Biochemical and Biophysical Research Communications **267,** 228–235 (2000)
doi:10.1006/bbrc.1999.1951, available online at http://www.idealibrary.com on IDEAL®

Tissue-Specific N-Terminal Isoforms from Overlapping Alternate Promoters of the Human AE2 Anion Exchanger Gene

Juan F. Medina,[1] Jon Lecanda, Alberto Acín, Pawel Ciesielczyk, and Jesús Prieto
Laboratory of Molecular Genetics, Unit of Hepatology and Gene Therapy, University Clinic and Medical School, University of Navarra, E-31008 Pamplona, Spain

Received November 26, 1999

Previously, we isolated the human AE2 (*SLC4A2*) gene, a member of the sodium-independent anion exchanger family. Rat ortholog of this gene was reported to drive alternative transcription yielding N-terminal variants of the AE2a message. We thus analyzed the human AE2 gene in this regard. Using HepG2 cells, two alternative first exons, each splicing to exon 3 in alternative transcripts, were found to be transcribed from overlapping sequences of intron 2. Exon $1b_1$ corresponds to the rat variant "b" and encodes three initial residues (MTQ) in $AE2b_1$ isoform that replace the first 17 amino acids of AE2a protein, while the novel exon $1b_2$ encodes eight initial residues (MDFLLRPQ) in $AE2b_2$ isoform. The relative abundance of $AE2b_1$ and $AE2b_2$ mRNAs was about 10% of AE2a mRNA each. Alternate promoter sequences have multiple potential binding motifs for liver-enriched factors, and dual-luciferase assays indicated that they possess the ability for driving transcription in transiently transfected HepG2 cells. Tissue survey showed that expression of human $AE2b_1$ and $AE2b_2$ transcripts is restricted to liver and kidney, while AE2a mRNA was encountered in all examined tissues. Our findings reveal a characteristic tissue-specific expression of two N-terminal variants of human AE2 from overlapping sequences within intron 2, one of which is a novel isoform. © 2000 Academic Press

Sodium-independent anion exchangers (AE) are a family of membrane proteins that mediate the electroneutral exchange of Cl^- for HCO_3^- ions across cell membranes (1). AE proteins occur in both polarized and non-polarized cells, and they may be involved in intracellular pH and cell volume regulation as well as in transepithelial acid/base transport. Thus far, three members of the AE family (AE1, AE2, and AE3) have been thoroughly identified (1, 2). In humans, the AE1 (*SLC4A1*) gene has been assigned to chromosome 17q21, the AE2 (*SLC4A2*) gene has been localized to chromosome 7q35–q36, and the AE3 (*SLC4A3*) gene has been mapped to chromosome 2q36 (see corresponding citations in Refs. (1, 3). Furthermore, the cloning of the human AE1 gene and, partially, that of the human AE3 gene have previously been reported (4–7). More recently, we have described the molecular cloning and characterization of the human AE2 gene (3).

While AE1 protein is mainly expressed in erythroid cells (and in kidney cells as a truncated isoform) and AE3 is predominantly encountered in excitable tissues such as brain and cardiac muscle, AE2 appears to be more widely distributed (1). The membrane location of AE2 in polarized cells was primarily described at the basolateral membrane in mouse choroid plexus (8). However, our immunohistochemical data in humans provided evidence for differences between tissues (9). A same monoclonal antibody against human AE2 was shown to stain the basolateral membrane of epithelial cells in human choroid plexus as well as canaliculi and apical membranes of bile duct epithelial cells in the human liver (9). Differential membrane targeting of human AE2 might result from the occurrence of AE2-protein isoforms due, for instance, to the use of alternate promoters and/or alternative splicing of the mRNA. In fact, the phenomenon of alternate promoters has already been reported for both human AE1 (4, 5, 10) and AE3 genes (7). In rat AE2 gene, the occurrence of alternates to the upstream promoter (named AE2a) has been described: sequences of intron 2 may drive transcription of one N-terminal variant named AE2b and sequences of intron 5 yield variants "c" in a

Abbreviations used: AE, sodium-independent anion exchanger; DMEM, Dulbecco's modified Eagle medium; nt, nucleotides; PCR, polymerase chain reaction; RACE, rapid amplification of cDNA ends; RT-PCR, reverse-transcription PCR.

[1] To whom correspondence should be addressed. Fax: +34 948 425649. E-mail: jfmedina@unav.es.

0006-291X/00 $35.00

tissue-specific manner (11). No data on alternate promoters for the human AE2 gene were available.

We now report that the human AE2 gene may also use alternate promoters in a peculiar tissue-specific manner. In addition to the AE2a message transcribed from the upstream promoter (3) which is expressed in most tissues, two N-terminal variants appear to be expressed from overlapping sequences of intron 2 in liver and HepG2 cells as well as in kidney. One of these variants corresponds to the rat AE2b transcript, hereby referred to as AE2b$_1$ mRNA; the other one is a novel isoform named AE2b$_2$ mRNA. Luciferase assays with transiently transfected HepG2 cells confirmed the functionality of promoter sequences.

MATERIALS AND METHODS

Genomic clones and DNA preparations. Clones with the human AE2 gene (cosmid HAE2co5 and others) have been described elsewhere (3). Nucleotide numbering for alternate promoters, alternative exons, and introns (see Fig. 1) are as follows: A of each ATG translation initiation codon within intron 2 is designated +1b$_1$ or +1b$_2$, with negative integers plus "b$_{1/2}$" proceeding 5′ in corresponding exon 1b$_1$ or 1b$_2$ and promoter sequences, and positive integers plus "b$_{1/2}$" proceeding 3′ until junction with alternative introns 1b$_1$ or 1b$_2$. Introns are numbered with positive integers proceeding from 5′ donor sites to 3′ acceptor sites. Genomic DNA was extracted with a commercial DNA extraction kit (Stratagene). Recombinant cosmid and plasmid DNAs were isolated with High Pure Plasmid Isolation kits (Roche Molecular Biochemicals) for minipreps, or with Nucleobond AX columns (Macherey–Nagel, Düren, Germany) for maxipreps. Sequencing of double-stranded DNA was with an ABI Prism Cycle Sequencing kit (PE Biosystems) followed by automated capillary electrophoresis in an ABI Prism 310 Genetic Analyzer.

RNA preparation, RT-PCR, and RACE. Total RNA was isolated from HepG2 cells and from eight human tissues (liver, stomach, rectum, prostate, kidney, and thyroid gland) with Ultraspec II kit (Biotecx). HepG2 RNA enriched for poly(A)$^+$ RNA with mRNA Isolation kit (Roche Molecular Biochemicals), was employed in RACE procedures (12) using the Marathon cDNA Amplification kit (Clontech). To determine 5′ ends, adaptor primer AP1 and the gene specific antisense primer JF1098 (nt 398–373 of the human AE2 cDNA, Ref. 3) were used for a first-round amplification (touchdown PCR) of adaptor-ligated double-stranded cDNA, followed by a subsequent nested amplification with adaptor primer AP2 and the specific antisense primer JF1096 (nt 233–208 in the AE2 cDNA). Resin extracted cDNA amplicons were ligated into pGEM-T vector (Promega) followed by transformation of *Escherichia coli* strain XL1-Blue. Plating of resultant 5′ RACE-amplicon libraries allowed for the screening, isolation, sequencing, and characterization of multiple AE2 cDNAs having the three different 5′ termini (see Fig. 1). The proportion between AE2 mRNA isoforms in HepG2 cells was estimated by colony hybridization of some of the 5′ RACE-amplicon libraries. In this methodology it was assumed that the efficacy of PCR amplification of all three isoforms using the same pair of primers is putatively very similar. After plating onto nitrocellulose filters and overnight growth, six replicas were obtained from each original, duplicates being hybridized with ^{32}P-labeled isoform-specific probes: oligonucleotides JF1003 (5′-TCCAGAGCGAGCGGGTTATG) for AE2a and JF1239 (5′-GCCATGGACTTCCTCCTGCGGCCTC) for AE2b$_2$, as well as a PCR fragment spanning nt +959/+1146 of intron 2 (amplified with flanking oligonucleotides on cosmid HAE2co5 as template) for AE2b$_1$. Probing with oligonucleotides used a particular protocol with SSC-rich solutions containing 0.05% of sodium pyrophosphate, as described (13). For autoradiograms, BioMax MS films with BioMax MS intensifying screens (Kodak) were exposed at −80°C for variable intervals of time. To inspect the 3′ side of the human AE2b$_1$ mRNA, a RACE procedure was also employed, using as sense primers only exon-1b$_1$ specific oligonucleotides. A first touchdown amplification was carried out on Marathon cDNA with oligonucleotides JF1209 (5′-GCTGGACCAAGGCTGCCTGGCCA) and AP1 (20 nM), followed by three consecutive nested amplifications: one with oligonucleotides 5′-CTGGCCAGGCGGCTGCCGCTTAGC and AP1 (15 cycles), then with 5′-CTTAGCTGGGCTGAGCTCTTACAAGC and AP2 (20 cycles), and finally with 5′-CGCCCGCAGGATGACTCA and AP2 (35 cycles). The final concentration of gene specific primers was each 200 nM versus 40 nM for adaptor-primers (14). Because only adaptor primers were employed as antisense primers, formation of PCR artifacts via coincident sequences with AE2a cDNA was circumvented. The expression of mRNA isoforms was evaluated in several tissues by semiquantitative RT-PCR (see Fig. 2). Reverse transcription of 3 μg of corresponding total RNA with random hexamers and M-MLV reverse transcriptase (Gibco) was performed in a total volume of 40 μl; PCR fragments for AE2a, AE2b$_1$, and AE2b$_2$ isoforms were produced with 4 μl of each cDNA pool, in a final volume of 50 μl of the reaction mixture, including a specific sense primer (JF1003, JF1209, and JF1239, respectively), and JF1096 as a common antisense primer for all isoforms (see above for details on these oligonucleotides). As normalizing internal control for each sample, amplification of a fragment of GAPDH cDNA was performed with oligonucleotides 5′-CCAAGGTCATCCATGACAAC and 5′-TGTCATACCAGGAAATGAGC (upstream and downstream primers, respectively). Aliquots were taken from amplification reactions at 20, 25, and 30 cycles for GAPDH, and at 30, 35, and 40 cycles for AE2 isoforms, and electrophoresed in 1% agarose gel stained with ethidium bromide. Bands were visualized with an ultraviolet lamp and analyzed with the software Molecular Analyst/PC (Bio-Rad). In Fig. 2, only amplified bands from aliquots that are in the exponential phase (i.e., with no plateau effect at the specified number of cycles), are shown. No contamination of genomic DNA in RNA preparations could have been responsible for any amplicon detected, since no introns were found in their sequences (introns 1 and 2 for the AE2a amplicon as well as introns 1b$_2$ or 1b$_1$ for corresponding alternative amplicons). Also, intron 3 intervenes in the sequence of oligonucleotide JF1096 employed as a common antisense primer. Routine PCR amplifications were performed in the reaction mixture supplied with the *Taq* DNA polymerase employed (either Roche Molecular Biochemicals, Mannheim, Germany or Eurobio, Les Ulis, France), supplemented most frequently with 5% dimethyl sulfoxide (Sigma). Particular PCRs such as touchdown PCR for RACE procedures, were performed with either an Advantage-GC cDNA PCR kit (Clontech) for the 5′ ends or a GeneAmp XL PCR kit (PE Biosystems) for the long 3′ side. Most touchdown amplifications followed the recommendations of Marathon kit manufacturers, although customized annealing temperatures were sometimes crucial for optimal results.

Construction of luciferase expression vectors and luciferase assays. Several stretches of AE2 promoter sequences (Table 1) were amplified with *Pfu* DNA polymerase (Stratagene) using flanking oligonucleotides on cosmid HAE2co5 as template, promoter fragments being further ligated to *Sma*I-linearized vector pGL3 (Promega). All inserts were sequenced to verify correct orientations and that no artifactual mutations were introduced during PCR procedures. Highly pure plasmid DNA prepared with Nucleobond AX columns was always employed for transfections. HepG2 cells were maintained in DMEM with 10% fetal calf serum (Gibco), 2 mM L-glutamine, antibiotics (penicillin and streptomycin), and 10 mM Hepes (final pH adjusted to 7.4), undergoing two passages during 14 days. We found that renewal of cells from the liquid nitrogen stock at regular intervals and regular passages of cell culture were crucial to obtain consistent results, as recently reported by others (15). Cells were plated in six-well culture plates (Greiner) at a density of 3×10^5 cells/well, and further cotransfected with 1 μg of corresponding pro-

moter construct plus 9 μg of pGL3-basic vector, and 2 ng of the internal control vector pRL-SV40 (the latter carrying *Renilla* luciferase driven by SV40 promoter), using the calcium phosphate precipitation method. Cells extracts were prepared 48 h after transfections and assayed for luciferase activity with the Dual-Luciferase Reporter Assay System (Promega), employing a Lumat LB 9507 luminometer (EG & G Berthold).

RESULTS AND DISCUSSION

Identification and Tissue Distribution of Human AE2 mRNA Isoforms

We previously reported the molecular cloning of the human AE2 gene, including the identification of the upstream AE2a promoter (3). Here, we describe how the human AE2 gene can also be expressed from alternate promoters ($AE2b_1$ and $AE2b_2$) in a tissue-specific manner. These alternate promoters are overlapping sequences which span different regions of intron 2 (and upstream DNA stretches, cf. Fig. 1). As deduced from 5′ RACE experiments on HepG2 cells poly(A)$^+$ RNA, transcribed $AE2b_1$ and $AE2b_2$ isoforms differ from AE2a mRNA at the 5′ side (exons 1 and 2 of the latter are replaced either by exon $1b_1$ or $1b_2$, respectively). Also, 5′ RACE experiments suggested corresponding putative transcription initiation sites. Similarly to what was reported for AE2a mRNA (3), the $AE2b_1$ message appeared to have multiple 5′ ends, most of which were located between positions $-134b_1$ and $-93b_1$, while minor sites were at positions T $-197b_1$ and T $-521b_1$ (Fig. 3). For $AE2b_2$ mRNA, however, a main start site was determined at position C $-20b_2$, although a rather minor start site was also found 2-nt upstream (Fig. 3). The 3′ side of the $AE2b_1$ message was ascertained by a RACE procedure involving four nested amplifications (with adaptor primers and only isoform-specific sense primers, cf. Materials and Methods), the last of which yielded a unique band of about 3.9 kb. In the case of $AE2b_2$ mRNA, the short length of exon $1b_2$ precluded appropriate design of multiple isoform-specific nested amplifications. But a unique band with the expected size of 3.8 kb could be obtained by RT-PCR, using oligonucleotides of exon $1b_2$ and exon 23 as sense and antisense primers, respectively. Complete sequence analyses of the 3.9- and 3.8-kb bands confirmed that exons 3–23 in $AE2b_1$ and $AE2b_2$ transcripts are coincident with those of AE2a mRNA. Isoform diversity at 5′ leads to different N-termini in corresponding encoded proteins AE2a, $AE2b_1$ and $AE2b_2$. Thus, the first 17 amino acids of AE2a (MSSAPRRPAKGADSFCT) are substituted by only three residues (MTQ) in $AE2b_1$, and by an eight-residue sequence (MDFLLRPQ) in $AE2b_2$ isoform.

Repeated experiments of colony hybridization with isoform-specific probes (see Materials and Methods) suggested that $AE2b_1$ and $AE2b_2$ transcripts account each for about 10% of AE2a mRNA in HepG2 cells.

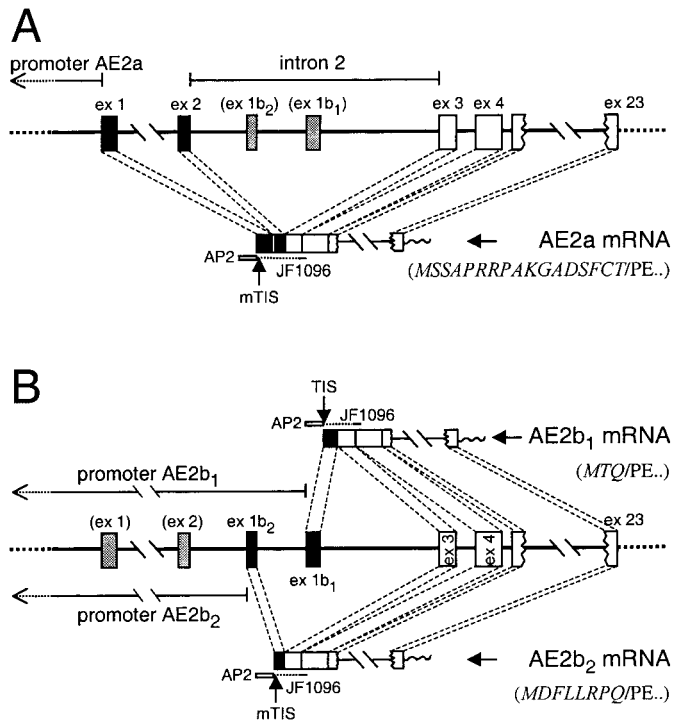

FIG. 1. Splicing patterns for AE2 mRNA isoforms. (A) Exon 1 transcribed from the upstream promoter AE2a is spliced to exon 2 in AE2a mRNA. (B) Alternative exons transcribed from overlapping sequences: Above, exon $1b_1$ transcribed from alternate promoter $AE2b_1$ (5′ half of intron 2 and upstream sequences) is spliced to exon 3 in $AE2b_1$ mRNA (the downstream half of intron 2 results in alternative intron $1b_1$); Below, exon $1b_2$ transcribed from alternate promoter $AE2b_2$ (first quarter of intron 2 and upstream sequences) is spliced to exon 3 in $AE2b_2$ mRNA (sequences of intron 2 downstream to exon $1b_2$ results in alternative intron $1b_2$). Schematic representations of 5′ RACE amplicons obtained with oligonucleotide pairs AP2 and JF1096 (cf. Materials and Methods) are depicted. Encoded residues (single-letter code) specific for each isoform are in italics. Open, black, and gray boxes, represent common exons, specific exons, and possible exons which are not transcribed, respectively, in corresponding transcripts. TIS, transcription initiation site; mTIS, multiple TIS.

Regarding the expression of AE2 isoforms in human tissues, semiquantitative RT-PCRs on tissular total RNA (from liver, stomach, rectum, prostate, kidney, and thyroid gland) indicated that in both liver and kidney, the expression patterns for $AE2b_1$ and $AE2b_2$ mRNAs were very similar to that in HepG2 cells (Fig. 2). The remaining analyzed tissues did not show expression of alternative mRNAs, whereas expression of AE2a mRNA was similarly detected in all tissues (Fig. 2).

Tissue-specific usage of alternate promoters has been reported in most genes of the AE family. In humans, an alternate promoter has been determined for AE1 gene (internal promoter in kidney that is alternative to upstream promoter in erythroid cells) (4, 5, 10) as well as for AE3 gene (internal promoter in the heart differing to upstream promoter in the brain) (7). Like-

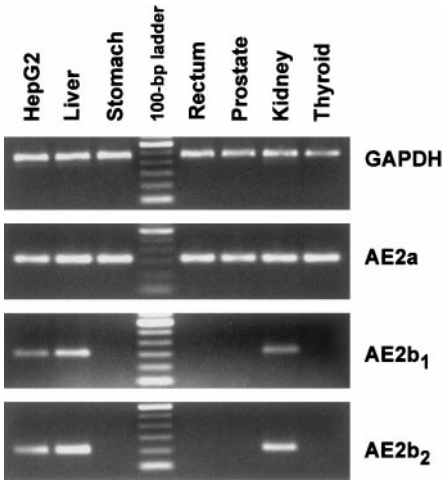

FIG. 2. Semiquantitative RT-PCR for AE2 isoforms in several human tissues and HepG2 cells. RT-PCR for GAPDH mRNA was used as a normalizing control. All bands shown correspond to amplicons which are in the linear phase of amplification at the specified number of cycles. Sizes of each amplicon and respective number of cycles are as follows: AE2a, 310 bp and 30 cycles; $AE2b_2$, 209 bp and 30 cycles; $AE2b_1$, 302 bp and 30 cycles; GAPDH, 465 bp and 25 cycles.

wise, equivalent features have been shown for murine and rat AE1 and AE3 counterparts. Rat ortholog of the AE2 gene has also been reported to drive transcription from alternate promoters (11), one of which locates in intron 2 and corresponds to the human $AE2b_1$ promoter. Thus far, a complete identity of human $AE2b_1$ mRNA with rat AE2b message can only be alleged, since the splicing pattern of rat AE2b downstream of exon 3 remains yet to be determined. Tissue expression of this isoform certainly shows some differences between the species, rat $AE2b_1$ transcript being highly expressed in the stomach (11), while human $AE2b_1$ mRNA cannot be found in this tissue by RT-PCR (Fig. 2). On the other hand, no counterpart of rat isoforms "c" (reported to be expressed only in the rat stomach, cf. Ref. 11), could be identified in humans. Finally, the human $AE2b_2$ isoform herein described is a novel isoform which has not yet been reported in any animal species.

Sequence Analyses of Alternate Promoters of the Human AE2 Gene

As mentioned, RACE experiments suggested the occurrence of alternate promoters from which AE2 isoforms are transcribed. Thus, sequences upstream of alternative exons $1b_1$ and $1b_2$ (i.e., part of intron 2, exon 2, intron 1, and exon 1, as well as AE2a promoter; cf. Fig. 1), were searched for putative *cis*-regulatory elements. Figure 3 shows that promoter regions within intron 2 indeed contain many potential binding motifs for transcription factors. Since these regions overlap to each other (for instance, position $-100b_2$ is coincident with position $-641b_1$, etc.), some elements may influence the transcription of both messages. Both overlapping promoter regions lack canonical TATA boxes and Inr consensus sequences (16). However, $AE2b_2$ promoter has two potential CCAAT elements close upstream to its start site at position $-20b_2$, the proximal region showing also several potential AP2 sites, E-boxes, and motifs for liver-enriched factors included in the HNF network (17) (Fig. 3). Regarding clustered start sites for $AE2b_1$ (between positions $-134b_1$ and $-93b_1$), there are upstream motifs for GCF, HNF3, and HNF6 (3-bp variants of the two latter, cf. Refs. 18 and 19), overlapping with each other, while several inverted CACCC boxes locate among clustered start sites (these boxes are not explicit in Fig. 3 for simplicity). Near downstream there is a striking 12-bp GC-rich palindromic sequence (positions $-86b_1$ to $-97b_1$), which overlaps with a 9-bp consensus sequence for a TNF-α response element (TNFRE). Moreover, three consecutive GCCG direct repeats (between positions $-19b_1$ and $-30b_1$), are found close upstream to a potential MED-1 site (a factor involved in transcription initiation at multiple sites from some TATA-less promoters) (20) that is located at the right distance from the 3′-most start site of multiple $AE2b_1$ start sites (Fig. 3). Additional GC-rich stretches with potential binding sites for GCF, ETF, Sp1, and MAZ (21–23) are widely distributed along alternate promoter overlapping sequences (similarly to what was observed for the upstream AE2a promoter, which is also a TATA-less promoter associated with multiple start sites), together with abundant potential sites for liver-enriched factors. In addition to those mentioned above, there are two slight variants of the HNF6 motif at positions $-81b_2$ and $-457b_2$, 1-bp variant of the 15-bp consensus sequence for HNF1 (24) at position $-113b_2$, and several HNF4/HNF4-core motifs (25). Furthermore, potential response elements (ERE, MRE, ARE, GRE core, CRE-like palindromic sequence, and IL-6 RE of both type I and type II) are also found (Fig. 3). There are other motifs within intron 2—for instance, several potential STAT consensus sequences TTN(4–6)AA as well as CT-rich regions, i.e., possible GAGA boxes—which, for simplicity, are not explicit in Fig. 3. Upstream to intron 2, in exon 2 and intron 1, there are also potential elements such as another TNFRE ($-825b_2$) and a HIP1 site ($-642b_2$), as well as several AP2 motifs.

Dual-Luciferase Assay with Promoter Sequences

At present, the biological relevance of nucleotide segments discussed above is mostly speculative, and much work is needed to elucidate exactly the pathways of AE2 gene expression in humans. To this end, constructs with putative promoter sequences for $AE2b_1$ and $AE2b_2$ as well as for AE2a (Table 1) were tested for

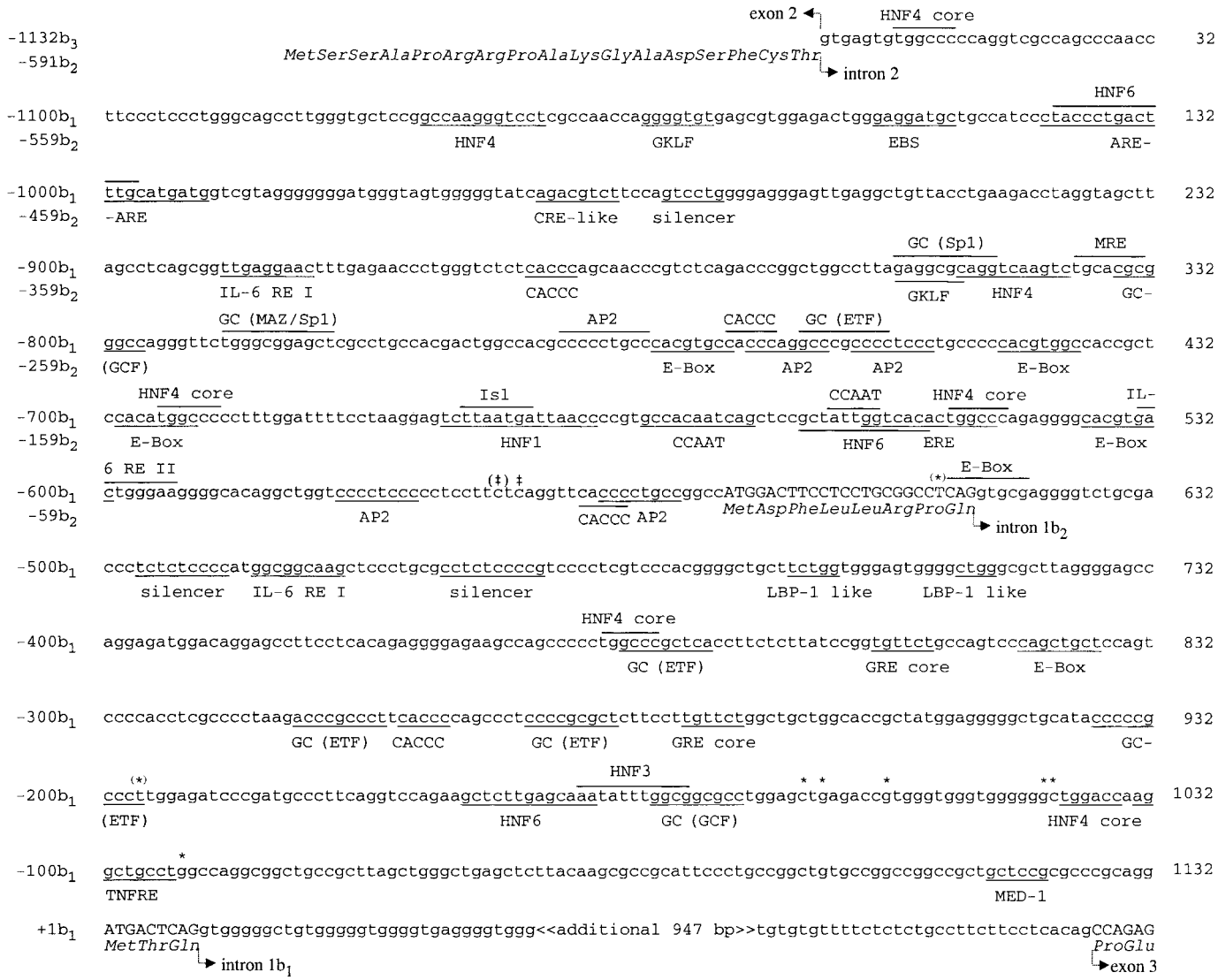

FIG. 3. Human overlapping alternate promoter sequences for $AE2b_1/AE2b_2$ within intron 2. Potential binding sites of transcription factors are either underlined or overlined (for simplicity, some more potential motifs mentioned in Results and Discussion are not explicit in the figure). The coding sequence of each alternative exon is shown in uppercase letters with encoded amino acids below. Superscripts * and ‡ indicate transcription start sites of exons $1b_1$ and $1b_2$, respectively (in parentheses, corresponding minor sites). It may be noted that intron 2 is 2152 bp in length.

their functionality in HepG2 cells (Fig. 4). All constructs yielded substantial luciferase activities, while background activity for the pGL3-basic vector with no insert was negligible. Among investigated constructs for alternate promoter regions, $I\text{-}b_2$ containing the $AE2b_2$ proximal promoter yielded the highest luciferase activity (18.77 ± 1.14 relative units), while a 5′ elongation in $II\text{-}b_2$ led to a decreased promoter activity (11.42 ± 0.62 units) (Fig. 4). This effect might be related to a potential motif located in the elongated sequence (position $-401b_2$, Fig. 3), which is identical to a silencer element identified in the human ANT2 promoter (26). Constructs relative to $AE2b_1$ promoter ($I\text{-}b_1$, $II\text{-}b_1$, and $III\text{-}b_1$) yielded lower activities (5.32 ± 0.20, 3.99 ± 0.14, and 2.74 ± 0.13 units, respectively) than $AE2b_2$ constructs (all $P < 0.001$). Since $AE2b_2$ promoter is overlapping with $AE2b_1$ promoter, sequences in $I\text{-}b_1$ and $II\text{-}b_1$ are both, in fact, 5′ and 3′ elongations of the short $I\text{-}b_2$ sequence (Fig. 4). Two potential silencer elements (each, 1-bp variant from the consensus sequence) (27, 28) are found immediately downstream from exon $1b_2$ at positions $-489b_1$ and $-459b_1$ (Fig. 3), and they could account for the observed further reduction in luciferase activity.

Regarding the AE2a promoter, construct I-a containing the longest analyzed insert yielded 5.26 ± 0.43 relative luciferase units. The activity dropped dramatically after deletion of as little as 176-bp at the proxi-

TABLE 1

Constructs of PCR Fragments of AE2 Promoters in Luciferase Vector pGL3-basic

Name	Promoter region	Insert boundaries (genomic sequence)	Oligonucleotides used in PCR (nt in the genomic sequence)	Insert length (bp)
I-a	AE2a	−1373 and ivs1 + 3	−1373/−1354 and ivs1(+3)/−81	1313
II-a	AE2a	−1373 and ivs1 + 3	−1373/−1354 and ivs1(+3)/−81	1137
III-a	AE2a	−539 and ivs1 + 3	−539/−521 and ivs1(+3)/−81	479
I-b_1	AE2b_1	ivs2 + 121 and ivs2 + 1146	ivs2(+121/+140) and ivs2(+1146/+128)	1026
II-b_1	AE2b_1	ivs2 + 121 and ivs2 + 821		701
III-b_1	AE2b_1	ivs2 + 822 and ivs2 + 1146		325
I-b_2	AE2b_2	ivs2 + 306 and ivs2 + 615	ivs2(+306/+325) and ivs2(+615/+592)	310
II-b_2	AE2b_2	ivs1 + 1934 and ivs2 + 615	ivs1(+1934/+1958) and ivs2(+615/+592)	957

Note. Insert in construct II-a had a deletion from positions −352 to −177 in the AE2a promoter, produced when no dimethyl sulfoxide was included in the PCR mixture. Deleted constructs II-b_1 and III-b_1 were obtained from I-b_1 by double digestions with *Pvu*II + *Hin*dIII and *Nhe*I + *Pvu*II, respectively, and further religations. In addition to promoter regions, both constructs I-b_1 and III-b_1 contain the first exon of AE2b_1 mRNA (including exon/intron junction at position +9b_1). Both constructs I-b_2 and II-b_2 contain the first exon of AE2b_2 mRNA, in addition to promoter regions; II-b_2 also includes the whole exon 2 and the preceding 3′ side of the 2161-bp-long intron (ivs) 1; cf. Fig. 4.

mal promoter region in II-a (2.12 ± 0.13 units) (Fig. 4), indicating that this short region accounts for a substantial proportion of the promoter activity. By contrast, deletion of a notable portion of upstream sequence in III-a led to a slight and nonsignificant increase in the promoter activity (5.64 ± 0.67 units) (Fig. 4). Should positive regulatory motifs occur in the deleted region, their effect could be counterbalanced by silencer elements to result in the almost null effect of the distal deletion. AT-rich sequences in the AE2a promoter (for instance, the sequence spanning positions −760 to −910) (3) are possible candidates as silencer elements, since negative regulatory effects have been reported for this type of sequence in some genes expressed in hepatocytes (29).

It may be noticed that these luciferase data obtained from *in vitro* experiments of transient transfections with promoter constructs do not harmonize exactly with our results of colony hybridizations which indicated that *in vivo* expression rates of AE2 mRNA isoforms in HepG2 cells follow the pattern AE2a > AE2b_1 = AE2b_2. Certainly, *in vivo* activities from promoter sequences in the gene context are very much dependent on additional interactions of *cis*- and

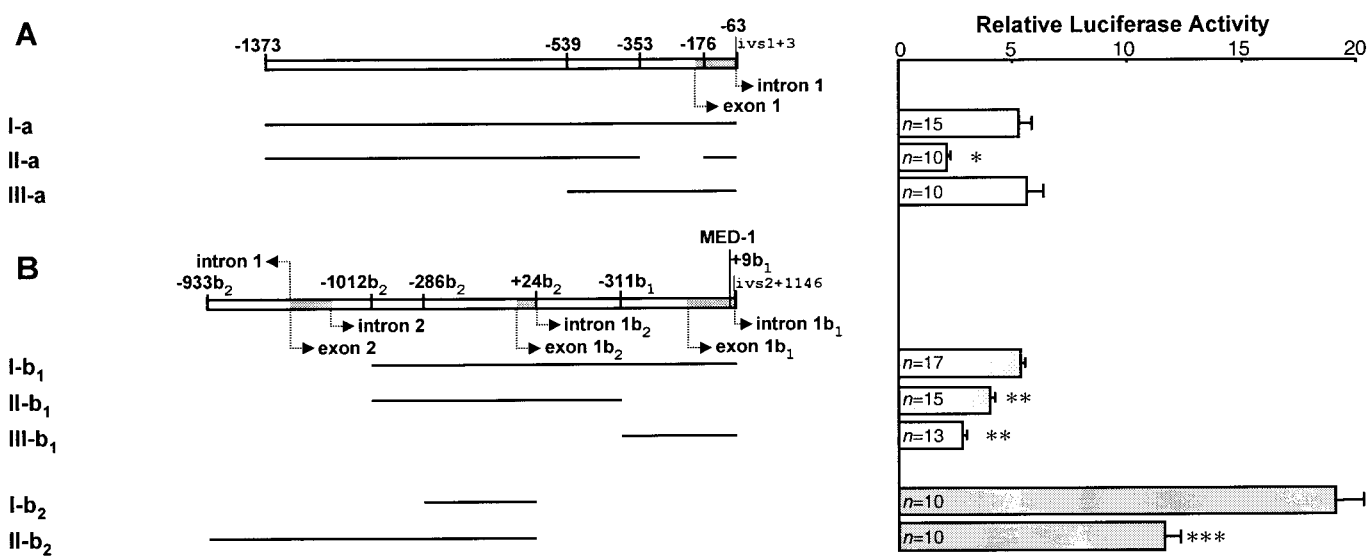

FIG. 4. Functional activity of human promoter for AE2a (A) and overlapping promoters for AE2b_1/AE2b_2 (B). Dual luciferase assays of HepG2 cells transiently transfected with corresponding constructs (left; cf. Table 1) were carried out in triplicate as described under Materials and Methods, and repeated 10 to 17 times (*n*). Values (right) are given as ratios of firefly luciferase activity of recombinant vectors to *Renilla* luciferase activity. Data for AE2a promoter were analyzed separately from those for AE2b_2/AE2b_1 overlapping promoters with the tests of Shapiro–Wilk, ANOVA, and Student–Newman–Keuls (all three were applied in a row). Mean values ± SEM are used. All *P* values are two-tailed. *, **, and *** indicate significances (always $P < 0.001$) versus the other values for AE2a, AE2b_1, and AE2b_2 constructs, respectively.

trans- regulatory elements, on methylation of some regions (it is now accepted that methylation of DNA represses transcription from genomic sequences with otherwise promoter capacities; cf. Refs. 30–32), etc. Moreover, interplays between transcriptional units resulting in promoter occlusion may take place here as described for other genes (33, 34). Additional experiments are under way in order to define the relationships between potential elements observed in AE2 promoters, methylation status, and promoter activities.

Possible Roles of Alternate Promoters and N-Terminal Variants of AE2

As for other genes (see review in Ref. 35), the physiological role for the occurrence of alternate promoters in the human AE2 gene in some tissues might be related to differential regulation of the gene expression, differential targeting of resultant AE2 protein isoforms, or both. At first glance, sequence differences between upstream and overlapping alternate promoters could suggest differential regulation as the main role, the upstream promoter possibly functioning in a more constitutive and widespread fashion. Also, this role is supported by the great similarity between encoded AE2 proteins, with only slight differences just at the beginning of their primary structure. However, little variations at the N-terminus should not be underestimated. In several proteins, such as human invariant chain Ii (36), murine ES cell differentiation inhibiting activity (DIA) factor (37), and *Saccharomyces cerevisiae* IPP transferase (38), little N-terminal variations have been shown to lead to alternative targeting of the protein.

In both human liver and kidney, using a specific AE2 monoclonal antibody against a peptide common to all isoforms, we localized immunoreactivity in polarized cells at only one of the poles. Liver immunoreactivity was located at the apical pole of hepatocytes (canaliculi) and cholangiocytes (9), which is compatible with an involvement of AE2 in transepithelial bicarbonate secretion into bile by apical chloride/bicarbonate exchange. Kidney immunoreactivity was located instead at the basolateral membrane of cells of thick ascending limbs and distal convoluted tubules (39), which concurs with these cells functioning to acidify the luminal fluid. Probably all three AE2 isoforms are only targeted at respective poles where they are detected. However, we cannot rule out the possibility that AE2 isoform(s) might also occur at cell locations (e.g., the corresponding opposite pole or associated with the intracellular membrane system) where it/they can remain "masked" from the antibody. In fact, a phenomenon of this kind has been suggested for a seemingly "hidden" apical AE1 protein in α-intercalated cells of renal collecting ducts (40).

We have previously reported that the expression of the human AE2 gene is decreased in the liver of patients with primary biliary cirrhosis (41, 42), which is in line with recent evidence for alterations of biliary bicarbonate secretion in these patients (43). Future studies should explore whether decreased AE2 is accounted for by one or more isoforms of the protein.

ACKNOWLEDGMENTS

The authors thank E. Sáez for excellent technical assistance. We are also grateful to Drs. C. D. Funk, F. J. Novo, and P. Fortes for comments on the manuscript and M. A. Martínez for help with statistics. This work was supported by grants from the Spanish FIS (Projects 96/0346 and 99/0647) and the "Fundación Echébano." Sequence data for the human AE2 gene have been deposited with the GenBank/EMBL Data Libraries under Accession Nos. U62290, U62531, and U76667-U76669.

REFERENCES

1. Alper, S. L. (1994) *Cell. Physiol. Biochem.* **4,** 265–281.
2. Kopito, R. R. (1990) *Int. Rev. Cytol.* **123,** 177–199.
3. Medina, J. F., Acín, A., and Prieto, J. (1997) *Genomics* **39,** 74–85. [published erratum appears in *Genomics* **46,** 535, 1997]
4. Sahr, K. E., Taylor, W. M., Daniels, B. P., Rubin, H. L., and Jarolim, P. (1994) *Genomics* **24,** 491–501.
5. Schofield, A. E., Martin, P. G., Spillett, D., and Tanner, M. J. A. (1994) *Blood* **84,** 2000–2012.
6. Su, Y. R., Klanke, C. A., Houseal, T. W., Linn, S. C., Burk, S. E., Varvil, T. S., Otterud, B. E., Shull, G. E., Leppert, M. F., and Menon, A. G. (1994) *Genomics* **22,** 605–609.
7. Linn, S. C., Askew, G., Menon, A. G., and Shull, G. E. (1995) *Circ. Res.* **76,** 584–591.
8. Lindsey, A. E., Schneider, K., Simmons, D. M., Baron, R., Lee, B. S., and Kopito, R. R. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 5278–5282.
9. Martínez-Ansó, E., Castillo, J. E., Díez, J., Medina, J. F., and Prieto, J. (1994) *Hepatology* **19,** 1400–1406.
10. Kollert-Jöns, A., Wagner, S., Hübner, S., Appelhans, H., and Drenckhahn, D. (1993) *Am. J. Physiol.* **34,** F813–F821.
11. Wang, Z., Schultheis, P. J., and Shull, G. E. (1996) *J. Biol. Chem.* **271,** 7835–7843.
12. Frohman, M. A. (1990) *in* PCR Protocols. A Guide to Methods and Applications, pp. 28–38.
13. Duby, A. (1993) *in* Current Protocols in Molecular Biology, pp. 6.4.1–6.4.3.
14. Bespalova, I. N., Adkins, S., and Burmeister, M. (1998) *Biotechniques* **24,** 575–577.
15. Bates, N. P., and Hurst, H. C. (1996) *in* Clinical Gene Analysis and Manipulation. Tools, Techniques and Troubleshooting, pp. 332–348.
16. Smale, S. T. (1997) *Biochim. Biophys. Acta* **1351,** 73–88.
17. Duncan, S. A., Navas, M. A., Dufort, D., Rossant, J., and Stoffel, M. (1998) *Science* **281,** 692–695.
18. Overdier, D. G., Porcella, A., and Costa, R. H. (1994) *Mol. Cell. Biol.* **14,** 2755–2766.
19. Samadani, U., and Costa, R. H. (1996) *Mol. Cell. Biol.* **16,** 6273–6284.
20. Ince, T. A., and Scotto, K. W. (1995) *J. Biol. Chem.* **270,** 30249–30252.

21. Kageyama, R., Merlino, G. T., and Pastan, I. (1989) *J. Biol. Chem.* **264,** 15508–15514.
22. Kageyama, R., and Pastan, I. (1989) *Cell* **59,** 815–825.
23. Parks, C. L., and Shenk, T. (1996) *J. Biol. Chem.* **271,** 4417–4430.
24. Tronche, F., Ringeisen, F., Blumenfeld, M., Yaniv, M., and Pontoglio, M. (1997) *J. Mol. Biol* **266,** 231–245.
25. Waltz, S. E., Gould, F. K., Air, E. L., McDowell, S. A., and Friezner Degen, S. J. (1996) *J. Biol. Chem.* **271,** 9024–9032.
26. Baráth, P., Albert-Fournier, B., Luciaková, K., and Nelson, B. D. (1999) *J. Biol. Chem.* **274,** 3378–3384.
27. Haecker, S. A., Muramatsu, T., Sensenbaugh, K. R., and Sanders, M. M. (1995) *Mol. Endocrinol.* **9,** 1113–1126.
28. Nagasawa, T., Takeda, T., Minemura, K., and DeGroot, L. J. (1997) *Mol. Cell. Endocrinol.* **130,** 153–165.
29. Yasuda, H., Mizuno, A., Tamaoki, T., and Morinaga, T. (1994) *Mol. Cell. Biol.* **14,** 1395–1401.
30. Kudo, S., and Fukuda, M. (1995) *J. Biol. Chem.* **270,** 13298–13302.
31. Lu, S., and Davies, P. J. A. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 4692–4697.
32. Siegfried, Z., Eden, S., Mendelsohn, M., Feng, X., Tsuberi, B.-Z., and Cedar, H. (1999) *Nat. Genet.* **22,** 203–206.
33. Proudfoot, N. J. (1986) *Nature* **322,** 562–565.
34. Bateman, E., and Paule, M. R. (1988) *Cell* **54,** 985–992.
35. Ayoubi, T. A., and Van-De-Ven, W. J. (1996) *FASEB J.* **10,** 453–460.
36. Schutze, M. P., Peterson, P. A., and Jackson, M. R. (1994) *EMBO J.* **13,** 1696–1705.
37. Rathjen, P. D., Toth, S., Willis, A., Heath, J. K., and Smith, A. G. (1990) *Cell* **21,** 1105–1114.
38. Slusher, L. B., Gillman, E. C., Martin, N. C., and Hopper, A. K. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 9789–9793.
39. Castillo, J. E., Martínez-Ansó, E., Malumbres, R., De Alava, E., García, C., Medina, J. F., and Prieto, J. (1999) *Cell Tissue Res.*, in press.
40. Al-Awqati, Q., Vijayakumar, S., Hikita, C., Chen, J., and Takito, J. (1998) *Am. J. Physiol.* **275,** F183–F190.
41. Prieto, J., Qian, C., García, N., Díez, J., and Medina, J. F. (1993) *Gastroenterology* **105,** 572–578.
42. Medina, J. F., Martínez-Ansó, E., Vázquez, J. J., and Prieto, J. (1997) *Hepatology* **25,** 12–17.
43. Prieto, J., García, N., Martí-Climent, J. M., Peñuelas, I., Richter, J. A., and Medina, J. F. (1999) *Gastroenterology* **117,** 167–172.