REVIEW

Mechanism and regulation of mRNA polyadenylation

Diana F. Colgan and James L. Manley[1]

Department of Biological Sciences, Columbia University, New York, New York 10027 USA

A poly(A) tail is found at the 3′ end of nearly every fully processed eukaryotic mRNA and has been suggested to influence virtually all aspects of mRNA metabolism. Its proposed functions include conferring mRNA stability, promoting an mRNA's translational efficiency, and having a role in transport of processed mRNA from the nucleus to the cytoplasm (for recent reviews, see Lewis et al. 1995; Sachs et al. 1997; Wickens et al. 1997). The reaction that catalyzes the addition of the poly(A) tail, an endonucleolytic cleavage followed by poly(A) synthesis, has also been the focus of intense investigation but, until recently, may have been viewed as a process that follows a predictable, isolated, and invariant path. Yet, as more is learned about 3′-end formation, it becomes clear that the function of the polyadenylation machinery extends beyond simply adding poly(A) tails to mRNAs.

The first report of a component of the mammalian cleavage and polyadenylation machinery was nearly 40 years ago in a paper describing an activity found in thymus nuclei extracts that could synthesize poly(A) from ATP (Edmonds and Abrams 1960). Ten years passed before poly(A) tails were identified as a post-transcriptionally added modification of mRNA 3′ termini and a possible function was assigned to poly(A) polymerase (Darnell et al. 1971; Edmonds et al. 1971; Lee et al. 1971). But nearly another decade elapsed before it was found that transcription proceeds past the polyadenylation site, revealing that a mechanism other than transcriptional termination generates mRNA 3′ ends (Ford and Hsu 1978; Nevins and Darnell 1978; Manley et al. 1982). The pace of discovery quickened with the development of cell extracts that reproduce the reaction, and this allowed the subsequent and still ongoing biochemical characterization of mRNA 3′-end formation (Manley 1983; Moore and Sharp 1984, 1985). The results from this work have shown that nuclear cleavage and poly(A) addition occurs in a coupled reaction and is carried out by a suprisingly large complex of multisubunit proteins (for recent reviews, see Keller 1995; Manley 1995). Several years were devoted to detailing the mechanism of 3′-end formation, assigning relatively simple functions to each separable factor of the complex polyadenylation machinery. With

the cloning of cDNAs encoding many of these factors, we have enjoyed an accelerated pace in understanding their precise functions, as well as the unexpected bonuses of finding that these basal factors link nuclear polyadenylation to a variety of cellular processes and that they can be important targets for regulating gene expression. Here we describe how information gained from studies done over the last few years has enhanced our understanding of the structure and function of the proteins catalyzing polyadenylation. We concentrate on mammalian systems but also highlight progress that points to both similarities and differences in yeast polyadenylation. From reviewing the latest events, we not only see how far we have come in 40 years but also become more aware of the rich path of discovery that lies ahead.

The polyadenylation element

In general, sequence elements that specify the site of cleavage and poly(A) addition in metazoan mRNAs flank the site of endonucleolytic attack (for review, see Proudfoot 1991). Nearly every known mRNA contains a polyadenylation signal sequence, the hexanucleotide AAUAAA, 10–30 bases upstream of the cleavage/polyadenylation site. AAUAAA is found in 90% of all sequenced polyadenylation elements and is one of the most highly conserved sequence elements known. Most of the 10% of mRNAs that do not have an exact match to the consensus differ by only a single substitution (an A → U conversion at the second position is the most common variant). A second sequence element is a GU-rich motif located 20–40 bases downstream of the cleavage site. In contrast to AAUAAA, this GU-rich element is found in ~70% of mammalian pre-mRNAs and is considerably more variable in sequence composition. It is frequently characterized by single or multiple stretches of up to five consecutive U residues, often interrupted by single G's (MacDonald et al. 1994; Takagaki and Manley 1997). These two sequence elements and their distance from one another not only specify the site of cleavage and poly(A) addition (MacDonald et al. 1994; Chen et al. 1995) but also largely determine the strength of a poly(A) signal. The precise site of endonucleolytic attack can be shifted by changing the distance between the two core

[1]Corresponding author.
E-MAIL jlm2@columbia.edu; FAX (212) 865-8246.

elements. A CA dinucleotide immediately 5′ to the site of cleavage is preferred but not required (e.g., Chen et al. 1995). The influence of the nucleotide sequences surrounding the cleavage site remains unclear (e.g., MacDonald et al. 1994; Gilmartin et al. 1995)

The general polyadenylation machinery

A large complex of factors is required to reconstitute the complete polyadenylation reaction in vitro, and most of these factors consist of several subunits. Although the functions of the proteins essential for the reaction are fairly well characterized (with the exception of the actual endonuclease, which remains unknown), indications of functions peripheral to the catalytic reactions have begun to emerge only recently. However, new data are shedding light on the reasons behind the unexpected multiplicity of participating factors and how the various proteins may carry out duties other than catalysis of cleavage and poly(A) addition. This is proving to be one of the most exciting areas of study in the field.

Cleavage/polyadenylation specificity factor

Cleavage and polyadenylation of newly synthesized transcripts is probably always performed in a tightly coupled reaction in vivo. However, the two phases of the reaction can be uncoupled and studied separately in vitro, allowing the functions of the participating factors to be examined individually. In vitro analyses of the two subreactions showed that AAUAAA is required during both phases, and a similar requirement for one factor, cleavage/polyadenylation specificity factor (CPSF), suggested that CPSF recognizes AAUAAA. Experimental evidence confirmed that CPSF does indeed specifically bind RNA containing the signal sequence, but binding is very weak with purified CPSF alone and is greatly enhanced by a cooperative interaction with a second component of the polyadenylation machinery, cleavage stimulation factor (CstF) (see below). The two factors bound to the precursor RNA form a stable ternary complex, which could function to recruit the other components of the polyadenylation machinery to the cleavage site (for recent reviews, see Manley 1995; Wahle and Keller 1996). Furthermore, recent evidence suggests that CPSF does more than participate in specifying the site of nuclear cleavage and poly(A) addition via its interactions with CstF and poly(A) polymerase (PAP) (Murthy and Manley 1995). Additional functions for CPSF may be found in the coordination of nuclear polyadenylation with transcription (Dantonel et al. 1997; McCracken et al. 1997), participation in cytoplasmic polyadenylation (Bilger et al. 1994), limitation of the length of the poly(A) tail (Wahle 1995), and the interaction with proteins associated with the splicing apparatus (Lutz et al. 1996).

Purified CPSF consists of four polypeptides with molecular masses of 160, 100, 73, and 30 kD (CPSF-160, CPSF-100, CPSF-73, and CPSF-30, respectively) (Bienroth et al. 1991; Murthy and Manley 1992), although the three largest subunits appear sufficient for activity of mammalian CPSF in in vitro assays (Murthy and Manley 1992; Gilmartin et al. 1995). Whereas CPSF-30 can be dispensable in vitro, its apparent homolog in yeast is essential for viability, and extracts from mutant strains display defects in 3′-end formation (Barabino et al. 1997). The amino acid sequence of CPSF-30 indicates that it contains five putative zinc-finger motifs and a carboxy-terminal zinc knuckle motif, which are characteristic of nucleic acid-binding proteins. Consistent with this observation, the protein has been shown to bind RNA with a preference for poly(U). However, the significance of this sequence preference and the precise function of CPSF-30 is not known. Cloning and sequencing of cDNAs encoding CPSF-100 and CPSF-73 have shown that the two proteins are related, with CPSF-73 displaying 23% identity to CPSF-100 (Jenny et al. 1994, 1996). The function(s) of these subunits is currently unknown.

CPSF-160 is the best characterized of the CPSF subunits (Murthy and Manley 1995). Analysis of its 1442 amino acid sequence (Jenny and Keller 1995; Murthy and Manley, 1995) shows a potential bipartite nuclear localization signal (NLS) and appropriately spaced sequences with similarity to the RNP1 and RNP2 motifs found in many RNA-binding proteins, which could conceivably constitute part of a divergent RNP-type RNA-binding domain (RBD). Consistent with earlier UV cross-linking studies, recombinant CPSF-160 alone can preferentially bind to AAUAAA-containing RNAs. But the interaction is weak and tolerant of mutations in the AAUAAA, leading to the conclusion that the participation of the other CPSF subunits may facilitate recognition of AAUAAA. Cooperative binding of CPSF and CstF to the poly(A) site has been partially reconstituted using human CstF and CPSF-160, in the absence of the other CPSF subunits. In addition, CPSF-160 has been shown to make contacts with both CstF-77 and PAP, suggesting that this subunit plays a key role in coordinating both cleavage and poly(A) addition subreactions. It is interesting to note that titration of excess CPSF-160 into nonspecific poly(A) synthesis reactions can inhibit the activity of PAP, implying that CPSF may also affect elongation by PAP in addition to positioning PAP at the appropriate 3′ end. Despite the fact that it makes contacts with PAP, CPSF-160 plus PAP are not sufficient to reconstitute AAUAAA-dependent poly(A) addition.

Unexpected new developments have provided evidence that in addition to playing a central role in cleavage and poly(A) addition, CPSF, and specifically CPSF-160, is involved in linking polyadenylation to transcription. It has been shown that CPSF can exist in a stable complex with the transcription factor TFIID, likely reflecting interactions between several subunits of TFIID and CPSF-160 (Dantonel et al. 1997). In a reconstituted transcription reaction this association recruits CPSF to the RNA polymerase II (Pol II) preinitiation complex, where, concomitant with transcription initiation and phosphorylation of the carboxy-terminal domain (CTD) of the Pol II largest subunit, CPSF dissociates from TFIID and becomes associated with the elongating Pol II (Dantonel et al. 1997). CPSF (as well as CstF; McCracken et

al. 1997) appears to travel with Pol II until reaching the polyadenylation element, where it is suggested to dissociate and define the poly(A) site. These findings have provided mechanistic insights and extended the original observations of McCracken et al. (1997). Those studies presented the unexpected findings that the CTD can be essential for efficient splicing and polyadenylation of newly transcribed pre-mRNA in transiently transfected cells, likely reflecting the observed presence of both CPSF and CstF in the Pol II holoenzyme complex via an association with the CTD. In addition to providing insights into how a poly(A) site might be defined in vivo, these results raise two additional intriguing possibilities: First, perhaps dissociation of the poly(A) factors influences the ability of Pol II to elongate, thereby providing at least a partial explanation for the requirement of a functional poly(A) site for transcription termination (for review, see Proudfoot 1989). Second, it might now be legitimate to consider the subunits of CPSF as TATA-binding protein-associated factors (TAFs). Given the apparent requirement of other TAFs for transcription from at least some promoters (for review, see Tansey and Herr 1997), perhaps changes in the concentration of polyadenylation factors (e.g., Takagaki et al. 1996) might influence transcription initiation!

In *Xenopus,* CPSF has also been shown to play a role in cytoplasmic polyadenylation of maternal mRNAs, which controls their translation during oocyte maturation and early embryogenesis (Bilger et al. 1994). To be recognized as substrates for cytoplasmic PAP, these mRNAs require both AAUAAA and a cytoplasmic polyadenylation element, which binds to developmentally specific proteins (for review, see Richter 1996). Other processes where CPSF participates, for example, in limiting the length of the poly(A) tail or interacting with factors involved in pre-mRNA splicing, are discussed in detail below.

CstF

CstF was first described as an activity required for efficient cleavage but not poly(A) addition. In fact, one of its subunits, CstF-64, was the first polyadenylation factor detected (with the exception of PAP), based on its strong and specific UV cross-linking to RNAs containing a functional poly(A) signal (Wilusz and Shenk 1988). CstF is a heterotrimeric protein with subunits of 77, 64, and 50 kD (CstF-77, CstF-64, and CstF-50). In the presence of CPSF, CstF binds the precursor RNA very tightly, and as expected, binding of CstF to the RNA is mediated by CstF-64. Cloning of cDNAs encoding CstF-64 indicated the presence of an RBD at the protein's amino terminus (Takagaki et al. 1992) and mapping of the site of UV cross-linking to the RNA in nuclear extracts confirmed that CstF-64 binds specifically to the GU-rich downstream motif (MacDonald et al. 1994). Recent work, including selection of high affinity binding sites (SELEX), has shown that the RBD region of CstF-64 alone is capable of sequence-specific binding to GU/U-rich sequences like those found in natural pre-mRNAs (Takagaki and Manley 1997).

In keeping with its early discovery, CstF-64 was the first (again with the exception of PAP) polyadenylation factor cloned (Takagaki et al. 1992). cDNAs encoding homologs of CstF-64 have since been isolated from mouse, human, and chicken (Takagaki et al. 1996). Comparison of their amino acid sequences shows several common features as well as potentially interesting differences (Fig. 1). As might be expected, the amino-terminal RBD is nearly invariant in all three species. The RBD is followed by a highly conserved ~100-residue region lacking any identifiable features (the hinge domain), which may interact with other polyadenylation factors. This precedes an ~300-residue carboxy-terminal region that consists of a possibly unstructured domain that is 20% proline and 20% glycine. However, imbedded within this domain in the mammalian homologs is a region that contains 12 consecutive repeats of the pentapeptide consensus MEARA/G, which potentially forms a long α-helix that is stabilized by salt bridges between neighboring acidic and basic residues. Although the human and mouse proteins are highly homologous thoughout their entire lengths, chicken CstF-64 displays several differences that may indicate species-specific
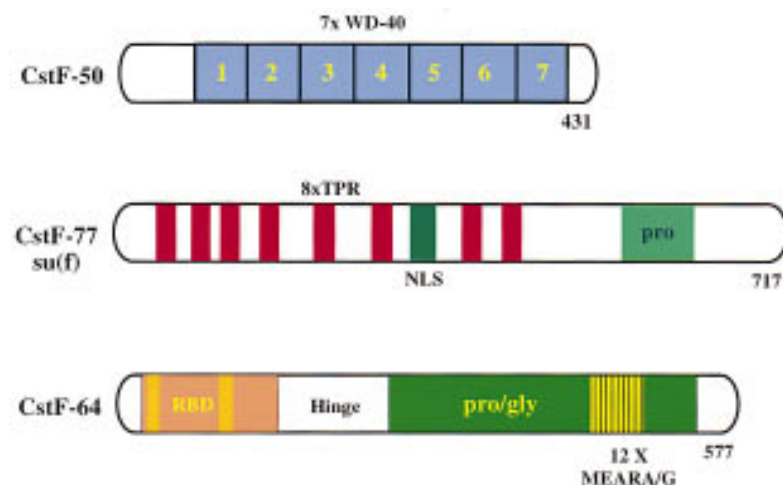

Figure 1. CstF subunits. Structural features of the three CstF subunits are indicated. See text for description and discussion of individual motifs. CstF-77, but neither of the other subunits, contains sequences that appear to constitute a bipartite NLS. Thus, CstF may assemble in the cytoplasm prior to transport to the nucleus.

variations in the function of the protein. The proline/glycine-rich region is increased to 45% and, more strikingly, the MEARA/G repeats contain an alanine-to-proline substitution in 7/11 of the central alanines, which could severely disrupt its predicted α-helical structure. Nothing further is currently known about the function of these potentially interesting motifs.

CstF-77 has been shown to bridge CstF-64 and CstF-50 (Takagaki and Manley 1994), as well as contact CPSF-160, possibly facilitating the cooperative RNA binding displayed by CPSF and CstF (Murthy and Manley 1995). Although not observed initially, CstF-77 contains eight repeats similar to tetratricopeptide repeat (TPR) motifs (C. Williams and K. O'Hare, pers. comm.; see Fig. 1), which have been proposed to mediate protein–protein interactions. Otherwise apparently unrelated proteins contain TPR motifs, and some have been genetically linked to proteins possessing transducin, or WD-40, repeats (Goebell and Yanagida 1991). This may be significant because CstF-50 contains seven WD-40 repeats (Takagaki and Manley 1992). The sequence of CstF-77 revealed that it is the mammalian homolog of the *Drosophila* Suppressor-of-forked [Su(f)] protein (Mitchelson et al. 1993; Takagaki and Manley 1994). Nonlethal mutations in *su(f)* can enhance or suppress the effects of transposable element insertions, probably because of changes in polyadenylation efficiency (for review, see O'Hare 1995). In addition to the TPR-like motifs, both CstF-77 and the Su(f) protein contain a carboxy-terminal proline-rich region in which 14/15 proline residues have been conserved, suggesting the importance of this domain for CstF-77 function. Nonlethal mutations in *su(f)* can affect the polyadenylation site utilization in a variety of unlinked genes. Some of these mutations are frequently found near the carboxyl terminus (Simonelig et al. 1996) and might be associated with a disruption in the function of the conserved proline-rich region. Little is known about the role of CstF-50, except that it is essential for CstF function in cleavage in vitro (Takagaki and Manley 1994). This subunit has also been shown to contact directly the Pol II CTD, perhaps anchoring CstF to the elongating polymerase (McCracken et al. 1997).

Cleavage factors

Two additional multisubunit factors, cleavage factors I and II (CF I and CF II) are required for the cleavage but not the poly(A) addition step of the reaction (Takagaki et al. 1989). These factors remain the most poorly characterized of the polyadenylaton machinery. Although endonuclease activity has yet to be demonstrated for any of the subunits of CF I or CF II, the process of elimination makes them likely candidates to be the proteins that perform the actual cleavage step. CF I has been highly purified and appears to consist of three subunits with molecular masses of 68, 59, and 25 kD (Ruegsegger et al. 1996). Unexpectedly, all three subunits could be UV cross-linked to substrate RNAs (Ruegsegger et al. 1996). Binding appeared to be specific, as it was not competed

by large amounts of nonspecific competitor RNA, and a precleaved substrate RNA did not bind CF I. Although CF I was able to stabilize CPSF on substrate RNA, CPSF did not enhance the RNA binding of any of the CF I subunits, nor did mutation of AAUAAA affect binding efficiency. Taken together, these results suggest that CF I recognizes sequence elements in the poly(A) site, but what these might be awaits further investigation. Characterization of CF II has not yet been reported.

PAP

PAP is probably the best understood of all of the components of the 3′ end-forming machinery. Its early discovery, homology to other proteins, and the ease with which polyadenylation assays can be performed have all helped characterization of the enzyme. Furthermore, because PAP is composed of a single polypeptide, matters are kept simple. One of the most suprising findings concerning PAP was the detection of several alternatively spliced forms (Raabe et al. 1991; Wahle et al. 1991; Ballantyne and Wickens 1995). Subsequent analysis of mouse genomic clones revealed that at least six isoforms can arise from a single alternatively spliced transcript (Zhao and Manley 1996). The full-length forms, PAP I, PAP II, and PAP IV, appear to be the enzymatically active forms (Raabe et al. 1991; Zhao and Manley 1996), with PAP II being isolated most commonly from cDNA libraries and probably the predominant full-length species in most cells.

PAP II is 740 amino acids in length. Initial sequence analysis predicted that its amino-terminal two-thirds contains the catalytic activity (Raabe et al. 1991). Recent studies have suggested two highly structured regions separated by a loop region (Martin and Keller 1996). These structures appear to place PAP in a larger family of nucleotidyl transferases (that includes DNA Pol β) and support a model in which catalysis is performed on the face of a five-stranded β-sheet supported by two α-helices. This model is supported by analysis of deletion and point mutants that have deleterious effects on PAP's activity (Martin and Keller 1996). The catalytic region is followed by a predicted ~20-residue primer binding domain that also encompasses one of two NLS sequences (Martin and Keller 1996; see also Zhelkovsky et al. 1995), both of which are required for complete nuclear localization (Raabe et al. 1994). The second NLS lies ~140 amino acids downstream of the first, and the two encompass an S/T-rich domain proposed to perform regulatory functions (Raabe et al. 1991), an idea for which there is now considerable support (see below). Following the second NLS, the long forms diverge because of alternative splicing. The function of the divergent carboxyl termini of the long forms is unknown, although differential interaction with other factors has been proposed (see below).

Three short forms of PAP would encode proteins that terminate prior to the first NLS and appear to arise from competition between polyadenylation and splicing (Zhao and Manley 1996). These short forms of PAP lack

the minimum required catalytic domain, and recombinant derivatives have no activity in several in vitro assays (Martin and Keller 1996; Zhao and Manley 1996). In addition, although the mRNAs have been detected, Western blot analysis has failed to confirm that these mRNAs are translated (Gebauer and Richter 1995; Zhao and Manley 1996). Because PAP is generally required to form an active cleavage complex, an attractive hypothesis is that the short species result from a negative feedback mechanism regulating PAP expression. High levels of full-length PAP could result in recognition of PAP pre-mRNA's weaker, upstream polyadenylation sites, causing expression of shorter, inactive, or unstable species. It is notable that there is evidence for a similar mechanism regulating the level of CstF-77/Su(f) in *Drosophila* (Simonelig et al. 1996), and such regulation may be typical of certain mRNA processing proteins (e.g., Zachar et al. 1987).

PAB II

There is an upper limit of 200–300 residues in the length of poly(A) tails found on newly synthesized RNA in vivo. This length restriction is mediated, in part, by poly(A)-binding protein II (PAB II), a nuclear protein with high affinity for poly(A) (Wahle et al. 1993). Once a short poly(A) tail has been synthesized, PAB II binds to it and forms a quaternary complex with CPSF, PAP, and the substrate RNA. This complex transiently stabilizes the binding of PAP to the RNA 3′ end, supporting processive synthesis of a long poly(A) tail in a single rapid step (Bienroth et al. 1993). Length control results from interruption of the interactions between the RNA, CPSF, and PAP (Wahle 1995); hence, the hypothesis that PAB II works by cooperating with CPSF to hold PAP in place on the primer RNA, and that these interactions are interrupted after the tail has reached a distinct length. Perhaps the ability of free CPSF-160 to inhibit PAP (Murthy and Manley 1995) reflects an interaction relevent to termination of poly(A) synthesis. In yeast, the highly conserved and abundant PAB I, shown previously to function in mRNA translation and stability, recently has been found to also participate in polyadenylation by controlling poly(A) tail length (Amrani et al. 1997; Minvielle-Sebastia et al. 1997). Perhaps this reflects the absence of a yeast PAB II homolog.

Similarities and differences in yeast polyadenylation

Yeast polyadenylation also involves a coupled two-step reaction of cleavage and poly(A) addition. However, the sequence elements governing yeast 3′-end processing appear decidedly different (Guo and Sherman 1996). Yeast signals contain no detectable AAUAAA or GU-rich sequence elements but, rather, appear to be defined by an A-rich positioning element and an AU-rich efficiency element, both of which lie upstream of the site of cleavage and are quite variable. It has been suggested that the efficiency element may be analogous to the mammalian GU-rich downstream element and is positioned up-

stream simply because of the relative proximity of genes in the yeast genome, whereas in mammals large distances between genes can readily accommodate downstream elements (Manley and Takagaki 1996).

The divergence of sequence elements governing polyadenylation in yeast and mammals had led to the anticipation that the participating factors would display comparable divergence. Yet, recent cloning of cDNAs encoding yeast polyadenylation proteins has revealed that although some yeast polyadenylation factors do not have known mammalian counterparts (such as FIP-1; Preker et al. 1995), several others are remarkably conserved (for recent reviews, see Manley and Takagaki 1996; Keller and Minvielle-Sebastia 1997). PAP, not unexpectedly, is highly conserved with an amino acid sequence that is 47% identical to mammalian PAP throughout the amino-terminal 450 amino acids, which contains the catalytic region (Lingner et al. 1991; Raabe et al. 1991; Wahle et al. 1991). However, yeast PAP is intriguingly different in that it is truncated after this region, lacking the S/T-rich regulatory region found in the mammalian enzyme. More unexpected (given the lack of AAUAAA and GU-rich motifs) is the observation that several polypeptides share significant similarity with subunits of CstF and CPSF. RNA14 and RNA15 are the apparent yeast homologs of CstF subunits (Takagaki and Manley 1994; Minvielle-Sebastia et al. 1994; Kessler et al. 1996). The amino acid sequence of RNA14 is 24% identical to CstF-77 throughout its entire length, maintaining the TPR-like motifs but lacking the proline-rich carboxy-terminal region, and RNA15 is 43% identical to CstF-64 in its RBD but divergent through the remainder of the protein, lacking all the features of CstF-64 referred to above (Takagaki and Manley 1994). The sequence divergence displayed by these factors could possibly reflect functional differences, and in fact, the complex containing RNA14 and RNA15 seems to function more like CPSF than CstF in that it is required for both cleavage and poly(A) addition in vitro (Minvielle-Sebastia et al. 1994; Kessler et al. 1996). However, this difference may be more apparent than real, perhaps reflecting only the relative positions of the binding sites in the pre-mRNA. The downstream position of the GU-rich element in mammalian mRNA precursors would essentially preclude its function in the poly(A) addition step of the reaction, at least as measured in vitro with precleaved RNA substrates.

Yeast proteins that are the apparent homologs of all four CPSF subunits have been identified and can be purified, along with several additional factors, as part of a large complex (Jenny et al. 1996; Zhao et al. 1997). Brr5/Ysh1 shares high similarity (53% identity) to the related protein CPSF-73 (as well as 23% identity to CPSF-100; Chanfreau et al. 1996; Jenny et al. 1996), CFT1 is the apparent homolog (24% identity) of CPSF-160 (Stumpf and Domdey 1996; Zhao et al. 1997), CFT2 is 24% identical to CPSF-100 (Zhao et al. 1997), and YTH1 is 40% identical to CPSF-30 (Barabino et al. 1997). It is as yet unknown how these proteins participate in the polyadenylation reaction.

These sequence similarities confirm a common evolutionary ancestry for the cleavage/polyadenylation reaction, and studies of 3′-end formation in yeast will therefore likely provide valuable insights into the function of homologous components in mammals. Yet the differences may be even more interesting. For example, recently it has been shown that the yeast protein Hrp1, which is related to mammalian heterogeneous nuclear RNP A1 (hnRNP A1), is required for 3′ cleavage in yeast, and, as a nuclear–cytoplasmic shuttling protein, may link polyadenylation and mRNA transport (Kessler et al. 1997). In mammals, however, there is no evidence that an hnRNP-like protein is involved in 3′-end formation, and it may be that metazoa have evolved a different mechanism to link mRNA processing and transport. It may also be informative to identify proteins, and protein domains, that are present in the mammalian machinery but absent in yeast, because they may give clues to systems in higher organisms that might be regulated at the level of polyadenylation.

Regulation of polyadenylation

One of the best-studied examples of regulation at the level of mRNA polyadenylation is found during early embryogenesis. Prior to the onset of zygotic transcription, gene expression is controlled by cytoplasmic poly(A) addition to fully processed maternal mRNAs. Although PAP and CPSF appear to participate in this reaction, it is distinct from the nuclear reaction and relies on several developmentally specific factors and sequence elements. The reader is referred to recent reviews for further discussion of this and related topics (Wormington 1994; Hake and Richter 1997; Wickens et al. 1997). Known examples of regulation at the level of nuclear polyadenylation of newly transcribed RNAs are relatively rare. Nevertheless, there is increasing evidence that 3′-end formation can be regulated (see Edwalds-Gilbert et al. 1997 for a compilation of possible targets), and we discuss a number of examples below.

Auxiliary elements

In addition to the AAUAAA- and the GU-rich motifs, some pre-mRNAs contain other genetic elements that influence the efficiency with which polyadenylation signals are utilized. The majority of examples of polyadenylation regulated by auxiliary signals have been detected in the genomes of viruses and are commonly associated with switches in the viral life cycle. Examples of viral auxiliary elements are found in the SV40, adenovirus, hepatitis B, and several retroviral pre-mRNAs. More recently, the human C2 complement gene and the mouse calcitonin/calcitonin gene-related peptide (CT/CGRP) have been identified as cellular genes that contain noncanonical polyadenylation signals requiring auxiliary elements to promote efficient utilization (for review, see Proudfoot 1996).

Because retroviral replication relies on the presence of terminally redundant sequences in the viral long terminal repeats (LTRs), duplicated polyadenylation sites are found at both the 3′ and 5′ ends of the viral transcript. Therefore, unique features outside of the LTRs must either repress recognition of the 5′-end poly(A) signal or enhance recognition of the 3′-end signal. HIV-1, for example, appears to have both features. Enhanced recognition of the 3′-end poly(A) site is promoted by a structural conformation assumed by RNA sequences lying upstream. These sequences appear to stabilize CPSF binding to the 3′ LTR's poly(A) site, thereby promoting its utilization (Gilmartin et al. 1995; Graveley and Gilmartin 1996; Graveley et al. 1996). However, mutation of these stabilizing sequences is not sufficient to cause use of the 5′–LTR poly(A) site, suggesting that recognition of this poly(A) site is somehow repressed. Analysis of proviral transcripts has shown that this repression is likely mediated by the presence of a strong 5′ splice site located just downstream of the 5′ LTR, as mutation of the splice site allowed recognition of the 5′ poly(A) site (Ashe et al. 1995; see Fig. 2). This case of a splice site influencing recognition of a poly(A) signal is one of a growing number of examples in which splicing factors and sequences affect cleavage and/or poly(A) addition.

Polyadenylation machinery can employ splicing factors and sequence elements

Early immunodepletion studies showed that 3′-end formation in crude extracts could be inhibited with antibodies recognizing small nuclear RNPs (snRNPs) (e.g., α-SM and α-trimethyl guanosine cap; Moore and Sharp 1985; Hashimoto and Steitz 1986). Although it is now clear that the basal polyadenylation machinery does not contain a snRNP, mounting evidence suggests that splicing and polyadenylation can be linked. For example, a functional polyadenylation signal can enhance splicing of the 5′-terminal intron in vitro, and vice versa, implying that excision of an mRNA's last intron and polyadenylation can be functionally linked (Niwa et al. 1990; Wasserman and Steitz 1993); in fact, recent studies suggest that splicing factors can play auxilary roles in poly(A) site selection.

The best-studied example of a spliceosomal protein influencing polyadenylation is the U1 snRNP A protein (U1A). The U1A pre-mRNA contains sequence elements upstream of its poly(A) site that resemble the U1A-binding site on the U1 snRNA. Expression of U1A mRNA is negatively autoregulated when two molecules of U1A protein bind the U1A pre-mRNA, resulting in an inhibition of polyadenylation (Boelens et al. 1993). When the two subreactions were examined independently, it was determined that the cleavage step was not affected by bound U1A. However, poly(A) addition was inhibited, even in nonspecific assays (see above), indicating that the repression mechanism was targeting PAP directly. Further examination suggested that this inhibition is the result of a repressive interaction between PAP and U1A (Gunderson et al. 1994). These studies propose that polyadenylation activity is inhibited through contact of PAP's regulatory carboxyl terminus with two molecules
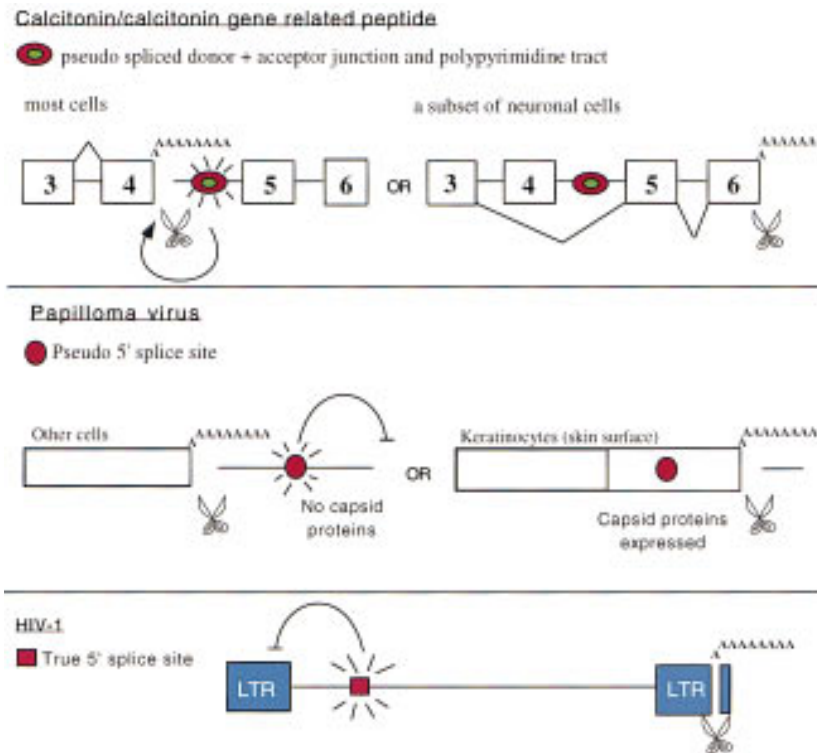
Figure 2. 5′ splice site-related sequences can influence poly(A) site selection. (*Top*) A polyadenylation enhancer affecting alternative splicing and polyadenyaltion of the CT/CGRP transcript contains sequences resembling the junction of spliced 5′ and 3′ splice sites. This element lies downstream of exon 4's alternatively selected polyadenylation site and is active in most tissues. Deletion of the element, as occurs during CT/CGRP expression in neuronal tissue, results in skipping of exon 4 and polyadenylation of exon 6. (*Middle*) In a papilloma virus transcript, a sequence resembling a 5′ splice site (and one that likely binds the U1 snRNP) can repress recognition of a downstream poly(A) site. Tissue-specific action of this pseudo-5′ splice site restricts papilloma virus production to terminally differentiated keratinocytes. (*Bottom*) In HIV-1, a major 5′ splice site represses selection of the upstream poly(A) site located in the 5′ UTR (see text for details).

of U1A (via U1 residues 103–119; Gunderson et al. 1997). In vitro experiments suggest that the carboxy-terminal 20 residues of PAP II are sufficient for inhibition (Gunderson et al. 1997). This is potentially interesting because alternatively spliced PAP I lacks sequences derived from the exon encoding PAP II's carboxyl terminus, suggesting a possible difference in the two forms of PAP.

Conversely, U1A has been shown to enhance polyadenylation of the SV40 late pre-mRNA (Lutz and Alwine 1994; Lutz et al. 1996). Maximal polyadenylation of the SV40 late transcript in vitro was found to be dependent on a previously described sequence lying upstream of its poly(A) site. U1A was shown to bind this element, and immunodepletion of U1A inhibited efficient utilization of the downstream poly(A) signal (Lutz and Alwine 1994). U1A likely functions by stabilizing CPSF binding to the RNA through a direct interaction between residues in the U1A amino terminus and CPSF-160 (Lutz et al. 1996). Stabilization of CPSF binding likely explains the observed enhancement of CPSF-dependent poly(A) addition by U1A. High concentrations of U1A, however, inhibited poly(A) addition, likely reflecting repression of PAP (see above). Thus, U1A has been suggested to have opposite effects on polyadenylation via interactions between distinct regions of the protein and two different components of the basal polyadenylation machinery.

The ability of a regulatory protein to function both positively and negatively is a theme well established in transcriptional control, and sequences resembling weak 5′ splice sites have likewise been suggested to have both positive and negative effects on the use of alternative poly(A) sites. For example, a "polyadenylation enhancer" that resembles abutting 5′ and 3′ splice sites, without an intervening sequence, has been identified in an intron of the CT/CGRP gene (see Fig. 2). This element lies within an intron, 168 nucleotides downstream of a weak alternative polyadenylation site, at the end of exon 4 (Lou et al. 1995). In a few cell types, including neuronal cells, exon 3 is spliced directly to exon 5, and a poly(A) site at the end of exon 6 is utilized. In most tissues exon 3 is spliced to exon 4, and a weak poly(A) site at the end of exon 4 is utilized, resulting in the exclusion of exons 5 and 6. This latter processing route was shown to be dependent on the presence of the intronic polyadenylation enhancer element lying between exons 4 and 5. In vitro experiments provided evidence that efficient recognition of exon 4's polyadenylation site was dependent on the presence of a functional U1 snRNP, and experiments with RNA competitors suggest that polypyrimidine tract-binding protein (PTB) is also required (Lou et al. 1996). PTB has been associated previously with splicing inhibition (e.g., Mulligan et al. 1992; Lin and Patton 1995; Singh et al. 1995), and its presence in a polyadenylation enhancer was thus unexpected. The enhancer was proposed to function by stabilization of CstF binding to the RNA, which could reflect the fact that the regulated poly(A) signal lacks a strong GU-rich region. However, it is conceivable that CPSF is the actual target, which, due to cooperative binding (see above), would result in stabilized CstF binding. By this scenario, the mechanism could be similar to that suggested for U1A activation of SV40 late polyadenylation.

Sequences resembling a 5′ splice site (referred to as a pseudo-5′ splice site) can also affect polyadenylation negatively. For example, in the case of 3′-end processing of the bovine papilloma virus transcript, repression of the late polyadenylation signal depends on the presence of an upstream pseudo-5′ splice site (Furth et al. 1994) . The inhibitory effect is tissue specific, functioning in cells other than terminally differentiated keratinocytes, thereby restricting virus production to specific skin cells (Spalhlolz and Howley 1989). Point mutations known to destroy the function of true splice sites also destroyed the repressive effect of the pseudo-5′ splice site (Furth et al. 1994). Furthermore, antisense oligonucleotides targeting U1 snRNA for RNase H degradation blocked the effect of the pseudo-5′ splice site. More compellingly, transfection of vectors expressing U1 RNAs with compensatory mutations partially restored the inhibitory activity of a mutated pseudo-5′ splice site sequence, confirming the association of the U1 snRNP with this inhibitory element. The investigators proposed that the presence of the 5′ splice site upstream of the polyadenylation site interferes with the ability of the splicing and polyadenylation machinery to define the 5′-terminal exon.

Regulation via basal factors

There are two well-documented examples in which the activity or abundance of a component of the basal polyadenylation machinery is implicated in regulation. One is found during B lymphocyte activation, when alternative pre-mRNA processing switches expression of IgM heavy chain from the membrane-bound form (µm) to the secreted form (µs). For years it had not been clear whether this process is controlled by alternative splicing or polyadenylation (for review, see Peterson 1994a), because utilization of the downstream (µm) poly(A) site occurs in conjunction with splicing out of the upstream (µs) poly(A) site (see Fig. 3). Regulation by cell-specific factors through *cis*-acting sequences appeared unlikely, as an artificial pre-mRNA transcribed from a heterologous gene, with the same arrangement of splicing and polyadenylation signals as the µ gene, was found to be regulated in a similar manner (Peterson 1994b). This suggested that regulation might be modulated by changes in a general RNA processing factor(s), and recent studies have suggested that the switch from µm to µs is regulated through changes in the abundance of CstF-64 (Takagaki et al. 1996). Earlier work provided a precedent for such a phenomenon, as CstF activity was reported to change during the course of adenovirus infection in HeLa cells, possibly altering 3′-end processing of viral pre-mRNA (Mann et al. 1993). Furthermore, nonlethal mutations in *su(f)* (which encodes the *Drosophila* homolog of CstF-77) can change the relative utilization of alternative polyadenylation signals in unlinked genes (for review, see O'Hare 1995).

In the µ gene, the upstream µs polyadenylation signal is relatively weak and increasing its strength by mutation can result in constitutive expression of the secreted
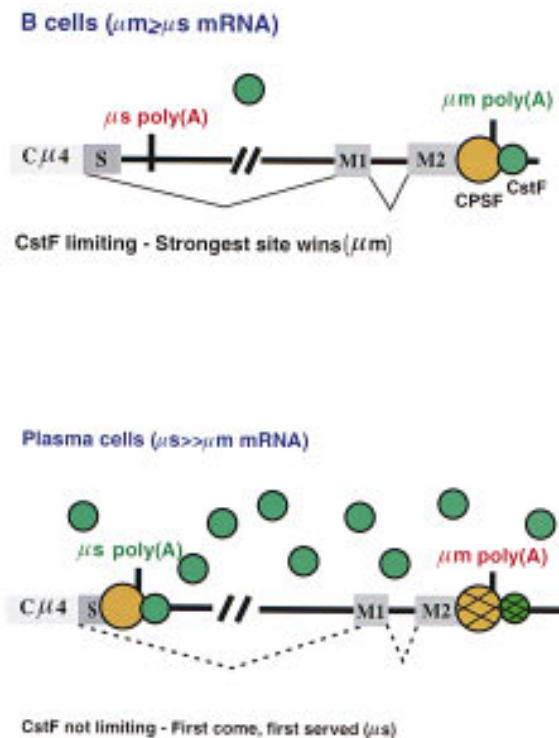

Figure 3. Regulation of mRNA processing by changes in the concentration of a basal polyadenylation factor. In B cells (*top*), the concentration of CstF-64 and, hence CstF (green circles), is low, resulting in inefficient utilization of the weak µs poly(A) site and preferential Cµ4–M1 splicing and use of the stronger µm poly(A) site, generating µm mRNA. In plasma cells (*bottom*), the concentration of CstF is elevated and no longer limiting for processing at the µs poly(A) site. Because of its 5′ location this site is thus used preferentially, leading to synthesis of µs mRNA. (The crosshatched circles indicate that the µm poly(A) site could still be recognized, but this would be unproductive when the upstream µs site is used. Likewise, use of the µs poly(A) site precludes the competing splicing event.) (See text for further discussion.) Adapted, with permission, from Takagaki et al. (1996).

form (for review, see Peterson 1994a). To examine whether changes in *trans*-acting factors might also influence poly(A) site choice, Western blot anlaysis was performed on mouse primary pre B cells activated by treatment with lipopolysaccharide (LPS). In resting cells, which preferentially utilize the strong µm poly(A) site, the concentration of CstF-64 was relatively low. But when cells were activated, the expression of CstF-64 was specifically increased 10- to 20-fold while the expression level of other factors tested remained constant. Thus, an increased level of CstF-64 correlated with use of the µs site and synthesis of secreted form mRNA. To obtain evidence that enhanced CstF-64 is sufficient to switch poly(A) site selection, a chicken B-cell line expressing a ratio of µs to µm similar to that in natural B cells was stably transfected with a vector encoding chicken CstF-64. In accordance with the hypothesis that the expression level of CstF-64 alters µ gene pre-mRNA processing,

lines expressing high levels of CstF-64 did indeed favor expression of μs, similar to the level seen when mouse primary B cells were induced to differentiate. Furthermore, there was a higher level of intact CstF complex formed in the CstF-64 overexpressing cells, suggesting that CstF-64 is limiting for complex formation.

This scenario is consistent with a model in which the CstF-64 levels control μ mRNA alternative processing (Fig. 3). In resting B cells, the limiting concentration of CstF results in preferred recognition of the strong downstream (μm) poly(A) site. Upon activation of the cells, some unknown event triggers enhanced CstF-64 expression, and because the other components of CstF are present in excess, this results in an increased level of CstF. This permits preferential use of the weaker upstream (μs) poly(A) site simply because it is transcribed first. Other studies have suggested that additional, unidentified factors may influence poly(A) site selection in B cells, implying that changes in CstF-64 levels may not tell the whole story (Edwalds-Gilbert and Milcarek 1995; Yan et al. 1995).

An important question that emerges is whether this is a widespread mechanism of gene control. A large number of genes give rise to transcripts that utilize alternative poly(A) sites, but in most cases the functional significance of this is unclear (for review, see Edwalds-Gilbert et al. 1997). An exception is the mouse CREM gene, where a switch from a downstream to an upstream site occurs during spermatogenesis (Foulkes et al. 1993), resulting in a more stable mRNA and enhanced CREM expression. It will be of interest to learn whether this switch also involves changes in the concentration of a basal polyadenylation factor.

Another example of regulation of a basal polyadenylation factor is M-phase repression of PAP, which, because of the target, likely affects synthesis of all polyadenylated transcripts (Colgan et al. 1996). M phase is driven by the cyclin-dependent kinase (cdk) p34^{cdc2}/cyclin B [mitosis or maturation promoting factor (MPF)], and many proteins have been shown to be phosphorylated, resulting in changes in their structure or activity. In addition to causing structural changes [such as chromosome condensation and dissociation of the nuclear lamina (for review, see Nigg 1991)], M phase results in repression of mRNA and protein synthesis. Recent studies have revealed that repressed gene expression is not a passive effect of the structural rearrangements but, rather, a specific shutdown via phosphorylation. Components of the transcriptional (for review, see Gottesfeld and Forbes 1997) and translational (see Ross 1997 and references therein) machinery are all targeted by MPF. Recently it was shown that PAP is also down-regulated during mitosis (Colgan et al. 1996). Three potential cdk sites, clustered in the S/T-rich regulatory domain of PAP, were found to constitute a target for MPF, resulting in hyperphosphorylation and strong repression of PAP activity. PAP hyperphosphorylated by MPF is repressed in both specific and nonspecific PAP assays, showing that the intrinsic activity of the polymerase is inhibited. Importantly, PAP isolated from mitotic HeLa cells was shown

to be similarly hyperphosphorylated and to have significantly reduced PAP activity. This down-regulation of PAP is reminiscent of the inhibitory effect of U1A protein on PAP via association with its carboxy-terminal 20 amino acids, and it will be of interest to determine whether both situations involve a similar mechanism of repression (see Fig. 4).

M-phase-specific phosphorylation of PAP is also observed during meiotic maturation of *Xenopus laevis* oocytes (Ballantyne and Wickens 1995) and again involves the cdk sites (Colgan et al. 1996). Reduced poly(A) (Wilt 1977; Sagata et al. 1980) and protein synthesis (Sagata et al. 1980; Kanki and Newport 1991) has been reported in unfertilized eggs, which contain high levels of active MPF and are arrested at M phase of meiosis II. Thus, down-regulation of PAP by MPF hyperphosphorylation probably contributes to the quiescent nature of both mitotic somatic cells and unfertilized eggs. This repression may be required in somatic cells to prevent inappropriate polyadenylation when the nuclear envelope disassembles. It is intriguing that this event does not occur in yeast, which could provide a rationale for why yeast PAP lacks an S/T-rich regulatory domain (see above). PAP repression may also be part of the shutdown of mRNA production required during a time of structural changes, possibly preventing interference with chromosome condensation and segregation. Furthermore, evolution may have chosen M phase as the point in the cell cycle to arrest unfertilzed eggs because their quiescence is conducive to the reprogramming of the genome (Fulka et al. 1996). It will be of interest in the future to determine whether repression of PAP activity by hyperphosphorylation is an essential feature of M-phase progression.

Conclusions

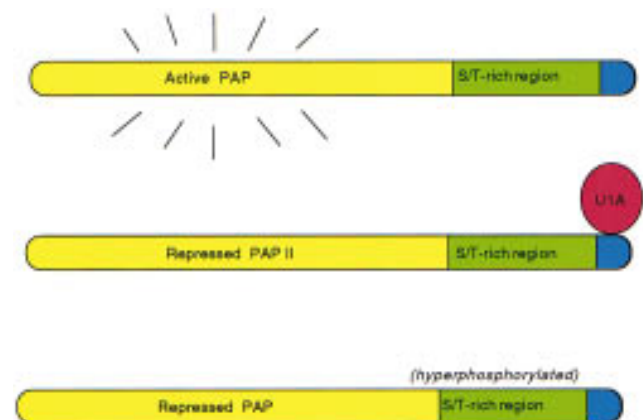Perhaps in hindsight it is not surprising that a central cellular activity such as polyadenylation should be quite

Figure 4. Inhibition of PAP activity via its carboxyl terminus. Phosphorylation of the PAP carboxyl terminus by p34^{cdc2}/cyclin B or interaction between the U1 snRNP A protein and the carboxy-terminal 20 amino acids of PAP II inhibits polyadenylation activity.

so complex. This complexity likely allows coordination of polyadenylation with transcription and splicing, and enables the transmission of information regarding the cell's activities to the polyadenylation machinery. In light of the newly observed interdependance of 3′-end formation and other cellular processes, the activities of polyadenylation factors with unknown functions may soon come to light. More surprises may be in store, and we may find that 3′-end formation does indeed participate in the regulation of several general, as well as tissue-specific, processes. Furthermore, cloning and characterization of the components of the remaining basal polyadenylation factors, including the endonuclease, are likely to result in additional interesting and perhaps unexpected findings. The reader can thus expect the field to continue to bring forth exciting developments that not only will shed more light on this essential celluar process but also provide important insights into other areas. We suspect that we will continue to be amazed at the ways the cell exploits this seemingly simple RNA processing reaction to regulate and coordinate its activities.

References

Amrani, N., M. Minet, M. Le Gouar, F. Lacroute, and F. Wyers. 1997. Yeast Pab1 interacts with RNA15 and participates in the control of the poly(A) tail length in vitro. *Mol. Cell. Biol.* **17:** 3694–3701.

Ashe, M.P., P. Griffin, W. James, and N.J. Proudfoot. 1995. Poly(A) site selection in the HIV-1 provirus: Inhibition of promoter-proximal polyadenylation by the downstream major splice donor site. *Genes & Dev.* **9:** 3008–3025.

Ballantyne, S. and M. Wickens. 1995. Poly(A) polymerase in the nucleus and cytoplasm of frog ooctyes: Dynamic changes during oocyte maturation and early development. *RNA* **1:** 64–78.

Barabino, S.M.L., W. Hubner, A. Jenny, L. Minvielle-Sebastia, and W. Keller. 1997. The 30-kD subunit of mammalian cleavage and polyadenylation specificity factor and its yeast homolog are RNA-binding zinc finger proteins. *Genes & Dev.* **11:** 1703–1716.

Bienroth, S., E. Wahle, C. Satler-Crazzolara, and W. Keller. 1991. Purification of the cleavage and polyadenylation factor involved in 3′ processing of mRNA precursors. *J. Biol. Chem.* **266:** 19768–19776.

Bienroth, S., W. Keller, and E. Wahle. 1993. Assembly of a processive messenger RNA polyadenylation complex. *EMBO J.* **12:** 585–594.

Bilger, A., C.A. Fox, E. Wahle, and M. Wickens. 1994. Nuclear polyadenylation factors recognize cytoplasmic polyadenylation elements. *Genes & Dev.* **8:** 1106–1116.

Boelens, W.C., E.J. Jansen, W.J. van Venrooij, R. Stripecke, I. Mattaj, and S.I. Gunderson. 1993. The human U1 snRNP-specific U1A protein inhibits polyadenylation of its own pre-mRNA. *Cell* **72:** 881–892.

Chanfreau, G., S. Nobel, and C. Guthrie. 1996. Essential yeast protein with unexpected similarity to subunits of mammalian cleavage and polyadenylation specificty factor (CPSF). *Science* **274:** 1511–1514.

Chen, F., C. Macdonald, and J. Wilusz. 1995. Cleavage site determinants in the mammalian polyadenylation signal. *Nucleic Acids Res.* **23:** 2614–2620.

Colgan, D.F., K.G.K. Murthy, C. Prives, and J.L. Manley. 1996.

Cell-cycle related regulation of poly(A) polymerase by phosphorylation. *Nature* **384:** 282–285.

Dantonel, J.C., K.G.K. Murthy, J.L. Manley, and L. Tora. 1997. CPSF links transcription and mRNA 3′ end formation. *Nature* **389:** 399–402.

Darnell, J.E., R. Wall, and R.J. Tushinski. 1971. An adenylic acid-rich sequence in messenger RNA of HeLa cells and its possible relationship to reiterated sites in DNA. *Proc. Natl. Acad. Sci.* **98:** 1321–1325.

Edmonds, M. and R. Abrams. 1960. Polynucleotide biosynthesis: Formation of a sequence of adenylate units from adenosine triphosphate by an enzyme form thymus nuclei. *J. Biol. Chem.* **235:** 1142–1148.

Edmonds, M., M.H. Vaughan, and H. Nakazato. 1971. Polyadenylic acid sequences in the heterogenous nuclear RNA and rapidly-labeled polyribosomal RNA of HeLa cells: possible evidence for a precursor relationship. *Proc. Natl. Acad. Sci.* **68:** 1336–1340.

Edwalds-Gilbert, G. and C. Milcarek. 1995. Regulation of poly(A) site use during mouse B-cell development involves a change in the binding of a general polyadenylation factor in a B-cell stage-specific manner. *Mol. Cell. Biol.* **15:** 6420–6429.

Edwalds-Gilbert, G., K.L. Veraldi, and C. Milcarek. 1997. Alternative poly(A) site selection in complex transcription units: Means to an end? *Nucleic Acids. Res.* **25:** 2547–2561.

Ford, J.P. and M.T. Hsu. 1978. Transcription pattern of in vivo-labeled late simian virus 40 RNA: Equimolar transcription beyond the mRNA 3′ terminus. *J. Virol.* **28:** 795–801.

Foulkes, N.S., F. Schlotter, P. Pevet, and P. Sassone-Corsi. 1993. Pituitary hormone FSH directs the CREM functional switch during spermatogenesis. *Nature* **362:** 264–267.

Fulka, J., N.L. First, and R.M. Moor. 1996. Nuclear transplantation in mammals: Remodeling of transplanted nuclei under the influence of maturation promoting factor. *BioEssays* **18:** 835–840.

Furth, P., W. Choe, J. Rex, J.C. Byrne, and C. Baker. 1994. Sequences homologous to 5′ splice sites are required for the inhibitory activity of papillomavirus late 3′ untranslated regions. *Mol. Cell. Biol.* **14:** 5278–5289.

Gebauer, F. and J.D. Richter. 1995. Cloning and characterization of a *Xenopus* poly(A) polymerase. *Mol. Cell. Biol.* **15:** 1422–1430.

Gilmartin, G.M., E.S. Fleming, J. Oetjen, and B.R. Graveley. 1995. CPSF recognition of an HIV-1 mRNA 3′ processing enhancer: multiple sequence contacts involved in poly(A) site definition. *Genes & Dev.* **9:** 72–83.

Goebell, M. and M. Yanagida. 1991. The TPR snap helix: A novel protein repeat motif form mitosis to transcription. *Trends Biochem. Sci.* **16:** 173–177.

Gottesfeld, J.M. and D.J. Forbes. 1997. Mitotic repression of the transcriptional machinery. *Trends Biochem. Sci.* **22:** 197–202.

Graveley, B.R. and G.M. Gilmartin. 1996. A common mechanism for the enhancement of mRNA 3′ processing by U3 sequences in two distantly related lentiviruses. *J. Virol.* **70:** 1612–1617.

Graveley, B.R., E.S. Fleming, and G. Gilmartin. 1996. RNA structure is a critical determinant of poly(A) site recognition by cleavage and polyadenylation specificity factor. *Mol. Cell. Biol.* **16:** 4942–4951.

Gunderson, S.I., K. Beyer, G. Martin, W. Keller, W.C. Boelens, and I.W. Mattaj. 1994. The human U1A snRNP protein regualtes polyadenylation via direct interaction with poly(A) polymerase. *Cell* **76:** 531–541.

Gunderson, S.I., S. Vagner, M. Polycarpou-Schwarz, and I. Mat-

taj. 1997. Involvement of the carboxyl terminus of vertebrate poly(A) polymerase in U1A autoregulation and in the coupling of splicing and polyadenylation. *Genes & Dev.* **11:** 761–773.

Guo, Z. and F. Sherman. 1996. 3′-End forming signals of yeast mRNA. *Trends Biochem. Sci.* **21:** 477–481.

Hake, L.E. and J.D. Richter. 1997. Translational regulation of maternal mRNA. *Biochim. Biophys. Acta* **1332:** M31–M38.

Hashimoto, C. and J.A. Steitz. 1986. A small nuclear ribonucleoprotein associates with the AAUAAA polyadenylation signal *in vitro*. *Cell* **45:** 581–591.

Jenny, A. and W. Keller. 1995. Cloning of cDNAs encoding the 160 kDa subunit of the bovine cleavage and polyadenylation specificity factor. *Nucleic Acids Res.* **23:** 2629–2635.

Jenny, A., H. Hauri, and W. Keller. 1994. Characterization of cleavage and polyadenylation specificity factor and cloning of its 100-kilodalton subunit. *Mol. Cell. Biol.* **14:** 8183–8190.

Jenny, A., L. Minvielle-Sebastia, P.J. Preker, and W. Keller. 1996. Sequence similarity between the 73-kilodalton protein of mammalian CPSF and a subunit of yeast polyadenylation factor I. *Science* **274:** 1511–1514.

Kanki, J.P. and J.W. Newport. 1991. The cell cycle dependence of protein synthesis during *Xenopus laevis* development. *Dev. Biol.* **146:** 191–213.

Keller, W. 1995. No end yet to messenger RNA 3′ processing! *Cell* **81:** 829–832.

Keller, W. and L. Minvielle-Sebastia. 1997. A comparison of mammalian and yeast pre-mRNA 3′-end processing. *Curr. Opin. Cell Biol.* **9:** 329–336.

Kessler, M.M., J. Zhao, and C.L. Moore. 1996. Purification of the *Saccharomyces cerevisiae* cleavage/polyadenylaton factor I. *J. Biol. Chem.* **271:** 27167–27175.

Kessler, M.M., M.F. Henry, E. Shen, J. Zhao, S. Gross, P.A. Silver, and C.L. Moore. 1997. Hrp1, a sequence-specific RNA-binding protein that shuttles between the nucleus and the cytoplasm, is required for mRNA 3′-end formation in yeast. *Genes & Dev.* **11:** 2545–2556.

Lee, S., J. Mendecki, and G. Brawerman. 1971. A polynucleotide sequence rich in adenylaic acid in the rapidly-labeled polyribosomal RNA component of mouse sarcoma 180 ascite cells. *Proc. Natl. Acad. Sci.* **68:** 1331–1335.

Lewis, J., S. Gunderson, and I.W. Mattaj. 1995. The influence of 5′ and 3′ end structures on pre-mRNA metabolism. *J. Cell Sci.* (Suppl.) **19:** 13–19.

Lin, C. and J.G. Patton. 1995. Regulation of alternative 3′ splice site selection by constitutive splicing factors. *RNA* **1:** 234–235.

Lingner, J., J. Kellerman, and W. Keller. 1991. Cloning and expression of the essential gene for poly(A) polymerase from yeast. *Nature* **354:** 496–498.

Lou, H., Y. Yang, G.J. Cote, and S.M. Berget. 1995. An intron splicing enhancer containing a 5′ splice site sequence in the human calcitonin/calcitonin gene related peptide. *Mol. Cell. Biol.* **15:** 7135–71442.

Lou, H., R.F. Gagel, and S.M. Berget. 1996. An intron enhancer recognized by splicing factors activates polyadenylation. *Genes & Dev.* **10:** 208–219.

Lutz, C.S. and J.C. Alwine. 1994. Direct interaction of the U1 snRNP-A protein with the upstream efficiency element of the SV40 late polyadenylation signal. *Genes & Dev.* **8:** 576–586.

Lutz, C.S., K.G.K. Murthy, N. Schek, J.P. O'Connor, J.L. Manley, and J.C. Alwine. 1996. Interaction between the U1 sn-RNP-A protein and the 160-kD subunit of cleavage–polyadenylation specificity factor increases polyadenylation efficiency in vitro. *Genes & Dev.* **10:** 325–337.

MacDonald, C., J. Wilusz, and T. Shenk. 1994. 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.* **14:** 6647–6654.

Manley, J.L. 1983. Accurate and specific polyadenylation of mRNA precursors in a soluable whole-cell lysate. *Cell* **33:** 595–605.

———. 1995. A complex protein assembly catalyzes polyadenylation of mRNA precursors. *Curr. Opin. Genet. Dev.* **5:** 222–228.

Manley, J.L. and Y. Takagaki. 1996. The end of the message-another link between yeast and mammals. *Science* **274:** 1481–1482.

Manley, J.L., P.A. Sharp, and M.L. Gefter. 1982. RNA synthesis in isolated nuclei, processing of adenovirus sero type 2 late messenger RNA precursors. *J. Mol. Biol.* **159:** 581–599.

Mann, K.P., E.A. Weiss, and J.R. Nevins. 1993. Alternative poly(A) site utilization during adenovirus infection coincides with a decrease in the acitivity of a poly(A) site processing factor. *Mol. Cell. Biol.* **13:** 2411–2419.

Martin, G. and W. Keller. 1996. Mutational analysis of mammalian poly(A) polymerase identifies a region for primer binding and a catalytic domain, homologous to the family X polymerases, and to other nucleotidyltransferases. *EMBO J.* **15:** 2593–2603.

McCracken, S., N. Fong, K. Yankulov, S. Ballantyne, G. Pan, J. Greenblatt, S.D. Patterson, M. Wickens, and D.L. Bentley. 1997. The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature* **385:** 357–361.

Minvielle-Sebastia, L., P.J. Preker, and W. Keller. 1994. RNA14 and RNA15 proteins as components of a yeast pre-mRNA 3′-end processing factor. *Science* **266:** 1702–1705.

Minvielle-Sebastia, L., P.J. Preker, T. Weiderkehr, Y. Strahm, and W. Keller. 1997. The major yeast poly(A)-binding protein is associated with cleavage factor IA and functions in pre-messenger RNA 3′-end formation. *Proc. Natl. Acad. Sci.* **94:** 7897–7902.

Mitchelson, A., M. Simonelig, C. Williams, and K. O'Hare. 1993. Homology with *Saccharomyces cerevisiae RNA 14* suggests that phenotypic suppression in *Drosophila melanogaster* by *suppressor of forked* occurs at the level of mRNA stability. *Genes & Dev.* **7:** 241–249.

Moore, C.L. and P.A. Sharp. 1984. Site-specific polyadenylation in a cell-free reaction. *Cell* **36:** 581–591.

———. 1985. Accurate cleavage and polyadenylation of exogenous RNA substrates. *Cell* **41:** 845–855.

Mulligan, G.J., W. Guo, S. Wormsley, and D.M. Helfman. 1992. Polypyrimidine tract binding protein interacts with sequences involved in alternative splicing of β-tropomyosin pre-mRNA. *J. Biol. Chem.* **267:** 25480–25487.

Murthy, K.G.K. and J.L. Manley. 1992. Characterization of the multisubunit cleavage and polyadenylation factor from calf thymus. *J. Biol. Chem.* **267:** 14804–14811.

———. 1995. The 160-kD subunit of human cleavage–polyadenylation specificity factor coordinates pre-mRNA 3′-end formation. *Genes & Dev.* **9:** 2672–2683.

Nevins, J.R. and J.E. Darnell. 1978. Steps in the processing of Ad2 mRNA: poly(A)$^+$ nuclear sequences are conserved and poly(A) addition precedes splicing. *Cell* **15:** 1477–1493.

Nigg, E.A. 1991. The substrates of the cdc2 kinase. *Semin. Cell Biol.* **2:** 261–270.

Niwa, M., S.D. Rose, and S.M. Berget. 1990. In vitro polyadenylation is stimulated by the presence of an upstream intron. *Genes & Dev.* **4:** 1552–1559.

O'Hare, K. 1995. mRNA 3′ ends in focus. *Trends Genet.* **11:** 255–257.

Peterson, M.L. 1994a. RNA processing and expression of immunoglobulin genes. In *Handbook of B and T lymphocytes* (ed. E.C. Snow), pp. 321–342. Academic Press, San Diego, CA.

———. 1994b. Regulated immunoglobulin (Ig) RNA processing does not require specific cis-active sequences: Non-Ig RNA can be alternatively processed in B cells and plasma cells. *Mol. Cell. Biol.* **14:** 7891–7898.

Preker, P.J., J. Lingner, L. Minivielle-Sebastia, and W. Keller. 1995. The *FIP1* gene encodes a component of a yeast pre-mRNA polyadenylation factor and directly interacts with poly(A) polymerase. *Cell* **81:** 379–389.

Proudfoot, N.J. 1989. How RNA polymerase II terminates transcription in higher eukaryotes. *Trends Biochem. Sci.* **14:** 105–110.

———. 1991. Poly(A) signals. *Cell* **64:** 671–674.

———. 1996. Ending the message is not so simple. *Cell* **87:** 779–781.

Raabe, T., F. Bollum, and J.L. Manley. 1991. Primary structure and expression of bovine poly(A) polymerase. *Nature* **353:** 229–234.

Raabe, T., K.G.K. Murthy, and J.L. Manley. 1994. Poly(A) polymerase contains multiple functional domains. *Mol. Cell. Biol.* **14:** 2946–2957.

Richter, J.D. 1996. Dynamics of poly(A) addition and removal during development. In *Translational control* (ed. M. Mathews, J. Hershey, and N. Sonenberg), pp. 481–503. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Ross, J. 1997. A hypothesis to explain why translation inhibitors stabilize mRNAs in mammalian cells: mRNA stability and mitosis. *BioEssays* **19:** 527–529.

Ruegsegger, U., K. Beyer, and W. Keller. 1996. Purification and characterization of human cleavage factor Im involved in the 3′ end processing of messenger RNA precursors. *J. Biol. Chem.* **271:** 6107–6113.

Sachs, A.B., P. Sarnow, and M.W. Hentze. 1997. Starting at the beginning, middle and end: Translation initiation in eukaryotes. *Cell* **89:** 831–838.

Sagata, N., K. Shiokawa, and K. Yamana. 1980. A study on the steady-state population of poly(A)+ RNA during early development of *Xenopus laevis. Dev. Biol.* **77:** 421–448.

Simonelig, M., K. Elliot, A. Mitchelson, and K. O'Hare. 1996. Interallelic complementation at the *suppressor of forked* locus of Drosophila reveals complementation between suppressor of forked proteins mutated in different regions. *Genetics* **142:** 1225–1235.

Singh, R., J. Valcarcel, and M.R. Green. 1995. Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. *Science* **268:** 1173–1176.

Spalhlolz, B.A. and P.M. Howley. 1989. Papillomavirus-host cell interactions. *Adv. Viral Oncol.* **8:** 27–53.

Stumpf, G. and H. Domdey. 1996. Dependence of yeast pre-mRNA 3′-end processing of CFT1: A sequence homolog of the mammalian AAUAAA binding factor. *Science* **274:** 1517–1520.

Takagaki, Y. and J.L. Manley. 1992. A human polyadenylation factor is a G protein β-subunit homologue. *J. Biol. Chem.* **267:** 23471–23474.

———. 1994. A polyadenylation factor subunit is the human homologue of the *Drosophila* suppressor of forked protein. *Nature* **372:** 471–474.

———. 1997. RNA recognition by the human polyadenylation factor CstF. *Mol. Cell. Biol.* **17:** 3907–3914.

Takagaki, Y., L.C. Ryner, and J.L. Manley. 1989. Four factors are required for 3′-end cleavage of pre-mRNAs. *Genes & Dev.* **3:** 1171–1124.

Takagaki, Y., C.C. MacDonald, T. Shenk, and J.L. Manley. 1992.

The human 64-kDa polyadenylation factor contains a ribonucleoprotein-type RNA binding domain and unusual auxillary motifs. *Proc. Natl. Acad. Sci.* **89:** 1403–1407.

Takagaki, Y., R.L. Seipelt, M.L. Peterson, and J.L. Manley. 1996. The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B-cell differentiation. *Cell* **87:** 941–952.

Tansey, W.P. and W. Herr. 1997 TAFs: Guilt by association? *Cell* **88:** 729–732.

Wahle, E. 1995. Poly(A) tail length is caused by termination of processive synthesis. *J. Biol. Chem.* **270:** 2800–2808.

Wahle, E. and W. Keller. 1996. The biochemistry of polyadenylation. *Trends Biochem. Sci.* **21:** 247–250.

Wahle, E., G. Martin, E. Schlitz, and W. Keller. 1991. Isolation and expression of cDNA clones encoding mammalian poly(A) polymerase. *EMBO J.* **10:** 4251–4257.

Wahle, E., A. Lustig, P. Jeno, and P. Maurer. 1993. Mammalian poly (A)-binding protein II. *J. Biol. Chem.* **268:** 2937–2945.

Wasserman, K.M. and J.A. Steitz. 1993. Association with terminal exons in pre-mRNAs: A new role of the U1 snRNA? *Genes & Dev.* **7:** 647–659.

Wickens, M., P. Anderson, and R.J. Jackson. 1997. Life and death in the cytoplasm: Messages from the 3′ end. *Curr. Opin. Genet. Dev.* **7:** 220–232.

Wilt, F. 1977. The dynamics of maternal poly(A)-containing mRNA in fertilized sea urchin eggs. *Cell* **11:** 673–681.

Wilusz, J. and T. Shenk. 1988. A 64kD nuclear protein binds to RNA segments that include the AAUAAA polyadenylation motif. *Cell* **52:** 221–228.

Wormington, M. 1994. Unmasking the role of the 3′ UTR in the cytoplasmic polyadenylation and translational regulation of maternal mRNAs. *BioEssays* **16:** 533–535.

Yan, D.-H., E.A. Weiss, and J.R. Nevins. 1995. Identification of an activity in B-cell extracts that selectively impairs the formation of an immunoglobulin μs poly(A) site processing complex. *Mol. Cell. Biol.* **15:** 1901–1906.

Zachar, Z., T. Chou, and P.M. Bingham. 1987. Evidence that a regulatory gene autoregulates splicing of its transcript. *EMBO J.* **6:** 4105–111.

Zhao, W. and J.L. Manley. 1996. Complex alternative RNA processing generates an unexpected diversity of poly(A) polymerase isoforms. *Mol. Cell. Biol.* **16:** 2378–2386.

Zhao, J., M.M. Kessler, and C.L. Moore. 1997. Cleavage factor II of *Saccharomyces cerevisiae* contains homologues to subunits of the mammalian cleavage/polyadenylation specificity factor and exhibits sequence-specific, ATP-dependent interaction with precursor RNA. *J. Biol. Chem.* **272:** 10831–10838.

Zhelkovsky, A.M., M.M. Kessler, and C.L. Moore. 1995. Structure-function relationships in the *Saccharomyces cerevisiae* poly(A) polymerase: Identification of a novel RNA binding site and a domain that interacts with specificity factor(s). *J. Biol. Chem.* **270:** 26715–26720.

Mechanism and regulation of mRNA polyadenylation

Diana F. Colgan and James L. Manley

References	This article cites 101 articles, 53 of which can be accessed free at: **http://genesdev.cshlp.org/content/11/21/2755.full.html#ref-list-1**
License	
Email Alerting Service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**