*Proc. Natl. Acad. Sci. USA*
Vol. 89, pp. 10188–10192, November 1992
Medical Sciences

# Cloning and expression of an NADP$^+$-dependent alcohol dehydrogenase gene of *Entamoeba histolytica*

(amebiasis/protozoan parasite/fermentation/inhibitors)

AJIT KUMAR*, PEI-SHEN SHEN†‡, STEVEN DESCOTEAUX*, JAN POHL§, GORDON BAILEY†, AND JOHN SAMUELSON*¶‖

*Department of Tropical Public Health, Harvard School of Public Health, Boston, MA 02115; †Department of Biochemistry, Morehouse School of Medicine, Atlanta, GA 30310; §Microchemical Facility, Emory University, Atlanta, GA 30322; and ¶Department of Pathology, New England Deaconess Hospital, Boston, MA 02215

*Communicated by Elkan Blout, June 29, 1992 (received for review May 1, 1992)*

ABSTRACT **Ethanol is the major metabolic product of glucose fermentation by the protozoan parasite *Entamoeba histolytica* under the anaerobic conditions found in the lumen of the colon. Here an internal peptide sequence determined from a major 39-kDa amoeba protein isolated by isoelectric focusing followed by SDS/PAGE was used to clone the gene for the *E. histolytica* NADP$^+$-dependent alcohol dehydrogenase (EhADH1; EC 1.1.1.2). The EhADH1 clone had an open reading frame that was 360 amino acids long and encoded a protein of ≈39 kDa (calculated size). EhADH1 showed a 62% amino acid identity with the tetrameric NADP$^+$-dependent alcohol dehydrogenase of *Thermoanaerobium brockii*. In contrast, EhADH1 showed a 15% amino acid identity with the closest human alcohol dehydrogenase. EhADH1 contained 18 of the 22 amino acids conserved in other alcohol dehydrogenases, including glycines involved in binding NAD(P)$^+$ as well as histidine and cysteine residues involved in binding the catalytic zinc ion. Like the *T. brockii* alcohol dehydrogenase, EhADH1 lacked a 23-amino acid stretch present in other alcohol dehydrogenases that includes four cysteines that bind a second noncatalytic zinc ion. An EhADH1–glutathione-*S*-transferase fusion protein showed the expected NADP$^+$-dependent alcohol dehydrogenase and NADPH-dependent acetaldehyde reductase activities. The enzymatic activities of the EhADH1 fusion protein were inhibited by pyrazole and 4-methylpyrazole.**

*Entamoeba histolytica* is a protozoan parasite that infects ≈500 million individuals in the developing world, resulting in ≈40 million cases of dysentery with a consequent annual mortality rate of ≈40,000 (1, 2). Entamoebae reproduce under anaerobic conditions in the colonic lumen and cause dysentery when they invade the bowel wall. Because poor countries cannot afford to make the changes in sanitary conditions that might eliminate fecal–oral spread of amoebic infection, amebiasis is primarily controlled by drug treatment of persons sick with the parasite (1–3).

Entamoebae are obligate fermenters because they lack mitochondria and the enzymes involved in respiration (4, 5). Amoebae use facilitated diffusion for the uptake of glucose, which is then metabolized to pyruvate via the Embden–Meyerhof pathway with the net gain of two molecules of ATP and two molecules of NADH (5). Pyruvate is converted to acetyl-CoA with the production of a ferredoxin free-radical and $CO_2$. Metronidazole, the most widely used drug to treat amebiasis, is activated by a one-electron reduction from the donor ferredoxin free-radical (3). Under anaerobic conditions, acetyl-CoA is sequentially reduced to acetaldehyde and ethanol with the production of two molecules of NAD(P)$^+$ so that the reducing equivalents are regenerated (4–7). Both NADP$^+$-dependent and NAD$^+$-dependent alcohol dehydrogenase (ADH) activities have been identified in *E. histolytica* trophozoites (4–7). The NADP$^+$-dependent ADH (EhADH1; EC 1.1.1.2) was found to be composed of four ≈30-kDa subunits, prefer secondary alcohols to primary alcohols, and be inhibited by the substrate analogue pyrazole (7). The amoeba NAD$^+$-dependent ADH (EC 1.1.1.1) is not well-characterized (4).

Human ADHs are cytoplasmic dimeric proteins composed of ≈40-kDa peptides (long chain) that bind NAD$^+$ and zinc as cofactors (8). Although the best-studied function of human ADHs is the oxidation of ethanol to acetaldehyde, human ADHs are capable of oxidizing multiple alcohols and of reducing numerous aldehydes and ketones (9). The tertiary structure of human ADH enzymes has been inferred by analogy with horse liver ADH, whose structure has been determined at atomic resolution by x-ray crystallography (10, 11).

With the long-term goal of developing antiamoebic drugs to complement those already available to treat human infection with *E. histolytica*, we are focusing on enzymes involved in the fermentation of glucose to ethanol. This metabolic pathway was chosen for four reasons. (*i*) This pathway is the target of the best antiamoebic drug metronidazole (3). (*ii*) The amoebae are dependent upon fermentation of glucose (4, 5), whereas the human host uses oxidative respiration. (*iii*) Multiple ADH and aldehyde dehydrogenase inhibitors are known (12, 13). (*iv*) One *E. histolytica* ADH is tetrameric and NADP$^+$-dependent (7) and so is likely to be quite different from human ADHs, which are dimeric and NAD$^+$-dependent (8). Here an internal peptide from an abundant 39-kDa amoeba protein isolated on two-dimensional protein gels was found serendipitously to contain a 16-amino acid sequence homologous to that of the tetrameric NADP$^+$-dependent ADH of the anaerobic thermophilic bacterium *Thermoanaerobium brockii* (14). We used the peptide sequence to clone, sequence, and express in *Escherichia coli* the NADP$^+$-dependent *E. histolytica* ADH (EhADH1) gene.**

## MATERIALS AND METHODS

**Separation of *E. histolytica* Proteins on Two-Dimensional Protein Gels.** Trophozoites of pathogenic *E. histolytica* HM-1:IHSS were cultured axenically in TYI-S-33 medium (15). For two-dimensional separation of proteins (16), parasites

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; EhADH1, NADP$^+$-dependent alcohol dehydrogenase of *E. histolytica*; GST, glutathione *S*-transferase.

‡Present address: Department of Tropical Public Health, Harvard School of Public Health, Boston, MA 02115.
‖To whom reprint requests should be addressed.

**The sequence reported in this paper has been deposited in the GenBank data base (accession no. M88600).

were grown to midlogarithmic phase, washed three times in phosphate-buffered saline, and lysed at an equivalent concentration of $2 \times 10^6$ amoebae per ml in 0.3% SDS/5% 2-mercaptoethanol/10 mM Tris, pH 8/5 mM iodoacetamide/0.2 mM diisopropyl fluorophosphate/0.2 mM 7-amino-1-chloro-3-tosylamido-2-heptanone. The lysate was boiled, incubated with RNase (7.5 μg/ml) and DNase (15 μg/ml), lyophilized, and solubilized in 9 M urea/2% 3-[(3-cholamidapropyl)dimethylammonio]-2-hydroxy-1-propanesulfonate/100 mM dithiothreitol/2% (wt/vol) ampholines ranging from pH 3 to pH 10 (Pharmacia LKB). About 300 μg of the amoeba protein was focused in isoelectric focusing gels in 1.5-mm diameter tubes. For the second dimension, the tube gels were incubated with SDS sample buffer, electrophoresed by SDS/PAGE on a slab gel containing 10% polyacrylamide with a 4% stacking gel, and stained with Coomassie brilliant blue. Major protein spots were excised and stored at −70°C in 50% (vol/vol) methanol.

Digestion of excised amoeba proteins with *Staphylococcus aureus* V8 protease was performed as described by Kennedy *et al.* (17). The isolated protein spots from four two-dimensional gels were combined, placed in SDS/PAGE buffer, digested at room temperature with V-8 protease (0.1 μg/ml) for 15 min in the sample loading well and then for 30 min at the stacking gel–running gel interface, and electrophoresed through a 15% polyacrylamide gel containing SDS. The gel was equilibrated in 10 mM 3-(cyclohexylamino)-1-propanesulfonic acid, pH 11/20% methanol, and electroblotted onto a poly(vinylidene difluoride) membrane (Bio-Rad), which was lightly stained with Coomassie blue. Peptides were excised and subjected to automated Edman degradation in an Applied Biosystems model 477A/120A sequencer/phenylthiohydantoin analyzer (18).

**Isolation of and Sequencing of the EhADH1 Gene.** Trophozoites were lysed by incubation in 1% Sarkosyl/100 mM EDTA/10 mM Tris, pH 8/proteinase K (100 μg/ml) for 2 hr at 45°C, and the genomic DNA was isolated on a cesium chloride gradient (19). A degenerate sense PCR primer [GG(TA)CAT(TC)GA(AG)GC(TA)G(TC)(TA)GG] was made in the codon usage of *E. histolytica* to a conserved amino acid sequence [GHEA(AV)G] located near the amino end of previously identified ADHs (10). A degenerate antisense [TTC(TA)AC(AG)AT(AG)TC(TA)CC(AG)TT] primer was made from a portion of the 16-amino acid sequence (NGDIVE) derived from the abundant 39-kDa amoeba protein, which was homologous to the *T. brockii* ADH (14). Thirty-five cycles of PCR were performed using these primers, 1 μg of amoeba DNA as a target, AmpliTaq polymerase (Perkin–Elmer/Cetus), and an annealing temperature of 40°C (20). The EhADH1 PCR product was cloned into the phagemid Bluescript [Stratagene (19)] and sequenced using deoxyadenosine 5′-[α-[$^{35}$S]thio]triphosphate, Sequenase (United States Biochemical), and the dideoxynucleotide chain-termination method (21). The EhADH1 PCR product was radiolabeled by random oligomer-priming (19) and used to screen an *E. histolytica* cDNA library made in λ ZAP (Stratagene), a generous gift from Egbert Tannich (Bernhard Nocht Institute for Tropical Medicine, Hamburg, F.R.G.). A phagemid (pAK311) that appeared to contain the entire coding sequence for EhADH1 was sequenced on both strands. The open reading frame of pAK311 was compared to the protein sequences in GenBank (June 9, 1992) by using the algorithm of Altschul *et al.* (22) at the Molecular Biology Computer Research Resource of the Dana–Farber Cancer Institute (Boston).

**Southern Blots.** To determine the number of copies of EhADH1 gene in the *E. histolytica* genome, 2 μg of *E. histolytica* DNA was cut with restriction enzymes (*Alu* I, *Eco*RV, *Sau*3AI, and *Dde* I), separated by size on a 1% agarose, transferred to nylon membranes (GeneScreen*Plus*; New England Nuclear), and hybridized with a radiolabeled probe containing the final 349 base pairs (bp) from the 3′ end of EhADH1 gene (19).

**Expression of EhADH1 as a Glutathione-*S*-Transferase (GST) Fusion Protein in *Escherichia coli*.** To express the EhADH1 enzyme in *Escherichia coli* XL1-Blue (Stratagene), restriction sites were added to ends of the coding sequence of the EhADH1 gene contained in pAK311 by using synthetic oligonucleotide primers, the PCR, and Vent polymerase (New England Biolabs). The PCR product was cloned in-frame at the 5′ end of the coding sequence of GST that was contained in the expression vector pGEX-2T (23). Isopropyl β-D-thiogalactoside (1 mM) was used to induce bacteria to make the GST–EhADH1 fusion protein containing the 26-kDa GST at the N terminus and the 39-kDa EhADH1 at the C terminus. Bacteria were lysed by sonication, and the GST–EhADH1 fusion protein was purified on glutathione-agarose beads (23). Protein concentrations were determined by the method of Bradford (24), and the purity of the fusion proteins was determined by SDS/PAGE. The size of the undenatured GST–EhADH1 fusion protein was determined by FPLC (Pharmacia) using a Superose 12 gel filtration column.

**Assays of ADH and Acetaldehyde Reductase Activities of the GST–EhADH1 Fusion Protein.** The ADH activity of the recombinant EhADH1 was determined by incubating enzyme extracts, substrates, and cofactors in 50 mM glycine·NaOH (pH 9.5), and monitoring the increase in absorbance at 334 nm with a Gilford Response spectrophotometer. The activities and Michaelis constants ($K_m$) of the enzyme for isopropanol and ethanol were determined by incubating various concentrations of each with 200 μM $NADP^+$ and plotting the data on an Eadie–Hofstee plot (25). The $K_m$ values for $NADP^+$ and $NAD^+$ were determined by incubating various concentrations of the cofactors with 1 mM isopropanol and plotting the data on an Eadie–Hofstee plot. One activity unit was defined as the amount of enzyme that produced 1 μmol of NADPH per min at 25°C. Inhibition constants ($K_i$) of ADH activity for pyrazole and 4-methylpyrazole were determined with a UVmax kinetic microplate reader (Molecular Devices, Palo Alto, CA) by adding various concentrations of the inhibitors to the purified fusion enzyme, 200 μM $NADP^+$, and various concentrations of isopropanol and plotting the data on a Dixon plot (25).

The acetaldehyde and acetone reductase activities and Michaelis constants of the GST–EhADH1 fusion protein were determined by incubating various concentrations of each substrate with 200 μM NADPH in 50 mM sodium phosphate (pH 6.5) and measuring the decrease in absorbance at 334 nm. The $K_m$ value for NADPH was determined by incubating various concentrations of the cofactor with 500 μM acetaldehyde. One unit of enzyme activity was defined as that which consumed 1 μmol of NADPH per min at 25°C.

## RESULTS

**Identification of a Major *E. histolytica* 39-kDa Protein on Two-Dimensional Protein Gels as a Likely ADH.** An ≈39-kDa protein was one of a number of major *E. histolytica* proteins excised from two-dimensional protein gels (Fig. 1). The 39-kDa protein was cleaved with *S. aureus* V8 protease, and a 16-amino acid long peptide sequence (YGATDIINYKNGDIVE) was obtained. Twelve of the 16 amino acids (underlined) were identical to amino acids 210–225 of *T. brockii* ADH (Fig. 2 and ref. 14).

**Cloning of the EhADH1 Gene.** A segment of the $NADP^+$-dependent EhADH1 gene was identified within *E. histolytica* DNA by using the PCR, a degenerate oligonucleotide antisense primer to the 16-amino acid sequence from the abundant 39-kDa amoeba protein, and a sense primer inferred

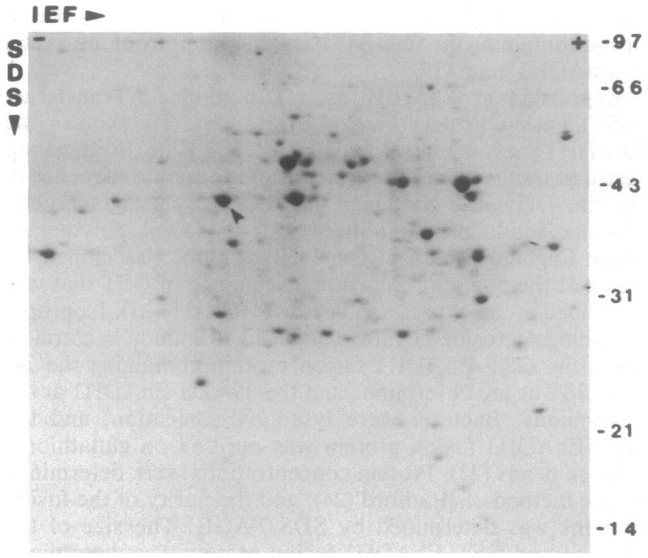


FIG. 1. Representative two-dimensional protein gel of *E. histolytica* proteins stained with Coomassie blue. The arrowhead indicates the ≈39-kDa protein that contained the 16-amino acid sequence (YGATDDIINYKNGDIVE) used to clone the $NADP^+$-dependent amoeba EhADH1 gene. IEF, isoelectric focusing. Positions of molecular mass markers (in kDa) are shown.

from amino acids [GHEA(AV)G] well-conserved in other ADH genes (Fig. 2 and ref. 10). EhADH1 PCR product was 504 bp long and contained an open reading frame that included the 16 amino acids identified in the abundant 39-kDa amoeba protein. Radiolabeled EhADH1 PCR product identified ≈50 plaques on a plate containing 50,000 *E. histolytica* cDNA phage. The cDNA clone pAK311 contained the entire 1080-bp EhADH1 coding sequence, 18 bp 5′ to the start codon, and 42 bp 3′ to the stop codon including an 18-bp poly(A) tail. The long open reading frame of EhADH1 beginning with a methionine was 360 amino acids long and encoded a protein with a calculated molecular weight of 38,709, the same size as the abundant 39-kDa amoeba protein on the two-dimensional protein gels (Fig. 2). Characteristic of previously identified amoeba genes, 92% of the third codons in the open reading frame of EhADH1 were A or T, and 8% were G or C (not counting methionine and tryptophan residues).

**EhADH1 Is a Single-Copy Gene.** When *E. histolytica* genomic DNA was digested with four restriction enzymes that cut once within the coding segment of the EhADH1 and then hybridized with a 349-bp probe from the 3′ end of EhADH1, only one band was seen in each lane (data not shown). This result strongly suggests that EhADH1 is a single-copy gene, as it is very unlikely that multiple copies of ADH1 have four restriction sites in the same location in the 3′ noncoding region.

**The Primary Structure of EhADH1 Is Most Like That of the Anaerobic Bacterium *T. brockii*.** The primary sequence of EhADH1 open reading frame showed a 62% positional identity with the tetrameric $NADP^+$-dependent ADH gene of *T. brockii* (Fig. 2 and ref. 14) and a 19% positional identity with $NAD^+$-dependent horse liver ADH, which has been crystallized (10, 11). EhADH1, like *T. brockii* ADH, contained 18 of the 22 amino acids conserved in all eukaryotic ADHs (in boxes and numbered in Fig. 2, see refs. 10 and 14). EhADH1 contained conserved residues Cys-3 and His-7 involved in binding the catalytic zinc ion (10, 11). EhADH1 contained 11 conserved glycine residues involved in protein folding, including three glycine residues (conserved amino acids 16–18 in Fig. 2) that form the $NADP^+$-binding $\beta$–$\alpha$–$\beta$ fold (10, 11). EhADH1 was different from the horse liver and mammalian ADHs in two ways (10). (*i*) Twenty-three amino acids including four cysteine residues (horse liver ADH residues 97, 100, 103, and 111) that bind a second noncatalytic zinc were missing from EhADH1 and the *T. brockii* ADH. (*ii*) EhADH1, like the *T. brockii* ADH, showed amino acid changes in the nucleotide-binding site previously associated with binding of $NADP^+$ rather than $NAD^+$ (10, 14). In EhADH1, Asp-223 in the horse liver ADH is replaced by an uncharged glycine (marked with an asterisk in Fig. 2) to fit the larger $NADP^+$ moiety, and Asn-225 in the horse liver ADH is replaced by the positively charged arginine (marked with a pound sign in Fig. 2) to bind the negatively charged phosphate. An $NADP^+$-dependent ADH of the anaerobic bacterium *Clostridium beijerinckii* showed a 60% positional identity with the EhADH1, was missing the 23 amino acid residues involved in binding noncatalytic zinc, and had the same changes in the nucleotide binding site as EhADH1 and

```
                                                  1                 3  4
EhADH1 : MKGLAMLGIGRIGWIEKKIP.......ECGFLDALVRPLAL..APCTSDTHTVWAGAIGD 51
TbADH1 : ---F---S--KV-----EK-.......AP--F--I-----V..-------I---FE----E
HLADH1 : STAGKVIKCKAAVLW-E-K-fsieevevAP-KAHE--IKMVATGI-R--D-V-SGTLVTP 60
                678  9      10 11    1213
EhADH1 : RHDMILGHEAVGQIVKVGSLVKRLKVGDKVI.....................VPAITP 88
TbADH1 : --N---------EV-E---E--DF-P--R-V.....................------
HLADH1 : .LPV-A----A-IVESI-EG-TTVRP-----plftpqcgkcrvckhpegnfclkNDLSM- 119

EhADH1 : DWGEEESQRGYPMHSGGMLGGWKFSNFKDGVFSEVFHVNEADANLALLPRDIKPEDAVML 148
TbADH1 : --RTS-V----HQ------A------V-----G-F----D--M---H--KE-PL-A---I
HLADH1 : RGTMQDGTSRFTCRGKPIYHFLGT-T-SQYTVVDEIS AKI--ASP-.....EKVCLIG 173
                              15      1617 18              * #
EhADH1 : SDMVTTGFHGAELANIKIGDTVCVIGIGPVGLMSVAGANHLGAGRIFAVGSRKHCCDIAL 208
TbADH1 : P--M----------D-E--A--A-I--------A----KLR-----I-----PV-V-A-K
HLADH1 : CGFS-GYGSAVKV-KVTQ-S-CA-F-I-G--SVIM-CKAA--A--IG-DIN-DKFAK-K 233
           20                          21
EhADH1 : EYGATDIINYKNGDIVE.QILKATDGKGVDKVVIAGGDVHTFAQAVKMIK.PGSDIGNV 265
TbADH1 : Y--------V---D-P-ES..--MNL-E-----AAI----NADIM-T---IV-.--GT-A--
HLADH1 : -V--ECV-PQDYKKPIQEVLTEMSN-.---FSFEVI-RLD-MVT-LSCCQEAYGVSVI- 292
                                    22
EhADH1 : NYLGEGDNIDIPRSEWGVGMGHKHIHGGLTPGGRVRMEKLASLISTGKLDTSKLITHRFE 325
TbADH1 : --F---EVLPV--L---C--A--T-K---C----L---R-ID-VFYKRV-P---V--V-R
HLADH1 : GVPPDSQ-LSMNPM-LLS-RTW-GAIF-.GFKSKDSVP--VADFMAK-FALDP----VLP 351

EhADH1 : GLEKVEDALMLMKNKPADLIKPVVRIHYDDEDTLH* 360
TbADH1 : -FDNI-K-F----D--K-------ILA*
HLADH1 : .F--INEGFD-LRS..GES-RTILTF*          374
```

FIG. 2. Alignment of the primary sequences of the ADHs of *E. histolytica* (EhADH1), *T. brockii* (TbADH), and horse liver (HLADH). Dashes mark positional identities of the *T. brockii* and horse liver ADHs with the EhADH1 sequence, periods mark absent amino acids, and an asterisk marks the stop codons. The 16-amino acid sequence from the peptide fragment of the major 39-kDa protein on two-dimensional protein gels is identified with a horizontal box. Amino acids corresponding to the primers used to make the 504-bp EhADH1 PCR product are underlined. Eighteen of the 22 amino acids conserved in most ADHs (10) are marked with vertical boxes and numbered from 1 to 22 (conserved residues 2, 5, 14, and 19 are missing from EhADH1). EhADH1 amino acids involved in binding $NADP^+$ rather than $NAD^+$ in EhADH1 are marked with an asterisk (Gly-198) and a pound sign (Arg-200).

*T. brockii* ADH (J.-S. Chen, Virginia Polytechnic Institute and State University, personal communication).

The hydropathy plot (26) of the EhADH1 open reading frame was nearly superimposable onto that of *T. brockii* ADH and was similar to that of the horse liver ADH (data not shown). The similarities of the hydropathy plots suggest that EhADH1 may have a tertiary structure like that of the crystallized horse liver ADH (11), even though the positional identity of the two enzymes is low.

**Phylogenetic Tree of ADHs.** EhADH1 was aligned with 20 other ≈40-kDa (long chain) ADHs present in the GenBank data base, and a phylogenetic tree was drawn (Fig. 3 and ref. 27). EhADH1 was most closely related to the *T. brockii* and *C. beijerinckii* ADHs, both of which are derived from an anaerobic organisms and are $NADP^+$-dependent (14). The next best match was with the $NAD^+$-dependent *Alcaligenes eutrophus* ADH, which is tetrameric and missing three of four cysteine residues that bind the noncatalytic zinc. EhADH1 was equally distant from two other bacterial ADHs (*Escherichia coli* and *Zymomonas mobilis*) and three fungal ADHs. EhADH1 was very distant from seven mammalian ADHs and bird, frog, and plant ADHs (Fig. 3). EhADH1 showed a 15% positional identity with the most similar human ADH (10).

**Expression of EhADH1 as a GST Fusion Protein in Bacteria.** The GST–EhADH1 fusion protein, composed of the 26-kDa GST and the 39-kDa EhADH1, had the expected 65-kDa apparent molecular mass by SDS/PAGE (data not shown). Both the bacterial lysate and the purified GST–EhADH1 rapidly reduced $NADP^+$ after addition of isopropanol, while the bacterial lysates containing GST alone and the purified GST failed to reduce $NADP^+$. The ADH activity of the purified GST–EhADH1 fusion enzyme with isopropanol and $NADP^+$ averaged 29 units/mg of protein (Table 1). The turnover number of the fusion enzyme ($K_{cat}$) was 1885 mol of isopropanol oxidized per mol of enzyme per min. The enrichment in activity observed after affinity purification on the glutathione-agarose beads was 11-fold with a yield of 51% of the total activity. The average $K_m$ values were 52 $\mu$M for isopropanol, 5 mM for ethanol, 31 $\mu$M for $NADP^+$, and >10 mM for $NAD^+$ (Table 1). The specificities of the enzyme, represented by the ratio of $K_{cat}$ to $K_m$, were $3.6 \times 10^7$ $M^{-1}\cdot min^{-1}$ for isopropanol and $6.1 \times 10^7$ $M^{-1}\cdot min^{-1}$ for $NADP^+$. The ADH activity of the enzyme with ethanol and $NADP^+$ was 1.9 units/mg of protein ($K_{cat}$ = 124 $min^{-1}$ and $K_{cat}/K_m = 2.5 \times 10^4$ $M^{-1}\cdot min^{-1}$). These values are compared in Table 1 to those reported (7) for the purified native amoeba $NADP^+$-dependent ADH. On FPLC, the glutathione-agarose-purified GST–EhADH1 fusion showed protein and activity peaks of >300 kDa, suggesting that the recombinant fusion enzyme formed tetramers [like the native amoeba $NADP^+$-dependent ADH (7)].

The acetaldehyde reductase activity of the purified GST–EhADH1 with NADPH averaged 49 units/mg of protein ($K_{cat}$ = 3185 $min^{-1}$), whereas GST alone had no acetaldehyde reductase activity. The $K_m$ of GST–EhADH1 averaged 10 $\mu$M for acetaldehyde and 91 $\mu$M for NADPH, so that the specificity constants $K_{cat}/K_m$ were $3.2 \times 10^8$ $M^{-1}\cdot min^{-1}$ for isopropanol and $3.5 \times 10^7$ $M^{-1}\cdot min^{-1}$ for NADPH. The acetaldehyde reductase activity, like the ADH activity of the GST–EhADH1 fusion protein, was greatly reduced when $NAD^+$ was substituted for $NADP^+$. The acetone reductase activity of the enzyme with NADPH was 70 units/mg ($K_{cat}$ = 4550 $min^{-1}$, the $K_m$ for acetone was 20 $\mu$M, and the specificity constant activity $K_{cat}/K_m$ was $3.5 \times 10^8$ $M^{-1}\cdot min^{-1}$). The acetaldehyde reductase activity of the recombinant GST–ADH fusion protein was similar to that previously reported for the purified native *E. histolytica* $NADP^+$-dependent ADH, whereas the $K_m$ values of the fusion enzyme for NADPH, acetaldehyde, and acetone were somewhat different (Table 1 and ref. 7).

**Inhibition of ADH Activity of GST–EhADH1 by Pyrazole and 4-Methylpyrazole.** Human ADHs are inhibited by pyrazole and 4-methylpyrazole (12, 13). The $K_i$ values of pyrazole and 4-methylpyrazole for the ADH activity of GST–EhADH1 fusion protein were 1.4 ± 0.7 $\mu$M and 21 ± 5 $\mu$M, respectively.

## DISCUSSION

The EhADH1 gene that we have cloned likely encodes the previously purified tetrameric $NADP^+$-dependent *E. his-*

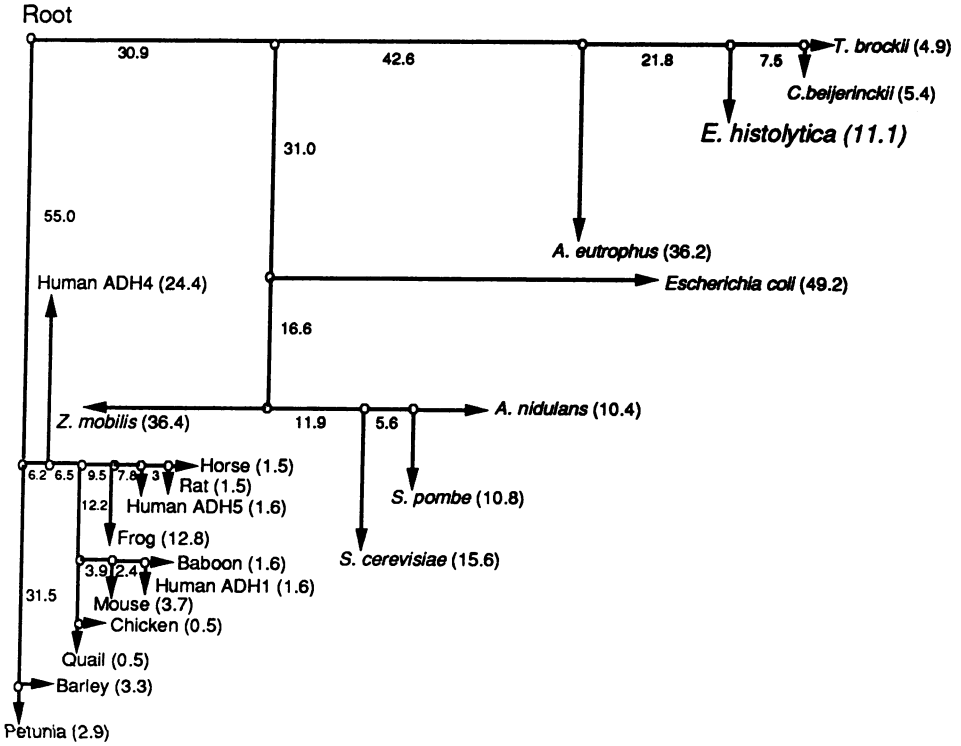


FIG. 3. Long-chain ADH phylogenetic tree generated by the progressive alignment method of Feng and Doolittle (27). Sequences were taken from GenBank. EhADH1 was aligned with five bacterial ADHs (*T. brockii, Alcaligenes eutrophus, Escherichia coli, C. beijerinckii*, and *Z. mobilis*), three fungal ADHs (*Saccharomyces cerevisiae, Schizosaccharomyces pombe*, and *Aspergillus nidulans*), three human ADHs [class I (ADH1), class II (ADH4), and class III (ADH5)], horse, mouse, rat, baboon, frog, bird, and plant ADHs. Open circles indicate nodes. The root node represents an average of all aligned sequences. Arrows indicate terminal sequences. Numbers indicate distance to previous node.

Table 1. Comparison of the NADP$^+$-dependent ADH activities of the recombinant fusion protein GST–EhADH1 and native ADH

| Parameter | GST–EhADH1 | Native ADH |
|---|---|---|
| ADH | | |
| Activity (isopropanol), units/mg | 29 ± 10 | ND |
| $K_m$ (isopropanol), μM | 52 ± 19 | 240 |
| $K_m$ (NADP$^+$), μM | 31 ± 11 | 3.5 |
| $K_m$ (NAD$^+$), mM | >10 | ND |
| Activity (ethanol), μM | 1.9 ± 1 | ND |
| $K_m$ (ethanol), mM | 5 ± 1 | 4 |
| Acetaldehyde and acetone reductase | | |
| Activity (acetaldehyde), units/mg | 49 ± 3 | 43 |
| $K_m$ (acetaldehyde), μM | 10 ± 4 | 53 |
| $K_m$ (NADPH), μm | 91 ± 27 | 40 |
| Activity (acetone), units/mg | 70 ± 32 | ND |
| $K_m$ (acetone), μM | 20 ± 1 | 38 |

Values for the native ADH are those reported by Lo and Chang (7).

*tolytica* ADH (7), because the recombinant GST–EhADH1 fusion protein like the natural enzyme preferred NADP$^+$ to NAD$^+$, preferred isopropanol to ethanol, and was inhibited by pyrazole. Consistent with its presumed but not proven function during fermentation of pyruvate to ethanol, GST–EhADH1 also reduced acetaldehyde in the presence of NADPH (7). However, the recombinant enzyme showed a high affinity and activity with acetone, suggesting that the parasite enzyme may reduce multiple substrates *in vivo*. Differences in Michaelis constants of the recombinant ADH enzyme vs. those of the purified native amoeba enzyme may be derived from different methods and machines used for determining these constants or may reflect real differences between the GST–ADH fusion enzyme and the native amoeba ADH. Differences in the length of the EhADH1 open reading frame and the size of the abundant protein spot on the two-dimensional protein gels (both ≈39 kDa) vs. that reported for the native amoeba enzyme [≈30 kDa (7)] may reflect a previous underestimate of size by SDS/PAGE in tube gels.

The primary structure of EhADH1 enzyme was most similar to the NADP$^+$-dependent ADHs of the anaerobic bacteria *T. brockii* and *C. beijerinckii*, with which it shares 62% and 60% amino acid identities, respectively (14). Similarly, the 6-kDa amoeba ferredoxin (28) is most like the 6-kDa ferredoxins of *Clostridium* and quite different from the 12-kDa human ferredoxin. EhADH1 showed a 15% amino acid identity with the human ADHs, which are dimeric and NAD$^+$-dependent (10). Whether these differences can be used to identify ADH inhibitors that selectively inhibit the amoeba enzyme remains to be determined.

Southern blots suggested that the EhADH1 gene is present as a single copy. Previously, the gene for a 120-kDa surface antigen was shown to be present as a single copy (29), whereas many other genes important for *E. histolytica* function are present in multiple copies including actin (five), P-glycoprotein (three), ferredoxin (two), and the Gal/GalNAc adherence lectin (two).

A segment of the amoeba EhADH1 gene was identified using the PCR and degenerate oligonucleotides made to a sequence of amino acids conserved in other ADHs and to a sequence from the ≈39-kDa protein identified on two-dimensional protein gels. This PCR method circumvents the peculiar A+T-rich codon usage of the parasite that makes it difficult to use heterologous DNA probes.

The substrate analogues pyrazole and 4-methylpyrazole inhibited the ADH activity of the recombinant GST–EhADH1 fusion protein. Inhibition of ADH activity with pyrazole has been reported for the native *E. histolytica* NADP$^+$-dependent ADH (7). There will be a better estimate of the efficacy of ADH inhibitors as amoebacidal drugs when a number of important questions concerning the amoeba ADHs are answered. For example, what is the relative importance of the NADP$^+$-dependent EhADH1 vs. the other NAD$^+$-dependent ADH previously identified in *E. histolytica*? What is the efficacy of ADH inhibitors against parasites growing anaerobically in the colon or microaerophilically within tissue abscesses? The cloning and expression of the *E. histolytica* NADP$^+$-dependent ADH is a first step to answering these questions and to finding additional antiamoebic drugs to complement those already used to treat human infection with *E. histolytica*.

Thanks to Dr. J.-S. Chen for use of his *C. beijerinckii* ADH sequence. Helpful advice concerning these experiments came from Drs. James Tobin and Barton Holmquist. This work was supported by grants from the National Institutes of Health, U.S. Agency for International Development, and the John D. and Catherine T. MacArthur Foundation. Additional support came from the Harvard Digestive Diseases Center. J.S. is an American Gastroenterology Association/Rorer Consumer Pharmaceuticals Research Scholar.

1. Walsh, J. A. (1986) *Rev. Infect. Dis.* **8**, 228–238.
2. Martinez-Palomo, A. & Martinez-Baez, M. (1985) *Rev. Infect. Dis.* **5**, 1093–1102.
3. Muller, M. (1986) *Biochem. Pharmacol.* **35**, 37–41.
4. Reeves, R. E. (1984) *Adv. Parasitol.* **23**, 105–142.
5. Muller, M. (1988) *Annu. Rev. Microbiol.* **42**, 465–488.
6. Reeves, R. E., Montalvo, F. E. & Lushbaugh, T. S. (1971) *Int. J. Biochem.* **2**, 55–64.
7. Lo, H.-S. & Chang, C.-J. (1982) *J. Parasitol.* **68**, 372–377.
8. Jornvall, H. (1985) *Alcohol* **2**, 61–66.
9. Deetz, J. S., Luehr, C. A. & Vallee, B. L. (1984) *Biochemistry* **23**, 6822–6828.
10. Jornvall, H., Persson, B. & Jeffrey, J. (1987) *Eur. J. Biochem.* **167**, 195–201.
11. Eklund, H., Samama, J.-P., Wallen, L., Branden, C.-I., Akeson, A. & Jones, A. J. (1981) *Mol. Biol.* **146**, 561–587.
12. Li, T. K. & Theorell, H. (1969) *Acta Chem. Scand.* **23**, 892–902.
13. Horjales, E., Eklund, H. & Branden, C. (1987) *J. Mol. Biol.* **197**, 685–694.
14. Peretz, M. & Burstein, Y. (1989) *Biochemistry* **28**, 6549–6555.
15. Diamond, L. S., Harlow, D. R. & Cunnick, C. C. (1978) *Trans. R. Soc. Trop. Med. Hyg.* **73**, 431–432.
16. Knowles, R. W. (1987) in *Histocompatibility Testing*, ed. Dupont, B. (Springer, New York).
17. Kennedy, T. E., Gawinowicz, M. A., Barzilai, A., Kandel, E. R. & Sweatt, J. D. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 7008–7012.
18. Hewick, R. M., Hinkapiller, M. W., Hood, L. E. & Dyer, W. J. (1981) *J. Biol. Chem.* **256**, 7990–7997.
19. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
20. Lee, C. C., Wu, X., Gibbs, R. A., Cook, R. G., Muzny, D. M. & Caskey, C. T. (1988) *Science* **239**, 1288–1291.
21. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
22. Altschul, S. F., Gish, W., Myers, E. W. & Lipman, D. J. (1990) *J. Mol. Biol.* **215**, 403–410.
23. Smith, D. & Johnson, K. (1988) *Gene* **67**, 31–40.
24. Bradford, M. (1976) *Anal. Biochem.* **72**, 248–254.
25. Dixon, M. & Webb, E. C. (1979) *Enzymes* (Academic, New York).
26. Kyte, J. & Doolittle, R. F. (1982) *J. Mol. Biol.* **157**, 105–132.
27. Feng, D.-F. & Doolittle, R. F. (1987) *J. Mol. Evol.* **25**, 351–360.
28. Huber, M., Garfinkel, L., Gitler, C., Mirelman, D., Revel, M. & Rozenblatt, S. (1988) *Mol. Biochem. Parasitol.* **31**, 27–34.
29. Edman, U., Meraz, M. A., Rausser, S., Agabian, N. & Meza, I. (1990) *J. Exp. Med.* **172**, 879–888.