A human alcohol dehydrogenase gene (ADH6) encoding an additional class of isozyme

MICHIO YASUNAMI, CHENG-SHENG CHEN, AND AKIRA YOSHIDA*

Beckman Research Institute of the City of Hope, 1450 East Duarte Road, Duarte, CA 91010

Communicated by Eugene Roberts, May 29, 1991 (received for review March 23, 1991)

ABSTRACT The human alcohol dehydrogenase (ADH; alcohol:NAD⁺ oxidoreductase, EC 1.1.1.1) gene family consists of five known loci (ADH1-ADH5), which have been mapped close together on chromosome 4 (4q21-25). ADH isozymes encoded by these genes are grouped in three distinct classes in terms of their enzymological properties. A moderate structural similarity is observed between the members of different classes. We isolated an additional member of the ADH gene family by means of crosshybridization with the ADH2 (class I) cDNA probe. cDNA clones corresponding to this gene were derived from PCR-amplified libraries as well. The coding sequence of a 368-amino-acid-long open reading frame was interrupted by introns into eight exons and spanned approximately 17 kilobases on the genome. The gene contains a glucocorticoid response element at the 5' region. The transcript was detected in the stomach and liver. The deduced amino acid sequence of the open reading frame showed about 60% positional identity with known human ADHs. This extent of homology is comparable to interclass similarity in the human ADH family. Thus, the newly identified gene, which is designated ADH6, governs the synthesis of an enzyme that belongs to another class of ADHs presumably with a distinct physiological role.

Human alcohol dehydrogenases (ADHs; alcohol:NAD⁺ oxidoreductase, EC 1.1.1.1) are a family of enzymes that catalyze oxidative conversion of various alcohols to the corresponding aldehydes in a reversible fashion. Five kinds of subunits (α , β , γ , π , and χ) governed by five nonallelic genes, ADH1, ADH2, ADH3, ADH4, and ADH5, respectively, have been identified. These genes were mapped close together on chromosome 4q21–25 (for review, see ref. 1). The expressions of these ADH genes are under the control of tissue-specific and developmental stage-specific mechanisms (for review, see ref. 2).

ADH isozymes of humans and other mammals are grouped in three distinctive "classes" based on their enzymatic characteristics. The isozymes, which exhibit high catalytic activity for oxidation of short-chain aliphatic alcohols, such as ethanol, and are strongly inactivated by pyrazole, are grouped as class I ADH (3). Another type (class II) of ADH isozymes exhibits a high activity for oxidation of long-chain aliphatic alcohols and aromatic alcohols and is less sensitive to pyrazole (4). ADH isozymes, which have virtually no activity for ethanol oxidation but exhibit high activity for oxidation of long-chain alcohols and are completely insensitive to pyrazole, are classified as class III ADH (5). Members within a class show a considerable structural similarity (>85% positional identity) in their amino acid sequences even in distal mammalian species, while a lesser similarity ($\approx 60\%$) exists between two different classes within the same species (6).

It has been reported that human stomach (7, 8), rat stomach (9), and baboon stomach and eye (10) contain other ADH isozymes that exhibit class II enzymatic characteristics but

differ from the liver class II ADH. The structure of the rat stomach ADH was substantially different from that of known rat class I, II, and III ADHs, and thus it was proposed that rat stomach contains another class (class IV) of ADH isozymes (11). The present study demonstrates the existence of an additional human ADH gene, designated ADH6, which encodes an additional class of ADHs.[†]

MATERIALS AND METHODS

Construction of Genomic Library and Isolation and Characterization of Genomic Clones. Human genomic DNA libraries were constructed either with cosmid vector pWE15 or with phage vector λ DASH (12). The cosmid library [\approx 40kilobase (kb) inserts] was screened with a full-length *ADH2* cDNA probe pUCADH β 14 (13) by the colony hybridization method. The λ DASH library (\approx 15-kb inserts) was screened by the plaque hybridization method with a 0.8-kb-long *Xba* I fragment derived from a cosmid clone cosADH98, which was obtained from the cosmid library (Fig. 1). Restriction mapping analysis, subcloning of the inserts into pBluescript II (Stratagene), and nucleotide sequencing were performed by standard procedures (14, 15).

Construction of cDNA Library and Isolation and Characterization of cDNA Clones. $Poly(A)^+$ RNA was prepared from a Japanese adult male liver by a standard method (14).

An amplified cDNA library corresponding to the 3' region of the transcript was constructed as described by Frohman *et al.* (16). Briefly, 1 μ g of the poly(A)⁺ RNA was reversetranscribed using (dT)₁₅ adaptor primer [20-mer adaptor GGACTCGAGCTCTAGAAGTT conjugated with poly-(T)₁₅]. Unincorporated primer was removed by polyacrylamide gel electrophoresis. One-thousandth of the (-)-strand 3' cDNA thus produced was used for PCR amplification with a specific 3' primer (20-mer ATGGTACCAGCAGGTT-TACC derived from the sequence of cosADH96 clone) and the adaptor primer described above.

For generation of single (-)-strand 5' cDNA, 1 μ g of the poly(A)⁺ RNA was reverse-transcribed using a specific reverse transcriptase primer (20-mer CTCTAGAGGAGC-GACTGCAT, derived from the sequence of cosADH96 clone). To improve annealing efficiency, the original procedures for construction of a 5' cDNA library (16) were modified as follows. The single (-)-strand 5' cDNA was tailed with a deoxycytidine homopolymer, instead of a deoxyadenosine homopolymer, by terminal deoxynucleotide transferase (Pharmacia) (17). pBluescript KS+ vector cleaved by Sst II was tailed with a deoxyguanosine homopolymer and used as a primer. The tailed cDNA and the primer were annealed and subjected to the Klenow polymerase reaction. One-twentieth of the reaction product thus produced was used for PCR amplification with a pair of

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: nt, nucleotide.

^{*}To whom reprint requests should be addressed.

[†]The sequence reported in this paper has been deposited in the GenBank data base (accession no. M68895).



primers—i.e., a specific 5' primer (21-mer GCTGGTAC-CATCAGACATCAG, derived from the sequence of cosADH96 clone) and M13 universal adaptor primer (17-mer GTAAAACGACGGCCAGT).

Biochemistry: Yasunami et al.

The products of the PCR were digested with Sst I and Kpn I and subcloned in pBluescript II KS+ (for 5' product) or in pBluescript II SK+ (for 3' product) (Stratagene). Randomly picked clones from the transformed *Escherichia coli* were subjected to analysis by restriction mapping and nucleotide sequencing.

FIG. 1. Restriction map of the *ADH6* gene. Cosmid clone cos-ADH96 and seven phage clones λ ADH61, -161, -31, -101, -52, -62, and -171 are shown by horizontal lines. Restriction cleavage sites are indicated by vertical bars. Exons are boxed. Coding regions are shaded, whereas 5' and 3' untranslated regions (UT) found in the cDNA clones are in open boxes. A 0.8-kb Xba I fragment used for screening the phage library is also indicated.

A full-length cDNA was constructed by conjugating the 5' cDNA clone A8 and the 3' cDNA clone 1.1, which overlapped at the Kpn I site, and cloned in pBluescript SK + (14). The procedures are schematically outlined in Fig. 2. The nucleotide sequence of the full-length cDNA was established by analysis of both strands.

Detection of mRNA. Total cellular RNA was prepared from a Japanese adult male liver and from a normal part of a surgically removed stomach of a Japanese adult female with adenocarcinoma. RNA preparations with serial 2-fold dilu-



FIG. 2. Strategy for cDNA cloning. Partial restriction map of the full-length cDNA clone is also shown. The structures of reverse transcriptase (RT) primer, specific 5' primer, 3' primer, 5' adaptor primer, and 3' adaptor primer are described in *Materials and Methods*.

tions (4–0.5 μ g per slot) were blotted onto a nitrocellulose filter with a slot blotter (Schleicher & Schuell). The filter was hybridized with the ADH6-specific oligonucleotide probe ACGTCCAGATTTCCAATGGC (16). The filter was deprobed and rehybridized with the ADH2-specific probe GT-CATCTGTGCGACAGATTCC.

One microgram of poly(A) RNA prepared from the liver or stomach was used for the PCR amplification by using the two sets of primers-i.e., the ADH6-specific ATGGTACCAG-CAGGTTTACC (5' primer) and ACGTCCAGATTTCCAAT-GGC (3' primer), which were expected to produce a 448base-pair (bp) product from the ADH6 mRNA, and the ADH2-specific GGAATCTGTCACACAGATGAC (5) primer) and AGCAGTAGTTGCTCTCCGGG (3' primer), which were expected to produce a 205-bp product from the ADH2 mRNA. One-twentieth of the amplification product was electrophoresed in an agarose gel (1.5%) and blotted onto a nitrocellulose filter. The filter was first hybridized with the ADH6-specific oligonucleotide CAGGATCATTGGAGTG-GATG (derived from the coding sequence of ADH6 shown in Fig. 2), and, after deprobing, the filter was rehybridized with the ADH2-specific probe GGGTCACCAGGTTGCCACTA.

RESULTS

Isolation of Genomic Clones. A genomic clone (cosADH96; insert size, 38 kb) was isolated from the cosmid genomic library by cross-hybridization with the ADH2 cDNA probe (13). Only the 0.8-kb fragment generated by *Xba* I digestion,

															t	tgat	aaac	tact	aga
gaaaatctgacaccatccttatctgtcaacaaggggcctacattcagaaacctgaaagaattactcaagtttagagac																			
aaatatgaccatttttgtattttgaaaacaacagaaaagaaatacttt <u>tgtacactctgttag</u> aaattttaagtttgg																			
acatttaaaagtccaaatttaaaaactcaaaaaaatggataataagagggacctgtttgattaagggagaaaaaaatag																			
tttgcattttcaccttttggctctttcactgagatgagcctatttcagattacacttaggaacttccatcaagcacgg																			
-92 gagageetaettttegtttaataattaecagaetaeagaaaggteggaeeageettetgatetACAGTCGCCTGTGT																			
-1 ACCTTTGACTTTCTACAG <u>TGA</u> AAGTTGCTACAGGATCTCCCTTTCTCAATAAATTCATCTGCGGTGGAGAAAATCAGC																			
l ATG Met	AGT Ser	ACT Thr	ACA Thr	GGC Gly	18 CAA Gln	gtac	gtgc	ag int	ron	tatt 1	ctat	ttcc	cag	19 GTC Val	ATC Ile	AGA Arg	TGC Cys	AAA Lys	GCA Ala
GCC Ala	ATA Ile	CTC Leu	TGG Trp	AAG Lys	CCT Pro	GGT Gly	GCA Ala	CCA Pro	TTT Phe	TCT Ser	ATT Ile	GAA Glu	GAG Glu	GTA Val	GAA Glu	GTG Val	GCC Ala	CCA Pro	CCA Pro
							120	ָ								121			
AAG Lys	GCA Ala	AAG Lys	GAG Glu	Val	Arg	Ile	AAG Lys	gtaa	aaaa	int	ron	2 2	CLLC	cata	cag	Val	Val	Ala	Thr
G GA Gly	CTG Leu	TGT Cys	GGT Gly	ACA Thr	GAG Glu	ATG Met	AAA Lys	GTG Val	TTG Leu	GGG Gly	AGT Ser	AAA Lys	CAC His	TTG Leu	GAC Glu	CTC Leu	TTG Leu	TAT Tyr	CCC Pro
ACC Thr	ATC Ile	TTG Leu	GGC Gly	CAT His	GAA Glu	GGG Gly	GCT Ala	GGA Gly	ATC Ile	GTT Val	GAG Glu	AGT Ser	ATT Ile	GGA Gly	G AA Glu	GGA Gly	GTA Val	AGC Ser	ACA Thr
		. :	262								263								
GTG Val	AAA Lys	CCA Pro	G gt G	agga	aatta in	a itror	n 3	ttgo	ctcti	tttaj	g GT ly	GAC Asp	AAA Lys	Val	Ile	ACA Thr	Leu	Phe	Leu
CCA Pro	CAG Gln	TCT Cys	GGA Gly	GAA Glu	TGT Cys	ACC Thr	TCT Ser	TGC Cys	CTG Leu	AAT Asn	TCT Ser	GAG Glu	GGC Gly	AAT Asn	TTT Phe	TGT Cys	ATA Ile	CAA Gln	TTC Phe
350 AA g Ly	taaş	gttt	tt int:	 ron	tttc 4	tcta	taaa	35 ag Á s	l CAG Gln	TCA Ser	AAA Lys	ACC Thr	CAA Gln	CTG Leu	ATG Met	TCT Ser	GAT Asp	GGT Gly	ACC Thr
AGC Ser	AGG Arg	<u>TTT</u> Phe	<u>ACC</u> Thr	TGC Cys	AAG Lys	GGA Gly	AAA Lys	TCA Ser	ATA Ile	TAT Tyr	CAC His	TTT Phe	GGT Gly	AAT Asn	ACC Thr	AGC Ser	ACC Thr	TTC Phe	TGT Cys
GAA Glu	TAC Tvr	ACA Thr	GTG Val	ATA Ile	AAG Lvs	GAA Glu	ATC Ile	TCA Ser	GTT Val	GCC Ala	AAG Lys	ATT Ile	GAT Asp	GCA Ala	GTC Val	GCT Ala	CCT Pro	<u>CTA</u> Leu	GAG Glu

Proc. Natl. Acad. Sci. USA 88 (1991)

which was closely located to the end of the insert, gave a positive hybridization signal in Southern blot hybridization with the ADH2 probe. Nucleotide sequencing revealed that a part of this fragment showed a significant homology with exon 6 of the *ADH2* gene (18). The 0.8-kb *Xba* I fragment was used as a probe in turn to obtain seven contiguous genomic clones λ ADH61, -161, -31, -101, -51, -62, and -171 from the λ DASH genomic library (Fig. 1). Southern blot hybridization of these genomic clones with the ADH2 cDNA probe indicated that λ ADH-61 and λ ADH-171 overlapped at their 3' and 5' regions and together spanned \approx 17 kb. Since only exon 5 and exon 6 of this gene were cross-hybridizable with the ADH2 cDNA probe, other exons of the gene were mapped by using its own cDNA as a probe as described below.

Isolation and Characterization of cDNA Clones. Twelve clones were picked from the 5' PCR-amplified cDNA library, and partial nucleotide sequences of their inserts were determined. Five of them had the nucleotide sequence corresponding to the genomic DNA flanking 5' primer sequence. The longest two clones (A8 and D7) were analyzed further. Both clones (insert \approx 450 bp) contained the initiation codon, the upstream nonsense codon, and 5' untranslated region of different lengths. Three clones (1.1, 2.10, and 3.2) obtained from the 3' amplified library had a 0.95-kb insert, which corresponded to the major PCR product. The length of poly(A) stretch at the 3' end differed in these three clones. A full-length cDNA, which can encode 368 amino acid residues including the initiation methionine, was constructed by conjugating the 5' A8 and 3' 1.1 clones.

AAA Lys	GTA Val	TGC Cys	CTA Leu	ATT Ile	AGC Ser	TGT Cys	GGC Gly	TTT Phe	TCC Ser	ACT Thr	GGG Gly	TTT Phe	GGT Gly	GCT Ala	GCA Ala	ATC Ile	AAT Asn	ACT Thr	GCC Ala
567									568								cmc	-	
lys	gtga	1998	gca- int	ron	5	ctat	gegei	cag	Val	Thr	Pro	Gly	Ser	Thr	Cys	Ala	Val	Phe	Gly
CTG Leu	GGA Glv	GGA Glv	GTC Val	GGC Glv	TTG Leu	TCT Ser	GTT Val	GTC Val	ATG Met	GGT Glv	TGT Cvs	AAA Lvs	GCA Ala	GCA Ala	GGA Gly	GCA Ala	GCC Ala	AGG Arg	ATC Ile
	,									,								0	
ATT Ile	GGA Gly	GTG Val	GAT Asp	GTC Val	AAC Asn	AAG Lys	GAG Glu	AAA Lys	TTT Phe	AAG Lys	AAG Lys	GCA Ala	CAG Gln	GAA Glu	TTG Leu	GGT Gly	GCG Ala	ACT Thr	GAG Glu
TGC Cys	CTC Leu	AAC Asn	CCT Pro	C AG Gln	GAC Asp	TTA Leu	AAG Lys	AAA Lys	CCC Pro	ATT Ile	CAA Gln	GAA Glu	GTT Val	TTA Leu	TTT Phe	GAT Asp	ATG Met	ACA Thr	GAT Asp
															8	28			
GCT Ala	GGT Gly	ATA Ile	GAC Asp	TTC Phe	TGC Cys	TTT Phe	GAG Glu	GCC Ala	ATT Ile	GGA Gly	AAT Asn	CTG Leu	GAC Asp	GTT Val	CTG Leu	gta	tgta	atc- intro	 on 6
				829															
cct	cact	tcga	acag	ĠCA Ala	GCT Ala	GCC Ala	CTC Leu	GCC Ala	Ser	TGC Cys	AA1 Asn	GAG Glu	AGC Ser	TA1 Tyr	GGG Gly	GTC Val	TGT Cys	GTG Val	GTT Val
GTT Val	GGG Gly	GTG Val	; TTG L Leu	CCT Pro	GCC Ala	AGT Ser	GTT Val	CAA Gln	. CTC . Leu	AAA Lys	ATC Ile	AGT Ser	GGC Gly	CAG Glr	TTG Leu	TTC Phe	TTC Phe	TCA Ser	GGA Gly
									964								96	5	
CGT Arg	TCT Ser	TTG Leu	5 AAG 1 Lys	GGT Gly	TCT Ser	GTT Val	TTT 1 Phe	GGA Gly	. G g 7 G	tatg	gatg	a int:	aa	tctg 7	tatt	tcca	g ĠC ly	TGG Trp	AAG Lys
AGC	AGA		G CAC	ATC	CCI		CTG	GTI	GCT	GAT	TAI	ATO	GCA	GAG	AAG	TTG	AAT	CTA	GAT
Ser	Arş	g GII	1 115	s i i e	Fre	, Lys	s Leu	i val	. Ald	, ve	, rài	net	. 110	1 010	Lys	Leu	ASI	Leu	пэр
CCA Pro	CT/ Lev	A AT	F ACT a Thr	CAT His	ACT Thr	CTC Leu	G AA1 Asn	CTI Leu	GAT Asp	AAA Lys	ATC Ile	AA1 Asr	GTI GIU	GC1 Ala	CGTI Val	GAA Glu	TTA Leu	. ATG Met	AAA Lys
ACT	GGA	A AA	110 A TGC)4 5 TA/	A AC	rgcco	CACA	[AAA	TTTT.	ATGT	CTGT	GTGT	TTGA	GACC	AAGA	GAAA	GGGT	GTAT	TGGC
TH1	CCM	у Бу. СТАС			rcan	°CCT1	CCTC	° 4 C 4	AGCT	GTAT	ACTG	ATA 4	TTGA	AGAG	GCTT	TCAG	GAAT	TTGT	`
110	OGA	JING	nnon	1000	1001		0010	SAUA	AUCI	oini	noro			10/10	17	06	0.011		
CAT	CTC	CTTC	СССТО	CTGC	ATTT	IGTT	TTAT	гтст	AGCT	AATA	<u>aa</u> at	ACAI	AATO	CTGA	AAGI	Att	aagt	gtto	acct
acc	gtt	actt	ttgc	caat	tage	attg	tatt	tcca	atat	ggat	tttt	ttt	t						
																		_	

FIG. 3. Nucleotide sequence of the *ADH6* gene and deduced amino acid sequence. Nucleotide residues are numbered on the basis of the cDNA sequence from +1 at the adenine of the initiator codon, and nucleotides in the 5' untranslated region are indicated by negative numbers. Upstream 5' in-frame stop codon (TGA) and glucocorticoid response element are underlined. Downstream 3' polyadenylylation signal (AATAAA) is also underlined. Poly(A) stretch starts from the A at nt 1306. The sequences of reverse transcriptase primer, 5' primer, and 3' primer used for construction of the amplified cDNA libraries (Fig. 2) are indicated by dashed lines. cDNA sequence is shown in capital letters, and the 5' region of the gene and the intron sequences are shown in lowercase letters.

T instead of A at nucleotide (nt) 12, G instead of A at nt 75, T instead of G at nt 223, and T instead of G at nt 714 were found in one of the cDNA clones but not in others and not in the genomic clones and genomic DNA samples obtained from five individuals. Except for $G \rightarrow T$ at nt 223, which can induce Val \rightarrow Phe substitution, all nucleotide discrepancies are silent and cannot induce any amino acid substitutions. Since the cDNA clones were derived from PCR amplified libraries, the discrepancies may be attributed to errors that occurred in the PCR. The nucleotide sequence shown in Fig. 3 can be considered to be the authentic normal cDNA structure.

Exon-Intron Organization. Using cDNA clones A8 and 1.1 as probes, the exons were mapped within the genomic regions, which were represented by contiguous phage and cosmid clones. The length of the gene was ≈ 17 kb (Fig. 1). The nucleotide sequences of exons and exon-intron junctions were determined. All junctions had the consensus GT-AG sequence. The gene consists of eight coding exons interrupted by seven introns and can encode 368 amino acid residues (Fig. 3). The human class I ADH genes (ADH), ADH2, and ADH3) consist of nine coding exons and encode 375 amino acid residues (13, 18). The newly discovered gene shares common insertion sites with the class I ADH genes (12, 18, 19) from the first through the seventh intron-i.e., the first insertion at positions 5/6, the second at 39/40, the third at 86, the fourth at 115, the fifth at 188/189, the sixth at 275/276, and the seventh at 321. However, the ADH6 gene lacks exon 9; instead, its exon 8 contains a stop codon

MINTTGOVIRCKAAILWKPGAPFSIEEVEVAPPKAKEVRIKVVATGLCGTEMKVLGSKHLEVLYP MSTAGKVIKCKAAUNHELKK<u>PFSIEEVEVAPPKAHEVRIKVVA</u>VGICCTDENVUSSTMUTP-L NSTAGKVIKCKAAUNHELKK<u>PFSIEDVEVAPPKAHEVRIKVVA</u>VGICCRDDENVSGTMUTP-L NSTAGKVIKCKAAUNEVKK<u>PFSIEDVEVAPPKA</u>YEVRIKVVAVGICRDDENVSGTUTP-L NSTAGKVIKCKAAUNEKK<u>PFSIEDVEVAPPKA</u>HEVRIKVAAGICRSDEHVVSGNLVTP-L NGTKGKVIKCKAAIMEAGKELCIEVVEVAPPKAHEVRIGIIATSLCHDASVIDSKEEGLAFE MANE<u>VI</u>K**CKAA**VA**W**EAGKPLSIEEIEVAPPKAHEVRIKIIATAVCHTDAYTLSGADPEGAFP * TILGHEGAGIVESIGEGVSTVKPGDKVITLFLPQCGECTSCLNSEGNFCIQFKQSKTQ-----: VILGHEAAGIVESVGEGVTTVKPGDKVIPLAIPQCGKCRICKNPESNYCLKNDVSNPQG----T : VILGHEAAGIVESVGEGVTTVKPGDKVIPLTPQCGKCRICKNPESNYCLKNDLGNPG----T : VILGHEAGIVESVGEGVTTVKPGDKVIPLTPQCGKCRICKNPESNYCLKNDLGNPG----T : VIVGHEAAGIVESIGPGVTNVKPGDKVIPLTPQCGKCRICKNPESNYCLKNDLSNPSG----T GHEGAGIVESVGEGVTKLKAGDTVIPLYIPOCGECKFCLNPKTNLCQKIRVTQGKG-MSDGTSRFTCKGKS I YHFGNTSTFCEYTVI KE I SVAK I DAVAP LEKVCL I SCGFSTGFGAAINT LODGTSRFTCRRKILHELGISTESOTTYVDENAVAKIDAASPLEKVOLIGGGFSTGYGSAVNV LODGTRFTCRGKILHELGISTESOTTYVDENAVAKIDAASPLEKVOLIGGGFSTGYGSAVNV LODGTRFTCRGKILHELGISTESOTTVDENAVAKIDAASPLEKVOLIGGGFSTGYGSAVNV MERKISFTCKGRVVHFFCJSTFSOTTVDENAVAKIDAASPLEKVOLIGGGFSTGYGAANN MPDGTSRFTCKGKTILHYMGTSTFSEYTVVADISVAKIDPLAPLDKVCLLGCGISTGYGAAVNT AKVTPGSTCAVFGLGGVGLSVVMGCKAAGAARIIGVDVNKEKFKKAOELGATECINPODLKKPI AKVTPGSTCAVFGLGGVGLSAIMGCKAAGAARIIAVDINKEFAKAKELGATECINPODYKKPI AKVTPGSTCAVFGLGGVGLSAIMGCKAAGAARIIAVDINKEFAKAKELGATECINPODYKKPI AKVTPGSTCAVFGLGGVGLSAVMGCKAAGAARIIAVDINKEFAKAKELGATECINPODYKKPI AKVTPGSTCAVFGLGGVGLSAVMGCKAAGASRIIGIDINSKFVKAKALGATECINPODYKRPI AKLEPGSVCAVFGLGGVGLAVIMGCKVAGASRIIGVDINKDKFARAKEFGATECINPODFSKPI • OEVLFDMTDAGIDFCFEAIGNLDVLAAALASCNESYGVCVVVGVLPASVOLKISGOLFFSGRSL QEVIKEMTDGVDFSFEVIGRLDTMMASLLCCHBACGTSVIVGVPPDSQNLSMNPMLLIGBTW QEVIKEMTDGVDFSFEVIGRLDTMMASLLCCHBACGTSVIVGVPPDSQNLSNNPMLLITGBTW QEVIKEMTDGVDFSFEVIGRLDTMMASLLCCHBACGTSVIVGVPPDSQNLSINPMLLITGBTW QEVILELTKGGVDFALDCAGGSETMKAALDCTTAGWGSCTFIGVAAGSKGTIFPEELIIGBTI • OEVLIEMTDGGVDYSFECIGNVKVMRAALEACHKGWGVSVVVGVAASGEEIATRPFQLVTGRTW KGSVFGGWKSROHIPKLVADYMAEKLNLDPLITHTLNLDKINEAVELMKTGKW-* KGSVFGGKKSKQHIFKLVADFMAKKFSLDALITHTLDKINEAVELMKTCKN-* KGAVIGGFKSKECGFKLVADFMAKKFSLDALITHVLPFEKINEGFDLLKSGKSIRTVLTF * KGAVIGGFKSKESVFKLVADFMAKKFSLDALITHVLPFEKINEGFDLLKSGKSIRTVLTF * KGFIFGGKKSVDSIFKLVADFMAKKFSLDALINILPFEKINEGFDLLKSGKSIRTVLTF * KGTAFGGHKSVDSIFKLVPLYNKKFNLDALVTHTLPFDKISEAFDLMKOGKSVRTVLTF * KGTAFGGHKSVESVFKLVSEYMSKKIKVDEFVTHNLSFDEINKAFELMHSGKSIRTVVKI

FIG. 4. Comparison of amino acid sequences of human ADHs. Amino acid sequence of the ADH6 product is shown on *n* lines. Sequences of ADH1 product (α subunit), ADH2 (β subunit), ADH3 (γ subunit), ADH4 (π subunit), and ADH5 (χ subunit) are shown on lines 1, 2, 3, 4, and 5, respectively. Identical residues existing in all sequences are indicated by underlined boldface letters. Gaps (–) are inserted to give maximal homology in these sequences. Numbering of amino acid residues is in accordance with that of the class I ADHs. The sequences for α , β , and γ subunits are from ref. 21, that of the π subunit is from ref. 22, and that of the χ subunit is from ref. 23. followed by a polyadenylylation signal sequence. The cDNA cloned has a poly(A) tail at the expected position. Therefore, the gene cloned is "full," not partial. The possibility of the existence of exon 9 in the *ADH6* gene and formation of mRNA with the extended 3' coding sequence, which might be produced by alternative splicing, was examined by screening the genomic libraries with the 3' region of the cloned *ADH6* gene and by amplifying poly(A) RNA using the 3' sequence of exon 8 as a 3' primer. The negative results observed suggested the absence of exon 9 in the *ADH6* gene. The gene contains a highly conserved sequence for a glucocorticoid response element (20) at the extended 5' region (Fig. 3).

Deduced Amino Acid Sequence. A 368-amino-acid-long sequence was deduced from the open reading frame of the gene. The initiation methionine codon was assigned because of the presence of the upstream in-frame stop codon (TGA) (Fig. 3). Compared with the class I ADHs, the deduced amino acid sequence had one insertion between positions 60 and 61, one deletion at position 121, and was shorter by seven residues due to the appearance of a stop codon (TAA) in exon 8 (Figs. 3 and 4). The calculated percentage of positional identity ranged from 56.8% with ADH4 (class II) to 62.8% with ADH2 (class I), which is comparable to an interclass similarity (59.5–62.8%) observed in human ADHs (Table 1).

Detection of mRNA. The slot-blot hybridization of cellular RNA with the ADH6-specific oligonucleotide probe and the ADH2-specific probe indicated that both liver and stomach contained the ADH6 mRNA as well as the ADH2 mRNA (data not shown). The analysis of PCR amplification products produced by two sets of primers, one specific for the ADH6 mRNA and the other specific for the ADH2 mRNA, confirmed the above conclusion. A PCR product (448 bp), hybridizable with the ADH6-specific probe, and a PCR product (205 bp), hybridizable with the ADH2-specific probe, were both produced from a stomach poly(A) RNA as well as a liver poly(A) RNA preparation (Fig. 5).

DISCUSSION

Thus far, ADH and other isozymes have been identified through examination of enzymatic characteristics and protein analysis. Since a group of enzymes (i.e., isozymes) have a certain degree of similarity in their coding nucleotide sequences, newly discovered isozymes may be identified and characterized through direct analysis of genomic DNA. The three human class I ADH isozyme subunits are highly similar to each other (\approx 94% positional identity). A moderate similarity (\approx 60% positional identity) exists between the human class I, II, and III ADH isozymes (Table 1). Therefore, a gene for a hitherto unknown member of a group of isozymes may be cloned by screening a genomic DNA library with nucleotide probes that contain conserved coding sequences of known isozyme members under a low stringent hybridization condition and/or by selecting weakly positive clones.

In the present study, a cosmid genomic library was screened by a full-length ADH2 cDNA. A genomic clone containing the nucleotide sequence, which was homologous,

Table 1. Similarity of amino acid sequences encoded by the human *ADH* genes

	Class I ADH1	Class I ADH2	Class I ADH3	Class II ADH4	Class III ADH5
ADH6	62.2	62.8	62.2	56.8	57.4
ADHI	100	93.6	93.1	59.5	59.7
ADH2		100	94.9	59.7	62.5
ADH3			100	59.5	62.8
ADH4				100	61.3

Numbers given are maximum positional identity between the ADH subunits encoded by the genes.



FIG. 5. Detection of ADH6-specific mRNA and ADH2-specific mRNA in the stomach and liver. Poly(A) RNA prepared from the tissues was amplified by using ADH6-specific 5' and 3' primers and ADH2-specific 5' and 3' primers. Amplified product was electrophoresed and blotted onto nitrocellulose filters. (A) Hybridized with ADH6-specific probe. (B) Hybridized with ADH2-specific probe. Lanes: L, product from liver poly(A) RNA; S, product from stomach poly(A) RNA.

but not identical, to exon 6 of any one of the known human ADH genes, was obtained. It should be mentioned that exon 6 is highly conserved (88% homology) in all five known human ADH genes. Using this region of the clone as a probe, overlapping genomic clones, which covered an entire gene, were obtained from a λ DASH genomic library (Fig. 1). Subsequently, a full-length cDNA was obtained from PCRamplified cDNA libraries, which were constructed based on the sequence data of the genomic clones (Fig. 2). The method outlined above is useful for the search of additional isozyme members of other enzymes. Another human aldehyde dehydrogenase gene was also cloned and characterized in this laboratory (1, 24).

The general architectural structure of the ADH6 gene is similar to that of other human ADH genes. However, the degree of resemblance between this gene and previously known genes is only $\approx 60\%$ at the coding nucleotide level and at the deduced amino acid sequence level (Table 1). Since the degree of similarity is much lower than the interisozyme class similarity observed even in different mammalian species (\approx 85%), the gene must be classified in a different class. Recently, partial amino acid sequence (165 amino acid residues) of rat stomach ADH isozyme was determined, and the sequence was found to be not highly similar to any one of rat class I, II, and III isozymes (11). Thus, the rat stomach ADH was classified in another class, ADH class IV (11). The similarity between the newly discovered human ADH (ADH6) and the rat stomach ADH is much lower (62%) identity at the amino acid sequence level) than the intraclass similarity of class I, II, and III isozymes between the human and rat (84-95%) (11). Therefore, the human ADH6 gene cannot be for human class IV isozyme. The isozyme encoded by the ADH6 gene probably exhibits unique enzymological properties, different from other classes of ADH, and may play a particular physiological role.

The human ADH6 gene is expressed in the stomach as well as in the liver, as evidenced by cDNA cloning, slot-blot RNA hybridization, and detection of specific mRNA by a PCR method (Fig. 5).

The ADH6 gene has the glucocorticoid response element, which is highly compatible (12 of 15 nucleotide bases) to the consensus sequence of the element (20) at the 5' region (Fig. 3), and thus its expression may be under hormonal control. The less homologous response elements existing in the ADH2 gene were suggested to be functional (25).

The recently reported stomach-specific human ADH isozyme(s) (μ - and/or σ -ADH) appears to correspond to the rat stomach-specific class IV ADH, rather than to the present ADH6 with respect to kinetic properties and tissue specificity (7, 8). However, further studies are required to clarify this point.

A substantial part of orally administered ethanol is metabolized in the stomachs of males, but not of females, and it was reported that the ADH activity of stomach mucus was higher in males than in females (26). It is not yet clear whether or not the stomach ADH (μ or σ) and class I ADH (γ) are involved in the sex difference of ethanol metabolism. The ADH isozyme produced in the stomach by the ADH6 gene with the hormone response element may play a role in this phenomenon.

We thank Dr. Akitaka Shibuya (Kitasato University, Japan) for providing human stomach tissue specimen and Arlene Collinwood for assistance in preparation of the manuscript. This work was supported by U.S. Public Health Service Grant AA05763.

- Yoshida, A., Hsu, L. C. & Yasunami, M. (1991) Prog. Nucleic 1. Acid Res. Mol. Biol. 40, 255-287.
- Smith, M. (1986) Adv. Hum. Genet. 15, 249-290.
- 3. Strydom, D. J. & Vallee, B. L. (1982) Anal. Biochem. 123, 422-429
- Li, T. K., Bosron, W. F., Defeldecker, W. P., Lange, L. G. & 4. Vallee, B. L. (1977) Proc. Natl. Acad. Sci. USA 74, 4378-4381.
- Adinolfi, A., Adinolfi, M. & Hopkinson, D. A. (1984) Ann. 5 Hum. Genet. 48, 1-10.
- Jörnvall, H., Höög, J.-O., von Bahr-Lindström, H. & Vallee, 6. B. L. (1987) Proc. Natl. Acad. Sci. USA 84, 2580-2584.
- Yin, S.-J., Wang, M.-F., Liao, C.-S., Chen, C.-M. & Wu, 7 C.-W. (1990) Biochem. Int. 22, 829-835.
- Moreno, A. & Parés, X. (1991) J. Biol. Chem. 266, 1128-1133.
- Julià, P., Farrés, J. & Parés, X. (1987) Eur. J. Biochem. 162, 9. 179-189.
- 10. Holmes, R. S. & Van de Berg, J. L. (1986) Exp. Eye Res. 43, 383-396.
- 11. Parés, X., Moreno, A., Cederlund, E., Höög, J.-O. & Jörnvall, J. (1990) FEBS Lett. 277, 115-118.
- Yasunami, M., Kikuchi, I., Sarapata, D. & Yoshida, A. (1990) 12 Genomics 7, 152–158.
- 13. Ikuta, T., Fujiyoshi, T., Kurachi, K. & Yoshida, A. (1985) Proc. Natl. Acad. Sci. USA 82, 2703-2707.
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Molecular 14. Cloning: A Laboratory Manual (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
- Feinberg, A. P. & Vogelstein, B. (1983) Anal. Biochem. 132, 15. 6 - 13.
- Frohman, M. A., Dusk, M. K. & Martin, G. R. (1988) Proc. 16. Natl. Acad. Sci. USA 85, 8998-9002.
- Okayama, H. & Berg, P. (1982) Mol. Cell. Biol. 2, 1422-1427. 17.
- Duester, G., Smith, M., Bilanchone, V. & Hatfield, G. W. 18 (1986) J. Biol. Chem. 261, 2027-2033.
- 19.
- Matsuo, Y. & Yokoyama, S. (1989) FEBS Lett. 243, 57-60. Jantzen, H.-M., Strahle, U., Gloss, B., Stewart, F., Schmid, 20. W., Boshart, M., Miksicek, R. & Schutz, G. (1987) Cell 49, 29-38.
- Ikuta, T., Szeto, S. & Yoshida, A. (1986) Proc. Natl. Acad. Sci. 21. USA 83, 634-638.
- 22. Kaiser, R., Holmquist, B., Hempel, J., Vallee, B. L. & Jörnvall, H. (1988) Biochemistry 27, 1132-1140.
- Julià, P., Parés, X. & Jörnvall, H. (1988) Eur. J. Biochem. 172, 23. 73-83.
- Hsu, L. C., Chang, W.-C. & Yoshida, A. (1989) Am. J. Hum. 24. Genet. 45, 196 (abstr.).
- 25. Winter, L. A., Stewart, M. J., Shean, M. L., Dong, Y., Poellinger, L., Okret, S., Gustafsson, J.-K. & Duester, G. (1990) Gene 91, 233-240.
- Frezza, M., di Padova, C., Pozzato, G., Terpin, M., Baraona, 26. E. & Lieber, C. S. (1990) N. Engl. J. Med. 322, 95-99.