Polyadenylation: alternative lifestyles of the A-rich (and famous?)

Alexa M Dickson and Jeffrey Wilusz*

Department of Microbiology, Immunology and Pathology, Colorado State University, Fort Collins, CO, USA

*Correspondence to: jeffrey.wilusz@colostate.edu

The EMBO Journal (2010) 29, 1473–1474. doi:10.1038/emboj.2010.67

Alternative polyadenylation has a major function in gene expression and is often mediated through signals that lack canonical signatures. In this issue, Nunes *et al* (2010) uncover a new upstream A-rich sequence element that in conjunction with a strong U/GU-rich downstream element may be responsible for up to a third of polyadenylation events in mammalian cells that are not mediated by a canonical AAUAAA hexamer.

In the not too distant past, the process of mRNA 3' end formation (aka polyadenylation) was largely considered a default event along the pathway of gene expression. It is now abundantly clear that polyadenylation is not only a fundamental step in mRNA biogenesis but is also highly regulated and networked with other aspects of gene expression. Interestingly, recent evidence indicates that the choice of polyadenylation site in many mRNAs changes in response to cell growth, developmental cues or oncogene activation (Ji and Tian, 2009; Mayr and Bartel, 2009). The resultant shortening of an mRNA's 3' untranslated region can remove a variety of regulatory elements, including miRNA target sites, and thereby dramatically alter gene expression patterns. Thus, how these alternative polyadenylation sites are selected in the ~50% of genes that possess them has emerged as a key question in the gene expression arena.

The textbook definition of a mammalian polyadenylation signal is an upstream AAUAAA hexamer plus a less conserved U-rich or GU-rich downstream element. Database analyses, however, show that up to 30% of polyadenylation events do not seem to involve an AAUAAA hexamer (or its close relative AUUAAA) and ~50% do not possess either core element. Curiously, many of these 'non-canonical' signals lie significantly 5' of the default polyadenylation signal in the 3' untranslated region of a pre-mRNA, making them prime targets for regulation and alternative polyadenylation events. Thus, to elucidate mechanisms of regulated alternative 3' end processing, we need to understand how non-canonical polyadenylation events are carried out. One fundamental aspect of this is to determine the nucleic acid elements or code(s) that mediate non-canonical polyadenylation events.

In studying 3' end formation of the melanocortin 4 receptor gene, Nunes *et al* (2010) in this issue of *The EMBO Journal* have uncovered what seems to be a major new class of non-canonical mammalian polyadenylation signal. This new 'A-rich' class of poly(A) signals is comprised of a critical

U/GU-rich downstream core element and an upstream A tract (Figure 1). Thus, it completely lacks an AAUAAA hexamer and is reminiscent of 3' end processing signals in plants and yeast. Furthermore, bioinformatic analysis suggests that A-rich poly(A) signals may account for up to a third of non-canonical 3' end processing events in mammalian cells. Finally, as changing AAUAAA to AAAAAA in the context of a standard poly(A) signal significantly reduces processing efficiency, the ability of an upstream A tract to function in 3' end processing is considerably context dependent. Thus, the bottom line is that it seems we have a new upstream core polyadenylation element and class of poly(A) signals in town.

These observations are of significant interest for three main reasons. First, the discovery of A-rich core upstream polyadenylation elements obviously provides considerable insight into the *cis*-acting sequences involved in alternative polyadenylation events that have a major, only recently appreciated function in determining the landscape of the transcriptome. The A-rich element is the first major new upstream core metazoan polyadenylation element to be identified since the Gilmartin laboratory uncovered a role for UGUAN sequences in the 3' end processing of select mRNAs (Venkataraman *et al*, 2005). Second, the study highlights the significant function played by the downstream

Figure 1 Alternative 3' end formation can require a choice between different types of polyadenylation signals. In a canonical poly(A) signal, the AAUAAA hexamer (bold) that interacts with CPSF is the dominant aspect, whereas in an A-rich poly(A) signal, the downstream U/GU element (bold) that interacts with CstF seems to be critical. The factors that are required to mediate usage of A-rich poly(A) signals are not clear, nor are the mechanisms that underlie alternative polyadenylation.

U/GU-rich element in determining the efficiency of 3' end processing. Therefore, this study should help focus an appropriate level of attention on downstream core elements as we develop a molecular understanding of alternative 3' end processing. Finally, although a significant evolutionary conservation in the core poly(A) factors (CPSF, CstF, CFI, CFII and PAP) among yeast, plants and mammals has been recognized for years, the reliance on AAUAAA hexamers for CPSF factor binding and initiation of 3' end processing is largely restricted to mammalian 3' end processing. Thus, this study suggests that there also remains a significant evolutionary relationship at the level of *cis*-acting elements as well.

The discovery of this A-rich class of polyadenylation element poses a number of interesting questions. How are A-rich polyadenylation signals recognized? Is it through the same general polyadenylation machinery (e.g. CPSF) that interacts

with canonical poly(A) signals or are specific factors involved that remain to be elucidated? Are there any functional consequences for an mRNA or for networked processes such as splicing and transcription termination when 3' end processing occurs through an A-rich versus AAUAAA element-containing polyadenylation signal? Do chromatin signatures that have recently been shown to be present around sites of 3' end processing (Spies *et al*, 2009) influence the selection of A-rich elements? Answers to these and other related questions will undoubtedly provide significant insight into the mechanisms that govern regulated alternative polyadenylation in mammalian cells.

Conflict of interest

The authors declare that they have no conflict of interest.

References

- Ji Z, Tian B (2009) Reprogramming of 3' untranslated regions of mRNAs by alternative polyadenylation in generation of pluripotent stem cells from different cell types. *PLoS One* **4**: e8419
- Mayr C, Bartel DP (2009) Widespread shortening of 3' UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* **138**: 673–684
- Nunes MN, Li W, Tian B, Furger A (2010) A functional human poly(A) site requires only a potent DSE and an A-rich upstream sequence. *EMBO J* **29**: 1523–1536
- Spies N, Nielsen CB, Padgett RA, Burge CB (2009) Biased chromatin signatures around polyadenylation sites and exons. *Mol Cell* **36**: 245–254
- Venkataraman K, Brown KM, Gilmartin GM (2005) Analysis of a noncanonical poly(A) site reveals a tripartite mechanism for vertebrate poly(A) signal recognition. *Genes Dev* **19**: 1315–1327