A model *in vitro* system for co-transcriptional splicing

Yong Yu, Rita Das, Eric G. Folco and Robin Reed*

Department of Cell Biology, Harvard Medical School, Boston, MA 02115, USA

ABSTRACT

A hallmark of metazoan RNA polymerase II transcripts is the presence of numerous small exons surrounded by large introns. Abundant evidence indicates that splicing to excise introns occurs co-transcriptionally, prior to release of the nascent transcript from RNAP II. Here, we established an efficient model system for co-transcriptional splicing *in vitro*. In this system, CMV-DNA constructs immobilized on beads generate RNAP II transcripts containing two exons and an intron. Consistent with previous work, our data indicate that elongating nascent transcripts are tethered to RNAP II on the immobilized DNA template. We show that nascent transcripts that reach full length, but are still attached to RNAP II, are efficiently spliced. When the nascent transcript is cleaved within the intron using RNase H, both the 5′ and 3′ cleavage fragments are detected in the bound fraction, where they undergo splicing. Together, our work establishes a system for co-transcriptional splicing *in vitro*, in which the spliceosome containing the 5′ and 3′ exons are tethered to RNAP II for splicing.

INTRODUCTION

In higher eukaryotes, RNAP II transcripts are typically thousands of nucleotides in length, containing numerous small exons separated by large introns (1–6). A long-standing problem in the splicing field has been to understand the mechanisms for ensuring that each upstream exon is ligated only to the adjacent downstream exon, without skipping any of the exons. Superimposed upon this problem, >95% of multi-exon transcripts undergo alternative splicing, in which different combinations of exons are used to generate a wide diversity of protein isoforms (5,6). A large body of work *in vivo* indicates that splicing occurs co-transcriptionally, prior to release of the nascent transcript from RNAP II [for reviews, see (2,7–10)]. Indeed, recent qPCR data showed that most constitutive exons are removed co-transcriptionally and that alternative exon choice also occurs during transcription (11). In addition, both *in vivo* and *in vitro* studies indicate that splicing is functionally coupled to RNAP II transcription (2,7–9). These results have raised the possibility that RNAP II plays a key role in correct pairing of adjacent exons in multi-exon transcripts. In one model, the nascent upstream exon is tethered to RNAP II until the adjacent downstream exon is synthesized. The two tethered exons are associated with one another via interactions within the spliceosome (12). In previous work *in vivo*, Dye and co-workers (13) obtained support for the RNAP II exon-tethering model by showing that a nascent RNAP II transcript that is co-transcriptionally cleaved within the intron is spliced as efficiently as the nascent transcript containing an intact intron. These data indicate that the 5′ and 3′ exons are tethered to one another via RNAP II (13).

Recently, *in vitro* systems for coupling transcription to splicing have been developed (14–18). In the coupled reaction, the nascent transcript is rapidly and quantitatively packaged into the spliceosome, which stabilizes the transcript and results in efficient splicing (14–17). However, at present, an efficient model *in vitro* system for co-transcriptional splicing, in which splicing occurs while the nascent transcript is still associated with RNAP II, remains to be established. Here, we combined our coupled transcription/splicing system with a well-established immobilized template assay in which emerging nascent transcripts remain attached to RNAP II bound on the immobilized DNA template [(19–26) and references therein]. Using this approach, we show that bound, full-length nascent transcripts undergo efficient splicing. Moreover, when RNase H is used to cleave the transcript within the intron, both the 5′ and 3′ portions of the transcript are detected in the bound fraction. We provide evidence that these 5′ and

*To whom correspondence should be addressed. Tel: +1 617 432 2844; Fax: +1 617 432 3091; Email: rreed@hms.harvard.edu

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

3′ portions undergo splicing while bound. Together, our data indicate that the spliceosome and the 5′ and 3′ exons associate (indirectly or directly) with RNAP II for co-transcriptional splicing.

MATERIALS AND METHODS

Plasmids

CMV-DoF1 plasmid was described previously (15). CMV-ATM (27) was a generous gift from M. Passoni (Trieste, Italy). RNAP II transcription templates were generated by PCR from each plasmid using a biotinylated 5′ primer (5′-biotin-TGG AGG TCG CTG AGT AGT GC-3′) or a biotinylated 3′ primer (5′-TAG AAG GCA CAG TCG AGG-biotin-3′). The biotin was linked to the oligonucleotide via a 15-atom spacer (TEG, Operon). The same oligonucleotides without biotin were used to generate non-biotinylated PCR products where indicated. The resulting PCR products were purified using phenol extraction followed ethanol precipitation and used as RNAP II transcription templates.

RNAP II transcription/splicing using immobilized CMV-DNA template

For *in vitro* RNAP II transcription/splicing reactions, 200 ng biotinylated CMV-DNA templates were bound to 10 μl of streptavidin magnetic particles (Roche) for 30 min and washed according to the manufacturer's protocol. The beads were incubated under RNAP II transcription/ splicing conditions at 30°C for the indicated times in 25 μl reaction mixtures containing 1 μl ^{32}P-UTP or CTP (800 Ci/mmol), 0.5 mM ATP, 3.2 mM MgCl$_2$ and 20 mM creatine phosphate (di–Tris salt) and 15 μl HeLa nuclear extract (28) (note that for nuclear extract preparation, the final centrifugation following dialysis is omitted, which results in more active extracts). Addition of extra NTPs does not affect the efficiency of transcription/splicing in these extracts, possibly because sufficient NTPs remain in the extract bound to proteins in an exchangeable manner. Actinomycin D (250 ng/25 μl reaction) was used to block transcription after 3, 5 or 10 min followed by continued incubation for times indicated in each figure. For PIC formation with PVA, the immobilized DNA template was incubated with 3.2 mM MgCl$_2$, 2 μl PVA and 15 μl HeLa nuclear extract at 30°C for 20 min, and then 20 mM creatine phosphate (di–Tris salt), 0.5 mM ATP and 1 μl ^{32}P-CTP (800 Ci/mmol) were added to initiate transcription and splicing. Bound and supernatant fractions were separated using a magnetic separator, and bound fractions were washed gently three times with 500 μl of immobilized wash buffer (20 mM HEPES pH 7.6, 20% glycerol, 0.2 mM EDTA, 0.1% Triton X-100 and 100 mM KCl). For RNase H cleavage, immobilized CMV-ATM was incubated under RNAP II transcription/splicing conditions for 3 or 5 min as indicated, followed by addition of actinomycin D and a 12-mer oligonucleotide (5′-AAGCT TCACCCG-3′) complementary to transcript. Incubation was continued at 30°C for indicated times. After incubation, bound and supernatant fractions were separated, and bound fractions were washed gently three times

with 500 μl of immobilized wash buffer containing 100 mM KCl. Total RNA was precipitated and analyzed on denaturing polyacrylamide gels. Bands were detected by phosphoimager.

RESULTS

To develop a model *in vitro* system for co-transcriptional splicing, we took advantage of an extensive body of work in the transcription field showing that ternary complexes containing the DNA template, RNAP II, and the nascent transcript assemble in nuclear extracts on DNA templates immobilized on a bead (18–26,29–32). For our work, we used a system in which the DNA template was attached to a magnetic streptavidin bead via a biotin moiety (24,25,30,31). To do this, a ~1.6 kb DNA template containing the CMV promoter and encoding the Ftz splicing substrate was amplified by PCR using primers containing either an upstream (5′b) or downstream (3′b) biotinylated nucleotide (Figure 1A). The same DNA template containing no biotin was used as a negative control (–b) (Figure 1A). Analysis of the PCR products by ethidium bromide staining (Figure 1B, lanes 1–3) or uniform labeling (Figure 1B, lanes 4–6) showed that one specific PCR product of the correct size was generated, and the levels of this product were similar for 5′b, 3′b and –b DNA templates. When these PCR products were mixed with streptavidin magnetic beads, the biotinylated, but not the non-biotinylated, DNA templates specifically bound to the beads (Figure 1C, lanes 1–9). Moreover, when the immobilized 5′b, 3′b and –b CMV-Ftz DNA templates were incubated in nuclear extract for 30 min under our standard RNAP II transcription/splicing conditions (15), the biotinylated DNA templates did not detach, as they were detected only in the bound and not in the supernatant fractions (Figure 1D, lanes 1–4). As expected, the –b DNA template was not detected in either fraction (lanes 5 and 6). Together, these data indicated that the 5′b and 3′b CMV-Ftz DNA templates were present in similar amounts, bound specifically to the streptavidin beads, and remained quantitatively bound to the beads under RNAP II transcription/splicing conditions.

To further characterize the biotinylated DNA templates, we carried out an RNAP II transcription/splicing reaction using the 5′b, 3′b or -b CMV-Ftz DNA templates free in solution. As shown in Figure 1E, bands of the expected sizes for the unspliced transcript (775 nt) and spliced transcript (608 nt) were detected with all of the templates, indicating that both RNAP II transcription (lanes 1, 3 and 5) and splicing (lanes 2, 4 and 6) occurred with similar efficiency with these templates [Note that U6 snRNA in the nuclear extract was labeled by the ^{32}P-UTP in the reaction mixture, as observed in previous work (33)]. We then carried out the same RNAP II transcription/splicing reaction using the immobilized DNA templates (Figure 1F). Importantly, these data revealed that both RNAP II transcription and splicing occur as efficiently using the immobilized DNA templates as with the DNA templates free in solution, indicating that the presence of the bead did not interfere
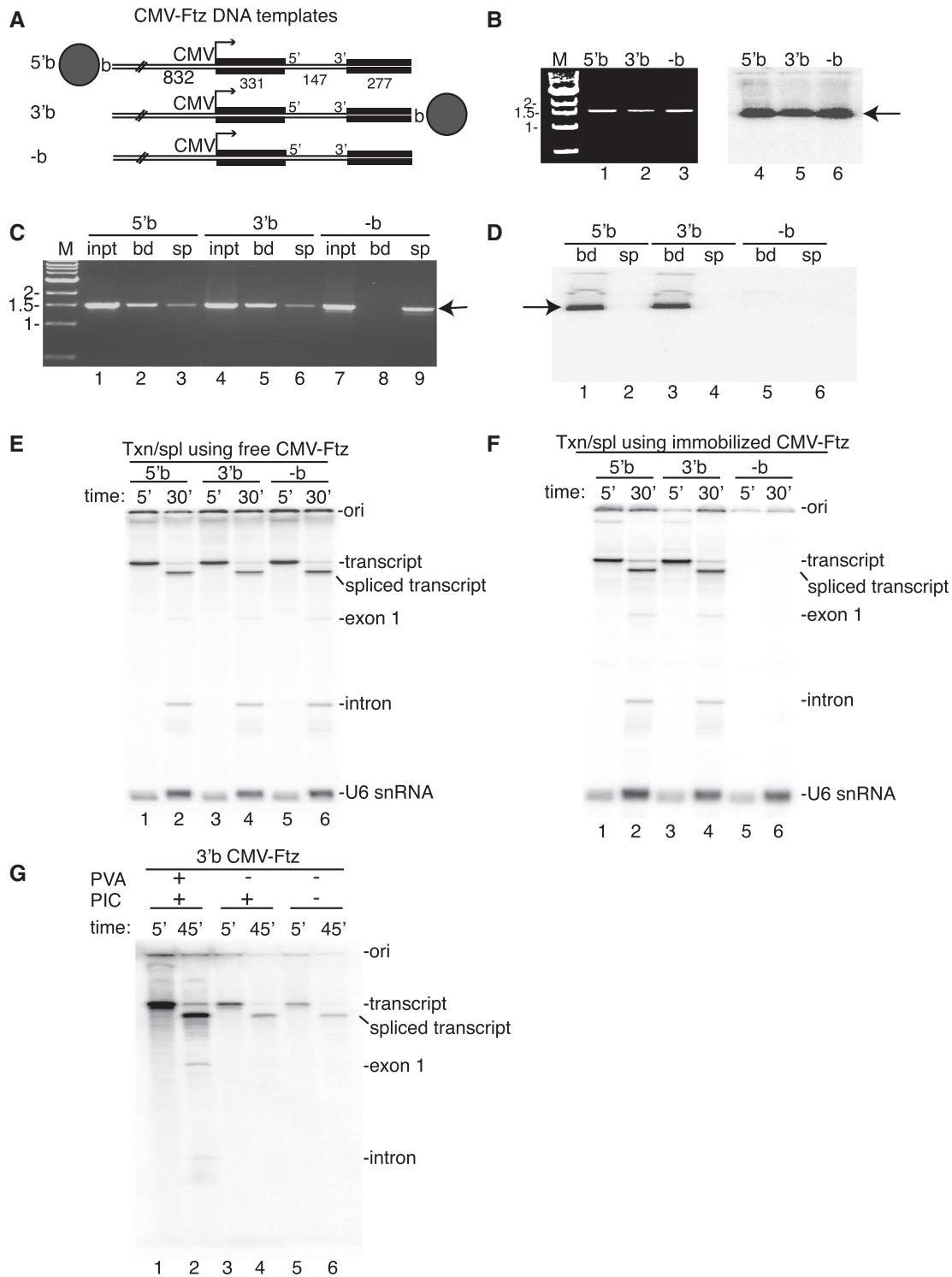
Figure 1. Efficient RNAP II transcription and splicing on immobilized CMV-Ftz DNA templates. (**A**) Schematic of CMV-Ftz DNA templates are shown, indicating the position of the biotin. The sizes of the CMV promoter and Ftz exons and intron are indicated. The gray circle represents the streptavidin magnetic bead. (**B**) CMV-Ftz DNA templates containing biotin at the 5′ (5′b) or 3′ end (3′b), or no biotin (-b) were constructed by PCR and detected on an ethidium bromide gel (lanes 1–3) or by uniform labeling (lanes 4–6). (**C**) Immobilized 5′b, 3′b and –b CMV-Ftz DNA templates were mixed with streptavidin magnetic beads, and after washing, bound (bnd) and supernatant (sup) fractions were analyzed on a 1% agarose gel stained with ethidium bromide. Total input is shown (lanes 1, 4 and 7). (**D**) Biotinylated CMV-Ftz DNA templates were uniformly labeled, bound to streptavidin magnetic beads, followed by incubation in nuclear extract at 30° for 30 min. Bound (1, 3 and 5) and supernatant (2, 4 and 6) fractions were analyzed on a 6.5% denaturing polyacrylamide gel. The arrow in panels B, C and D indicate the Ftz DNA template. Marker in panels B and C are in kilobases. (**E** and **F**) Free (E) or immobilized (F) CMV-Ftz DNA templates were incubated under RNAP II transcription/splicing conditions for 5 min, actinomycin D was added, and incubation was continued for another 25 min. Total RNA was fractionated on a 6.5% denaturing polyacrylamide gel. The splicing intermediates and products, tRNA, and U6 snRNA are indicated. Ori designates the gel origin. (**G**) PIC conditions with or without PVA were compared with standard RNAP II transcription/splicing conditions. Total bound (lane 1, 3 and 5) and supernatant (lane 2, 4, 6) fractions were analyzed on a 6.5% denaturing polyacrylamide gel.

with the reactions (compare Figure 1E and F, lanes 1–4). As expected, no transcript was detected with the -b DNA template, because it did not bind to the streptavidin beads (Figure 1F, lanes 5 and 6). In previous immobilized template assay systems, formation of a pre-initiation complex (PIC) and/or inclusion of polyvinyl alcohol (PVA) in nuclear extract stimulated transcription or polyadenylation (25,34,35). Thus, we next compared these conditions in our transcription/splicing system using 3′b immobilized CMV-Ftz. As shown in Figure 1G, PIC formation enhanced the levels of Ftz unspliced and spliced transcript (lanes 3 and 4) relative to our standard conditions (lanes 5 and 6), and PVA provided a further increase (lanes 1 and 2). Thus, PIC conditions with PVA were used to establish the co-transcriptional splicing system.

In prior studies [(19,22–26,35,36) and references therein], nascent transcripts were shown to remain quantitatively attached to RNAP II if the transcription reaction was stopped prior to RNAP II 'running off' the end of the DNA template. By definition, for 'co-transcriptional splicing', the splicing reaction cannot take place *while* the transcript is being synthesized because both exon 1, the intron, and exon 2 must be present in the transcript in order to undergo splicing. Thus, we next carried out a transcription time course using immobilized 5′b or 3′b CMV-Ftz DNA template (Figure 2A and C) to identify conditions in which the nascent transcript was mostly full length (and therefore a functional splicing substrate), but still attached to RNAP II. As shown in Figure 2B and D (lanes 1–6; see right panels for longer exposure), when the total RNA in the bound versus supernatant fractions was compared at the 30, 60 and 90 s time points, transcript ladders of increasing length were detected over time and were detected only in the bound fraction, indicating that transcribing RNAP II was associated with the immobilized DNA template [(25) and references therein]. At later time points, the levels of the full-length transcript
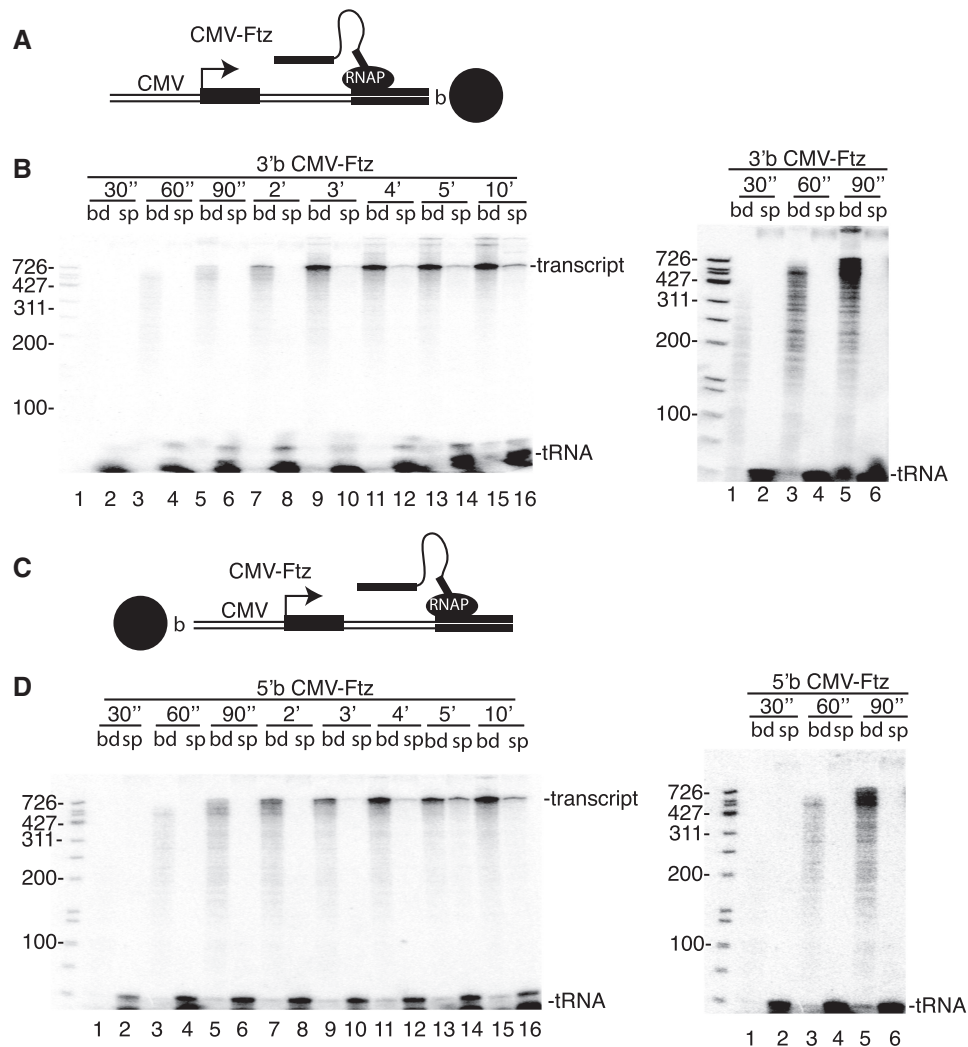

Figure 2. Transcription time course to analyze association of nascent transcripts with RNAP II bound to the DNA template. (**A** and **C**) Schematics showing transcription of 3′b (A) or 5′b (C) CMV-Ftz DNA template. (**B** and **D**) Immobilized 3′b (B) or 5′b (D) CMV-Ftz DNA template was incubated under PIC conditions with PVA for 30″, 60″, 90″, 2′, 3′, 4′, 5′and 10′. Total bound and supernatant fractions were separated, and the bound fractions were washed in 100 mM KCl, 0.1% triton. Bound (lanes 1, 3, 5, 7, 9 11, 13 and 15) and supernatant (lanes 2, 4, 6, 8, 10 12, 14 and 16) fractions were analyzed on a 6.5% denaturing polyacrylamide gel. (B, D) right. Longer exposure of lanes 1–6 for (B) and (D).

accumulated (Figure 2B and D, lanes 7–16). At the 2–10 min time points, transcripts were mostly detected in the bound fraction, but some run-off transcript was also detected in the supernatant at the later time points (Figure 2B and D, lanes 3–16). Thus, consistent with previous work, our data indicate that the nascent transcript remains associated with active RNAP II on the DNA template at early time points. We chose early (3–5 min) time points for studies of co-transcriptional splicing because we reproducibly found that most of the full-length transcript remained bound at this time point.

Co-transcriptional splicing *in vitro*

We next asked whether splicing could occur while the nascent transcript was associated with RNAP II on the immobilized DNA template. Accordingly, we assembled PICs on the immobilized 3′b CMV-Ftz DNA template, followed by transcription for 5 min, addition of actinomycin D (to stop further transcription), and then continued incubation for a total of 20 and 50 min to allow splicing. After washing the beads, total RNA in the bound and supernatant fractions was analyzed (Figure 3, lanes 1–6). As expected, at the 5-min time point, only unspliced transcript was detected, and it was present primarily in the bound fraction (Figure 3, lanes 1 and 2). In contrast, the tRNA present in the nuclear extract was detected only in the supernatant (Figure 3, lanes 1 and 2). Significantly, when incubation was continued, the splicing intermediates (lariat-exon 2 and exon 1) and spliced transcript were detected (Figure 3, lanes 3–6), and these RNA species were primarily present in the bound fraction (Figure 3, lanes 3–6). These data indicate that the nascent transcript can undergo splicing while associated with RNAP II on the DNA template. Thus, this system can be used as a simple *in vitro* model for co-transcriptional splicing. We also carried out the co-transcriptional splicing assay without adding actinomycin D. In this case, splicing began with the same kinetics as in the presence of actinomycin D (data not shown). However, at later times of incubation, transcription continues to occur, but splicing does not, possibly because splicing factors become limiting in the nuclear extract more quickly than transcription factors. In addition, more of the transcripts 'run-off' the end of the DNA template, as expected from the data shown in Figure 2. Thus, in the studies presented below, we used actinomycin D in our co-transcriptional splicing system. In a previous study, spliced transcripts were detected in association with an immobilized DNA template, but only when PCR was used for the assay (14). The difference between their study and ours could be that most of their transcript had run-off the DNA template.

Evidence that the 5′ and 3′ portions of the nascent transcript are associated with RNAP II after RNase H cleavage within the intron

In recent studies *in vivo*, Dye and co-workers (13) generated a transcript that could be co-transcriptionally cleaved within the intron by a hammerhead ribozyme. Analysis of splicing using this system revealed that the
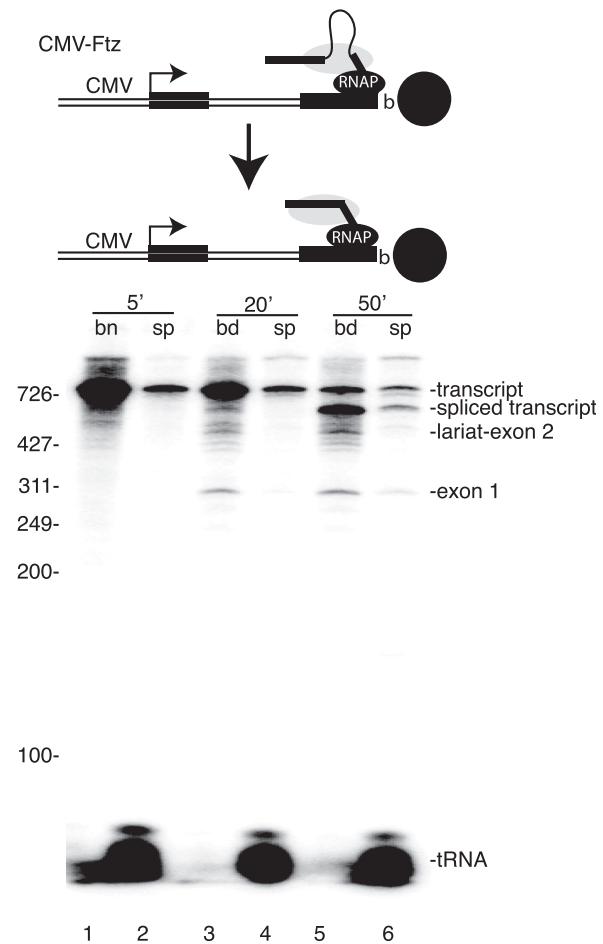

Figure 3. Evidence that nascent transcripts are spliced while tethered to RNAP II. Immobilized 3′b CMV-Ftz DNA template was incubated under PIC conditions with PVA for 5 min, actinomycin D was added, and incubation was continued for a total of 20 and 50 min. The bound and supernatant fractions were separated, and the bound fractions were washed in 100 mM KCl, 0.1% triton. Total RNA was isolated from the bound and supernatant fractions and run on a 6.5% denaturing polyacrylamide gel.

transcript was spliced with similar efficiency to that of an intact intron, providing evidence for a molecular tether between RNAP II and adjacent exons in the transcript (13). To investigate whether this type of molecular tether could be detected using our *in vitro* system, we employed oligonucleotide-directed RNase H digestion to cleave the nascent transcript within the intron. Initially, we attempted to use CMV-Ftz but encountered technical difficulties due to the small size of the intron. To circumvent this problem, we used a construct derived from the ATM gene, which contains a 339 nt intron (Figure 4A, see schematic) (27,37). As shown in Figure 4A, the CMV-ATM construct was transcribed by 3 min, and by 30 min splicing intermediates and products were detected (spliced transcript is 428 nt) (lanes 1 and 2). The intron-containing species were identified using oligonucleotide-directed RNase H cleavage with an oligonucleotide complementary to the intron (data not shown). When the CMV-ATM construct was immobilized on beads,
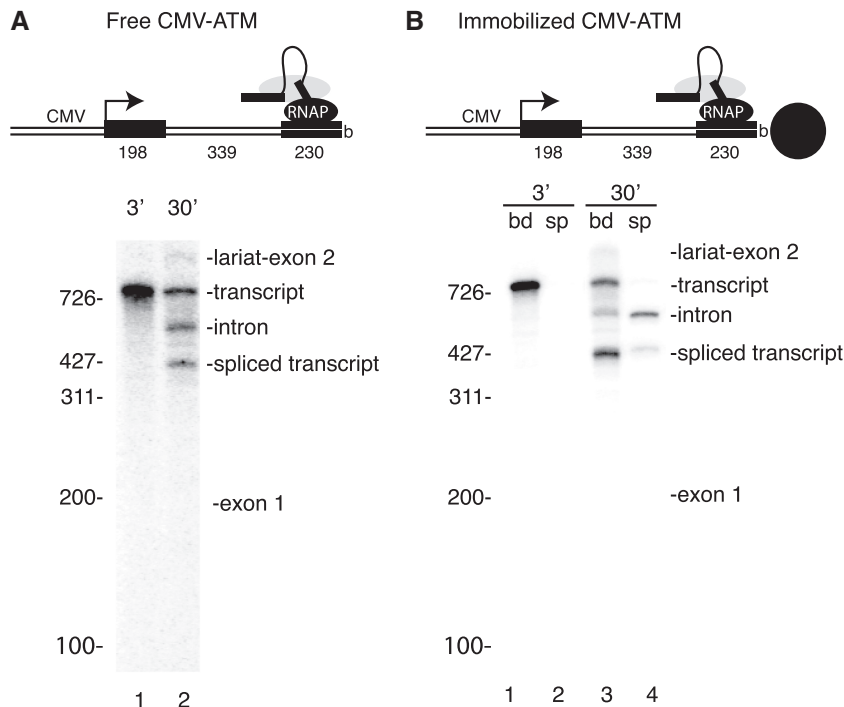
Figure 4. Splicing of free and immobilized CMV-ATM DNA template. (**A**) Free CMV-ATM DNA template was incubated under RNAP II transcription/splicing conditions for the times indicated and analyzed on a 6% denaturing polyacrylamide gel. The unspliced transcript, splicing intermediates and products are indicated. (**B**) 3′b CMV-ATM DNA template was immobilized on streptavidin magnetic beads and then incubated under PIC conditions with PVA. Total bound (lanes 1 and 3) and supernatant (lanes 2 and 4) fractions were analyzed on a 6% denaturing polyacrylamide gel.

unspliced ATM transcript, lariat-exon 2, exon 1 and the spliced transcript were largely detected in the bound fraction, indicating that, like Ftz, ATM transcripts are spliced while attached to RNAP II on the immobilized DNA template (Figure 4B, lanes 1–4). The lariat intron was largely detected in the supernatant, indicating that it is not tethered to the RNAP II/spliceosome complex (the lariat was identified by cleavage with RNase H, data not shown).

To determine whether the ATM transcript could be cleaved within the intron by RNase H, we used an oligo-nucleotide that should generate a 5′ fragment of ∼228 nt and a 3′ fragment of ∼539 nt (Figure 5A). As shown in Figure 5B, the ATM transcript (lane 1) was cleaved to completion to generate the expected 5′ and 3′ fragments after 5 (lane 2) or 10 (lane 3) min of incubation with the oligonucleotide. We next investigated the possibility that the cleaved 5′ or 3′ fragments remained associated with RNAP II bound on the DNA template. To do this, the CMV-ATM DNA template was immobilized on beads, transcription was carried out for 5 min, followed by addition of the oligonucleotide for 5 min. After washing the beads, RNA in the total bound and supernatant fractions was analyzed. Significantly, this analysis revealed that not only the 3′ fragment but also the 5′ fragment were largely present in the bound fraction (Figure 5C, lanes 1–4), indicating association of these fragments with RNAP II on the DNA template.

In our system, only picogram levels of transcript are synthesized from 200 ng of the DNA template. Thus,

further studies are needed to optimize the efficiency of the system for biochemical analyses of the factors that function in co-transcriptional splicing and in tethering the 5′ portion of the transcript to RNAP II.

Evidence that the 5′ and 3′ portions of nascent transcripts undergo splicing when associated with RNAP II

In light of the observation that both the 5′ and 3′ RNase H cleavage fragments are present in the bound fraction, we next asked whether spliced transcript could be generated from these tethered fragments. To do this, the immobilized CMV-ATM DNA template was incubated in the RNAP II transcription/splicing system, cleaved with RNase H, and incubation was continued. The bound and supernatant fractions were then analyzed. As shown in Figure 5C, after PIC formation and 5 min of transcription, the nascent transcript was detected in the bound fraction (lanes 9 and 10). After cleavage with RNase H for 5 min, the 5′ and 3′ fragments were also largely present in the bound fraction (Figure 5C, lanes 11 and 12). Significantly, when incubation was continued for a total of 30 or 45 min, a band that co-migrated with the spliced transcript generated from intact unspliced transcript (Figure 5C, lanes 5 and 7) was generated from the RNase H cleavage fragments (lanes 13 and 15). RT–PCR of these samples revealed a band of the expected size for the spliced transcript (data not shown), consistent with conclusion that the cleaved 5′ and 3′ fragments underwent splicing. Moreover, as observed with the spliced product generated from the intact transcript (Figure 5C,
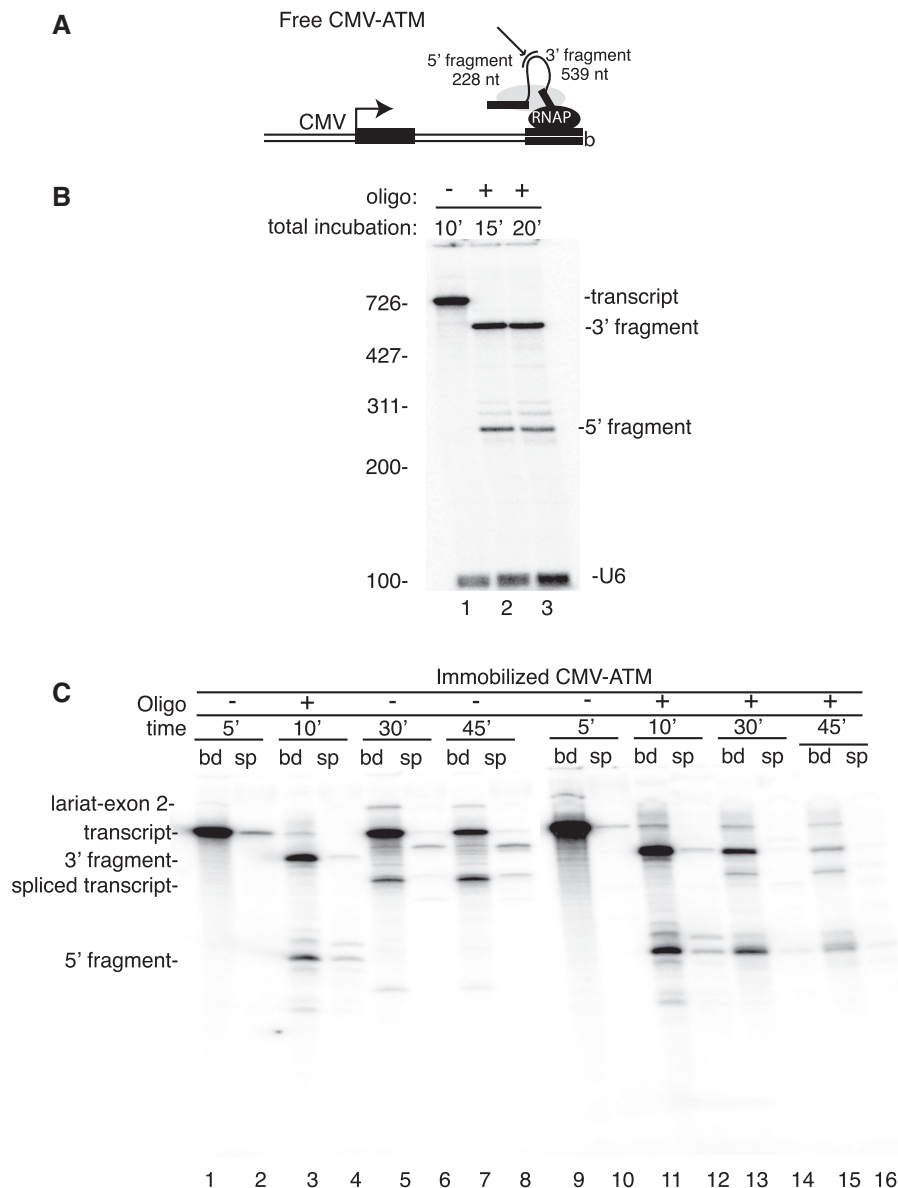
Figure 5. The 5′ and 3′ portions of the transcript cleaved within the intron by RNase H are tethered to RNAP II and undergo splicing. (**A**) Schematic showing 5′ and 3′ portions of the ATM transcript cleaved within the intron. The short line indicates the oligonucleotide used for RNase H cleavage, and the sizes of the 5′ and 3′ cleavage products are indicated. (**B**) 3′b CMV-ATM DNA template was incubated under RNAP II transcription/splicing conditions for 10 min (lanes 1). The oligonucleotide was then added and incubation was continued for 5 (lane 2) or 10 min (lane 3). The RNA was cleaved by the endogenous RNase H in the nuclear extract, and the 5′ and 3′ cleavage fragments are indicated. (**C**) 3′b CMV-ATM DNA template was immobilized on streptavidin magnetic beads and then incubated under PIC conditions with PVA for 5 min (lanes 1 and 2) followed by addition of actinomycin D and oligonucleotide for 5 min (lanes 3 and 4), or only actinomycin D for 25 min (lanes 5 and 6) and 40 min (lanes 7 and 8). Lanes 9–12 are the same as lanes 1–4. An aliquot of the sample shown in lane 11 was incubated further, for a total of 30 and 45 min, and then RNA from total bound and supernatant fractions was analyzed on a 5.75% denaturing polyacrylamide gel (lanes 13–16).

lanes 5–8), the spliced product generated from the RNase H-cleaved fragments was present only in the bound fraction (Figure 5C, lanes 13–16). Together, these data indicate that the tethered 5′ and 3′ portions of the transcript underwent splicing and that the spliced transcript remained tethered even after splicing. Although the efficiency of splicing of the RNase H-cleaved transcript was significantly less than that observed with the intact transcript (Figure 5C, lanes 5 and 7), this is due, at least in part, to degradation of the 5′ and 3′ fragments (Figure 5C,

lanes 13 and 15). Together, our data indicate that the 5′ and 3′ fragments of the cleaved transcript are tethered to RNAP II on the DNA template and undergo splicing while tethered.

DISCUSSION

In this study, we provided evidence that nascent transcripts are efficiently spliced *in vitro* while tethered to RNAP II on an immobilized DNA template. As splicing

occurs prior to release of the transcript from RNAP II, this study provides the first simple efficient model *in vitro* system for co-transcriptional splicing. Recently, an elegant study *in vivo* showed that transcription occurs at the rate of 3800 nt per minute and that co-transcriptional splicing occurs within 5–10 min of synthesis (11). In our *in vitro* system, both transcription and co-transcriptional splicing are about 5–10 times slower, possibly due to dilution of the extracts, limiting factors, or packaging of the DNA. Nevertheless, the observation that our rates are not that far off from those observed *in vivo* indicates that the *in vitro* system is a useful model for co-transcriptional splicing studies.

Most transcripts in higher eukaryotes contain numerous exons flanked by large introns. The exon-tethering model, in which the upstream exon is tethered to RNAP II until the downstream exon is synthesized, has been proposed to explain how exons are correctly joined to each other in sequential order. Previous work *in vivo* showed that that nascent transcripts are efficiently spliced when the intron is cleaved in the middle with a hammerhead ribozyme (13). Thus, these data suggested the existence of a molecular tether between the exons and RNAP II, supporting the exon-tethering model (13). This model has gained additional support by the observation that splicing *in vivo* occurs at similar rates whether the intron is 1 kb or >200 kb (11). This similarity in rates could occur if the upstream exon is tethered to RNAP II until the downstream exon is synthesized (11). Studies using our *in vitro* co-transcriptional splicing system also support the exon-tethering model and provide new insights. When we cleaved the tethered transcript within the intron using an oligonucleotide and RNase H, both the 5′ and 3′ fragments of the transcript were present in the bound fraction, indicating that they are associated with the RNAP II on the DNA template. Previous work indicates that the 3′ portion of the transcript is directly attached to the elongating RNAP II [(25) and references therein]. It is not known how the 5′ portion of the transcript is associated with RNAP II, i.e. whether the 5′ portion is directly tethered to RNAP II or associates with it indirectly via other factors, such as spliceosome components. Recently, Bentley and co-workers found that splicing *in vivo* did not occur if an intron was cleaved by a fast ribozyme (hepatitis δ) whereas splicing did occur if a slower cleaving hammerhead ribozyme was used (38). These data raise the possibility that spliceosome components are able to associate with the transcript when the slower ribozyme is used, and these spliceosome components may play a role in tethering the 5′ exon to RNAP II (38). Our studies in which we do detect splicing after cleaving the intron with RNase H appear to be more similar to the slower cleaving ribozyme work *in vivo* (13,38) and indicate that spliceosome components mediate the interaction between the RNase H-cleaved transcript and RNAP II. In a previous study, in which transcription and polyadenylation occurred on an immobilized template, Price and co-workers found that the 5′ portion of an RNase H-cleaved transcript was associated with the DNA template in a functional polyadenylation-competent complex (35). As this transcript lacked splice sites, polyadenylation factors likely mediate tethering of the 5′ fragment to RNAP II (35).

Our data show that the spliced transcript containing exons 1 and 2 remains tethered to RNAP II. Thus, it is possible that in multi-exon transcripts, all of the upstream spliced exons, or the upstream pair of spliced exons, remain tethered until the next exon is synthesized, thereby ensuring correct joining to the emerging downstream exon. Considering that the vast majority of RNAP II transcripts undergo alternative splicing, specific mechanisms must exist to ensure that correct ordering of exons occurs in the context of this greater level of complexity. In the simplest view, alternative exons would be looped out similar to intron sequences and thereby not associate with RNAP II. The *in vitro* co-transcriptional splicing system described here provides a new model for elucidating the mechanisms involved in co-transcriptional splicing.

ACKNOWLEDGEMENTS

FUNDING

REFERENCES

1. Black,D.L. (2003) Mechanisms of alternative pre-messenger RNA splicing. *Annu. Rev. Biochem.*, **72**, 291–336.
2. Kornblihtt,A.R., de la Mata,M., Fededa,J.P., Munoz,M.J. and Nogues,G. (2004) Multiple links between transcription and splicing. *RNA*, **10**, 1489–1498.
3. Pandit,S., Wang,D. and Fu,X.D. (2008) Functional integration of transcriptional and RNA processing machineries. *Curr. Opin. Cell Biol.*, **20**, 260–265.
4. Wang,E.T., Sandberg,R., Luo,S., Khrebtukova,I., Zhang,L., Mayr,C., Kingsmore,S.F., Schroth,G.P. and Burge,C.B. (2008) Alternative isoform regulation in human tissue transcriptomes. *Nature*, **456**, 470–476.
5. Nilsen,T.W. and Graveley,B.R. (2010) Expansion of the eukaryotic proteome by alternative splicing. *Nature*, **463**, 457–463.
6. Pan,Q., Shai,O., Lee,L.J., Frey,B.J. and Blencowe,B.J. (2008) Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.*, **40**, 1413–1415.
7. Proudfoot,N. (2000) Connecting transcription to messenger RNA processing. *Trends Biochem Sci*, **25**, 290–293.
8. Moore,M.J. and Proudfoot,N.J. (2009) Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell*, **136**, 688–700.
9. Bentley,D.L. (2005) Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. *Curr. Opin. Cell Biol.*, **17**, 251–256.

10. Wahl,M.C., Will,C.L. and Luhrmann,R. (2009) The spliceosome: design principles of a dynamic RNP machine. *Cell*, **136**, 701–718.

11. Singh,J. and Padgett,R.A. (2009) Rates of in situ transcription and splicing in large human genes. *Nat. Struct. Mol. Biol.*, **16**, 1128–1133.

12. Maniatis,T. and Reed,R. (2002) An extensive network of coupling among gene expression machines. *Nature*, **416**, 499–506.

13. Dye,M.J., Gromak,N. and Proudfoot,N.J. (2006) Exon tethering in transcription by RNA polymerase II. *Mol. Cell*, **21**, 849–859.

14. Ghosh,S. and Garcia-Blanco,M.A. (2000) Coupled in vitro synthesis and splicing of RNA polymerase II transcripts. *RNA*, **6**, 1325–1334.

15. Das,R., Dufu,K., Romney,B., Feldt,M., Elenko,M. and Reed,R. (2006) Functional coupling of RNAP II transcription to spliceosome assembly. *Genes Dev.*, **20**, 1100–1109.

16. Hicks,M.J., Yang,C.R., Kotlajich,M.V. and Hertel,K.J. (2006) Linking splicing to Pol II transcription stabilizes pre-mRNAs and influences splicing patterns. *PLoS Biol.*, **4**, e147.

17. Lazarev,D. and Manley,J.L. (2007) Concurrent splicing and transcription are not sufficient to enhance splicing efficiency. *RNA*, **13**, 1546–1557.

18. Rigo,F. and Martinson,H.G. (2008) Functional coupling of last-intron splicing and 3′-end processing to transcription in vitro: the poly(A) signal couples to splicing before committing to cleavage. *Mol. Cell. Biol.*, **28**, 849–862.

19. Maderious,A. and Chen-Kiang,S. (1984) Pausing and premature termination of human RNA polymerase II during transcription of adenovirus in vivo and in vitro. *Proc. Natl Acad. Sci. USA*, **81**, 5931–5935.

20. ten Heggeler,B. and Wahli,W. (1985) Visualization of RNA polymerase II ternary transcription complexes formed in vitro on a Xenopus laevis vitellogenin gene. *EMBO J.*, **4**, 2269–2273.

21. Cai,H. and Luse,D.S. (1987) Variations in template protection by the RNA polymerase II transcription complex during the initiation process. *Mol. Cell. Biol.*, **7**, 3371–3379.

22. Reines,D., Chamberlin,M.J. and Kane,C.M. (1989) Transcription elongation factor SII (TFIIS) enables RNA polymerase II to elongate through a block to transcription in a human gene in vitro. *J. Biol. Chem.*, **264**, 10799–10809.

23. Izban,M.G., Samkurashvili,I. and Luse,D.S. (1995) RNA polymerase II ternary complexes may become arrested after transcribing to within 10 bases of the end of linear templates. *J. Biol. Chem.*, **270**, 2290–2297.

24. Xie,Z. and Price,D. (1997) Drosophila factor 2, an RNA polymerase II transcript release factor, has DNA-dependent ATPase activity. *J. Biol. Chem.*, **272**, 31902–31907.

25. Adamson,T.E., Shore,S.M. and Price,D.H. (2003) Analysis of RNA polymerase II elongation in vitro. *Methods Enzymol*, **371**, 264–275.

26. Cheng,B. and Price,D.H. (2007) Properties of RNA polymerase II elongation complexes before and after the P-TEFb-mediated transition into productive elongation. *J. Biol. Chem.*, **282**, 21901–21912.

27. Pagani,F., Buratti,E., Stuani,C., Bendix,R., Dork,T. and Baralle,F.E. (2002) A new type of mutation causes a splicing defect in ATM. *Nat. Genet.*, **30**, 426–429.

28. Dignam,J.D., Lebovitz,R.M. and Roeder,R.G. (1983) Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.*, **11**, 1475–1489.

29. Marble,H.A. and Davis,R.H. (1995) RNA transcription from immobilized DNA templates. *Biotechnol. Prog.*, **11**, 393–396.

30. Ranish,J.A., Yudkovsky,N. and Hahn,S. (1999) Intermediates in formation and activity of the RNA polymerase II preinitiation complex: holoenzyme recruitment and a postrecruitment role for the TATA box and TFIIB. *Genes Dev.*, **13**, 49–63.

31. Choy,B. and Green,M.R. (1993) Eukaryotic activators function during multiple steps of preinitiation complex assembly. *Nature*, **366**, 531–536.

32. Kazerouninia,A., Ngo,B. and Martinson,H.G. Poly(A) signal-dependent degradation of unprocessed nascent transcripts accompanies poly(A) signal-dependent transcriptional pausing in vitro. *RNA*, **16**, 197–210.

33. Reddy,R., Henning,D., Das,G., Harless,M. and Wright,D. (1987) The capped U6 small nuclear RNA is transcribed by RNA polymerase III. *J. Biol. Chem.*, **262**, 75–81.

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

35. Adamson,T.E., Shutt,D.C. and Price,D.H. (2005) Functional coupling of cleavage and polyadenylation with transcription of mRNA. *J. Biol. Chem.*, **280**, 32262–32271.

36. Luse,D.S., Kochel,T., Kuempel,E.D., Coppola,J.A. and Cai,H. (1987) Transcription initiation by RNA polymerase II in vitro. At least two nucleotides must be added to form a stable ternary complex. *J. Biol. Chem.*, **262**, 289–297.

37. Lewandowska,M.A., Stuani,C., Parvizpur,A., Baralle,F.E. and Pagani,F. (2005) Functional studies on the ATM intronic splicing processing element. *Nucleic Acids Res*, **33**, 4007–4015.

38. Fong,N., Ohman,M. and Bentley,D.L. (2009) Fast ribozyme cleavage releases transcripts from RNA polymerase II and aborts co-transcriptional pre-mRNA processing. *Nat. Struct. Mol. Biol.*, **16**, 916–922.