Circular RNAs from transcripts of the rat cytochrome P450 2C24 gene: Correlation with exon skipping

PETER G. ZAPHIROPOULOS

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

Communicated by Minor J. Coon, The University of Michigan, Ann Arbor, MI, March 4, 1996 (received for review October 2, 1995)

ABSTRACT The cytochrome P450 2C4 gene is characterized by the capability to generate, in rat kidney, a transcript containing exons 2 and 4 spliced at correct sites but having the donor site of exon 4 directly joined to the acceptor site of exon 2 (exon scrambling). By reverse transcriptase-PCR analysis, it is now shown that the only exons present in the scrambled transcript are exons 2, 3, and 4 and that this molecule lacks a poly(A) + tail. Furthermore, the use of PCR primers in both orientations of either exon 2 or exon 4 revealed that the orders of the exons in the scrambled transcript are 2-3-4-2 and 4-2-3-4, respectively. These results, combined with the observation that P450 2C24 is a single-copy gene, with no duplication of the exon 2 to exon 4 segment, suggest that the scrambled transcript has properties consistent with that of a circular molecule. In line with this is the observation of an increased resistance of the transcript to phosphodiesterase 1, a 3’-exonuclease. Moreover, an alternatively processed cytochrome P450 2C24 mRNA, lacking the three scrambled exons and having exon 1 directly joined to exon 5, has been identified in kidney and liver, tissues that express the scrambled transcript. This complete identity of the exons that are absent in the alternatively processed mRNA but present in the scrambled transcript is interpreted as indicative of the possibility that exon scrambling and exon skipping might be interrelated phenomena. It is therefore proposed that alternative premRNA processing has the potential to generate not only mRNAs lacking one or more exons but also circular RNA molecules.

Exon scrambling is a recently described phenomenon in eukaryotic gene expression where exons are found to be spliced at correct sites but joined in an order different from that present in genomic DNA (1). In every known case, the position of a single exon relative to the other exons has been changed, and is found 5’ of its original location (1, 2). It has been proposed that scrambled exons result from a loop structure that the pre-mRNA can assume, which allows nonsequential exons to be spliced together (1, 3). This would result in the generation of circular RNAs, and indeed scrambled transcripts have properties of circular molecules (3). Furthermore, and in agreement with being circular molecules, these transcripts appear to lack a poly(A) + tail (1, 2). In addition, circular RNA molecules composed of a single exon, joined at the 5’- and 3’-splice junctions, have also been reported (4). The cellular localization of these circular RNAs has been found to be predominantly in the cytoplasm (3, 4).

Cytochrome P450 2C24 is a member of a large group of enzymes involved not only in drug and xenobiotic metabolism but also in the synthesis of physiologically important biomolecules, including steroid hormones and oxygenated arachidonic acid metabolites (5, 6). The subclass to which P450 2C24 belongs, subfamily 2C, is characterized by high structural identity, conserved intron/exon organization, and a total of nine coding exons (7, 8). During efforts to isolate the 5’ end of the 2C4 mRNA, by the rapid amplification of cDNA ends (RACE) methodology on rat kidney RNA, several transcripts containing exons 2, 3, and 4 spliced at correct sites but having the donor site of exon 4 directly joined to the acceptor site of exon 2 were identified (9).

In this work, evidence is presented that the generation of the 2C4 scrambled exons is related to the observed exon skipping in certain mRNAs from this gene. The exons composing the scrambled transcript (exons 2, 3, and 4) are the ones that are present in the P450 2C24 mRNA that is alternatively processed by exon skipping. The biological implications of this novel observation of an association between exon scrambling and exon skipping as well as mechanistic interpretations of this phenomenon are discussed.

MATERIALS AND METHODS
cDNA Synthesis. Total rat kidney and liver RNA were prepared by the guanidinium isothiocyanate method (10), and cDNA was synthesized as described (9) except that 15 pmol of the random hexamers and 5 pmol of the oligo dT primer or 1 pmol of the exon 4 and the exon 5 antisense primers were used. The exon 4 and exon 5 antisense primers were 5’-GGGGA-TCCCTATAATCACAAACGATTTAGGAAACAACGT, positions 496–526 of the P450 2C24 sequence (8), with a BamHI site at the 5’-end (underlined), and 5’-TGACCTCC-CGAGAGATAATCG, positions 525–545 of the P450 2C24 sequence (8), respectively. Of the 20-μl cDNA synthesis mixture, 1 μl was directly used for PCR amplification.

PCR Amplification, Cloning, and Sequencing. For the identification of possible trans-spliced transcripts (see Fig. 1), the primers used for PCR amplification (see Fig. 2) were as follows: 5’-GGGATACCTGTCAAGGGATGCCAGTCTGG, positions 54–73 of the 2C24 sequence (9), with a Kpn I site at the 5’-end (exon 1 sense; underlined) and 5’-GGGATACCCCAA-GTGCTCTCGTTTAAGG, positions 1425–1445 of the 2C24 sequence (8), with a Sal I site at the 5’-end (underlined; exon 9 antisense). Denaturation was for 45 sec at 94°C, with annealing for 25 sec at 55°C and extension for 3 min at 72°C. Thirty-five amplification cycles were performed in a total volume of 100 μl, with 100 pmol of each of the primers. The PCR products were digested with restriction enzymes Kpn I and Sal I (the corresponding restriction sites are present at the 5’-end of the primers) before cloning in pGEM-3Z (Promega) and sequencing with the use of Sequenase (United States Biochemical).

The primers for PCR amplification used to determine the borders of the scrambled transcript (see Fig. 3) were as follows: exon 1 sense, same as above; exon 3 antisense, 5’-GGGAT-CCGCTCTTCCAAAAGACAGGAG, positions 403–423 of the 2C4 sequence (8), with a BamHI site at the 5’-end (underlined); exon 4 sense, 5’-TGGATACCTGGTTGCTTTCGTAATAC, positions 625–645 of the P450 2C24 sequence (8), respectively. Of the 20-μl cDNA synthesis mixture, 1 μl was directly used for PCR amplification.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.
FIG. 1. Schematic representation of a possible trans-splicing event that would result in a P450 2C24 mRNA molecule having the donor site of exon 4 joined to the acceptor site of exon 2. The intermolecular association of the two pre-mRNAs might be facilitated by sequences within intron 1 and intron 4. This scheme takes into account reports showing that the efficiency of trans-splicing of artificial mammalian pre-mRNAs in vitro depends on the sequence complementarity between intronic segments (11, 12). Exons are indicated by open boxes and introns by solid lines in the upper pre-mRNA and by closed boxes and wavy lines in the lower pre-mRNA. Each exon is numbered at the top.

In addition to these primers, a nested exon 3 antisense primer, 5'-GCGGATCCGAAGCTCCTTCCATTTTG, positions 309–328 of the 2C24 sequence (8) with a BamHI site at the 5'-end (underlined), and a nested exon 4 sense primer, 5'-GAGAACTCAATGAAAACTTTAAAATC, positions 551–577 of the 2C24 sequence (8), were also used to determine whether the scrambled transcript is polyadenylated (see Fig. 4a). In Fig. 4b the primers used were as follows: exon 2 sense, 5'-AAGCAGTAAAGGAAGCCCTGGA, positions 198–219 of the 2C24 sequence (8); exon 2 antisense, 5'-GGGATC- CACCCACAGTAGGCTTCGA, positions 166–186 of the 2C24 sequence (8) with a BamHI site at the 5'-end (underlined); exon 4 sense, same as the nested exon 4 primer used in Fig. 4a; and exon 4 antisense, same as the primer used for cDNA synthesis. Denaturation was for 1 min at 94°C, with annealing for 1 min at 59°C and extension for 1 min at 72°C. Thirty-five amplification cycles were performed in a total volume of 100 μL, with 100 pmol of each of the primers. In some experiments the annealing temperature was increased to 62°C to minimize the nonspecific bands observed. When the nested primers were used (see Fig. 4a, lanes 1 and 2), only 25 amplification cycles were performed. The PCR products from the exon 3 antisense/exon 4 sense amplification (lane 4 in Fig. 3a) were digested with restriction enzyme BamHI (this restriction site is present at the 5'-end of the primers) before cloning and sequencing. In Fig. 4c, 1 μg of kidney RNA was incubated for 10 min at room temperature with 3 × 10−8 to 3 × 10−7 units of phosphodiesterase I (Boehringer Mannheim), extracted with phenol/chloroform, precipitated with ethanol, and subjected to cDNA synthesis and PCR amplification with the same primers as in Fig. 3a, lane 2. For the detection of 2C24 transcripts in liver and kidney (see Fig. 5) the primers used were as follows: for the scrambled transcript, same as in Fig.

FIG. 2. (a) Agarose gel electrophoresis of PCR amplification products using an exon 1 sense and an exon 9 antisense primer on oligo(dT) reverse-transcribed kidney RNA. Lane 1, molecular weight markers; and lane 2, PCR products. The primer strategy for this experiment is also shown in a diagram. (b) The sequence of the 1-kb PCR product (lower band on the gel) revealed that it contains exon 1 directly joined to exon 5.
Fig. 3. (a) Comparison of the exon 4 and the exon 5 antisense primer for the ability to produce, from kidney RNA, cDNAs corresponding to the scrambled transcript. The individual lanes on agarose gel electrophoresis are as follows: lane 1, molecular weight markers; lane 6, PCR amplification products using the exon 1 sense and the exon 3 antisense primer, to detect the canonical transcript, on cDNA generated by the exon 4 antisense primer; lane 5, PCR amplification products using the same primers as in lane 6, on cDNA generated by the exon 5 antisense primer; lane 4, PCR amplification products using the exon 4 sense and the exon 3 antisense primers, to detect the scrambled transcript, on cDNA generated by the exon 4 antisense primer; lane 3, PCR amplification products using the same primers as in lane 4, on cDNA generated by the exon 5 antisense primer; lane 2, PCR amplification products using the exon 1 and the exon 4 sense primers in combination with the exon 3 antisense primer, to detect both the canonical and the scrambled transcript, on cDNA generated by the exon 4 antisense primer; and lane 1, PCR amplification products using the same three primers as in lane 2, on cDNA generated by the exon 5 antisense primer. The additional minor bands detected on the gel may represent misprimed PCR amplification products. The primer strategy for this experiment is also shown in a diagram. (b) The sequence of the 450-bp PCR product from lane 4 (major band in that lane) revealed that it contains the donor site of exon 4 joined to the acceptor site of exon 2.

4a; for the canonical/exon-skipped transcript, initial amplification, same as in Fig. 2, nested amplification, exon 1 sense, 5'-GGGATCCGTTCAAGTTTCTGTCTTCTCTTTCTCACTG, positions 92–114 of the 2C24 sequence (9) with a BamHI site at the 5'-end (underlined), and exon 5 antisense, 5'-GGCTAGCT-TATGGCTTCTGGGAGATAATCG, positions 625–648 of the 2C24 sequence (8) with a SalI site at the 5'-end (underlined). The annealing temperature for this nested amplification was increased to 68°C.

RESULTS

The observation of a cytochrome P450 transcript containing the donor site of exon 4 joined to the acceptor site of exon 2 prompted the experimental analysis of several possible interpretations of this phenomenon. Originally, this transcript was thought to result from either the expression of a duplicate exon 2 to exon 4 genomic segment or an intermolecular splicing (trans-splicing) of two 2C24 pre-mRNAs (9).

To determine whether a 2C24 gene duplication is present in the genome, rat chromosomal DNA was digested individually with several restriction enzymes and hybridized with an exon 2 probe. The results of the Southern blot always revealed a single hybridizing band, suggesting that this segment of the gene is not duplicated (data not shown). If, on the other hand, the scrambled transcript results from a trans-splicing event, it is likely that this would generate two products: a long mRNA having a duplicate exon 2 to exon 4 segment and a short mRNA lacking that segment and having exon 1 directly joined to exon 5 (Fig. 1). To investigate whether the proposed trans-spliced transcripts are indeed synthesized, PCR amplification was performed on oligo(dT) reverse-transcribed kidney RNA, using exon 1 sense and exon 9 antisense primers. In addition to the canonical transcript of nine exons, a significant proportion of the resulting PCR products (8 out of 51 cloned PCR products) were found to lack exons 2, 3, and 4 and had exon 1 directly joined to exon 5 (Fig. 2a). None of the cloned PCR products, however, had a duplication of the exon 2 to exon 4 segment. Furthermore, no such duplication was detected when nested primers were used in a second amplification of PCR products longer than the canonical transcript, in an effort to increase the sensitivity of the assay.

Because an mRNA having a duplicate segment could not be detected with the methodology used, it became imperative to determine the exact borders of the scrambled transcript. For that purpose, an exon 4 and an exon 5 antisense primer were compared for their ability to produce, from kidney RNA, cDNA corresponding to the scrambled transcript. Both primers were shown to permit the detection, by PCR, of the canonical mRNA. On the other hand, only the use of the exon 4 antisense primer for cDNA synthesis allowed the detection of the scrambled transcript, with no PCR product being observed when the exon 5 antisense primer was used. These results therefore suggest that the 3' limit of the transcript is exon 4 (Fig. 3a). Moreover, the level of expression of the scrambled transcript appears to be significantly lower in comparison with that of the canonical mRNA (Fig. 3a, lane 2), which correlates well with the fact that Northern analysis has revealed the presence of only the canonical transcript (8). Furthermore, only the use of random hexamers, and not of an oligo(dT) primer, during reverse transcription allowed the detection of the scrambled transcript, suggesting that this molecule is not polyadenylated (Fig. 4a). Finally, PCR amplification was attempted on the cDNA generated by the exon 4 antisense primer, using
Fig. 4. (a) Comparison of random hexamers and an oligo(dT) primer for their ability to produce, from kidney RNA, cDNAs corresponding to the scrambled transcript. The individual lanes on agarose gel electrophoresis are as follows: lane 7, molecular weight markers; lane 6, PCR amplification products using the exon 1 sense and the exon 3 antisense primers, on cDNA generated by the oligo dT primer; lane 5, PCR amplification products using the same primers as in lane 6, on cDNA generated by the random hexamers; lane 4, PCR amplification products using the exon 4 sense and the exon 3 antisense primers, on cDNA generated by the oligo dT primer; lane 3, PCR amplification products using the same primers as in lane 4, on cDNA generated by the random hexamers; lane 2, PCR amplification products using the nested exon 4 sense and the nested exon 3 antisense primers on 1 μl of the PCR products shown in lane 4; and lane 1, PCR amplification products using the same nested primers as in lane 2, on 1 μl of the PCR products shown in lane 3. Note that lane 1, but not lane 2, reveals the presence of a DNA band of the size expected for the scrambled transcript. The identity of this band with the scrambled transcript was confirmed by restriction enzyme Sau3AI fingerprinting as well as by sequencing, after cloning into the pGEM-T vector. The primer strategy for this experiment is also shown in a diagram. (b) Autoradiography of PCR products generated by sense and antisense primers corresponding to either exon 2 (lane 1) or exon 4 (lane 2), on kidney RNA reverse-transcribed by the use of the exon 4 antisense primer. After gel electrophoresis, the DNA was transferred to Hybond filter paper and hybridized with a full-length P450 2C24 cDNA, under conditions similar to the ones described before (7). The horizontal arrow indicates the presence of the additional hybridizing band (lane 2). Sequencing of the major hybridizing PCR products from lanes 1 and 2, after cloning into the pGEM-T vector, established the presence of exons 2, 3, and 4, spliced at correct sites, and the joining of the donor

primers in both orientations of exon 2 and exon 4. The exon 2 sense-antisense amplification revealed that the order of the exons in the scrambled transcript is 2-3-4-2, whereas the exon site of exon 4 to the acceptor site of exon 2. The primer strategy for this experiment is also shown in a diagram. (c) Increased resistance of the scrambled transcript to phosphodiesterase I treatment. Kidney RNA was exposed to increasing amounts of phosphodiesterase I, then reverse-transcribed with the exon 4 antisense primer and PCR amplified using as primers the exon 1 sense, the exon 3 antisense, and the exon 4 sense. A sample of the resulting PCR products was run on a 2% agarose gel. The individual lanes are as follows: lane 1, no cDNA added; lane 2, kidney PCR products; lane 3, liver PCR products; and lane 4, molecular weight markers. The identity of the major band observed with the canonical and the exon-skipped mRNA was confirmed by sequencing, after cloning into the pGEM-T vector. (d) Scrambled transcript. Two rounds of PCR amplification were performed on the cDNA generated with random hexamers using the same primers as in Fig. 4a (initial amplification, 30 cycles; nested amplification, 30 cycles). Of the products of the nested amplification, 10 μl were run on a 2% agarose gel. The individual lanes are as follows: lane 1, no cDNA added; lane 2, kidney PCR products; lane 3, liver PCR products; and lane 4, molecular weight markers. The identity of the major band observed with the scrambled transcript was confirmed by sequencing, after cloning into the pGEM-T vector. The primer strategy for this experiment is also shown in a diagram.
Trans-splicing was recently shown to have the potential to occur in vivo in higher eukaryotes (13), and there are several reports of possible trans-splicing phenomena in mammalian cells (14–17). However, the 2C24 scrambled transcript is composed of only exons 2, 3, and 4 and does not contain the remaining exonic segments that a trans-splicing event would be likely to generate. In addition, the other product of such a trans-splicing event, the observed mRNA that has exon 1 directly joined to exon 5, could simply be the result of exon skipping of the 2C24 pre-mRNA. This, in fact, would represent the first reported observation of exons being deleted in the expression of cytochrome P450 genes.

Furthermore, the capability of PCR primers, corresponding to both orientations of either exon 2 or exon 4, to detect the scrambled transcript (Fig. 4b) provides evidence that this molecule could be a circular RNA. The alternative interpretation is that the transcript is a linear molecule containing a tandem repeat of the exon 2 to exon 4 segment. However, the latter possibility is inconsistent with the genomic hybridization data, the lack of detection of such a duplication by PCR using exon 1 and exon 9 primers, and the absence of polyadenylation in the scrambled transcript. Moreover, as shown in Fig. 4b, an additional hybridizing band of ~900 bp can be detected. This might represent amplification from a longer cDNA generated by an additional round of reverse transcription of the circular template (474 bases). In addition, the observed increased resistance of the scrambled transcript to the 3′-exonuclease phosphodiesterase I (Fig. 4c) is indicative of the absence of free 3′-OH ends and consistent with the possibility that this molecule could be circular.

Finally, it was investigated whether the presence of the scrambled transcript correlates with that of the exon-skipped mRNA not only in kidney but also in liver, the other tissue with relatively high levels of expression of the canonical P450 2C24 mRNA containing nine exons (8). It was found that liver expresses both the exon-skipped and the scrambled transcript and at levels that are comparable to the ones of the kidney, as seen in Fig. 5a and b.

**DISCUSSION**

The experiments described in this paper reveal that the exons composing the scrambled transcript in cytochrome P450 2C24 are identical to the ones that are absent in a transcript differentially processed by exon skipping. This novel observation of an association between the exons present in a scrambled transcript and the exons deleted during differential splicing suggests that a relationship between the two phenomena might exist.

A possible mechanistic interpretation of these experimental data could be related to the recently described phenomenon of inverse splicing of group II introns (18). In this scheme, intron
4 would form a 2'-5' phosphodiester bond with intron 1 (branched intermediate), allowing the donor site of exon 4 to be spliced together with the acceptor site of exon 2, resulting in a circular molecule. Subsequently, the branched intron could be removed by a splicing event that generates the mRNA in which the donor site of exon 1 is linked to the acceptor site of exon 5 (Fig. 6).

An additional interpretation could involve the mechanism of alternative processing of the 2C24 pre-mRNA, which would be likely to generate a lariat containing exons 2, 3, and 4. Removal of the introns from the lariat would result in the joining of exons 2, 3, and 4 and the formation of a circular molecule (Fig. 6). In line with this, it is worth noting that synthetic circular pre-mRNAs have been shown to be spliced correctly in vitro (19). Furthermore, the demonstration that higher eukaryotic cells have the potential for trans-splicing both in vivo and in vitro (11–13) is consistent with the possibility of branched intermediates being capable to undergo additional splicing events.

Although at present these two interpretations cannot be distinguished, both the inverse splicing and the lariat splicing hypotheses suggest that exon scrambling is an event that occurs concomitantly with exon skipping. Accordingly, it would be expected that two or more exons skipped during alternative pre-mRNA processing could be present in a circular molecule that has the donor site of the 3'-exon joined to the acceptor site of the 5'-exon. Furthermore, if only one exon is skipped, this might result in the generation of a circular molecule composed of that single exon, joined at the 5' - and 3' -splice junctions, as is the case of the circular Sry transcript in adult mouse testis (4). Even though P450 2C24 is the first example of a gene subjected to a reciprocal exon-scrambling/exon-skipping event, this could represent a generalized phenomenon that might be revealed in additional cases where genes are known to undergo exon skipping or exon scrambling. Possibly, variable structural characteristics of the pre-mRNAs, such as inverted repeats in the introns flanking the exons that are to be skipped, might stabilize the intermediates involved, allowing exon skipping and exon scrambling to occur (20, 4). However, the 1:1 ratio between the scrambled and the corresponding exon-skipped transcript suggested on mechanistic grounds might not remain constant, because of a likely differential stability between the linear and the circular molecule. Furthermore, assuming that the lariat splicing pathway is indeed occurring, it is not certain that all lariats would splice out their introns, resulting in circularization. Possibly most of them would be debranched and degraded before an intron removal event (21).

Finally, an obvious question is whether there is a biological function associated with the circular scrambled transcripts. Although at present there is no direct evidence suggesting this, it should be noted that such molecules have the capability of being transported to the cytoplasm, where they accumulate (3, 4). It is also worth noting that although exon skipping has the potential to generate proteins with novel functions (22), it is unlikely that the P450 2C24 exon-skipped product could act as a P450 catalyst because of the large deletion that it harbors. However, a possible role of the scrambled transcripts as a feedback regulator of alternative pre-mRNA splicing cannot be excluded.

Professor Rune Toftgård is acknowledged for support and encouragement. This research was supported by grants from the Lars Hierts Minne Foundation, the Åke Wibergs Foundation, and the Karolinska Institute.