BIOCELL
2022 46(6): 1453-1457

Tech Science Press

How to make the end of a gene, the simple way

Karel H. M. van WELY*

Centro Nacional de Biotecnología/CSIC, Madrid, 28049, Spain

Key words: RNA transcription, RNA polymerase II pausing, Splicing, Termination, Polyadenylation

Abstract: Transcription termination of nearly all protein-coding genes in mammals requires 3' end processing by a multiprotein complex that will cleave and polyadenylate the messenger RNA precursor. Because a variety of enzyme complexes intervene, 3' end processing was thought to be fundamentally complex and subject to a multitude of regulatory effects. The possibility to select just one out of several polyadenylation sites, in particular, has caused much questioning and speculation. What appear to be separate mechanisms however can be combined into a defined set of rules, allowing for a relatively simple interpretation of 3' end processing. Ultimately, readiness of the terminal exon splice site determines when a transcript reaches the maturity to select a nearby polyadenylation signal. Transcriptional pausing then acts in concert, extending the timeframe during which the transcription complex is close to polyadenylation sites. Since RNA polymerase pausing is governed by the same type of sequences in bacteria and metazoans, mammalian transcription termination resembles its prokaryote counterpart more than generally thought.

Introduction

Most genes in mammals have evolved into large transcription units, bearing relatively small protein-coding exons separated by very large introns. These genes span 10 to 15 kilobases on average, but occasionally reach a size of more than 1000 kilobases. To transcribe these long templates, RNA polymerase II is among the most processive enzyme complexes known. Pol II on occasion has been compared to a juggernaut, a practically unstoppable massive machine (Proudfoot, 2016). Individual genes nonetheless must be functionally separated, which implies that transcription termination at the 3' end is strictly controlled and should be coordinated with other aspects of messenger RNA production.

Transcription termination in bacteria is a comparatively simple process, in which a stretch of uridines provokes pausing of the RNA polymerase. Backtracking is prevented by a hairpin structure, leading to the dissolution of the polymerase and liberation of the messenger RNA. The same process in mammals generally requires a combination of enzyme complexes, in particular for protein-coding genes transcribed by RNA Polymerase II (Pol II). Nearly all protein-coding transcripts in mammals contain a polyadenylated tail, which has essential functions in nuclear export, translation, and stability. Notwithstanding its importance, the poly(A) tail of a mature messenger is not directly encoded in the genome. Instead of being transcribed by Pol II, the poly(A) tail is synthesized by a dedicated enzyme complex, to which the immature messenger is transferred at the end of transcription (Humphrey *et al.*, 1987).

The core signal that is responsible for cleavage and polyadenylation (AAUAAA) is present in our genome in many more copies than transcripts are formed. Thus, additional signals must control polyadenylation during RNA maturation *in vivo*. Exactly how the different enzyme complexes cooperate to control polyadenylation of immature messengers has been the topic of numerous studies, but even today remains a matter of debate. What has become clear, however, is the requirement to temporarily stall transcription in order to achieve polyadenylation (Proudfoot, 2016). Recent advances have allowed for genome-wide analysis of *in vivo* transcription, but the individual sources of these data have not yet distilled a unified mechanism. Since each technique exposes just one aspect of polyadenylation without providing a complete picture, we will test the compatibility between the underlying mechanisms in this viewpoint.

In order to understand the role of RNA Pol II in transcription termination, several studies have looked at the complexes responsible for mRNA synthesis. In agreement with the proposed requirements for 3' end processing, RNA and chromatin immunoprecipitation (CHIP) sequencing have shown frequent Pol II pausing in terminal exons and downstream of genes (Carrillo Oesterreich *et al.*, 2010; Cortazar *et al.*, 2019). Pol II pausing is considered extremely

*Address correspondence to: Karel H. M. van Wely, kvanwely@cnb.csic.es
Received: 25 August 2021; Accepted: 24 November 2021

Doi: 10.32604/biocell.2022.018939

important, as it gives time to recruit enzyme complexes not directly associated with the transcription holoenzyme. While the recognition of Pol II pausing at the 3' end of genes has been important for enumerating the prerequisites for polyadenylation, it left some key issues unanswered. One of the open questions concerns the mechanisms that may regulate Pol II pausing at the 3' end of genes. In conclusion, the reasons why and where Pol II pauses might been answered but the way how it does so still need explaining.

Pol II Pausing in Terminal Introns

Although Pol II pausing in terminal exons has been linked to 3' end processing itself, it likely performs a much wider function. A striking example can be found in mutants of *DIDO* (*Death Inducer Obliterator*), a gene producing several proteins that connect Pol II to the splicing apparatus (Kinkelin *et al.*, 2013; Mora Gallardo *et al.*, 2019). *DIDO* mutants not only show gross defects in exon inclusion, but also readthrough at the 3' end of genes and alternative polyadenylation (Mora Gallardo *et al.*, 2021). Importantly, genes in which Pol II pausing around terminal exons is more pronounced, are less sensitive to *DIDO* mutation. These recent results highlight the intimate connection between splicing and 3' end processing; through pausing, upstream RNA splicing has more time to complete, promoting overall messenger maturity before polyadenylation and export. Recent studies suggest that messenger maturity is controlled by a checkpoint that involves the upstream splice sites (Leader *et al.*, 2021). Binding of the U1 spliceosome subunit to the 5' splice site (SS) suppresses polyadenylation some distance downstream, a phenomenon termed telescripting (Berg *et al.*, 2012).

By suppressing the processing of internal fortuitous polyadenylation sites (PAS) located mostly in introns, telescripting prevents premature polyadenylation of partially transcribed genes. The 5' SS and U1 spliceosome subunit remain attached to Pol II while the intron is transcribed (Leader *et al.*, 2021), but are ejected from the RNA as soon as the spliceosome assembles after 3' SS recognition by U2 (Lee and Rio, 2015). The binding and ejection of U1 thus provide a simple mechanism to promote elimination of the terminal intron before activation of the PAS which defines the terminal exon; pausing in this region promotes U2 binding (Mora Gallardo *et al.*, 2019) and subsequent processing. *In vitro*, destruction of the 3' SS in the presence of an 5' SS suppresses polyadenylation (Rigo and Martinson, 2008). *In vivo*, reducing the capacity to recruit splicing factors important for 3' SS processing has exactly the same effect (Mora Gallardo *et al.*, 2021). In conclusion, suppressing U1 and the 5' SS leads to premature polyadenylation (Oh *et al.*, 2017) while reducing the contribution of U2 and the 3' SS to splicing causes transcriptional readthrough (Mora Gallardo *et al.*, 2021). Thus, the frequently observed Pol II pausing in terminal exons (Carrillo Oesterreich *et al.*, 2010) is not just a curiosity but provides time to assemble the U1-U2 interface and correctly define the upstream intron as terminal. Without removal of the terminal intron and ejection of U1 from the nascent RNA, telescripting suppresses processing of transcribed PAS (Fig. 1). Noteworthy, polyadenylation of genes containing a single exon is not restricted by telescripting, because neither 5' SS nor 3' SS are present.

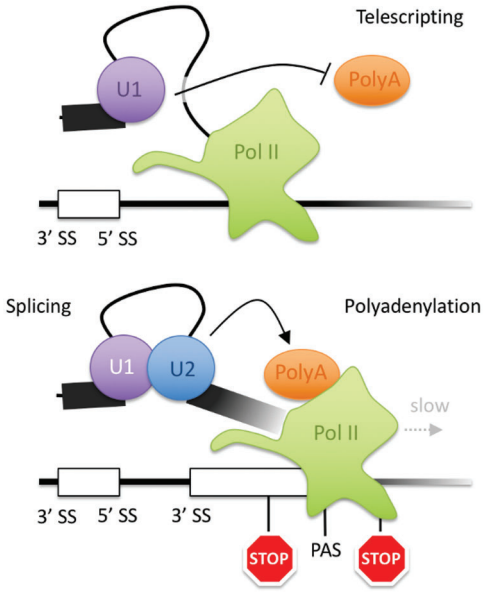

FIGURE 1. Splicing efficiency of the terminal exon determines polyadenylation site (PAS) selection. As long as the terminal 3' splice site (SS) is not recognized, telescripting of the upstream 5' SS may suppress polyadenylation. Note that the terminal intron in mammals typically is much bigger than the terminal exon, and that the U1 spliceosome subunit is present in more copies than U2. Only when the 3' SS is recognized by U2, terminal exon readiness is achieved. Pol II pausing, typically on TRSM, extends the time available for U2, temporarily retaining Pol II downstream of the 3' SS and allowing selection of nearby PAS.

Pol II Pausing Downstream of Genes

In addition to RNA maturation, Pol II pausing has a more direct role in 3' end processing, by preventing unnecessary synthesis of downstream RNA after the mature messenger is severed off (Cortazar *et al.*, 2019). In this "sitting duck" model, Pol II pausing occurs some distance downstream of the actual PAS. The stretch of RNA between the PAS and the stalled Pol II can then be digested by the XRN2 exonuclease (Rat1p in yeast), dislodging the holoenzyme from the template DNA and finally resolving the transcription complex. Indeed, a shift in the balance between pausing and 3' end processing not only promotes the use of PAS located further downstream but also the formation of RNA fusions between tandem genes. The absence of intergenic pause sites in short intergenic regions provide ideal conditions for Pol II to continue transcribing across multiple PAS (Mora Gallardo *et al.*, 2021). Accordingly, the majority of natural fusion products consist of closely spaced tandem genes.

Stopping the Unstoppable Machine

One of the most studied ways in which Pol II speed can be modulated is the nucleosome around which the template DNA is wrapped. Due to its close contact with the DNA, the nucleosome is thought to form a physical barrier for transcription (Chiu *et al.*, 2018). Accordingly, the average size of internal exons matches the number of bases than can be wrapped around a single nucleosome. In addition, internal exons are marked by specific epigenetic modifications, several of which are associated with RNA splicing (Tilgner and Guigo, 2010; Leung *et al.*, 2019). In

contrast to internal exons, terminal exons do not adhere to the nucleosome-associated length restrictions. Although some evidence for a positioned nucleosome close to the 3' acceptor site of terminal exons has been found, this occupancy is lower than in internal exons. Finally, the pausing in terminal exons normally occurs much further downstream than the 3' SS, typically some 250 base-pairs upstream of a canonical PAS (Carrillo Oesterreich *et al.*, 2010). Together with the finding that pausing also occurs downstream of the actual gene, inside the so-called "junk DNA" (Cortazar *et al.*, 2019; Mora Gallardo *et al.*, 2021), the distribution towards the 3' end of terminal exons indicates that other mechanisms contribute to Pol II pausing.

In addition to physical barriers, transcription speed is influenced by sequence composition (Zamft *et al.*, 2012). Transcription by RNA polymerases in general is slow across AT-rich templates as compared to GC-rich templates. The underlying mechanism is not fully understood, but likely depends on the hybridization strength of the bases in the RNA-DNA hybrid and possible structures formed by the nascent RNA (Palangat *et al.*, 2004). Massive sequencing of native RNA showed that *in vivo*, too, sequence composition accounts for the majority of Pol II pause sites (Mora Gallardo *et al.*, 2021). The latter study showed that a set of interrelated Thymidine-rich sequence motifs (TRSM) are responsible for Pol II pausing. The same set of TSRM, when located at the 3' end of genes, favored the use of nearby PAS over alternative sites. TRSM frequency in metazoans increases from the 5' to 3' end of terminal exons and is especially high downstream of canonical PAS. When considering the multiple roles of pausing, TRSM distribution correlates better with termination than reported nucleosome position.

A single stretch of 6 Thymidine residues in the template strand efficiently increased both Pol II pausing frequency and duration in model transcripts (Palangat *et al.*, 2004). Oligo-thymidine stretches, in collaboration with an upstream hairpin, also act as terminators in bacteria (Proudfoot, 2016). In the NET-seq data, TRSM bearing other bases were identified, but a typical motif contained at least 80% Thymidine. Not all TRSM seemed to produce long lasting pauses, as the same data revealed numerous closely spaced tandem peaks. Thus, transcriptional pausing seems to be governed by gradual changes as the template is read, rather than a strict on-off mechanism. Importantly, 3' end processing entails significant changes in Pol II conformation (Proudfoot, 2016). Although binding of external factors such as XRN2 might contribute, TRSM-dependent pausing itself is enough to produce an allosteric change in Pol II (Palangat *et al.*, 2004). Comparable to bacteria, mammalian transcription *in vitro* can be terminated by a conformational change alone, without PAS cleavage or degradation of excess RNA (Zhang *et al.*, 2015). The properties of termination *in vivo* suggests that pausing-associated conformational changes contribute *in vivo*, too.

Cause and Consequence

While different lines of evidence indicate a direct link between template sequence composition and Pol II pausing, additional properties of transcription likely contribute to termination *in vivo*. First of all, a central role has been attributed to the Pol II carboxyterminal domain (CTD). The Pol II CTD is a heptad repeat, which is conserved among all eukaryotes and differentially modified during the various stages of transcription. The CTD has been shown to recruit a variety of RNA processing factors, including proteins involved in splicing —albeit not SPFQ— and transcription termination (Fong and Bentley, 2001). The modified CTD promotes recruitment of a variety of RNA processing factors, so disruption of this interaction understandably impairs transcription termination and polyadenylation. Likewise, depletion of protein complexes required for RNA cleavage at termination sites or addition of poly(A) tails directly impact production of a mature messenger. Effects on RNA maturation however can be interpreted in several ways, and do not automatically mean that Pol II transcription itself is modulated. Because Pol II pausing is a kinetic process governed by efficiency and half-life, reduced recruitment means that pausing simply is too short to promote assembly of a functional termination complex. Accordingly, directed mutation of the CTD can be used to separate termination and pausing (Fusby *et al.*, 2016).

The CTD is made up of multiple heptad repeats, and has approximately doubled in size in vertebrates with respect to yeast. Since this size doubling lacks any increase in complexity, that vertebrate evolution seems to respond to a capacity problem instead of a more intricate regulation. Moreover, the CTD comprises an unstructured flexible domain, making it hard to envision how this region could transmit allosteric pausing signals to the Pol II core. Accordingly, 3' end processing factors can be immunoprecipitated with Pol II lacking the CTD (Nag *et al.*, 2007); direct modulation of the Pol II body by the polyadenylation apparatus thus may contribute significantly to pausing. Combining these data into a model, the CTD seems to recruit RNA processing factors which in turn act on the Pol II body in a feedback loop. Transcriptional pausing will expose the flexible CTD to a confined genomic and nuclear domain, allowing for accretion of locally required processing factors. The combination of CTD modifications make it a structure ideally suited to respond to the basic properties of the Pol II holoenzyme, but only indirectly active in the control over transcription.

In addition to the Pol II CTD, epigenetic modifications of histone tails have been suggested to control transcription. For example, trimethylation of lysine 4 of histone H3 is particularly abundant on the 5' end of genes that are actively transcribed by Pol II. This epigenetic mark mostly associates with the boundary of first exon and intron, precisely the region where many *processing* factors first associate with Pol II (Bieberstein *et al.*, 2012; Mora Gallardo *et al.*, 2019). H3K4Me3 distribution and abundance suggest that this mark may function as a bulk reservoir for proteins that are required along the gene body, one could say a kind of "pick-up zone". One of the benefits of promotor-proximal pausing, a near-universal feature of Pol II transcription, would be a prolonged stay in this zone and optimal recruitment of processing factors. Another histone H3 modification, H3K36Me3, is believed to regulate RNA processing further towards the 3' end of genes. H3K36Me3 may function similarly to H3K4Me3, through the binding of processing factors (in this case EAF3) close to the site where they are needed (Leung *et al.*, 2019). Even though manipulation of H3K36Me3 content has an effect on polyadenylation, its

prime target is the RNA splicing machinery just as the DIDO-splicing axis. Most importantly, histone H3 trimethylations are deposited by enzymes such as SETD1 and SETD2, which in turn are recruited by the Pol II CTD. This means that local Pol II pausing might be a strong signal for CTD-dependent SETD2 recruitment and subsequent histone H3 trimethylation. Histone H3 trimethylation thus is the ultimate consequence of transcriptional pausing, providing a local memory mark used during subsequent passes of Pol II. Likewise, higher order chromatin structures such as gene loops have been attributed a memory function, aiding recycling of processing factors between 5' and 3' gene ends (Hampsey *et al.*, 2011). In the cascade of transcriptional signals, TRSM indeed could control distribution of epigenetic marks along the gene body, whereas the opposite —DNA sequence acquisition through epigenetic modifications— is less likely.

Discussion

Recent advancements in sequencing technology enabled the genome-wide analysis of various aspects of transcription. Among other findings, these advancements have indicated that RNA splicing is fast for most introns and that transcription slows down at the 3' end of many genes. Although massive sequencing has generated a wealth of data, this has not directly lead to deeper mechanistic understanding. Since most of the massive sequencing techniques focused on the transcription holoenzyme itself, a unilateral view was created. Through this single focus, the role of Pol II easily could be overestimated. The mutation of *DIDO*, a gene whose products link splicing to transcription, offered a much needed different angle (Mora Gallardo *et al.*, 2019). Deletion of most splicing factors is incompatible with cell survival and leads to incomplete transcription of longer genes (Takeuchi *et al.*, 2018). *DIDO* mutation however causes a redistribution of splicing efficiencies due to differential recruitment of the general splicing facilitator SFPQ. In these mutants, splicing of a subset of exons in fact is enhanced, because reduced availability in other exons liberates SFPQ from a stable but limited total pool. In natural tissues, the relative expression of DIDO isoforms likely determines subnuclear SFPQ distribution, because some but not all DIDO isoforms catalyze SFPQ recruitment. The use of *DIDO* mutants has shown the importance of splicing in the genome-wide regulation of alternative poly-adenylation; depending on the time needed for activation of the upstream 3' SS, Pol II transcription may proceed less or more downstream before a nearby PAS is processed. Thus, in genes where the terminal 3' SS is weak, the chance of using downstream PAS is higher. Vice versa, an efficient upstream 3' SS means that Pol II proceeds only briefly before the terminal exon acquires polyadenylation readiness. TRSM-induced pausing in terminal exons, even upstream of PAS, may thus promote polyadenylation by extending the time window during which the terminal intron 3' SS is processed. Modulation of the terminal exon 3' SS by splicing factors thus provides a flexible mechanism for PAS selection.

The second observation that can be deduced from the *DIDO* mutants is how pausing is regulated. Just as 3' SS, the PAS that are processed most efficiently are flanked by Pol II pause sites (Mora Gallardo *et al.*, 2021). A combination of various techniques has shown that both pausing and termination are promoted by relatively short (6 to 20 base-pairs) thymidine-rich motifs. One of the most striking conclusions is the evolutionary conservation of transcription termination, induced by the same sequences in organisms ranging from bacteria to mammals. Prokaryote transcription termination occurs at sites containing a hairpin and an oligo-uridine stretch. While no evidence for a hairpin has been uncovered yet in mammals, the TRSM that are transcribed into oligo-uridine appear frequently in noncoding regions of metazoan genomes, including the terminal exon and directly downstream of genes. Metazoan 3' end processing factors, involved in cleavage, polyadenylation, and digestion of residual RNA, may have taken over the role of the prokaryote hairpin (Fig. 2). The contribution of TRSM *in vivo* thus shows that transcription termination in mammals resembles the prokaryote mechanism more closely than generally thought. When considered carefully, archaic TRSM define a gene end just as much as the evolutionary recent PAS. Ultimately, the relative placing of 3' SS, PAS, and TRSM determines if paused Pol II, through its CTD, can pair with the cleavage and polyadenylation machinery and thereby achieve successful termination.

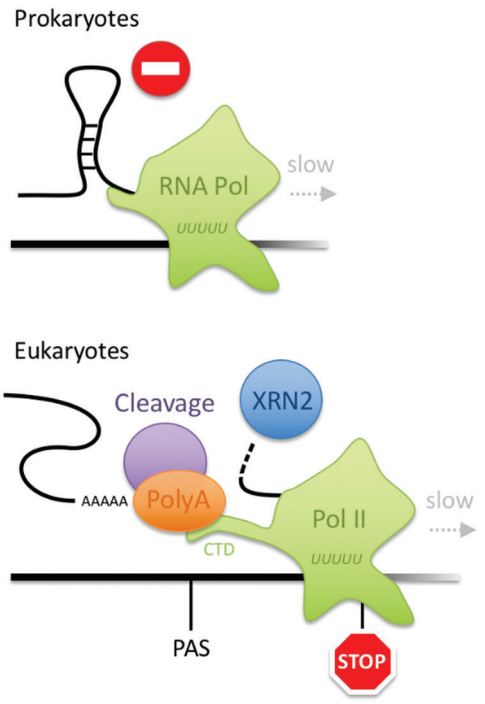

FIGURE 2. Comparison of basic 3' end processing in prokaryotes and eukaryotes. While a hairpin prevents backtracking in prokaryote terminators, its role has been taken over by a combination of enzyme complexes in eukaryotes. The fundamental sequence requirements for transcriptional pausing however have remained the same throughout evolution. Note that modifications of the flexible Pol II carboy-terminus (CTD) and histone modifications might promote RNA processing without modifying transcription speed, but could be a consequence of intrinsic pausing close to polyadenylation sites (PAS).

Author Contribution: The author confirms sole responsibility for the following: Study conception and design, data collection, analysis and interpretation of results, and manuscript preparation.

Funding Statement: The author is financed in part by the Spanish Government (Ministerio de Ciencia e Innovación, Grant MCIN/AEI PID2019-110574RB-I00).

Conflicts of Interest: The author declares that they have no conflicts of interest to report regarding the present study.

References

Berg MG, Singh LN, Younis I, Liu Q, Pinto AM et al. (2012). U1 snRNP determines mRNA length and regulates isoform expression. *Cell* **150**: 53–64. DOI 10.1016/j.cell.2012.05.029.

Bieberstein NI, Carrillo Oesterreich F, Straube K, Neugebauer KM (2012). First exon length controls active chromatin signatures and transcription. *Cell Reports* **2**: 62–68. DOI 10.1016/j.celrep.2012.05.019.

Carrillo Oesterreich F, Preibisch S, Neugebauer KM (2010). Global analysis of nascent RNA reveals transcriptional pausing in terminal exons. *Molecular Cell* **40**: 571–581. DOI 10.1016/j.molcel.2010.11.004.

Chiu AC, Suzuki HI, Wu X, Mahat DB, Kriz AJ, Sharp PA (2018). Transcriptional pause sites delineate stable nucleosome-associated premature polyadenylation suppressed by U1 snRNP. *Molecular Cell* **69**: 648–663. DOI 10.1016/j.molcel.2018.01.006.

Cortazar MA, Sheridan RM, Erickson B, Fong N, Glover-Cutter K et al. (2019). Control of RNA Pol II speed by PNUTS-PP1 and Spt5 dephosphorylation facilitates termination by a sitting duck torpedo mechanism. *Molecular Cell* **76**: 896–908. DOI 10.1016/j.molcel.2019.09.031.

Fong N, Bentley DL (2001). Capping, splicing, and 3' processing are independently stimulated by RNA polymerase II: Different functions for different segments of the CTD. *Genes and Development* **15**: 1783–1795. DOI 10.1101/gad.889101.

Fusby B, Kim S, Erickson B, Kim H, Peterson ML, Bentley DL (2016). Coordination of RNA Polymerase II pausing and 3' end processing factor recruitment with alternative polyadenylation. *Molecular and Cellular Biology* **36**: 295–303. DOI 10.1128/MCB.00898-15.

Hampsey M, Singh BN, Ansari A, Laine JP, Krishnamurthy S (2011). Control of eukaryotic gene expression: Gene loops and transcriptional memory. *Advances in Enzyme Regulation* **51**: 118–125. DOI 10.1016/j.advenzreg.2010.10.001.

Humphrey T, Christofori G, Lucijanic V, Keller W (1987). Cleavage and polyadenylation of messenger RNA precursors *in vitro* occurs within large and specific 3' processing complexes. *EMBO Journal* **6**: 4159–4168. DOI 10.1002/j.1460-2075.1987.tb02762.x.

Kinkelin K, Wozniak GG, Rothbart SB, Lidschreiber M, Strahl BD et al. (2013). Structures of RNA polymerase II complexes with Bye1, a chromatin-binding PHF3/DIDO homologue. *Proceedings of the National Academy of Sciences of the United States of America* **110**: 15277–15282. DOI 10.1073/pnas.1311010110.

Leader Y, Lev Maor G, Sorek M, Shayevitch R, Hussein M et al. (2021). The upstream 5' splice site remains associated to the transcription machinery during intron synthesis. *Nature Communications* **12**: 4545. DOI 10.1038/s41467-021-24774-6.

Lee Y, Rio DC (2015). Mechanisms and regulation of alternative pre-mRNA splicing. *Annual Reviews in Biochemistry* **84**: 291–323. DOI 10.1146/annurev-biochem-060614-034316.

Leung CS, Douglass SM, Morselli M, Obusan MB, Pavlyukov MS et al. (2019). H3K36 methylation and the chromodomain protein Eaf3 are required for proper cotranscriptional spliceosome assembly. *Cell Reports* **27**: 3760–3769. DOI 10.1016/j.celrep.2019.05.100.

Mora Gallardo C, Sanchez de Diego A, Gutierrez Hernandez J, Talavera-Gutierrez A, Fischer T et al. (2019). Dido3-dependent SFPQ recruitment maintains efficiency in mammalian alternative splicing. *Nucleic Acids Research* **47**: 5381–5394. DOI 10.1093/nar/gkz235.

Mora Gallardo C, Sanchez de Diego A, Martinez AC, van Wely KHM (2021). Interplay between splicing and transcriptional pausing exerts genome-wide control over alternative polyadenylation. *Transcription* **12**: 55–71. DOI 10.1080/21541264.2021.1959244.

Nag A, Narsinh K, Martinson HG (2007). The poly(A)-dependent transcriptional pause is mediated by CPSF acting on the body of the polymerase. *Nature Structural and Molecular Biology* **14**: 662–669. DOI 10.1038/nsmb1253.

Oh JM, Di C, Venters CC, Guo J, Arai C et al. (2017). U1 snRNP telescripting regulates a size-function-stratified human genome. *Nature Structural and Molecular Biology* **24**: 993–999. DOI 10.1038/nsmb.3473.

Palangat M, Hittinger CT, Landick R (2004). Downstream DNA selectively affects a paused conformation of human RNA polymerase II. *Journal of Molecular Biology* **341**: 429–442. DOI 10.1016/j.jmb.2004.06.009.

Proudfoot NJ (2016). Transcriptional termination in mammals: Stopping the RNA polymerase II juggernaut. *Science* **352**: aad9926. DOI 10.1126/science.aad9926.

Rigo F, Martinson HG (2008). Functional coupling of last-intron splicing and 3'-end processing to transcription *in vitro*: The poly(A) signal couples to splicing before committing to cleavage. *Molecular and Cellular Biology* **28**: 849–862. DOI 10.1128/MCB.01410-07.

Takeuchi A, Iida K, Tsubota T, Hosokawa M, Denawa M et al. (2018). Loss of SFPQ causes long-gene transcriptopathy in the brain. *Cell Reports* **23**: 1326–1341. DOI 10.1016/j.celrep.2018.03.141.

Tilgner H, Guigo R (2010). From chromatin to splicing: RNA-processing as a total artwork. *Epigenetics* **5**: 180–184. DOI 10.4161/epi.5.3.11319.

Zamft B, Bintu L, Ishibashi T, Bustamante C (2012). Nascent RNA structure modulates the transcriptional dynamics of RNA polymerases. *Proceedings of the National Academy of Sciences of the United States of America* **109**: 8948–8953. DOI 10.1073/pnas.1205063109.

Zhang H, Rigo F, Martinson HG (2015). Poly(A) Signal-dependent transcription termination occurs through a conformational change mechanism that does not require cleavage at the poly(A) site. *Molecular Cell* **59**: 437–448. DOI 10.1016/j.molcel.2015.06.008.