro* from recombinant subunits. *Molecular Cell* 1998;1:243–253. [PubMed: 9659921]
- Ryan K, Calvo O, Manley JL. Evidence that polyadenylation factor CPSF-73 is the mRNA 3' processing endonuclease. *RNA* 2004;10:565–573. [PubMed: 15037765]

- Sanchez R, Marzluff WF. The stem-loop binding protein is required for efficient translation of histone mRNA *in vivo* and *in vitro*. *Mol. Cell Biol* 2002;22:7093–7104. [PubMed: 12242288]
- Sanchez R, Marzluff WF. The oligo(A) tail on histone mRNA plays an active role in translational silencing of histone mRNA during *Xenopus* oogenesis. *Mol. Cell Biol* 2004;24:2513–2525. [PubMed: 14993288]
- Scharl EC, Steitz JA. Length suppression in histone messenger RNA 3'-end maturation: Processing defects of insertion mutant pre-messenger RNAs can be compensated by insertions into the U7 small nuclear RNA. *Proc. Natl. Acad. Sci. USA* 1996;93:14659–14664. [PubMed: 8962110]
- Scharl EC, Steitz JA. The site of 3' end formation of histone messenger RNA is a fixed distance from the downstream element recognized by the U7 snRNP. *EMBO J* 1994;13:2432–2440. [PubMed: 8194533]
- Schafele F, Gilmartin GM, Bannwarth W, Birnstiel ML. Compensatory mutations suggest that base-pairing with a small nuclear RNA is required to form the 3' end of H3 messenger RNA. *Nature* 1986;323:777–781. [PubMed: 3022153]
- Schul W, Groenhout B, Koberna K, Takagaki Y, Jenny A, Manders EMM, Raska I, Van Driel R, De Jong L. The RNA 3' cleavage factors CstF 64 kDa and CPSF 100 kDa are concentrated in nuclear domains closely associated with coiled bodies and newly synthesized RNA. *EMBO J* 1996;15:2883–2892. [PubMed: 8654386]
- Schul W, van Driel R, Materna AG, Van Driel R, De Jong L. Nuclear domains enriched in RNA 3'-processing factors associate with coiled bodies and histone genes in a cell cycle-dependent manner. *Mol. Biol. Cell* 1999;10:3815–3824. [PubMed: 10564273]
- Schumperli D, Pillai RS. The special Sm core structure of the U7 snRNP: far-reaching significance of a small nuclear ribonucleoprotein. *Cell Mol. Life Sci* 2004;61:2560–2570. [PubMed: 15526162]
- SenGupta DJ, Zhang BL, Kraemer B, Prochart P, Fields S, Wickens M. A three-hybrid system to detect RNA-protein interactions *in vivo*. *Proc. Natl. Acad. Sci. USA* 1996;93:8496–8501. [PubMed: 8710898]
- Smith HO, Tabiti K, Schaffner G, Soldati D, Albrecht U, Birnstiel ML. Two-step affinity purification of U7 small nuclear ribonucleoprotein particles using complementary biotinylated 2'-*O*-methyl oligoribonucleotides. *Proc. Natl. Acad. Sci. USA* 1991;88:9784–9788. [PubMed: 1835087]
- Soldati D, Schumperli D. Structural and functional characterization of mouse U7 small nuclear RNA active in 3' processing of histone pre-mRNA. *Mol. Cell. Biol* 1988;8:1518–1524. [PubMed: 3380087]
- Spycher C, Streit A, Stefanovic B, Albrecht D, Koning THW, Schumperli D. 3' end processing of mouse histone pre-mRNA: Evidence for additional base-pairing between U7 snRNA and pre-mRNA. *Nucleic Acids Res* 1994;22:4023–4030. [PubMed: 7937126]
- Stanek D, Neugebauer KM. The Cajal body: a meeting place for spliceosomal snRNPs in the nuclear maze. *Chromosoma* 2006;115:343–354. [PubMed: 16575476]
- Stauber C, Schumperli D. 3' processing of pre-mRNA plays a major role in proliferation-dependent regulation of histone gene expression. *Nucleic Acids Res* 1988;16:9399–9413. [PubMed: 3141900]
- Stefanovic B, Hackl W, Lüthmann R, Schumperli D. Assembly, nuclear import and function of U7 snRNPs studied by microinjection of synthetic U7 RNA into *Xenopus* oocytes. *Nucleic Acids Res* 1995;23:3141–3151. [PubMed: 7667090]
- Stein GS, Stein JL, Van Wijnen AJ, Lian JB. Transcriptional control of cell cycle progression: The histone gene is a paradigm for the G1/S phase and proliferation differentiation transitions. *Cell Biol. Int* 1996;20:41–49. [PubMed: 8936406]
- Streit A, Koning TW, Soldati D, Melin L, Schumperli D. Variable effects of the conserved RNA hairpin element upon 3' end processing of histone pre-mRNA *in vitro*. *Nucleic Acids Res* 1993;21:1569–1575. [PubMed: 8479907]
- Strub K, Birnstiel ML. Genetic complementation in the *Xenopus* oocyte: co-expression of sea urchin histone and U7 RNAs restores 3' processing of H3 pre-mRNA in the oocyte. *EMBO J* 1986;5:1675–1682. [PubMed: 2943587]
- Strub K, Galli G, Busslinger M, Birnstiel ML. The cDNA sequences of the sea urchin U7 small nuclear RNA suggest specific contacts between histone mRNA precursor and U7 RNA during RNA processing. *EMBO J* 1984;3:2801–2807. [PubMed: 6084590]

- Sullivan E, Santiago C, Parker ED, Dominski Z, Yang X, Lanzotti DJ, Ingledue TC, Marzluff WF, Duronio RJ. Drosophila stem loop binding protein coordinates accumulation of mature histone mRNA with cell cycle progression. *Genes Dev* 2001;15:173–187. [PubMed: 11157774]
- Takagaki Y, Manley JL. Complex protein interactions within the human polyadenylation machinery identify a novel component. *Mol. Cell. Biol* 2000;20:1515–1525. [PubMed: 10669729]
- Takagaki Y, Ryner LC, Manley JL. Separation and characterization of a poly(A) polymerase and a cleavage/specificity factor required for pre-mRNA polyadenylation. *Cell* 1988;52:731–742. [PubMed: 2830992]
- Terns MP, Jacob ST. Role of poly(A) polymerase in the cleavage and polyadenylation of mRNA precursor. *Mol. Cell. Biol* 1989;9:1435–1444. [PubMed: 2566910]
- Tian B, Hu J, Zhang H, Lutz CS. A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Res* 2005;33:201–212. [PubMed: 15647503]
- Tollervey D. Molecular biology: termination by torpedo. *Nature* 2004;432:456–457. [PubMed: 15565140]
- Uguen P, Murphy S. The 3' ends of human pre-snRNAs are produced by RNA polymerase II CTD-dependent RNA processing. *EMBO J* 2003;22:4544–4554. [PubMed: 12941705]
- Urlaub H, Raker VA, Kostka S, Lührmann R. Sm protein-Sm site RNA interactions within the inner ring of the spliceosomal snRNP core structure. *EMBO J* 2001;20:187–196. [PubMed: 11226169]
- Vasserot AP, Schaufele FJ, Birnstiel ML. Conserved terminal hairpin sequences of histone mRNA precursors are not involved in duplex formation with the U7 RNA but act as a target site for a distinct processing factor. *Proc. Natl. Acad. Sci. USA* 1989;86:4345–4349. [PubMed: 2734288]
- Venkataraman K, Brown KM, Gilmartin GM. Analysis of a noncanonical poly(A) site reveals a tripartite mechanism for vertebrate poly(A) site recognition. *Genes Dev* 2005;19:1315–1327. [PubMed: 15937220]
- Vogel A, Schilling O, Spath B, Marchfelder A. The tRNase Z family of proteins: physiological functions, substrate specificity and structural properties. *Biol. Chem* 2005;386:1253–1264. [PubMed: 16336119]
- Wagner EJ, Marzluff WF. ZFP100, a component of the active U7 snRNP limiting for histone pre-mRNA processing, is required for entry into S phase. *Mol. Cell Biol* 2006;26:6702–6712. [PubMed: 16914750]
- Wagner EJ, Ospina JK, Hu Y, Dundr M, Matera AG, Marzluff WF. Conserved zinc fingers mediate multiple functions of ZFP100, a U7snRNP associated protein. *RNA* 2006;12:1206–1218. [PubMed: 16714279]
- Wahle E. Poly(A) tail length control is caused by termination of processive synthesis. *J. Biol. Chem* 1995;270:2800–2808. [PubMed: 7852352]
- Wahle E, Ruegsegger U. 3'-End processing of pre-mRNA in eukaryotes. *FEMS Microbiol. Rev* 1999;23:277–295. [PubMed: 10371034]
- Walther TN, Wittop KT, Schümperli D, Müller B. A 5'-3' exonuclease activity involved in forming the 3' products of histone pre-mRNA processing in vitro. *RNA* 1998;4:1034–1046. [PubMed: 9740123]
- Wang Z, Fast W, Valentine AM, Benkovic SJ. Metallo-beta-lactamase: structure and mechanism. *Curr. Opin. Chem. Biol* 1999a;3:614–622. [PubMed: 10508665]
- Wang Z-F, Ingledue TC, Dominski Z, Sanchez R, Marzluff WF. Two *Xenopus* proteins that bind the 3' end of histone mRNA: Implications for translational control of histone synthesis during oogenesis. *Mol. Cell. Biol* 1999b;19:835–845. [PubMed: 9858606]
- Wang Z-F, Whitfield ML, Ingledue TI, Dominski Z, Marzluff WF. The protein which binds the 3' end of histone mRNA: a novel RNA-binding protein required for histone pre-mRNA processing. *Genes and Development* 1996;10:3028–3040. [PubMed: 8957003]
- Weiner AM. E Pluribus Unum: 3' end formation of polyadenylated mRNAs, histone mRNAs, and U snRNAs. *Mol. Cell* 2005;20:168–170. [PubMed: 16246719]
- Whitfield ML, Kaygun H, Erkmann JA, Townley-Tilson WH, Dominski Z, Marzluff WF. SLBP is associated with histone mRNA on polyribosomes as a component of the histone mRNP. *Nucleic Acids Res* 2004;32:4833–4842. [PubMed: 15358832]

- Whitfield ML, Zheng L-X, Baldwin A, Ohta T, Hurt MM, Marzluff WF. Stem-loop binding protein, the protein that binds the 3' end of histone mRNA, is cell cycle regulated by both translational and posttranslational mechanisms. *Mol. Cell. Biol* 2000;20:4188–4198. [PubMed: 10825184]
- Wickens M, Bernstein DS, Kimble J, Parker R. A PUF family portrait: 3'UTR regulation as a way of life. *TIG* 2002;18:150–157. [PubMed: 11858839]
- Wickens M, Gonzalez TN. Molecular biology. Knives, accomplices, and RNA. *Science* 2004;306:1299–1300. [PubMed: 15550648]
- Will CL, Luhrmann R. Spliceosomal UsnRNP biogenesis, structure and function. *Curr. Opin. Cell Biol* 2001;13:290–301. [PubMed: 11343899]
- Williams AS, Marzluff WF. The sequence of the stem and flanking sequences at the 3' end of histone mRNA are critical determinants for the binding of the stem-loop binding protein. *Nucleic Acids Res* 1995;23:654–662. [PubMed: 7899087]
- Wu C-HH, Gall JG. U7 small nuclear RNA in C snurposomes of the *Xenopus* germinal vesicle. *Proc. Natl. Acad. Sci. USA* 1993;90:6257–6259. [PubMed: 8327506]
- Yang XC, Purdy M, Marzluff WF, Dominski Z. Characterization of 3'hExo, a 3' exonuclease specifically interacting with the 3' end of histone mRNA. *J. Biol. Chem* 2006;281:30447–30454. [PubMed: 16912046]
- Ye X, Wei Y, Nalepa G, Harper JW. The cyclin E/Cdk2 substrate p220(NPAT) is required for S-phase entry, histone gene expression, and Cajal body maintenance in human somatic cells. *Mol. Cell Biol* 2003;23:8586–8600. [PubMed: 14612403]
- Zhang BL, Kraemer B, SenGupta D, Fields S, Wickens M. Yeast three-hybrid system to detect and analyze interactions between RNA and protein. *Methods Enzymol* 1999;306:93–113. [PubMed: 10432449]
- Zhao J. Coordination of DNA synthesis and histone gene expression during normal cell cycle progression and after DNA damage. *Cell Cycle* 2004;3:695–697. [PubMed: 15153807]
- Zhao J, Hyman L, Moore C. Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol. Mol. Biol. Rev* 1999a;63:405–445. [PubMed: 10357856]
- Zhao J, Kessler M, Helmling S, O'Connor JP, Moore C. Pta1, a component of yeast CFII, is required for both cleavage and poly(A) addition of mRNA precursor. *Mol. Cell. Biol* 1999b;19:7733–7740. [PubMed: 10523662]
- Zhao JY, Kennedy BK, Lawrence BD, Barbie DA, Matera AG, Fletcher JA, Harlow E. NPAT links cyclin E-Cdk2 to the regulation of replication-dependent histone gene transcription. *Genes and Development* 2000;14:2283–2297. [PubMed: 10995386]
- Zheng L-X, Dominski Z, Yang X, Elms P, Raska CS, Borchers CH, Marzluff WF. Phosphorylation of SLBP on two threonines triggers degradation of SLBP, the sole cell-cycle regulated factor required for regulation of histone mRNA processing, at the end of S-phase. *Mol. Cell Biol* 2003;23:1590–1601. [PubMed: 12588979]

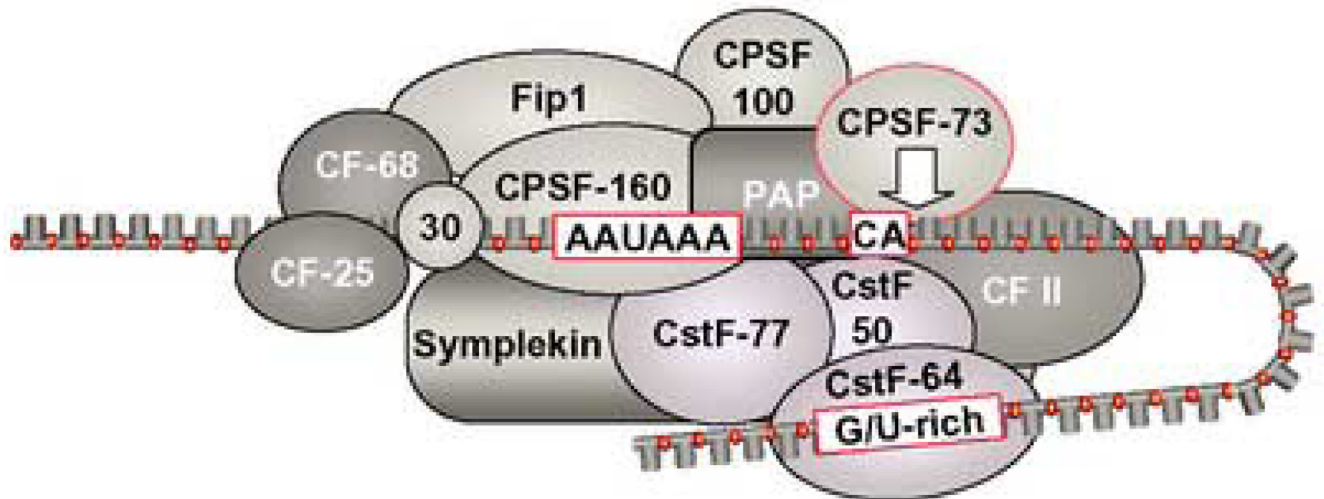


Figure 1.

The processing machinery involved in cleavage/polyadenylation. The arrow indicates the site of cleavage by CPSF-73. The position of some of the components in the complex has not been experimentally supported. The model does not include the CTD that is required for *in vitro* cleavage/polyadenylation under certain experimental conditions.

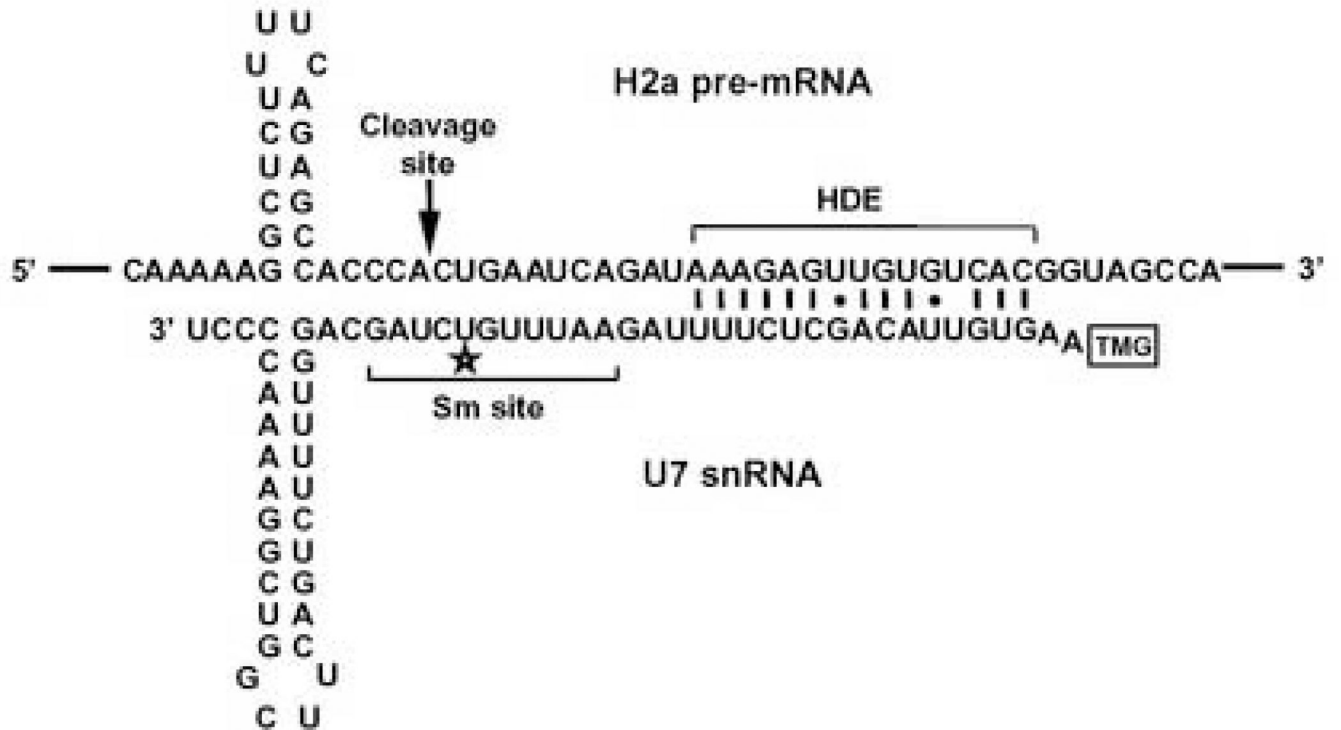


Figure 2.

The sequence and predicted configuration of the mouse H2a-614 pre-mRNA (top) and the mouse U7 snRNA (bottom) during 3' end processing. The base pair interaction between the two RNAs is indicated with vertical lines (Watson-Crick base pairs) and dots (GU base pairs). The tri-methyl guanosine cap (TMG) at the 5' end of the U7 snRNA is framed. The unique U7 Sm site, the histone downstream element (HDE) and the site of cleavage are indicated. The region of the U7 Sm site that UV cross-links with Lsm11 is labeled with the asterisk.

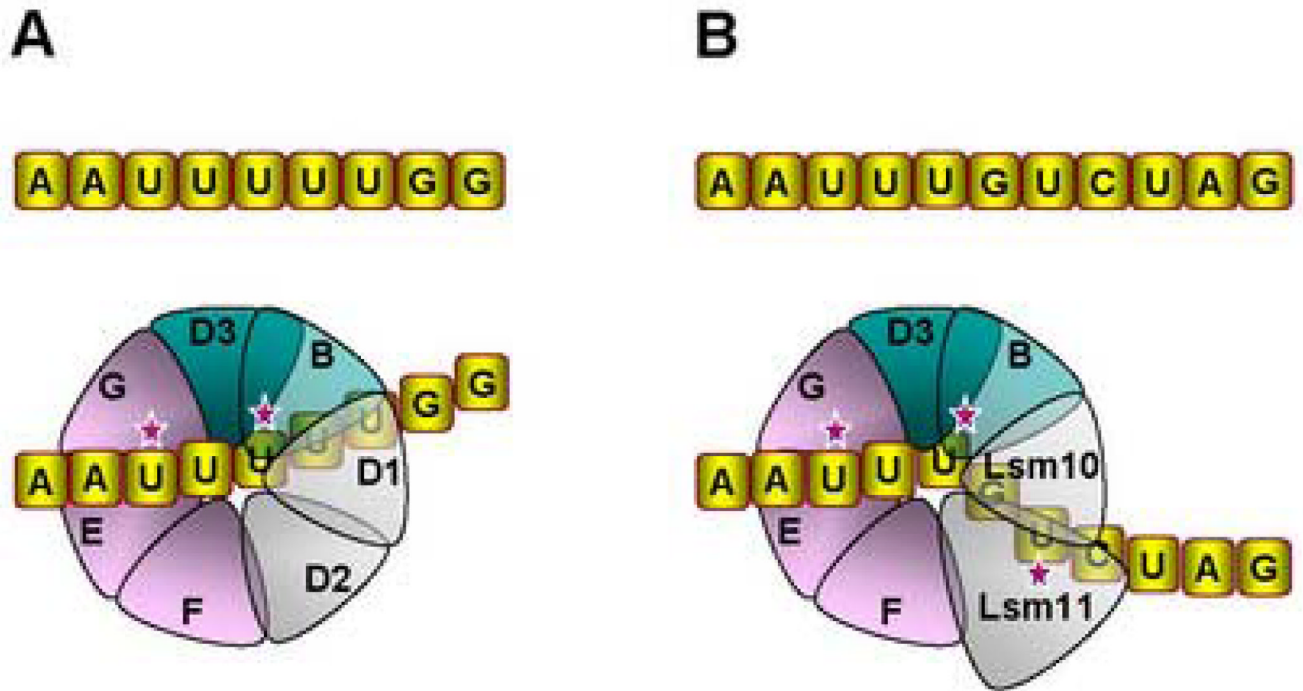


Figure 3. Two types of the Sm core in snRNPs. Proteins that make direct contacts with nucleotides of the Sm site are marked with asterisks. **A.** The composition of the canonical Sm core assembled on the Sm site in the spliceosomal snRNAs. **B.** The unique composition of the Sm core assembled on the U7-specific Sm site.



Figure 4.
A schematic mode of binding of SLBP and 3'hExo to the 3' end of histone mRNA.

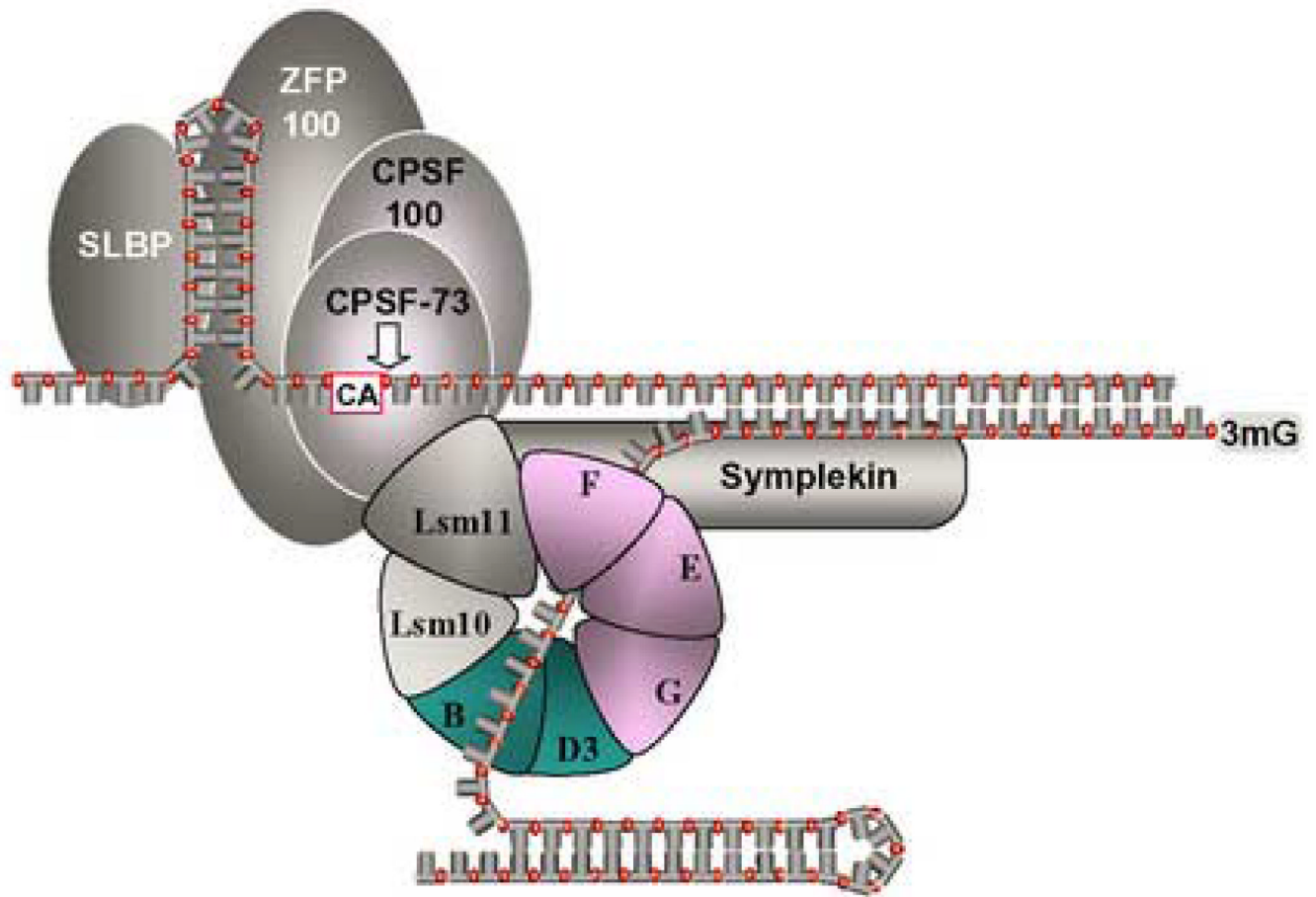


Figure 5. Known components of the 3' end processing machinery cleaving histone pre-mRNAs. The position of symplekin in the complex is hypothetical and the presence of CPSF-100 has not been supported experimentally.