Title

Sequential Polyadenylation to Enable Alternative mRNA 3’ End Formation

Journal

Molecules and Cells, 46(1)

ISSN

1016-8478

Authors

Hao, Yajing

Cai, Ting

Liu, Chang

et al.

Publication Date

2023

DOI

10.14348/molcells.2023.2176

Molecules and Cells

Minireview

Sequential Polyadenylation to Enable Alternative mRNA 3′ End Formation

Yajing Hao[1], Ting Cai[1], Chang Liu[1], Xuan Zhang[1], and Xiang-Dong Fu[1,2,*]

[1]Department of Cellular and Molecular Medicine, Institute of Genomic Medicine, University of California San Diego, La Jolla, CA 92093, USA, [2]Present address: Westlake Laboratory of Life Sciences and Biomedicine, School of Life Sciences, Westlake University, Hangzhou 310024, China
*Correspondence: fuxiangdong@westlake.edu.cn
https://doi.org/10.14348/molcells.2023.2176
www.molcells.org

In eukaryotic cells, a key RNA processing step to generate mature mRNA is the coupled reaction for cleavage and polyadenylation (CPA) at the 3′ end of individual transcripts. Many transcripts are alternatively polyadenylated (APA) to produce mRNAs with different 3′ ends that may either alter protein coding sequence (CDS-APA) or create different lengths of 3′UTR (tandem-APA). As the CPA reaction is intimately associated with transcriptional termination, it has been widely assumed that APA is regulated co-transcriptionally. Isoforms terminated at different regions may have distinct RNA stability under different conditions, thus altering the ratio of APA isoforms. Such differential impacts on different isoforms have been considered as post-transcriptional APA, but strictly speaking, this can only be considered "apparent" APA, as the choice is not made during the CPA reaction. Interestingly, a recent study reveals sequential APA as a new mechanism for post-transcriptional APA. This minireview will focus on this new mechanism to provide insights into various documented regulatory paradigms.

Keywords: alternative polyadenylation, cleavage and polyadenylation, co-transcriptional alternative polyadenylation, mRNA 3′ end formation, RNA processing, sequential alternative polyadenylation

INTRODUCTION

Most protein coding genes in eukaryotic cells are transcribed by RNA polymerase II (RNAPII). With the exception of the canonical, replication-dependent transcripts that encode histones in metazoans, the maturation of mRNA 3′ end involves endonucleolytic cleavage of nascent RNA followed by synthesis of a poly(A) tail at the 3′ terminus of the cleaved product by poly(A) polymerase (PAP) (Fig. 1A). Multiple *cis*-acting elements, including upstream UGUA elements, core AAUAAA motif, and downstream GU-rich sequences, collectively define the functional poly(A) site (PAS), which is recognized by the poly(A) machinery (Fig. 1B) to carry out the coupled cleavage and polyadenylation (CPA) reaction (Elkon et al., 2013; Gruber and Zavolan, 2019; Tian and Manley, 2017).

The poly(A) machinery has been biochemically elucidated (Shi and Manley, 2015; Shi et al., 2009). Briefly, besides PAP, the poly(A) machinery contains 4 subcomplexes: Cleavage factor I (CFI; in mammalian cells, CFIm) binds to the upstream UGUA element. Cleavage and polyadenylation specificity factor (CPSF) recognizes the core AAUAAA motif and CPSF also tightly couples with cleavage factor II (CFII) consisting of CLP1 and PCF11. Finally, cleavage stimulation factor (CSTF) binds the downstream GU-rich sequences. Within the CPSF complex, CPSF73 is the endonuclease, and the cleavage reaction is aided by CLP1 to cut RNA ~21 nt downstream of the AAUAAA motif followed by PAP-catalyzed poly(A) addition.

Received November 10, 2022; revised November 21, 2022; accepted November 22, 2022; published online January 3, 2023

eISSN: 0219-1032

The length of the poly(A) tail is controlled by the nuclear poly(A) binding protein PABPN1 (Kuhn et al., 2009; Wahle, 1995).

Multiple lines of evidence demonstrate that the CPA reaction takes place co-transcriptionally (Mitschka and Mayr, 2022; Tian and Manley, 2017). Most importantly, CPA is tied to transcriptional termination, because RNAPII does not have a defined transcription termination site. According to the popular torpedo model, the CPA reaction generates a free 5′ end to be attached by the exonuclease XRN2 (Fig. 1A), which tailgates the elongating RNAPII to cause eventual drop-off (Connelly and Manley, 1988; Proudfoot, 2004). The alternative allosteric model suggests that CPA coupled with certain *cis*-acting elements may cause transient RNAPII pausing near the PAS, which may induce conformational changes in the elongating RNAPII complex to become increasingly prone to drop-off (Proudfoot, 2016; Zhang et al., 2015). More recent studies suggest combined flavors of both models during the coupling between CPA and transcriptional termination and defects in this process cause global transcriptional read-through (Eaton et al., 2020). Mechanistically, the poly(A) machinery is physically associated with RNAPII, which may even be loaded onto the elongating RNAPII complex at the beginning of transcription (Glover-Cutter et al., 2008).

WIDESPREAD ALTERNATIVE POLYADENYLATION IN HIGHER EUKARYOTIC CELLS

An array of technologies has been developed to enable the global profiling of PASs (Zhou et al., 2014). Interestingly, over 70% of genes have multiple PASs in both *Drosophila* (Liu et al., 2017) and humans (Derti et al., 2012). Therefore, be-

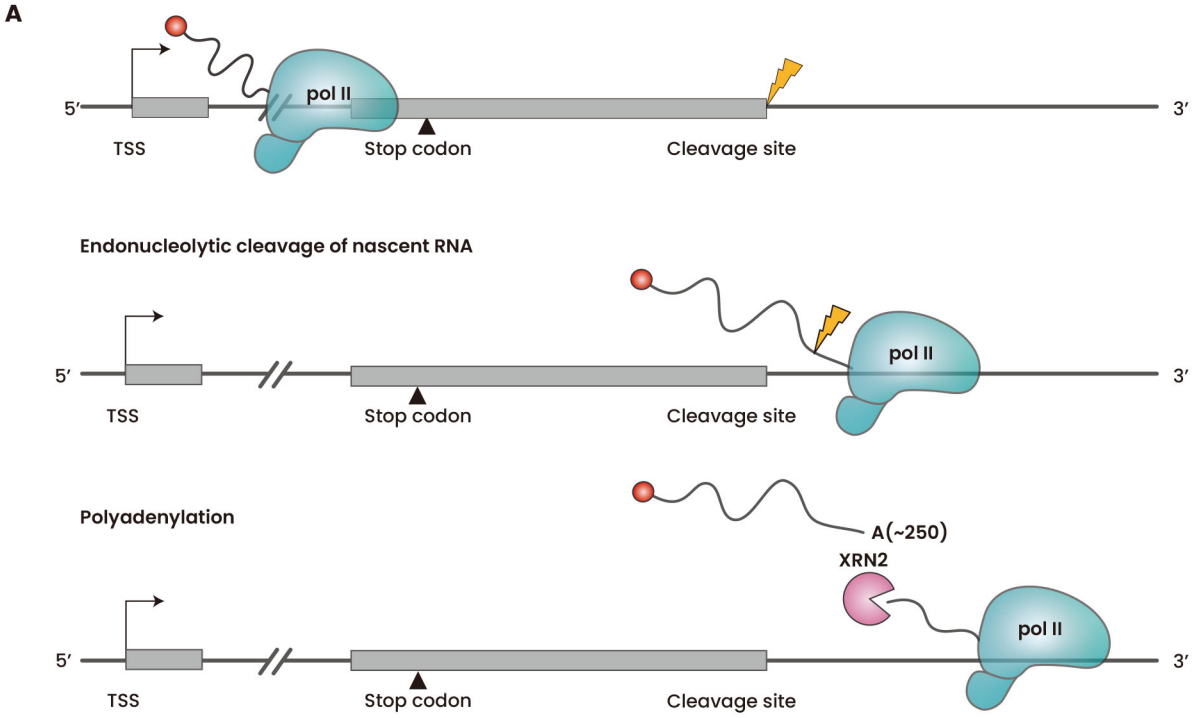

Fig. 1. mRNA 3′ end formation is controlled by *cis*-acting elements and poly(A) machinery. (A) The diagram for the cleavage and polyadenylation (CPA) reaction. Transcription is executed by RNAPII using a template DNA strand (top). Nascent RNA is cleaved at the cleavage site (middle) and then poly(A) tail is added (bottom). After the CPA reaction, the exonuclease XRN2 degrades the RNA strand with a 5′ free end and terminates the transcription (bottom). TSS, transcription start site. (B) *Cis*-acting elements and the core poly(A) machinery for the CPA reaction. A functional PAS consists of three major *cis*-elements: upstream UGUA bound by cleavage factor I (CFI), core AAUAAA element recognized by cleavage and polyadenylation specificity factor (CPSF), and downstream GU/U rich sequences targeted by CSTF. AAUAAA are located ~21 nt upstream of the cleavage site (top). The poly(A) machinery includes 4 subcomplexes and several single proteins (bottom). The complexes are CFI, which contains two CFI-25 subunits and two larger subunits of 68 kDa (CFI-68) and/or 59 kDa (CFI-59); CPSF, which contains CPSF-30/73/160/100 (based on their molecular weight in kDa), Fip1, and Wdr33; cleavage factor II (CFII), which consists Clp1 (also known as CLP1) and Pcf11; and cleavage stimulation factor (CSTF) binding to the GU/U-rich sequence, which contains CSTF 77 kDa subunit (also known as CstF-77), CSTF 50 kDa subunit (also known as CstF-50), CSTF 64 kDa subunit (also known as CstF-64), and its paralogue, τCstF-64. Single proteins include Symplekin and poly(A) polymerase (PAP). Poly(A) binding protein 1 (PABPN1) binds to the growing poly(A) tail, disrupting the interaction between CPSF and PAP when the tail is ~250 nt. Several pieces of evidence show CTD of RNAPII could directly interact with a part of poly(A) machinery.

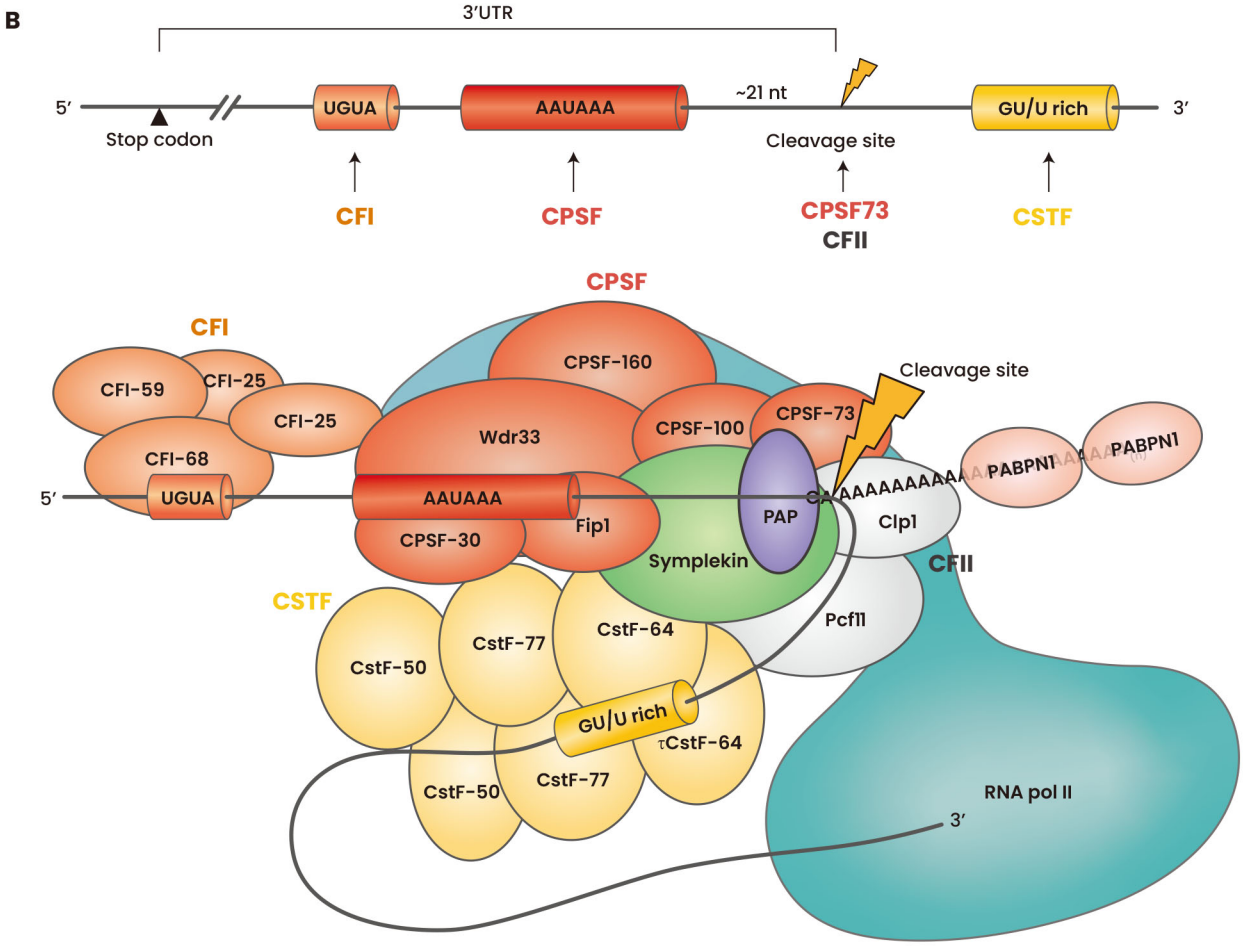

Fig. 1. Continued.

sides alternative promoters to generate mRNA isoforms with different 5′ ends and alternative splicing to produce mRNA isoforms with different sequences in gene bodies, alternative polyadenylation (APA) represents another key mechanism to produce mRNA isoforms with different 3′ ends from single genes. Growing evidence also suggests that many mRNA isoforms generated by APA may have distinct biological functions, which are subjected to regulation in development, differentiation, and disease (Gruber and Zavolan, 2019; Mitschka and Mayr, 2022; Tian and Manley, 2017). As APA that occur in an intronic region tends to truncate the C-terminus of the encoded protein, this type of APA is thus referred to as IPA (intronic polyadenylation) or CDS-APA (coding sequence-APA). A more frequent mode of APA is tandem APA with two or more PASs in the last exon, thereby generating mRNA isoforms with the same protein-coding capacity but with different lengths of 3′ untranslated region (3′UTR).

Characterization of *cis*-acting elements at proximal and distal PASs suggests a general rule underlying APA. Although the AAUAAA motif is the core signal for the CPA reaction, it is often insufficient to define a functional PAS in mammalian genomes. Both the upstream UGUA and downstream GU-rich sequences are critical, which together define the strength of a given PAS in a combinatorial fashion (Tian and Manley, 2017). In general, the proximal PAS is weaker compared to the downstream distal PAS (Gruber and Zavolan, 2019). Thus, the distal PAS may be considered a constitutive PAS while the upstream PAS a regulated PAS in APA, the latter of which may be specifically activated or repressed by a variety of mechanisms (see below). Conceptually, because the CPA reaction is mechanistically and functionally coupled with transcriptional termination, it has been generally thought that APA regulation also takes place co-transcriptionally. However, we wish to emphasize here that while there are multiple lines of evidence to support various co-transcriptional mechanisms, most studies published to date simply assume that is the case, especially those regulated by specific RNA binding proteins (listed in Mitschka and Mayr [2022]). However, in light with the recent finding of sequential APA (Tang et al., 2022), many data may be alternatively interpreted (see examples below), suggesting the need to revisit various regulatory paradigms previously thought to occur at the co-transcriptional levels.

CO- AND POST-TRANSCRIPTIONAL REGULATION OF APA

Initial reporter-based studies show that specific transcription blockage between the proximal and distal PASs is sufficient to induce APA to favor the use of the proximal PAS (Gromak et al., 2006; Yonaha and Proudfoot, 1999). Such transcription blockage may be instituted with G-rich sequences (Beaudoin and Perreault, 2013), which play a key role in initiating R-loop formation (Skourti-Stathaki et al., 2011). This mechanism may account for the frequent pausing of RNAPII at PASs (Glover-Cutter et al., 2008), especially when coupled with a mutant RNAPII with reduced elongation rate (Pinto et al., 2011).

It has been widely accepted that transcription is intimately coupled with various mRNA processing steps (Glover-Cutter et al., 2008) and such coupling may be initiated even at the beginning of transcription, as indicated by gene promoters as the hotspots for the ChIP-seq signals of many RNA binding proteins involved in diverse RNA processing pathways (Xiao et al., 2019), including those involved in CPA. However, the specific coupling mechanism(s) has remained unclear because different transcription factors, including the C-terminal domain (CTD) of the largest subunit of RNAPII (McCracken et al., 1997), TFIID (transcription factor II D) (Dantonel et al., 1997), and the elongation factor PAF1c (Nagaike et al., 2011), etc. have been implicated in the recruitment of specific components of the poly(A) machinery. Such recruitment may be further enhanced by transcription enhancers, via a variety of transcription factors and co-activators (Kwon et al., 2022). These processes may also be modulated by various epigenetic mechanisms (Kaczmarek Michaels et al., 2020; Lin et al., 2020; Nanavaty et al., 2020; Soles and Shi, 2021; Wood et al., 2008).

Despite clearly recordable effects of transcription factors on APA, a major challenge in understanding the mechanism(s) is to determine how the proximal or distal PAS is differentially impacted to account for the observed switch in APA. Two recent studies shed critical light on this: One study characterized a pair of transcriptional anti-terminators SCAF4 and SCAF8 (Gregersen et al., 2019). The homologs of these SR protein-related RNA binding proteins have been genetically implicated in transcription termination in yeast (Yuryev et al., 1996). Importantly, depletion of these proteins in mammalian cells was found to enhance transcription termination at the proximal PAS of many genes, thereby favoring the use of the weak proximal PAS. Interestingly, this process appears to be coupled with the interaction of these RNA binding proteins with specific RNA elements in nascent RNAs, thus providing a mechanism for context-dependent effects. However, how exactly such RNA binding activity contributes to the anti-termination effect remains to be worked out. It also remains to be determined how broadly this mechanism may apply to the impacted APA events. In another study, inactivation of DNA methyltransferases was found to enhance CTCF binding, which in turn impacts APA (Nanavaty et al., 2020). While this effect on APA may enlist multiple direct or indirect mechanisms, on specific gene models examined, enhanced CTCF binding was found to induce RNAPII pausing between the proximal and distal PAS, thus favoring the use of the proximal PAS. Together, these studies illustrate the "first come, first served" model for the relatively weak proximal PAS to gain the advantage over the strong downstream PAS in recognition by the poly(A) machinery (Fig. 2A).

Besides the co-transcriptional regulation of APA, several mechanisms have been suggested to represent the post-transcriptional mechanisms for APA regulation, largely through differential degradation of mRNA isoforms with different

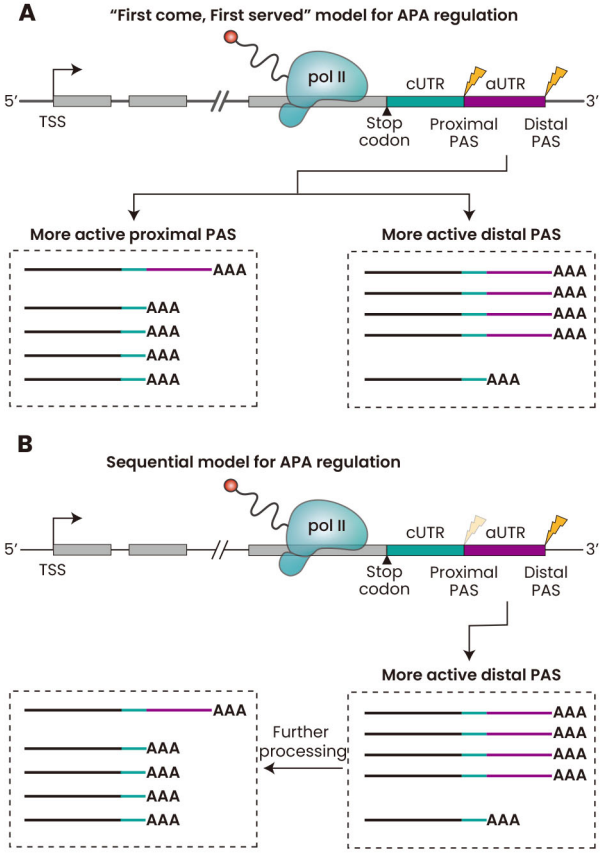

Fig. 2. Two models for alternative polyadenylation (APA) regulation. (A) "First come, First served" model for APA regulation. In the "First come, First served" model, APA is regulated by the activity of poly(A) site (PAS). A more active proximal poly(A) site (PAS) can produce mRNA isoforms with shorter 3′UTR, whereas a more active distal PAS favors longer 3′UTR. Multiple *cis*-elements and *trans*-factors could influence the strength of PAS. Different colors were added for common UTR (cUTR, the 3′UTR region upstream of the proximal polyadenylation site, which is common between the short and long isoforms) and alternative UTR (aUTR, the 3′UTR region downstream of the proximal polyadenylation site, which is present only in the long isoform). (B) Sequential model for APA regulation. In the sequential model, cleavage and polyadenylation (CPA) is first carried out at an active distal PAS, producing transcripts with longer 3′UTR. Subsequent CPA occurs at the proximal PAS, generating isoforms with shorter 3′UTR. TSS, transcription start site.

3′UTRs (Mitschka and Mayr, 2022; Tian and Manley, 2017). For example, mRNA isoforms with different 3′ ends may be differentially exported out of the nucleus, localized in different cytoplasmic compartments, or recruited to the translation machinery. As a result, different mRNA isoforms may be selectively targeted by microRNAs or other nucleases due to isoform-specific cellular compartmentalization mechanisms. These mechanisms may contribute to the global APA regulation during tumorigenesis, as cancer cells are in general associated with 3′UTR shortening (Mayr and Bartel, 2009; Xia et al., 2014) and the opposite appears to be the case with cell differentiation (Cheng et al., 2020; Ji and Tian, 2009). Here, we would like to argue that these modes of regulation cannot be considered post-transcriptional APA because alternative polyadenylation is supposed to take place during the CPA reaction to select one PAS over the other. Selective degradation of one mRNA isoform relative to the other can thus only be regarded as "apparent" switch in APA.

SEQUENTIAL POLYADENYLATION AS A NEW MECHANISM FOR POST-TRANSCRIPTIONAL APA

It has been documented that mRNAs with longer 3′UTR tend to accumulate in the nucleus compared to their isoforms with shorter 3′UTR (Djebali et al., 2012), likely due to various nuclear retention mechanisms granted by additional *cis*-elements in those longer mRNA isoforms, such as the hairpins for microRNA processing (Neve et al., 2016). Interestingly, a recent study further pursued this phenomenon, providing multiple lines of evidence that those nuclear detained long APA isoforms ended in the distal PASs can in many cases serve as the precursors to generate short APA isoforms through CPA at the proximal PASs (Tang et al., 2022). Because distal PASs are normally stronger than proximal PASs, the distal ones would be preferentially used unless the proximal ones employ certain gene-specific mechanism(s) to enhance their selection. Thus, according to the "first come, first served" model, the selection of the proximal PAS would be independent of the selection of the distal PAS. This proves not to be the case. By engineering a ribozyme between the proximal and distal PAS, the authors demonstrated that the selection of the proximal PAS for CPA depended on the presence of the distally polyadenylated sequence. Furthermore, a novel Cleave-seq was developed to provide direct evidence that the longer APA isoform is the intermediate of the shorter APA isoform. In this assay, the exonuclease XRN2 was first inactivated, thus protecting the 5′ end of the downstream RNA fragment after the cleavage reaction at the proximal PAS. By oligo(dT) selection and then ligating an unphosphorylated DNA linker with 3′-OH followed by deep sequencing, this captured the cleaved products.

This study suggests a general model for sequential polyadenylation as a key mechanism for post-transcriptional APA (Fig. 2B). In this model, the CPA reaction at the distal PAS is first carried out co-transcriptionally by the poly(A) machinery, which is coupled with transcriptional termination. The resulting products may either be exported out of the nucleus or transiently detained in the nucleus. This transient period depends on the combined strength of nuclear retention elements, such as pseudo-splice sites in individual transcripts. These detained RNAs thus serve as precursors for CPA at the proximal PAS. Interestingly, this mechanism may also allow an additional slow splicing reaction(s) to occur, which may also contribute to the nuclear retention of the precursor mRNA (Tang et al., 2022). It is conceivable that this mechanism enables various RNA binding proteins to influence APA at either co- or post-transcriptional levels. Thus, future studies need to revisit this widely presumed assertion instead of assuming that regulated APA by RNA binding proteins is always a co-transcriptional event. This resembles the mechanism for regulated alternative splicing, which predominately takes place post-transcriptionally, and as proposed earlier (Han et al., 2011), this mechanism may even also allow co-transcriptional commitment but post-transcriptional catalytic reaction.

NEW MODEL GENERATING NEW INSIGHTS INTO DOCUMENTED APA REGULATORY PARADIGMS

CFI-25 (also known as NUTD21) and CFI-68 (also known as CPSF6) are key components of the CFI complex. Because the CF1 complex is responsible for recognizing the UGUA element, which is more prevalently associated with distal PASs, knockdown of these factors has been shown to induce a global switch of APA to proximal PASs (Li et al., 2015). It has remained elusive how such switch occurred. Unaltered proximal PASs and compromised distal PASs would contribute to "apparent" APA switch. Alternatively, many proximal PASs may become reactivated. How would this occur, especially with those PASs without any sequences that resemble the UGUA motif? According to the sequential CPA model (Fig. 3A), impaired CPA at distal PASs in CFI knockdown cells would cause transcriptional readthrough as documented earlier (Zhu et al., 2018). Such aberrant transcripts (red lines in Fig. 3A) are more likely detained in the nucleus, which may then serve as the precursors for CPA at proximal sites. Future studies are needed to directly test this model in cells depleted of CFI-25 or CFI-68.

A recent study revealed a key role of NXF1, a critical adaptor for mRNA nuclear export, in the regulation of APA (Chen et al., 2019). NXF1 depletion switched APA to proximal PASs, which may result from its functional interaction with CFI-68. However, CFI-68 knockdown does not always phenocopy the effect of NXF1 knockdown, suggesting additional mechanisms in operation for both factors. NXF1 depletion-induced APA switch was proposed as a co-transcriptional event based on the ability of NXF1 to co-IP with RNAPII, and more importantly, on increased RNAPII pausing at PASs in response to NXF1 depletion. Unfortunately, it has not been determined whether NXF1 depletion selectively induced RNAPII pausing at the proximal PASs, as with SCAF4/SCAF8 (Gregersen et al., 2019) and CTCF (Nanavaty et al., 2020). Alternatively, NXF1 deficiency is known to induce R-loop formation, which is tied to RNAPII pausing (Niehrs and Luke, 2020; Santos-Pereira and Aguilera, 2015). This may account for increased RNAPII near PASs, as observed in NXF1 knockdown cells (Chen et al., 2019). Impaired NXF1 could detain distal isoforms in the nucleus, thus providing precursors for sequential CPA at proximal PASs (Fig. 3B). Future work will test this alternative

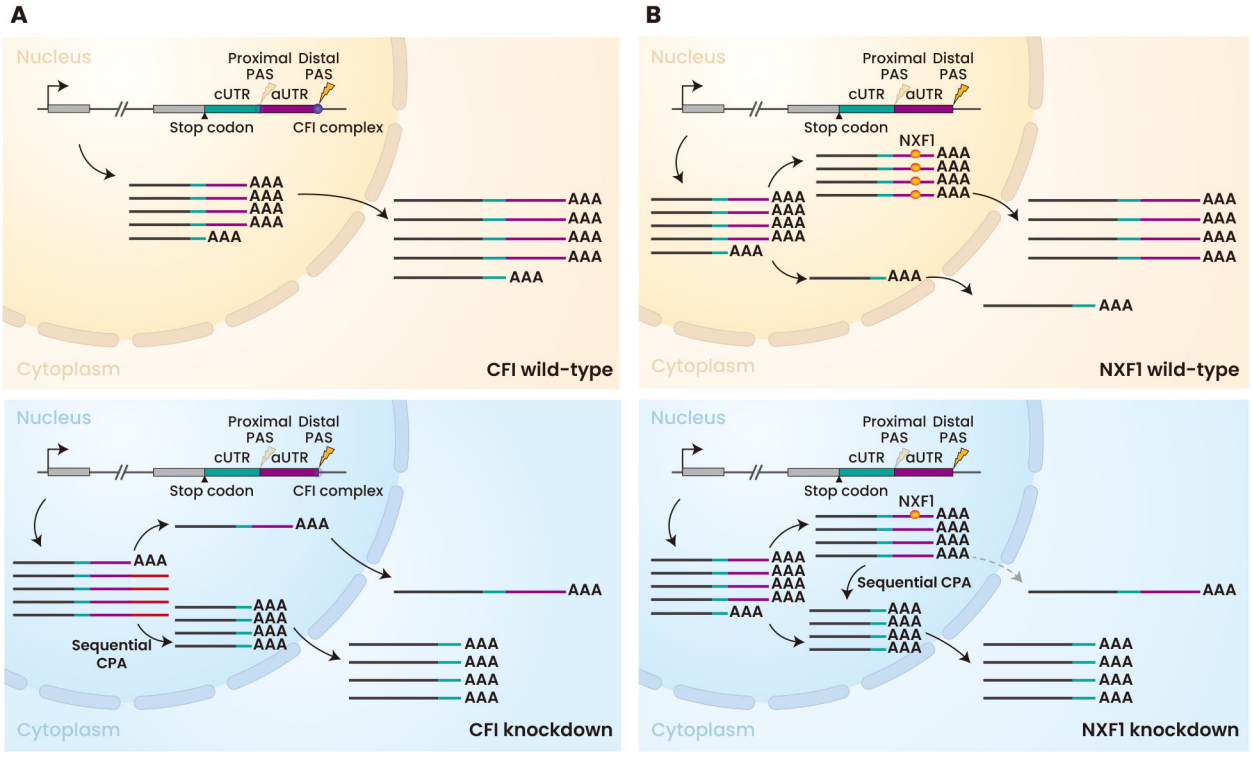

Fig. 3. The sequential model cast new light on documented alternative polyadenylation (APA) regulatory paradigms. (A) New insights on molecular mechanism for cleavage factor I (CFI)-mediated regulation on APA. CFI shows a strong binding preference for distal poly(A) site (PAS) compared to proximal PAS, facilitating the production of mRNA isoform with longer 3′UTR at the co-transcriptional level (top). Depletion of CFI causes transcriptional readthrough (red lines), generating aberrant transcripts which could be detained at the nucleus and provide the intermediates for sequential cleavage and polyadenylation (CPA) at the proximal sites (bottom). (B) The alternative APA regulatory mechanism by knockdown NXF1. NXF1 plays an important role in exporting mRNA isoforms with longer 3′UTR from the nucleus to the cytoplasm. Depletion of NXF1 leads to detain of longer isoforms at the nucleus, which then served as the precursor for sequential CPA, resulting in producing of more short isoforms.

interpretation on NXF1 deficiency-induced APA switch.

One may argue that many components of different types of RNA processing machinery, including the poly(A) machinery, are associated with RNAPII, which has been considered as key evidence for co-transcriptional RNA processing. In these processes, the most prevailing model is the loading of various RNA processing factors to RNAPII CTD. However, to date, no studies directly tested the functional requirement of RNAPII CTD for the observed interactions in cells. Given the ability of RNAPII CTD to form condensates and nascent RNAs likely play key roles in such condensate formation to create transcription hubs where transcription and RNA processing are functionally integrated (Boehning et al., 2018; Guo et al., 2019). In this regard, the CTD of RNAPII may play a general role as a key organizer of such gene expression hubs, rather than directly involved in binding and loading specific RNA processing factors and such hubs may not only enable co-transcriptional RNA processing but also provide tight links between transcription and post-transcriptional RNA processing events. Thus, co-IP with RNAPII does not necessarily reflect co-transcriptional RNA processing.

CONCLUDING REMARKS

APA is a key strategy to regulate gene expression by creating mRNA isoforms from single genes that have the potential to be selectively regulated at the levels of RNA stability, export, location, and translation. Because of tight link of the CPA reaction to transcription termination, it has been generally thought that most regulated APA events also take place in a co-transcriptional fashion. However, a recent study suggests that the CPA reaction at the distal PAS is co-transcriptional, but the CPA reaction at the alternative proximal PAS(s) can be either co- or post-transcriptional. If post-transcriptional, the proposed sequential polyadenylation model provides a new framework to understand specific regulatory mechanisms. It is our hope that this review will serve as a catalyst to further enrich this new conceptual framework.

ACKNOWLEDGMENTS

Y.H., T.C., C.L., and X.Z. were supported by NIH grants HG004659 and DK098808.

AUTHOR CONTRIBUTIONS

Y.H. and X.-D.F. wrote the manuscript. T.C., C.L., and X.Z. gave valuable inputs for improving manuscript modification.

CONFLICT OF INTEREST

The authors have no potential conflicts of interest to disclose.

ORCID

Yajing Hao https://orcid.org/0000-0003-1384-4176
Ting Cai https://orcid.org/0000-0003-3064-1733
Chang Liu https://orcid.org/0000-0001-5219-1089
Xuan Zhang https://orcid.org/0000-0002-3761-3789
Xiang-Dong Fu https://orcid.org/0000-0001-5499-8732

REFERENCES

Beaudoin, J.D. and Perreault, J.P. (2013). Exploring mRNA 3'-UTR G-quadruplexes: evidence of roles in both alternative polyadenylation and mRNA shortening. Nucleic Acids Res. *41*, 5898-5911.

Boehning, M., Dugast-Darzacq, C., Rankovic, M., Hansen, A.S., Yu, T., Marie-Nelly, H., McSwiggen, D.T., Kokic, G., Dailey, G.M., Cramer, P., et al. (2018). RNA polymerase II clustering through carboxy-terminal domain phase separation. Nat. Struct. Mol. Biol. *25*, 833-840.

Chen, S., Wang, R., Zheng, D., Zhang, H., Chang, X., Wang, K., Li, W., Fan, J., Tian, B., and Cheng, H. (2019). The mRNA export receptor NXF1 coordinates transcriptional dynamics, alternative polyadenylation, and mRNA export. Mol. Cell *74*, 118-131.e7.

Cheng, L.C., Zheng, D., Baljinnyam, E., Sun, F., Ogami, K., Yeung, P.L., Hoque, M., Lu, C.W., Manley, J.L., and Tian, B. (2020). Widespread transcript shortening through alternative polyadenylation in secretory cell differentiation. Nat. Commun. *11*, 3182.

Connelly, S. and Manley, J.L. (1988). A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. Genes Dev. *2*, 440-452.

Dantonel, J.C., Murthy, K.G., Manley, J.L., and Tora, L. (1997). Transcription factor TFIID recruits factor CPSF for formation of 3' end of mRNA. Nature *389*, 399-402.

Derti, A., Garrett-Engele, P., Macisaac, K.D., Stevens, R.C., Sriram, S., Chen, R., Rohl, C.A., Johnson, J.M., and Babak, T. (2012). A quantitative atlas of polyadenylation in five mammals. Genome Res. *22*, 1173-1183.

Djebali, S., Davis, C.A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., et al. (2012). Landscape of transcription in human cells. Nature *489*, 101-108.

Eaton, J.D., Francis, L., Davidson, L., and West, S. (2020). A unified allosteric/torpedo mechanism for transcriptional termination on human protein-coding genes. Genes Dev. *34*, 132-145.

Elkon, R., Ugalde, A.P., and Agami, R. (2013). Alternative cleavage and polyadenylation: extent, regulation and function. Nat. Rev. Genet. *14*, 496-506.

Glover-Cutter, K., Kim, S., Espinosa, J., and Bentley, D.L. (2008). RNA polymerase II pauses and associates with pre-mRNA processing factors at both ends of genes. Nat. Struct. Mol. Biol. *15*, 71-78.

Gregersen, L.H., Mitter, R., Ugalde, A.P., Nojima, T., Proudfoot, N.J., Agami, R., Stewart, A., and Svejstrup, J.Q. (2019). SCAF4 and SCAF8, mRNA anti-terminator proteins. Cell *177*, 1797-1813.e18.

Gromak, N., West, S., and Proudfoot, N.J. (2006). Pause sites promote transcriptional termination of mammalian RNA polymerase II. Mol. Cell. Biol. *26*, 3986-3996.

Gruber, A.J. and Zavolan, M. (2019). Alternative cleavage and polyadenylation in health and disease. Nat. Rev. Genet. *20*, 599-614.

Guo, Y.E., Manteiga, J.C., Henninger, J.E., Sabari, B.R., Dall'Agnese, A., Hannett, N.M., Spille, J.H., Afeyan, L.K., Zamudio, A.V., Shrinivas, K., et al. (2019). Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. Nature *572*, 543-548.

Han, J., Xiong, J., Wang, D., and Fu, X.D. (2011). Pre-mRNA splicing: where and when in the nucleus. Trends Cell Biol. *21*, 336-343.

Ji, Z. and 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.

Kaczmarek Michaels, K., Mohd Mostafa, S., Ruiz Capella, J., and Moore, C.L. (2020). Regulation of alternative polyadenylation in the yeast Saccharomyces cerevisiae by histone H3K4 and H3K36 methyltransferases. Nucleic Acids Res. *48*, 5407-5425.

Kuhn, U., Gundel, M., Knoth, A., Kerwitz, Y., Rudel, S., and Wahle, E. (2009). Poly(A) tail length is controlled by the nuclear poly(A)-binding protein regulating the interaction between poly(A) polymerase and the cleavage and polyadenylation specificity factor. J. Biol. Chem. *284*, 22803-22814.

Kwon, B., Fansler, M.M., Patel, N.D., Lee, J., Ma, W., and Mayr, C. (2022). Enhancers regulate 3' end processing activity to control expression of alternative 3'UTR isoforms. Nat. Commun. *13*, 2709.

Li, W., You, B., Hoque, M., Zheng, D., Luo, W., Ji, Z., Park, J.Y., Gunderson, S.I., Kalsotra, A., Manley, J.L., et al. (2015). Systematic profiling of poly(A)+ transcripts modulated by core 3' end processing and splicing factors reveals regulatory rules of alternative cleavage and polyadenylation. PLoS Genet. *11*, e1005166.

Lin, J., Hung, F.Y., Ye, C., Hong, L., Shih, Y.H., Wu, K., and Li, Q.Q. (2020). HDA6-dependent histone deacetylation regulates mRNA polyadenylation in Arabidopsis. Genome Res. *30*, 1407-1417.

Liu, X., Freitas, J., Zheng, D., Oliveira, M.S., Hoque, M., Martins, T., Henriques, T., Tian, B., and Moreira, A. (2017). Transcription elongation rate has a tissue-specific impact on alternative cleavage and polyadenylation in Drosophila melanogaster. RNA *23*, 1807-1816.

Mayr, C. and Bartel, D.P. (2009). Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. Cell *138*, 673-684.

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

Mitschka, S. and Mayr, C. (2022). Context-specific regulation and function of mRNA alternative polyadenylation. Nat. Rev. Mol. Cell Biol. *23*, 779-796.

Nagaike, T., Logan, C., Hotta, I., Rozenblatt-Rosen, O., Meyerson, M., and Manley, J.L. (2011). Transcriptional activators enhance polyadenylation of mRNA precursors. Mol. Cell *41*, 409-418.

Nanavaty, V., Abrash, E.W., Hong, C., Park, S., Fink, E.E., Li, Z., Sweet, T.J., Bhasin, J.M., Singuri, S., Lee, B.H., et al. (2020). DNA methylation regulates alternative polyadenylation via CTCF and the cohesin complex. Mol. Cell *78*, 752-764.e6.

Neve, J., Burger, K., Li, W., Hoque, M., Patel, R., Tian, B., Gullerova, M., and Furger, A. (2016). Subcellular RNA profiling links splicing and nuclear DICER1 to alternative cleavage and polyadenylation. Genome Res. *26*, 24-35.

Niehrs, C. and Luke, B. (2020). Regulatory R-loops as facilitators of gene expression and genome stability. Nat. Rev. Mol. Cell Biol. *21*, 167-178.

Pinto, P.A., Henriques, T., Freitas, M.O., Martins, T., Domingues, R.G., Wyrzykowska, P.S., Coelho, P.A., Carmo, A.M., Sunkel, C.E., Proudfoot, N.J., et al. (2011). RNA polymerase II kinetics in polo polyadenylation signal selection. EMBO J. *30*, 2431-2444.

Proudfoot, N. (2004). New perspectives on connecting messenger RNA 3' end formation to transcription. Curr. Opin. Cell Biol. *16*, 272-278.

Proudfoot, N.J. (2016). Transcriptional termination in mammals: stopping

the RNA polymerase II juggernaut. Science *352*, aad9926.

Santos-Pereira, J.M. and Aguilera, A. (2015). R loops: new modulators of genome dynamics and function. Nat. Rev. Genet. *16*, 583-597.

Shi, Y., Di Giammartino, D.C., Taylor, D., Sarkeshik, A., Rice, W.J., Yates, J.R., 3rd, Frank, J., and Manley, J.L. (2009). Molecular architecture of the human pre-mRNA 3' processing complex. Mol. Cell *33*, 365-376.

Shi, Y. and Manley, J.L. (2015). The end of the message: multiple protein-RNA interactions define the mRNA polyadenylation site. Genes Dev. *29*, 889-897.

Skourti-Stathaki, K., Proudfoot, N.J., and Gromak, N. (2011). Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. Mol. Cell *42*, 794-805.

Soles, L.V. and Shi, Y. (2021). Crosstalk between mRNA 3'-end processing and epigenetics. Front. Genet. *12*, 637705.

Tang, P., Yang, Y., Li, G., Huang, L., Wen, M., Ruan, W., Guo, X., Zhang, C., Zuo, X., Luo, D., et al. (2022). Alternative polyadenylation by sequential activation of distal and proximal PolyA sites. Nat. Struct. Mol. Biol. *29*, 21-31.

Tian, B. and Manley, J.L. (2017). Alternative polyadenylation of mRNA precursors. Nat. Rev. Mol. Cell Biol. *18*, 18-30.

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

Wood, A.J., Schulz, R., Woodfine, K., Koltowska, K., Beechey, C.V., Peters, J., Bourc'his, D., and Oakey, R.J. (2008). Regulation of alternative polyadenylation by genomic imprinting. Genes Dev. *22*, 1141-1146.

Xia, Z., Donehower, L.A., Cooper, T.A., Neilson, J.R., Wheeler, D.A., Wagner, E.J., and Li, W. (2014). Dynamic analyses of alternative polyadenylation from RNA-seq reveal a 3'-UTR landscape across seven tumour types. Nat. Commun. *5*, 5274.

Xiao, R., Chen, J.Y., Liang, Z., Luo, D., Chen, G., Lu, Z.J., Chen, Y., Zhou, B., Li, H., Du, X., et al. (2019). Pervasive chromatin-RNA binding protein interactions enable RNA-based regulation of transcription. Cell *178*, 107-121.e18.

Yonaha, M. and Proudfoot, N.J. (1999). Specific transcriptional pausing activates polyadenylation in a coupled in vitro system. Mol. Cell *3*, 593-600.

Yuryev, A., Patturajan, M., Litingtung, Y., Joshi, R.V., Gentile, C., Gebara, M., and Corden, J.L. (1996). The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins. Proc. Natl. Acad. Sci. U. S. A. *93*, 6975-6980.

Zhang, H., Rigo, F., and Martinson, H.G. (2015). Poly(A) signal-dependent transcription termination occurs through a conformational change mechanism that does not require cleavage at the poly(A) site. Mol. Cell *59*, 437-448.

Zhou, Y., Li, H.R., Huang, J., Jin, G., and Fu, X.D. (2014). Multiplex analysis of polyA-linked sequences (MAPS): an RNA-seq strategy to profile poly(A+) RNA. Methods Mol. Biol. *1125*, 169-178.

Zhu, Y., Wang, X., Forouzmand, E., Jeong, J., Qiao, F., Sowd, G.A., Engelman, A.N., Xie, X., Hertel, K.J., and Shi, Y. (2018). Molecular mechanisms for CFIm-mediated regulation of mRNA alternative polyadenylation. Mol. Cell *69*, 62-74.e4.