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

Frank Rigo and Harold G. Martinson*

Department of Chemistry and Biochemistry and the Molecular Biology Institute, University of California at Los Angeles, Los Angeles, California 90095-1569

Received 7 August 2007/Returned for modification 14 September 2007/Accepted 18 October 2007

We have developed an in vitro transcription system, using HeLa nuclear extract, that supports not only efficient splicing of a multiexon transcript but also efficient cleavage and polyadenylation. In this system, both last-intron splicing and cleavage/polyadenylation are functionally coupled to transcription via the tether of nascent RNA that extends from the terminal exon to the transcribing polymerase downstream. Communication between the 3' splice site and the poly(A) site across the terminal exon is established within minutes of their transcription, and multiple steps leading up to 3'-end processing of this exon can be distinguished. First, the 3' splice site establishes connections to enhance 3'-end processing, while the nascent 3'end processing apparatus makes reciprocal functional connections to enhance splicing. Then, commitment to poly(A) site cleavage itself occurs and the connections of the 3'-end processing apparatus to the transcribing polymerase are strengthened. Finally, the chemical steps in the processing of the terminal exon take place, beginning with poly(A) site cleavage, continuing with polyadenylation of the 3' end, and then finishing with splicing of the last intron.

Most vertebrate messenger RNAs are capped, spliced, and polyadenylated, and the molecular machineries responsible for these posttranscriptional modifications are intimately coupled both with each other and with the transcriptional apparatus (3, 4, 10, 34, 44, 58). Striking examples of the collaboration between these machineries are the splicing and the cleavage and polyadenylation proteins that can be recruited to the polymerase in vitro prior to transcription and then transferred to the RNA (14, 69). In vivo, nuclear receptor coregulators and other transcription factors can also act at the promoter to establish splicing patterns for RNA molecules whose transcription has not yet even begun (3, 28, 34, 62).

For all mRNAs, upon initiation of transcription, capping factors are recruited to the elongation complex and the transcript is capped (12, 51). For the majority of vertebrate genes which contain introns, completion of capping is communicated to the splicing machinery via the cap-binding complex, and together, these protein assemblies, with the help of serine-arginine-rich (SR) proteins, define the first exon (16, 40). Then, as transcription proceeds, exons are successively defined along the transcript by communication between the 3' and 5' splice sites (5), probably in collaboration with the polymerase large subunit C-terminal domain (CTD) (19, 31, 50, 74), until the end of the message is reached. There, the splicing and the cleavage/polyadenylation machineries cooperate to define the terminal exon (36, 48, 70). The coupling between splicing and cleavage/polyadenylation that defines the terminal exon also results in the mutual stimulation of both of these reactions (2, 41, 56, 57), and recognition of the 3' splice site is coupled to transcription as well (20). Indeed, as stated at the outset, the machinery for every step of processing from capping to polyadenylation is intimately and functionally interconnected with transcription (6, 15–17, 23, 25, 33, 38, 42, 43, 45, 49, 60).

Once the terminal exon is defined, cleavage at the poly(A) site ensues (77) and the 3' end of the mRNA is irrevocably determined. Events leading up to cleavage include assembly of the cleavage and polyadenylation apparatus on the CTD of the polymerase (1, 22, 30, 45, 46, 65, 75) through a process that may be stepwise (11, 27, 54, 60) and initially very tenuous (60). Indeed, if the tether of nascent RNA that holds the poly(A) signal close to the polymerase is cut before the apparatus is sufficiently mature, the apparatus loses its grip on the polymerase, assembly ceases, and cleavage/polyadenylation is inhibited (6, 60). However, in the normal course of events, assembly proceeds and the poly(A) signal becomes more strongly associated with the polymerase, to which it remains attached even after cleavage at the poly(A) site (1, 60).

Surprisingly, although the splicing and the cleavage/polyadenylation reactions are known to be coupled strongly to each other (2, 36, 41, 48, 56, 57, 70) and in vitro transcription has successfully been coupled individually to efficient splicing (15, 16, 26, 29, 55) and to efficient cleavage and polyadenylation (1, 47, 60, 73), it has not yet been possible in vitro to obtain concerted reactions that robustly and reproducibly recapitulate all three activities. Here we describe an in vitro system in which efficient splicing and efficient cleavage and polyadenylation are both coupled to transcription and to each other. Splicing in this system proceeds almost exclusively via exon definition. Using this system, we demonstrated a dependence of both last-intron splicing and poly(A) site cleavage on the tether of nascent RNA that connects them to the polymerase, and we dissect the...
progression of events leading up to 3’-end processing of the terminal exon into several steps.

**MATERIALS AND METHODS**

**Plasmids.** The founding construct for our experiments, pβ-β, is diagrammed in Fig. S1B in the supplemental material. For this construct, the 2,173-bp HindIII-PstI fragment of SP64-HΔ6 (containing the human beta-globin gene) (23) was placed under the control of the simian virus 40 (SV40) early promoter by insertion into Stul-Clal-digested pAP117/cat (67). For the p(βA) signal mutant version of this plasmid, the AATAAAA hexamer of the β-globin poly(A) signal was converted to AgTAct by site-directed mutagenesis.

For some experiments, we used a construct with a shortened second intron, pβ-β. To shorten the intron, HindIII sites were inserted 96 bp and 707 bp downstream of the 5’ splice site of the second intron. The removal of the resulting HindIII segment from the plasmid resulted in a 611-bp deletion from the second intron.

We also used a construct, pβ-β-L, in which the strong SV40 late poly(A) signal replaced the weaker β-globin one. For this construct, a 328-bp PCR fragment from pβ40E6L (60) (produced using oligonucleotide 1 and oligonucleotide 2) was ligated to a 3,491-bp PCR fragment (oligonucleotide 3 and oligonucleotide 4) from pβ40E6L (23). The sequence of and mutation version of the p(βA) signal for this construct were AATAAA and AAGtAc. Mutations in the 5’ and 3’ splice sites of these plasmids were produced by site-directed mutagenesis.

For the tether-cutting experiments for which data are shown in Fig. 3 through 5, we wanted to be able to compare the effects of the SV40 late and the β-globin poly(A) signals. To eliminate any differences that might be attributable to downstream events, we used the vector sequence beginning 47 bp downstream of the SV40 late poly(A) cleavage site to an identical position downstream of the β-globin poly(A) site. This was done by creating a BglII site 47 bp downstream of the β-globin poly(A) cleavage site in pβ-β and pβ-β-L, and then replacing this segment with a 1,907-bp BamHI-PvuI fragment taken from the SV40 L late construct to create pβ-β-L.

**DNA oligomers (5’ to 3’).** The oligonucleotides used were as follows: oligonucleotide 1, CTCACTCTGARGCCGTCCAGCCCGGATCATAG; oligonucleotide 2, GCAACACGGCTTTAGCTCGTAACTCGTCCAGGAC; oligonucleotide 3, GTGTAACAACCGCTGTGAGGAAAGCCCTCGAGC; oligonucleotide 4, GCTAGCTCGAGGCAAGAAAGCGAGCTTA; oligonucleotide 5, CCGCCGCTATTGTTG; oligonucleotide 6, CCCACCCCTTCTGAGCAGCAGAAGCTAAG; oligonucleotide 7, CCGCTTTTTGAAATATCTCGGAGCTTGC; oligonucleotide 8, AACGATATTGAATGAGCTTGGT; oligonucleotide 9, GGGGAGCCTGAGCAGAGAGT; oligonucleotide 10, CCGCTTTTTGAAATATCTCG.
Although cleavage/polyadenylation takes place before the chemical steps of splicing shown in Fig. 1A, we assume that transcription-coupled spliceosome assembly commences much earlier, as previously reported (15) (see below). We also note that the introns do not appear to be spliced sequentially in this experiment, as suggested by the significant presence of partially spliced RNA in which only the second intron has been removed. However, in agreement with previous reports (37, 55), we found first-intron splicing to uniformly precede second-intron splicing for the native β-globin gene, which has a longer second intron and a weaker poly(A) signal, factors which combine to reduce the rate of second-intron splicing compared to that of the construct in Fig. 1 (data not shown).

To confirm by sequencing the identities of the species in Fig. 1A, we needed to generate discrete bands that could be cut out of the gel for reverse transcription-PCR. To generate discrete bands, we repeated the reaction producing the result shown in Fig. 1A, lane 3, but with the addition of 3’-dATP in the later part of the reaction to prevent poly(A) tail growth and the attendant heterogeneity in length. Since 3’-dATP interferes with transcription, it was added only after a significant fraction of the pulse-labeled transcripts had passed the poly(A) site but before significant polyadenylation had occurred (60). The discrete bands generated by splicing in the presence of 3’-dATP are shown in Fig. 1B, lane 2. For comparison, a control reaction carried out in parallel in the absence of 3’-dATP is shown in lane 1. Since 3’-dATP blocks both transcription and poly(A) tail growth, α-amanitin was also added to both reactions at 4 min so that transcription would not occur in either sample after 4 min. Thus, the only difference between lane 1 and lane 2 in Fig. 1B is the presence or absence of poly(A) tail growth.

To confirm by sequencing the identities of the species in Fig. 1A, we needed to generate discrete bands that could be cut out of the gel for reverse transcription-PCR. To generate discrete bands, we repeated the reaction producing the result shown in Fig. 1A, lane 3, but with the addition of 3’-dATP in the later part of the reaction to prevent poly(A) tail growth and the attendant heterogeneity in length. Since 3’-dATP interferes with transcription, it was added only after a significant fraction of the pulse-labeled transcripts had passed the poly(A) site but before significant polyadenylation had occurred (60). The discrete bands generated by splicing in the presence of 3’-dATP are shown in Fig. 1B, lane 2. For comparison, a control reaction carried out in parallel in the absence of 3’-dATP is shown in lane 1. Since 3’-dATP blocks both transcription and poly(A) tail growth, α-amanitin was also added to both reactions at 4 min so that transcription would not occur in either sample after 4 min. Thus, the only difference between lane 1 and lane 2 in Fig. 1B is the presence or absence of poly(A) tail growth.

To confirm by sequencing the identities of the species in Fig. 1A, we needed to generate discrete bands that could be cut out of the gel for reverse transcription-PCR. To generate discrete bands, we repeated the reaction producing the result shown in Fig. 1A, lane 3, but with the addition of 3’-dATP in the later part of the reaction to prevent poly(A) tail growth and the attendant heterogeneity in length. Since 3’-dATP interferes with transcription, it was added only after a significant fraction of the pulse-labeled transcripts had passed the poly(A) site but before significant polyadenylation had occurred (60). The discrete bands generated by splicing in the presence of 3’-dATP are shown in Fig. 1B, lane 2. For comparison, a control reaction carried out in parallel in the absence of 3’-dATP is shown in lane 1. Since 3’-dATP blocks both transcription and poly(A) tail growth, α-amanitin was also added to both reactions at 4 min so that transcription would not occur in either sample after 4 min. Thus, the only difference between lane 1 and lane 2 in Fig. 1B is the presence or absence of poly(A) tail growth.

The identities of the bands in Fig. 1B, lane 2, and the accuracy of the in vitro splicing were confirmed by sequencing in a scaled-up experiment, as shown in Fig. S1D in the supplemental material.

Splicing is appropriately coupled with 3’-end processing. Having established that splicing is efficient and accurate under conditions in which 3’-end processing is coupled to transcription, we next wanted to confirm that the known interconnections between these two processing reactions also are faithfully reproduced under the same coupling conditions. Coupling between splicing and 3’-end processing in vitro is delicately dependent on experimental conditions (13). Nevertheless, as pointed out earlier, to the extent that the coupling of 3’-end processing to transcription in our system is physiologically relevant, we expected the same system to support coupling between the two processing reactions as well.

Under appropriate conditions, in reactions not coupled to transcription, the poly(A) signal participates in defining terminal exons for the purpose of splicing and, in so doing, enhances
the efficiency of last-intron removal (13, 56). We sought to confirm this for the last intron of the β-globin transcripts in our coupled system by comparing the efficiency of second-intron splicing for transcripts with wild-type or mutant poly(A) signals (Fig. 2A).

A technical issue that arises is that the splicing status of transcripts with mutant poly(A) signals cannot be inferred directly from mobilities in a gel. This is because, without a functional poly(A) signal, such transcripts lack a defined 3′ end to serve as the basis for correlating transcript length with splicing status (e.g., see transcripts diagrammed on the right of the gel in Fig. 1B). To overcome this limitation, we purified the RNA and then cut it with RNase H in the presence of a DNA oligonucleotide that targeted the cutting to the poly(A) site. This ensured that even mutant transcripts would have 3′ ends at the position of the poly(A) site at the time of analysis (see, e.g., Fig. 2A, lane 4). For consistency, this step was also carried out for the purified wild-type transcripts so that they as well would all terminate uniformly at the poly(A) site, regardless of whether the 3′ end was generated by processing in the extract or through postcutting by RNase H.

The gel in Fig. 2A confirms that transcripts with a mutant poly(A) signal (lanes 4 to 6) accumulate less second-intron-spliced RNA (bands 3 and 4) than transcripts with a functional poly(A) signal do (lanes 1 to 3). At the same time, the singly spliced second-intron-containing intermediate, band 2, accumulates for the mutant to become the most prominent species present in lane 6 at 60 min (however, relative to that in lane 5, the accumulation of this intermediate in lane 6 is not apparent because of reduced recovery and less-efficient RNase H postcutting for the sample in lane 6). Quantitative analysis confirms the qualitative observations described above. Graphs 1 and 2 in Fig. 2A show that a functional poly(A) signal dramatically enhances second-intron splicing (for which it can help define the terminal exon) while having little effect on splicing of the first intron. Graph 3 shows that, conversely, mutating the poly(A) signal causes accumulation of the singly spliced second-intron-containing intermediate up to nearly 50% of the total, reflecting the loss in ability to splice the second intron. In contrast, for the wild-type poly(A) signal, the singly spliced second-intron-containing intermediate never exceeds 15% of the total, because this species is essentially in steady state, with the rate of its production by first-intron splicing being matched by the rate of its removal via second-intron splicing.

The experiment described above was for transcripts carrying the SV40 late poly(A) signal. Coupled transcription, splicing, and 3′-end processing for transcripts with a β-globin poly(A) signal gave similar results (see Fig. S2A in the supplemental material). Minor differences in the case of the β-globin poly(A) signal reflect the fact that it is weaker than the SV40 late poly(A) signal. For example, in the case of the β-globin poly(A) signal, there is relatively more band 2 splicing intermediate but less band 3 because of reduced enhancement of second-intron splicing.

Just as the poly(A) signal enhances splicing by promoting definition of the terminal exon, so also its partner in exon definition, the 3′ splice site, reciprocates by enhancing cleavage at the poly(A) site (20, 57). To confirm this aspect of coupling in our system, we determined the effect of a 3′-splice site AG→CT mutation on poly(A) site cleavage. Figure 2B shows that mutation of the second-intron 3′ splice site reduces the efficiency of poly(A) site cleavage by about a factor of three (see the percentages given at the bottom of Fig. 2B). This is most apparent in the gel at the earliest time point (lanes 1 and 4), before splicing occurs, when the only discrete species present is the poly(A) site-cleaved, but unspliced, precursor. Thus, in comparison to the wild type (Fig. 2B, lane 1), when the 3′ splice site is mutated (lane 4), there is less poly(A) site-cleaved RNA in band 1 and more uncleaved RNA higher in the gel. Moreover, as pointed out previously (2, 57), since enhancement of poly(A) site cleavage precedes splicing (the RNA of band 1 in Fig. 2B lanes 1 and 4 is unspliced), this result confirms that the enhancing activity of the 3′ splice site occurs during recruitment of the splicing factors (48) and before the chemical steps of splicing take place.

Of course, the enhancing activity of the 3′ splice site on poly(A) site cleavage also occurs, necessarily, before the chemical step of poly(A) site cleavage. Since cleavage at the poly(A) site occurs with a lag of only 3 to 5 min after transcription of the third exon (60; unpublished data), this means that recognition of the 3′ splice site and establishment of the connections to the poly(A) signal occur rapidly (<5 min) in the wake of transcription.

Results shown in Fig. 2A and B confirmed that second-intron splicing and cleavage at the poly(A) site are strongly coupled to each other. Next, we wanted to confirm that splicing of the first intron requires definition of the second exon, as would be expected for vertebrate splicing (5, 24, 66), especially when coupled to transcription (74). According to the exon definition model, internal exons must be defined to allow splicing of both flanking introns (5). Therefore, in our construct, mutation of the 5′ splice site of the second intron should result in the inhibition of both first- and second-intron splicing as, in fact, is observed for the human β-globin gene in vivo (68). Lanes 5 and 6 of Fig. 2C show that both first- and second-intron splicing are indeed reduced dramatically (compared to lanes 2 and 3) when the 5′ splice site of the second intron is mutated. Thus, bands 3 and 4 (the two second-intron-spliced species) are completely absent because there is no 5′ splice site for the second intron, and there is no corresponding accumulation of first-intron-spliced material (band 2) because first-intron splicing is also severely reduced, owing to the lack of second-exon definition. Instead, the second exon is skipped, and the first exon is ligated directly to exon 3 to yield band 5 (lanes 5 and 6). This shows that mutation of the second-intron 5′ splice site has no effect on the definition of exon 3, whose 3′ splice site and poly(A) signal both remain intact. Moreover, the 3′ splice site remains fully capable of enhancing poly(A) site cleavage (compare lanes 1 and 4).

The experiments described above were for transcripts with the SV40 late poly(A) signal. Similar results, though less pronounced, were obtained for transcripts with the weaker β-globin poly(A) signal (see Fig. S2B and C in the supplemental material).

3′-end processing is coupled to transcription. We have previously shown that the functional coupling of transcription and 3′-end processing in vitro is sustained by the tether of nascent RNA that extends from the poly(A) signal to the active site of the polymerase downstream (Fig. 3A) (60). This tether, which is also required for efficient 3′-end processing in vivo (6), is a
FIG. 2. Connections between splicing and 3'-end processing. (A) The poly(A) signal enhances second-intron splicing. The percent splicing was calculated as the sum of bands 3 and 4 (graph 1), bands 2 and 4 (graph 2), or band 2 alone (graph 3) expressed as a percentage of the sum of all the bands (i.e., 1 through 4). The averages and ranges for two independent experiments are shown. The isolated RNA was postcut at the poly(A) site by using RNase H and oligonucleotide 5. The cutting efficiency for lanes 1 to 5 was 72% ± 2% (average ± standard deviation), but that for lane 6 was only 50%, which exaggerates the impression of RNA loss for this lane. RNA recovery at 60 min compared to 10 min for all data is summarized in panel A and Fig. S2A in the supplemental material was 64% ± 9% (average ± standard deviation). (B) The 3' splice site of the second intron enhances 3'-end processing. No RNase H was used in this experiment. (C) First-intron splicing requires second-exon definition. No RNase H was used in this experiment.
FIG. 3. Poly(A) site cleavage is inhibited by cutting the tether, but cleavage at the poly(A) site is not required for the enhanced second-intron splicing conferred by the poly(A) signal. (A) Cartoon of the core poly(A) site cleavage complex showing the RNA tether, together with diagrams of the construct and the transcription protocol used. (Cartoon adapted from references 54 and 60 with permission.) (B and C) These experiments were all done using extract A, but for the experiment corresponding to panel C, the protocol for extract B was used. The precutting oligonucleotide (77 oligo) directs RNase H (72) to cut the tether during transcription within the vector sequence predominantly at a position about 77 nucleotides downstream of the SV40 or β-globin poly(A) cleavage sites. The control oligonucleotide (lanes 1 to 5) is the complement to 77 oligo. Splicing of the RNase H-cut RNA in lanes 6 to 15 is quantitated in Fig. S5 in the supplemental material.
unique feature of the transcription elongation complex, and establishing its role in processing is direct evidence for a functional connection between processing and transcription. Using templates that lack introns, we previously showed that cutting this tether with RNase H during transcription in vitro severely impairs 3'end processing (60).

To determine whether the tethering criterion for coupling applies to constructs which also splice, we carried out a tether-cutting experiment using the same splicing template as that described for Fig. 1 and 2. Transcription was carried out in the presence of a DNA oligonucleotide whose sequence was chosen to direct RNase H to cut the RNA downstream of the poly(A) signal before assembly of the processing complex is complete (called precutting; Fig. 3A) (60). As in previous experiments, we added 3' dATP during the processing phase of the reaction to prevent poly(A) tail growth and thereby ensure sharp bands on the gel.

The first five lanes of Fig. 3B show the results of a control reaction in which the added DNA oligonucleotide was not complementary to any part of the transcript. Just as for Fig. 1 and 2, one can see the initial production of 3'-end processed precursor in the presence of this control oligonucleotide (band pA1 in lane 1 of Fig. 3B) followed by splicing, eventually to yield fully spliced and poly(A) site cleaved RNA as the predominant species (band pA4 in lane 5). In the following two paragraphs, we focus on the role of the tether in the initial poly(A) site cleavage reaction only; we later return to this figure to consider the role of the tether in the splicing that occurs subsequently.

In the reactions corresponding to lanes 6 to 15 in Fig. 3B, a DNA oligonucleotide directed RNase H to cut 77 nucleotides downstream of the poly(A) site. In these lanes, a strong band of RNase H-cut RNA appears (band H1) in place of the broad smear of longer material otherwise present higher in the gel (lanes 1 to 5). This precut RNA, cut by RNase H in vector sequences downstream of the cloned poly(A) signal (Fig. 3A), is clearly impaired in its ability to support subsequent cleavage at its poly(A) site. This is easiest to see by comparing the amount of unspliced, poly(A) site-cleaved precursor (band pA1) in lanes 1 to 3 of Fig. 3B with the amount in lanes 6 to 8. Band pA1 in lanes 1 to 3 is substantial, but this band is much reduced beneath the RNase H-cut RNA (band H1) in lanes 6 to 8. At later time points, the effect of cutting the tether can be seen by looking at the fully spliced species. While band pA4 in lanes 3 to 5 of Fig. 3B is fairly prominent, this band in lanes 8 to 10 is much diminished. These results confirm that splicing-competent transcripts require a tether for coupled 3'-end processing, just as do the nontron transcripts studied previously (60).

The template used for Fig. 3B carries the SV40 late poly(A) signal. Figure 3C shows that running the same experiment using a template that retains the original, but weaker, β-globin poly(A) signal leads to essentially the same results: band pA1 is clearly present in lanes 1 to 3 but very weak to absent in lanes 6 to 8, and band pA4 is clearly present in lanes 3 to 5 (best seen in the enhanced panel of Fig. 3C) but very weak or absent in lanes 8 to 10. For completeness, a similar experiment was carried out on the β-globin gene with unshortened introns and unaltered sequences downstream of the poly(A) signal, again with equivalent results (see Fig. S3 in the supplemental material). We conclude that poly(A) site cleavage is functionally coupled to transcription for the splicing templates used here, just as for the nonsplicing templates used previously (60).

Splicing of the second intron is coupled to transcription.

Since a tether is required for efficient poly(A) site cleavage (Fig. 3; see Fig. S3 in the supplemental material) and a functional poly(A) signal is required for efficient second-intron splicing (Fig. 2A; see Fig. S2A in the supplemental material), we wondered whether severing the tether downstream of the poly(A) site would also impair second-intron splicing. To answer this, we needed to carry out a tether-cutting experiment similar to that illustrated in Fig. 3 but in a way that detects all spliced RNA in the population regardless of whether 3'-end processing has occurred. In particular, we wanted to ensure detection of any spliced species contained within the non-3'-processed RNA of control samples like those shown in lanes 1 to 5 in Fig. 3B and C. To achieve this, we adopted the approach used in Fig. 2A, in which, after completing the experiment but before loading the purified RNA on the gel, we postcut all of the RNA at the poly(A) site by using RNase H. This facilitated a straightforward analysis of the efficiency of splicing for the entire population by ensuring a common 3' end at the poly(A) site for all species in the gel, whether the end was generated by RNase H or by the 3'-end processing machinery.

Figure 4A shows that second-intron splicing is indeed coupled to transcription via the tether. Transcription and processing were carried out in the presence of a control oligonucleotide that was not complementary to any part of the transcript (Fig. 4A, lanes 1 to 5). In lanes 6 to 10, the same oligonucleotide used for the experiments whose results are shown in Fig. 3 was included in the reaction to direct RNase H cutting to a position 77 nucleotides downstream of the poly(A) site for all species in the gel, whether the end was generated by RNase H or by the 3'-end processing machinery.

Figure 4A shows that second-intron splicing is indeed coupled to transcription via the tether. Transcription and processing were carried out in the presence of a control oligonucleotide that was not complementary to any part of the transcript (Fig. 4A, lanes 1 to 5). In lanes 6 to 10, the same oligonucleotide used for the experiments whose results are shown in Fig. 3 was included in the reaction to direct RNase H cutting to a position 77 nucleotides downstream of the poly(A) site for all species in the gel, whether the end was generated by RNase H or by the 3'-end processing machinery.

Figure 4B shows that second-intron splicing is indeed coupled to transcription via the tether. Transcription and processing were carried out in the presence of a control oligonucleotide that was not complementary to any part of the transcript (Fig. 4A, lanes 1 to 5). In lanes 6 to 10, the same oligonucleotide used for the experiments whose results are shown in Fig. 3 was included in the reaction to direct RNase H cutting to a position 77 nucleotides downstream of the poly(A) site for all species in the gel, whether the end was generated by RNase H or by the 3'-end processing machinery.

Figure 4A shows that second-intron splicing is indeed coupled to transcription via the tether. Transcription and processing were carried out in the presence of a control oligonucleotide that was not complementary to any part of the transcript (Fig. 4A, lanes 1 to 5). In lanes 6 to 10, the same oligonucleotide used for the experiments whose results are shown in Fig. 3 was included in the reaction to direct RNase H cutting to a position 77 nucleotides downstream of the poly(A) site for all species in the gel, whether the end was generated by RNase H or by the 3'-end processing machinery.

Figure 4A shows that second-intron splicing is indeed coupled to transcription via the tether. Transcription and processing were carried out in the presence of a control oligonucleotide that was not complementary to any part of the transcript (Fig. 4A, lanes 1 to 5). In lanes 6 to 10, the same oligonucleotide used for the experiments whose results are shown in Fig. 3 was included in the reaction to direct RNase H cutting to a position 77 nucleotides downstream of the poly(A) site for all species in the gel, whether the end was generated by RNase H or by the 3'-end processing machinery.
FIG. 4. A tether is required for efficient second-intron splicing when the β-globin poly(A) signal defines the terminal exon. (A) This experiment was done as described in the legend to Fig. 3C except that transcripts were postcut at the poly(A) site with RNase H by using oligonucleotide 7 before loading onto the gel. (B and C) Quantitations are as described in the legend to Fig. 2A. Averages and ranges are for two independent experiments, including the experiment shown in panel A. wt, wild type; mt, mutant; 77, 77 oligo.
spliced species illustrated by line 1 in Fig. 4B. First, there is no consistent pattern of RNA loss in experiments of this type, and when there is loss, it affects the samples in the presence of the control oligonucleotide and those in the presence of 77 oligo indiscriminately. Second, even when the results for the particular experiments of Fig. 4B are recalculated according to worst-case assumptions for the effects of RNA loss, the negative effect of tether cutting on second-intron splicing persists (see Fig. S4A in the supplemental material).

In contrast to second-intron splicing, first-intron splicing is little affected by tether cutting (Fig. 4C). This does not rule out a tether requirement for first-intron splicing, however. The first intron lies far upstream of where the tether is cut in these experiments, so we think it most likely that commitment to splicing has already occurred by the time cutting takes place. Importantly, the continued uninhibited splicing of the first intron after RNase H cutting (Fig. 4C) shows that the inhibition of second-intron splicing (Fig. 4B) and inhibition of poly(A) site cleavage (Fig. 3; see Fig. S3 in the supplemental material) (60) by tether cutting are direct effects on these processing events, not a consequence of some general damage to the transcripts.

Finally, note that in these experiments the tether is cut at a position that lies within the cloning vector. Therefore, the inhibitory effect of cutting is on the integrity and function of the transcription-processing complex and not on any putative functional sequence in the RNA.

The poly(A) signal enhances splicing and commits to poly(A) site cleavage in separate steps. Figure 4B not only establishes that second-intron splicing is coupled to transcription via the tether but it also suggests the existence of at least two separate steps in the assembly of the cleavage and polyadenylation apparatus. Thus, just as Fig. 2B shows that enhancement of poly(A) site cleavage by the 3′ splice site occurs before splicing (compare lane 1 and lane 4), see also Fig. 4B argues that the enhancement of second-intron splicing by the poly(A) signal is based on interactions that are established before poly(A) site cleavage. This conclusion follows from the observation that in spite of the inhibition caused by precutting with RNase H, second-intron splicing still remains more efficient in the presence of a functional poly(A) signal than in its absence (Fig. 4B, compare lines 2 and 3). This observation shows that, even after most poly(A) site cleavage has been blocked by the uncoupling effect of cutting the tether (Fig. 3; see Fig. S3 in the supplemental material) (60), interactions established by the poly(A) signal in an earlier step, before commitment to cleavage, remain capable of enhancing splicing on these RNase H-cut transcripts.

Closer analysis of Fig. 3B and C, for which the corresponding experiments did not include postcutting by RNase H, establishes directly the two separate events in 3′-end processing complex assembly inferred from Fig. 4B, first the enhancement of second-intron splicing and then the commitment to 3′-end cleavage. The ability of the poly(A) signal to enhance second-intron splicing without actually cleaving at the poly(A) site can be seen in both Fig. 3B and C by comparing band H4 in lanes 8 to 10 with band H4 in lanes 13 to 15 (best seen in the enhanced panel). Band H4 consists of fully spliced RNA and an uncleaved poly(A) signal that is impaired in its ability to cleave because the tether has been cut [a comparison of band pA4 in lane 10 with that in lane 5 shows that the tether-cut RNA in band H4 does not undergo efficient poly(A) site cleavage to yield band pA4]. In lanes 8 to 10, this uncleaved poly(A) signal is wild type, whereas in lanes 13 to 15, the signal is mutant. Band H4 in lanes 8 to 10 of Fig. 3B is fivefold stronger than the signal of band H4 in lanes 13 to 15 because in lanes 8 to 10, the wild-type SV40 poly(A) signal enhances second-intron splicing to produce the larger amount of fully spliced band H4 RNA. This effect is twofold for the weaker β-globin poly(A) signal shown in Fig. 3C. A complete quantitative analysis is presented in Fig. S5 in the supplemental material. These results show that a functional poly(A) signal can enhance second-intron splicing even without cleavage at the poly(A) site. We conclude that the ability of the poly(A) signal to enhance splicing can be uncoupled from its ability to commit to cleavage. This suggests, in turn, that the splicing enhancement conferred by the poly(A) signal occurs prior to the commitment to cleavage which is blocked by cutting the tether. This is consistent with the recent demonstration that enhancement of splicing does not require actual cleavage at the poly(A) site (36).

Recall from Fig. 4B that the tether is required for fully efficient second-intron splicing of the β-globin gene (left panel, compare line 2 to line 1), yet, as has just been discussed, the poly(A) signal remains capable of significant enhancement of second-intron splicing even after the tether has been cut and poly(A) site cleavage has been blocked (Fig. 3). This suggests that for some transcripts, the step in 3′-end processing apparatus assembly that enhances splicing occurred before the transcripts were cut by RNase H. If this interpretation is correct, it would predict that a poly(A) signal that is capable of more-rapid assembly of the 3′-end processing apparatus would be able to complete this enhancing step for a greater fraction of transcripts before RNase H cutting occurs. We tested this prediction by using the SV40 late poly(A) signal, which is an unusually strong one that has been shown to assemble the cleavage and polyadenylation apparatus more rapidly than other poly(A) signals in vivo (11). Quantitation of various data presented in this report shows that the SV40 late poly(A) signal is about three times as strong as the β-globin poly(A) signal (based on data in Fig. 2B and C, Fig. S2B and C in the supplemental material, and Fig. 3). The β-globin poly(A) signal is the one present in the construct used for the experiments illustrated in Fig. 4. When the experiments illustrated in Fig. 4 were repeated using the SV40 late poly(A) signal (Fig. 5A), there was, indeed, a substantially reduced effect of cutting the tether on second-intron splicing (Fig. 5B, left panel, compare line 2 with line 1). As expected, there was also a reduced effect on the accumulation of partially spliced band 2 material (Fig. 5B, right panel, line 2) which accumulates to a level much more similar to that of the wild-type poly(A) signal (line 1) than to that of the mutant (line 3).

Under our conditions, the tether is cut by RNase H on average about 2 min following extrusion of the poly(A) signal from the polymerase. Evidently, communication to the splicing apparatus from the strong SV40 late poly(A) signal is substantially complete within 2 min so that the enhancement of splicing is nearly undiminished. This same conclusion was reached when RNase H postcutting was directed to a different position in the transcript, upstream of the poly(A) site (see Fig. S4B in the supplemental material).
FIG. 5. RNase H cutting has much less effect on second-intron splicing when the SV40 late poly(A) signal defines the terminal exon. (A) This experiment was done as described in the legend to Fig. 3B except that transcripts were postcut at the poly(A) site with RNase H by using oligonucleotide 5 before loading onto the gel. (B and C) Quantitations are as described in the legend to Fig. 2A. wt, wild type; mt, mutant; 77, 77 oligo.
DISCUSSION

The in vitro system used here was first employed in studies of poly(A)-dependent transcriptional pausing (53, 67), then it was optimized for transcription-coupled 3'-end processing (60), and now we have shown that it robustly recapitulates all major activities of the “mRNA factory” (79), including splicing. Although we have not explicitly characterized capping, mammalian capping has been reproduced numerous times in vitro under a wide variety of experimental conditions (12, 32, 43, 51, 78), so we assume that it occurs normally here as well. The splicing efficiencies we report are comparable to those of recently reported transcription-processing systems optimized exclusively for splicing (15, 16, 26, 29, 39), and the cleavage/polyadenylation activity reported here is comparable to that of transcription-processing systems optimized exclusively for cleavage and polyadenylation (1, 47, 60, 73). Thus, it appears that no compromises have been made to the individual processing activities in order to drive both of them simultaneously, along with transcription, under the in vitro conditions reported here. We believe that this is an important indication of the physiological significance of the functional interactions observed by use of this system because in vivo, of course, all of these processes share the same conditions.

As emphasized previously (60), simply achieving high rates and efficiencies of concurrent reactions in vitro does not establish that the reactions are functionally coupled. To establish functional coupling, it is necessary to show that the individual reactions affect each other according to meaningful functional criteria. Figure 6A summarizes the functionally coupled relationships (arrows 1 to 5) established experimentally for the transcription/processing system described here. Four of these coupling interactions (arrows 1 to 4) involve second-intron splicing. Of these, the first three involve coupling between second-intron splicing and upstream (first-intron splicing) or downstream (cleavage/polyadenylation) processing reactions and have previously been described for systems that do not include concurrent transcription (5, 77). However, arrow 4 in Fig. 6A denotes a new coupling interaction in which second-intron splicing is linked directly to the integrity of the ternary transcription complex. The fifth coupling interaction in Fig. 6A is that between the 3'-end processing apparatus and the transcription complex, which occurs here, in the presence of splic-
ing, just as with the nonsplicing substrates described previously (60). We now consider individually each of the coupling interactions noted in Fig. 6A.

Arrow 1 in Fig. 6A refers to the disabling effect on first-intron splicing caused by mutating the 5′ splice site of the second intron (Fig. 2C). This shows that the 5′ splice site of the second intron couples first- and second-intron splicing via exon definition (5) in this in vitro system. Most transcription-splicing studies of which we are aware have been done under conditions of intron definition (15, 16, 26, 29, 39). However, exon definition is the predominant mode of vertebrate splicing (5, 24, 66, 74). According to the exon definition model, 3′ splice sites do not commit to splicing until the polymerase traverses the entire downstream exon, including its flanking 5′ splice site. Thus, for example, the splice sites flanking exon 2 in the globin pre-mRNA direct formation of an exon definition complex, mediated by U2AF, SR proteins, and the U1 snRNP, before being juxtaposed for splicing with the previously transcribed exon 1 upstream (5, 8). Importantly, exon definition is robust in our coupled system (Fig. 2C) even though the first intron is short and fully capable of intron definition (29, 39) and is transcribed in its entirety well before the downstream 5′ splice site of exon 2 emerges from the polymerase.

Arrow 2 in Fig. 6A refers to the impaired second-intron splicing caused by mutation of the poly(A) signal (Fig. 2A). This reflects participation of the poly(A) signal in second-intron splicing via its role in third-exon definition (56). This situation is analogous to the role of the downstream 5′ splice site in second-exon definition discussed above (5). Arrow 3 refers to the need for an intact 3′ splice site in the second intron to achieve full activity of the poly(A) signal (Fig. 2B). This confirms the expected coupling of second-intron splicing to 3′-end processing in this system (20, 57). The mutual enhancements of splicing and 3′-end processing summarized by arrows 2 and 3 reflect the joint participation of these two activities in definition of the third exon. The molecular interactions allowing this cooperation appear to be based on binding between the U2AF-U2 snRNP complex of the spliceosome and the CFIm-CPSF complex of the cleavage/polyadenylation apparatus (Fig. 6B) (36, 41, 48). Presumably, these four proteins constitute the core of the exon definition complex for exon 3 (Fig. 6B, red).

Arrows 2 and 3 of Fig. 6A reflect interactions that are established very early during transcription and before the processing events themselves occur, consistent with findings from previous work (36, 48). Thus, a 3′ splice site enhances poly(A) site cleavage on transcripts that have not been spliced (Fig. 2B, lane 1), and a poly(A) signal enhances second-intron splicing on transcripts that have not been cleaved (band H4 in Fig. 3B and C, lanes 8 to 10; see Fig. S5 in the supplemental material). The ability of the 3′ splice site to enhance poly(A) site cleavage in this coupled system shows that the splicing apparatus forges connections with the cleavage apparatus, not only prior to splicing but also prior to poly(A) site cleavage. Thus, in less than 5 min after transcription of the third exon, the 3′ splice site is recognized, the poly(A) signal is recognized, and the two establish functional connections with each other. Indeed, as pointed out in Results, the SV40 late poly(A) signal establishes connections with the 3′ splice site in less than 2 min.

Arrow 4 in Fig. 6A refers to the impairment of second-intron splicing that results from cutting the tether of nascent RNA between the poly(A) signal and the polymerase during transcription (Fig. 4B and 6B). This tether is a unique feature of the coupled state in which it physically links processing and transcription. The role of the tether has been discussed in detail previously in the context of 3′-end processing (60), for which it was shown to facilitate assembly of the nascent 3′-end processing complex by keeping the poly(A) signal close to the polymerase. If the tether is cut prior to stable attachment of the cleavage apparatus to the polymerase, then the attachment does not occur and cleavage does not take place (60). Similarly, in the present study, the tether apparently facilitates assembly during transcription of the exon definition complex across the third exon and, therefore, allows for efficient second-intron splicing (Fig. 4B).

Assembly of the exon definition complex is presumably initiated by factors associated with the 3′ splice site, such as U2AF and the U2 snRNP. U2AF65 is known to be associated with the polymerase (61), and the two probably interact directly (69). However, this binding appears to be weakened upon the interaction of U2AF65 with RNA (42, 69). In this context, the tether may serve to promote the 3′ splice site/ U2AF/polymerase association while also retaining the poly(A) signal in the same vicinity (Fig. 6B). This would facilitate recruitment of CFIm and CPSF to the nascent exon definition complex assembling on the polymerase (Fig. 6B). This scenario is consistent with various suggestions that CFIm and CPSF are among the first of the cleavage and polyadenylation factors to be recruited to the poly(A) signal (54, 64, 71). If the tether is cut before the necessary interactions have been established, then the polymerase cannot participate in the assembly of the exon definition complex, and splicing of the last intron is impaired.

This model predicts that second-intron splicing will be rescued from the adverse effects of tether cutting if a stronger poly(A) signal that can recruit the exon definition factors faster, before the tether is cut, is used. This is the result that was obtained in the experiment illustrated in Fig. 5B (and Fig. S4B in the supplemental material) when the strong SV40 late poly(A) signal was used in place of the β-globin poly(A) signal from the experiment illustrated in Fig. 4. This is reminiscent of in vivo data showing that strong poly(A) signals are faster than weak poly(A) signals to establish resistance to inactivation by antisense elements in the RNA (11). Hence, splicing gives the appearance of being mostly tether independent for strong poly(A) signals like the SV40 late signal (Fig. 5B) but largely tether dependent for weaker ones like the β-globin signal (Fig. 4B). Apparently, only about half the normal number of the slower β-globin poly(A) signals are able to contribute to the formation of an exon definition complex before the process is interrupted by the tether being cut with RNase H (Fig. 4).

Arrow 5 in Fig. 6A refers to the tether-cutting experiments illustrated in Fig. 3 (and Fig. S3 in the supplemental material) which confirm, as described previously (60), that the tether is required for efficient poly(A) site cleavage. More importantly, the tether-cutting experiments illustrated in Fig. 3 indicate that the two tether-dependent steps indicated by arrows 4 and 5 in Fig. 6A can be experimentally distinguished. This follows from the observation that although cutting the tether blocks poly(A) site cleavage (Fig. 3B and C and 6, arrow 5) (60), a significant
proportion of the RNase H-cut RNAs continue to exhibit poly(A) signal-mediated enhancement of second-intron splicing (Fig. 3B and C; see Fig. S5 in the supplemental material), a step which was shown in separate experiments also to be tether dependent (Fig. 4B, lane 2, and 6B, arrow 4). Thus, the tether-dependent step that defines the exon and enhances splicing (Fig. 6B, arrow 4) occurs at a distinctly earlier time than the tether-dependent step required for commitment to cleave at the poly(A) site (arrow 5). These results agree well with in vivo results showing that cleavage/polyadenylation complex assembly goes through at least two stages (11) and that cleavage/polyadenylation is more vulnerable than splicing to cutting of the tether (6).

The second of the two steps described above, commitment to poly(A) site cleavage, coincides with attachment of the cleavage/polyadenylation apparatus to the polymerase (60). This suggests that the nascent apparatus commits to cleavage (after having previously committed to definition of the terminal exon) as a composite exon definition/cleavage/polyadenylation apparatus attached to the polymerase CTD (Fig. 6B). This idea is consistent with the recent demonstration that Pcf11 (a component of CFIm) (Fig. 6B) is not required for enhancement of splicing but is required for cleavage at the poly(A) site (in a transcriptionally uncoupled system) (36). As illustrated by the RNA molecules of band H4 in Fig. 3B and C, the events that enhance splicing and that do not require Pcf11 (36) have already occurred in the immature cleavage/polyadenylation complexes whose cleavage is abrogated by cutting the tether. The CTD is an essential participant in poly(A) site cleavage (30), and Pcf11 is a polymerase CTD-binding protein (46, 75). Perhaps Pcf11 is recruited late to the maturing cleavage/polyadenylation apparatus (75, 76) to provide the final cleavage/polyadenylation factors at the promoter. Genes Dev. 17:1321–1327.


Fong, N., and D. L. Bentley. 2001. Capping, splicing, and 3′ processing are independently stimulated by RNA polymerase II: different functions for different segments of the CTD. Genes Dev. 15:1785–1793.


Lang, K. M., and R. A. Spritz. 1987. In vitro splicing pathways of pre-


