Proc. Natl. Acad. Sci. USA
Vol. 82, pp. 2262–2266, April 1985
Biochemistry

Cloning and sequencing of cDNA encoding the complete mouse liver alcohol dehydrogenase

(recombinant DNA/deduced amino acid sequence/oligonucleotide probe/initiation of protein synthesis)

HOWARD J. EDENBERG, KE ZHANG, KENNETH FONG*, WILLIAM F. BOSRON, AND TING-KAI LI

Department of Biochemistry, Indiana University School of Medicine, 635 Barnhill Drive, Indianapolis, IN 46223

Communicated by Bert L. Vallee, December 3, 1984

ABSTRACT **The main ethanol-active alcohol dehydrogenase (ADH; alcohol:NAD$^+$ oxidoreductase, EC 1.1.1.1) in mouse liver (ADH-AA) is similar in catalytic and molecular properties to horse liver ADH-EE and to the human class I ADHs. We have isolated cDNA clones encoding the entire mouse liver enzyme plus flanking regions. A mixture of 16 different oligonucleotides, each 14 bases long, was used to screen a liver cDNA library made from a DBA/2J mouse. A strongly hybridizing clone was found and identified as an ADH-encoding cDNA by partial DNA sequencing. This clone was used as a probe to identify others. Two overlapping cDNA clones together contained the entire protein-encoding region plus 100 nucleotides of the 5′ noncoding region and 133 nucleotides of the 3′ noncoding region culminating in a short poly(dA) tail. The amino acid sequence of the mouse liver enzyme deduced from this cDNA closely resembles that of horse liver ADH-E: 316 of 374 residues are identical, and 29 of the differences are conservative substitutions. The 5′ region of this cDNA is interesting: the AUG that initiates the ADH polypeptide is preceded by an AUG that would encode the first amino acid of a tripeptide. Presumably termination of this tripeptide is followed by reinitiation at the AUG immediately preceding the sequence of the mature ADH polypeptide.**

The effects of ethanol consumption depend upon tissue concentrations of the alcohol and its metabolites and upon duration of exposure; these in turn are largely determined by the rate of ethanol oxidation in the liver (1). In mammals, the rate-limiting step of ethanol oxidation is catalyzed by alcohol dehydrogenase (ADH; alcohol:NAD$^+$ oxidoreductase, EC 1.1.1.1). Mammalian ADHs are dimeric molecules with subunits of about 40,000 daltons. The horse liver ADH-E and ADH-S subunits have been sequenced (2), and the three-dimensional structure of ADH-EE has been determined to 2.4-Å resolution by x-ray crystallography (3). Structural comparisons of the enzymes from horse, rat, humans, yeast, and *Bacillus stearothermophilus* have been made (4–6).

There are many isozymes of ADH in humans, each representing homodimeric or heterodimeric combinations of subunits (1, 7). It has been proposed that the human class I isozymes are the products of alleles at three loci (7); presumably classes II and III isozymes arise from yet other loci (8). These ADH isozymes differ substantially in kinetic properties (1, 8–13). Thus variations in the structure and expression of the genes for ADH subunits might underlie the differences among individuals in ethanol metabolism (1, 13) and in physiological responses to ethanol.

We are interested in studying the control of ADH gene expression in a system which is similar to that in the human and amenable to experimental manipulation. Mice, like humans, have three classes of ADHs; three loci for their structural genes have been proposed (14–16). The ADH-AA isozyme, assigned to the putative *Adh-1* locus, is found in liver and other tissues and resembles the human class I isozymes in catalytic properties. The B and C subunits, assigned to the *Adh-2* and *Adh-3* loci, respectively, are immunologically distinct from the A subunit; ADH-BB and ADH-CC have catalytic properties resembling the human class III and class II isozymes (16). These similarities and the availability of inbred mouse strains to study gene expression make the mouse an excellent model system.

We began our study by cloning and sequencing a cDNA encoding a portion of a mouse liver ADH (17). We report here the cloning and sequencing of the cDNA for the entire ADH subunit and the amino acid sequence deduced from this cDNA sequence.

MATERIALS AND METHODS

cDNA Library. The cDNAs were made from poly(A)$^+$ RNA extracted from the liver of a DBA/2J mouse (18). They were tailed with poly(dC) and inserted into poly(dG)-tailed, *Pst* I-cleaved pBR322. The resulting molecules were used to transform *Escherichia coli* strain MC1061; about 100,000 independent transformants were generated.

Selection of Positive Clones. The library was screened by using a mixed oligonucleotide probe (19). The amino acid sequence of mouse liver ADH has not been determined. Therefore, we searched the sequence of horse ADH-E (2, 5) and the partial sequences of rat liver ADH (6) for a conserved region in the putative catalytic domain. We designed a mixed oligonucleotide probe (19) to be complementary to all 16 mRNA sequences that could encode amino acids 334 to 338 of the horse sequence (2, 5). The mixed 14-nucleotide probe shown in Fig. 1 was synthesized by P-L Biochemicals. The probe was made radioactive by using T4 polynucleotide kinase and [γ-^{32}P]ATP (20).

Aliquots of the library were spread at densities of 200–500 colonies per dish on nutrient agar (10 g of tryptone, 5 g of yeast extract, 5 g of NaCl, 15 g of agar per liter, adjusted to pH 7.5) containing tetracycline (15 μg/ml) and transferred by contact to Whatman 541 or 42 filters (21). Filters were successively exposed to 0.5 M NaOH/1.5 M NaCl (twice), 0.5 M Tris·HCl, pH 7.0/1.5 M NaCl (twice), and 2× NaCl/Cit (1× NaCl/Cit is 0.15 M NaCl/0.015 M sodium citrate), then air dried and heated to 80°C under reduced pressure.

DNA on filters was hybridized to the oligonucleotide probe for 1 hr at 33°C in 6× NaCl/Cit containing 0.1% sodium pyrophosphate and 5× Denhardt's solution (22). Filters were washed in 6× NaCl/Cit twice at room temperature, followed by 5 min at 33°C. The colony giving the strongest signal (pZK6-6) was characterized.

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.

Abbreviations: ADH, alcohol dehydrogenase; bp, base pair(s).

*Present address: Clontech Laboratories Inc., 922 Industrial Avenue, Palo Alto, CA 94303.

	334	335	336	337	338
AMINO ACIDS:	ASP	PHE	MET	ALA	LYS
CODONS:	5′-GA$^{U}_{C}$	UU$^{U}_{C}$	AUG	GCN	AA$^{A}_{G}$-3′
PROBE:	3′-CT$^{A}_{G}$	AA$^{A}_{G}$	TAC	CGN	TT--5′

FIG. 1. Probe design. Amino acids 334 to 338 of horse liver ADH-E (2, 5) and the mRNA sequences that could encode them. The 16 different 14-nucleotide-long oligomers complementary to each potential mRNA sequence are shown. N represents a mixture of all four nucleotides.

The library was rescreened, using as probe the entire insert from pZK6-6 (Fig. 2) excised from pBR322 sequences with *Pst* I and nick-translated (20). Hybridization was at 65°C in 6× NaCl/Cit/14× Denhardt's solution/1.5 mM EDTA/0.1% $NaDodSO_4$. After two washes at 22°C in 2× NaCl/Cit/0.1% $NaDodSO_4$, filters were incubated 2 hr at 65°C in 1× NaCl/Cit/0.1% $NaDodSO_4$. The positive clone with the longest insert was characterized (pZK7). A 330-base-pair (bp) fragment extending from the 5′ end of pZK7 to the *Bgl* I site (Fig. 2) was used as a probe for a further screening of the library, under similar conditions.

DNA Purification and Restriction Analysis. Selected colonies were grown in nutrient broth and amplified overnight in the presence of chloramphenicol (170 μg/ml). Plasmid DNA was isolated by alkaline lysis (23) and purified by centrifugation to equilibrium in CsCl/ethidium bromide. Restriction maps were assembled by standard techniques. Detection of fragments was by one of three techniques: ethidium bromide fluorescence, transfer to nitrocellulose (24) and hybridization to the mixed oligonucleotide probe or to nick-translated restriction fragments, or autoradiography of end-labeled and redigested restriction fragments (25).

DNA Sequencing. The mixed oligonucleotide probe was annealed with double-stranded pZK6-6 DNA by heating to 93°C for 7 min and slowly cooling to 30°C, at which temperature the dideoxynucleotide sequencing reactions (26) were performed.

Further sequencing was done after subcloning portions of the cDNAs (Fig. 2) in M13 vectors (27). Single-stranded DNA extracted from the phage was annealed with a 15-nucleotide "universal primer" (Bethesda Research Laboratories) for dideoxynucleotide sequencing (26, 27). Multiple independent isolates of many M13 subclones were sequenced (not shown); in some cases, the same region was sequenced from several independent clones from the liver cDNA library. Nucleotide sequences were aligned and compared by using a microcomputer (28, 29).

RESULTS

Isolation of pZK6-6, an ADH cDNA Clone. The mixed oligonucleotide probe shown in Fig. 1 was used to identify a strongly hybridizing colony, pZK6-6. Confirmation that pZK6-6 contained an ADH cDNA was obtained by DNA sequence analysis. Since the mixture of oligonucleotides hybridized to a single restriction fragment in pZK6-6, we used it as a primer for dideoxynucleotide sequencing of the double-stranded clone itself. We thus determined the sequence of 105 nucleotides (17), which encoded 35 amino acids in the expected direction from the primer site; all but two matched the sequence of horse ADH-E (2). This is a useful extension of the technique of Smith *et al.* (30), who used individual single-stranded oligonucleotides as primers for sequencing a duplex template. The ability to sequence *double-stranded* DNA by using as a primer the same *mixed* oligonucleotide originally used as probe is a great convenience: it is rapid, avoids extra cloning steps, and ensures that a portion of the coding region will be sequenced. pZK6-6, containing an insert of only 700 bp, was too small to encode the entire mouse ADH (Fig. 2).

Other ADH cDNA Clones. We used the 700-bp cDNA insert of pZK6-6 as a probe to screen the library, and we characterized the largest cDNA obtained, pZK7. Restriction mapping indicated that pZK7 overlapped pZK6-6 and extended further toward the sequences encoding the amino terminus of ADH (Fig. 2). This cDNA was 1100 bp long, still too short to encode the entire ADH. We rescreened the library, using the entire cDNA insert of pZK7 as a probe, but did not find colonies containing longer inserts.

We then screened the library by using as a probe the portion of pZK7 encoding the amino acids nearest to the amino terminus of ADH; this was a *Pst* I/*Bgl* I fragment (Fig. 2). Restriction mapping showed that several clones contained a new *Pst* I site; among them were two clones (pZK102-21 and pZK105-36) extending 97 and 100 nucleotides beyond the initiating ATG.

DNA Sequencing. The sequences of pZK6-6 and pZK7 (Fig. 2) were determined in their entirety: pZK7 extends from nucleotide 306 (Fig. 3) to the 3′ end, and pZK6-6 extends from nucleotide 717 to the 3′ end. There are two differences between these cDNAs: nucleotides 1281 and 1334 are both T in pZK6-6. This eliminates from pZK6-6 the *Hin*fI site shown in Fig. 2 (and underlined in Fig. 3), an alteration confirmed by restriction mapping.

pZK105-36 (Fig. 2) provided the sequence of the 5′ portion of the cDNA. There is a single difference at nucleotide 323 (T in pZK7), which does not alter the amino acid encoded there. Two independent clones (not shown) from the cDNA library provided further confirmation of this region: portions of pZK102-21 were sequenced to confirm nucleotides 4–87 and 289–450 (with the same T at nucleotide 323 as in pZK7) and pZK112-72 confirmed nucleotides 128–294. The *Alu* I site underlined at nucleotides 2–5 (Fig. 3) was directly confirmed by restriction mapping (25) of DNA labeled at the *Bgl* II site

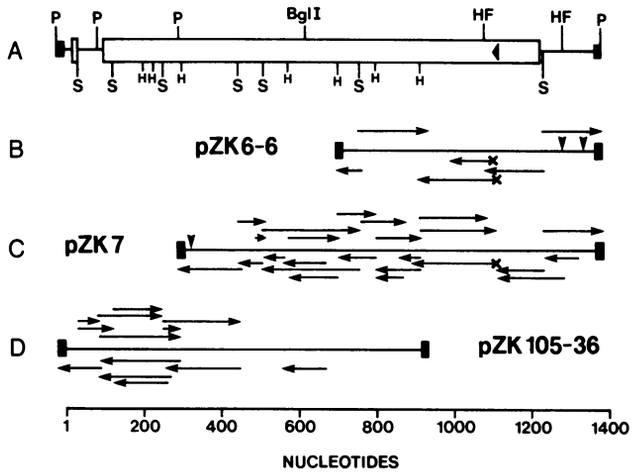

FIG. 2. Sequencing strategy. (*A*) Map of the (reconstructed) cDNA insert. Solid blocks are poly(dG)·poly(dC) introduced during cloning. Open blocks are coding regions: a small tripeptide at the left and mouse ADH extending over most of the insert. The triangle within the coding region marks the position of the probe sequence shown in Fig. 1. P, *Pst* I sites; HF, *Hin*fI sites; S, *Sau*3AI sites; H, *Hae* III sites. (*B*) pZK6-6 and regions sequenced after subcloning in phage M13 derivatives. × at the beginning of an arrow denotes priming by the mixed oligonucleotide probe. Vertical arrowheads represent single nucleotides that differ from those in pZK7 (see Fig. 3). (*C*) pZK7, primary reference for sequences beyond nucleotide 450; arrowhead is nucleotide 323, which differs from that in pZK105-36. (*D*) pZK105-36, primary reference for sequences from nucleotide 1 to 450. The location of the right end of pZK105-36 is approximate, based upon restriction mapping.

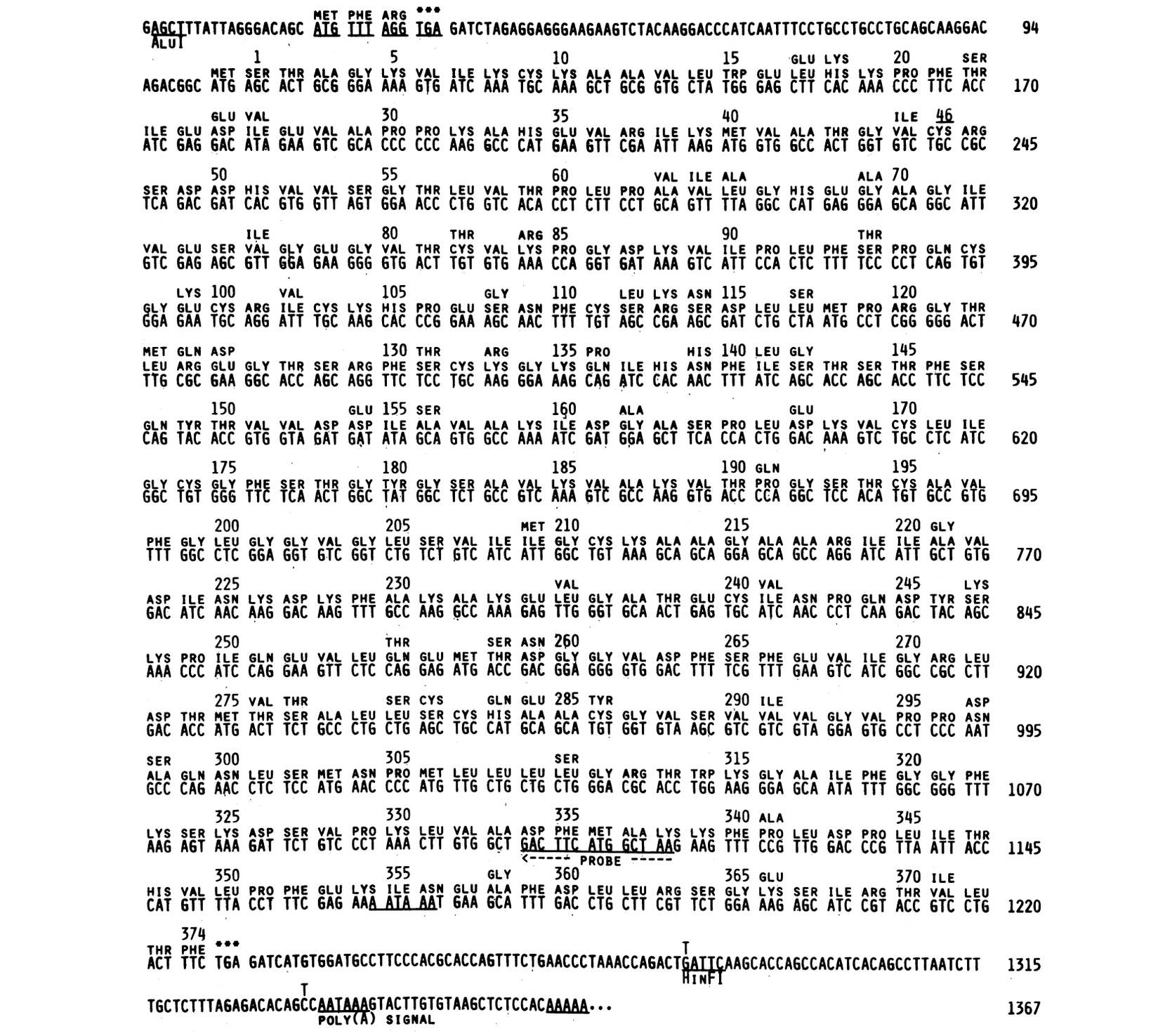

FIG. 3. DNA sequence of the complete mouse liver ADH cDNA: Composite of pZK105-36 and pZK7. Nucleotide numbering begins at the last dG of the 5′ poly(dG) and continues through the very short (5-bp) 3′ poly(dA); 16 dC residues added during cloning were omitted. The nucleotide sequence, presented in large type, is numbered at the right of the line. Just above is the amino acid translation of this cDNA. The line above that numbers the amino acid sequence according to the horse liver ADH-E sequence (2) and also gives the amino acid present in horse ADH wherever it differs from the mouse. Also on that line (at nucleotides 1281 and 1334) are two of the three single-nucleotide differences among clones from the library: T in pZK6-6 at both positions. The third nucleotide difference is a T in pZK7 substituted for the C at 323. Underlined are an *Alu* I and *Hin*fI site noted in the text, the potential tripeptide within the 5′ leader, the region complementary to the probe, two A-A-T-A-A-A sequences (one internal to the coding region), and the poly(dA).

(nucleotides 31–36). The genesis of the three single-nucleotide differences is not known: since the DBA/2J mouse is supposed to be inbred, one would not expect polymorphism at the ADH gene. It is possible that a copying error occurred in the reverse transcription of the mRNA (cf. refs. 31 and 32).

cDNA Structure and Amino Acid Sequence. The entire cDNA sequence (Fig. 3) was derived by overlapping the sequences of pZK105-36 and pZK7. The length of the cDNA is 1361 nucleotides, not including the poly(dG) at the 5′ side and a short poly(dA) followed by poly(dC) at the 3′ side. It encodes a 375-amino acid polypeptide (Fig. 3) that very

closely resembles horse ADH-E (2). There are 100 nucleotides of 5′ nontranslated region between the end of the poly(dG) (nucleotide 1) and the initiating *A*TG (nucleotide 102), and 133 nucleotides of 3′ nontranslated region between the TG*A* terminator codon and the start of the poly(dA). Within the 5′ region, it is noteworthy that there is an ATG preceding the ATG that initiates the ADH polypeptide.

The amino terminus of the horse ADH-E is acetylserine (2). The mouse cDNA sequence has the initiating methionine immediately adjacent to the corresponding serine; there is no potential leader or signal sequence. We cannot determine from the DNA sequence whether one or more amino-terminal residues are cleaved or acetylated (33).

Codon Usage. The coding region of this cDNA (1125 nucleotides encoding 375 amino acids) contains 24.3% A, 25.6% C, 26.8% G, and 23.4% T. Codon usage is nonrandom ($P < 0.001$ by χ^2). C is found in the third position of the codons 35.5% of the time ($P < 0.001$). Among the 193 codons in which all four nucleotides in the third position encode the same amino acid, C is used in 35.8% ($P < 0.01$). Neither the 5′ nor the 3′ nontranslated region shows significant bias in any of the three potential reading frames ($P > 0.2$). The short coding region of human ADH thus far reported (34), representing the carboxyl-terminal 91 amino acids, shows no significant codon bias ($P > 0.7$).

DISCUSSION

We have isolated cDNA clones encoding the entire mouse liver ADH and deduced the previously unknown amino acid sequence of the mouse liver ADH (Fig. 3). Comparisons of the amino acid sequences (2, 4–6, 34, 35) and the catalytic properties of the human (1, 8–13), rat (13), horse (5), and mouse liver ADH (14–16, 36) indicate that the mouse liver ADH-AA enzyme resembles horse liver ADH-EE, the class I ADHs of humans, and rat liver ADH. The amino acid sequence of the ADH encoded by our mouse cDNA is very similar to the horse ADH-E subunit (2, 5): 316 of 374 positions (84.5%) are identical and 29 of the 58 differences are conservative substitutions. The sequence similarity strongly suggests that this cDNA encodes the mouse ADH-A subunit. In a comparison of our mouse sequence with that of horse ADH-E (5), conserved residues include all seven zinc ligands and most sites that interact with the ADP-ribose moiety (15 of 16; Ser-277 for Thr), line the active site pocket (11 of 14; Leu-117 for Ser, Ile-141 for Leu, Leu-310 for Ser), are involved in interactions between the two subunits (12 of 13; Val-291 for Ile), or form hydrophobic cores (81 of 90, all substitutions hydrophobic). Thus the mouse ADH-AA should closely resemble the horse ADH-EE in three-dimensional structure (3). The frequency of identities in the catalytic domain (193/231, 83.5%) is similar to that in the coenzyme binding domain (123/143, 86.0%).

The mouse ADH-A encoded by our cDNA is slightly more similar to rat liver ADH than to horse ADH-E [321 of 362 positions identical, 88.7%; some peptides in the rat sequence were placed by homology to the horse ADH and 13 positions were not identified (7)]. It should be noted that rat liver ADH has an insert of one cysteine residue immediately after Cys-111 (6); the mouse enzyme whose sequence is deduced here has no such insertion.

One region of the ADH molecule, residues 108–125, shows much variability among horse, rat, and mouse sequences. In this region, there are 8 differences between mouse ADH-A and horse ADH-E, 10 between mouse ADH-A and horse ADH-S, 12 between mouse ADH-A and rat liver ADH, 13 between rat liver ADH and horse ADH-E (5), and 12 between rat liver ADH and horse ADH-S (5). Part of this region, residues 95–113, forms a projection at the surface of the molecule that binds the noncatalytic zinc (3, 5), and the remainder is involved in no apparent secondary structure (3); residues 123–126 form a reverse bend (3). The structural or catalytic significance of variations in this region is unclear.

The sequence of the mouse ADH-A subunit is also similar to that of the human ADH-β_1 subunit: 23 of 30 amino acids around the catalytic site (residues 41–70; ref. 35) are identical, as are 75 of 91 residues at the carboxyl terminus (34). Comparison with the recently published sequence of the only other reported mammalian ADH cDNA, encoding the carboxyl-terminal 91 amino acids of the human ADH-β_1 cDNA (34), shows that the similarity between nucleotide sequences in the coding region (including TGA terminator) is 81.9%, close to the 82.6% similarity in amino acid sequence. The 3′ end of the mouse cDNA (133 nucleotides) is much shorter than that of the human (590 nucleotides). Introducing three single-nucleotide gaps in the mouse and one in the human 3′ sequences aligns them: they are identical at 84 of 136 sites (62%). With this alignment, the A-A-T-A-A-A of the mouse cDNA and the first of the four A-A-T-A-A-As of the human cDNA are identically located. The similarity in nucleotide sequence, extending into the noncoding region, argues that mouse ADH-A and human ADH-β_1 genes have descended from a common ancestral gene.

Translation of the mouse liver ADH from this mRNA would begin at the second AUG, in contrast to 90–95% of the mRNAs thus far reported (37). The sequences surrounding both AUGs are similar: G-A-C-R-G-C-*A-U-G*, in which R indicates purine. The first AUG is, in fact, slightly more similar to Kozak's (37) consensus sequence (C-C-R-C-C-*A-U-G*-G) in having A as the purine at position -3 (as in 79% of the sequences), rather than G (18%). Kozak noted that of the mRNAs that do not follow the "first AUG" rule, several contain stop codons in frame with the first AUG, and she suggested that the main message may be translated by ribosomes that have first made the short polypeptide and then continued moving down the RNA before encountering the second AUG (37). Liu *et al.* (38) inserted AUGs upstream of the normal AUG of the hepatitis B surface antigen gene and found that they depressed correct initiation in monkey cells; this depression is partly suppressed by inserting in-frame termination codons. The ADH mRNA from which this cDNA arose fits this hypothesis: it could make a tripeptide and then initiate the ADH after passing 63 nucleotides (Fig. 3).

The mouse ADH cDNA contains the A-A-T-A-A-A consensus polyadenylylation signal (39) separated by 21 nucleotides from the short poly(A) stretch. There is also an A-A-T-A-A-A within the coding region (nucleotides 1166–1171), as has been reported for other genes (e.g., see ref. 36). Berget (39) examined genomic sequences around 61 vertebrate poly(A) sites and discerned two classes of consensus sequence: class I genes had C-A-C-U-G immediately upstream of the poly(A) site, and class II had abutted copies of C-A-U-U-G downstream of the site. We do not find the class I sequence immediately upstream of the poly(A); not having a genomic clone, we cannot examine for class II sequences.

Our isolation of the cDNA for the entire mouse liver ADH-A subunit provides the starting point for a number of studies. These include isolation of the genomic ADH clones from mouse and from humans, examination of the control of ADH gene expression in different mouse and human tissues, and searches for restriction fragment length polymorphisms. The latter should provide a method for determining ADH genotype in humans and allow studies correlating genotype with physiological and pathological effects of alcohol consumption in individuals.

We thank R. E. Jerome and G. Gamboa for technical assistance. Portions of this work were supported by grants from the Grace M. Showalter Trust, the Project Development Program of Indiana

University School of Medicine, and grants from the National Institute on Alcohol Abuse and Alcoholism (AA02342, AA04307, and AA06460).

1. Li, T.-K. (1983) in *Medical and Social Aspects of Alcohol Abuse*, eds. Tabakoff, B., Sutker, P. B. & Randall, C. L. (Plenum, New York), pp. 44–77.
2. Jörnvall, H. (1970) *Eur. J. Biochem.* **16,** 25–49.
3. Eklund, H., Nordström, B., Zeppezauer, E., Söderlund, G., Ohlsson, I., Boiwe, T., Söderberg, B.-O., Tapio, O., Bränden, C.-I. & Åkerson, Å. (1976) *J. Mol. Biol.* **102,** 27–59.
4. Jörnvall, H. (1977) *Eur. J. Biochem.* **72,** 443–452.
5. Brändén, C.-I., Jörnvall, H., Eklund, H. & Furugren, B. (1975) in *The Enzymes*, ed. Boyer, P. D. (Academic, New York), 3rd Ed., Vol. 11, pp. 103–190.
6. Jörnvall, H. (1974) in *Alcohol and Aldehyde Metabolizing Systems*, eds. Thurman, R. G., Yonetani, T., Williamson, J. R. & Chance, B. (Academic, New York), pp. 23–32.
7. Smith, M., Hopkinson, D. A. & Harris, H. (1971) *Ann. Hum. Genet.* **34,** 251–271.
8. Vallee, B. L. & Bazzone, T. J. (1983) in *Isoenzymes: Current Topics in Biological and Medical Research*, eds. Rattazzi, M. C., Scandalios, J. G. & Whitt, G. S. (Liss, New York), Vol. 8, pp. 219–244.
9. Schenker, T. M., Teeple, L. J. & von Wartburg, J.-P. (1971) *Eur. J. Biochem.* **24,** 271–279.
10. Blair, H. A. & Vallee, B. L. (1966) *Biochemistry* **5,** 2026–2034.
11. Pietruszko, R., Theorell, H. & DeZalenski, C. (1972) *Arch. Biochem. Biophys.* **153,** 279–293.
12. Bosron, W. F., Magnes, L. J. & Li, T.-K. (1983) *Biochemistry* **22,** 1852–1857.
13. Bosron, W. F., Crabb, D. W. & Li, T.-K. (1983) *Pharmacol., Biochem. Behav.* Suppl. 1, **18,** 223–227.
14. Holmes, R. S., Duley, J. A. & Burnell, J. N. (1983) in *Isozymes: Current Topics in Biological and Medical Research*, eds. Rattazzi, M. C., Scandalios, J. G. & Whitt, G. S. (Liss, New York), Vol. 8, pp. 155–174.
15. Felder, M. R., Burnett, K. G. & Balak, K. J. (1983) in *Isozymes: Current Topics in Biochemical and Medical Research*, eds. Rattazzi, M. C., Scandalios, J. G. & Whitt, G. S. (Liss, New York), Vol. 9, pp. 143–161.
16. Algar, E. M., Seeley, T.-L. & Holmes, R. S. (1983) *Eur. J. Biochem.* **137,** 139–147.
17. Edenberg, H. J. & Zhang, K. (1984) *ICSU Short Rep.* **1,** 162–163.
18. Goodman, H. M. & MacDonald, R. J. (1979) *Methods Enzymol.* **68,** 75–89.
19. Wallace, R. B., Johnson, M. J., Hirose, T., Miyake, T., Kawashima, E. H. & Itakura, K. (1981) *Nucleic Acids Res.* **9,** 879–894.
20. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
21. Gergen, P., Stern, R. H. & Wensink, P. C. (1979) *Nucleic Acids Res.* **7,** 2115–2136.
22. Denhardt, D. (1966) *Biochem. Biophys. Res. Commun.* **23,** 641–646.
23. Birnboim, H. C. & Doly, J. (1979) *Nucleic Acids Res.* **7,** 1513–1523.
24. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 502–517.
25. Smith, H. O. & Birnstiel, M. L. (1976) *Nucleic Acids Res.* **3,** 2387–2398.
26. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
27. Messing, J. (1983) *Methods Enzymol.* **101,** 20–78.
28. Larson, R. & Messing, J. (1982) *Nucleic Acids Res.* **10,** 39–49.
29. Queen, C. & Korn, L. J. (1984) *Nucleic Acids Res.* **12,** 581–599.
30. Smith, M., Leung, D. W., Gillam, S., Astell, C. R., Montgomery, D. L. & Hall, B. D. (1979) *Cell* **16,** 753–761.
31. Browne, J. K., Paddock, G. V., Liu, A., Clarke, P., Heindell, H. C. & Salser, W. (1977) *Science* **195,** 389–391.
32. Kakidani, H., Furutani, Y., Takahashi, H., Noda, M., Morimoto, Y., Hirose, T., Asai, M., Inayama, S., Nakanishi, S. & Numa, S. (1982) *Nature (London)* **298,** 245–249.
33. Wold, F. (1981) *Annu. Rev. Biochem.* **50,** 783–814.
34. Duester, G., Hatfield, G. W., Bühler, R., Hempel, J., Jörnvall, H. & Smith, M. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 4055–4059.
35. Jörnvall, H., Hempel, J., Vallee, B. L., Bosron, W. F. & Li, T.-K. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 3024–3028.
36. Rex, D. K., Bosron, W. F. & Li, T.-K. (1984) *Biochem. Genet.* **22,** 115–124.
37. Kozak, M. (1984) *Nucleic Acids Res.* **12,** 857–872.
38. Liu, C.-C., Simonsen, C. C. & Levinson, A. D. (1984) *Nature (London)* **309,** 82–85.
39. Berget, S. M. (1984) *Nature (London)* **309,** 179–182.