AMERICAN SOCIETY FOR MICROBIOLOGY | Molecular and Cellular Biology

Coordination of RNA Polymerase II Pausing and 3′ End Processing Factor Recruitment with Alternative Polyadenylation

Becky Fusby,[a] Soojin Kim,[a] Benjamin Erickson,[a] Hyunmin Kim,[a] Martha L. Peterson,[b] David L. Bentley[a]

Department of Biochemistry and Molecular Genetics, University of Colorado School of Medicine, Aurora, Colorado, USA[a]; Department of Microbiology, Immunology & Molecular Genetics, University of Kentucky Medical Center, Lexington, Kentucky, USA[b]

Most mammalian genes produce transcripts whose 3′ ends are processed at multiple alternative positions by cleavage/polyadenylation (CPA). Poly(A) site cleavage frequently occurs cotranscriptionally and is facilitated by CPA factor binding to the RNA polymerase II (Pol II) C-terminal domain (CTD) phosphorylated on Ser2 residues of its heptad repeats (YS_2PTSPS). The function of cotranscriptional events in the selection of alternative poly(A) sites is poorly understood. We investigated Pol II pausing, CTD Ser2 phosphorylation, and processing factor CstF recruitment at wild-type and mutant IgM transgenes that use alternative poly(A) sites to produce mRNAs encoding the secreted and membrane-bound forms of the immunoglobulin (Ig) heavy chain. The results show that the sites of Pol II pausing and processing factor recruitment change depending on which poly(A) site is utilized. In contrast, the extent of Pol II CTD Ser2 phosphorylation does not closely correlate with poly(A) site selection. We conclude that changes in properties of the transcription elongation complex closely correlate with utilization of different poly(A) sites, suggesting that cotranscriptional events may influence the decision between alternative modes of pre-mRNA 3′ end processing.

RNA 3′ end formation by cleavage/polyadenylation (CPA) is an essential maturation step for most mRNAs and many long noncoding RNAs. The AAUAAA and GU-rich components of the consensus poly(A) site are recognized by the CPSF and CstF complexes, and endonucleolytic cleavage is catalyzed by CPSF (1–3). The cleavage reaction exposes a 3′ OH terminus that is polyadenylated and a 5′ phosphate end, which is degraded. CPA usually occurs cotranscriptionally and is necessary for RNA polymerase II (Pol II) transcription termination, probably coupled to exonucleolytic degradation of the cleaved nascent RNA by a torpedo mechanism (4, 5). Efficient CPA requires the C-terminal domain (CTD) of the Pol II large subunit, which comprises conserved heptad repeats with the consensus sequence YS_2PTS_5PS (6, 7). The CTD is reversibly phosphorylated at different positions within the heptad repeats, in a manner that is coordinated with the initiation, elongation, and termination phases of the transcription cycle (8–10). Phosphorylation of CTD Ser2 residues on Pol II during the elongation phase enhances RNA 3′ end processing and binding to CPA factors (11–13). Pol II pauses approximately 1 to 5 kb downstream of poly(A) sites in human cells, in a state that is hyperphosphorylated at CTD Ser2 residues and associated with the cleavage/polyadenylation factors CstF and CPSF (14–19). We have speculated that cleavage of the nascent transcript is carried out by CPSF within the paused transcription elongation complex (17), but the relationships between pausing, Ser2 hyperphosphorylation, and cleavage factor recruitment are poorly understood. It is not known whether the poly(A) site in the nascent RNA is necessary or sufficient to induce Pol II pausing, Ser2 phosphorylation (Ser2P), or 3′ end processing factor recruitment. However, on a human β-globin reporter gene, mutation of the AATAAA poly(A) site consensus sequence inhibited both Pol II pausing and Ser2 hyperphosphorylation at the normal position 1 to 2 kb downstream of the gene (20, 21). Knockdown of the cleavage factor CPSF73 also inhibited Ser2 hyperphosphorylation downstream of the β-globin gene. These results suggest a model in which Ser2 phosphorylation promotes recruitment of 3′ end processing factors that in turn reinforce Ser2 phosphorylation in a positive-feedback loop (20).

Alternative polyadenylation (APA) is a major regulator of mRNA isoforms that affects the expression of most human genes (22–24). The first example of APA was discovered for IgM heavy chain mRNA, whose 3′ end is processed in a cell-type-specific way. Mature plasma cells use a poly(A) site (μS) in the intron between the Cμ4 and M1 exons (Fig. 1A) to produce an mRNA coding for the secreted isoform. Immature B cells, on the other hand, use a more distal poly(A) site (μM) in the M2 exon (25, 26) to produce an mRNA coding for membrane-bound IgM. The cell-type-specific selection of alternative immunoglobulin (Ig) μ poly(A) sites is regulated by differential expression of the CstF64 subunit (27, 28), splicing factors (29, 30), and transcription elongation factors (31). Competition between splicing of the Cμ4-M1 intron and processing at the μS poly(A) site that lies within it is a major determinant of the decision between alternative Ig μ poly(A) sites (32). Thus, strengthening the weak 5′ splice site of this intron inhibits μS poly(A) site processing and favors the alternative μM site (33). A putative pause site where Pol II density accumulates has been identified downstream of the μS poly(A) site, and the delay associated with transcription elongation through the

Received 30 September 2015 **Returned for modification** 22 October 2015
Accepted 28 October 2015

Accepted manuscript posted online 2 November 2015

Citation 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. Mol Cell Biol 36:295–303. doi:10.1128/MCB.00898-15.

Address correspondence to David L. Bentley, david.bentley@ucdenver.edu.

B.F. and S.K. contributed equally to this article.

Supplemental material for this article may be found at http://dx.doi.org/10.1128/MCB.00898-15.

Copyright © 2016, American Society for Microbiology. All Rights Reserved.

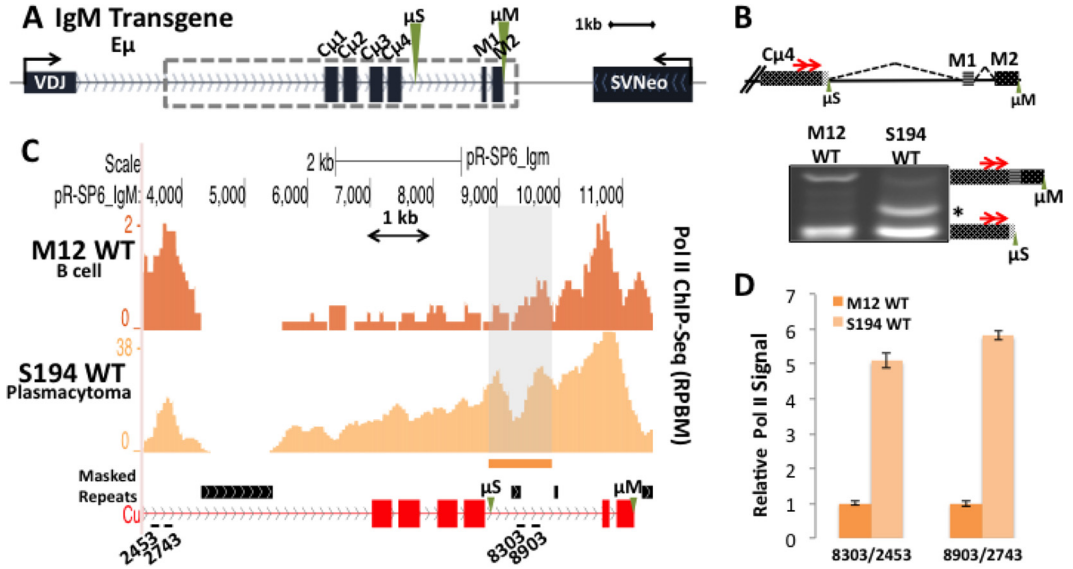

FIG 1 Pol II pausing in the Cμ4-M1 intron is coupled to μS poly(A) site use. (A) Diagram of the WT IgM transgene showing the locations of the μS and μM poly(A) sites. The plasmid (pR-SP6) also contains a convergent SVNeo gene. The dashed box indicates the region of the transgene shown in panel C. The intronic Eμ enhancer is marked. (B) Results of 3′ RACE with RNAs from S194 and M12 cells with the WT IgM transgene. Oligo(dT)-primed cDNA was amplified using nested forward PCR primers (red arrows) (see Materials and Methods). PCR fragments were visualized by ethidium bromide staining. *, unidentified reverse transcription-PCR (RT-PCR) product specific to S194 cells. (C) Anti-Pol II ChIP-seq of IgM transgenes in S194 plasmacytoma and M12 B cells revealed a Pol II pause (shaded box) specific to S194 cells, centered ~500 bp downstream of the μS poly(A) site (green arrow). A UCSC genome browser screen shot is shown. The "masked repeats" track is a UCSC track indicating highly repetitive DNA regions. Note that the dip in ChIP-seq reads in the center of the pause region is due to the presence of repeated sequences that were masked out during mapping. (D) Biological replicates of the anti-Pol II ChIPs used for panel C were analyzed by qPCR with primers specific to the μS + 500 pause site (amplicons 8303 and 8903) and normalized to upstream amplicons (amplicons 2453 and 2743). Values are expressed relative to those for the M12 line for each pair of amplicons. Means and standard errors of the means (SEM) for ≥2 PCRs are shown. The standard deviation of the ratios was calculated using the BRR formula to calculate the SEM, which is represented by error bars. Note the increased pausing in the S194 WT cells.

Cμ4-M1 intron appears to provide a competitive advantage for the μS over the μM poly(A) site (32, 34).

In principle, the decision between alternative poly(A) sites could be made by cotranscriptional and/or posttranscriptional mechanisms. It is not well understood whether alternative poly(A) site usage is associated with changes in the state of Pol II transcription elongation complexes, such as their pausing, CTD Ser2 phosphorylation, or recruitment of 3′ end processing factors. In this study, we examined several features of the transcription elongation complex at IgM transgenes modified such that either μS or μM poly(A) site utilization is favored. We report that sites of Pol II pausing and CstF recruitment to the transcription elongation complex correlate closely with which poly(A) site is utilized, whereas CTD Ser2 phosphorylation does not. These results suggest that cotranscriptional events make an important contribution to the decision regarding processing at alternative poly(A) sites.

MATERIALS AND METHODS

IgM transgenes and cell lines. M12 B cells and S194 plasmacytoma cells stably transfected with the wild-type (WT) Ig Cμ gene were described previously (33). The S194 cells were also transfected with Ig Cμ pA21 (34) and with Ig Cμ 5′-SP (33). To minimize any effects of individual integration sites on gene expression, multiple individual transformants were pooled. Quantitative PCR (qPCR) analysis of genomic DNAs from the stable polyclonal cell lines (see Fig. S1C in the supplemental material) relative to those from parental S194 cells revealed average transgene copy numbers of 1 to 5, assuming that the S194 parent is diploid. It was previously shown that multiple gene copies do not disrupt normal regulation of transiently transfected IgM reporter genes (32). CHO Flp-in cells with integrated WT and AAGAAA (A2GA3) mutant human β-globin genes were described previously (21).

Antibodies. Rabbit anti-pan-Pol II CTD was described previously (35). Rat monoclonal anti-Ser2P-CTD (3E10; Chromotek) was used with rabbit anti-rat IgG (Jackson ImmunoResearch) for immunoprecipitation (IP). Anti-CstF77 was raised in rabbits against a glutathione *S*-transferase (GST) fusion to the C terminus of human CstF77 (amino acids 539 to 717) and was affinity purified.

ChIP-seq. Chromatin immunoprecipitation sequencing (ChIP-seq), qPCR, and Illumina library preparation were performed as described previously (21). IgM primer pairs used for qPCR are listed in Table S1 in the supplemental material, and those for β-globin were described previously (21). Amplicons are named by the position of the middle of the amplicon relative to the start of the gene. Libraries were sequenced on the Illumina Hi-Seq platform. Single-end 50-base reads were mapped to the mm10 UCSC mouse genome (December 2011) or to a custom genome corresponding to the wild-type IgM SVneo plasmid pR-SP6 (36) by use of Bowtie, version 0.12.5 (37). Repeats were masked using Repeat Masker. We generated BED and WIG profiles by using 50-bp bins and 200-bp windows, assuming a 180-bp fragment size shifting effect. Libraries were normalized by the total number of aligned reads (reads per bin per million reads [RPBM]), and results were viewed with the UCSC genome browser. Reads mapped to the IgM custom genome were normalized to the total number of reads aligned to the mm10 mouse genome. Higher ChIP signals obtained for the WT IgM transgene in S194 cells than those for the pA21 and 5′-SP genes were consistent with the higher average copy number of the WT transgene (see Fig. S1C in the supplemental material).

RNA preparation and 3′ RACE. Total RNA was extracted by use of TRIzol. cDNA was synthesized using Protoscript II reverse transcriptase

(NEB) with the dT18-XbaKpnBam primer for 3′ rapid amplification of cDNA ends (RACE) (38, 39). 3′ RACE was carried out using nested forward primers and the XbaKpnBam reverse primer, as previously described (38, 39). Oligonucleotide sequences are listed in Table S1 in the supplemental material.

IgM copy number quantification. Copy number was determined by real-time PCR analysis of genomic DNA from each cell line, using primer pair 2743 for the IgM gene (see Table S1), relative to a standard curve made from the linearized pR-SP6 IgM Cμ plasmid. Copy number was normalized to that in parental S194 cells.

Accession number. ChIP-seq data sets have been deposited in the GEO database under accession number GSE75301.

RESULTS

Pol II pausing correlates with poly(A) site usage. To investigate the relationship between transcription and coupled mRNA 3′ end processing, we examined the mouse IgM gene, in which alternative poly(A) sites are used to encode the secreted (μS) and membrane-bound (μM) forms of the Ig μ heavy chain. μM expressed in B cells is processed at the downstream membrane poly(A) site, whereas μS expressed in plasma cells uses the upstream poly(A) site, located in the intron between the Cμ4 and M1 exons. The locations of these poly(A) sites in the IgM gene are diagramed in Fig. 1A. We compared M12 B cells and S194 plasmacytoma cells, each containing integrated intact rearranged IgM transgenes (pR-SP6) (33, 34, 36). 3′ RACE confirmed the previously reported shift in poly(A) site use from almost exclusively μS in S194 plasmacytoma cells to a mix of μS and μM in M12 B cells (34) (Fig. 1B). To investigate whether Pol II pausing changes when alternative poly(A) sites are used, we performed Pol II ChIP-seq on these cell lines. In these experiments, we interpreted an accumulation of the Pol II ChIP signal at a defined region as a pause. For the S194 plasmacytoma cells, there was a discrete Pol II pause, spanning a 1-kb region downstream of the μS poly(A) site (shaded bar in Fig. 1C), that was absent in the M12 B cells. A biological replicate Pol II ChIP-seq experiment confirmed that this pause was specific to S194 plasmacytoma cells and was absent or much reduced in M12 B cells (34) (see Fig. S1A, top two tracks, in the supplemental material). Note that the peak of Pol II ChIP-seq reads at this pause has a dip in the center due to repeated sequences (Fig. 1C) that were masked out during mapping. The center of this pause detected by Pol II ChIP is approximately 300 bp downstream of a pause site mapped previously by nuclear run-on analysis (34). We refer to this accumulation of Pol II density centered 500 bases downstream of the poly(A) site as the "μS + 500 pause." The previously identified pause is likely represented by the peak of Pol II density directly downstream of the μS poly(A) site, which is within the μS + 500 pause region. Large peaks of Pol II density were consistently observed in the intron between the VDJ and Cμ1 exons (Fig. 1C; see Fig. S1A), and we speculate that they may correspond to paused Pol II at cryptic transcription start sites. Promoter activity of the Eμ enhancer in this intron was documented previously (40). Substantial Pol II ChIP signals were also observed downstream of the μM poly(A) site, but interpretation of these signals is confounded by the convergent SVneo transcription unit in the transgene (Fig. 1A; see Fig. S1A). Our conclusions are therefore limited to analysis of the pause within the Cμ4-M1 intron, which is well downstream of SVneo. SVneo serves as an internal control for transgene activity, and as expected, ChIP signals over the IgM sequences relative to those for SVneo were comparable between the different cell lines (see Fig. S1A).

To validate and quantify pausing at the μS + 500 site within the Cμ4-M1 intron, we performed independent ChIP-qPCR experiments on the IgM transgene in S194 and M12 cells. Anti-Pol II ChIP signals were quantified for two amplicons amplified from regions near the μS + 500 pause site (Fig. 1A) (amplicons 8303 and 8903) and for control upstream amplicons from the VDJ-Cμ1 intron (Fig. 1A) (amplicons 2453 and 2743). qPCR confirmed the μS + 500 pause site identified by ChIP-seq and showed that its signal was approximately 5-fold higher in S194 plasmacytoma cells than in M12 B cells (Fig. 1D). In summary, these experiments show that use of the μS poly(A) site in S194 cells results in an early and distinct Pol II μS + 500 pause that is diminished or absent when the downstream membrane poly(A) site is used in B cells.

The μS poly(A) site is necessary for the μS + 500 Pol II pause. Since the μS + 500 pause is correlated with the μS poly(A) site in plasmacytoma cells, we asked whether this poly(A) site is necessary to establish Pol II pausing. For these experiments, we used the previously characterized mutant pA21, with a 21-nucleotide deletion of the μS AATAAA consensus sequence and the surrounding AU/A-rich sequence (Fig. 2A) (34, 41). We confirmed by 3′ RACE that this mutation causes a profound switch from the μS to the μM poly(A) site (Fig. 2B, lanes 1 and 2). Anti-Pol II ChIP-seq of the pA21 and WT IgM transgenes in S194 cells (Fig. 2C) revealed a marked decrease in the Pol II μS + 500 pause in the pA21 μS poly(A) site deletion relative to that in the WT (Fig. 2B, shaded box), which was confirmed in a replicate ChIP-seq experiment (see Fig. S1A in the supplemental material).

To validate and quantify altered pausing at the μS + 500 site in the pA21 mutant, we performed independent anti-Pol II ChIP-qPCR experiments. Pol II ChIP signals for two amplicons amplified from regions near the μS + 500 pause site (Fig. 2C, amplicons 8303 and 8903) were quantified relative to a control amplicon amplified from an upstream region in the Cμ4 exon (Fig. 2C, amplicon 7036). This analysis showed that the ratio of Pol II at the μS + 500 site relative to that at the upstream position was approximately 6-fold higher for the WT IgM gene than for the pA21 mutant (Fig. 2D). We concluded that the μS poly(A) site is necessary for Pol II pausing at the intronic μS + 500 site.

The μS poly(A) site is not sufficient to induce pausing. We next investigated whether the μS poly(A) site is sufficient to induce the prominent Pol II μS + 500 pause in the Cμ4-M1 intron. To address this question, we used the 5′-SP mutant, in which the 5′ splice site of the Cμ4-M1 intron is mutated to a strong canonical splice site sequence (Fig. 2A). Efficient splicing of the Cμ4-M1 intron in the 5′-SP mutant outcompetes any use of the μS poly(A) site even though the μS poly(A) site sequence is left intact (33), as we confirmed by 3′ RACE (Fig. 2B, lanes 1 and 3). We performed replicate Pol II ChIP-seq experiments with S194 cells expressing the 5′-SP transgene (Fig. 2C) and reproducibly observed a loss of the μS + 500 pause in this mutant relative to the WT (Fig. 2C; see Fig. S1A in the supplemental material).

To independently verify the reduced pausing in the 5′-SP mutant, we quantified Pol II ChIP signals by qPCR as described above for the pA21 mutant. Ratios of Pol II occupancy in the μS + 500 region (amplicons 8303 and 8903) relative to the control upstream region (amplicon 7036) were determined for the 5′-SP mutant and the WT (Fig. 2D). By this criterion, the 5′-SP mutant had a nearly 5-fold decrease in μS + 500 Pol II pausing compared to that of the WT. Together, the experiments in Fig. 2 show that while the μS poly(A) site is necessary, it is not sufficient for downstream Pol

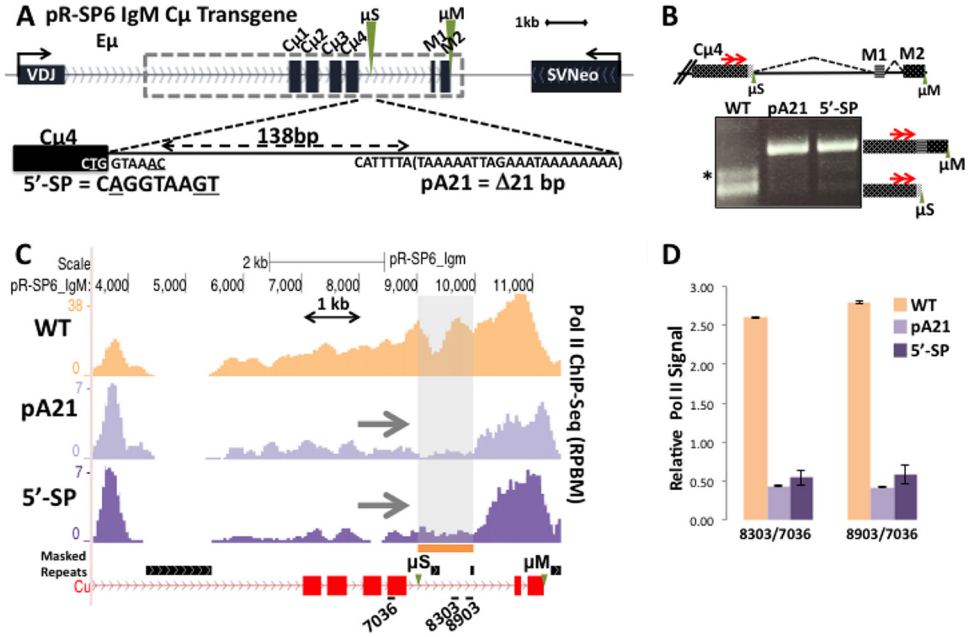

FIG 2 The μS poly(A) site is necessary but not sufficient for the μS + 500 Pol II pause. (A) Diagram of the IgM Cμ mutant transgenes, including pA21, which deletes the μS poly(A) site, and 5′-SP, which strengthens the 5′ splice site. The dashed box indicates the region of the transgene shown in panel C. (B) Results of 3′ RACE with WT and mutant IgM transgene transcripts in S194 cell lines, as described in the legend to Fig. 1B. *, unidentified RT-PCR product specific to S194 cells. Note that the pA21 and 5′-SP mutants shifted their poly(A) site use in favor of μM. (C) UCSC genome browser screen shot of anti-Pol II ChIP-seq of WT and mutant IgM transgenes in S194 cells, as described in the legend to Fig. 1C. Note that the μS + 500 pause (shaded gray box) was reduced or absent in the pA21 and 5′-SP mutants (gray arrows). (D) Biological replicates of the anti-Pol II ChIPs used for panel C were analyzed by qPCR with primers specific to the μS + 500 pause site (amplicons 8303 and 8903) and normalized to an upstream amplicon (amplicon 7036). Means and SEM for ≥3 PCRs are shown. Note the reduced pausing for both the pA21 and 5′-SP mutants relative to that for the WT.

II pausing. When μS poly(A) site usage is outcompeted by efficient splicing of the Cμ4-M1 intron, pausing within the intron is strongly inhibited.

CTD Ser2 phosphorylation is uncoupled from Pol II pausing at the μS + 500 site. We next investigated the relationship between Pol II pausing at the μS + 500 site and the Pol II CTD Ser2 phosphorylation (Ser2P) that is associated with cotranscriptional mRNA 3′ end processing. To measure Ser2P in the Cμ4-M1 intron, we carried out ChIP-seq with an antibody specific to this phospho-isoform in S194 plasmacytoma cells with the WT, pA21, and 5′-SP IgM transgenes. As expected, Pol II paused at the μS + 500 site in the WT IgM transgene was associated with a strong Ser2P ChIP signal (Fig. 3A, top track, shaded box). Unexpectedly, we also detected clear evidence of Ser2-phosphorylated Pol II at the same position in the pA21 and 5′-SP mutants (Fig. 3A, shaded box), even though Pol II pausing at the μS + 500 site was much reduced in these mutants (Fig. 2C). These results were reproduced in biological replicates that also revealed Ser2P ChIP signals at the μS + 500 Cμ4-M1 pause site in the WT, pA21, and 5′-SP transgenes (see Fig. S2A in the supplemental material).

To validate and quantify CTD Ser2 phosphorylation relative to total Pol II at the μS + 500 pause site, we performed qPCR on independent ChIP samples of three amplicons amplified from regions near the μS + 500 pause (amplicons 8303, 8584, and 8734) and several others from regions further downstream (Fig. 3B). Ratios of Ser2P to total Pol II ChIP signals for each amplicon were normalized to the Ser2P/total Pol II ratio for the 3′ end of the Actb gene, which served as an internal control. This experiment showed that the level of CTD Ser2 phosphorylation relative to total Pol II was not diminished in the region of the μS + 500 pause in either the pA21 or 5′-SP mutant (Fig. 3B), even though μS poly(A) site usage and Pol II pausing were much reduced relative to those of the WT. Surprisingly, for amplicons 8303 and 8584 from the μS + 500 pause region, the Ser2P/total Pol II ratio was actually higher for the pA21 mutant than for the WT. Together, these results suggest that CTD Ser2 phosphorylation can be uncoupled from poly(A) site processing and Pol II pausing downstream of a poly(A) site. Furthermore, the maintenance of relatively high Ser2P/total Pol II ratios near the μS poly(A) site in the 5′-SP mutant compared to the WT suggests that CTD Ser2 hyperphosphorylation is not sufficient to induce processing at the poly(A) site.

The β-globin poly(A) site is necessary for CTD Ser2 hyperphosphorylation. Our finding that the level of CTD Ser2 phosphorylation downstream of the μS poly(A) site is not well correlated with processing at that site is surprising in light of a recent study showing that CTD Ser2 hyperphosphorylation at the 3′ end of a human β-globin reporter gene is dependent on a functional poly(A) site (20). To independently assess the relationship between CTD Ser2 phosphorylation and the use of the β-globin poly(A) site, we examined single-copy human β-globin transgenes in CHO cells (21), with the poly(A) site kept WT or mutated at a single position, from AATAAA to AAGAAA (A2GA3), which prevents 3′ end processing (see Fig. S3A in the supplemental material) and causes a downstream shift of Pol II pausing relative to that with the WT (see Fig. S3B) (21). We analyzed levels of CTD Ser2 phosphorylation relative to total Pol II by ChIP-qPCR with the WT and A2GA3 mutant β-globin genes. These experiments

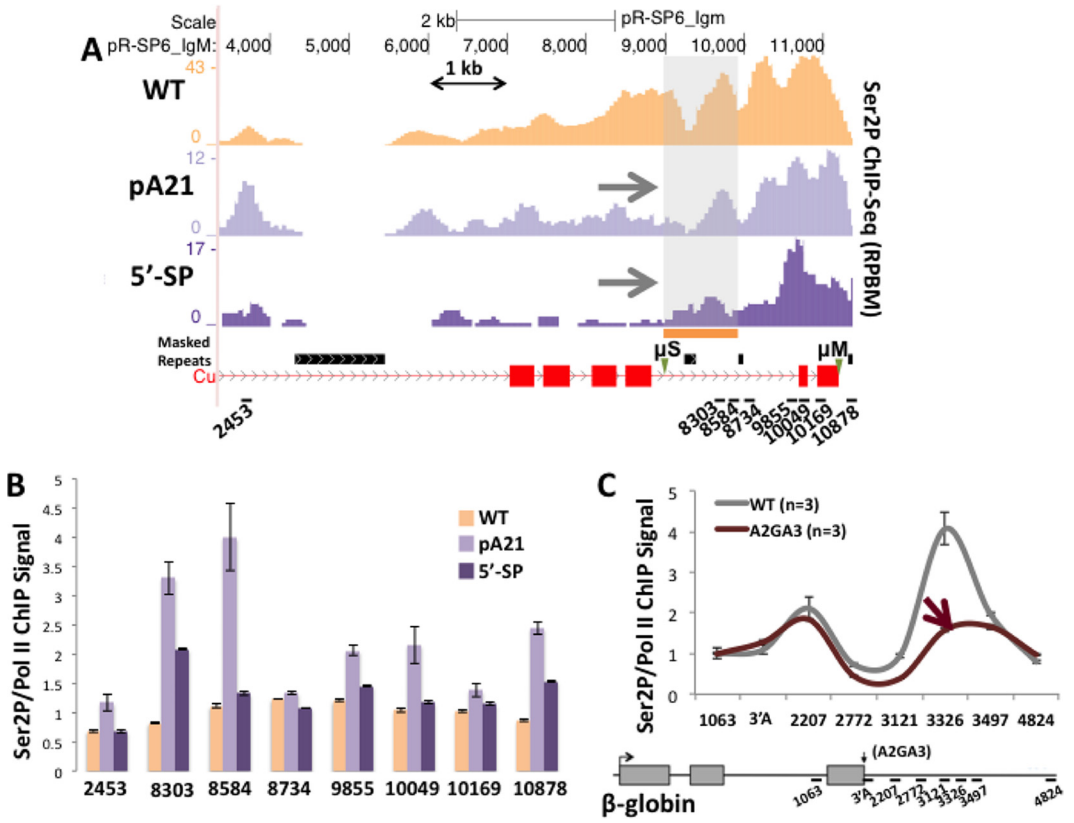

FIG 3 Uncoupling of pausing from Pol II CTD Ser2 hyperphosphorylation in mutants with altered μS poly(A) site usage. (A) UCSC genome browser screen shot of anti-Ser2P-CTD ChIP-seq of IgM transgenes in S194 cells, as described in the legend to Fig. 2C. Note the CTD Ser2P ChIP signals (arrows) at the μS + 500 pause, marked by the shaded gray box. (B) Biological replicates of the anti-Ser2P-CTD ChIPs used for panel A were quantified by real-time PCR. Ser2P/total Pol II ratios for each amplicon were normalized to the values for the 3′ end of the Actb gene (see Fig. S2B in the supplemental material). Means and SEM for ≥3 PCRs are shown, as described in the legend to Fig. 1D. (C) Poly(A) site mutation reduces CTD Ser2P at the β-globin gene, in contrast to the case at the IgM μS poly(A) site. Anti-Ser2P/total Pol II values for each amplicon were normalized to those for position 1063, upstream of the poly(A) site, in the WT and A2GA3 poly(A) site mutant β-globin transgenes in CHO Flp-in cells. Means and SEM are shown. Note the decreased Ser2P/total Pol II ratio for the A2GA3 mutant (arrow), consistent with previous results (20).

showed that Ser2P levels were decreased in the A2GA3 mutant, in agreement with previous results (Fig. 3C) (20). Together, these observations suggest that the relationship between poly(A) site use and CTD phosphorylation on Ser2 differs between genes and may depend on whether a single major poly(A) site or multiple alternative sites are present.

Alternative poly(A) site use and CstF recruitment. To investigate the relationship between μS poly(A) site usage and recruitment of the cleavage/polyadenylation factor CstF, we conducted anti-CstF77 ChIP-seq of WT and mutant IgM transgenes in S194 cells. As expected, there was strong recruitment of CstF77 to WT IgM in the region of the μS + 500 Pol II pause (Fig. 4A). In contrast, relative CstF77 ChIP signals in the region of the pause site were lower in the pA21 and 5′-SP mutants than in the WT, consistent with reduced Pol II occupancy at these positions (Fig. 4A). For comparison, levels of CstF77 recruitment to the Actb gene were very similar in all three cell lines (see Fig. S4 in the supplemental material).

To quantify CstF77 relative to Pol II, we performed an independent ChIP-qPCR experiment and quantified the results for 16 primer pairs spanning the IgM and SVneo transcription units (Fig. 4B). CstF77/Pol II values were normalized to those for a control region of maximal recruitment downstream of the SVneo gene. This experiment showed that, consistent with the ChIP-seq results, CstF77 recruitment downstream of the μS poly(A) site was lower in both the pA21 and 5′-SP mutants (Fig. 4B, arrow), in which processing at this site was sharply reduced. The greatest reduction in CstF77 recruitment occurred in the pA21 poly(A) site deletion mutant, consistent with the idea that recognition of the consensus RNA processing element is important for stable association of CstF with the transcription complex. These results also suggest that maintenance of WT or even higher levels of CTD Ser2 phosphorylation in the pA21 and 5′-SP mutants (Fig. 3B) is not sufficient to optimally recruit the CstF77 3′ end processing factor.

DISCUSSION

In this report, we show that use of the alternative poly(A) sites in the IgM gene is associated with distinct cotranscriptional events manifested by different properties of the Pol II transcription elongation complexes on the gene. The salient results are as follows. (i) Pol II pausing downstream of the μS poly(A) site and recruitment of the 3′ end processing factor CstF correlate with processing at that site, suggesting a functional coupling (Fig. 2 and 4). (ii) A functional poly(A) site is necessary but not sufficient for pausing. When the μS poly(A) site and the necessary factors in plasmacy-

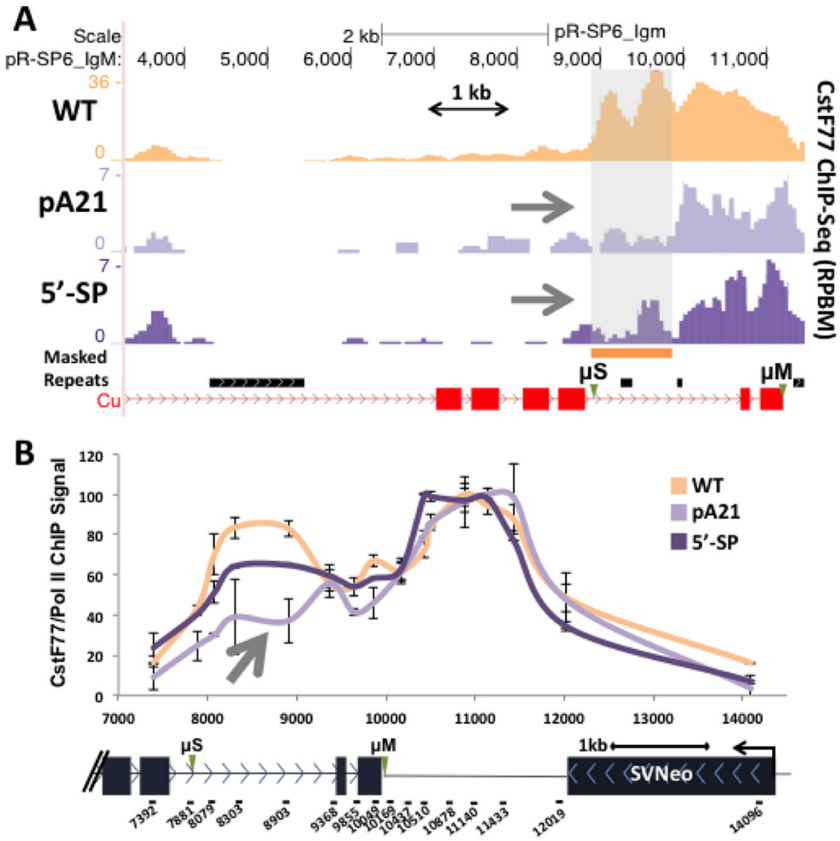

FIG 4 Reduced CstF recruitment in mutants with decreased μS poly(A) site use. (A) UCSC genome browser screen shot of anti-CstF77 ChIP-seq of IgM transgenes in S194 cells, as described in the legend to Fig. 2C. Note the reduced CstF77 recruitment at the μS + 500 pause (shaded box) for the pA21 and 5′-SP mutants (gray arrows) relative to the WT. (B) Biological replicate anti-CstF77 ChIPs quantified by real-time PCR. The CstF77/total Pol II ratio for each amplicon was normalized to the maximum value for the region between the convergent IgM and SVneo genes. Means and SEM for 2 PCRs are shown. Note the reduced CstF77 recruitment relative to that of Pol II in the pA21 and 5′-SP mutants (arrow).

toma cells are present but the site is not utilized, as in the 5′-SP mutant, pausing is suppressed (Fig. 2). (iii) In contrast to the case for the β-globin gene, with a single strong poly(A) site, the level of Pol II CTD Ser2 phosphorylation downstream of the alternative IgM μS poly(A) site did not correlate closely with the extent of processing at that site.

A previously reported Pol II pause site located 50 to 200 nucleotides downstream of the μS poly(A) site was mapped by nuclear run-on analysis (34). The higher-resolution Pol II ChIP-seq experiments reported here suggest that this pause extends several hundred bases further downstream, into the middle of the Cμ4-M1 intron, and is centered approximately 500 bases downstream of the poly(A) site. Interestingly, both deletion of the μS poly(A) site (in the pA21 mutant) and reduced processing at this site without changing its sequence (in the 5′-SP mutant) strongly inhibit Pol II pausing at the μS + 500 position (Fig. 2). This paused peak of Pol II in the WT gene is also more prominent in plasmacytoma cells, where processing at the μS poly(A) site predominates, than in B cells, where it is used less frequently (Fig. 1). We conclude that the 3′ end Pol II pause is strongly associated with functional processing at the μS poly(A) site rather than just the presence of the consensus cleavage/polyadenylation sequences. This result is consistent with the hypothesis that poly(A) site processing within the context of a Pol II transcription elongation complex causes the polymerase to pause (Fig. 5A). This idea is also supported by previous results implicating the cleavage factor CPSF in promoting pausing *in vitro* and *in vivo* (20, 42). In addition, pausing could feed back to facilitate cotranscriptional μS poly(A) site processing, as previously suggested (16, 43). It is also possible that binding of U1 snRNP to the 5′ splice site inhibits pausing within the intron. We are not aware of any evidence for such a mechanism, but it would be consistent with the proposed inhibition of premature cleavage polyadenylation by U1 snRNP (38, 39). Our observations are not easily reconciled with the idea that pausing is caused by intrinsic, dominantly acting DNA sequence elements within the Cμ4-M1 intron, because such sequences are unaffected by the poly(A) site (pA21) and 5′ splice site (5′-SP) mutations that severely reduce pausing (Fig. 2).

CTD Ser2 hyperphosphorylation correlates with Pol II pausing and recruitment of 3′ end processing factors downstream of poly(A) sites on human genes (17, 20). However, what this correlation signifies for the mechanism relating 3′ end processing, processing factor recruitment, and CTD phosphorylation remains unclear. Deletion of the single major poly(A) site in the human β-globin gene caused a marked loss of CTD Ser2 phosphorylation in the 3′ flanking region (20), and we repeated this observation in independent reporter cell lines (Fig. 3C). These results suggest some form of coupling between poly(A) site recognition and CTD Ser2 phosphorylation. In contrast to the case for β-globin, we found that CTD Ser2 phosphorylation downstream of the IgM μS

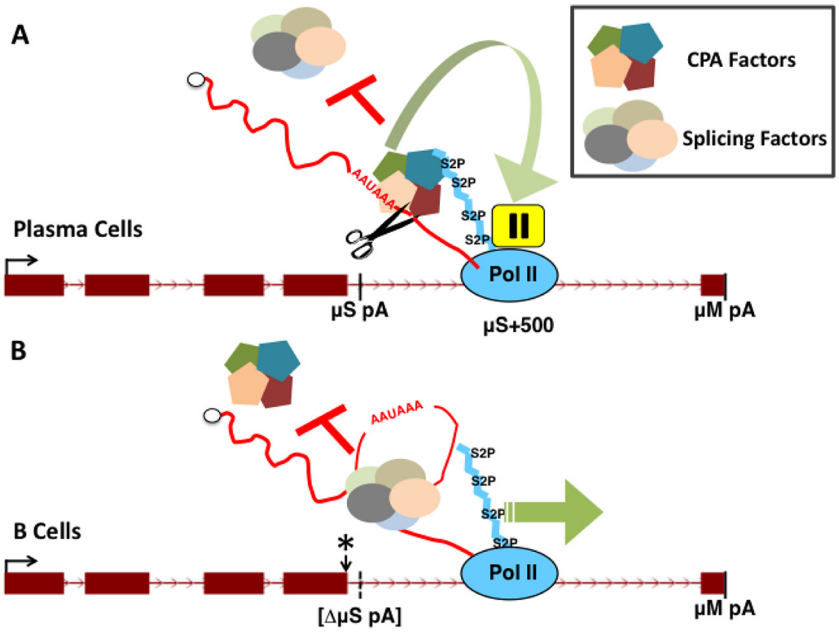

FIG 5 Model for coordination of pausing and 3′ end processing factor recruitment with alternative poly(A) site choice at the IgM gene. (A) In plasma cells, processing of the μS poly(A) site outcompetes splicing of the Cμ4-M1 intron and is associated with recruitment of cleavage/polyadenylation (CPA) factors to the transcription elongation complex. Our results suggest that recognition of the poly(A) site results in Pol II pausing. It is not known whether splicing factors are excluded as shown or present in the complex but unable to function. (B) In B cells with the poly(A) site deletion mutant (pA21 [ΔμS pA]) or the consensus 5′ splice site mutant (5′-SP [*]), splicing outcompetes the μS poly(A) site. When splicing factors have the competitive advantage, CPA factors are excluded and Pol II does not pause at the μS + 500 site within the Cμ4-M1 intron.

alternative poly(A) site was maintained and even enhanced in mutants with severely reduced processing at this site (Fig. 3B and 5B). These results therefore suggest that maintenance of Ser2 phosphorylation can be uncoupled from poly(A) site processing and pausing, but they do not eliminate the possibility that pausing can enhance Ser2 phosphorylation. Additionally, we cannot exclude the possibility that phospho-CTD ChIP signals are affected by epitope masking in some contexts or that poly(A) site usage could influence the reactivity of phospho-CTD epitopes. In summary, the relationship between Ser2 hyperphosphorylation and RNA 3′ end processing can differ between genes and may depend on gene length and whether a single strong poly(A) site or multiple alternative sites are present. It remains possible that inherent poly(A) site strength might influence the relationship between 3′ end processing and CTD Ser2 phosphorylation. It is not known whether the β-globin and IgM μS poly(A) sites have different strengths, though both have consensus AAUAAA elements and the μS site has functional AU- and GU-rich consensus elements (41, 44).

The roles of the poly(A) site consensus sequences, Ser2 phosphorylation of the CTD, and transcriptional pausing in recruitment of cleavage/polyadenylation factors to the Pol II elongation complex remain to be resolved. We observed that cotranscriptional recruitment of CstF to elongation complexes on the IgM gene changed when different alternative poly(A) sites were utilized. Processing at the μS poly(A) site was associated with strong recruitment of CstF77 to Pol II complexes at the μS + 500 pause site (Fig. 4A). Inhibition of processing at this site, either by deletion of the consensus AATAAA sequence (pA21) or by strengthening of the competing 5′ splice site (5′-SP), diminished CstF77 recruitment (Fig. 4B), although CTD Ser2 phosphorylation levels were maintained relative to those of the WT (Fig. 3B). These results argue that recognition of 3′ end processing signals in the nascent transcript is necessary for maximal association of CstF with the transcription elongation complex. Recruitment of CstF may also be enhanced by the Ser2-hyperphosphorylated CTD; however, Ser2 hyperphosphorylation alone appears not to be sufficient for formation of a stable complex. Although CstF77 recruitment downstream of the μS poly(A) site approximately parallels the level of processing at this site (Fig. 2B and 4B), we cannot exclude the possibility that other factors might regulate processing at a step after CstF binding. For example, it is possible that U1 snRNP, which inhibits processing at premature poly(A) sites (38, 39), might inhibit μS poly(A) site processing at a stage after CstF77 recruitment.

In summary, our results show that changes in alternative poly(A) site choice are associated with changes in Pol II pausing and 3′ end processing factor recruitment to the elongation complex. The results suggest that the decision between alternative poly(A) sites is likely made in part at the cotranscriptional level. In future, it will be of interest to investigate how properties of the Pol II transcription elongation complex, such as its elongation rate and CTD phosphorylation state, may influence the decision between alternative poly(A) sites and to identify factors responsible for CTD Ser2 hyperphosphorylation independent of poly(A) site use.

ACKNOWLEDGMENTS

This work was supported by NIH grant GM58613 to D.L.B. and NSF grant MCB-0919099 to M.L.P. S.K. was supported by a postdoctoral fellowship (grant PF-07-297-01-GMC) from the American Cancer Society and the Clark family. B.F. was supported by the Victor W. Bolie and Earleen D.

Bolie Graduate Scholarship Fund through the UC Denver Molecular Biology Program.

We thank K. Diener, B. Gao, and the UC Denver Nextgen sequencing facility for sequencing, P. Vanlaarhoven for antibody purification, and T. Saldi, R. Sheridan, and M. Cortazar for helpful discussions and comments on the manuscript.

FUNDING INFORMATION

HHS | NIH | National Institute of General Medical Sciences (NIGMS) provided funding to David Bentley under grant number GM58613. NSF | BIO | Division of Molecular and Cellular Biosciences (MCB) provided funding to Martha L. Peterson under grant number MCB-0919099.

REFERENCES

1. **Chan SL, Huppertz I, Yao C, Weng L, Moresco JJ, Yates JR, III, Ule J, Manley JL, Shi Y.** 2014. CPSF30 and Wdr33 directly bind to AAUAAA in mammalian mRNA 3′ processing. Genes Dev **28:**2370–2380. http://dx.doi.org/10.1101/gad.250993.114.
2. **Mandel CR, Bai Y, Tong L.** 2008. Protein factors in pre-mRNA 3′-end processing. Cell Mol Life Sci **65:**1099–1122. http://dx.doi.org/10.1007/s00018-007-7474-3.
3. **Schönemann L, Kühn U, Martin G, Schäfer P, Gruber AR, Keller W, Zavolan M, Wahle E.** 2014. Reconstitution of CPSF active in polyadenylation: recognition of the polyadenylation signal by WDR33. Genes Dev **28:**2381–2393. http://dx.doi.org/10.1101/gad.250985.114.
4. **Kuehner JN, Pearson EL, Moore C.** 2011. Unravelling the means to an end: RNA polymerase II transcription termination. Nat Rev Mol Cell Biol **12:**283–294. http://dx.doi.org/10.1038/nrm3098.
5. **Richard P, Manley JL.** 2009. Transcription termination by nuclear RNA polymerases. Genes Dev **23:**1247–1269. http://dx.doi.org/10.1101/gad.1792809.
6. **Hirose Y, Manley JL.** 1998. RNA polymerase II is an essential mRNA polyadenylation factor. Nature **395:**93–96. http://dx.doi.org/10.1038/25786.
7. **McCracken S, Fong N, Yankulov K, Ballantyne S, Pan G, Greenblatt J, Patterson SD, Wickens M, Bentley DL.** 1997. The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. Nature **385:**357–361. http://dx.doi.org/10.1038/385357a0.
8. **Buratowski S.** 2009. Progression through the RNA polymerase II CTD cycle. Mol Cell **36:**541–546. http://dx.doi.org/10.1016/j.molcel.2009.10.019.
9. **Heidemann M, Hintermair C, Voss K, Eick D.** 2013. Dynamic phosphorylation patterns of RNA polymerase II CTD during transcription. Biochim Biophys Acta **1829:**55–62. http://dx.doi.org/10.1016/j.bbagrm.2012.08.013.
10. **Komarnitsky P, Cho EJ, Buratowski S.** 2000. Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. Genes Dev **14:**2452–2460. http://dx.doi.org/10.1101/gad.824700.
11. **Ahn SH, Kim M, Buratowski S.** 2004. Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3′ end processing. Mol Cell **13:**67–76. http://dx.doi.org/10.1016/S1097-2765(03)00492-1.
12. **Licatalosi DD, Geiger G, Minet M, Schroeder S, Cilli K, McNeil JB, Bentley DL.** 2002. Functional interaction of yeast pre-mRNA 3′ end processing factors with RNA polymerase II. Mol Cell **9:**1101–1111. http://dx.doi.org/10.1016/S1097-2765(02)00518-X.
13. **Ni Z, Schwartz BE, Werner J, Suarez JR, Lis JT.** 2004. Coordination of transcription, RNA processing, and surveillance by P-TEFb kinase on heat shock genes. Mol Cell **13:**55–65. http://dx.doi.org/10.1016/S1097-2765(03)00526-4.
14. **Anamika K, Gyenis A, Poidevin L, Poch O, Tora L.** 2012. RNA polymerase II pausing downstream of core histone genes is different from genes producing polyadenylated transcripts. PLoS One **7:**e38769. http://dx.doi.org/10.1371/journal.pone.0038769.
15. **Dye MJ, Gromak N, Proudfoot NJ.** 2006. Exon tethering in transcription by RNA polymerase II. Mol Cell **21:**849–859. http://dx.doi.org/10.1016/j.molcel.2006.01.032.
16. **Enriquez-Harris P, Levitt N, Briggs D, Proudfoot NJ.** 1991. A pause site for RNA polymerase II is associated with termination of transcription. EMBO J **10:**1833–1842.
17. **Glover-Cutter K, Kim S, Espinosa J, Bentley DL.** 2008. RNA polymerase II pauses and associates with pre-mRNA processing factors at both ends of genes. Nat Struct Mol Biol **15:**71–78. http://dx.doi.org/10.1038/nsmb1352.
18. **Gromak N, West S, Proudfoot NJ.** 2006. Pause sites promote transcriptional termination of mammalian RNA polymerase II. Mol Cell Biol **26:**3986–3996. http://dx.doi.org/10.1128/MCB.26.10.3986-3996.2006.
19. **Lian Z, Karpikov A, Lian J, Mahajan MC, Hartman S, Gerstein M, Snyder M, Weissman SM.** 2008. A genomic analysis of RNA polymerase II modification and chromatin architecture related to 3′ end RNA polyadenylation. Genome Res **18:**1224–1237. http://dx.doi.org/10.1101/gr.075804.107.
20. **Davidson L, Muniz L, West S.** 2014. 3′ end formation of pre-mRNA and phosphorylation of Ser2 on the RNA polymerase II CTD are reciprocally coupled in human cells. Genes Dev **28:**342–356. http://dx.doi.org/10.1101/gad.231274.113.
21. **Kim S, Kim H, Fong N, Erickson B, Bentley DL.** 2011. Pre-mRNA splicing is a determinant of histone H3K36 methylation. Proc Natl Acad Sci U S A **108:**13564–13569. http://dx.doi.org/10.1073/pnas.1109475108.
22. **Di Giammartino DC, Nishida K, Manley JL.** 2011. Mechanisms and consequences of alternative polyadenylation. Mol Cell **43:**853–866. http://dx.doi.org/10.1016/j.molcel.2011.08.017.
23. **Hoque M, Ji Z, Zheng D, Luo W, Li W, You B, Park JY, Yehia G, Tian B.** 2013. Analysis of alternative cleavage and polyadenylation by 3′ region extraction and deep sequencing. Nat Methods **10:**133–139. http://dx.doi.org/10.1038/nchembio.1406.
24. **Lianoglou S, Garg V, Yang JL, Leslie CS, Mayr C.** 2013. Ubiquitously transcribed genes use alternative polyadenylation to achieve tissue-specific expression. Genes Dev **27:**2380–2396. http://dx.doi.org/10.1101/gad.229328.113.
25. **Alt FW, Bothwell AL, Knapp M, Siden E, Mather E, Koshland M, Baltimore D.** 1980. Synthesis of secreted and membrane-bound immunoglobulin mu heavy chains is directed by mRNAs that differ at their 3′ ends. Cell **20:**293–301. http://dx.doi.org/10.1016/0092-8674(80)90615-7.
26. **Early P, Rogers J, Davis M, Calame K, Bond M, Wall R, Hood L.** 1980. Two mRNAs can be produced from a single immunoglobulin mu gene by alternative RNA processing pathways. Cell **20:**313–319. http://dx.doi.org/10.1016/0092-8674(80)90617-0.
27. **Takagaki Y, Manley JL.** 1998. Levels of polyadenylation factor CstF-64 control IgM heavy chain mRNA accumulation and other events associated with B cell differentiation. Mol Cell **2:**761–771. http://dx.doi.org/10.1016/S1097-2765(00)80291-9.
28. **Takagaki Y, Seipelt RL, Peterson ML, Manley JL.** 1996. The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation. Cell **87:**941–952. http://dx.doi.org/10.1016/S0092-8674(00)82000-0.
29. **Bruce SR, Dingle RW, Peterson ML.** 2003. B-cell and plasma-cell splicing differences: a potential role in regulated immunoglobulin RNA processing. RNA **9:**1264–1273. http://dx.doi.org/10.1261/rna.5820103.
30. **Ma J, Gunderson SI, Phillips C.** 2006. Non-snRNP U1A levels decrease during mammalian B-cell differentiation and release the IgM secretory poly(A) site from repression. RNA **12:**122–132. http://dx.doi.org/10.1261/rna.2159506.
31. **Martincic K, Alkan SA, Cheatle A, Borghesi L, Milcarek C.** 2009. Transcription elongation factor ELL2 directs immunoglobulin secretion in plasma cells by stimulating altered RNA processing. Nat Immunol **10:**1102–1109. http://dx.doi.org/10.1038/ni.1786.
32. **Peterson ML, Perry RP.** 1986. Regulated production of mu m and mu s mRNA requires linkage of the poly(A) addition sites and is dependent on the length of the mu s-mu m intron. Proc Natl Acad Sci U S A **83:**8883–8887. http://dx.doi.org/10.1073/pnas.83.23.8883.
33. **Peterson ML, Perry RP.** 1989. The regulated production of mu m and mu s mRNA is dependent on the relative efficiencies of mu s poly(A) site usage and the c mu 4-to-M1 splice. Mol Cell Biol **9:**726–738. http://dx.doi.org/10.1128/MCB.9.2.726.
34. **Peterson ML, Bertolino S, Davis F.** 2002. An RNA polymerase pause site is associated with the immunoglobulin mus poly(A) site. Mol Cell Biol **22:**5606–5615. http://dx.doi.org/10.1128/MCB.22.15.5606-5615.2002.
35. **Schroeder SC, Schwer B, Shuman S, Bentley D.** 2000. Dynamic association of capping enzymes with transcribing RNA polymerase II. Genes Dev **14:**2435–2440. http://dx.doi.org/10.1101/gad.836300.
36. **Ochi A, Hawley RG, Hawley T, Shulman MJ, Traunecker A, Kohler G, Hozumi N.** 1983. Functional immunoglobulin M production after transfection of cloned immunoglobulin heavy and light chain genes into lym-

phoid cells. Proc Natl Acad Sci U S A **80:**6351–6355. http://dx.doi.org/10.1073/pnas.80.20.6351.
37. **Langmead B, Trapnell C, Pop M, Salzberg SL.** 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol **10:**R25. http://dx.doi.org/10.1186/gb-2009-10-3-r25.
38. **Berg MG, Singh LN, Younis I, Kaida D, Zhang Z, Wan L, Dreyfuss G.** 2012. U1 snRNP determines mRNA length and regulates isoform expression. Cell **150:**53–64. http://dx.doi.org/10.1016/j.cell.2012.05.029.
39. **Kaida D, Berg MG, Younis I, Kasim M, Singh LN, Wan L, Dreyfuss G.** 2010. U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. Nature **468:**664–668. http://dx.doi.org/10.1038/nature09479.
40. **Lennon GG, Perry RP.** 1985. C mu-containing transcripts initiate heterogeneously within the IgH enhancer region and contain a novel 5′-nontranslatable exon. Nature **318:**475–478. http://dx.doi.org/10.1038/318475a0.
41. **Peterson ML, Bingham GL, Cowan C.** 2006. Multiple features contribute to the use of the immunoglobulin M secretion-specific poly(A) signal but are not required for developmental regulation. Mol Cell Biol **26:**6762–6771. http://dx.doi.org/10.1128/MCB.00889-06.
42. **Nag A, Narsinh K, Martinson HG.** 2007. The poly(A)-dependent transcriptional pause is mediated by CPSF acting on the body of the polymerase. Nat Struct Mol Biol **14:**662–669. http://dx.doi.org/10.1038/nsmb1253.
43. **Yonaha M, Proudfoot NJ.** 1999. Specific transcriptional pausing activates polyadenylation in a coupled in vitro system. Mol Cell **3:**593–600. http://dx.doi.org/10.1016/S1097-2765(00)80352-4.
44. **Phillips C, Virtanen A.** 1997. The murine IgM secretory poly(A) site contains dual upstream and downstream elements which affect polyadenylation. Nucleic Acids Res **25:**2344–2351. http://dx.doi.org/10.1093/nar/25.12.2344.