Volume 16 Number 15 1988 Nucleic Acids Research

Molecular consequences of truncations of the first exon for *in vitro* splicing of yeast actin pre-mRNA

Michael Duchêne, Angela Löw, Andrea Schweizer and Horst Domdey*

Laboratorium für Molekulare Biologie, Genzentrum-der-Ludwig-Maximilians-Universität, München, D-8033 Martinsried, FRG

Received June 14, 1988; Accepted July 11, 1988

ABSTRACT

A defined minimum length of the first exon is required for the generation of spliced products from a synthetic yeast actin mRNA-precursor in vitro. If the first exon is 1, 2, 3 or 5 nucleotides long, only the first step of the splicing reaction can take place. A transcript starting with the first nucleotide of the intron does not get converted into any of the normally obtained splicing products or intermediates. On the other hand, spliceosome assembly does not depend on the presence of a first exon.

INTRODUCTION

The primary transcripts of many eukaryotic genes contain intervening sequences which are removed by RNA splicing. For the splicing reaction of pre-mRNAs a two-step model has been proposed (1, 2): In the first step the phosphodiester bond at the 5' splice site of the intron is cleaved and the 5' end of the intron becomes linked to a nucleotide in the 3' proximal region of the intron via a 2',5' phosphodiester bond, thus generating a lariat structure of the intron-plus-exon-2 intermediate. In the second step, the phosphodiester bond at the 3' splice site of the intron is cleaved and the exons are ligated.

Whereas a number of details has been collected in the past about minimum length requirements of intron sequences (3-7), only scarce information is available about the minimum length requirements of exon sequences for accurate splicing and splicing complex formation: Rymond and Rosbash reported that splicing of a yeast RP51 gene construct was not inhibited by a reduction of the second exon to 10 nucleotides (8). Even after deleting the entire 3' exon and approximately 15-22 nucleotides from the 3' end of the intron, the first step of the splicing reaction still took place. Conversely, a 3' exon with more than 24 nucleotides was required for accurate splicing of a human β-globin pre-mRNA in vitro (9), although 24 nucleotides (10), or even less when other mammalian substrates were used (11, 12), were sufficient for the formation of splicing complexes. Even less information is available on the sequence and minimum length requirements for exon 1: For splicing and spliceosome formation of a mammalian pre-mRNA, an exon 1 length of 20 nucleotides has been shown to be sufficient (12). Furthermore, the nucleotides preceding directly the intron of a mammalian pre-mRNA might play a role in a possible interaction with the U1-snRNP (13, 14). When

examining the sequences of intron containing yeast genes, we were not able to find an exon 1 smaller than 16 nucleotides. This finding might reflect that there is a minimum length required for efficient and accurate splicing of a yeast pre-mRNA.

MATERIALS AND METHODS

Generation of SP6-actin transcripts with reduced lengths of exon 1

An AluI fragment of the yeast actin gene (15), containing the entire intron of 309 nucleotides and parts of the flanking exon sequences, i.e., 73 nucleotides of exon 1 and 162 nucleotides of exon 2 (16), was cloned first into the unique SmaI site of the SP6 transcription vector Proteus6 (Fig. 1A). From this construct (Proteus6-act) a BamHI-EcoRI fragment was transferred into the transcription vector pSP64. An SphI-EcoRI fragment of this construct (spact 64), containing the SP6 promoter, part of the polylinker and the above described part of the actin gene, was finally cloned into M13mp18. Single stranded DNA from this construct (M13mp18act) was used as template for oligonucleotide directed mutagenesis according to Newman et al. (17). Several exon 1 deletion mutants (M13mp18actΔE1), leaving exon 1 sequences from 0 to 19 nucleotides, were created with the help of the oligonucleotides shown in Fig. 1B. All deletions were confirmed by DNA sequence analysis according to Sanger et al. (18). This set of template DNAs was linearized with EcoRI and transcribed in vitro according to Melton et al. (19) to yield transcripts with identical intron and exon 2 sequences but different lengths of exon 1 (Fig. 1C). Two sets of RNA precursors differing in the presence or absence of a 5' terminal cap structure were synthesized.

In vitro splicing reactions

The radioactively labeled capped or uncapped mRNA precursors were incubated in a soluble yeast splicing extract (20) under conditions as described by Lin et al. (20). 10 µl reactions were carried out at 23°C for 0, 15, 30, 45, and 60 minutes and were analyzed by electrophoresis in 8 M urea / 8% polyacrylamide gels.

Glycerol Gradient Sedimentation Analysis

Around 20,000 cpm of each transcript were subjected to a yeast splicing reaction in the presence or the absence of ATP in a total reaction volume of 30 µl for 15 min at 23°C, chilled on ice, and sedimented through a 15 to 40 % glycerol gradient for 16 h at 30,000 rpm in a Beckman SW41 TI rotor as described by Brody and Abelson (21). Gradients were fractionated from bottom to top in 400 µl aliquots, the radioactivity of each fraction was determined as Cerenkov cpm and the percentage of radioactivity per fraction was calculated.

RESULTS

In order to study the significance of the length of the first exon of a yeast mRNA-precursor for a yeast in vitro splicing reaction, several SP6 derived yeast actin pre-mRNA transcripts with exon 1 lengths from 0 to 19 nucleotides and identical intron and exon 2

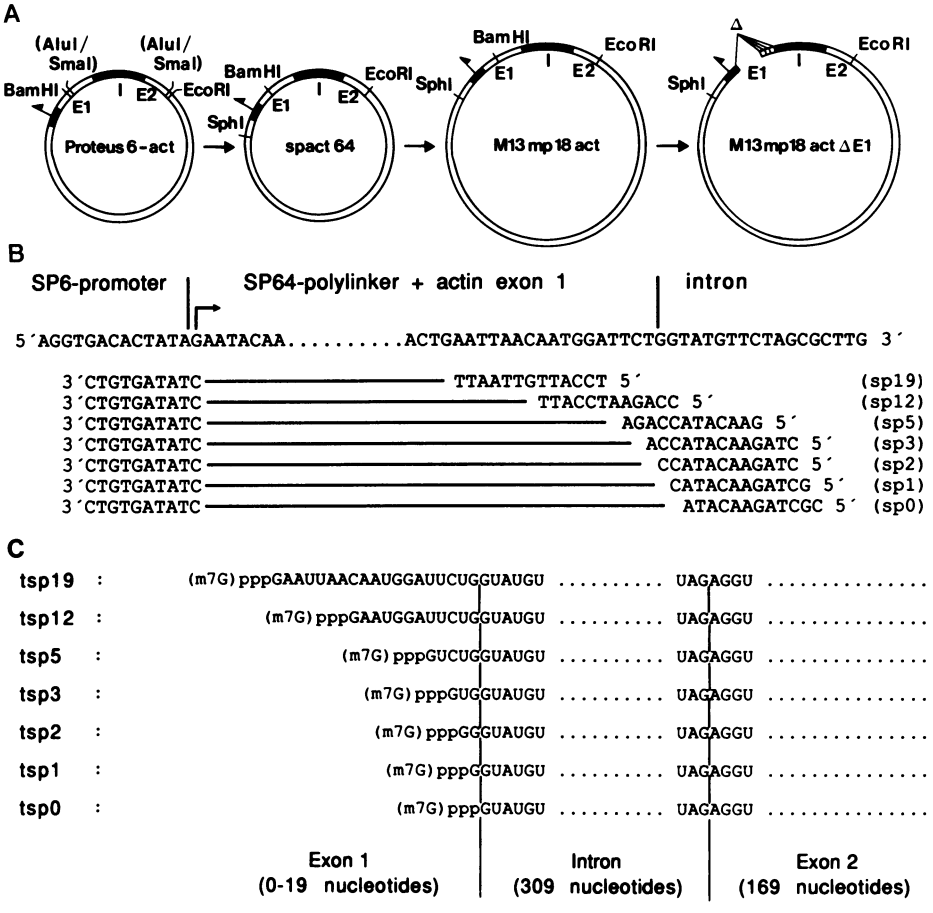

FIG. 1. Generation of SP6-actin transcripts with reduced lengths of exon 1.
A, construction of SP6-actin fusions for oligodeoxynucleotide directed mutagenesis and SP6 transcription (see Materials and Methods). E1, exon 1; I, intron; E2, exon 2.
B, sequence of the seven oligodeoxynucleotides (sp19 to sp0) that were used for deleting increasing parts of the first exon. The horizontal lines in the oligodeoxynucleotide sequences connect actually adjacent deoxynucleotides and are presented to show which part of the SP6-polylinker and actin sequences (top) were deleted, respectively. The arrow indicates the SP6 transcription start site.
C, general structure of the capped SP6 runoff transcripts tsp19 to tsp0.

sequences were synthesized (Fig. 1) and incubated in a soluble yeast splicing extract. The products of the splicing reaction were analyzed by polyacrylamide gel electrophoresis (Fig. 2).

By all criteria the transcript tsp19 with an exon 1 length of 19 nucleotides was correctly spliced. Both of the expected end products of the reaction, the spliced mRNA and the excised intron lariat, could be easily identified (Fig. 2A). Furthermore one of the two splicing intermediates, the intron-plus-exon-2 lariat, was identified. The other splicing intermediate,

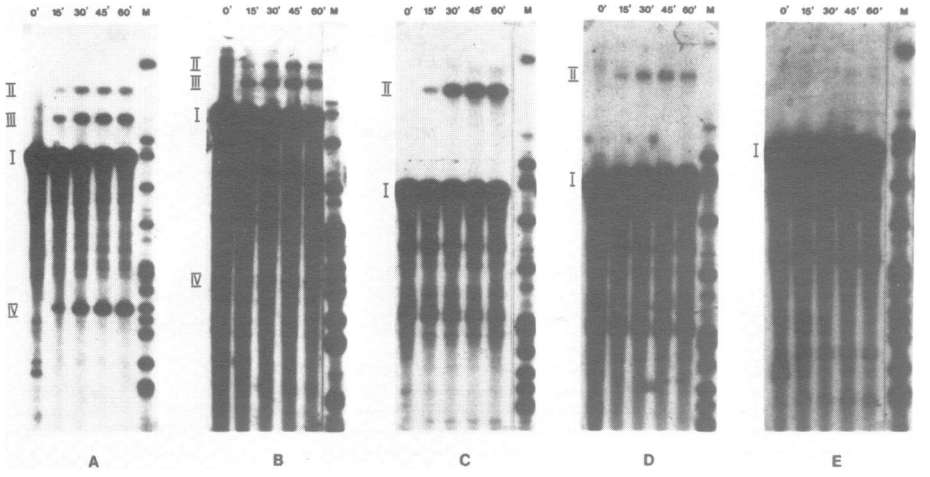

FIG. 2. Analysis of in vitro splicing reactions by polyacrylamide gel electrophoresis. Results are shown only for capped tsp19 (A), tsp12 (B), tsp5 (C), tsp1 (D) and tsp0 (E). Reactions were carried out for 0, 15, 30, 45, and 60 minutes and analyzed on 8 M urea / 8% polyacrylamide gels. pBR322 fragments obtained through cleavage with Hpa II were used as length markers. Band I represents precursor RNA, band II intron-exon-2 lariat, band III excised intron-lariat and band IV spliced mRNA.

exon 1, was not detected due to its small size. A similar pattern was obtained with tsp12 (Fig. 2B). Although it was not possible to immediately identify the correctly spliced mRNA, a primer extension analysis of the splicing products of tsp12 with reverse transcriptase proved that the splicing reaction had generated both of the expected end products (data not shown).

From tsp5 in which exon 1 consists of only five nucleotides, none of the usual final products of the splicing reaction was obtained. Instead, only the intron-plus-exon-2 intermediate was generated (Fig. 2C). Even ten-fold longer exposure times of the gels did not reveal a band corresponding to the excised intron-lariat. Due to its characteristic migration behavior, this band would have been easily detected even if it had been present in only extremely small amounts. Its presence would have indicated that the reaction proceeded into the second step of a complete splicing reaction. Analogous results were obtained when tsp3 and tsp2 were subjected to the splicing extract (data not shown). Even tsp1 in which the first exon is represented by only one nucleotide was converted to the intron-plus-exon-2 lariat, however, to a significantly smaller extent (Fig. 2D). Conversely, no spliced product or intermediate of the splicing reaction could be obtained from transcript tsp0 which starts with the first nucleotide of the intron (Fig. 2E). Fig. 2 shows the results obtained with capped mRNA precursors; analogous results were obtained with uncapped precursors (data not shown).

To test whether the shortened transcripts, especially tsp0, can assemble proper

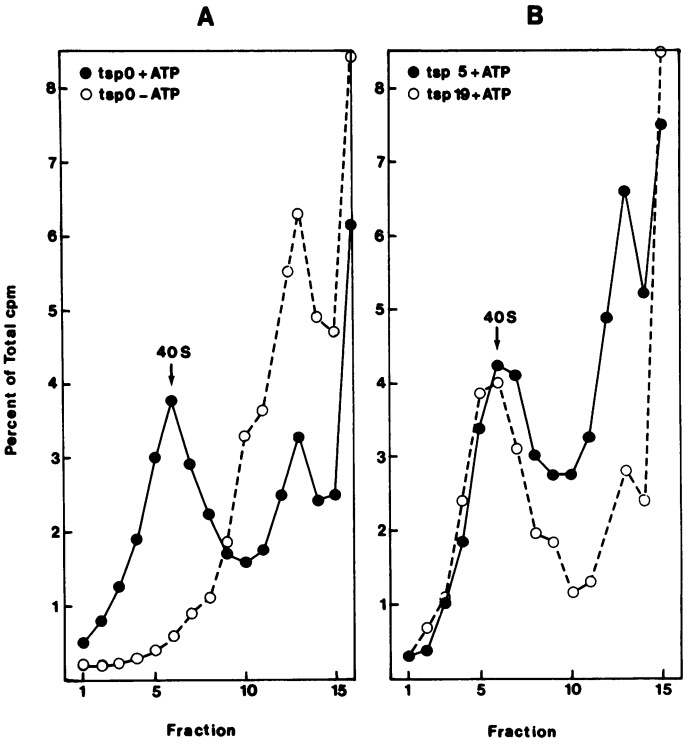

FIG. 3. Glycerol Gradient Sedimentation Analysis of tsp0, tsp5, and tsp19.
A, tsp0 incubated in the presence (●) or the absence (○) of ATP; B, tsp19 (●) and tsp5 (○) incubated in the presence of ATP. The spliceosome specific 40S peak is indicated with an arrow.

spliceosomes, splicing reactions were analyzed for spliceosome formation by glycerol gradient centrifugation. All the transcripts led to an ATP-dependent formation of a 40S spliceosome peak. Even the tsp0 transcript which lacks completely exon 1 sequences and does not enter into a splicing reaction at all led to the formation of 40S spliceosomes (Fig. 3).

DISCUSSION

Our results demonstrate that the first and the second step of the splicing reaction have different requirements for exon 1 sequences. Whereas a lariat intermediate can still be formed from a transcript with an exon 1 of only one nucleotide, a transcript with an exon 1 of five nucleotides cannot enter into the second step of the splicing reaction. This might be caused simply by a loss of an extremely small exon 1 through diffusion. We have tried to test this assumption by adding a large molar excess (up to 10^5 fold) of di- and triribonucleotide

TABLE 1. Correlation of the length of the first exon of the yeast actin gene with the formation of spliceosomes, splicing intermediates, and splicing end products.

length of exon 1 (nuclefotides)	spliceosome formation	splicing intermediates	splicing endproducts
19	+	+	+
12	+	+	+
5	+	+	-
3	+	+	-
2	+	+	-
1	+	+	-
0	+	-	-

(UpG^{OH} and $ApUpG^{OH}$) to a splicing reaction with the tsp2 and tsp3 transcripts. It was not possible, however, to drive the reaction forward in this way; no excised intron lariat or spliced mRNA could be obtained (results not shown). From a physiological point of view it is of course not surprising that exogenously added oligoribonucleotides cannot enter into an mRNA splicing reaction.

Interestingly, exon 1, although indispensable for the splicing reaction, is not required for spliceosome formation. This result at least questions a possible involvement of exon 1 sequences in aligning the 5' splice site of a yeast mRNA precursor with the 5' end of snR19, the yeast analog of the mammalian U1 RNA (19, 22). On the other hand, a minimum size of exon 1, although dispensable for splicing complex formation, still could be essential for keeping the cleaved off exon 1 attached to the spliceosome after the first step of the reaction, possibly by an interaction with the U1 snRNP.

The main conclusion from our data which are summarized in Table 1 is that only a very limited amount of exon 1 sequences (at least 6 but not more than 12 nucleotides) is required for a complete splicing reaction to take place. This minimum size requirement agrees well with the lengths of the first exons found in intron-containing yeast genes, in which no exon 1 smaller than 16 nucleotides has been described until now. If the first exon has a length of five or less nucleotides, the second step of the splicing reaction is no longer possible. Spliceosome analysis has shown that no exon 1 sequences are required for the ATP-dependent formation of spliceosomes. These findings demonstrate again the dominance of intron sequences in the recognition of precursor mRNAs by the splicing machinery. Whether the primary structure of the intron-adjacent exon sequences is of any significance for a yeast splicing reaction, or at least for its efficiency, is currently under investigation.

ACKNOWLEDGMENTS

We thank Ronald Mertz for the synthesis of oligodeoxynucleotides. This work was supported by a grant from the Bundesministerium für Forschung und Technologie and by donations from Hoechst AG and Wacker GmbH.

*To whom correspondence should be addressed

REFERENCES

1. Padgett, R.A., Konarska, M.M., Grabowski, P.J., Hardy, S.F. and Sharp, P.A. (1984) Science **225**, 898 - 903.
2. Ruskin, B., Krainer, A.R., Maniatis, T. and Green, M.R. (1984) Cell **38**, 317 - 331.
3. Gallwitz, D. (1982) Proc. Natl. Acad. Sci. USA **79**, 3493 - 3497.
4. Langford, C.J., Klinz, F.-J., Donath, C. and Gallwitz, D. (1984) Cell **36**, 645 - 653.
5. Thompson-Jäger, S. and Domdey, H. (1987) Mol. Cell. Biol., **7**, 4010 - 4016.
6. van Santen, V.L. and Spritz, R.A. (1985) Proc. Natl. Acad. Sci. USA **82**, 2885 - 2889.
7. Wieringa, B., Hofer, E. and Weissmann, C. (1984) Cell **37**, 915 - 925.
8. Rymond, B.C. and Rosbash, M. (1985) Nature **317**, 735 - 737.
9. Furdon, P.J. and Kole, R. (1986) Proc. Natl. Acad. Sci. USA **83**, 927 - 931.
10. Bindereif, A. and Green, M.R. (1987) EMBO J. **6**, 2415 - 2424.
11. Frendewey, D. and Keller, W. (1985) Cell **42**, 355 - 367.
12. Parent, A., Zeitlin, S. and Efstratiadis, A. (1987) J. Biol. Chem. **262**, 11284 - 11291.
13. Aebi, M., Hornig, H. and Weissmann, C. (1987) Cell **50**, 237 - 246.
14. Siliciano, P.G., Jones, M.H. and Guthrie, C. (1987) Science **237**, 1484 - 1487.
15. Ng, R. and Abelson, J. (1980) Proc. Natl. Acad. Sci. USA **77**, 3912 - 3916.
16. Domdey, H., Apostol, B., Lin, R.-J., Newman, A., Brody, E. and Abelson, J. (1984) Cell **39**, 611 - 621.
17. Newman, A.J., Ogden, R.C. and Abelson, J. (1983) Cell **35**, 117 - 125.
18. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA **74**, 5463 - 5467.
19. Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K. and Green, M.R. (1984) Nucleic Acids Res. **12**, 7035 - 7056.
20. Lin, R.-J., Newman, A.J., Cheng, S.-C. and Abelson, J. (1985) J. Biol. Chem. **260**, 14780 - 14792.
21. Brody, E. and Abelson, J. (1985) Science **228**, 963 - 967.
22. Kretzner, L., Rymond, B.C. and Rosbash, M. (1987) Cell **50**, 593 - 602.