Proc. Natl. Acad. Sci. USA
Vol. 96, pp. 5400–5405, May 1999
Biochemistry

Exon repetition in mRNA

SIMON A. FRANTZ*, AMRIK S. THIARA*, DAVID LODWICK†, LEONG L. NG†, IAN C. EPERON‡, AND NILESH J. SAMANI*§

Departments of *Cardiology, †Medicine and Therapeutics, and ‡Biochemistry, University of Leicester, Leicester LE3 9QP, United Kingdom

Edited by Joan A. Steitz, Yale University, New Haven, CT, and approved February 26, 1999 (received for review October 26, 1998)

ABSTRACT **The production of different transcripts (transcript heterogeneity) is a feature of many genes that may result in phenotypic variation. Several mechanisms, that occur at both the DNA and RNA level have been shown to contribute to this transcript heterogeneity in mammals, all of which involve either the rearrangement of sequences within a genome or the use of alternative signals in linear, contiguous DNA or RNA. Here we describe tissue-specific repetition of selective exons in transcripts of a rat gene (SA) with a normal exon–intron organization. We conclude that nonlinear mRNA processing can generate tissue-specific transcripts.**

Many genes express tissue-specific isoforms of mRNA that may result in an alteration in the production and/or function of its protein product. Several mechanisms that occur at both the DNA and RNA level may contribute to this transcript heterogeneity in mammals. At the DNA level, these include rearrangements of the DNA itself or the use of multiple transcriptional start sites. At the RNA level, alternative splicing of the primary transcript and variability in the site and length of the 3′ poly(A)+ tail are important mechanisms. All of these processes involve either the rearrangement of sequences within a genome or the use of alternative signals in linear, contiguous DNA or RNA. The only major exceptions occur during recombination or in trans-splicing. Trans-splicing has not been demonstrated conclusively to occur with endogenous gene products in mammals, nor has it been implicated in tissue-specific processing.

The rat SA gene, a putative hypertension-related gene of as-yet-undefined function (1–3), shows markedly greater expression in the kidney and liver of the spontaneously hypertensive rat (SHR) compared with respective tissues of the normotensive Wistar-Kyoto (WKY) rat (1, 4). In the course of studies on the regulation of the gene, we observed additional, major transcripts in the kidney of the WKY rat that were not present in either its liver or in the kidney or liver of SHR. This indicated the occurrence of strain and tissue-specific heterogeneity of transcripts that were not related to the steady-state abundance of the mRNA. In this report, we demonstrate that the additional transcripts in the WKY kidney contain tandem repetition of specific exons and provide evidence that this could not occur through any known linear processing mechanism. Exon repetition has been described previously only in products of the rat carnitine octanoyltransferase gene and has been attributed to trans-splicing (5). Our investigation of the SA gene confirms the existence of exon repetition and shows that it arises at the RNA level and that it can be regulated in specific tissues.

MATERIALS AND METHODS

Northern Blot Analysis. Total SHR and WKY kidney and liver RNAs were extracted as described (6). Northern blot hybridization was carried out using 60 μg of RNAs probed with a 1.6-kb rat SA cDNA fragment (2).

Reverse Transcription–PCR (RT-PCR) Analysis. Total RNAs were reverse transcribed with oligo $d(T)_{12–18}$ using Moloney murine leukemia virus reverse transcriptase (Life Technologies, Grand Island, NY). cDNAs were amplified with *Taq* polymerase (Bioline, London) using primer pairs as indicated. After size fractionating on 1% agarose gels, the RT-PCR products were isolated and cloned into pGEM-T (Promega). Positive recombinant plasmids were isolated and fully sequenced (7) with a T7 sequencing kit (Amersham Pharmacia). The 2-kb WKY cDNA was the source of probe in subsequent Southern, dot-blot, and RNase H analyses.

RNase Protection Assay. The 195-bp SA12/SAX5 RT-PCR product (see Fig. 2*c*) was cloned into pGEM-T, and radiolabeled probe was prepared by *in vitro* transcription (Promega). Poly(A)+ RNA was isolated from each tissue and hybridized to 5×10^5 cpm of each probe overnight at 55°C. Unhybridized nucleic acids were digested with 5 units of RNase One (Promega). After RNase inactivation and precipitation of the protected RNA fragments, each sample was analyzed on a 6% polyacrylamide gel, and bands were detected by using a PhosphorImager (Molecular Dynamics).

RNase H Mapping Analysis. Total RNA was mapped as described (8). Oligonucleotide primer SA4 (150 pmol) and total RNA (80 μg) were mixed and denatured at 58°C for 1 min. Digestion was then performed with 4 units of RNase H (GIBCO) at 37°C for 20 min. After RNase H inactivation, precipitated mRNA products were analyzed by Northern blotting and probed with the 2-kb SA cDNA fragment.

Determination of SA Genomic Structure. The 5′ and 3′ ends of SHR SA cDNA were first determined by using 5′ and 3′ rapid amplification of cDNA ends. The 5′ end was congruous with the published cDNA sequence (1, 9), the 3′ sequence extended by 500 bp (GenBank accession no. AF027188). An SHR splenic genomic library was prepared by cloning DNA partially digested with *Mbo*I into the *Bam*HI site of the phage EMBL-3 SP6/T7 (CLONTECH). Clones containing SA gene fragments were identified by colony hybridization and subcloned into pBluescript (Stratagene). Exon–intron boundaries were determined by sequencing the clones by using primers spanning the cDNA sequence. Intron sizes were determined by PCR amplification of the clones using primer pairs flanking individual introns. In addition, a genomic fragment of 8.8 kb flanking exons 2–4 was fully sequenced to verify the exon–intron boundaries in this region of the gene.

PCR Comparison of SHR and WKY SA Genes. Regions of the SA gene were amplified under long-range PCR conditions (10) using the primer pairs SA1/SAX5′ (Fig. 4*b*, arrow a), SA1/SA10 (Fig. 4*b*, arrow b), SA3/SA8 (Fig. 4*b*, arrow c),

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.

PNAS is available online at www.pnas.org.

This paper was submitted directly (Track II) to the *Proceedings* office.
Abbreviations: SHR, spontaneously hypertensive rat; WKY, Wistar Kyoto rat; PT, proximal tubule; RT-PCR, reverse transcription–PCR.
Data deposition: The sequence reported in this paper has been deposited in the GenBank database (accession no. AF027188).
§To whom reprint requests should be addressed at: Department of Cardiology, Glenfield General Hospital, Groby Road, Leicester LE3 9QP, United Kingdom. e-mail njs@le.ac.uk.

SA7/SAEX2 (Fig. 4*b*, arrow d) and SF4/SA3′ (Fig. 2*a*). The region between exons 1 and 5 was studied in more detail by amplifying SHR and WKY kidney and liver genomic DNA with the primer pairs SA1/SAX5 (exon 1–2), SA1/SA4 (exon 1–3), SA1/SA6 (exon 1–4), SA12/SA10 (exon 2–5), SA3/SA10 (exon 3–5), and SA5/SA10 (exon 4–5). Products were separated on 0.5% agarose gels, transferred to Hybond N membrane (Amersham Pharmacia) and hybridized to the 2-kb SA cDNA probe.

Restriction Mapping of the SHR and WKY SA Genes. SHR and WKY kidney genomic DNA (5 μg) were digested with the restriction enzymes *Bgl*II, *Stu*I, *Sac*I (*Sst*I), and *Xba*I. After Southern blotting, products were hybridized initially to the 2-kb SA cDNA probe, followed by hybridization to probes derived by PCR from SA cDNA with the following primers: SA1/SAX5 (exon 1–2), SA12/SA4 (exon 2–3), SA3/SA6 (exon 3–4), SA5-SA8 (exon 4–10), and SA7/SAEX2 (exon 10–15).

Dot-Blot Analysis. Serial dilutions of 4, 2, 1, 0.5, and 0.25 μg of SHR and WKY kidney and liver genomic DNA were applied onto a Hybond N membrane and hybridized sequentially to an SA cDNA probe and then hybridized to a glyceraldehyde-3-phosphate dehydrogenase probe. Autoradiographs were analyzed by scanning, and pixel intensities were measured.

Isolation and Analysis of Proximal Tubular Cells. Freshly dissected kidney cortices from SHR and WKY rat were finely minced and enzymatically digested as described (11). Crude tubule segments were recovered by brief centrifugation, and layered onto a preformed 43% Percoll gradient. After centrifugation at 1,100 $\times$ *g* for 20 min, the band containing proximal tubule (PT) segments was recovered at a density of 1.076–1.088 g/ml (11). Phase contrast microscopy was performed on a Nikon TMS microscope to confirm the morphology of the isolated segments. An aliquot of each tubule preparation was checked for the brush border enzyme markers characteristic of proximal tubules by using established methods (12) (data not shown). In addition, enrichment for PT segments was analyzed by using Western blotting for the newly described sodium bicarbonate cotransporter NBC3, which is only present in PT cells (13). Protein samples (50 μg) were resolved on SDS/7.5% PAGE gels and, after blotting, were incubated with the NBC3-specific antibody G186, followed by detection with horseradish peroxidase-conjugated goat-anti-rabbit IgG and enhanced chemiluminescence (Amersham Pharmacia). The antibody G186 was raised in a rabbit by using a keyhole limpet hemocyanin conjugate with the peptide sequence IRIEPPKSLPSSDKR [amino acids 245–359 of the rat NBC3 sequence (13)]. Total cellular RNA was extracted from PT cells and analyzed by using RT-PCR (see above) for exon repetition in SA mRNA. Finally, genomic DNA was extracted and analyzed by using Southern blotting of restriction fragments as described above to look for genomic rearrangement.

Primer Sequences. All primers used were designed by using OLIGO (14) and based on the published SHR SA cDNA sequence (1, 9) or our 3′ rapid amplification of cDNA ends sequence (GenBank accession no. AF027188). The numbers in parenthesis represent the nucleotide position relative to the transcriptional start site. SA1, TGGCTTTGTGTGGGATTAAG (7–26); SA2, TGTTGTCCATTCCTTTCTCC (2025–2006); SAEX2, GGGCCAAGTTTAAGTTGTTG (2040–2021); SA3, GTGTTTTCACCGCCTAGCAATCCCTGATCC (312–341); SA3′, GCTGCTCTATTATCTTTGACATTG (2504–2481); SA4, GGATCAGGGATTGCTAGGCGGTGAAAACAC (341–312); SF4, TCCTGATTACAAACTCC (1851–1867); SA5, GGTGATTCTGCCCAAGATCCCAGAGTGGTG (642–671); SAX5, TCCTGGTGTCCACCTCCTTGTGTGAGAAG (157–129); SA6, CACCACTCTGGGATCTTGGGCAGAATCACC (671–642); SA7, CAATGGGGAAGCCCTCTCCTGCTTTTAATG (1460–1489); SA8, CATTAAAAGCAGGAGAGGGCTTCCCCATTG (1489–1460); SA10, CAGGAGTGCTGAGACACAAT(875–856); and SA12, AGATCACTGACTTGTGAGCT (111–130).

RESULTS

Identification of Exon Repetition. The SA gene in the rat is expressed primarily in the proximal tubules of the kidney (15, 16) and at a lower level in liver hepatocytes. Northern blot analysis showed that over half of the SA mRNA in the kidney of the WKY rat appeared to be longer than the corresponding mRNA from the liver (Fig. 1). In contrast, the much more abundant SA mRNA from both tissues of the SHR rat lacked the longer transcripts.

Further analysis using RT-PCR with primers directed to the ends of the published cDNA sequence (1) showed two additional products only in the kidney of the WKY rat (Fig. 1*b*, bands A and B). Sequence analysis of the two products confirmed they originated from SA mRNA and that each contained additional sequences that were precise tandem repetitions of regions of the cDNA. Based on our sequence analysis of the SA gene from the SHR rat (Table 1), the additional sequences appeared to be precise duplications of specific exons. PCR product A contained a tandem duplication of exon 2, whereas product B contained a duplication of exons 2–4 compared with the predicted transcript (Fig. 1).

Confirmation of SA Exon Repetition. To confirm that the additional products seen in the WKY kidney were genuine transcripts and not some form of PCR priming artifact, three approaches were used. First, RT-PCR amplification of RNAs was carried out by using primers specifically designed to produce unique fragments only in the presence of the exon duplications. By using primer pairs located in exons 1 and 2, two additional products of expected sizes were only seen in the WKY kidney (Fig. 2*a*). In addition, by using either a 5′ primer

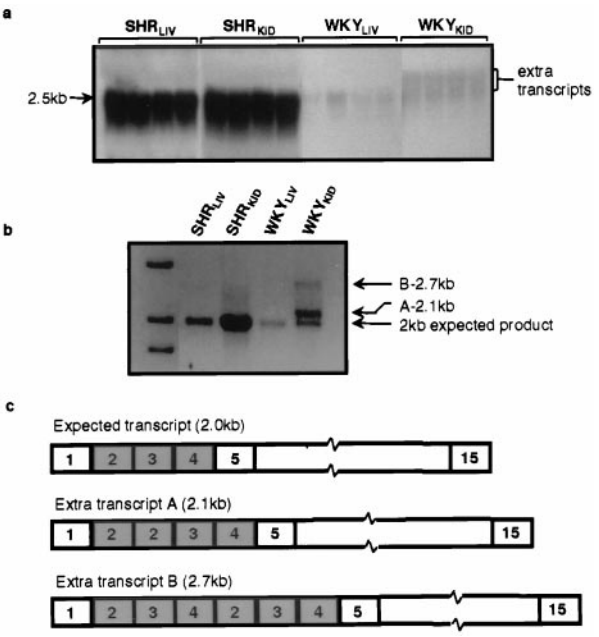

FIG. 1. Analysis of SA transcript heterogeneity. (*a*) Northern blot analysis of SA expression in SHR and WKY liver and kidney RNA. (*b*) RT-PCR amplification of SA transcripts. Oligo(dT)-primed cDNA from SHR (2 μg) and WKY (5 μg) liver and kidney total RNA were amplified by using the SA-specific primers SA1 and SA2. The extra products in WKY kidney are marked by arrows *A* and *B*. (*c*) Sequence analysis of WKY kidney SA transcripts. Schematic shows a summary of the sequence data from the three transcripts seen in WKY kidney (see Fig. 1*b*). Numbered boxes correspond to exons.

Table 1. Exon/Intron sizes and boundaries of SHR SA gene

Exon	Length, bp	Exon sequence	Intron sequence, 5′ end	Intron length, bp	Intron sequence, 3′ end	Exon sequence	Exon
1	67	AGTATCTTTA	**GTGAGT**	**6500**	**TCTTTGCAG**	GTGAAAAACA	2
2	148	CCAAATCC**AT**	**GTAAG**A	834	**GTCTTTTAG**	CTGGAGTAGT	3
3	274	TACAGAAAA**AG**	**GTATG**G	611	**TTCCCTCAG**	ACTGGAAAAA	4
4	211	CTGCGAAC**AG**	**GTTC**GT	**4500**	**TTCCTTTAG**	GGACAGTTTT	5
5	208	AGATGATGAA	**GTGAGT**	**560**	**CTTTGGCAG**	ATATGCCAGT	6
6	144	TCAACGGAA**G**	**GTATTT**	206	**GTATTCCAG**	GTTCTGGCTG	7
7	157	CATCTTGCAA	**GTAAG**G	**320**	**CAACCGCAG**	ACCCTCTCCA	8
8	80	ACATAACC**AG**	**GTAAG**A	**1530**	**TTGGAGAAG**	CTATAAGTTC	9
9	124	GACAGAAAC**G**	**GTACCT**	**1000**	**CTTGCCTAG**	GTGCTGATCT	10
10	81	TAATGTGG**AG**	**GTTTGT**	**1230**	**TTATTCTAG**	ATTTTAGATG	11
11	102	TCATTATGTA	**GTAAG**A	**1050**	**TTTTTGCAG**	GATAATCCTT	12
12	128	TATCCTCTG**G**	**GTAATT**	80	**TTTTCCTAG**	TTACCGAATT	13
13	100	CAGAGGAG**AG**	**GTAAGC**	**5440**	**TTACTTTAG**	GTAGTAAAGG	14
14	120	CCCCAGGA**AG**	**GTAGGT**	**410**	**TACCAACAG**	ATAGAATTTA	15

Exon and intron sizes are shown in their respective boxes. Intron sizes in **bold** are approximate sizes as determined by agarose gel analysis of PCR products; otherwise sizes were determined by sequencing of phages. Exon/intron nucleotides obeying splice site consensus sequences (17) are shown in **bold**.

located in exon 4 and a 3′ primer located in exon 3 (Fig. 2*b*) or both primers located in exon 2 (Fig. 2*c*), we found products of the predicted sizes in the WKY kidney only. Sequence analysis of these additional products confirmed their origin from transcripts containing the duplicated exons.

In a second approach, one of the unique RT-PCR products was used as a probe in RNase-protection experiments. The 195-bp product containing the exon 2 duplication (Fig. 2*c*) was cloned and transcribed in the complementary sense with nonspecific flanking sequence. This probe contained the whole of exon 2 (148 bp) in two parts, with a duplication of 47 bp (Fig. 3*a*) The [^{32}P]RNA was then annealed to poly(A)+ RNA prepared from the various tissues and digested with RNase. The protected

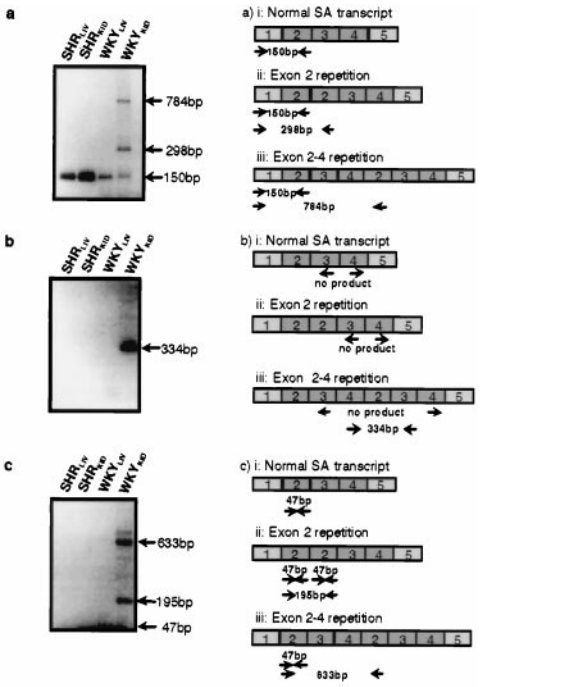

FIG. 2. RT-PCR demonstration of exon duplication in WKY kidney SA mRNA. The diagrams on the *Right* show RT-PCR products predicted by using exon 1 (SA1) and exon 2 (SAX5) primers (*a*), exon 5 (SA5) and exon 4 (SA4) primers (*b*), and exon 2 (SA12 and SAX5) primers (*c*), if the WKY kidney transcripts are as summarized in Fig. 1*c*. *Left* shows corresponding RT-PCR analysis of SHR and WKY liver and kidney total RNA.

products produced by normal mRNA were predicted to be 105 and 90 nt long, but a product of 148 nt could be formed if exon 2 of the mRNA looped around to base pair with both consecutive portions of the exon 2 in the probe. The SHR mRNA samples produced products of 105 and 148 nt (Fig. 3*b*). In contrast, a major product formed by WKY kidney mRNA was reproducibly about 195 nt, the length predicted only if the mRNA contained a duplication of exon 2.

Proof of the repetitions of exons 2–4 was demonstrated by a third method. Total RNA was digested with RNase H in the presence of an oligonucleotide complementary to exon 3 and then SA fragments detected after Northern blotting with the 2-kb SA cDNA probe. In the absence of exon repetition, two fragments should be formed (Fig. 3*bi* and *bii*, lane SHR_{KID}). Repetition of exons 2–4 would lead to the liberation of a novel RNA fragment of approximately 581 nt (Fig. 3*bi*). A band of this size was detected only in WKY kidney RNA, as predicted (Fig. 3*bii*, lane WKY_{KID}). The fainter band at 489 nt in WKY kidney RNA (Fig. 3*bii*) is of the size expected from the 5′ portion of mRNA containing the exon 2 duplication.

These results demonstrate through several independent methods that the WKY kidney SA mRNA exhibits tandem duplication of specific exons, a phenomenon we term exon repetition.

Analysis of SA Gene Structure. Exon repetition in mRNA could arise from changes in genomic organization in the WKY rat. Specifically, exons 2–4 might have been duplicated, with the duplicated exons being skipped during splicing of some of the transcripts, or the entire gene might have been duplicated and the two copies transcribed in one pre-mRNA (cf. ref. 18), with splicing removing most of the exons and the intergenic sequence. These rearrangements might be incorporated in the germ line, with pre-mRNA splicing hiding any manifestations in the liver, or the duplications might occur somatically in the kidney proximal tubules.

Genomic exon duplication in the WKY rat was tested by using PCR analysis in which the SHR and WKY SA genes were compared (Fig. 4*a*). Intron A of the SHR SA gene contains a LINE insertion (19). Aside from a difference in fragment size between SHR and WKY of 1.3 kb with some primer pairs as a consequence of this, we found no other difference to suggest exon duplication at the DNA level in the WKY rat. Specifically, primers located in exons 1 and 5 of the gene gave predicted products of 13.2 kb and 11.8 kb in SHR and WKY, respectively (Fig. 4 *a* and *b*), whereas if there had been duplication of the genomic region between exons 2 and 4 in the WKY strain, one would have expected a minimum fragment of 13.9 kb [the

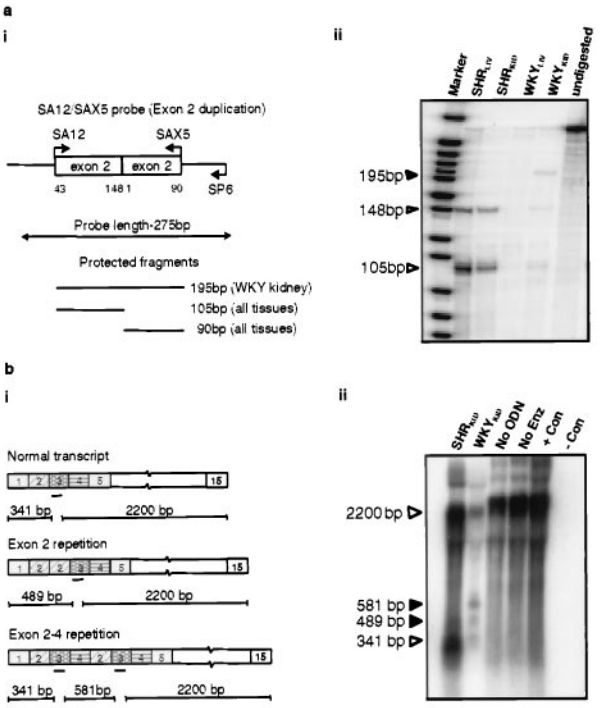

FIG. 3. Confirmation of exon duplication in WKY kidney SA mRNA. (*a*) RNase protection assay confirming repetition of exon 2. (*i*) Schematic of the riboprobe used; the cloned 195-bp fragment from Fig. 2*c* containing duplicated exon 2. Open boxes denote SA exons, and lines correspond to polylinker sequence of the pGEM-T vector. The numbers denote the portions of exon 2 present, numbered from the 5′ end of the exon. Below are the expected protected products from the assay and their location. (*ii*) Analysis of corresponding RNase protection assay products from 5 μg of SHR and WKY poly(A)+ RNA. The 195-nt protected product unique to the WKY kidney is denoted by a black arrowhead, whereas the products expected in all tissues are shown by clear arrowheads. The formation of a product of 148 nt, instead of the 90-nt product predicted, is discussed in the text. (*b*) RNase H digestion assay confirming repetition of exons 2–4. (*i*) The schematic shows the products expected from the various SA transcripts after annealing an oligonucleotide primer located in exon 3 (SA4) to RNA and digesting with RNase H. (*ii*) Corresponding analysis of RNase H digested products from SHR and WKY kidney RNA. The 581-nt product produced as a result of exon 2–4 repetition in the WKY kidney only is shown by the upper black arrowhead. The 341-nt and 2,200-nt products derived from the canonically organized transcript are shown by clear arrowheads. A fainter 489-nt product from repetition of exon 2 also is seen in WKY kidney (lower black arrowhead). The controls (lanes 3–6) include reactions with no oligonucleotide (No ODN), no RNase H (No Enz), and a positive control (+Con) for the Northern blot, all of which show the undigested 2.5-kb SHR SA transcript and a negative yeast control lane (−Con).

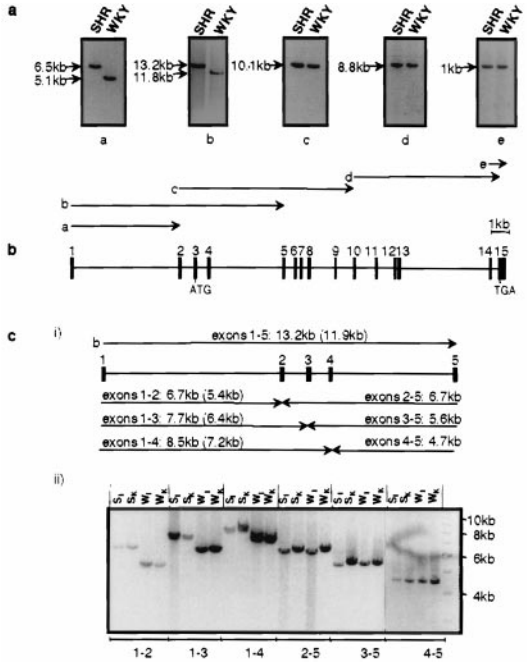

FIG. 4. Comparison by using PCR of the SA gene of SHR and WKY rat. (*a*) Products obtained by amplification across different exon as indicated by arrows a-e in *b* (see *Results* for explanation of size differences). (*b*) The exon/intron organization of the SHR SA gene. Filled rectangles represent exons (see Table 1 for exon and intron sizes). (*c*) Detailed PCR analysis of exons 1–5. (*i*) Diagram of SHR SA exons 1–5 with arrow b corresponding to the PCR product shown in Fig. 5*a*. The arrows below the exon/intron map denote the products expected by amplification across specific exons. Figures in brackets correspond to the (smaller) WKY product sizes caused by the absence of the LINE insertion in intron A. (*ii*) Corresponding PCR Analysis. All PCR products agreed with their predicted sizes and show no evidence of intragenic exon duplication.

additional 2.1 kb being the size of the gene between exons 2–4 (Table 1)].

To exclude the possibility that there might be small duplications of individual exons, more detailed PCR analysis was done as shown in Fig. 4*c*. If an exon (for example, exon 2) had been duplicated, then PCR from exon 1 to the nearest copy of exon 2 (hypothetical exon 2a) and from exon 5 to the nearest copy of exon 2 (hypothetical exon 2b) might not reveal that there were two copies if the distances involved were, fortuitously, very similar to those in the SHR rat. However, the sum of the sizes of the two PCR products would not equal the distance between exons 1 and 5. The results shown in Fig. 4*c* demonstrate that the greater distance is the sum of the two separate distances for all three potentially duplicated exons.

The possibility that the entire gene had been duplicated in the WKY rat was tested by using Southern blotting. A restriction map was generated of the SHR SA gene (Fig. 5), and genomic fragments including exons 2–4 and significant proportions of the flanking introns were fully sequenced (8.8 kb, marked in Fig. 5). Fragments generated with several restriction enzymes for both strains could all be placed within the single copy structure of the gene, allowing for the LINE element unique to the SHR rat and for a known *Stu*I polymorphism (2) (Fig. 5*bi*). The 5′ end of the gene lies within the bands of 3.3 kb and 3 kb in the *Bgl*II and *Stu*I digestions, respectively, both of which are represented equally in the SHR and WKY kidney DNA. The 3′ end can be seen most clearly in the 2.8-kb *Bgl*II, 4.2-kb *Stu*I (upper band) and 11-kb *Xba*I fragments, which again show no indications that there might be a second copy of the gene in the WKY kidney in a different environment. The two enzymes that cut within exon 2 [*Sst*I (*Sac*I) and *Xba*I] gave single bands when probed with an exon 2–3-specific cDNA probe (Fig. 5*bii*). The results confirm the PCR evidence in Fig. 4 showing that there are no intragenic exon duplications, and they argue that the existence of tandem copies of the gene in the germ line is extremely unlikely. An additional test for gene duplication was done by quantitative dot blots (Fig. 5*c*). The ratio of SA to glyceraldehyde-3-phosphate dehydrogenase signals is no greater for the WKY kidney than for the other tissues. We conclude that neither the specific exons nor the entire gene are duplicated in the WKY germ line.

Analysis of Exon Repetition in Isolated Proximal Tubular Cells. *In situ* hybridization (15, 16) and immunocytochemistry (N.J.S. and D.L., unpublished data) show that the SA gene is expressed in PT cells, which comprise about 5% of kidney cells. To exclude the possibility that genomic rearrangement within these specific cells accounts for the observed transcript heterogeneity in the WKY rat, PT segments were isolated from

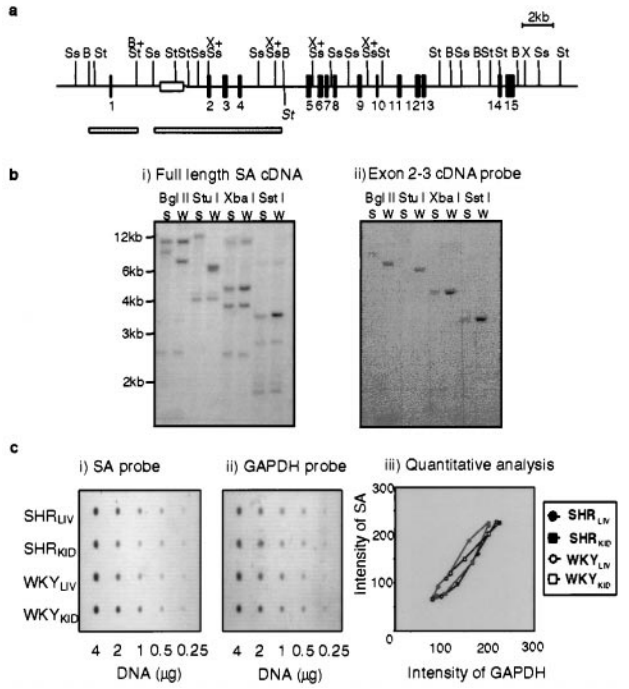

FIG. 5. Restriction mapping of SHR and WKY SA genes. (*a*) Restriction map of the SHR SA gene. The map was constructed by positioning fragments generated by digestion of SHR DNA with *Bgl*II (B), *Stu*I (St), *Xba*I (X), and *Sst*I (Ss) and probing with the 2-kb SA cDNA probe or subprobes corresponding to specific regions of the SA cDNA. Black boxes correpond to exons. The white box in intron A represents a previously characterized LINE insertion (19); the hatched boxes underneath restriction map denotes regions of the SA gene that were fully sequenced, and the *Stu*I point mutation in intron D is shown in *italics*. (*b*) Southern blot analysis of SHR and WKY kidney SA gene structures. All bands could be mapped to the single copy structure of the gene shown in *a*. Because of the small size of exon 1 (67 bp), neither the restriction fragments that contained only exon 1 nor the 500-bp *Stu*I fragment of exon 13 was easily visible on the autoradiograph. The blot was hybridized with a 2-kb SA cDNA probe (*i*) and an exon 2–3-specific SA cDNA probe (*ii*). (*c*) Dot-blot analysis of SHR and WKY liver and kidney SA DNA. SA genomic DNA was quantified by successive hybridization of serial dilutions of each DNA to an SA cDNA probe (*i*), a probe for glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (*ii*). (*iii*) Quantitative analysis of the dot blots is shown.

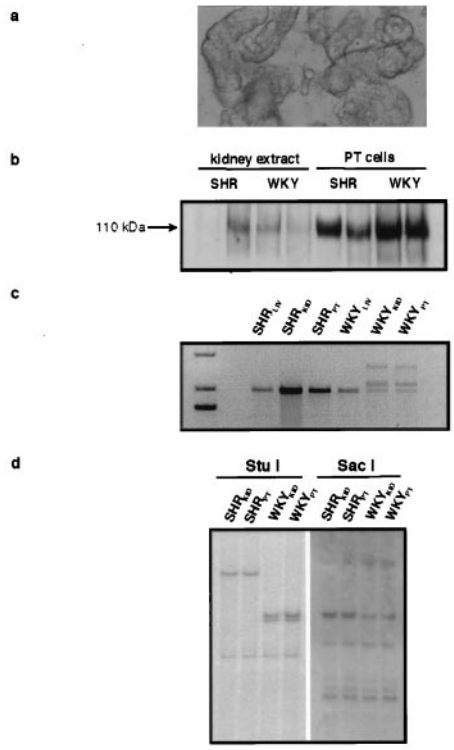

FIG. 6. Analysis of isolated proximal tubular segments. (*a*) Phase contrast microscopy (×400) of isolated segments showing a homogeneous population of large diameter segments with morphology typical of PT cells with granular cytoplasm. (*b*) Western blot of protein extracted from either whole kidney (lanes 1–4) or purified PT segments (lanes 5–8) from SHR (lanes 1, 2, 5, and 6) and WKY (lanes 3, 4, 7, and 8) rats probed for presence of the specific PT protein NBC3. (*c*) RT-PCR analysis of SA mRNA from liver ($_{LIV}$), kidney ($_{KID}$), and isolated PT segments ($_{PT}$) of SHR and WKY. Note the presence of additional PCR products showing exon repetition in RNAs from WKY kidney (WKY_{KID}) and WKY PT cells (WKY_{PT}) only. (*d*) *Stu*I and *Sac*I restriction analysis of SA gene structure in whole kidney and PT segments of SHR and WKY excluding somatic rearrangement of the SA gene in WKY PT cells.

SHR and WKY kidneys to >95% purity by using isopycnic centrifugation (Fig. 6*a*). A >10-fold enrichment for the highly specific PT marker NBC3 was demonstrated by using Western blotting (Fig. 6*b*). RT-PCR analysis of cellular RNAs isolated from these segments confirmed the presence of exon duplication only in the WKY cells (Fig. 6*c*). Analysis of genomic DNA from SHR and WKY rat PT segments (Fig. 6*d*) showed no difference in genetic structure apart from the known polymorphisms (see above). In addition, in each strain there was no difference in restriction pattern between DNAs from the PT and from whole kidney. Specifically, duplication of exons 2–4 in the WKY PT cells would have been expected to produce at least one novel *Sac*I (*Sst*I) fragment because exon 2 contains an *Sst*I restriction site (see Fig. 5*bii*). We conclude that the cells expressing the novel isoforms of SA mRNA in WKY kidney do not have any detectable SA gene rearrangement. Exon repetition is therefore a postgenetic process.

DISCUSSION

The major site of SA mRNA expression in rats is in the renal proximal tubules. It was most unexpected to observe that the majority of SA transcripts in the WKY rat contain extra exons. Nonetheless, the data from Northern blots (Fig. 1*a* and 3*b*), RT-PCR (Figs. 1*b* and 2), and RNase protection (Fig. 3*a*) indicate the unusual transcripts are not minor byproducts of an error-prone reaction or a PCR artifact. Similarly, the absence of the transcripts in the liver, which contains less mRNA overall, argues that the reaction is not a consequence of scarcity. Instead, a very efficient process generates the transcripts, containing repeated exons, specifically in the kidney in a strain-specific manner. A recent report by Caudevilla *et al.* (5) demonstrated the presence of two extra transcripts of the rat carnitine octanoyltransferase gene in which either exon 2 or exon 2 and 3 were repeated. This gene also was found to be single-copy, with no internal duplication of exons. Our findings confirm the phenomenon of exon repetition and furthermore demonstrate that this could be a new mechanism contributing to tissue-specific transcript heterogeneity in mammals.

Exon 2 of the SA gene is located upstream of the putative translation start site (1) and therefore the presence of the duplication would not be expected to alter the protein product. However, the exon 2–4 duplication would alter the reading frame, resulting in a truncated, altered protein product of 157 aa. The function of the SA protein, although possibly involved in blood pressure regulation (20), is currently unknown, and therefore any pathophysiological significance of these modified transcripts in the WKY kidney remains speculative. We have examined some other rat strains and crosses. The transcripts have been detected also in Milan hypertensive and Dahl salt-sensitive strains of rats, and it appears that the property

cosegregates with the chromosome (e.g., with the absence of the LINE insertion) rather than with the normotensive phenotype (S.F. and N.J.S., data not shown).

There have been several previous reports of unexpected products of splicing in which exons have been joined in a scrambled order, although not duplicated (21–23). The transcripts in these cases have usually been minor products expressed at levels of about 1% of the normal mRNA (21, 22) and lacking a poly(A)+ tail (21, 23). Several lines of evidence suggest that these sequences are excised during splicing as circular molecules (21,23–25) either as a consequence of error-prone splicing (25) or as byproducts of exon skipping (23). The existence of circular SA mRNA containing the duplicated exons could not be wholly excluded by using RT-PCR. The PCR products in Fig. 2 *b* and *c* could be derived from amplification of circular SA RNA exons, comprising exon 2 or exons 2–4. We can, however, exclude the presence of circular transcripts on the basis of other results. We infer from the absence of PCR products that if oligo(dT) was omitted during reverse transcription (data not shown), then the SA transcripts were polyadenylated. In addition, small exon circles would not produce the lower mobility bands on the Northern blots, and the RNase protection and RNase H results can only be interpreted as evidence for exon repetition in WKY kidney SA mRNA.

Given that the mRNA contained tandem repeats of exon 2 or exons 2–4, which would not be expected to arise by any known RNA processing reaction, it was important to exclude genomic or somatic rearrangements that could account for the effect. Both by PCR and Southern blotting (Figs. 4–6), we can exclude the possibility of intragenic exon duplications in the germ line of WKY rats or somatically in the proximal tubule cells. An alternative explanation for exon repetition might be that the whole gene had been duplicated, showing transcriptional read-through and splicing of a transcript encompassing both genes (18). This can be excluded in the germ line from the Southern blots and dot blots in Fig. 5. Tissue-specific recombination has been proposed as a mechanism for generating hybrid transcripts in *Drosophila* (26), but is unlikely to account for SA exon repetition because the Southern blots of PT segment DNA show no evidence of novel junctions flanking any potential duplication (Fig. 6*d*).

Other possible mechanisms for production of mRNA with exon repetitions include extensive RNA polymerase dislocation or RNA processing. The latter possibility is more attractive, given that tissue-specific splicing is widespread, but it implies a transfer of RNA from one pre-mRNA molecule to another. The only immediately obvious mechanism for achieving this is by trans-splicing, in which the 5′ portion of one molecule (exons 1–2 or 1–4) splices to the 3′ splice site of exon 2 in another molecule transcribed from the same gene. Trans-splicing is a major mechanism in a number of lower eukaryotes for introducing the 5′ exon into mRNA (27) but would be unprecedented for an endogenous mammalian gene. Trans-splicing of separate RNA molecules has been seen *in vitro* and *in vivo*, but only with partial RNA substrates or transfected genes, respectively, in which, 5′ and 3′ splice sites on individual molecules lack a cis-partner (28–34). Similar experiments were done *in vitro* with partial substrates of the rat carnitine octanoyltransferase gene (5) but it was not shown whether these were more adept than normal partial substrates at trans-splicing. It remains to be seen whether exon repetition transcripts can be produced efficiently from the canonically organized SA gene by trans-splicing. A first prerequisite is to establish whether all of the natural pre-mRNA has the expected sequence organization. If trans-splicing did occur with authentic SA pre-mRNA, the Northern blots suggest that the 5′ portion of the first molecule is constrained to react only with homologous targets. This might be a consequence of cotranscriptional splicing (32, 35), in which transcripts might be in close proximity by virtue of their common attachment to a gene; this raises questions about splicing between products of adjacent active genes and the effects of intron length. The specificity for the rat strain and cell type may be caused by differences in the concentrations of splicing factors or regulators or it may reflect differences in the loading or spacing of RNA polymerases.

We conclude that our findings provide strong evidence for exon repetition in mRNA. The biological role, if any, and the frequency of exon repetition are still unknown, but the phenomenon itself appears to reveal an unexpected capability in mammalian cells.

We thank Dr. John Armour for his help regarding the long-range PCR protocol. This work was funded by the Wellcome Trust.

1. Iwai, N. & Inagami, T. (1991) *Hypertension* **17,** 161–169.
2. Samani, N. J., Lodwick, D., Vincent, M., Dubay, C., Kaiser, M. A., Kelly, M. P., Lo, M., Harris, J., Sassard, J., Lathrop, M., *et al.* (1993) *J. Clin. Invest.* **92,** 1099–1103.
3. Iwai, N., Ohmichi, N., Hanai, K., Nakamura, Y. & Kinoshita, M. (1994) *Hypertension* **23,** 375–380.
4. Kaiser, M. A., Lodwick, D. & Samani, N. J. (1994) *Clin. Sci.* **87,** 1–4.
5. Caudevilla, C., Serra, D., Miliar, A., Codony, C., Asins, G., Bach, M. & Hegardt, F. G. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 12185–12190.
6. Samani, N. J., Brammar, W. J. & Swales, J. D. (1988) *Biochem. J.* **253,** 907–910.
7. Sanger, F., Nicklen, S. & Coulsen, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
8. Porter, D. & Curthoys N. P. (1997) *Anal. Biochem.* **247,** 279–286.
9. Harris, E. L., Dene, H. & Rapp, J. P. (1993) *Am. J. Hypertens.* **6,** 330–334.
10. Cheng, S., Fockler, C., Barnes, W. M. & Higuchi, R. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 5695–5699.
11. Kelly, M. P., Quinn, P. A., Davies, J. E. & Ng, L. L. (1997) *Circ. Res.* **80,** 853–860.
12. Hjelle, J. T., Morin, J.-P. & Trouet, A. (1981) *Kidney Int.* **20,** 71–77.
13. Romero, M. F., Fong, P., Berger, U. V., Hediger, M. A. & Boron, W. F. (1998) *Am. J. Physiol.* **274,** F425–F432.
14. Rychlik, W. & Rhoads, R. E. (1989) *Nucleic Acids Res.* **17,** 8543–8551.
15. Patel, H. R, Thiara, A. S., West, K. P., Lodwick, D. L. & Samani, N. J. (1994) *J. Hypertens.* **12,** 1347–1352.
16. Yang, T., Hassan, S. A., Singh, I., Smart, A., Brosius, F. C., Holzman, L. B., Schnermann, J. B. & Briggs, J. P. (1996) *Hypertension* **27,** 541–545.
17. Breathnach, R. & Chambon, P. (1981) *Annu. Rev. Biochem.* **50,** 349–383.
18. Magrangeas, F., Pitiot, G., Dubois, S., Bragado-Nilsson, E., Cherel, M., Jobert, S., Lebeau, B., Boisteau, O., Lethe, B., Mallet, J., *et al.(1998) J. Biol. Chem.* **273,** 16005–16010.
19. Frantz, S. A., Thiara, A. S., Lodwick, D. & Samani, N. J. (1996) *Mamm. Genome* **7,** 865–866.
20. Samani, N. J. & Lodwick, D. (1995) *J. Hum. Hypertens.* **9,** 501–503.
21. Nigro, J. M., Cho, K. R., Fearon, E. R., Kern, S. E., Ruppert, J. M., Oliner, J. D., Kinzler, K. W. & Vogelstein, B. (1991) *Cell* **64,** 607–613.
22. Cocquerelle, C., Daubersies, P., Majerus, M.-A., Kerckaert, J.-P. & Bailleul, B. (1992) *EMBO J.* **11,** 1095–1098.
23. Zaphiropoulous, P. G. (1997) *Mol. Cell. Biol.* **17,** 2985–2993.
24. Capel, B., Swain, A., Nicolis, S., Hacker, A., Walter, M., Koopman, P., Goodfellow, P. & Lovell-Badge, R. (1993) *Cell* **73,** 1019–1030.
25. Bailleul, B. (1996) *Nucleic Acids Res.* **24,** 1015–1019.
26. Ngan, V., Vacano, G., Bond, U. & James, T. (1995) *Insect Biochem. Mol. Biol.* **25,** 331–346.
27. Nilsen, T. W. (1995) *Mol. Biochem. Parasitol.* **73,** 1–6.
28. Bruzik, J. P. & Maniatis, T. (1992) *Nature (London)* **360,** 692–695.
29. Bruzik, J. P. & Maniatis, T. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 7056–7059.
30. Chiara, M. D. & Reed, R. (1995) *Nature (London)* **375,** 510–513.
31. Eul, J., Graessmann, M. & Graessmann, A. (1995) *EMBO J.* **14,** 3226–3235.
32. Misteli, T. & Spector, D. L. (1997) *Trends Cell. Biol.* **7,** 135–138.
33. Solnick, D. (1985) *Cell* **42,** 157–164.
34. Konarska, M. M. & Sharp, P. A. (1985) *Cell* **42,** 165–171.
35. Steinmetz, E. J. (1997) *Cell* **89,** 491–494.