Structure of the Angiotensin I-converting Enzyme Gene

Two alternate promoters correspond to evolutionary steps of a duplicated gene*  

(Received for publication, November 1, 1990)

Christine Hubert†, Anne-Marie Houot, Pierre Corvol, and Florent Soubrier

From the Institut National de la Santé et de la Recherche Médicale Unit 36, Collège de France, 3 Rue d'Ulm, 75005 Paris, France

Overlapping genomic clones containing the entire sequence of the human angiotensin I-converting enzyme (ACE) gene were isolated from a λ phage human DNA library. This gene spans 21 kilobases (kb) and comprises 26 exons, ranging in size from 88 to 481 base pairs. Intron-exon boundaries were sequenced and the relative positions of the exons were mapped. The two different mRNAs transcribed from the ACE gene were assigned to their respective exons. The large endogenous type ACE mRNA (4.3 kb long) is transcribed from exon 1 to exon 26, excluding exon 13. The 3-kb long testicular ACE mRNA is transcribed from exon 13 to exon 26. Exon 13 encodes for the 87 amino acids of the NH₂-terminal region of the testicular ACE, whereas downstream exons encode a sequence common to both isozymes. The gene duplication suggested by the internal homology of the endothelial ACE mRNA is now confirmed by the presence of two homologous clusters of eight exons (exons 4-11 and exons 17-24) having similar sizes and codon phases at exon-intron boundaries. The presence of two alternate promoters was investigated by ribonuclease protection assays. The different 5' ends of the two ACE transcripts revealed a promoter for the endothelial ACE mRNA in the 5'-flanking region of the first exon and a promoter for the testicular ACE mRNA situated in intron 12.

Angiotensin I-converting enzyme (ACE)† (dipeptidyl carboxypeptidase \( \mathrm{I}, \text{EC } 3.4.15.1 \)) is a peptidylpeptide hydrodase belonging to the class of zinc metalloproteases. It is, with renin, one of the two major enzymes of the renin-angiotensin system and converts the decapeptide angiotensin I into the active octapeptide angiotensin II. ACE is also able to hydrolyze the vasodilator peptide bradykinin. The wide distribution of ACE in many body tissues suggests that, in addition to its major role in the metabolism of vasoactive peptides, ACE is probably involved in the metabolism of other peptides, such as neurotransmitters (Erős and Skidgel, 1987).

The structures of the two forms of the human enzyme have been determined by molecular cloning of their respective cDNAs (Soubrier et al., 1988; Lattion et al., 1989; Ehlers et al., 1989). A large form (\( M_r = 170,000 \)) is present in endothelial cells and in the brush border of epithelial cells. A smaller form (\( M_r = 90,000 \)) is present in the germinal cells of the testis and is encoded by a shorter mRNA (3 instead of 4.3 kb). The large form of the enzyme has a 2-fold internal homology, and each of the homologous domains bears a putative active site, suggesting gene duplication. The small form contains only one of these domains and has a different amino-terminal region. Although the specific substrate of the fully active ACE in germinal cells is not known, it must be important as this peculiar form is highly conserved, being found in all the mammalian species studied to date, including humans (El Dorry et al., 1982, Bernstein et al., 1989; Velletri et al., 1985; Strittmatter et al., 1985). The expressions of the two forms are also under different hormonal regulations: the endothelial enzyme is induced by glucocorticoids, whereas the testicular form is stimulated by androgens (Krulewitz et al., 1984; Velletri et al., 1985).

Southern blot experiments indicated that a single gene encodes for both the endothelial and testicular ACE mRNAs (Soubrier et al., 1988). This prompted us to elucidate which of two possible mechanisms, differential splicing or alternate promoters, are responsible for transcription of the two mRNAs, and to determine the structural basis for their different tissue specificities and hormonal regulation.

We report here the structural organization of the human ACE gene and show that it contains two overlapping transcriptional units. The sequences of ACE cDNAs in other mammalian species have been used to make intragenic and interspecies sequence comparisons and thus provide insight into the evolutionary features of the ACE gene.

MATERIALS AND METHODS

Screening of the Human Genomic DNA Library—A human genomic library, prepared by partial AliI-HaeIII digests of human fetal liver DNA and cloned in λ phage charon 4A with EcoRI linkers was kindly provided by Drs. Lawn and Maniatis (Lawn et al., 1978). Two million clones were plated on Escherichia coli strain LE392 and screened by plaque hybridization (Benton and Davis, 1977). Filters were hybridized in duplicate with inserts of phage clones A1922 and A2111 containing nearly the whole endothelial ACE cDNA sequence (Soubrier et al., 1988). These DNA inserts were labeled with \( [\gamma-\!^{32}\!] \) PCTP by the random primer labeling method (Feinberg and Vogelstein, 1983) using a commercial kit (Amersham Corp.). Hybridization was carried out at 42 °C in 50% formamide, 5 × SSPE (1 × SSPE is 150 mM NaCl, 10 mM NaH₂PO₄, 1 mM EDTA, pH 7.4), 5 × Denhardt's solution, 0.1% sodium dodecyl sulfate, and 190 µg/ml denatured salmon sperm DNA. Filters were washed at 65 °C in 0.1 SSC (1 × SSC is 150 mM NaCl, 15 mM sodium citrate, pH 7), 0.1% sodium dodecyl sulfate and autoradiographed at ~80 °C with an intensifying screen. Positive clones were isolated by repeated plaque purification. Phage DNA was purified by the standard method of cesium density gradient centrifugation (Maniatis et al., 1982).
weight DNA was isolated from human placenta (Maniatis et al., 1982). DNA samples (14 μg) were incubated with restriction endonucleases (New England Biolabs). Digested DNAs were electrophoresed through a 0.7% agarose gel in 40 mM Tris, 20 mM sodium acetate, 2 mM EDTA, pH 7.8, transferred to nylon filter and hybridized with α-32P-labeled Egl-I-EcoRI restriction fragment of ACE cDNA, corresponding to positions 3778-4024 of the endothelial ACE cDNA. The filter was hybridized and washed as described in the screening section.

Mapping the 5' Ends of the Endothelial and Testicular Transcripts by Ribonuclease Protection Assay—Total RNAs from human endothelial cells and human testis were prepared by the method of Chirgwin et al. (1979). Poly(A)+ RNAs were purified on oligo(dT) cellulose. Two restriction fragments were subcloned into plasmid vector Bluescript in order to synthesize both sense and antisense transcripts corresponding to the putative 5' ends of the endothelial and testicular ACE transcripts. A 617-bp Apal-PouII fragment bearing 225 bp of the 5' end of the endothelial ACE cDNA sequence and the upstream genomic sequence was subcloned into BlueScript SK (pEH18-16). A 244-bp BanII-BanII fragment corresponding to 219 bp of the 5' extremity of the testis ACE cDNA sequence and to 25 bp of the upstream genomic sequence was treated with T4 DNA polymerase and subcloned into BlueScript SK (pTH35-1). The conditions used for the following experiments were as in Zinn et al. (1983) unless otherwise stated. In vitro transcriptions were performed in the presence of [α-32P]UTP to synthesize labeled antisense RNA. Approximately 5 × 10⁶ cpm of the probe was coprecipitated with RNAs. Hybridization was performed at 45 and 35 °C for pTH35-1 and pEH18-16-derived RNAs, respectively, in a hybridization mix for 24 h, and was followed by digestion with ribonucleases T1 and A. After phenol-chloroform extraction and ethanol precipitation, RNAs were applied to a 5% acrylamide, 7 M urea sequencing gel. Sequencing reactions were used as size markers together with the undigested labeled antisense RNA.

Sequence Alignment—The 1484-bp upstream from the tsp of the endothelial ACE mRNA and intron 12, which corresponds to the DNA region upstream from the tsp of the ACE testicular transcript, were optimally aligned with the program NUCALN (Wilbur and Lipman, 1983). The same program was used to compare the sequences of the human, mouse, and rabbit ACE genes.

RESULTS

Isolation and General Organization of the Human ACE Gene—Out of 2 million phage clones screened with the human endothelial ACE cDNA probe, 18 phages were selected. The isolated genomic clones were segregated into four overlapping λ phages (λ1011, λ1915, λ311, and λ621) spanning a 35 kb region of continuous human genomic DNA (Fig. 1). No other

FIG. 1. Organization of the human ACE gene and restriction map of the genomic clones. Top line, scale in kilobases. Second line, location of the 26 numbered exons (vertical bars). Exon 13 (open bar) is specific to testicular ACE. The cleavage sites of seven restriction endonucleases are indicated. Those marked with an asterisk are also present in cDNAs. The position of four overlapping λ phage inserts (hatched bars) and the plasmid subclones (lines) are indicated at the bottom of the figure. Plasmid numberings correspond to their respective lengths.

FIG. 2. Southern blot analysis of the human ACE gene. Human genomic DNA was digested with the restriction endonucleases indicated on the top. The filter was hybridized with a human ACE cDNA fragment corresponding to positions 3778 to 4024 (see "Materials and Methods"). HindIII fragments of λ phage DNA were used as size markers and are indicated in kilobase pairs.

Restriction Endonuclease Mapping of Genomic DNA Clones—Phage DNAs were digested with one or more restriction endonucleases and the resulting fragments were separated by agarose gel (1%) electrophoresis, denatured in alkali, and transferred to nylon membrane (Hybond N, Amersham Corp.) as described by Southern (1975). Filters were hybridized with selected oligonucleotides, designed from the ACE cDNA sequence. The filter was hybridized and washed as described in the screening section.

Subcloning and DNA Sequencing—Selected restriction fragments of genomic DNA inserts were subcloned into the plasmid vectors Bluescript (Stratagene), pTZ19R, and pTZ18R (Pharmacia LKB Biotechnology Inc.). Detailed restriction maps of each subclone were prepared by restriction enzyme digestions and Southern blotting. DNA sequencing was performed by the dideoxy chain termination method (Sanger et al., 1977) with T7 (Sequenase, United States Biochemical) or E. coli DNA polymerase (Klenow fragment, Boehringer Mannheim) and a set of synthetic oligonucleotide primers designed from the ACE cDNA sequence.

Southern Blot Analysis of Human Genomic DNA—High molecular weight DNA was isolated from human placenta (Maniatis et al., 1982). DNA samples (14 μg) were incubated with restriction endonucleases (New England Biolabs). Digested DNAs were electrophoresed through a 0.7% agarose gel in 40 mM Tris, 20 mM sodium acetate, 2 mM EDTA, pH 7.8, transferred to nylon filter and hybridized with α-32P-labeled Egl-I-EcoRI restriction fragment of ACE cDNA, corresponding to positions 3778-4024 of the endothelial ACE cDNA. The filter was hybridized and washed as described in the screening section.

Mapping the 5' Ends of the Endothelial and Testicular Transcripts by Ribonuclease Protection Assay—Total RNAs from human endothelial cells and human testis were prepared by the method of Chirgwin et al. (1979). Poly(A)+ RNAs were purified on oligo(dT) cellulose. Two restriction fragments were subcloned into plasmid vector Bluescript in order to synthesize both sense and antisense transcripts corresponding to the putative 5' ends of the endothelial and testicular ACE transcripts. A 617-bp Apal-PouII fragment bearing 225 bp of the 5' end of the endothelial ACE cDNA sequence and the upstream genomic sequence was subcloned into BlueScript SK (pEH18-16). A 244-bp BanII-BanII fragment corresponding to 219 bp of the 5' extremity of the testis ACE cDNA sequence and to 25 bp of the upstream genomic sequence was treated with T4 DNA polymerase and subcloned into BlueScript SK (pTH35-1). The conditions used for the following experiments were as in Zinn et al. (1983) unless otherwise stated. In vitro transcriptions were performed in the presence of [α-32P]UTP to synthesize labeled antisense RNA. Approximately 5 × 10⁶ cpm of the probe was coprecipitated with RNAs. Hybridization was performed at 45 and 35 °C for pTH35-1 and pEH18-16-derived RNAs, respectively, in a hybridization mix for 24 h, and was followed by digestion with ribonucleases T1 and A. After phenol-chloroform extraction and ethanol precipitation, RNAs were applied to a 5% acrylamide, 7 M urea sequencing gel. Sequencing reactions were used as size markers together with the undigested labeled antisense RNA.

Sequence Alignment—The 1484-bp upstream from the tsp of the endothelial ACE mRNA and intron 12, which corresponds to the DNA region upstream from the tsp of the ACE testicular transcript, were optimally aligned with the program NUCALN (Wilbur and Lipman, 1983). The same program was used to compare the sequences of the human, mouse, and rabbit ACE genes.

RESULTS

Isolation and General Organization of the Human ACE Gene—Out of 2 million phage clones screened with the human endothelial ACE cDNA probe, 18 phages were selected. The isolated genomic clones were segregated into four overlapping λ phages (λ1011, λ1915, λ311, and λ621) spanning a 35 kb region of continuous human genomic DNA (Fig. 1). No other
Human Angiotensin I-converting Enzyme Gene

Nucleotide sequence of the intron-exon junctions in the human ACE gene

Exon sequences are in uppercase letters; intron sequences are in lowercase. The positions of cDNA exons are relative to the start site of the endothelial ACE mRNA, except for exon 13 which is positioned relative to the testicular ACE mRNA start site. The sizes of the introns were estimated from restriction endonuclease analysis or from nucleotide sequences. The codon phase refers to the position of intron in the codon triplet. O intron occurs between codons, I intron occurs after the first nucleotide, and II intron occurs after the second nucleotide in the codon.

<table>
<thead>
<tr>
<th>Exon no. (size)</th>
<th>cDNA position of endothelial exon</th>
<th>5' splice donor</th>
<th>Intron no. (size)</th>
<th>3' splice acceptor</th>
<th>Codon phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>bp</td>
<td></td>
<td></td>
<td>bp</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (271)</td>
<td>1-271</td>
<td>CCAAGgtgggc</td>
<td>1 (-600)</td>
<td>tgccaatag6AGG</td>
<td>O</td>
</tr>
<tr>
<td>2 (168)</td>
<td>272-439</td>
<td>CCAAGgtgggc</td>
<td>2 (-800)</td>
<td>gactcccccgTCA</td>
<td>O</td>
</tr>
<tr>
<td>3 (94)</td>
<td>440-635</td>
<td>CCAAGgtgggc</td>
<td>3 (-700)</td>
<td>tgtttctcggATCT</td>
<td>I</td>
</tr>
<tr>
<td>4 (144)</td>
<td>534-677</td>
<td>GAGGgtgggc</td>
<td>4 (-450)</td>
<td>agcccccacgGCTT</td>
<td>I</td>
</tr>
<tr>
<td>5 (192)</td>
<td>678-869</td>
<td>CTTGgttaag</td>
<td>5 (-600)</td>
<td>ccttactcagGAGA</td>
<td>I</td>
</tr>
<tr>
<td>6 (98)</td>
<td>870-967</td>
<td>GCAAGgttaag</td>
<td>6 (-300)</td>
<td>cctctccacggGCT</td>
<td>O</td>
</tr>
<tr>
<td>7 (173)</td>
<td>968-1140</td>
<td>TCAGgttcaag</td>
<td>7 (-700)</td>
<td>taggccccacgGATC</td>
<td>II</td>
</tr>
<tr>
<td>8 (224)</td>
<td>1141-1364</td>
<td>AGGggttaag</td>
<td>8 (-300)</td>
<td>cgccacagAGG</td>
<td>I</td>
</tr>
<tr>
<td>9 (145)</td>
<td>1365-1509</td>
<td>TCTGgttaag</td>
<td>9 (-300)</td>
<td>ccttctcagAGAAC</td>
<td>II</td>
</tr>
<tr>
<td>10 (99)</td>
<td>1510-1608</td>
<td>TCAAGgttata</td>
<td>10 (-300)</td>
<td>tgctgctcagGTAC</td>
<td>II</td>
</tr>
<tr>
<td>11 (123)</td>
<td>1609-1731</td>
<td>TCTGgttata</td>
<td>11 (-350)</td>
<td>gcacccacagGAGAG</td>
<td>II</td>
</tr>
<tr>
<td>12 (212)</td>
<td>1732-1943</td>
<td>ATAGgttaag</td>
<td>12 (290)</td>
<td>ggggccagGACT</td>
<td>I</td>
</tr>
<tr>
<td>13 (228)</td>
<td>1828-2088</td>
<td>CCAAGgttgaag</td>
<td>13 (170)</td>
<td>tgtccacagACCT</td>
<td>I</td>
</tr>
<tr>
<td>14 (157)</td>
<td>1944-2088</td>
<td>TCTGgttgaag</td>
<td>14 (-150)</td>
<td>tgtccctcagGTC</td>
<td>O</td>
</tr>
<tr>
<td>15 (159)</td>
<td>2081-2239</td>
<td>GAGGgtttcgct</td>
<td>15 (-250)</td>
<td>ccttctcagATCA</td>
<td>O</td>
</tr>
<tr>
<td>16 (88)</td>
<td>2240-2327</td>
<td>CCAAGgttgaag</td>
<td>16 (-1850)</td>
<td>cccctccagATCT</td>
<td>I</td>
</tr>
<tr>
<td>17 (144)</td>
<td>2328-2471</td>
<td>ATAGgttgaag</td>
<td>17 (150)</td>
<td>ccttctcagGCTA</td>
<td>I</td>
</tr>
<tr>
<td>18 (192)</td>
<td>2472-2663</td>
<td>TCAGgttcaag</td>
<td>18 (-1800)</td>
<td>tgcgctcagGAGAA</td>
<td>II</td>
</tr>
<tr>
<td>19 (100)</td>
<td>2584-2763</td>
<td>CAGGgtttcgct</td>
<td>19 (150)</td>
<td>tgcgctcagGTC</td>
<td>II</td>
</tr>
<tr>
<td>20 (171)</td>
<td>2764-2954</td>
<td>TCTGgttcat</td>
<td>20 (-2000)</td>
<td>ctattcagGATC</td>
<td>II</td>
</tr>
<tr>
<td>21 (224)</td>
<td>2935-3158</td>
<td>GACGgttcaag</td>
<td>21 (-250)</td>
<td>tgacGAGCA</td>
<td>I</td>
</tr>
<tr>
<td>22 (145)</td>
<td>3159-3363</td>
<td>TCAGgttctcg</td>
<td>22 (310)</td>
<td>cactgacagGCTG</td>
<td>II</td>
</tr>
<tr>
<td>23 (99)</td>
<td>3304-3402</td>
<td>TCAGgttcaag</td>
<td>23 (-1900)</td>
<td>tgcgctcagGTAC</td>
<td>II</td>
</tr>
<tr>
<td>24 (122)</td>
<td>3403-3524</td>
<td>CCCGgttcaag</td>
<td>24 (-250)</td>
<td>gcccacacagGAC</td>
<td>I</td>
</tr>
<tr>
<td>25 (189)</td>
<td>3525-3713</td>
<td>TCTGgttcaag</td>
<td>25 (150)</td>
<td>tccacacagCTCG</td>
<td>I</td>
</tr>
</tbody>
</table>

TABLE II
Similarity between exons from the first and the second domains of the ACE gene

Codon phases are explained in Table I.

<table>
<thead>
<tr>
<th>Exon no.</th>
<th>Sizes</th>
<th>Nucleotide sequence similarity</th>
<th>Codon phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>bp</td>
<td>%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 and 17</td>
<td>144/144</td>
<td>64.8</td>
<td>1 Gly/I Gly</td>
</tr>
<tr>
<td>5 and 18</td>
<td>192/192</td>
<td>70.3</td>
<td>1 Gly/I Gly</td>
</tr>
<tr>
<td>6 and 19</td>
<td>98/100</td>
<td>67.3</td>
<td>O Gln/I Gly</td>
</tr>
<tr>
<td>7 and 20</td>
<td>173/171</td>
<td>80.1</td>
<td>II Arg/II Arg</td>
</tr>
<tr>
<td>8 and 21</td>
<td>224/224</td>
<td>68.3</td>
<td>II Arg/I Glu</td>
</tr>
<tr>
<td>9 and 22</td>
<td>145/145</td>
<td>64.8</td>
<td>II Arg/I Arg</td>
</tr>
<tr>
<td>10 and 23</td>
<td>99/99</td>
<td>60.6</td>
<td>II Arg/I Arg</td>
</tr>
<tr>
<td>11 and 24</td>
<td>123/123</td>
<td>73.7</td>
<td>II Arg/I Ala</td>
</tr>
</tbody>
</table>

Type of hybridizing phage was isolated from this library. Several fragments were detected by Southern blot hybridization of human genomic DNA with the complete ACE cDNA as probe and were all identified in isolated phases. In contrast, only a single DNA band was hybridized when a short restriction fragment of the cDNA, located within exon 26, was used as probe (Fig. 2). These results confirm that a single gene encodes for ACE in the human genome.

A set of oligonucleotides designed from the coding sequence of the endothelial and testicular ACE cDNAs was used to determine the gene structure of ACE by hybridization with cloned genomic sequences. Intron-exon boundaries were sequenced (Table I). The DNA was sequenced using ACE cDNA-specific oligonucleotides to prime chain elongation on subclones containing the genomic fragments. Intron sizes were determined on genomic subclones by restriction endonuclease mapping or DNA sequencing (Table I). Some intron sizes were verified on isolated phages by DNA amplification using primers of the flanking coding region. The human ACE gene contains 26 exons interrupted by 25 introns and spans approximately 21 kb of DNA (Fig. 1). All the exon-intron junction sequences follow the GT/AG rule of Breathnach and Chambon (1981). The 26 exons range in size between 88 bp (exon 16) and 481 bp (exon 26). The 25 intervening sequences vary considerably in length, ranging from 150 bp (introns 17 and 25) to 2 kb (intron 20) (Table I). The polyadenylation site has been determined for the testicular ACE mRNA (Lettion et al., 1989; Ehlers et al., 1989). The sequence surrounding the stop codon in the genomic fragment is identical to the published sequences until the polyadenylation signal which is localized 238 bp downstream from the stop codon. For the endothelial ACE mRNA the polyadenylation site was not reported, other site(s), downstream from the testicular polyadenylation site, cannot be ruled out.

Endothelial ACE Results from Transcription of a Duplicated Cluster of Exons—Previous molecular cloning of the endothelial ACE cDNA revealed a high degree of internal homology between two large domains representing the major part of the cDNA. In the gene, each homologous domain is encoded by a group of eight exons: exons 4–11 for the first domain, and exons 17–24 for the second domain. The symmetrical exons in each group have very similar sequences and sizes, and their codon phases are conserved (Table II). Exons 8 and exon 21, which contain the consensus sequence of metalloproteinase active site (H-E-X-X-H) are 68% similar, while exon 7 and
part A,

15380 Human Angiotensin I-converting Enzyme Gene bases downstream from the probe. The restriction enzyme sites of the ACE gene used to isolate sense probe.

and sense probe, respectively. The calculated size of the protected

used as RNA size marker, is obtained by adding the length of the T3

sense testis probe. The total length of the antisense testicular probe (325 bases),

with 15 pg of cultured endothelial cell poly(A)' RNA, 15 pg of tRNA,

fragment (225 b) is indicated on the right. Lanes 3-5,

undigested endothelial promoter antisense RNA probe, Lane 1, undigested testicular promoter antisense RNA probe. Lane 2, undigested endothelial promoter antisense probe. Lanes 3-5, RNase-protected antisense probes hybridized with 15 µg of cultured endothelial cell poly(A)+ RNA, 15 µg of tRNA, and sense probe, respectively. The calculated size of the protected fragment (225 b) is indicated on the right. Part A, mapping of the endothelial ACE RNA. On the left, sequencing reactions used as size markers. Lane 1, undigested testicular promoter antisense RNA probe. Lane 2, undigested endothelial promoter antisense probe. Lanes 3-5, RNase-protected antisense probe hybridized with 15 µg of cultured endothelial cell poly(A)+ RNA, 15 µg of tRNA, and sense probe, respectively. The calculated size of the protected fragment (225 b) is indicated on the right. Part B, on the left, sequencing reactions used as size markers. Lane 1, undigested antisense testis probe. Lanes 2 and 3, RNase-protected antisense probe hybridized with 10 µg of testis RNA and with sense RNA probe, exons 20 are 80% similar. In contrast to the exons, the introns in the two halves of the gene differ in size. Four introns in the 3' part of the gene (introns 16, 18, 20, 23) are much larger than the homologous introns in the 5' part of the gene (introns 3, 5, 7, 10).

Endothelial ACE Promoter—The transcription start point (tsp) of the endothelial ACE mRNA was determined using the ribonuclease protection assay (see “Materials and Methods”). The main tsp observed was a purine nucleotide (G) 22 bases upstream from the ATG start codon; this tsp corresponds to the 5' end of endothelial ACE mRNA that was previously determined by cDNA cloning using a specifically primed cDNA library (Fig. 3A) (Soubrier et al., 1988). The DNA sequence of a 1484-bp region upstream of the tsp was determined, and a search for core promoter elements was carried out (Fig. 4). A typical TATA box (TATAAA) was found 37 bases upstream from the tsp, but no consensus CCAAT element was found. Three Sp1 potential binding sites are present and two of them are in tandem (positions −62, −68, −131) within a GC-rich region, reaching 75% on the 200 nucleotides upstream the start site (Dyrrna and Tjian, 1985).

We also searched for hormone-responsive elements, based on the known in vitro regulation of ACE expression in endothelial cells (Krulwit et al., 1984). The hexanucleotide of the consensus glucocorticoid-responsive element was found at position −1213 (TGTCTC), −935 (TGTCTC), −794 (TGTCTC), and −866 (TGTCTC) (Karin et al., 1984).

The Testicular Transcript Is Derived from the 3’ Moiety of the ACE Gene—The 5' end sequence of the testicular mRNA, absent from the endothelial ACE mRNA, contains a non-translated DNA sequence and encodes for 67 amino acids, including a putative signal peptide. The entire testicular-specific DNA sequence is encoded by exon 13 (228 bp long), which is in the middle part of the gene, and is surrounded by exons which are present in the endothelial transcript (Fig. 1).

All the exons following exon 13 (exons 14–26) are present in both the testis mRNA and the endothelial transcript, including the last exon, which encodes the membrane-anchoring segment. We have studied the transcription start site of the testicular mRNA using the ribonuclease protection assay. Human testis poly(A)+ RNA was hybridized with a labeled anti-sense RNA transcribed from a restriction fragment overlapping the 3' part of intron 12 and 5' part of exon 13. A major transcription start point, at 28 nucleotides upstream from the translation initiation site, was found which corresponds to the 5' end determined on the cDNA clone obtained by screening a specifically primed cDNA library (Fig. 3B).

Sequence Analysis of Intron 12 as an ACE Testicular Promoter—The ribonuclease protection assay showed that the 5' end of testis ACE mRNA corresponds to the 5' end of exon 13, supporting the hypothesis that a promoter is located upstream, within intron 12. The complete sequence of intron 12 is shown in Fig. 4. We detected a TATA-like sequence (TATTT) 15 bases upstream from the major testis mRNA tsp. A sequence very similar to the consensus sequence of the cAMP-responsive element (TGGAGGTCGA) is present at position −44 from the tsp (Roessler et al., 1988). A sequence homologous to the steroid-responsive element was found at position −190 (TGTCTC).

A computer search for homology between the endothelial promoter and intron 12 was carried out with the program NUCALN (Wilbur and Lipman, 1983). Two regions of the endothelial promoter had sequences similar to intron 12. A region between positions −1238 and −1138 was 54% homologous respectively. Sizes of the undigested probe (325 b) and of the protected fragment (219 b) are indicated on the right.
ogous to the sequence between positions -216 and -107 of intron 12 and another region between positions -489 and -367 of the endothelial promoter and between positions +1 and +81 of intron 12 was 46% homologous. Using the same alignment a short 20-nucleotide region, separated from the promoters (positions -874 of the endothelial promoter and two other regions by large gaps, was 80% similar in the two genes.

**Interspecies and Intragenic Nucleotide Sequence Comparisons**—We calculated the approximate rate of divergence per million years between human, mouse, and rabbit at the ACE gene locus by comparing the nucleotide sequences of exons 7 and exons 20, which are the highest homologous regions between the two domains of the ACE gene (Bernstein et al., 1989; Kumar et al., 1989). The divergence rate for exon 20 was 13.9% between human and mouse, 8.2% between human and rabbit, and 13.9% between mouse and rabbit. The divergence rate for exon 7 between human and mouse was 15.9%, showing that there is no significant difference between the divergence rates of the two parts of the gene. This calculation could not be done for the rabbit as the endothelial form of rabbit cDNA has not been described. Assuming that the mammalian radiation occurred 80 million years ago, a 1% divergence rate at the exon 20 ACE gene locus is reached in 80/13.9 = 5.9 million years from the human-rabbit comparison, in 80/13.9 = 5.7 million years from mouse/rabbit comparison, and in 80/13.5 = 5.9 million years from human-mouse comparison. Intragenic sequence comparisons show that exon 7 presents a 18.9% divergence rate from its homologous exon 20 in the human ACE gene. The same comparison shows a 24.6% divergence rate in the mouse ACE gene. It is assumed from these calculations that ACE gene duplication occurred between 118 and 194 million years ago, from human/mouse and human/rabbit comparisons, respectively.

**DISCUSSION**

The structural organization of the entire gene coding for the two ACE isozymes was established. Out of 26 exons, the endothelial ACE is encoded by 25 exons and the testicular ACE is constituted by 14 exons whose 13 are common with the endothelial enzyme. Molecular cloning of ACE endothelial and subsequently of testicular cDNA showed that the structure of the gene resulted from gene duplication (Soubrier et al., 1988; Lattion et al., 1989). This result is strengthened and extended to the structure of the ACE gene described in this paper. Not only are the sequences of the two domains constituting the major part of the molecule similar, but the homologous exons encoding the two domains have similar sizes and similar codon phases at the exon-intron junctions.

The gene duplication appears to have occurred very early in evolution. Two types of observations support this hypothesis. First, the molecular weight of the circulating enzyme or the molecular structure of the ACE mRNA indicates that the ACE gene is probably duplicated in pig, mouse, rat, and rabbit (Hooper et al., 1987; Velletri et al., 1985; El Dorry et al., 1982). Second, interspecies and intragenic nucleotide sequence comparisons (see "Results") suggest that ACE gene duplication occurred before the mammalian radiation. A more precise dating of the duplication will be possible once the structures of the ACE genes in non-mammalian vertebrates have been determined.
determined. Enzymes similar to ACE and inhibited by the ACE inhibitor, captopril, have been found in both reptiles and birds (Fregly et al., 1981; Gervitz et al., 1987).

Such duplicated structures occur in other brush-border anchored enzymes, including sucrase-isomaltase and lactase-phlorizin hydrolase (Hunziker et al., 1986; Mantei et al., 1988).

In the latter, 4-fold internal homology suggests two cycles of duplication. Each of the two domains in these two enzymes bears an active site having a separate catalytic activity (Schlegel-Haueter et al., 1972; Sjöström et al., 1980). The activities of each domain of ACE were studied by in vitro expression in our laboratory (Wei et al., 1991). Both domains possess enzymatic activity, but they have different catalytic properties. Identification of residues responsible for these differences will be of great value for understanding the catalytic mechanisms of ACE. The amino-terminal catalytic site may well have another substrate specificity, as it does in lactase-phlorizin hydrolase and sucrase-isomaltase.

No overall sequence similarity was found between ACE and other metallopeptidases, such as neutral endopeptidase and collagenase. Only short amino acid sequences were found to fit the consensus sequences of the catalytic site of zinc metallopeptidases (Soubrier et al., 1988). This sequence H-E-X-X-H, in which the 2 histidines are believed to bind the zinc atom, occurs twice in the ACE molecule. Shuffling of an exon bearing this consensus sequence in metallopeptidase genes is not supported by the structure of ACE, collagenase, and neutral endopeptidase genes, as there is no sequence similarity between exons concerned (Collier et al., 1988; D'Adamio et al., 1989).

A remarkable feature of the ACE gene is the existence of two mRNA species, one corresponding to the duplicated gene and present in several tissues, and the other, corresponding to the '3' half of the gene, expressed solely in the testis. Besides their different tissue specificity, expression of the two ACE transcripts is regulated by different developmental and hormonal controls. The level of expression is also very different; the testicular ACE mRNA species is much more abundant in the testis than in the larger form in other tissues, assuming that the mRNAs have similar stabilities (Soubrier et al., 1988). DNA elements responsive to glucocorticoids, which induce ACE secretion by cultured endothelial cells (Krulwitz et al., 1984), were found in the upstream region of the endothelial promoter. In vivo studies in the rat have shown that ACE expression in the testis is positively regulated by androgens and luteinizing hormone (Vellutri et al., 1985). The testicular transcript could be formed by alternate splicing of a large ACE primary transcript or by transcription from an internal promoter. Mapping of the 5' end of the testicular ACE mRNA by RNase protection showed that it starts at the 3' edge of intron 12, indicating an internal promoter. We also found a TATA-like sequence and short DNA sequences very similar to cAMP and consensus steroid hormone-responsive element in intron 12. We therefore propose that the testicular transcript is transcribed under the control of an internal promoter and suggest that intron 12 is probably responsible for tissue specificities of expression and hormonal regulation, although we cannot rule out a role of other regions of the gene. Similarly, Howard et al. (1990) recently showed that the mouse ACE testicular transcript originates from an internal promoter.

We identified two similar (50% identity) sequences, separated by large gaps, in the two putative promoters of the ACE gene. Their presence could indicate that they are derived from a common ancestral sequence and that they diverged toward different regulatory and tissue specificities after the duplication.

Both the structure of the testis ACE mRNA, transcribed from the 3' half of the gene and the location of its promoter in the central region of the gene support the hypothesis that this mRNA arises from the ancestral preduplication form of the gene. Alternative promoters have been described in many other genes (for review see Schibler and Sierra, 1987). In these cases they ensure different developmental regulation and tissue-specific expression, but in none of them does transcription from the alternative promoter appear to produce an "ancestral" mRNA, anterior to duplication of the whole functional gene. Interestingly, rat farnesyl pyrophosphate synthetase mRNA also appears to have a testis-specific alternative promoter which produces abundant transcription of an mRNA with a different 5' end (Teruya et al., 1990).

Transfection experiments, with either the endothelial ACE gene promoter or intron 12 driving transcription of a reporter gene, will be necessary to study the tissue specificities and hormonal regulation of the two alternative promoters.

Acknowledgments—We thank J. Allegrini, A. L. Lattion, T. Sardo, and P. Testut for their help during achievement of this work, N. Braun for typing the manuscript, A. Depardieu for artwork, and J-L. Sikorav for critical reading of the manuscript.

REFERENCES


Human Angiotensin I-converting Enzyme Gene