MOLECULAR AND CELLULAR BIOLOGY, June 1997, p. 2985–2993 Vol. 17, No. 6
0270-7306/97/$04.00+0
Copyright © 1997, American Society for Microbiology

Exon Skipping and Circular RNA Formation in Transcripts of the Human Cytochrome P-450 2C18 Gene in Epidermis and of the Rat Androgen Binding Protein Gene in Testis

PETER G. ZAPHIROPOULOS*

Department of Bioscience, Center for Nutrition and Toxicology, Karolinska Institute, Novum, 141 57 Huddinge, Sweden

Received 2 December 1996/Returned for modification 21 January 1997/Accepted 4 March 1997

The cytochrome P-450 2C18 gene was found by reverse transcription-PCR to represent the most abundantly expressed gene of the P-450 2C subfamily in human epidermis. However, in addition to the canonical mRNA of nine exons, transcripts that have skipped exon 4 or 5, exons 4, 5, and 6, or exons 4, 5, 6, and 7 were also identified in this tissue. Remarkably, circular RNA transcripts synthesized by the joining of the donor and acceptor splice sites of the same exon were detected in human epidermis for exons 4 and 5. Moreover, molecules composed of exons 4, 5, and 6 with the donor splice site of exon 6 joined to the acceptor splice site of exon 4 or composed of exons 4, 5, 6, and 7 with the donor splice site of exon 7 joined to the acceptor splice site of exon 4 were also found to be present in this tissue. In rat testis, a similar analysis allowed the detection of a circular RNA molecule composed of exons 6 and 7 of the androgen binding protein (ABP) gene, with the donor splice site of exon 7 joined to the acceptor splice site of exon 6, and of an ABP mRNA which had skipped exons 6 and 7. These results apparently substantiate the hypothesis that alternative pre-mRNA splicing has the potential to generate not only mRNAs that lack one or more exons but also circular RNA molecules that are composed of the exons that are skipped. However, additional 2C18 circular species containing various combinations of exons were also detected in human epidermis, and an exon 6-skipped ABP mRNA molecule was identified in rat testis. This observation is interpreted as indicative that at low frequency, numerous circular RNA formation and exon skipping events may occur, allowing the joining of a variety of different combinations of exons. Moreover, the relative stability of these molecules is apparently the key factor that determines the relative ease of their detection.

Although at the genomic level genes contain exons that are arranged in an ordered array, gene transcripts characterized by changes in this ordered array of exons have recently been observed in eukaryotic cells (8, 9, 16). These scrambled transcripts contain the donor splice site of a 3′ exon joined to the acceptor splice site of a 5′ exon, lack a poly(A)$^+$ tail, accumulate in the cytoplasm, and have properties consistent with circular molecules. Large introns adjacent to the exons that are scrambled have been implicated in the process that generates these molecules (8). Moreover, a transcript composed of a single exon joined head to tail was unambiguously shown to represent a circular species by the use of RNase H mapping (5). In this case, the presence of large inverted repeats in the *Sry* locus was postulated to allow the formation of a stem-loop structure that could facilitate the synthesis of that molecule. However, circular, single-exon-containing RNA molecules have also been observed during the expression of the ETS-1 gene (1). In addition, it has been shown that circular RNA molecules can be readily generated in vitro from both linear or circular yeast pre-mRNAs (18, 19).

The rat cytochrome P-450 2C24 gene represents an example of a gene capable of producing a transcript containing scrambled exons (23). The 2C24 scrambled transcript has the donor splice site of exon 4 joined to the acceptor splice site of exon 2 in a molecule composed of exons 2, 3, and 4, lacks a poly(A)$^+$ tail, and has additional properties that are consistent with a circular RNA. Moreover, in the same tissues where this transcript has been detected (liver and kidney), an exon-skipped P-450 2C24 mRNA that lacks exons 2, 3, and 4 has also been identified. This observed identity between the exons that are skipped and the exons that are present in the circular species prompted the formulation of the hypothesis that exon skipping and circular RNA formation might be interrelated events. It was therefore suggested that during the process of exon skipping, circular RNA molecules composed of the skipped exons could also be produced (24).

In this study, the correlation between exon skipping and circular RNA formation was investigated during the expression of the cytochrome P-450 2C18 gene in human epidermis and of the androgen binding protein (ABP) gene in rat testis. This resulted in extending the original observation for cytochrome P-450 2C24 and also in identifying cases where either exon-skipped or circular RNAs are relatively abundant without their corresponding species being present at comparable levels.

MATERIALS AND METHODS

Human epidermal and rat testis RNAs were extracted as described before (26) by the guanidinium isothiocyanate method (7). cDNA synthesis was performed from total RNA, using as primers random hexamers, oligo(dT), or oligonucleotide ABP6r, as described previously (24). PCR amplifications were performed on a model 2400 Perkin-Elmer thermocycler, using *Taq*, Expand (Boehringer Mannheim), or KlenTaq (Clontech) polymerase. The primers used for the PCR are indicated in Table 1, and their positions on the P-450 2C18 or the ABP cDNA are shown in Fig. 1. The annealing temperature was chosen to be 1 to 5 degrees below the melting temperatures of the primers. Amplification using 30 cycles was employed in all cases except with the set of primers 4F5 and 4R6 or when primer ABP4F was used in combination with primers ABP7R5R, ABP8R5R, and ABPNC5R, in which case only 20 cycles were performed. Cycling was for 1 min at 94°C, 1 min at the desired annealing temperature, and 1 min at 72°C. For the nested amplifications, 1 μl of the products of the initial

* Phone: 46 8 6089151. Fax: 46 8 6081501. E-mail: peza@cnt.ki.se.

TABLE 1. Sequences of the PCR primers used in the amplification experiments

Primer	Sequence
consF, sense	5′ TGT GCT CCI TGC AAI GTI ATC TGC
consR, antisense	5′ ATC IAG AAA GTG ICC IGG GTC AAA
1F, sense	5′ TCA ATG GAT CCA GCT GTG GCT
9R, antisense	5′ CTT CAG ACA GGA ATG AAG CAG
4f, sense	5′ ACT TGA TGG AAA AAT TCA ATG
4F, sense (nested)	5′ CTC AGG ATT CTG AGC TCT
4F5, sense (nested)	5′ CTC CAT GGA TCC AGG TCT G
4r, antisense	5′ AAA CCT CTG ATC TTT ATA ATC
4R, antisense (nested)	5′ AAT CGA TCA TGG AAA ATA ACA
4R6, antisense (nested)	5′ ACA GGG TGA GGC TGT GAC C
5f, sense	5′ GGA GAG AAT AAA AGA ACA TC
5F, sense (nested)	5′ AAG AAT CCC TGG ACA TGA A
5r, antisense	5′ ATA CAT AAC TTT TAA TGT AAG C
5R, antisense (nested)	5′ TTT TAT TAT GAC TTC CTG GG
6f, sense	5′ GCT GGA ACA GAG ACA ACG
6F, sense (nested)	5′ CAC CAC TCT GAG ATA TGG A
ABP6f, sense	5′ AGG ACG ACT CCT TAC TCT
ABP6F, sense (nested)	5′ TGG GAC AGG GAC AAA TTC
ABP6r, antisense	5′ CTC CAT CCA CCA GCT TAA
ABP6R, antisense (nested)	5′ ATC CCA GCT CCA GAG AAA
ABP6RR, antisense	5′ ATC CCA GCT CCA GAG AAA AGG TC
ABP6RRR, antisense (nested)	5′ TCT GTG TGA GGC TGG GG
ABP4f, sense	5′ CCA GCT CAG CAT GAG GA
ABP4F, sense (nested)	5′ CCC CAC TTC CAA ACT TCG
ABP8r, antisense	5′ AAA GGG GTT TAG TGG GAG
ABP8R5R, antisense (nested)	5′ AGT CCT CTC CTT GGA GAC
ABP7R5R, antisense (nested)	5′ AAC CAC CGT CTT GGA GAC
ABPNC5R, antisense (nested)	5′ CAC TGT CAG CTT GGA GAC

amplification (total volume of 50 μl) was used directly in the second amplification reaction. When the CapFinder technology (Clontech) was used, only 20 amplification cycles were performed, with PCR primer ABP6RR in combination with the CapFinder 5′ primer. In this case, the manufacturer's recommendations concerning the optimal temperatures and the duration of each cycling step were followed. The PCR products were electrophoresed on agarose gels of various concentrations, depending on their expected size, prior to cloning into the pGEM T vector (Promega). Plasmid DNA from individual colonies was isolated by using a JetPrep kit (Genomed) and sequenced with fluorescent dye terminators either by Cybergene AB (Huddinge, Sweden) or by the sequencing facility of the Department of Molecular Medicine at the Karolinska Hospital. Sequencing comparisons were performed by using the GCG program of the University of Wisconsin.

RESULTS

Identification of cytochrome P-450 2C18 as the major P-450 of subfamily 2C that is expressed in human epidermis: detection of an exon 5-skipped product. Reverse transcription-PCR (RT-PCR) analysis of human epidermal RNA was performed with PCR primers consF and consR (Table 1; Fig. 1). These primers have previously been used successfully to detect mRNAs coding for 2C P-450s that are expressed in rat brain and pig ovaries and are derived from two conserved regions of the cytochrome P-450s of the 2C subfamily (25, 27). All of the 16 independent PCR clones that were sequenced represented P-450 2C18, suggesting that this P-450 is the most abundantly expressed member of the 2C subfamily in human epidermis. However, an additional PCR product with a molecular weight lower than that of the expected product of 728 bp was also identified in this tissue. Sequencing of three independent PCR clones of the lower-molecular-weight band revealed that these clones also represented P-450 2C18 but without exon 5 (Fig. 2).

Identification of a circular cytochrome P-450 2C18 transcript having the donor and acceptor splice sites of exon 5 joined together (head-to-tail joining). According to the previously formulated exon skipping-circular RNA hypothesis, the exons that are skipped during pre-mRNA splicing will form a circular molecule that has the donor splice site of the 3′ skipped exon joined to the acceptor splice site of the 5′ skipped exon. Since in the case of P-450 2C18 only exon 5 was found to be skipped, this would result in the joining of the donor and acceptor splice sites of exon 5, forming a circularized, single-exon molecule (head-to-tail joining). To test for the presence of such molecules in human epidermis, the RT-PCR analysis was used, with primers 5f and 5r initially and then with primers 5F and 5R (Table 1; Fig. 1). This resulted in the identification of a PCR product of the expected size of 119 bp from cDNA generated by random primers. When oligo(dT) primers were used for cDNA synthesis, no such molecule could be detected. Sequencing of four independent PCR clones of that size confirmed that these clones represented exon 5 joined at its donor and acceptor splice sites (Fig. 3).

Identification of cytochrome P-450 2C18 mRNAs that lack exon 4 or exons 4, 5, and 6. Although an exon 5-skipped P-450 2C18 product had been identified with the use of the conserved PCR primers consF and consR, which map to exon 4 and exon 8, respectively (Fig. 1), the presence of mRNAs that have skipped additional exons had not thoroughly been investigated. For that purpose, the RT-PCR analysis was carried out with primers 1F and 9R (Table 1; Fig. 1). This experiment clearly revealed the presence of the canonically spliced, nine-exon-containing mRNA (band A) and the exon 5-skipped mRNA (band B, six of eight independent PCR clones of that size). However, an exon 4-skipped mRNA (band B, the two remaining PCR clones of that size) was also found to be present in human epidermis. Moreover, a P-450 2C18 mRNA that had skipped exons 4, 5, and 6 (band C, all 19 independent PCR clones of that size) was also detected in this tissue (Fig. 4).

Identification of circular cytochrome P-450 2C18 transcripts with the donor and acceptor splice sites of exon 4 joined together (head-to-tail joining) or the donor splice site of exon 6 joined to the acceptor splice site of exon 4. The finding of the exon 4 skipped P-450 2C18 mRNA prompted the search for a circular transcript having the donor splice site of exon 4 joined to its acceptor splice site (head-to-tail joining) in human epidermis. By RT-PCR analysis using primers 4f and 4r and subsequently primers 4F and 4R (Table 1; Fig. 1), a PCR product of the expected size of 108 bp was detected from cDNA generated by random primers. When oligo(dT) primers were used for cDNA synthesis, no such molecule could be detected. Sequencing of five independent PCR clones of that size confirmed that these clones represented exon 4 joined at its donor and acceptor splice sites (Fig. 5). Similarly RT-PCR analysis using primers 6f and 4r and subsequently primers 6F and 4R (Table 1; Fig. 1) resulted in the identification in this tissue of a PCR product of a size of 131 bp from cDNA generated by random but not oligo(dT) primers. Sequencing of three independent PCR clones of that size revealed that these clones contained the donor splice site of exon 6 joined to the acceptor splice site of exon 4 (Fig. 5). To confirm that this exon 6-exon 4 scrambled product results from a circular transcript that is composed of exons 4, 5, and 6, efforts were focused on detecting a PCR product having a 4-5-6-4 order of exons. For that purpose, the products of the exon 4 sense-antisense amplification (Fig. 5A) were subjected to an additional amplification using primers 4F5 and 4R6. Primer 4F5 contains 14

A

1 2 3 4 5 6 7 8 9

ATG 168 331 481 642 819 961 1149 1291 TGA

-3 1F

514 consF

697 5R

755 5F

909 6F

1242 consR

1475 9R

560 4R

613 4F

492 4R6

628 4F5

4R6

B

1 2 3 4 5 6 7 8

ATG 114 206 396 568 718 855 1063 TAA

787 ABP6r

789 ABP6f

495 ABP4f

769 ABP6R

807 ABP6F

1220 ABP8r

769 ABP6RR

740 ABP6RRR

534 ABP4F

ABP8R5R

1072 ABP8R5R

864

FIG. 1. Localization of PCR primers in the P-450 2C18 (A) and ABP (B) cDNAs. The positions of the exon junctions in the P-450 2C18 and ABP cDNAs are indicated. The primers used for PCR amplifications are shown, with the position of the first nucleotide indicated.

nucleotides of the 3′ end of exon 4 followed by 5 nucleotides of the 5′ end of exon 5, while primer 4R6 contains 11 nucleotides of the 5′ end of exon 4 followed by 8 nucleotides of the 3′ end of exon 6 (Table 1; Fig. 1), thus allowing the direct identification of a PCR product with a 4-5-6-4 order of exons (Fig. 6).

Identification of additional circular P-450 2C18 RNA molecules containing various combinations of exons and of an additional exon-skipped 2C18 mRNA. While attempting to identify the circularized single exon 5 molecule, additional PCR products were also detected and originally thought to

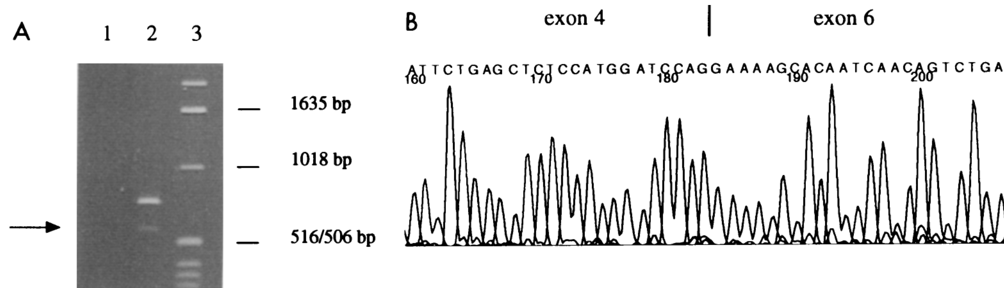

FIG. 2. (A) Agarose gel electrophoresis of the RT-PCR products generated with the use of primers consF and consR on human epidermal RNA. Lane 1, RT-PCR products with no addition of RNA; lane 2, RT-PCR products with added epidermal RNA; lane 3, molecular weight markers. (B) Sequencing of PCR clones corresponding to the lower-molecular-weight band (indicated by an arrow) established that it represents P-450 2C18 with exon 5 being skipped.

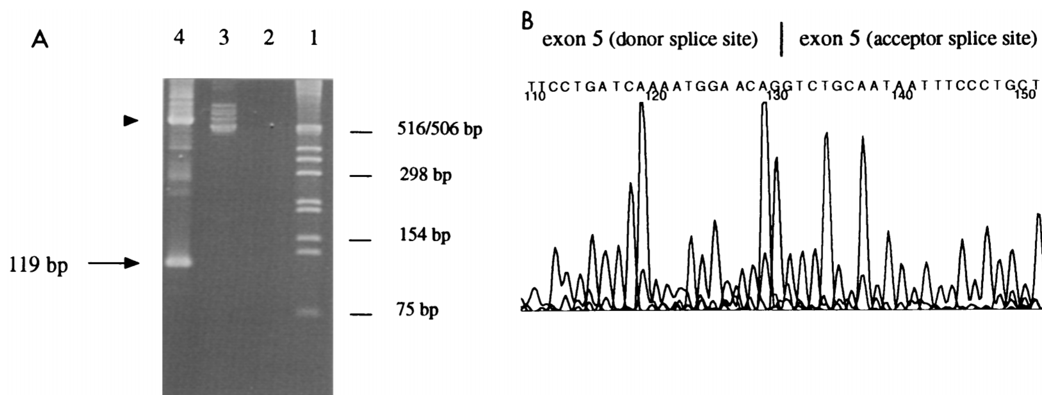

FIG. 3. (A) Agarose gel electrophoresis of the RT-PCR products generated with the use of initial primers 5f and 5r and subsequently with nested primers 5F and 5R on human epidermal RNA. Lane 1, molecular weight markers; lane 2, RT-PCR products with no addition of RNA; lane 3, RT-PCR products with the use of an oligo(dT) primer during reverse transcription of epidermal RNA; lane 4, RT-PCR products with the use of random hexamers during reverse transcription of epidermal RNA. [Both the oligo(dT) and random hexamer cDNA preparations allow the generation of comparable amounts of PCR products corresponding to the canonical 2C18 mRNA (data not shown).] Note that besides the major band of 119 bp, additional PCR products are present in lane 4, with the arrowhead indicating circularized species composed of four exons that have an approximate size of about 600 bp (Table 2). The bands observed in lane 3 as well as a small percentage of the cloned products of lane 4 were found to represent 2C18 species that contained incomplete exons joined at various sites that do not conform with the canonical splice site consensus. At this point, these products are considered to represent PCR artifacts. (B) Sequencing of PCR clones corresponding to the 119-bp band established that it represents the exon 5 of P-450 2C18 joined at its donor and acceptor splice sites.

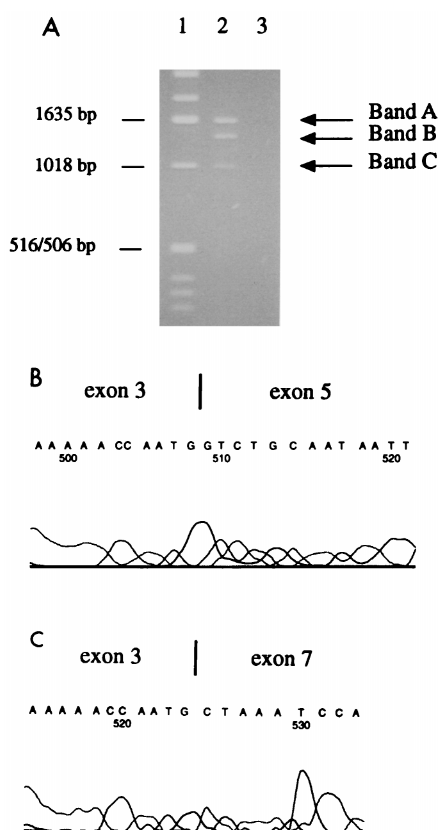

FIG. 4. (A) Agarose gel electrophoresis of the RT-PCR products generated with the use of primers 1F and 9R on human epidermal RNA. Lane 1, molecular weight markers; lane 2, RT-PCR products with added epidermal RNA; lane 3, RT-PCR products with no addition of RNA. (B) Sequencing of eight PCR clones corresponding to band B revealed that two of these clones represented P-450 2C18 with exon 4 being skipped. (C) Sequencing of PCR clones corresponding to band C revealed that it represents P-450 2C18 with exons 4, 5, and 6 being skipped.

indicate that mispriming had occurred during the amplification (Fig. 3a). However, cloning and sequencing of several of them established that they represent additional cases of circular RNAs. These 2C18 molecules contained various combinations of exons starting from a three-exon species up to a large molecule composed of six exons that also included a small, 17-bp sequence that had the properties of an unspliced intron (conserved GT-AG dinucleotides) (Table 2). Interestingly, the previously identified circular molecule that was composed of exons 4, 5, and 6 was not detected in this analysis, suggesting that it is not as abundant. Furthermore, an additional P-450 2C18 mRNA that had skipped exons 4, 5, 6, and 7 (the same exons present in one of the newly identified circular molecules) was also detected when nested amplification was performed on the products of the exon 1 sense-exon 9 antisense PCR amplification (Fig. 4).

Identification of a circular ABP RNA molecule composed of exons 6 and 7 and of an ABP mRNA that has skipped exons 6 and 7 in rat testis. The ABP gene is known to undergo alternative splicing of exon 6 in fetal rat liver (20). To investigate whether this phenomenon could occur in the testis and permit the generation of a circular RNA molecule that has the donor and acceptor splice sites of exon 6 joined together, RT-PCR analysis was performed with primers ABP6f and ABP6r and subsequently primers ABP6F and ABP6R (Table 1; Fig. 1). However, instead of the expected circularized exon 6 (size of 100 bp), the major PCR product identified was found to represent a transcript containing exons 6 and 7 with the donor splice site of exon 7 joined to the acceptor splice site of exon 6 (Fig. 7). This scrambled transcript was detected only when random and not oligo(dT)-primed cDNA was used, in line with what was previously observed with the 2C18 circular molecules. To investigate whether an mRNA that has skipped exons 6 and 7 is also present in rat testis, RT-PCR analysis was performed with primers ABP4f and ABP8r (Table 1; Fig. 1). Since the major product detected was the canonical, non-exon-skipped mRNA (Fig. 7), efforts were focused on developing an amplification scheme that would directly detect an mRNA that has skipped exons 6 and 7. For that purpose, the products of the ABP4f-ABP8r amplification were subjected to an additional PCR amplification using as the sense primer ABP4F and as the

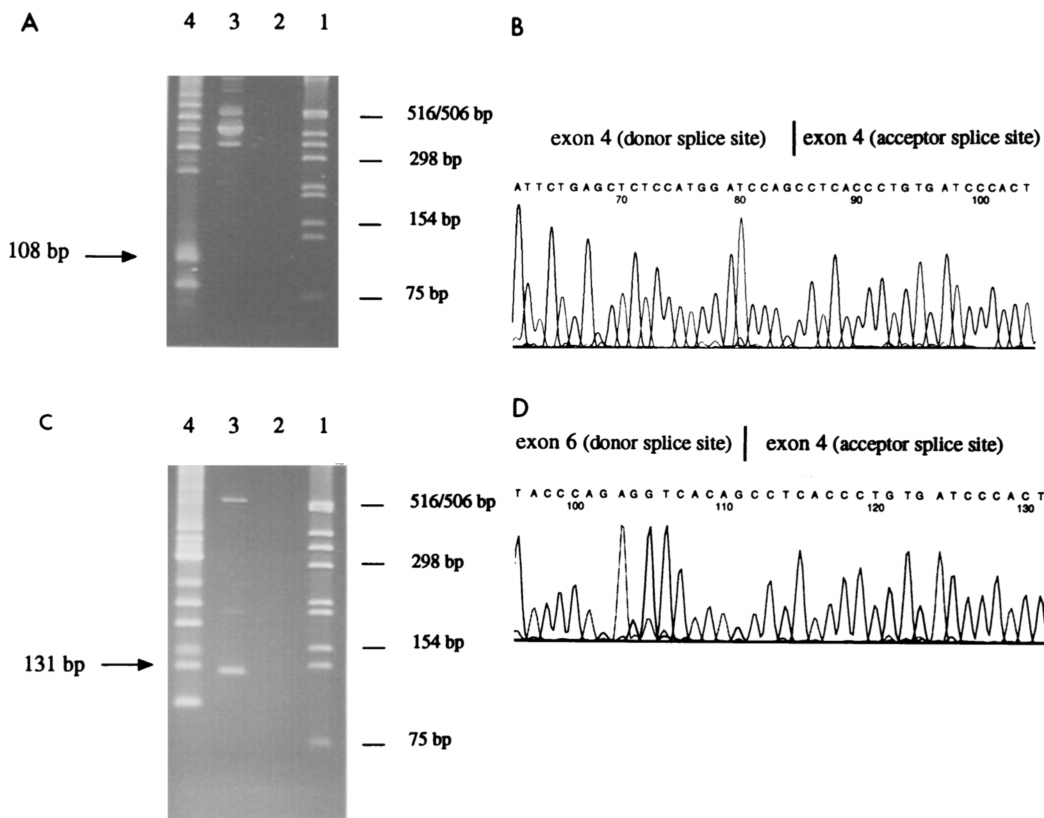

FIG. 5. (A) Agarose gel electrophoresis of the RT-PCR products generated with the use of initial primers 4f and 4r and subsequently with nested primers 4F and 4R on human epidermal RNA. Lane 1, molecular weight markers; lane 2, RT-PCR products with no addition of RNA; lane 3, RT-PCR products with the use of an oligo(dT) primer during reverse transcription of epidermal RNA; lane 4, RT-PCR products with the use of random hexamers during reverse transcription of epidermal RNA. (B) Sequencing of PCR clones corresponding to the 108-bp band established that it represented the exon 4 of P-450 2C18 joined at its donor and acceptor splice sites. (C) Agarose gel electrophoresis of the RT-PCR products generated with the use of initial primers 4r and 6f and subsequently with nested primers 4R and 6F on human epidermal RNA. Lane 1, molecular weight markers; lane 2, RT-PCR products with no addition of RNA; lane 3, RT-PCR products with the use of an oligo(dT) primer during reverse transcription of epidermal RNA; lane 4, RT-PCR products with the use of random hexamers during reverse transcription of epidermal RNA. (D) Sequencing of PCR clones corresponding to the 131-bp band established that it represents a P-450 2C18 transcript that has the donor splice site of exon 6 joined to the acceptor splice site of exon 4. Most of the extra bands seen in lanes 4 of panels A and C could represent additional circular species, as was the case in Fig. 3. However, some of these bands as well as the PCR products seen in lanes 3 of panels A and C were found to represent 2C18 species containing incomplete exons, as was the case in Fig. 3.

antisense primer ABP8R5R. Primer ABP8R5R contains nine nucleotides of the 5′ end of exon 8 followed by nine nucleotides of the 3′ end of exon 5 (Table 1; Fig. 1). This approach established that an ABP mRNA that has skipped exons 6 and 7 is expressed in rat testis (Fig. 7). However, when primer ABP8R5R was substituted by primer ABP7R5R, which contains nine nucleotides of the 5′ end of exon 7 followed by nine nucleotides of the 3′ end of exon 5 (Table 1; Fig. 1), an even more abundant PCR product was detected. These findings indicate that an exon 6-skipped ABP mRNA is expressed at higher levels than the exon 6- and 7-skipped mRNA; this is in contrast to the observation that the major product revealed by the exon 6 sense-antisense amplification is the exon 6 and 7 circular molecule and not the putative exon 6 circularized species.

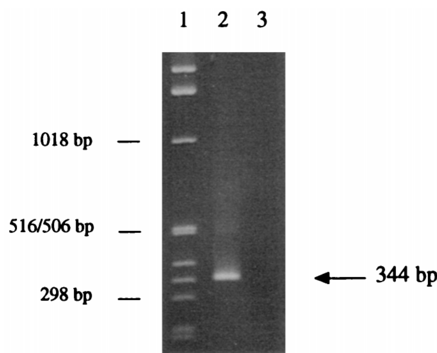

FIG. 6. Agarose gel electrophoresis of the PCR products generated with the use of primers 4F5 and 4R6 on the amplification products of the exon 4 sense-antisense RT-PCR (Fig. 5A). Lane 1, molecular weight markers; lane 2, PCR products with an added sample of the exon 4 sense-antisense RT-PCR; lane 3, PCR products with no added sample of the exon 4 sense-antisense RT-PCR. Note the presence of a major PCR product of the expected size of 344 bp in lane 2.

TABLE 2. Circular P-450 2C18 molecules detected by the use of exon 5 sense-antisense primers (in addition to the circularized single exon 5)

Exon order	No. of independent clones analyzed
5-2-3-5	3
5-2-3-4-5	17
5-6-7-4-5	14
5-6-7-8-5	7
5-6-7-2-3-5	2
5-6-small intron-7-2-3-4-5	2

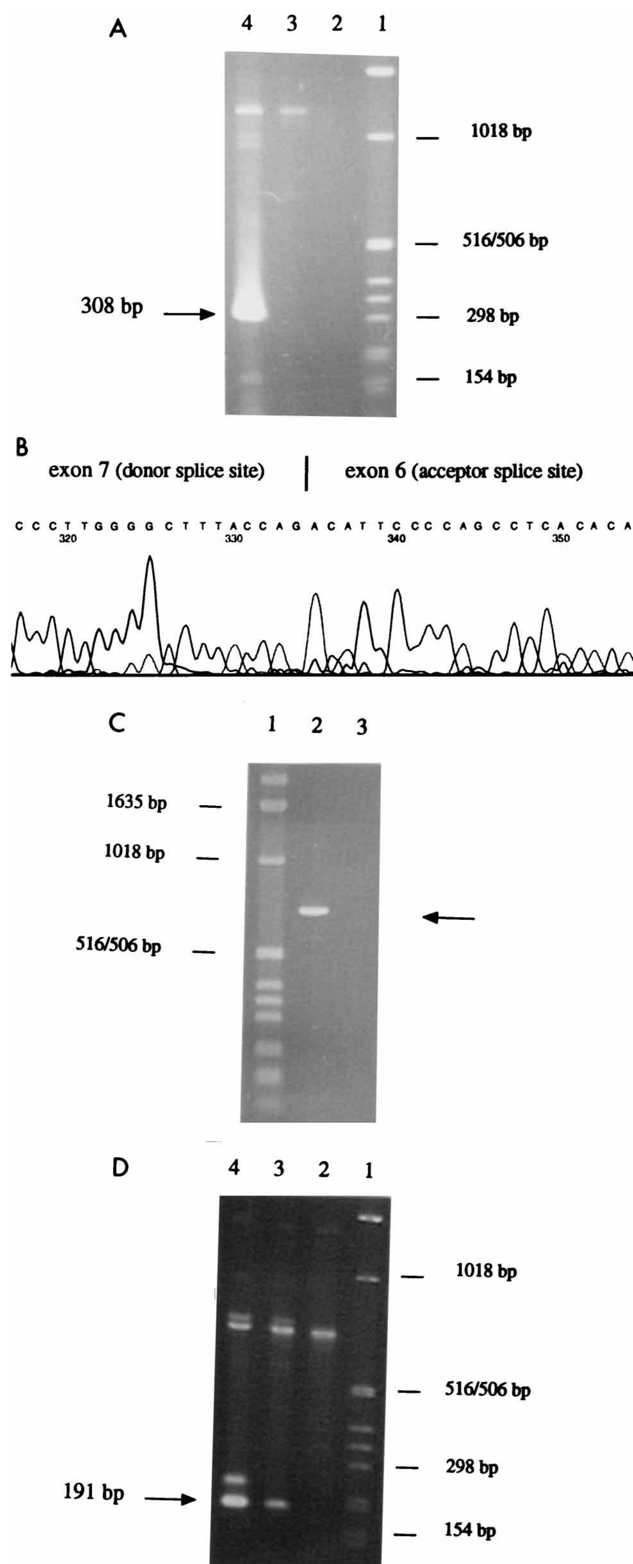

FIG. 7. (A) Agarose gel electrophoresis of the RT-PCR products generated with the use of initial primers ABP6f and ABP6r and subsequently with nested primers ABP6F and ABP6R on rat testicular RNA. Lane 1, molecular weight markers; lane 2, RT-PCR products with no addition of RNA; lane 3, RT-PCR products with the use of an oligo(dT) primer during reverse transcription of testicular RNA; lane 4, RT-PCR products with the use of random hexamers during reverse transcription of testicular RNA. [Both the oligo(dT) and the random hexamer cDNA preparations allow the generation of comparable amounts of PCR products corresponding to the canonical ABP mRNA (data not shown).] The major band present in lane 3 was cloned and sequenced, verifying that it represents a misprimed PCR product. (B) Sequencing of PCR clones corresponding to the 308-bp band of lane 4 revealed that it represents an ABP transcript that has the donor splice site of exon 7 joined to the acceptor splice site of exon 6. (C) Agarose gel electrophoresis of the RT-PCR products generated with the use of primers ABP4f and ABP8r on rat testis RNA. Lane 1, molecular weight markers; lane 2, RT-PCR products with added testicular RNA; lane 3, RT-PCR products with no addition of RNA. (D) Agarose gel electrophoresis of the PCR products generated with the use of nested primer ABP4F in combinations with primer ABP8R5R, ABP7R5R, or ABPNC5R on the products of the exon 4 sense-exon 8 antisense RT-PCR (C, lane 2). Lane 1, molecular weight markers; lane 2, PCR products obtained by using primer ABPNC5R (this primer contains the nine nucleotides of the 5′ of exon 5, similar to primers ABP7R5R and ABP8R5R, followed by nine random nucleotides [Table 1]); lane 3, PCR products obtained by using primer ABP8R5R; lane 4, PCR products obtained by using primer ABP7R5R. Note the presence of a major PCR product of the expected size of 191 bp in lanes 3 and 4.

Absence of a cap structure in the exon 6 and 7 scrambled ABP molecule. To investigate whether a cap structure is present in the exon 6 and 7 scrambled ABP molecule, the CapFinder technology (Clontech) was used. Briefly, testicular cDNA was synthesized by using primer ABP6r, in the presence of the CapSwitch oligonucleotide, and subsequently amplified with the use of primer ABP6RR (Table 1; Fig. 1) in combination with a 5′ primer encompassing sequences from the CapSwitch oligonucleotide. This amplified cDNA preparation, which is enriched in transcripts containing a cap structure, was compared with the nonamplified testicular cDNA for the presence of the exon 6 and 7 scrambled ABP molecule (Fig. 8). As seen in Fig. 8, only the nonamplified cDNA allowed the detection of the scrambled ABP transcript, while sequences corresponding to the canonical ABP mRNA were found to be present at comparable amounts in both cDNA preparations. These observations are therefore consistent with the possibility that scrambled transcripts are indeed circular RNAs, as this would preclude incorporation of a cap structure in these molecules.

DISCUSSION

The finding that there is a tight (albeit not absolute) correlation between exon skipping and circular RNA formation in human P-450 2C18 as well as in rat ABP extends the original findings for rat cytochrome P-450 2C24, suggesting that this phenomenon is not restricted to a single gene but also characterizes the expression of non P-450 genes (Fig. 9). Exon skipping is a splicing mechanism which allows the generation of different protein products with diverse functions from a single gene (14). However the biological role of the circularized, scrambled exon-containing transcripts has not yet been elucidated. Their apparent correlation with exon skipping could indicate that these molecules have functional roles in mechanisms regulating the process of alternative pre-mRNA splicing. Moreover, the recent finding that lariats from group II introns can insert themselves into double-stranded DNA (22) is suggestive that uncharacterized functions might be associated with circular RNA species and that evolutionary could have included RNA mobility and exon shuffling.

How certain is the notion that these identified scrambled transcripts represent circular molecules? Is it possible that putative *trans*-splicing events can account for the generation of these species? Certainly *trans*-splicing of individual pre-mRNAs could result in molecules having downstream donor splice sites joined to upstream acceptor splice sites. In fact, this phenomenon has been shown to occur in vitro (3, 6), and furthermore, mammalian cells were found to have the ability to generate

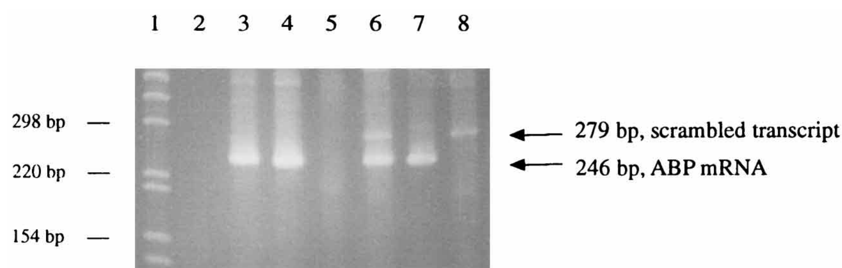

FIG. 8. Comparison of cap-enriched versus non-cap-enriched testicular cDNAs for the presence of the exon 6 and 7 scrambled ABP transcript. Lane 8, PCR products obtained by using primers ABP6F and ABP6RRR after an initial amplification with primers ABP6f and ABPR on the non-cap-enriched cDNAs; lane 7, PCR products obtained by using primers ABP4f and ABP6RRR after an initial amplification with primers ABP6f and ABPR on the non-cap-enriched cDNAs; lane 6, PCR products obtained by using primers ABP6F and ABP4f in combination with primer ABP6RRR after an initial amplification with primers ABP6f and ABPR on the non-cap-enriched cDNAs; lane 5, PCR products obtained by using the same primers as in lane 8 on the cap-enriched cDNAs; lane 4, PCR products obtained by using the same primers as in lane 7 on the cap-enriched cDNAs; lane 3, PCR products obtained by using the same primers as in lane 6 on the cap-enriched cDNAs; lane 2, PCR products obtained by using the same primers as in lanes 3 and 6 with no addition of cDNA; lane 1, molecular weight markers.

functional mRNA molecules by *trans* splicing (12). However, in vivo-generated *trans*-spliced molecules are polyadenylated like canonical mRNAs. The lack of polyadenylation observed in all known scrambled species is consistent with either circular molecules or nonpolyadenylated *trans*-spliced products. In the case of the P-450 2C24 scrambled species, it has furthermore been shown that an exon 5 antisense reverse transcriptase primer does not allow the detection of the scrambled molecule, suggesting that the 3′ limit of this transcript is exon 4. Therefore, if *trans* splicing was indeed occurring, it would not encompass the complete pre-mRNA, an event that is difficult to rationalize (24). In addition, the observed lack of a cap structure in the ABP scrambled molecule is also consistent with this species being a circular RNA and not a linear, *trans*-spliced product of two canonical pre-mRNAs. Moreover in the case of the *Sry* circularized exon, the circular nature of this molecule was demonstrated by a non-PCR method, RNase H mapping (5). This approach took advantage of the relatively high levels of expression of the *Sry* circular species in the cell lines used and is technically difficult to be performed with the low expression levels of the 2C18 or the ABP scrambled molecules.

If scrambled species are indeed circular RNA molecules, could they represent only a by-product of splicing and therefore be limited to nuclear RNA? In all cases examined thus far, from the original DCC scrambled species to the circular Sry molecule, these transcripts have been found to accumulate in the cytoplasm (5, 9, 16), suggesting a biological role beyond the splicing process for these molecules.

Is there a requirement for the presence of large introns adjacent to the exons that produce scrambled transcripts? The introns linking exon 4 to exon 5 and exon 5 to exon 6 in the P-450 2C18 gene are over 10 kb, with the complete gene spanning about 55 kb (10); however, in the ABP gene, the exon 1-to-exon 8 region is confined within 3 kb. It should be noted that several alternative exon 1 ABP sequences are also present within 25 kb 5′ of the canonical exon 1 (21). It appears therefore that although the exons that are scrambled do not have to be immediately flanked by large introns, extended intronic sequences are present proximal to these exons. Whether this observation has a functional significance in the process of generating scrambled molecules remains to be seen.

Are specific structural features, such as inverted repeats, involved in the formation of scrambled molecules? Although it has been shown that complementary sequences can promote exon circularization in mammalian nuclear extracts (17) and that inverted repeats are necessary for circularization of the mouse testis *Sry* transcript in transfected cells (11), the 376-base 5′ intron and the 302-base 3′ intron that flank the ABP exons 6 and 7 do not contain inverted repeats longer than five bases. However, the 5′ intron contains two copies of an ID element that is also present, in the opposite orientation, at the 3′ untranslated region of that gene (20), allowing the formation of a stem-loop structure that would encompass the scrambled exons. Considering the fact that the *Sry* circular transcript is very abundant and the ABP circular species, as judged from the semiquantitative PCR experiments, appears to be more abundant than any of the 2C18 circular molecules, it seems that formation of a stem-loop structure may facilitate circularization. However, the preference for the joining of the donor splice site of exon 7 with the acceptor splice site of exon 6, from all of the other splice sites that are also present in the postulated loop structure of the ABP pre-mRNA, is likely to be dependent on the association of specific splicing factors (see below).

The RT-PCR technology used throughout this study revealed the presence of various combinations of exons in the circular P-450 2C18 molecules as well as of various exon-skipped 2C18 mRNAs in a gene family previously characterized by the concept of one gene-one mRNA-one protein. This exemplifies the unique sensitivity of the PCR technique in analyzing low-abundance, alternatively spliced transcripts. Despite the fact that pre-mRNA splicing is a major process in eukaryotic gene expression, and even though significant progress in the molecular understanding of this phenomenon has been achieved in recent years (for a review, see reference [15]), the control mechanisms that regulate alternative splicing are still poorly understood. Although SR proteins appear to be essential for both constitutive and regulated splicing, a role of hnRNP proteins in modulating alternative splicing by antagonizing SR proteins has also been proposed (4). Moreover, SR proteins have been found to function as activators or repressors of splicing, depending on the positions of their binding sites on the pre-mRNA (13).

In addition, in accordance with the exon recognition model (2), the splicing machinery is considered to initially recognize a pair of closely spaced splice sites in an exonic polarity. Once the exons are defined by the binding of U1 and U2 snRNPs and associated splicing factors, exon juxtaposition events must occur, presumably via interactions of SR and hnRNP proteins. Although regulated splicing would favor juxtaposition of certain exons, constitutive splicing might allow additional juxtaposition events to occur, and it is conceivable that at a low frequency, these events could include a variety of different (and maybe all) combinations of exons. Such low-abundance

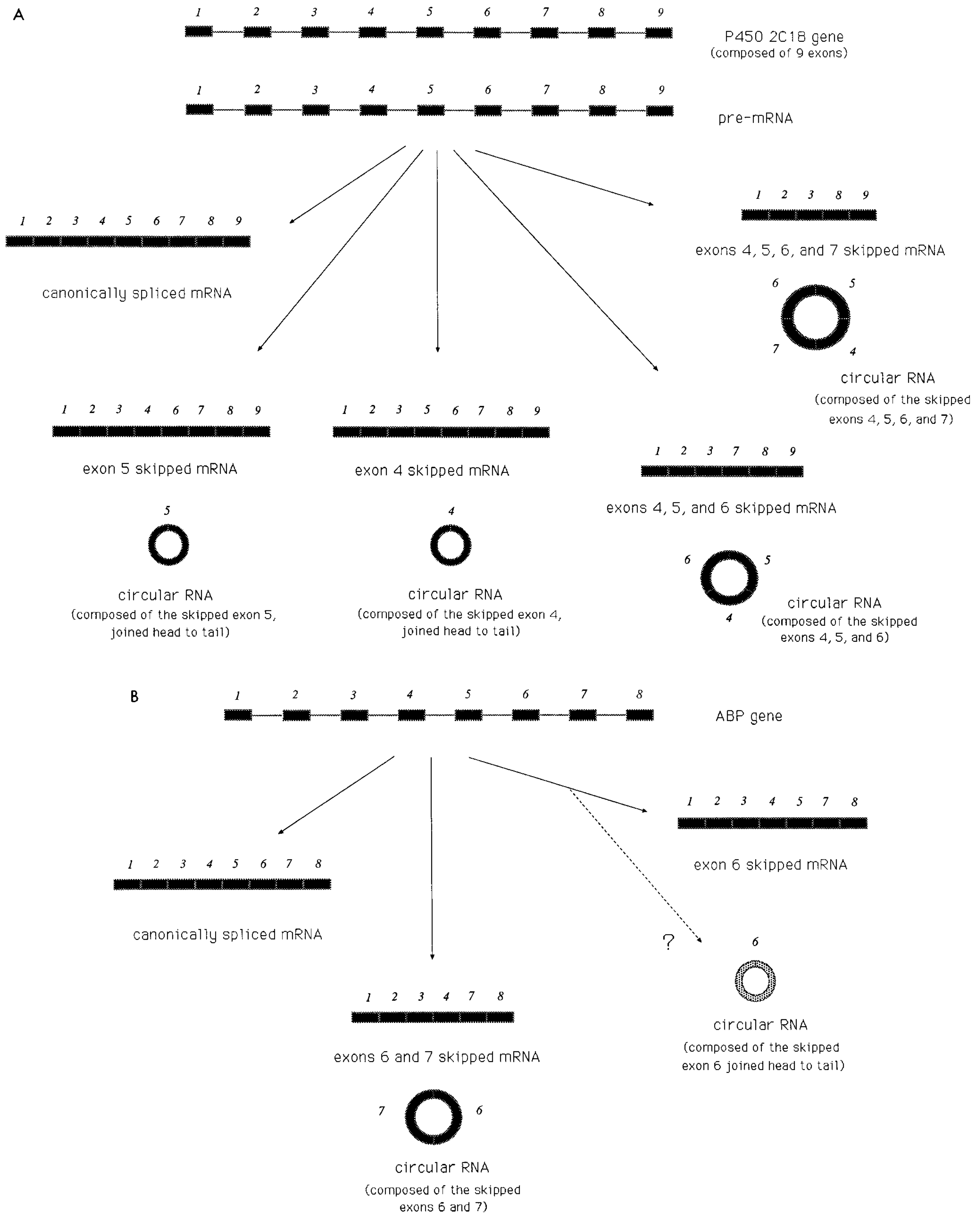

FIG. 9. Schematic diagram of the exon-skipped mRNAs and the circularized RNA molecules of cytochrome P-450 2C18 in human epidermis (A) and of the ABP gene in rat testis (B).

transcripts would be detected only by the most sensitive method available today, namely, PCR.

Is it likely that formation of an exon-skipped mRNA is always accompanied by the generation of the corresponding circularized RNA molecule and vice versa? The tight correlation between exon skipping and circular RNA formation observed in P-450 2C24 seems to favor this hypothesis; however, the more loose correlation observed in P-450 2C18 and ABP calls into question the absolute validity of this statement. A possible scenario that might underlie this phenomenon, however, could be that when a juxtaposition event of noncontiguous exons is occurring (exon skipping), this event may bring in proximity the enclosed exons with their associated splicing factors in an active state that has the potential to allow excision of the internal introns (circular RNA formation). Certainly, this process would require functional interactions between tissue-specific and constitutive splicing factors. Moreover, depending on the gene in question and the exons involved, there might be a favored production of exon-skipped mRNAs, of circular RNAs, or of both types of molecules at comparable levels. A key factor for the detection of such rare RNA species would also be their relative stabilities. In this context, it is not at all certain that all circular molecules generated from one gene would be equally stable. What is likely to be true, though, is that a circular species, once synthesized, would generally be more stable than a corresponding linear molecule, due to the absence of free ends for exonuclease attack. This might account for the fact that circular 2C18 RNAs have been identified, with no corresponding exon-skipped mRNA yet being detected. However, in the case of ABP, a major splicing event that apparently occurs is the production of an exon 6-skipped ABP mRNA and not of a putative exon 6 circularized molecule. It is possible though, that the lack of detection of such a circular transcript might simply reflect the overwhelming predominance of the exon 6 and exon 7 circularized molecule which, in the PCR assay used, would hinder the identification of a less abundant exon 6 species (Fig. 7A).

Accumulated evidence therefore indicates that during the process of pre-mRNA splicing, events that allow 3′ exons to be joined with 5′ exons occur. The frequency of such events, which result in the generation RNAs that have properties of circular molecules, shows a significant but not absolute correlation with the synthesis of mRNAs that have skipped the exons that are present in these circular species. This observed interrelation between exon skipping and circular RNA formation may therefore suggest that the two phenomena are mechanistically linked and rationalized by processes such as lariat or inverse splicing (24).

In summary, circular RNA molecules containing various combinations of exons are expressed in eukaryotic cells, and their production appears to correlate with the generation of mRNAs that have skipped the exons present in these circularized species.

ACKNOWLEDGMENTS

Rune Toftgård is acknowledged for his continuous support, and Darren Krause is acknowledged for a critical reading of the manuscript.

This study was supported by grants from the Lars Hiertas Minne Foundation, the Åke Wibergs Foundation, and the Karolinska Institute.

REFERENCES

1. **Bailleul, B.** 1996. During *in vivo* maturation of eukaryotic nuclear mRNA, splicing yields excised exon circles. Nucleic Acids Res. **24:**1015–1019.
2. **Berget, S. M.** 1995. Exon recognition in vertebrate splicing. J. Biol. Chem. **270:**2411–2414.
3. **Bruzik, J. P., and T. Maniatis.** 1995. Enhancer-dependent interaction between 5′ and 3′ splice sites in trans. Proc. Natl. Acad. Sci. USA **92:**7056–7059.
4. **Cáceres, J. F., S. Stamm, D. M. Helfman, and A. R. Krainer.** 1994. Regulation of alternative splicing in vivo by overexpression of antagonistic splicing factors. Science **265:**1706–1709.
5. **Capel, B., A. Swain, S. Nicolis, A. Hacker, M. Walter, P. Koopman, P. Goodfellow, and R. Lovell-Badge.** 1993. Circular transcripts of the testis-determining gene *Sry* in adult mouse testis. Cell **73:**1019–1030.
6. **Chiara, M. D., and R. Reed.** 1995. A two step mechanism for 5′ and 3′ splice-site pairing. Nature **375:**510–513.
7. **Chomczynski, P., and N. Sacchi.** 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162:**156–159.
8. **Cocquerelle, C., P. Daubersies, M.-A. Majérus, J.-P. Kerckaert, and B. Bailleul.** 1992. Splicing with inverted order of exons occurs proximal to large introns. EMBO J. **11:**1095–1098.
9. **Cocquerelle, C., B. Mascrez, D. Hétuin, and B. Bailleul.** 1993. Mis-splicing yields circular RNA molecules. FASEB J. **7:**155–160.
10. **de Morais, S. M. F., H. Schweikl, J. Blaisdell, and J. A. Goldstein.** 1993. Gene structure and upstream regulatory regions of human CYP2C9 and CYP2C18. Biochem. Biophys. Res. Commun. **194:**194–201.
11. **Dubin, R. A., M. A. Kazmi, and H. Ostrer.** 1995. Inverted repeats are necessary for circularization of the mouse testis *Sry* transcript. Gene **167:**245–248.
12. **Eul, J., M. Graessmann, and A. Graessmann.** 1995. Experimental evidence for RNA *trans*-splicing in mammalian cells. EMBO J. **14:**3226–3235.
13. **Kanopka, A., O. Mühlemann, and G. Akusjärvi.** 1996. Inhibition by SR proteins of splicing of a regulated adenovirus pre-mRNA. Nature **381:**535–538.
14. **Maniatis, T.** 1991. Mechanisms of alternative pre-mRNA splicing. Science **251:**33–34.
15. **Manley, J. L., and R. Tacke.** 1996. SR proteins and splicing control. Genes Dev. **10:**1569–1579.
16. **Nigro, J. M., K. R. Cho, E. R. Fearon, S. E. Kern, J. M. Ruppert, J. D. Oliner, K. W. Kinzler, and B. Vogelstein.** 1991. Scrambled exons. Cell **64:**607–613.
17. **Pasman, Z., M. D. Been, and M. A. Garcia-Blanco.** 1996. Exon circularization in mammalian nuclear extracts. RNA **2:**603–610.
18. **Schindewolf, C. A., and H. Domdey.** 1995. Splicing of a circular yeast pre-mRNA *in vitro*. Nucleic Acids Res. **23:**1133–1139.
19. **Schindewolf, C. A., S. Braun, and H. Domdey.** 1996. *In vitro* generation of a circular exon from a linear pre-mRNA transcript. Nucleic Acids Res. **24:**1260–1266.
20. **Sullivan, P. M., P. Petrusz, C. Szpirer, and D. R. Joseph.** 1991. Alternative processing of androgen-binding protein RNA transcripts in fetal rat liver. J. Biol. Chem. **266:**143–154.
21. **Sullivan, P. M., Y.-M. Wang, and D. R. Joseph.** 1993. Identification of an alternate promoter in the rat androgen-binding protein/sex hormone-binding globulin gene that regulates synthesis of a messenger RNA encoding a protein with altered function. Mol. Endocrinol. **7:**702–715.
22. **Yang, J., S. Zimmerly, P. S. Perlman, and A. M. Lambowitz.** 1996. Efficient integration of an intron RNA into double-stranded DNA by reverse splicing. Nature **381:**332–335.
23. **Zaphiropoulos, P. G.** 1993. Differential expression of cytochrome P450 2C24 transcripts in rat kidney and prostate: evidence indicative of alternative and possibly trans-splicing events. Biochem. Biophys. Res. Commun. **192:**778–786.
24. **Zaphiropoulos, P. G.** 1996. Circular RNAs from transcripts of the rat cytochrome P450 2C24 gene: correlation with exon skipping. Proc. Natl. Acad. Sci. USA **93:**6536–6541.
25. **Zaphiropoulos, P. G., and T. Wood.** 1993. Identification of the major cytochrome P450s of subfamily 2C that are expressed in brain of female rats and in olfactory lobes of ethanol treated male rats. Biochem. Biophys. Res. Commun. **193:**1006–1013.
26. **Zaphiropoulos, P. G., and R. Toftgård.** 1996. cDNA cloning of a novel WD repeat protein mapping to the 9q22.3 chromosomal region. DNA Cell Biol. **15:**1049–1056.
27. **Zaphiropoulos, P. G., A. Skantz, M. Eliasson, and M. Bengtsson Ahlberg.** 1995. Cytochrome P450 genes expressed in porcine ovaries: identification of novel forms, evidence for gene conversion, and evolutionary relationships. Biochem. Biophys. Res. Commun. **212:**433–441.