

Eukaryotic mRNA 3' processing: a common means to different ends

Gregory M. Gilmartin¹

University of Vermont, Burlington, Vermont 05405, USA

In principle, the formation of the 3' end of a eukaryotic mRNA is a simple process. At a minimum, it involves the hydrolysis of a single phosphodiester bond in the nascent transcript. For one unique class of mRNAs, transcripts of the metazoan replication-dependent histone genes, this is indeed the only requirement. All other eukaryotic mRNAs require both cleavage and subsequent poly(A) addition to the newly generated 3' hydroxyl. Despite this apparent simplicity, and more than two decades of work, basic questions concerning the mechanisms of 3'-end formation for both classes of transcripts persist. The accumulated evidence indicates that the mechanisms by which the ends of histone and polyadenylated mRNA are formed are fundamentally distinct, both in the factors responsible for processing and the RNA sequence elements that direct their action (for review, see Marzluff 2005; Zhao et al. 1999a). The 3' processing of polyadenylated mRNAs requires a complex of over a dozen proteins, nearly all of which are conserved from yeast to humans. In contrast, the formation of the 3' ends of the metazoan replication-dependent histone transcripts not only requires a unique set of proteins, but an essential snRNA component as well. In both cases, the identity of the endonuclease has remained largely a matter of speculation.

Two reports—one from Kolev and Steitz (2005) in this issue of *Genes & Development*, and the second from Dominski et al. (2005a) have revealed that the mechanisms of poly(A) site processing and histone 3' processing are not quite as unique as they appear. Surprisingly, the two processing machines share a common core of proteins that almost certainly includes the endonuclease. The most exciting aspect of these reports, however, is that they provide a fresh insight into the evolution of what had seemed to be needlessly redundant mechanisms for generating mRNA 3' ends. In essence, these reports indicate that the enzymatic machinery responsible for cleaving all nascent pre-mRNAs is likely to be identical; the differences lie in the elaborate mechanisms that have evolved to recruit this machinery to the two classes of transcripts. The driving forces responsible

for the evolution of two distinct 3' processing complexes are likely to involve the conflicting pressures of post-transcriptional regulation: coordinate cell cycle control of the replication-dependent histone mRNAs versus the developmental and cell-type-specific regulation of alternative poly(A) site selection.

Similarities among contrasts

The RNA sequence elements that direct the 3' processing of histone and polyadenylated mRNAs could hardly be more different (Fig. 1A). Three elements contribute to mammalian poly(A) site recognition, each of which is recognized by a distinct protein complex (Zhao et al. 1999a). The most well conserved element is the AAUAAA hexamer that generally resides between 10 and 30 nucleotides (nt) upstream of the cleavage site. This element is recognized by a five-subunit complex termed cleavage and polyadenylation specificity factor (CPSF). Downstream of the cleavage site is an amorphous U-rich element that is recognized by the heterotrimeric cleavage stimulatory factor (CstF). A third element of the form UGUA is often present in one or more copies at a variable distance upstream of the cleavage site (Hu et al. 2005), and is recognized by the heterodimeric cleavage factor I_m (CFI_m) (Brown and Gilmartin 2003; Venkataraman et al. 2005). No conserved RNA structural elements appear to be shared among poly(A) sites; on the contrary, studies have indicated that poly(A) site recognition is facilitated by an "open" or unstructured conformation (Gimmi et al. 1989; Graveley et al. 1996).

In contrast to polyadenylated transcripts, the 3' processing of histone pre-mRNAs requires both conserved sequence and structural elements (Marzluff 2005). The 3' end of the mature mammalian histone mRNA possesses a highly conserved 26-nt sequence, encompassing a 16-nt stem-loop, located 24–70 nt downstream of the stop codon (Fig. 1B). This element is bound by the stem-loop-binding protein (SLBP) that functions in 3' processing, as well as translation and the coupling of message stability to DNA replication and the cell cycle. Endonucleolytic cleavage occurs 5 nt downstream of the stem-loop. Nine to 12 nt downstream of the cleavage site is the histone downstream element (HDE) that functions in 3' process-

¹Correspondence.

E-MAIL ggilmart@zoo.uvm.edu; FAX (802) 656-8808.

Article and publication are at <http://www.genesdev.org/cgi/doi/10.1101/gad.1378105>.

stop codon and the 3' end of the transcript (Kaygun and Marzluff 2005b). The primary candidate for the nuclease responsible for regulated histone mRNA decay is a 3'-to-5' exonuclease, termed 3' hExo (Dominski et al. 2003). High-affinity binding of 3' hExo to the histone mRNA requires both the stem-loop and the highly conserved 3'-terminal ACCCA sequence. Extension of the RNA beyond the ACCCA terminus reduced the binding of 3' hExo by two orders of magnitude. 3' hExo binds the stem-loop independently as well as simultaneously with SLBP. When bound together with SLBP, 3' hExo degradation of the histone mRNA is blocked.

The common set of 3' ends shared by the replication-dependent histone mRNAs is essential for their regulation. Precise pre-mRNA processing is required to generate a class of transcripts in which the RNA-binding proteins required for post-transcriptional regulation are strictly positioned with respect to the 3' end. Cleavage site precision is achieved by the base-pairing of U7 snRNA with the histone pre-mRNA HDE within a very narrowly defined window downstream of the cleavage site (Scharl and Steitz 1994, 1996; Cho et al. 1995). Scharl and Steitz (1994) proposed that the HDE:U7 snRNP interaction functions as a "molecular ruler" to precisely determine the site of cleavage. The binding of U7 snRNP to the pre-mRNA is necessary, and in some cases sufficient, for the recruitment of a functional 3' processing complex, as well as for cleavage site precision. The use of RNA:RNA base-pairing in *trans* to confer catalytic specificity can be seen in a wide variety of processes, including pre-mRNA splicing, translation, snoRNA-directed rRNA processing and modification, and the action of miRNA and siRNAs.

A comparable set of rigid constraints is not observed in the 3' processing of polyadenylated mRNAs. Poly(A) sites exhibit a relatively wide range of both sequence and spatial variability. As depicted in Figure 1A, the AAUAAA hexamer resides within a window of ~20 nt upstream of the cleavage site. Although the AAUAAA hexamer is the most conserved element among mammalian poly(A) sites, only ~53% of human poly(A) sites contain this sequence. An AUUAAA hexamer is present in ~17% of human poly(A) sites, and the remaining ~30% contain other hexamer variants (Tian et al. 2005). The position of the downstream element, which lacks an identified consensus, varies within an ~30-nt window with respect to the cleavage site. Although flexible, the spatial arrangement of these sequences appears to be more constrained than those of the UGUA elements, which vary not only in their position relative to the cleavage site, but in copy number as well (Brown and Gilmartin 2003; Hu et al. 2005; Venkataraman et al. 2005). The variability in poly(A) site organization is reflected in the variability of cleavage site choice. Bioinformatic studies indicate that cleavage site choice within a poly(A) site is inherently imprecise (Pauws et al. 2001; Tian et al. 2005). In addition to cleavage site heterogeneity, alternative selection among different poly(A) sites (that range over distances in the kilobase range) is observed within the majority of human pre-mRNAs

(Iseli et al. 2002; Tian et al. 2005). The consequences of alternative poly(A) site choice may impact the protein coding capacity of the message, as well as its localization, translation efficiency, and stability (Edwards-Gilbert et al. 1997). Metazoan 3' UTRs have been found to function as repositories for a wide variety of post-transcriptional regulatory elements (recognized by both proteins and miRNAs) that may function in a developmental or tissue-specific manner (Kuersten and Goodwin 2003; Xie et al. 2005). While there are cases of alternative 3' processing of histone genes, the only choice observed is between the nonadenylated processing site and a downstream poly(A) site (Dominski and Marzluff 1999).

An intrinsic difference between the 3' processing of replication-dependent histone mRNAs and polyadenylated mRNAs, as observed by Birnstiel et al. (1985), is therefore the precision and uniformity of the former class and the diversity and flexibility of the latter. The metazoan replication-dependent histone mRNA 3' processing machinery appears to have evolved to support the coordinate post-transcriptional regulation of an entire class of transcripts through its ability to generate a common set of 3' ends. In contrast, the polyadenylation machinery appears to have evolved to support a diverse array of developmental and tissue-specific post-transcriptional regulatory responses. Thus while the evidence indicates that the 3' processing complexes of these two classes of transcripts share a common catalytic core, the mechanisms that have been developed to specify the site of catalysis are distinct. The mechanisms for the recruitment of the catalytic core to the pre-mRNA are not restricted to the factors bound at the 3' ends, but also involve both splicing and transcription. Whereas poly(A) site processing is clearly coupled to splicing (Minviella-Sebastia and Keller 1999), the 3' processing of metazoan replication-dependent histone pre-mRNAs (which uniformly lack introns) appears to be inherently incompatible with splicing (Pandey et al. 1990). Furthermore, whereas poly(A) site processing is functionally coupled to transcription (Adamson et al. 2005), this does not appear to be the case for histone mRNA 3' processing (Adamson and Price 2003).

A common catalytic core

Kolev and Steitz (2005) have identified all five subunits of CPSF, two subunits of CstF, and symplekin as constituents of a factor termed HLF (heat labile factor) that is required for histone 3' processing *in vitro*. CPSF and symplekin not only function in poly(A) site processing in the nucleus (Zhao et al. 1999a,b; Xing et al. 2004), they are also required for developmentally regulated poly(A) addition in the cytoplasm (Barnard et al. 2004). Dominiski et al. (2005a) have captured the 73-kDa subunit of CPSF at the site of cleavage of a mouse H2A pre-mRNA by UV cross-linking. Their evidence suggests that CPSF-73 functions as both the endonuclease that cleaves the nascent RNA and as the 5'-to-3' exonuclease that initiates the degradation of the downstream cleavage product. The presence of CPSF-73 at the histone pre-mRNA

3' cleavage site provides new support for the previously proposed role of this protein as the poly(A) site endonuclease. An endonucleolytic function for CPSF-73 was initially suggested by the identification of a domain shared with a group of predicted or proven nucleases, termed the β -CASP domain (Callebaut et al. 2002). The work of Ryan et al. (2004) provided additional support for CPSF-73 as the poly(A) site endonuclease, but conclusive evidence has remained elusive.

Several key questions remain to be addressed. The histone mRNA 3' processing and polyadenylation complexes clearly share a core set of proteins that likely encompasses the endonuclease, but the composition of this core remains to be defined. Symplekin, a proposed assembly/scaffolding factor (Takagaki and Manley 2000), is clearly an essential component, and CPSF-73 is very likely to be as well, but the additional CPSF and CstF subunits identified as constituents of HLF are as yet guilty by association. Thus the complete complement of factors required for 3' processing has yet to be conclusively identified for either metazoan histone mRNAs or for polyadenylated mRNAs. Furthermore, the absence of nonpolyadenylated histone mRNAs in all organisms other than metazoans prompts the question of what mechanisms these organisms have evolved to regulate histone gene expression. The apparent conservation of a 5'-to-3' exonuclease activity along with an endonuclease activity in the processing of these two classes of transcripts raises the question as to whether it may have a role in transcription termination. Exonucleases have been proposed to contribute to RNA polymerase II transcription termination in both yeast and humans (Kim et al. 2004; West et al. 2004), and transcription termination has been shown to be coupled to the 3' processing of both histone and polyadenylated mRNAs (Chodchoy et al. 1991; Proudfoot et al. 2002).

Taken together, the reports of Kolev and Steitz (2005) and Dominski et al. (2005a) demonstrate that while the 3' processing of metazoan replication-dependent histone genes is unique in its outcome, its mechanism is not. These findings indicate that the 3' processing machinery of all eukaryotic mRNA likely share a common evolutionary origin. Furthermore, the divergence of histone and poly(A) site processing may simply reflect a selection for precision versus flexibility. The essential difference between the two classes of mRNAs appears to be that they have evolved distinct mechanisms for the recruitment and activation of a shared catalytic activity. Which just goes to show, it's how you get there that makes all the difference in the end.

Acknowledgments

I thank Zbigniew Dominski, Xiao-cui Yang, and Bill Marzluff for sharing their work prior to publication.

References

Adamson, T.E. and Price, D.H. 2003. Cotranscriptional processing of *Drosophila* histone mRNAs. *Mol. Cell. Biol.* **23**: 4046–4055.

- Adamson, T.E., Shutt, D.C., and Price, D.H. 2005. Functional coupling of cleavage and polyadenylation with transcription of mRNA. *J. Biol. Chem.* **280**: 32262–32271.
- Barnard, D.C., Ryan, K., Manley, J.L., and Richter, J.D. 2004. Symplekin and xGLD-2 are required for CPEB-mediated cytoplasmic polyadenylation. *Cell* **119**: 641–651.
- Birnstiel, M.L., Busslinger, M., and Strub, K. 1985. Transcription termination and 3' processing: The end is in site! *Cell* **41**: 349–359.
- Brown, K.M. and Gilmartin, G.M. 2003. A mechanism for the regulation of pre-mRNA 3' processing by human cleavage factor I_m. *Mol. Cell* **12**: 1467–1476.
- Callebaut, I., Moshous, D., Mormon, J.P., and De Villartay, J.P. 2002. Metallo- β -lactamase fold within nucleic acids processing enzymes: The β -CASP family. *Nucleic Acids. Res.* **30**: 3592–3601.
- Cho, D.C., Scharl, E.C., and Steitz, J.A. 1995. Decreasing the distance between the two conserved sequence elements of histone pre-messenger RNA interferes with 3' processing in vitro. *RNA* **1**: 905–914.
- Chodchoy, N., Pandey, N.B., and Marzluff, W.F. 1991. An intact histone 3'-processing site is required for transcription termination in a mouse histone H2a gene. *Mol. Cell. Biol.* **11**: 497–509.
- Dominski, Z. and Marzluff, W.F. 1999. Formation of the 3' end of histone mRNA. *Gene* **239**: 1–14.
- Dominski, Z., Yang, X.C., Kaygun, H., Dadlez, M., and Marzluff, W.F. 2003. A 3' exonuclease that specifically interacts with the 3' end of histone mRNA. *Mol. Cell* **12**: 295–305.
- Dominski, Z., Yang, X.C., and Marzluff, W.F. 2005a. The polyadenylation factor CPSF-73 is involved in histone-pre-mRNA processing. *Cell* **123**: 37–48.
- Dominski, Z., Yang, X.C., Purdy, M., Wagner, E.J., and Marzluff, W.F. 2005b. 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.* **25**: 1489–1500.
- Edwards-Gilbert, G., Veraldi, K.L., and Milcarek, C. 1997. Alternative poly(A) site selection in complex transcription units: Means to an end? *Nucleic Acids Res.* **25**: 2547–2561.
- Gimmi, E.R., Reff, M.E., and Deckman, I.C. 1989. Alterations in the pre-mRNA topology of the bovine growth hormone polyadenylation region decrease poly(A) site efficiency. *Nucleic Acids Res.* **17**: 6983–6998.
- Gorgoni, B., Andrews, S., Schaller, A., Schumperli, D., Gray, N.K., and Muller, B. 2005. The stem-loop binding protein stimulates histone translation at an early step in the initiation pathway. *RNA* **11**: 1030–1042.
- Graveley, B.R., Fleming, E.S., and Gilmartin, G.M. 1996. RNA structure is a critical determinant of poly(A) site recognition by cleavage and polyadenylation specificity factor. *Mol. Cell. Biol.* **16**: 4942–4951.
- Harris, M.E., Bohni, R., Schneiderman, M.H., Ramamurthy, L., Schumperli, D., and Marzluff, W.F. 1991. Regulation of histone mRNA in the unperturbed cell cycle: Evidence suggesting control at two posttranscriptional steps. *Mol. Cell. Biol.* **11**: 2416–2424.
- Hernandez, N. and Weiner, A. 1986. Formation of the 3' end of U1 snRNA requires compatible snRNA promoter elements. *Cell* **47**: 249–258.
- Hu, J., Lutz, C.S., Wilusz, J., and Tian, B. 2005. Bioinformatic identification of candidate *cis*-regulatory elements involved in human mRNA polyadenylation. *RNA* **11**: 1485–1493.
- Iseli, C., Stevenson, B.J., deSouza, S.J., Samaia, H.B., Camargo, A.A., Buetow, K.H., Strausberg, R.L., Simpson, A.J.G., Bucher, P., and Jongeneel, C.V. 2002. Long-range heterogeneity at the 3' ends of human mRNAs. *Genome Res.* **12**: 1068–1074.

- Kaygun, H. and Marzluff, W.F. 2005a. Regulated degradation of replication-dependent histone mRNAs requires both ATR and Upf1. *Nat. Struct. Mol. Biol.* **12**: 794–800.
- . 2005b. Translation termination is involved in histone mRNA degradation when DNA replication is inhibited. *Mol. Cell. Biol.* **25**: 6879–6888.
- Kim, M., Krogan, N.J., Vasiljeva, L., Rando, O.J., Nedeá, E., Greenblatt, J.F., and Buratowski, S. 2004. The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature* **432**: 517–522.
- Kolev, N.G. and Steitz, J.A. 2005. Symplekin and multiple other polyadenylation factors participate in 3'-end maturation of histone mRNAs. *Genes & Dev.* (this issue).
- Kuersten, S. and Goodwin, E.B. 2003. The power of the 3' UTR: Translational control and development. *Nat. Rev. Genet.* **4**: 626–637.
- Marzluff, W.F. 2005. Metazoan replication-dependent histone mRNAs: A distinct set of RNA polymerase II transcripts. *Curr. Opin. Cell Biol.* **17**: 274–280.
- Minvielle-Sebastia, L. and Keller, W. 1999. mRNA polyadenylation and its coupling to other RNA processing reactions and to transcription. *Curr. Opin. Cell Biol.* **11**: 352–357.
- Mowry, K.L. and Steitz, J.A. 1988. snRNP mediators of 3' end processing: Functional fossils? *Trends Biochem. Sci.* **13**: 447–451.
- Pandey, N.B., Chodchoy, N., Liu, T.J., and Marzluff, W.F. 1990. Introns in histone genes alter the distribution of 3' ends. *Nucleic Acids Res.* **18**: 3161–3170.
- Pauws, E., van Kampen, A.H., van de Graaf, S.A., Vijlder, J.J., and Ris-Stalpers, C. 2001. Heterogeneity in polyadenylation cleavage sites in mammalian mRNA sequences: Implications for SAGE analysis. *Nucleic Acids Res.* **29**: 1690–1694.
- Pillai, R.S., Grimmmer, M., Meister, G., Will, C.L., Luhrmann, R., Fischer, U., and Schumperli, D. 2003. 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.* **17**: 2321–2333.
- Proudfoot, N.J., Furger, A., and Dye, M.J. 2002. Integrating mRNA processing with transcription. *Cell* **108**: 501–512.
- Ryan, K., Calvo, O., and Manley, J.L. 2004. Evidence that polyadenylation factor CPSF-73 is the mRNA 3' processing endonuclease. *RNA* **10**: 565–573.
- Scharl, E.C. and Steitz, J.A. 1994. 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.* **13**: 2432–2440.
- . 1996. Length suppression in histone messenger RNA 3'-end maturation: Processing defects of insertion mutant premessenger RNAs can be compensated by insertions into the U7 small nuclear RNA. *Proc. Natl. Acad. Sci.* **93**: 14659–14664.
- Takagaki, Y. and Manley, J.L. 2000. Complex protein interactions within the human polyadenylation machinery identify a novel component. *Mol. Cell. Biol.* **20**: 1515–1525.
- Tian, B., Hu, J., Zhang, H.B., and Lutz, C.S. 2005. A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Res.* **33**: 201–212.
- Venkataraman, K., Brown, K.M., and Gilmartin, G.M. 2005. Analysis of a noncanonical poly(A) site reveals a tripartite mechanism for vertebrate poly(A) site recognition. *Genes & Dev.* **19**: 1315–1327.
- Wang, Z.F., Ingledue, T.C., Dominski, Z., Sanchez, R., and Marzluff, W.F. 1999. Two *Xenopus* proteins that bind the 3' end of histone mRNA: Implications for translational control of histone synthesis during oogenesis. *Mol. Cell. Biol.* **19**: 835–845.
- West, S., Gromak, N., and Proudfoot, N.J. 2004. Human 5'→3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. *Nature* **432**: 522–525.
- Xie, X.H., Lu, J., Kulbokas, E.J., Golub, T.R., Mootha, V., Lindblad-Toh, K., Lander, E.S., and Kellis, M. 2005. Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. *Nature* **434**: 338–345.
- Xing, H., Mayhew, C.N., Cullen, K.E., Park-Sarge, O.K., and Sarge, K.D. 2004. HSF1 modulation of Hsp70 mRNA polyadenylation via interaction with symplekin. *J. Biol. Chem.* **279**: 10551–10555.
- Zeng, L.-X., Dominski, Z., Yang, X., Elms, P., Raska, C.S., Borchers, C.H., and Marzluff, W.F. 2003. 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.* **23**: 1590–1601.
- Zhao, J., Hyman, L., and Moore, C. 1999a. Formation of mRNA 3' ends in eukaryotes: Mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol. Mol. Biol. Rev.* **63**: 405–445.
- Zhao, J., Kessler, M., Helming, S., O'Connor, J.P., and Moore, C. 1999b. Pta1, a component of yeast CF II, is required for both cleavage and poly(A) addition of mRNA precursor. *Mol. Cell. Biol.* **19**: 7733–7740.