GENOMICS 26, 55–62 (1995)

Structural Organization of the Human Sorbitol Dehydrogenase Gene (SORD)

TAKESHI IWATA,* NICHOLAS C. POPESCU,† DRAZEN B. ZIMONJIC,† CHRISTINA KARLSSON,‡ JAN-OLOV HÖÖG,‡ GERALDO VACA,§ IGNACIO R. RODRIGUEZ¶ AND DEBORAH CARPER*,[1]

*Laboratory of Mechanisms of Ocular Diseases and ¶Laboratory of Retinal Cell and Molecular Biology, National Eye Institute, and †Laboratory of Biology, Division of Cancer Etiology, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892; ‡Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden; and §Instituto Mexicano del Seguro Social, Centro de Investigacion Biomedia de Occidente, Guadalajara, Jalisco, Mexico

Received June 29, 1994; revised November 22, 1994

The primary structure of human sorbitol dehydrogenase (SORD) was determined by cDNA and genomic cloning. The nucleotide sequence of the mRNA covers 2471 bp including an open reading frame that yields a protein of 356 amino acid residues. The gene structure of SORD spans approximately 30 kb divided into 9 exons and 8 introns. The gene was localized to chromosome 15q21.1 by *in situ* hybridization. Two transcription initiation sites were detected. Three Sp1 sites and a repetitive sequence $(CAAA)_5$ were observed in the 5′ noncoding region; no classical TATAA or CCAAT elements were found. The related alcohol dehydrogenases and ζ-crystallin have the same gene organization split by 8 introns, but no splice points coincide between SORD and these gene types. The deduced amino acid sequence of the SORD structure differs at a few positions from the directly determined protein sequence, suggesting allelic forms of the enzyme. High levels of SORD transcripts were observed in lens and kidney, as judged from Northern blot analysis.

INTRODUCTION

Sorbitol dehydrogenase (SORD) (EC 1.1.1.14) is a zinc-containing enzyme that catalyzes the conversion of sorbitol to fructose with NAD^+ as coenzyme. It is a member of the multigene family that includes alcohol dehydrogenases (Jörnvall *et al.*, 1984), ζ-crystallin (Borras *et al.*, 1989), and threonine dehydrogenase (Aronson *et al.*, 1989). SORD is involved in the metabolism of different polyols and is believed to cooperate with aldose reductase in osmotic regulation (Burg, 1988). The regulation through the polyol pathway is thought to affect the accumulation of sorbitol that is associated with diabetes mellitus and its complications, such as neuropathy (Gabbay, 1973), retinopathy (Robison *et al.*, 1983), and cataracts (Kinoshita, 1974). SORD is also thought to affect the cytosolic ratio of NADH/NAD^+ impacting on several metabolic pathways (Williamson *et al.*, 1993). The SORD activity in normal human lens has been observed to be higher than that in rat, rabbit, and calf lens (Jedziniak *et al.*, 1981). A deficiency in SORD activity has been linked to cataract formation in nondiabetics (Vaca *et al.*, 1982; Shin *et al.*, 1984).

The SORD enzyme has been purified and characterized from various species (Jeffery and Jörnvall, 1988; Maret and Auld, 1988; Wiesinger and Hamprecht, 1989). The nature of the zinc-binding site has been studied by computer modeling (Eklund *et al.*, 1985) and site-directed mutagenesis (Karlsson and Höög, 1993).

Recently cDNAs from rat (Karlsson *et al.*, 1991; Wen and Bekhor, 1993), the silkworm *Bombyx mori* (Niimi *et al.*, 1993), the yeast *Saccharomyces cerevisiae* (Sarthy *et al.*, 1993), and partial sequence of human SORD (Lee *et al.*, 1994) have been cloned. The cloning of the gene (gut B) from *Bacillus subtilis* (Ng *et al.*, 1992) has also been reported recently. A detailed study of the tissue distribution of SORD expression has been performed in rat (Estonius *et al.*, 1993). However, the characterization of the human SORD gene is novel and will be important for evaluating how the regulation of this gene is related to the formation of human cataract and complications connected to diabetes mellitus.

MATERIALS AND METHODS

Isolation and sequencing of human SORD cDNA. A partial rat cDNA clone (λ SDH2, Karlsson *et al.*, 1991) was used as a probe to screen a human liver cDNA library (Stratagene, La Jolla, CA) by the method of Young and Davis (1983). Isolated clones were directly sequenced in both directions by cycle sequencing (*f*mol DNA sequencing system, Promega, Madison, WI; dsDNA cycle sequencing system, GIBCO BRL, Gaithersburg, MD) or by the fluorescence autosequencing system (*Taq* DyeDeoxy Terminator cycle sequencing kit, 370A DNA sequencer, Applied Biosystems, Foster City, CA). The 5′ end

[1] To whom correspondence should be addressed at Room 232, Bldg. 6, 9000 Rockville Pike, Laboratory of Mechanisms of Ocular Diseases, National Eye Institute, National Institutes of Health, Bethesda, MD 20892. Telephone: (301) 496-2144. Fax: (301) 496-1759.

0888-7543/95 $6.00

sequence was obtained by using the 5′-RACE (rapid amplification of cDNA ends) method (5′-RACE system, GIBCO BRL; 5′-AmpliFINDER RACE kit, CLONTECH Laboratories, Palo Alto, CA). One microgram of poly(A) RNA from human liver and 10 pmol of antisense primer 1 (Fig. 1) were used for reverse transcription. The RACE products from the PCR reaction were subcloned into the pCR II vector (Invitrogen, San Diego, CA), amplified, and sequenced using the M13 forward and reverse primers (Invitrogen).

Northern blot analysis. Total RNA from normal human lens (NDRI, Philadelphia, PA) was isolated by the acid guanidinium isothiocyanate–phenol–chloroform extraction method (Chomczynski and Sacchi, 1987). Total RNAs from human brain, heart, kidney, liver, lung, placenta, retina, skeletal muscle, and testis were obtained from Clontech. Total RNA (10 μg) was separated on a 1% agarose gel run at 30 V for 12 h and blotted onto a nylon membrane (Boehringer Mannheim Co., Indianapolis, IN) for Northern analysis (Sambrook *et al.,* 1989) using human SORD probe B (Fig. 4). The hybridization bands were digitally recorded into a Macintosh Quadra950 computer (Apple Computer Inc., Cupertino, CA) by scanning (UC1200S, UMAX Data System Inc., Hsinchu, Taiwan) using PhotoShop software (Adobe Systems Inc., Mountain View, CA), and the intensity of each band was quantified using the NIH Image software (developed by W. Reisband, NIH, Bethesda, MD).

Isolation and sequence analysis of SORD genomic clones. A partial cDNA clone coding for human SORD (probe B, Fig. 4) was used to screen a human genomic phage library (Stratagene) and a human genomic cosmid library (Stratagene). Separately, a human genomic P1 library was screened using PCR with primers 2, 4 and primers 5, 6 (Fig. 1), respectively (service by Genome Systems Inc., St. Louis, MO). The exon/intron boundaries and intron sequences were determined by direct sequencing of isolated clones in both directions. The sequence assembly and alignment were performed with PileUp, Pretty, Gap (sequence analysis software package, GCG, Madison, WI), MacVector, AssemblyLIGN (International Biotechnologies Inc., New Haven, CT), and INHERIT (Applied Biosystems) softwares. The 5′ flanking sequence was analyzed with MacSignalScan software (developed by D. S. Prestridge, Los Alamos National Laboratory, Los Alamos, NM).

Primer extension analysis for determining the transcription start sites. The oligonucleotide primer 7 (Fig. 1) complementary to human SORD mRNA was used for primer extension (primer extension systems, Promega). Three to five micrograms of poly(A) RNA from brain, retina, liver, kidney, testis, and yeast were used.

Chromosome in situ hybridization. To identify the chromosomal locus, the phage 1 genomic clone (probe A, Fig. 4) was labeled by nick-translation with biotin-11–dUTP and used for fluorescence *in situ* hybridization (FISH). Digital image acquisition and analysis were performed as previously described (Popescu *et al.,* 1994; Zimonjic *et al.,* 1994).

RESULTS

The Complete Sequence of Human SORD mRNA

A human liver cDNA library was screened with a partial rat cDNA clone (λSDH2). Two cDNA clones (Phage-3 and -4) (Fig. 4) were isolated from a total of 3×10^5 plaques and were subsequently sequenced on both strands. The deduced amino acid sequence of the cDNA was compared with the sequence directly determined at the protein level (Karlsson *et al.,* 1989). The assembled nucleotide sequence of the two clones covered 80% of the coding region and a complete 3′-untranslated region (1311 bp). To obtain the 5′ end sequence of the mRNA, the 5′-RACE method was performed on liver poly(A) RNA. After 60 cycles of PCR, a RACE product was subcloned for sequencing. The additional 5′ end region from three separate clones completed the nucleotide sequence of the human SORD cDNA (2471 bp). The open reading frame encodes 356 amino acid residues (Fig. 1). We found five differences in the deduced amino acid sequence as compared to that previously reported (Karlsson *et al.,* 1989). One additional alanine (GCG) was observed at codon 1. Furthermore, an aspartate to asparagine difference at codon 58, a glutamine to methionine difference at codon 185, a serine to threonine difference at codon 280, and a threonine to isoleucine difference at codon 288 were found.

The human SORD nucleotide sequence was compared with published nucleotide sequence data for other species using the Gap (GCG) program. The positional identities between human and rat, silkworm, *Bacillus subtilis,* and *Saccharomyces cerevisiae* were 79, 49, 47, and 43%, respectively. The deduced amino acid sequence of human SORD was also compared for homology between sheep, rat, silkworm, *Bacillus subtilis, Saccharomyces cerevisiae,* and *Pichia stipitis* xylitol dehydrogenases using the PileUp (GCG) program (Fig. 2).

SORD Expression in Human Tissues

The level of SORD transcripts in 10 different human tissues was examined by Northern analysis (Fig. 3) using probe B (Fig. 4). A 2.5-kb transcript was observed, which is consistent with the cDNA sequence (2471 bp). Although eye tissue was limited, RNA adequate to show expression of SORD in lens was obtained. After normalization of the blot by hybridization to the 18S ribosomal DNA probe, the highest level of steady-state expression was observed in the lens and kidney (Fig. 3). Much less expression was seen in liver.

The Isolation and Characterization of Human SORD Gene Clones

Several positive genomic clones from different libraries were isolated either by hybridization with probe B in phage and cosmid libraries or by PCR screening of a P1 library using primers 2, 4 and primers 5, 6 (Fig. 1), respectively. The clones obtained overlapped with each other (Fig. 4) and covered the entire gene plus the 5′ flanking region. The human SORD gene spans approximately 30 kb and is composed of 9 exons. Exon 1 contains a translation initiation codon and 21 amino acids of the encoded sequence. Exon 9 contains 54 amino acids, a stop codon, and the 3′-untranslated region (1308 bp). The exon/intron boundaries and the distance between the exons (Table 1) were determined by sequencing, restriction mapping, and PCR. All of the splice junction sequences fit the donor–acceptor rule (Mount, 1982).

Chromosomal Localization by in Situ Hybridization

The efficiency of hybridization with probe A (Fig. 4) was high. Ninety-five of 100 randomly selected meta-

FIG. 1. Nucleotide sequence and encoded polypeptide of human SORD and 5′, 3′ untranslated sequences. The nucleotide sequence extends from the transcription initiation site through the translation initiation site and stop codon (***) to the poly(A) tail. The numbering to the right refers to nucleotides. The boldface sequences with italic numbers over them are the newly identified sequences. The first nucleotide of each exon is indicated by underlining and boldface. The arrows under the nucleotide sequence indicate the primer sequences used for screening the P1 genomic library, reverse transcription, and the 5′-RACE method.

phase chromosomes exhibited fluorescent signals on the long arm of a large acrocentric chromosome (Fig. 5). Furthermore, 80 chromosomal spreads exhibited a clear signal at a homologous site on both chromatids of two acrocentric chromosomes. Symmetrical fluorescent spots were not observed on any other chromosomes. Twenty-five chromosomal spreads with symmetrical fluorescent doublets were reexamined after G-banding by trypsin treatment. Enlarged digital images of the labeled and banded chromosomes were measured and compared on screen. Based on localization of the doublets from 35 individual chromosomes on a 400-band idiogram, the locus of the gene was assigned to chromosome 15q21.1, the most likely location of the gene.

Two Transcription Initiation Sites

The primer extension analysis determined the transcriptional initiation sites in brain, retina, liver, kidney, and testis (Fig. 6). The two transcriptional initiation signals were found 16 and 89 bp upstream of the translation initiation site in all tissues analyzed. How-

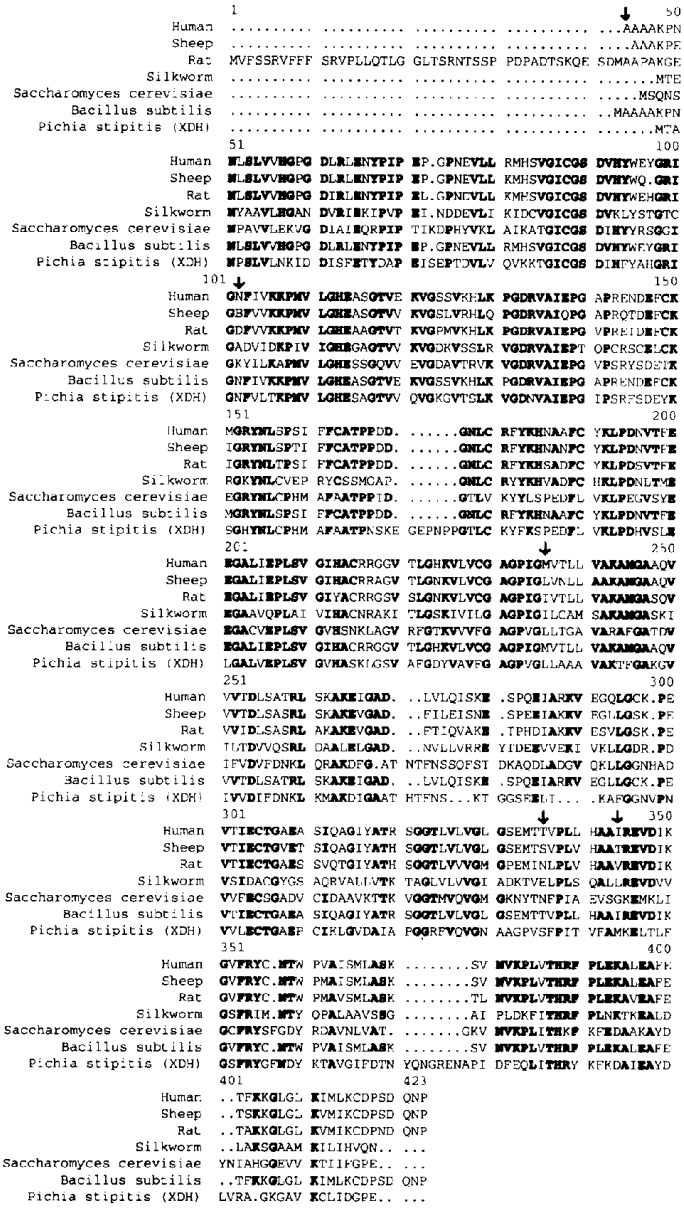

FIG. 2. Amino acid sequence comparison of human SORD with sheep, rat, silkworm (*Bombyx mori*), *S. cerevisiae*, *B. subtilis*, and *P. stipitis* xylitol dehydrogenase. Amino acids conserved in more than five species are in boldface. The corrected deduced amino acid sequences for the human SORD mRNA are indicated by arrows.

ever, the majority of the transcripts started 16 bp upstream for all tissues except liver. No diversity was observed in the translational initiation codon.

Analysis of the Promoter Region of the Putative Functional Human SORD Gene

The sequence of the 5′ flanking region was searched for *cis*-acting elements important for the transcriptional initiation of the SORD gene using the MacSignalScan transcription factor database computer program (Fig. 7). No obvious TATAA or CCAAT boxes were found in this region, but three Sp1 sites and a CACCC box (Yu *et al.*, 1991) were found. Upstream of these sites a $(CAAA)_5$ repeat sequence was observed.

DISCUSSION

The complete cDNA sequence and the genomic organization of human SORD were determined. The first exon contains 89 bp (liver) of 5′-untranslated sequence, and exon 9 contains 1311 bp of 3′-untranslated sequence. The 3′-untranslated sequence encompasses 53% of the total cDNA sequence. This long stretch of 3′-untranslated sequence could be involved in mRNA stability and/or translational regulation. Unlike in rat, an upstream ATG codon possibly generating a pre-SORD protein (Wen and Bekhor, 1993) was not observed in human SORD. The homologous human alcohol dehydrogenase genes and the human ζ-crystallin gene are also arranged into 9 exons and 8 introns, but none of the splicing points coincides exactly with the splicing points of the SORD gene (Duester *et al.*, 1986; von Bahr-Lindstrom *et al.*, 1991; Gonzalez *et al.*, 1994).

One additional codon and four different codons were found when the deduced translated amino acid sequence was compared to the directly determined amino acid sequence (Karlsson *et al.*, 1989), a finding in agreement with a recent report by Lee *et al.*, (1994). The additional alanine at position 1 gives higher ho-

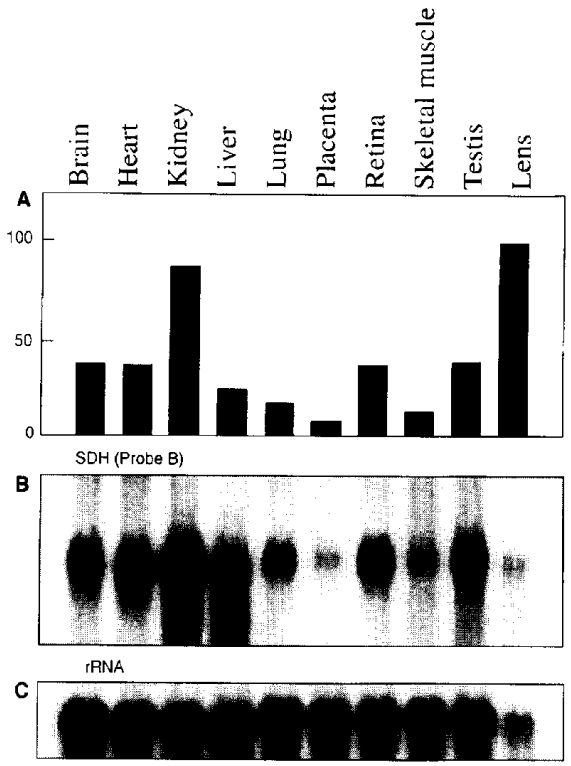

FIG. 3. Expression of the human SORD gene in various human tissues. **(A)** The relative amount of SORD expression in various tissues is indicated by the ratio of SORD mRNA/18S ribosomal RNA setting lens at 100%. **(B)** Northern blot of the human tissue panel using probe B (Fig. 4). **(C)** Northern blot for standardization using the 18S ribosomal DNA probe.

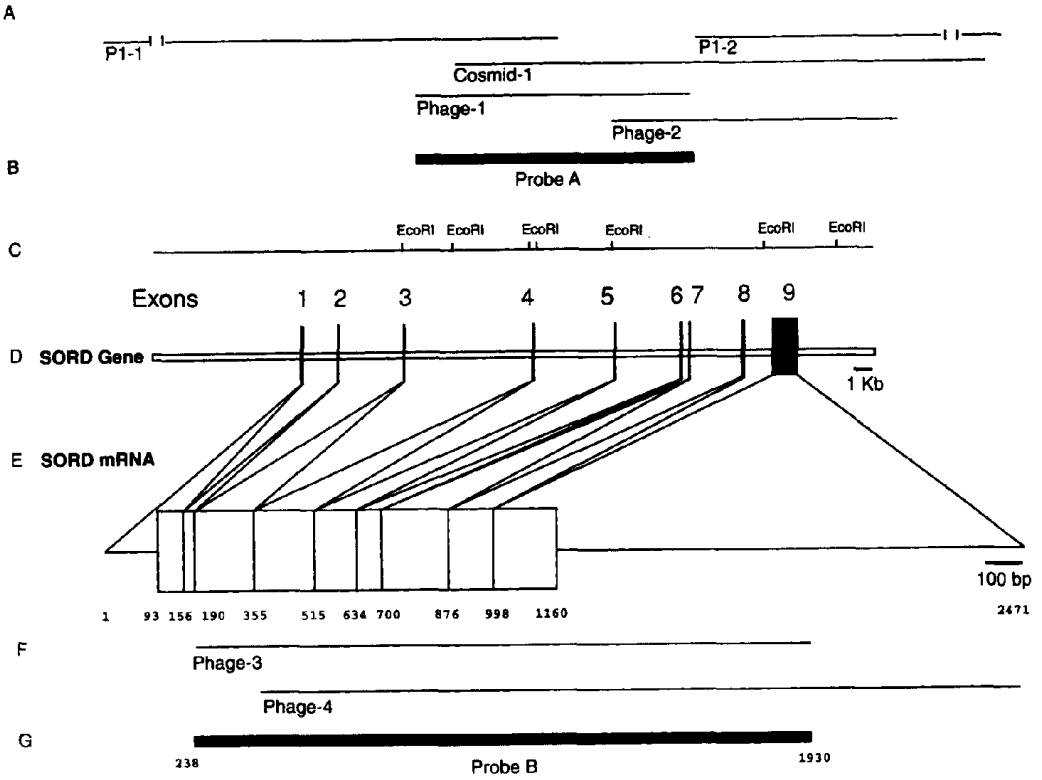

FIG. 4. Structure of the human SORD gene and mRNA. **(A)** Diagram of the screened phage, cosmid, and P1 genomic clones. **(B)** The Phage-1 insert was used as probe A for chromosomal sublocalization by *in situ* hybridization. **(C)** Enzyme restriction sites (*Eco*RI) of the SORD gene. **(D)** Structure of the SORD gene. Exons are shown in closed boxes. **(E)** The SORD mRNA is shown at a scale of 13:1 in relation to that of the SORD gene. Open boxes indicate open reading frames. **(F)** Schematic diagram of screened cDNA clones. **(G)** The Phage-3 insert was used as probe B for Northern blot analysis.

mology to the rat N-terminus. The other four amino acids do not strongly affect the previous homology analysis performed with other species (Karlsson *et al.*, 1991) nor do they have any likely impact on the proposed function of the enzyme. These differences can provide evidence for allelic variation at the SORD gene locus.

The Northern blot analysis clearly demonstrates that the highest expression of human SORD mRNA is in lens and kidney with much lower expression in liver. This result is similar to a study on rat tissue (Wen and Bekhor, 1993) where SORD mRNA was lower in liver than in lens, but differs from another rat study (Estonius *et al.*, 1993) where liver SORD transcripts were high compared to many other tissues. SORD has been used by many investigators as a marker enzyme for liver since SORD activity is high in this tissue. Our results suggest that although the activity may be higher, the level of steady-state mRNA in human liver is not appreciably high, when compared to lens, kidney, retina, or brain. The high expression of SORD observed

TABLE 1

Exon–Intron Organization of the SORD Gene

Exon		Intron		Exon
			AGTCC	1
1	GCCTG | GTAAG·····	(~2, 500 bp)	·····TTTAG | GAGAA	2
2	AAATG | GTAAG·····	(~3, 000 bp)	·····ATCAG | AGGTC	3
3	ACCAG | GTCAG·····	(~11,000 bp)	·····TCCAG | GTGAT	4
4	TACAA | GTTAG·····	(4,045 bp)	·····CTCAG | GCTTC	5
5	AGCTG | GTAAG·····	(2,692 bp)	·····TTTAG | GGCCA	6
6	GACTG | GTAAG·····	(628 bp)	·····TTCAG | ATCTG	7
7	TCTAC | GTGAG·····	(~3,200 bp)	·····TACAG | GCCAC	8
8	AACAC | GTGAG·····	(~1,400 bp)	·····TACAG | GTGGC	9
9	ATCTG |			

Note. Over 300 bp of 5′ and 3′ flanking region surrounding each exon was sequenced. Ninety percent of the intron sequence from exon 3 to exon 7 was determined. Five nucleotides on both sides of the boundaries are indicated. The numbering on both sides indicates the exons.

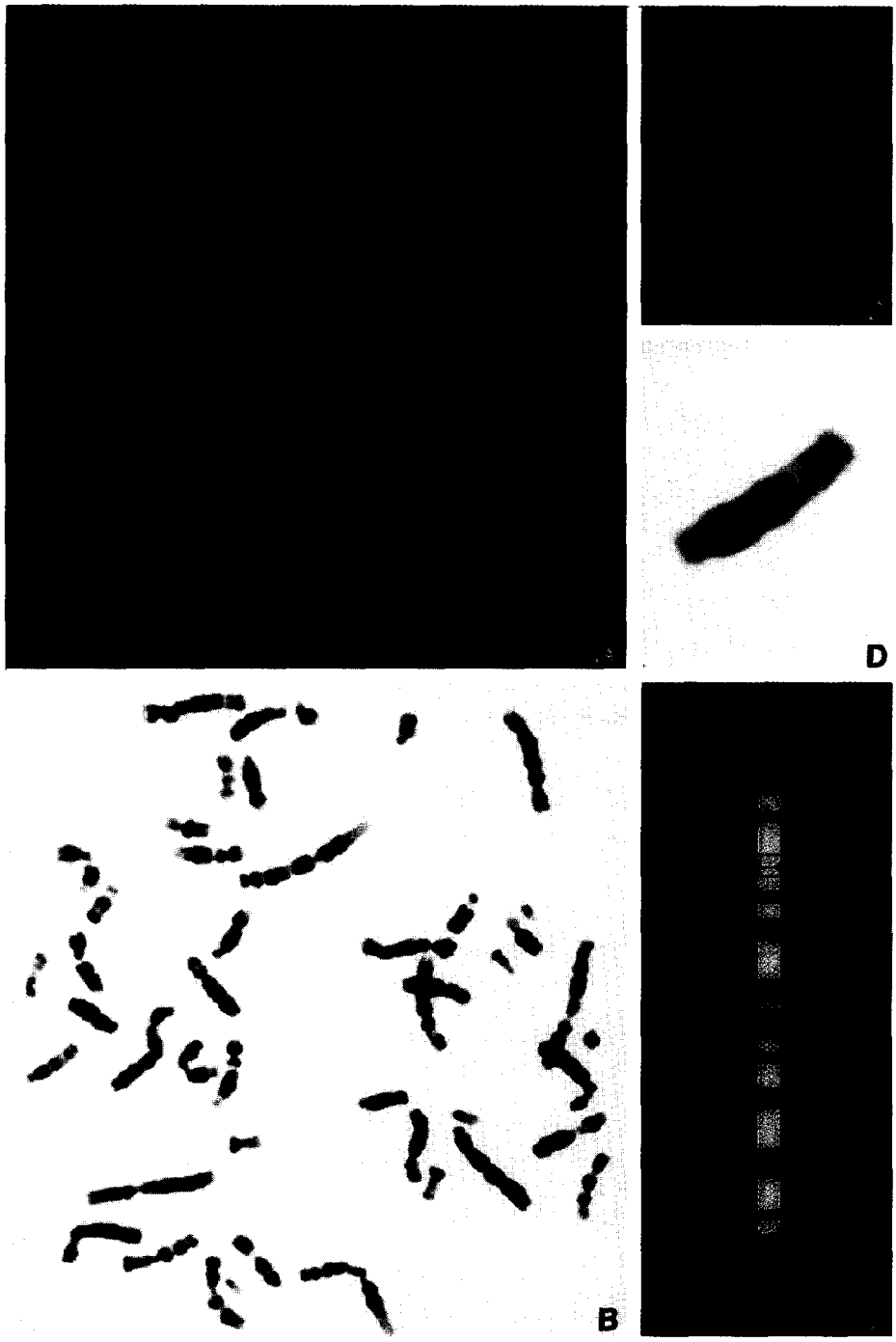

FIG. 5. Localization by fluorescence *in situ* hybridization (FISH) of the SORD gene to human chromosome 15q21.1. **(A)** Metaphase chromosomes after FISH with biotinylated genomic DNA probe A (Fig. 4), exhibiting a fluorescent signal on both chromatids of the medium size acrocentric chromosome's long arm. **(B)** The same G-banded metaphase permits the identification of the labeled chromosome as chromosome 15 with localization of the signal at band 15q21.1. **(C, D)** Enlarged chromosome 15 from **A** and **B**. **(E)** The ideogram of human chromosome 15, indicating the position of the gene.

in the human lens is of great interest in view of previous reports showing lowered SORD activity in red blood cells of a family where some members have congenital cataract (Vaca *et al.*, 1982).

Human SORD was previously localized to chromosome 15 by an indirect method using an enzyme assay on human–hamster somatic hybrid cell lines (Donald *et al.*, 1980) and recently localized to 15q15 by a standard fluorescence *in situ* hybridization technique (Lee *et al.*, 1994). We have independently performed a fluorescence *in situ* hybridization (Fig. 5) and digital-image analysis to localize the SORD gene precisely. Current localization to chromosome 15q21.1 is one band more distal to the centromere than the previous assignment.

Linkage studies of families with inherited cataract have been reported (Armitage *et al.*, 1993; Bateman *et*

al., 1993; Kojis *et al.*, 1993). Recently, a correlation was found between a γ-crystallin abnormality and a family with cataract (Brakenhoff *et al.*, 1994). Individuals in families with inherited SORD deficiency are more likely to develop cataract (Vaca *et al.*, 1982; Shin *et al.*, 1984). The localization of the human SORD gene will provide an additional locus for linkage studies of families with inherited cataract and may provide some insight into the cause of cataract in SORD-deficient individuals.

At the promoter region of human SORD, no obvious TATA or CCAAT box was found. Two transcription initiation sites were observed with the shorter transcript predominating in all tissues examined except liver, where the longer transcript was in higher abundance. These different transcription initiation sites do not affect the translational initiation site (ATG codon). At the 5′ flanking region of this gene three Sp1 sites and a CACCC box, which is reported to bind the Sp1 transcriptional factor protein (Yu *et al.*, 1991), were found. The sequential motif of this region resembles the promoter region of duck lactate dehydrogenase B/ϵ-crystallin (Kraft *et al.*, 1993), which is highly expressed in heart as an enzyme and in lens as a crystallin. The promoter of this enzyme/crystallin gene consists of three Sp1 sites and two transcriptional initiation sites and lacks a TATAA box. The CAAA repeat, which was found upstream of the Sp1 and CACCC box of human SORD, was searched by computer (INHERET). Several genes that contained identical sequences at the 5′ flanking region were found. A report on negative regulation of the H-2K^b class I antigene by a transcriptional factor that binds to the CAA(A) repeat (Ozawa *et al.*, 1993) provides an interesting possibility for further study of the regulation of human SORD.

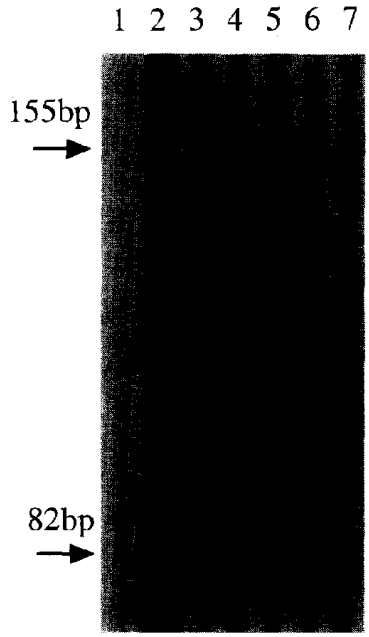

FIG. 6. Determination of the transcriptional initiation site of human SORD by primer extension analysis. The primer extension was performed by reverse transcribing 3–5 μg of poly(A) RNA using antisense primer 7 (Fig. 1) as described under Materials and Methods. Lane **1**, yeast negative control; **2**, sequence T ladder; **3**, brain; **4**, retina; **5**, liver; **6**, kidney; **7**, testis. The arrows and numbers indicate two major bands and sizes.

```
-288
AAAATACGTGGTGGTGGGCGCCTGTAGTCCCAGCTATTCGGGAGGCTGAGGCAGGAGAA

-238
TGGCTTGAACCCGGGAGGCAGAGCTTGCAGTGAGCCGGATCGCACACTGTACTCCAGCC

-188
TCGGCGACAAACAAACAAACAAACAAAAAACAGCGCGCTGCAGATGGAGCCAGCACCC
                                                    CACCC-BOX
-138
GGGCTTCTCGCTCCCAGCCCCCTCCCGAAGCCCGCCCCGCCTCCGGGTCTGGAGGAGG
                                    Sp1
-88
CTGGGAAGCCCCGCCCCTCCACGCTAGCGCCGCCCAGGCTGGCACAAAGGAGGAAGCCT
          Sp1
-38                                                  ┌──
AGTCCCGCCCCTGCGTGCGGCGCTTCTCCCAGGCCCCACCTTCCATCCAGTGCCCTGGA
    Sp1                                          +1(A)
+12
CCCTCGGCTGGGTAGCGCCACCAGAGCGACCAAACGTCCCGCGCCTTCCAGGCCGCACT

+62 ┌──
CCAGAGCCAAAAGAGCTCCATGGCGGCGGCGGCCAAGCCCAACAACCTTTCCCTGGTGG
   +66(B)                          Exon 1
```

FIG. 7. Nucleotide sequence of the 5′-flanking region of the human SORD gene. The numbers to the left refer to the position of each nucleotide from the cap site. The different initiation sites for **(A)** liver and **(B)** brain, retina, kidney, and testis are shown by arrows. The Sp1 sites and CACCC box are indicated by open boxes. The translational initiation site (ATG) is in boldface. The unique $(CAAA)_5$ repeat is indicated by underline.

We have demonstrated high expression of SORD in human lens compared to other tissues. A previous study reported higher enzyme activity in human lens compared to other species (Jedziniak *et al.*, 1981). These data suggest that SORD may play an important role in the human lens and that dysfunction of this enzyme may lead to alterations in the polyol pathway.

ACKNOWLEDGMENTS

The authors greatly appreciate Dr. Peter Frederikse for discussion on promoter analysis and Mr. Pedro J. Miret for helping with the initial contact with Mexico. This publication includes data presented at the Association for Research in Vision and Ophthalmology meeting in Sarasota, Florida (Iwata *et al.*, 1993). The cDNA sequence and the total 15 kb of the gene sequence are currently available from GenBank, the EMBL Data Bank, and the DNA Data Bank of Japan under Accession Nos. L29008, L29249, L29250, L29251, L29252, L29253, and L29254.

REFERENCES

Armitage, M. M., Kivlin, J. D., and Ferrell, R. E. (1993). Exclusion of primary candidate genes in a Cerulean cataract family. *Am. J. Hum. Genet.* **53:** 1684.

Aronson, B. D., Somerville, R. L., Epperly, B. R., and Dekker, E. E. (1989). The primary structure of *Escherichia coli* L-threonine dehydrogenase. *J. Biol. Chem.* **264:** 5226–5232.

Bateman, J. B., Kojis, T. L., Heinzmann, C., Flodman, P., Mullen, M. L., Spence, M. A., and Sparkes, R. S. (1993). Linkage relation-

ships of hereditary cataracts to human chromosomal regions 1q21–q25 and 16q22.1. *Am. J. Hum. Genet.* **53:** 1686.

Borras, T., Persson, B., and Jörnvall, H. (1989). Eye lens ζ-crystallin relationships to the family of "long-chain" alcohol/polyol dehydrogenases. *Biochemistry* **28:** 6133–6139.

Brakenhoff, R. H., Henskens, H. A. M., van Rossum, M. W. P. C., Lubsen, N. H., and Schoenmakers, J. G. G. (1994). Activation of the gammaE-crystallin psuedogene in the human hereditary Coppock-like cataract. *Hum. Mol. Genet.* **3:** 279–283.

Burg, M. B. (1988). Role of aldose reductase and sorbitol in maintaining the medullary intracellular milieu. *Kidney Int.* **33:** 635–641.

Chomczynski, P., and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. *Anal. Biochem.* **162:** 156–159.

Donald, L. J., Wang, H. S., and Hamerton, J. L. (1980). Assignment of the sorbitol dehydrogenase locus to human chromosome 15 pter leads to q21. *Biochem. Genet.* **18:** 425–431.

Duester, G., Smith, M., Bilanchone, V., and Hatfield, G. W. (1986). Molecular analysis of the human class I alcohol dehydrogenase. *J. Biol. Chem.* **261:** 2027–2033.

Eklund, H., Horjales, E., Jörnvall, H., Branden, C., and Jeffery, J. (1985). Molecular aspect of functional differences between alcohol and sorbitol dehydrogenases. *Biochemistry* **24:** 805–812.

Estonius, M., Danielsson, O., Karlsson, C., Persson, H., Jörnvall, H., and Höög, J. (1993). Distribution of alcohol and sorbitol dehydrogenase. *Eur. J. Biol.* **215:** 497–503.

Gabbay, K. H. (1973). The sorbitol pathway and the complications of diabetes. *N. Engl. J. Med.* **288:** 831–836.

Gonzalez, P., Rao, P. V., and Zigler, J. S., Jr. (1994). Organization of the human ζ-crystallin/quinone reductase gene (CRYZ). *Genomics* **21:** 317–324.

Iwata, T., Höög, J. O., Reddy, V. N., and Carper, D. (1993). Cloning of the human sorbitol dehydrogenase gene. *Invest. Ophthal. Vis. Sci. Suppl.* **34:** 712.

Jedziniak, J. A., Chylack, L. T., Jr., Cheng, H.-M. M., Grillis, K., Kalustian, A. A., and Tung, W. H. (1981). The sorbitol pathway in the human lens. *Invest. Ophthalmol. Vis. Sci.* **20:** 314–326.

Jeffery, J., and Jörnvall, H. (1988). Sorbitol dehydrogenase *Adv. Enzymol.* **61:** 47–106.

Jörnvall, H., von Bahr-Lindström, H., and Jeffery, J. (1984). Extensive variations and basic features in the alcohol dehydrogenase–sorbitol dehydrogenase family. *Eur. J. Biochem.* **140:** 17–23.

Karlsson, C., Maret, W., Auld, D. S., Höög, J. O., and Jörnvall, H. (1989). Variability within mammalian sorbitol dehydrogenases. *Eur. J. Biochem.* **186:** 543–550.

Karlsson, C., Jörnvall, H., and Höög, J. O. (1991). Sorbitol dehydrogenase: cDNA coding for the rat enzyme. *Eur. J. Biochem.* **198:** 761–765.

Karlsson, C., and Höög, J. O. (1993). Zinc coordination in mammalian sorbitol dehydrogenase. *Eur. J. Biochem.* **216:** 103–107.

Kinoshita, J. H. (1974). Mechanisms initiating cataract formation. Proctor Lecture. *Invest. Ophthalmol.* **13:** 713–724.

Kojis, T. L., Heinzmann, C., Flodman, P., Mullen, M. L., Sparkes, R. S., Spence, M. A., and Bateman, J. B. (1993). Linkage analysis of crystallin genes and hereditary cataracts. *Am. J. Hum. Genet.* **53:** 1699.

Kraft, H. J., Hendriks, W., de Jong, W. W., Lubsen, N. H., and Schoenmakers, J. G. G. (1993). Duck lactate dehydrogenase B/ε-crystallin gene: Lens recruitment of a GC-promoter. *J. Mol. Biol.* **229:** 849–859.

Lee, F. K., Cheung, M. C., and Chung, S. (1994). The human sorbitol dehydrogenase gene: cDNA cloning, sequence determination, and mapping by fluorescence *in situ* hybridization. *Genomics* **21:** 354–358.

Maret, W., and Auld, D. S. (1988). Purification and characterization of human liver sorbitol dehydrogenase. *Biochemistry* **27:** 1622–1628.

Mount, S. M. (1982). A catalogue of splice junction sequences. *Nucleic Acids Res.* **10:** 459–472.

Niimi, T., Yamashita, O., and Yaginuma, T. (1993). A cold-inducible *Bombyx* gene encoding a protein similar to mammalian sorbitol dehydrogenase. *Eur. J. Biochem.* **213:** 1125–1131.

Ng, K., Ye, R., Wu, X., and Wong, S. (1992). Sorbitol dehydrogenase from *Bacillus subtillis*. *J. Biol. Chem.* **267:** 24989–24994.

Ozawa, K., Hagiwara, H., Tang, X., Saka, F., Kitabayashi, I., Shiroki, K., Fujinaga, K., Israël, A., Gachelin, G., and Yokoyama, K. (1993). Negative regulation of the gene H-2K^b class I antigen by adenovirus 12-E1A is mediated by a CAA repeated element. *J. Biol. Chem.* **268:** 27258–27268.

Popescu, N., Zimonjic, D., Hatch, C., and Bonner, W. (1994). Chromosomal mapping of the human histone H2A.Z gene to 4q24 by *in situ* fluorescence hybridization. *Genomics* **20:** 333–335.

Robison, W. G., Jr., Kador, P. F., and Kinoshita, J. H. (1983). Retinal capillaries: Basement membrane thickening by galactosemia prevented with aldose reductase inhibitor. *Science* **221:** 1177–1179.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual," 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sarthy, A. V., Schopp, C., and Idler, K. B. (1993). Cloning and sequence determination of the encoding sorbitol dehydrogenase from *Saccharomyces cerevisiae*. *Gene* **140:** 121–126.

Shin, Y. S., Rieth, M., and Endres, W. (1984). Sorbitol dehydrogenase deficiency in a family with congenital cataracts. *J. Inher. Metab. Dis.* **7:** 151–152.

Vaca, G., Ibarra, B., Bracamontes, M., Garcia-Cruz, D., Sanchez-Corona, J., Medina, C., Wunsch, C., Gonzales-Quiroga, G., and Cantu, J. M. (1982). Red blood cell sorbitol dehydrogenase deficiency in a family with cataract. *Hum. Genet.* **61:** 338–341.

Von Bahr-Lindstrom, H., Jörnvall, H., and Höög, J. O. (1991). Cloning and characterization of the human ADH4 gene. *Gene* **22:** 269–274.

Wen, Y., and Bekhor, I. (1993). Full-length cDNA sequencing reveals a mRNA coding for a protein containing an additional 42 amino acids at the N-terminal end. *Eur. J. Biochem.* **217:** 83–87.

Wiesinger, H., and Hamprecht, B. (1989). Purification and characterization of sorbitol dehydrogenase from bovine brain. *J. Neurochem.* **52:** 342–348.

Williamson, J. R., Chang, K., Frangos, M., Hasan, K. S., Ido, Y., Kawamura, T., Nyengaard, J. R., van den Enden, M., Kilo, C., and Tilton, R. G. (1993). Hyperglycemia pseudohypoxia and diabetic complications. *Diabetes* **42:** 801–813.

Young, R. A., and Davis, R. W. (1983). Efficient isolation of genes by using antibody probes. *Proc. Natl. Acad. Sci. USA* **80:** 1194–1198.

Yu, C.-Y., Motamed, K., Chen, J., Bailey, A. and Shen, C.-K. J. (1991). The CACC box upstream from human embryonic ε-globin gene binds Sp1 and is a functional promoter element *in vitro* and *in vivo*. *J. Biol. Chem.* **266:** 8907–8915.

Zimonjic, D. B., Popescu, N. C., Matsui, T., Ito, M., and Chihara, K. (1994). Localization of the human cholecystokinin-B/gastrin receptor gene to chromosome 11p15.4–15.5 by fluorescence *in situ* hybridization. *Cytogenet. Cell Genet.* **65:** 184–185.