Structure of the bovine ET$_B$ endothelin receptor gene

Takeshi MIZUNO,* Yasuyoshi SAITO,* Makoto ITAKURA,* Futoshi ITO,* Teizo ITO,* Etsuko N. MORIYAMA,† Hiromi HAGIWARA* and Shigehisa HIROSE†

*Department of Biological Sciences, Tokyo Institute of Technology, Ookayama, Meguroku, Tokyo 152, and †National Institute of Genetics, Mishima, Shizuoka 411, Japan

The structure of the gene encoding the bovine type B endothelin receptor (ET$_B$) has been established and compared with those of other heptahelical receptors. The gene is present as a single copy in the bovine genome, as demonstrated by Southern blot analysis, and spans at least 36 kb. The coding region is divided into 7 exons separated by 6 introns, one of which is more than 23 kb in length. The exons correspond well to the structural domains of the receptor: the first exon encodes the first and second transmembrane domains, and each of the following transmembrane domains is encoded by a separate exon. The portion of the ET$_B$ protein sequence encoded by exon 3 is quite different from the corresponding ET$_A$ sequence, suggesting that this region is responsible for the distinct ligand specificities of the two receptor subtypes. The second intron interrupts the canonical Asp-Arg-Tyr sequence, which is located at the end of the third transmembrane domain of the heptahelical receptors, as with the substance P, substance K, dopamine D$_2$ and dopamine D$_3$ receptor genes. To map the 5' region of the gene and determine the start of transcription, primer-extended cDNAs were cloned and sequenced: multiple start sites were deduced with no apparent TATA box in the expected upstream region. Similar results were obtained by ribonuclease protection analysis.

INTRODUCTION

The presence of an endothelial-cell-derived vasoconstrictor peptide was first demonstrated by Hickey et al. [1] and later confirmed by O'Brien et al. [2]. Soon after the isolation of the substance (endothelin) by Yanagisawa et al. [3], it became apparent that the endothelin family has three members: endothelin-1, endothelin-2 and endothelin-3. Two types of receptors that mediate the smooth-muscle-contracting activity of endothelins have been identified and their primary structures determined by cDNA cloning and sequencing [4,5]. The receptor subtypes, ET$_A$ and ET$_B$, have characteristic ligand specificities and patterns of tissue expression. For example, the ET$_A$ endothelin receptor which binds endothelin-1 and endothelin-2 much more strongly than endothelin-3, predominates in arterial smooth muscle cells [6] and testicular myoid cells [7], whereas the ET$_B$ receptor, which recognizes the three ligands equally, appears to be the major form in endothelial cells [8,9] and in the ureter [10]. Elucidation of the structural basis for this differential expression requires a detailed analysis of their genes.

Our interest in the endothelin receptor genes also stems from an evolutionary point of view. Both ET$_A$ and ET$_B$ receptors belong to the G-protein-coupled receptor family and have seven membrane-spanning domains. The members of this heptahelical receptor family can be divided into two groups in terms of their gene structure: one has introns, whereas the other is intronless. To elucidate which type of gene encodes the endothelin receptor, we carried out a structural and comparative analysis of the bovine ET$_B$ receptor gene.

MATERIALS AND METHODS

Materials

Fresh bovine lung and adrenal gland were obtained from the Shibaura Abattoir Sanitation Inspection Station, Tokyo, Japan. Restriction enzymes and T4 polynucleotide kinase were from Toyobo, Osaka, Japan; ribonuclease A was from Boehringer Mannheim, Mannheim, Germany; DNA ligation kit and random primer DNA labelling kit were from Takara, Kyoto, Japan; You-Prime cDNA synthesis kit, mRNA purification kit and double-stranded Nested Deletion kit were from Pharmacia, Uppsala, Sweden; λ ZAP II, Gigapack Gold in vitro packaging kit, pBluescript II and RNA transcription kit were from Stratagene, San Diego, CA, U.S.A.; Sequenase version 2.0 sequencing kit was from United States Biochemical Corp., Cleveland, OH, U.S.A.; a bovine genomic DNA library was from Clontech, Palo Alto, CA, U.S.A.; nitrocellulose filters were from Schleicher & Schuell, Dassel, Germany; Gene Screen Plus was from Du Pont–New England Nuclear, Boston, MA, U.S.A.; GeneAmp DNA amplification reagent kit was from Perkin-Elmer Cetus, Norwalk, CT, U.S.A.; ribonuclease T1 and Superscript MMLV reverse transcriptase were from Life Technologies, Gaithersburg, MD, U.S.A.; 35P-labelled nucleotides were from Amersham International, Amersham, Bucks., U.K. Oligonucleotides were synthesized with a Milligen/Biosearch Cyclone Plus DNA synthesizer. DNA sequences were analysed using the computer software GENETYX purchased from Software Development Co., Tokyo, Japan.

Isolation of bovine ET$_B$ receptor cDNA clones

To obtain cDNA clones of the bovine ET$_B$ receptor covering the 5' end of the mRNA, a cDNA library was constructed with a specific primer. Poly(A)$^+$ RNA was prepared from bovine lung as described previously [11] and reverse-transcribed with the synthetic oligonucleotide 5'--GTCCTCCGGAAGTGGCTAGGT--3', corresponding to the 5' region of the partial cDNA clone bETR5 [11], using a Pharmacia You-Prime cDNA synthesis kit according to the directions provided. The double-stranded, blunt-ended and adaptor-ligated cDNA obtained was cloned into λ ZAP II. Approx. 6 × 10$^4$ phage clones were screened with

Abbreviations used: ET$_A$, type A endothelin receptor; ET$_B$, isopeptide-nonselective type B endothelin receptor.

† To whom correspondence and reprint requests should be sent.

The nucleotide sequence data reported in this paper will appear in the DDBJ, EMBL and GenBank Nucleotide Sequence Databases under accession nos. D10989–D10994.
a random-primed labelled 127 bp BamH1/XbaI restriction fragment of the previously cloned bETR5 [11] by the plaque hybridization method [12]. Positive clones were characterized and sequenced [12].

Isolation and characterization of bovine ET\textsubscript{B} receptor gene

The bovine ET\textsubscript{B} receptor gene was isolated from a bovine genomic library constructed in EMBL-3. Approx. 3 × 10\textsuperscript{6} phage clones were screened with various random-primed labelled restriction fragments of bovine ET\textsubscript{B} receptor cDNA (bETR5) [11]. Positive clones were purified, characterized by restriction mapping, subcloned into the pBluescript II vector and sequenced by the dideoxy chain-termination method [12] using Sequenase version 2.0. All exons and intron/exon junctions with adjacent intronic sequences were subjected to sequence analyses. The nucleotide sequence of 1000 bp of 5'-flanking DNA was also determined.

Genomic Southern blot analysis

Bovine genomic DNA (20 μg) prepared from cultured endothelial cells was digested with 200 units of EcoRI and HindIII, electrophoresed and blotted on to Gene Screen Plus by the standard method [12]. Hybridization was performed with a random-primed labelled 500 bp EcoRI/EcoRI fragment of \(\lambda\)BerG1 as described previously [13].

Primer extension

A synthetic oligonucleotide corresponding to nucleotides -210 to -191 and -128 to -112 in the 5'-flanking region of the gene was labelled with \(\gamma\text{-}^{32}\text{P}\)ATP, and 2 × 10\textsuperscript{6} c.p.m. was annealed to 2 μg of bovine adrenal poly(A)\textsuperscript{+} RNA for 16 h at 30°C in 30 μl of hybridization buffer containing 40 mM-NaCl, 40 mM-Pipes, pH 6.4, 1 mM-EDTA and 80% (v/v) formamide. The annealed RNA/primer mixture was precipitated and resuspended in 20 μl of 50 mM-Tris/HCl, pH 8.3, 75 mM-KCl, 3 mM-MgCl\(_2\), 10 mM-dithiothreitol, 50 μg of actinomycin D/ml and 0.5 mM-deoxyribonucleotide triphosphate, and extended with 100 units of Superscript MMLV reverse transcriptase at 37°C for 60 min. Remaining RNA was digested with DNAAse-free ribonuclease A, extracted with phenol/chloroform and ethanol-precipitated. The precipitated extension products were re-suspended in sample buffer and electrophoresed through a 5% (w/v) denaturing acrylamide sequencing gel. The size of the extended product was determined by reference to sequence reactions run on the same gel.

RNAase protection analysis

A 315 bp HindIII/XbaI genomic fragment containing the 5'-flanking region of the gene was subcloned into pBluescript II. Antisense riboprobes were generated with T3 RNA polymerase in the presence of [\(\alpha\text{-}^{32}\text{P}\)]UTP according to the specifications of the manufacturer (Stratagene). Bovine lung total RNA and the \(\alpha\)P-labelled cRNA was hybridized in 30 μl of 40 mM-Pipes containing 80% (v/v) formamide, 40 mM-NaCl and 1 mM-EDTA at 45°C or 55°C for 16 h. Non-annealed RNA was digested for 1 h at 30°C with 20 μg of ribonuclease A/ml and 2 μg of ribonuclease T1/ml in 300 μl of digestion buffer containing 10 mM-Tris/HCl, pH 7.4, 300 mM-NaCl and 5 mM-EDTA, followed by digestion with proteinase K and ethanol precipitation. Protected fragments were analysed by electrophoresis.

RESULTS AND DISCUSSION

Isolation and mapping of genomic clones

A bovine genomic library in EMBL-3 was screened with bovine ET\textsubscript{B} endothelin receptor cDNA and numerous positive phage clones were obtained. Four overlapping clones, \(\lambda\)BerG1, \(\lambda\)BerG2, \(\lambda\)BerG3 and \(\lambda\)BerG4, were selected for mapping (Fig. 1) and analysis of intron/exon organization (Fig. 1c).

Intron/exon structure

The location of 7 exons of the ET\textsubscript{B} receptor gene was determined by Southern hybridization analysis and DNA sequencing; the positions of splice junctions were deduced by comparing the nucleotide sequence of the gene with that of the receptor cDNA (Fig. 2). The gene, spanning more than 36 kb, is relatively large and composed of 7 exons and 6 introns, one of which is more than 23 kb. The other introns ranged from 130 bp to 4.2 kb. The exons varied in size from 109 to 2000 bp. The

![Fig. 1. Restriction enzyme map and organization of the bovine ET\textsubscript{B} receptor gene](image)

(a) The four lines represent cDNA clones used to determine the 5'-flanking region. All clones were obtained from a specific-primed cDNA library using a synthetic oligonucleotide corresponding to bases 107-126 (arrow); however, two clones, bETR13 and bETR15, turned out to be primed non-specifically, a phenomenon often observed. (b) Schematic representation of the cDNA structure of the bovine ET\textsubscript{B} receptor. Non-coding regions are shown by hatched boxes. Putative membrane-spanning domains are shown by closed boxes and are numbered I-VII. (c) The structure of the bovine ET\textsubscript{B} gene. The exons are indicated by boxes and are numbered 1-7. Beneath the genomic map is shown the restriction enzyme map; EcoRI and HindIII sites are indicated by vertical lines. (d) The five lines labelled \(\lambda\)BerG1, \(\lambda\)BerG2, \(\lambda\)BerG3 and \(\lambda\)BerG4 represent the overlapping phage clones used to determine the gene structure.

![Fig. 2. Intron/exon splice sites of the bovine ET\textsubscript{B} receptor gene](image)

The sizes of the introns and exons in bp and the nucleotide sequences around the splice sites are indicated. The putative membrane-spanning domains (MSD) in each exon are shown, and the percentages of identical amino acids in ET\textsubscript{B} and ET\textsubscript{A} are also indicated in parentheses, although the non-homologous N-terminal region in the first exon is omitted. We have demonstrated that this N-terminal region is not responsible for ligand specificity [28]. The sequences at the beginning of the first exon correspond to the major transcription start site determined by RNAase protection analysis.
intron/exon boundaries conform to the consensus splice sequences [14]. Two of the junctions split codons (type 2 introns [15]), whereas codons are preserved in the four remainders (type 0 introns).

Comparison of genomic and protein organization has revealed that there is considerable correspondence between individual exons and the functional domains of the protein. The exon structure determined here, therefore, may provide useful information for determining the ligand-specificity-conferring region by, for example, making chimaeric receptors of ETₐ and ET₇, i.e. the exons or the locations of introns can effectively be used to determine the portions to be exchanged between the two receptor subtypes. We have thus included in Fig. 2 the amino acid sequence similarities of the portions encoded by the individual exons of ETₐ compared with the corresponding regions of ET₇. The region corresponding to exon 3 showed the lowest similarity (52.9%), and therefore is the region most likely to be responsible for the ligand specificity. This point will be resolved by future studies using chimaeric constructions.

**Nucleotide sequence of 5'-untranslated region and 5'-genomic flanking sequence**

The nucleotide sequence upstream of the translation initiation site was determined from a BamH1/XbaI subclone of the λBerG3 genomic clone. Approx. 1000 bases upstream were sequenced (Fig. 3). The sequence of the 5'-untranslated region was determined exclusively from the genomic clones for bases −1019 to −148, and from the genomic and cDNA clones for bases −147 to −1. When the 5'-untranslated region of the genomic clone λBerG3 was sequenced, it turned out to be completely different from the previously reported cDNA sequence [11] at nucleotide −148 and further upstream. The sequence surrounding nucleotide −148 did not fit the consensus of the splice site. We suspected that the difference might result from an artifact arising by rearrangement of DNA sequences during construction of the cDNA library, and sought additional cDNA clones. New clones containing a longer 5'-region of the cDNA sequence were obtained by screening a newly prepared specific-primer-extended bovine lung cDNA library. Four independent clones were sequenced; the longest clone, bETR15, extended 269 additional bases in the 5' direction from position −148. The nucleotide sequence of the extended region was consistent with that of the genomic clone. The previously published cDNA sequence was therefore revised in its 5'-untranslated region, deleting the sequence corresponding to nucleotides 1–924 of the previous clone bETR5 [11].

The above specific-primer-extended cDNAs may also help to determine the start site of transcription. Three out of the four cDNAs had similar 5'-termini (Fig. 1a) which correspond to bases −239, −232 and −226, suggesting that transcriptional initiation begins close to these sites. To confirm this, RNAase protection analysis was performed. Bovine lung total RNA was hybridized to a radiolabelled 352 bp antisense cRNA probe corresponding to exon 1 and sequence upstream from this exon, and digested with ribonuclease A and T1. In complete agreement with the result of the primer-extended cDNA analysis, multiple protected fragments were recovered (Fig. 4), which located the transcription initiation site at positions −237 to −229. The hybridization conditions we used were stringent [55 °C in 80% (v/v) formamide] and this pattern was reproducible with two different preparations of bovine lung total RNA. No apparent TATA box was found in the promoter region (−350 to −260; Fig. 3).

In addition to these major sites there must be an additional upstream transcription start site(s), since we obtained, although as a minor species, a longer cDNA (bETR15, Fig. 1a) and the corresponding cRNA–mRNA hybrid was protected in the RNAase protection analysis (Fig. 4). However, it is not clear whether the upstream start site is located close to the 5' end of the longest cDNA clone bETR15 or further upstream because, in primer extension analysis using the primers corresponding to bases −210 to −191 and −128 to −112 (Fig. 3), the appearance of specific transcripts was influenced by the locations of the primer binding region, probably due to a stem-and-loop structure in the bovine ETₐ mRNA.

**Southern blot analysis of total bovine genomic DNA**

To estimate the copy number of the bovine ETₐ receptor gene, high-molecular-mass DNA extracted from bovine cultured endothelial cells was digested separately with two restriction enzymes, electrophoresed on a 1% (w/v) agarose gel, transferred to a nylon membrane and probed under stringent conditions with the radiolabelled EcoRI/EcoRI fragment derived from exon 7. As shown in Fig. 5, the probe hybridized to single genomic restriction fragments whose sizes agree with those predicted from the restriction map of the genomic clone λBerGl (Figs. 1 and 5), suggesting that the ETₐ receptor gene is present as a single copy in the bovine genome.
Comparison of the \( \text{ET}_B \) gene with other intron-containing heptahelical receptor genes

Many G-protein-coupled receptor genes lack introns; in fact, until 1990, only two exceptions were known: the genes for opsin [16] and the dopamine \( D_2 \) receptor [17,18]. However, a number of precedents now exist for the presence of introns within the coding region of such receptor genes [19]. Fig. 6 compares the splice sites of the heptahelial receptors encoded by the intron-containing genes. The splice sites occur at or near the borders of putative membrane-spanning domains. Similar cases have been reported for other membrane proteins such as the atrial natriuretic peptide receptors [13], the Cl\(^-\)/HCO\(_3\)\(^-\) exchanger [20], the nicotinic acetylcholine receptor [21] and glucose transporters [22]. It is noteworthy that the second intron interrupts the canonical Asp-Arg-Tyr sequence at the end of the third transmembrane domain, since similar interruptions have been demonstrated in the case of substance P [23], substance K [24], dopamine \( D_2 \) [25,26] and dopamine \( D_3 \) [27] receptor genes. The location of the sixth intron of the bovine \( \text{ET}_B \) receptor gene is unique in that it interrupts the coding sequence at the external border of the seventh transmembrane domain, where no introns occur in the other members of the family so far analysed.

Accumulation of information on the gene structure of the heptahelial receptors will help us to understand how this family of genes evolved through duplication, recombination, loss or insertion of introns and nucleotide substitution. The present study will also provide a starting point for the analysis of the mechanism of regulation of \( \text{ET}_B \) gene expression.
We thank Setsuko Satoh, Kazuko Tanaka, Yohko Suzuki and Rieko Yamamoto for their secretarial and technical assistance. This work was supported by Grants-in-Aid for Scientific Research on Priority Areas from the Ministry of Education, Science and Culture, Japan, and research grants from Chichibu Cement, the Monbusho International Scientific Research Program, the Kanagawa Academy of Science and Technology, the Tokyo Hypertension Conference and the Japan Shipbuilding Industry Foundation.

REFERENCES