JOURNAL OF BACTERIOLOGY, June 1997, p. 3736–3745 Vol. 179, No. 11
0021-9193/97/$04.00+0


# Evidence for the Bacterial Origin of Genes Encoding Fermentation Enzymes of the Amitochondriate Protozoan Parasite *Entamoeba histolytica*

BEN ROSENTHAL,[1] ZHIMING MAI,[1] DANIEL CAPLIVSKI,[1] SUDIP GHOSH,[1] HUMBERTO DE LA VEGA,[2] THOMAS GRAF,[3] AND JOHN SAMUELSON[1]*

*Department of Tropical Public Health, Harvard School of Public Health, Boston, Massachusetts 02115[1]; Center for Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139[2]; and BioMolecular Engineering Research Center, Boston University, Boston, Massachusetts 02215[3]*

Received 18 December 1996/Accepted 21 March 1997

***Entamoeba histolytica*** **is an amitochondriate protozoan parasite with numerous bacterium-like fermentation enzymes including the pyruvate:ferredoxin oxidoreductase (POR), ferredoxin (FD), and alcohol dehydrogenase E (ADHE). The goal of this study was to determine whether the genes encoding these cytosolic *E. histolytica* fermentation enzymes might derive from a bacterium by horizontal transfer, as has previously been suggested for *E. histolytica* genes encoding heat shock protein 60, nicotinamide nucleotide transhydrogenase, and superoxide dismutase. In this study, the *E. histolytica por* gene and the *adhE* gene of a second amitochondriate protozoan parasite, *Giardia lamblia*, were sequenced, and their phylogenetic positions were estimated in relation to POR, ADHE, and FD cloned from eukaryotic and eubacterial organisms. The *E. histolytica por* gene encodes a 1,620-amino-acid peptide that contained conserved iron-sulfur- and thiamine pyrophosphate-binding sites. The predicted *E. histolytica* POR showed fewer positional identities to the POR of *G. lamblia* (34%) than to the POR of the enterobacterium *Klebsiella pneumoniae* (49%), the cyanobacterium *Anabaena* sp. (44%), and the protozoan *Trichomonas vaginalis* (46%), which targets its POR to anaerobic organelles called hydrogenosomes. Maximum-likelihood, neighbor-joining, and parsimony analyses also suggested as less likely *E. histolytica* POR sharing more recent common ancestry with *G. lamblia* POR than with POR of bacteria and the *T. vaginalis* hydrogenosome. The *G. lamblia adhE* encodes an 888-amino-acid fusion peptide with an aldehyde dehydrogenase at its amino half and an iron-dependent (class 3) ADH at its carboxy half. The predicted *G. lamblia* ADHE showed extensive positional identities to ADHE of *Escherichia coli* (49%), *Clostridium acetobutylicum* (44%), and *E. histolytica* (43%) and lesser identities to the class 3 ADH of eubacteria and yeast (19 to 36%). Phylogenetic analyses inferred a closer relationship of the *E. histolytica* ADHE to bacterial ADHE than to the *G. lamblia* ADHE. The 6-kDa FD of *E. histolytica* and *G. lamblia* were most similar to those of the archaebacterium *Methanosarcina barkeri* and the δ-purple bacterium *Desulfovibrio desulfuricans*, respectively, while the 12-kDa FD of the *T. vaginalis* hydrogenosome was most similar to the 12-kDa FD of γ-purple bacterium *Pseudomonas putida*. *E. histolytica* genes (and probably *G. lamblia* genes) encoding fermentation enzymes therefore likely derive from bacteria by horizontal transfer, although it is not clear from which bacteria these amebic genes derive. These are the first nonorganellar fermentation enzymes of eukaryotes implicated to have derived from bacteria.**

The endosymbiont hypothesis or serial endosymbiosis theory, one of the most exciting ideas in evolutionary biology, postulates that the mitochondria and chloroplasts of eukaryotic cells derive from an endocytosed α-purple bacterium and cyanobacterium, respectively (Fig. 1) (18, 19, 35, 63). Although bacterial rRNA genes as well as those encoding some structural proteins and enzymes remain in mitochondria and chloroplasts, many of the bacterially derived genes have relocated over time from the organelles to the eukaryotic nucleus (20). Signal sequences at the amino termini of nucleus-encoded proteins of bacterial origin target them to mitochondria and chloroplasts (49). The eukaryotic nucleus then is chimeric, made up of (i) genes from a eukaryotic ancestor encoding proteins involved in protein synthesis, growth control and replication, morphogenesis, and secretion and (ii) genes from bacterial ancestors that encode organellar proteins involved in energy metabolism, including oxidative phosphorylation (mitochondria) and photosynthesis (chloroplasts), and transport of proteins into the organelles, including heat shock proteins (e.g., HSP10 and HSP60).

*Entamoeba histolytica* is a protozoan parasite that consumes bacteria in the colonic lumen and causes dysentery when it invades the colonic wall and lyses epithelial cells and erythrocytes (43). In both locations, *E. histolytica* is an obligate fermenter, as it lacks mitochondria and enzymes of oxidative phosphorylation (6, 39, 44). However, several observations suggest the lineage giving rise to *E. histolytica* once possessed mitochondria and acquired other bacterial genes via endosymbiosis or transfection (Fig. 1). First, phylogenies based on small-subunit rRNA genes suggest that *E. histolytica* shared with protozoa of the apicomplexa and kinetoplastida a common ancestor that possessed mitochondria (52). Second, the gene trees that implicate α-purple bacteria as the source for mitochondria also indicate that the *E. histolytica* HSP60 is characteristic of those HSP60 belonging to mitochondrion-possessing eukaryotes (13). Third, *E. histolytica* contains a gene encoding a nicotinamide nucleotide transhydrogenase (NNT;

* Corresponding author. Mailing address: Department of Tropical Public Health, Harvard School of Public Health, 665 Huntington Ave., Boston, MA 02115. Phone: (617) 432-4670. Fax: (617) 739-4914. E-mail: descot@mbcrr.harvard.edu.

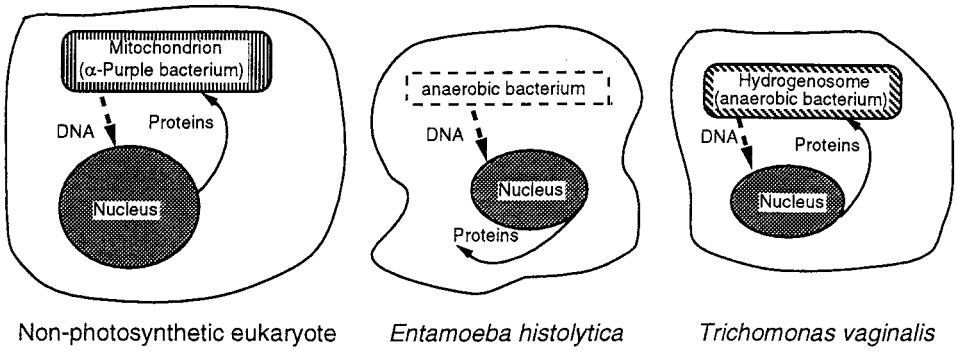



FIG. 1. Cartoon outlining endosymbiont hypotheses for mitochondria of nonphotosynthetic eukaryotes, *E. histolytica* fermentation enzymes, and *T. vaginalis* hydrogenosome. Dotted arrows refer to evolutionary movement of genes from endosymbiont bacteria to the eukaryotic nuclei. Solid arrows refer to targeting of nucleus-encoded proteins to organelles (nonphotosynthetic eukaryote and *T. vaginalis*) or to cytosol and nonorganellar membranes (*E. histolytica*). Note that it is also possible that the *E. histolytica* genes were acquired via bacterial plasmids or transposons rather than by a bacterial endosymbiont.

46% positional identity with *Escherichia coli* NNT), which has previously been associated only with mitochondria and bacteria (13, 68). The structure of the *E. histolytica* NNT, which lacks amino-terminal mitochondrion-targeting sequences, parallels that of bacterial rather than bovine mitochondrial NNT in that this fusion protein has homologs of the bacterial α and β peptides at the carboxy and amino halves, respectively. In contrast, the bovine mitochondrion NNT is a fusion protein with homologs of the bacterial α and β peptides in reverse order, at the amino and carboxy halves, respectively. Fourth, the *E. histolytica* gene encoding the iron-dependent superoxide dismutase (SOD; 57% identity with that of *Pseudomonas* sp.) also appears to be derived by horizontal transfer from bacteria (50, 55). Fifth, *E. histolytica* contains two bacterium-like pyrophosphate ($PP_i$)-dependent glycolytic enzymes, a phosphofructokinase (PPFK) and pyruvate phosphate dikinase, which substitute for ATP-dependent glycolytic enzymes present in most other eubacteria and eukaryotes (3, 8, 26). Sixth, another amitochondriate protozoan parasite, *Trichomonas vaginalis*, appears to contain genes derived from a bacterial endosymbiont (Fig. 1) (25, 28, 39, 40). *T. vaginalis* has a double-membrane-bound organelle called a hydrogenosome that contains numerous fermentation enzymes. The hydrogenosome appears to be endosymbiont derived because the *T. vaginalis* HSP60 groups with mitochondrial HSP60 and nucleus-encoded hydrogenosome proteins have leader sequences similar to those of mitochondrial proteins (11, 20, 24, 25).

We wondered whether other *E. histolytica* genes encoding bacterium-like fermentation enzymes, which resemble those of the *T. vaginalis* hydrogenosome and the amitochondriate protozoan parasite *Giardia lamblia* (which causes duodenitis), might also be derived by horizontal transfer from bacteria (Fig. 1 and 2) (1, 8, 9, 17, 22, 25–28, 31, 39, 41, 44, 45, 48, 53, 56, 57, 61, 64). The pyruvate:ferredoxin oxidoreductase (POR) catalyzes the oxidative decarboxylation of pyruvate to acetyl coenzyme A (acetyl-CoA), utilizing ferredoxin (FD) (or flavodoxin in the case of *nifJ*-encoded proteins of enterobacteria) as an electron acceptor (4). Metronidazole, a potent drug against *E. histolytica*, *G. lamblia*, *T. vaginalis*, and anaerobic bacteria, is activated when it receives an electron from FD that was reduced during the reaction catalyzed by the POR (39). Humans, most eukaryotes, and most eubacteria fail to activate metronidazole, because pyruvate is converted to acetyl-CoA via pyruvate dehydrogenase, producing NADH rather than reduced FD (60).

*E. histolytica* converts acetyl-CoA to acetaldehyde via a CoA-dependent acetaldehyde dehydrogenase, which is at the amino half of a fusion protein called alcohol dehydrogenase E (ADHE) (9, 44, 64). *E. histolytica* converts acetaldehyde to ethanol via one of three ADH: ADHE, a zinc-dependent ADH1, or an iron-dependent ADH3 (31, 45). As all three *E. histolytica* ADH are absent from humans, these enzymes may be targets for new antiamebic drugs.

To determine how similar *E. histolytica* fermentation enzymes were to those of *G. lamblia*, *T. vaginalis*, and anaerobic bacteria, molecular cloning techniques were used to obtain the *E. histolytica por* gene and the *G. lamblia adhE* gene (32). The phylogenetic positions of POR, ADHE, and FD sequences were estimated by using maximum-parsimony (P), neighbor-joining (NJ), and maximum-likelihood (ML) methods to determine whether *E. histolytica* and *G. lamblia* fermentation

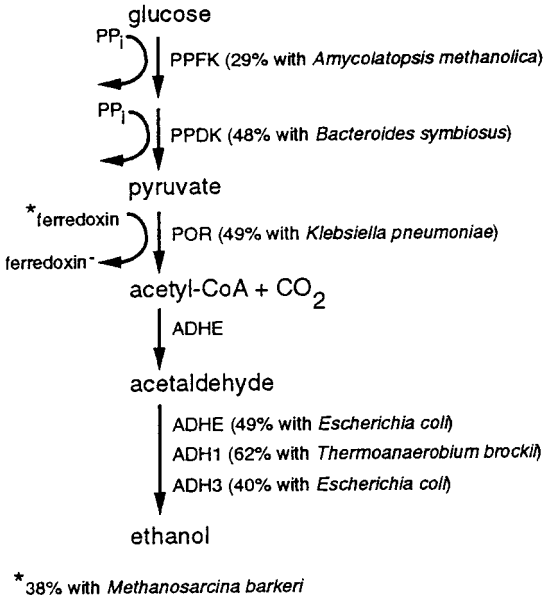



FIG. 2. Bacterium-like enzymes of *E. histolytica* involved in metabolism of glucose to ethanol include PPFK (EC 2.7.1.90; GenBank accession no. U12513), pyruvate phosphate dikinase (PPDK; X74596), POR (EC 1.2.7.1; U30149), FD (J03870), ADHE (EC 1.1.1.1; X77132), ADH1 (EC 1.1.1.2; M88600), and ADH3 (EC 1.1.1.1; Z48752) (8, 9, 25, 26, 31, 44, 45, 64). Positional identities of each enzyme with those of the most similar bacterium are shown in parentheses.

proteins may have derived from a common eukaryotic ancestor or whether instead they may have been acquired from bacteria by independent horizontal transfers, as is believed to have occurred with enzymes associated with the *T. vaginalis* hydrogenosome (14–16, 21).

## MATERIALS AND METHODS

**Cloning of the *E. histolytica por* gene.** A 408-bp segment of the *E. histolytica por* gene was isolated from *E. histolytica* HM-1:IMSS DNA by using PCR and degenerate primers to conserved POR peptides (Fig. 3) (32). The sense PCR primer to IANATGC was AT(AT)GC(AT)AA(TC)GC(AT)AC(AT)GG(AT)TG. The antisense PCR primer to GGDGWAY, which includes the Gly-Asp-Gly tripeptide conserved in thiamine $PP_i$ (TPP)-binding sites, was TA(AT)ATCCA(AT)CC(AG)TC(AT)CC(AT)CC (4, 5, 23, 25, 29). The *E. histolytica por* PCR product was used to screen cDNA and genomic DNA (gDNA) libraries of *E. histolytica* HM-1:IMSS. The PCR product, three cDNA clones, and three genomic DNA clones were identical in overlapping regions.

**Cloning of the *G. lamblia adhE* gene.** A 575-bp segment of the *G. lamblia adhE* gene, which was identified by a search of GenBank peptide sequences at the National Center for Biotechnology Information by using BLAST, was isolated serendipitously, using *G. lamblia* DNA and PCR primers designed for other purposes (2). This PCR product was used to obtain the entire *adhE* gene from a gDNA library of *G. lamblia*. Five *G. lamblia adhE* gDNA clones were identical in overlapping regions. Interestingly, 44 bp downstream from the stop codon of the *G. lamblia adhE* was an open reading frame (ORF) starting with Met that encoded a peptide of at least 171 amino acids, for which no match was found in GenBank (data not shown).

**Phylogenetic analyses.** Sequences similar to the *E. histolytica* POR, ADHE, and FD sequences were obtained from GenBank by using BLAST (2). These sequences were aligned by using PIMA2, and regions of ambiguous alignment were removed by using MASE (15, 51). Three independent means were used to reconstruct the phylogeny of these proteins, each employing the Dayhoff matrix of amino acid substitution probabilities (14). Unrooted networks (trees) of minimum total length were constructed by the NJ method, using the protdist algorithm of PHYLIP (15), which groups taxa according to the proportion of amino acid positions that are shared or of similar state. Additionally, unrooted P trees based on a cladistic approach, which minimizes the number of amino acid substitutions required to produce the observed alignment given a hypothesized common ancestor, were constructed by using PHYLIP's protpars program. NJ and P algorithms were each applied to 100 bootstrap replicates of the alignment to investigate the robustness of the inferred phylogeny to sampling variance. A phenogram resolving only those nodes present in a majority of both NJ and P methods was constructed. The log-likelihood of this tree (the probability of the series of amino acid changes implied by this tree) was calculated by using an ML method (PROTML) (21). Using the star decomposition algorithm in PROTML, we failed to identify trees with significantly greater likelihood than the consensus NJ/P tree. Accepting this tree as a reliable, conservative representation of the inferred phylogeny, we derived branch lengths of this topology by employing PROTML and drew this tree, labeling each node by the frequency with which it appeared in the replicated NJ and P analyses. Any unusual case where one analysis identified >50% bootstrap support for a node not so indicated by the other analysis, and where this difference could affect the interpretation of the relationship between fermentation enzymes in eukaryotes, is noted in the text.

**Nucleotide sequence accession numbers.** Nucleotide and derived amino acid sequences for the *E. histolytica por* gene and the *G. lamblia adhE* gene have been submitted to GenBank with accession no. U30149 and U93353, respectively.

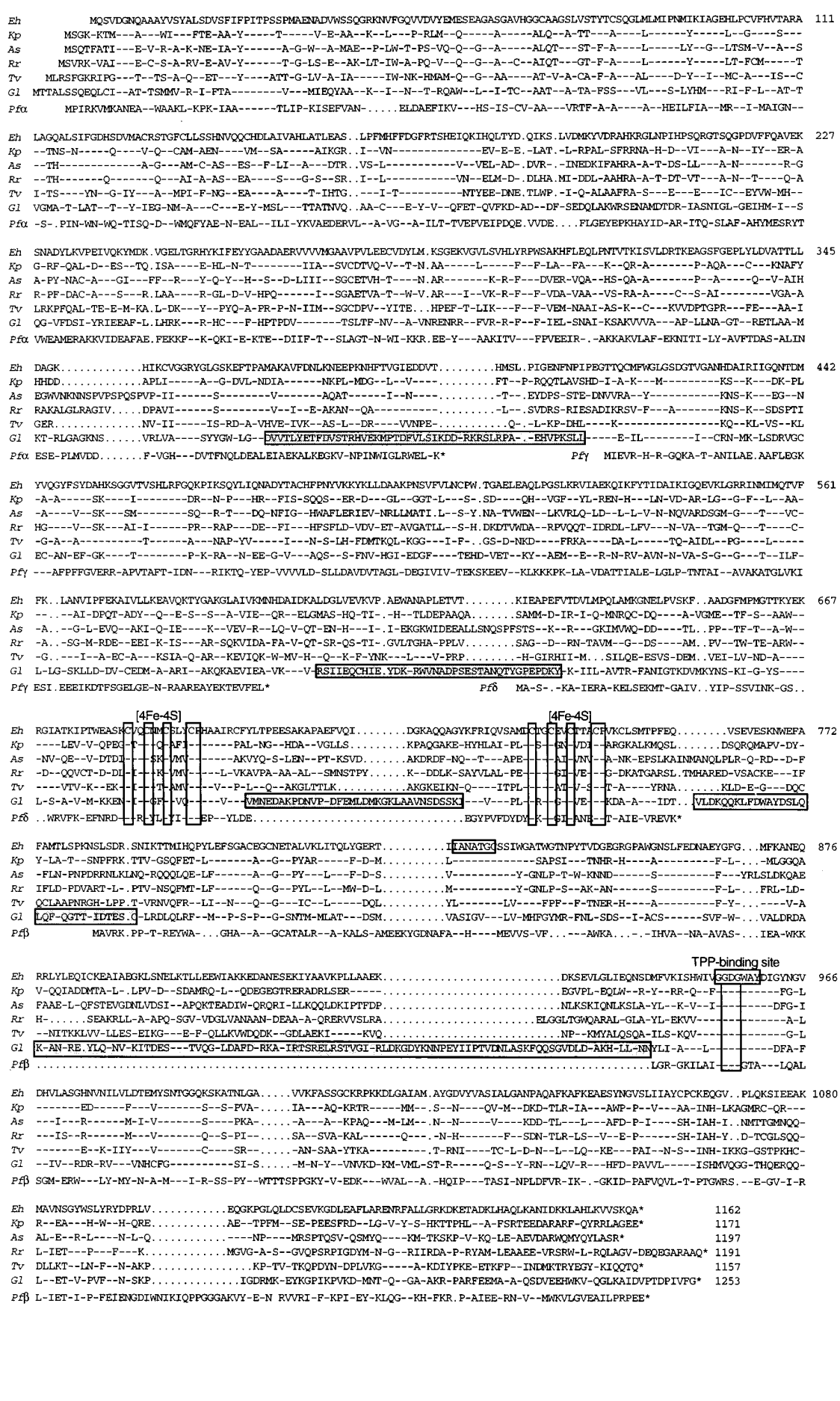
## RESULTS

**Primary structure of the *E. histolytica* POR.** The *E. histolytica por* gene encoded a 1,620-amino-acid peptide ($M_r$, 127,492), the expected size for eubacterial and protozoan POR (Fig. 3). The Met residue chosen as the start codon was likely correct because there was an in-frame stop codon before the next ATG in the 5′ untranslated region, and the next potential start codon was not until $Met^{31}$, well into the overlap with other POR (4, 25, 29). Unlike the *T. vaginalis* POR, the *E. histolytica* POR lacked mitochondrion- or hydrogenosome-targeting sequences at its amino terminus (25). As is typical of the 5′ untranslated sequences of other *E. histolytica* genes, there was a putative ribosome-binding sequence (ATTC) beginning at −10 and a TATA box at −40 (TATTTAA), no introns were present, and 88% of the third positions in the codons were A or T (excluding Met and Trp) (7). The *E. histolytica por* gene might be single copy, because overlapping sequences were identical among PCR products, cDNA clones, and gDNA clones.

The *E. histolytica* POR was used to search the database in GenBank, and 10 other protozoan, archaebacterial, and eubacterial POR (*nifJ* gene products) were identified (Fig. 3) (2, 4, 25, 29). The *E. histolytica* POR closely resembled the POR of the nitrogen-fixing γ-purple bacterium *Klebsiella pneumoniae* (49% positional identity), the *T. vaginalis* hydrogenosome (46%), the cyanobacterium *Anabaena* sp. (46%), and the α-purple bacterium *Rhodospirillum rubrum* (44%) (51). The *E. histolytica* POR was less similar to the POR of *G. lamblia* (32%) and the archaebacterium *Pyrococcus furiosus* (23%), which is composed of four peptides (29). Similarities between POR were in the putative TPP-binding center located in the β subunit of the *P. furiosus* POR and in the 2[4Fe-4S]-binding center located in the δ subunit of the *P. furiosus* POR (5, 23). The connector regions between and within subunits of the *P. furiosus* POR were similar for *E. histolytica*, *T. vaginalis*, *K. pneumoniae*, *Anabaena* sp., and *R. rubrum*. In contrast, the *G. lamblia* connector regions were longer than those of the other eubacterial and eukaryotic POR and were unrelated in their sequences (Fig. 3).

**Phylogeny of the *E. histolytica* POR.** ML, NJ, and P trees were drawn from an alignment of eight eukaryotic and eubacterial POR, from which connector regions and any regions lacking easy alignment were removed, leaving 967 amino acids (77% of its total) for comparison (14–16, 21) (Fig. 4). Each method identified a clade comprising the POR of *K. pneumoniae* and a second γ-purple bacterium, *Enterobacter agglomerans*. The enterobacterial POR were part of a larger clade that included the POR of *E. histolytica*, *G. lamblia*, and *T. vaginalis*. This larger clade was distinct from a third clade composed of the POR of *Anabaena* sp., a second cyanobacterium (*Synechocystis* sp.), and *R. rubrum*. The node linking the members of the protozoan and enterobacterial POR clade in Fig. 4 was collapsed and marked with an asterisk, because the relationship between these POR was unresolved (no ≥50% majority in NJ and P analyses). However, implied monophyly of *E. histolytica* and *G. lamblia* POR occurred in only 6.5% (ML), 3% (NJ), and 7% (P) of bootstrap replicates. In contrast, monophyly of the *E. histolytica* and eubacterial POR was inferred in 48% of the ML replicates, while monophyly of the *E. histolytica* and *T. vaginalis* hydrogenosome POR was inferred in 34% (NJ) and 40% (P) of the bootstrap replicates. These results therefore provide no evidence to support the belief that *E. histolytica* and *G. lamblia por* genes derive from a common eukaryotic ancestor. While not definitive, these results suggest that the *E. histolytica por* gene may more likely derive from an enterobacterial donor or from the same (as yet unidentified) donor as the *T. vaginalis* hydrogenosome *por* gene.

FIG. 3. Alignment (single-letter code) of the *E. histolytica* (*Eh*) POR (*nifJ* gene products) with those of *K. pneumoniae* (*Kp*; GenBank accession no. X13303), *Anabaena* sp. (*As*; L14925), *R. rubrum* (*Rr*; X77515), *T. vaginalis* (*Tv*; U16823), *G. lamblia* (*Gl*; L27221), and *P. furiosus* (*Pf*; X82850) (4, 25, 29, 57). Dashes mark amino acid identities with *E. histolytica* POR, and periods mark gaps. Unshaded horizontal boxes mark locations of degenerate oligonucleotides used to amplify a segment of the *E. histolytica por* gene, while shaded boxes mark unique connector regions in the *G. lamblia* POR. Unshaded vertical boxes mark 2[4Fe-4S]-binding sites, while the lightly shaded vertical box marks a TPP-binding site (5, 23). A second *E. histolytica por* sequence recently submitted to GenBank (accession no. Z50193) predicted ~56 different amino acids, many of which appeared to have been secondary to sequencing errors (46).


[4Fe-4S]
[4Fe-4S]
TPP-binding site

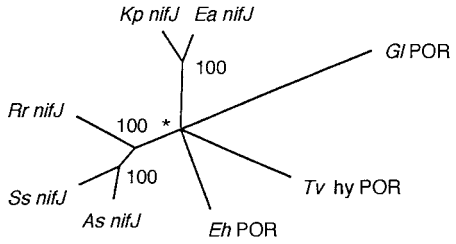



FIG. 4. Fifty percent phenogram of POR (*nifJ* gene products) of eubacteria and eukaryotes shown in Fig. 3, as well as those of *E. agglomerans* (*Ea*; GenBank accession no. X78558) and *Synechocystis* sp. (*Ss*; D64005). Branch lengths in this composite, unrooted tree derive from the ML analysis, while bootstrap values at nodes derive from NJ trees and P trees (both were 100 at the nodes marked) (14–16, 21). Asterisks mark nodes where either NJ or P analysis failed to identify the same bifurcating branch in over 50% of the bootstrap replicates. For abbreviations not given above, see the legend to Fig. 3.

**Primary structure of the *G. lamblia* ADHE.** The *G. lamblia adhE* encoded an 888-amino-acid peptide ($M_r$, 96,832) (Fig. 5). The Met residue chosen as the start codon was likely correct because there was an in-frame stop codon before the next ATG in the 5′ untranslated region and the next potential start codon was not until $Met^{62}$, well into the overlap with other eukaryotic and prokaryotic ADHE (9, 17, 41, 64). There was a putative ribosome-binding sequence (ATTC) beginning at −7, no introns were present, and 70% of the third positions in the codons were G or C (excluding Met and Trp). The *G. lamblia adhE* gene is likely single copy, because overlapping sequences were identical among the PCR product and gDNA clones.

The *G. lamblia* ADHE closely resembled the ADHE of the γ-purple bacterium *E. coli* (49% overall positional identity), gram-positive bacterium *Clostridium acetobutylicum* (44%), and *E. histolytica* (43%) (Fig. 5) (9, 17, 41, 64). Like the *E. histolytica* ADHE, the *G. lamblia* ADHE lacked mitochondrion- or hydrogenosome-targeting sequences at its amino terminus (49). The amino half of *G. lamblia* ADHE was more similar to the CoA- and $NADP^+$-dependent succinate semialdehyde dehydrogenase (encoded by *sucD*) of *Clostridium kluyveri* (32%) than to the CoA-independent aldehyde dehydrogenase of *E. histolytica*, yeasts, and eubacteria (all with <20% identity) (33, 54, 69). Throughout the amino half of the *G. lamblia* ADHE were numerous residues conserved in CoA-independent aldehyde dehydrogenases, including Glu and Cys residues that may be involved in catalysis. However, ALDH sequences defined by the PROSITE algorithm were absent in ADHE of *G. lamblia*, *E. histolytica*, *E. coli*, and *C. acetobutylicum* (5).

The carboxy half of the *G. lamblia* ADHE showed extensive positional identities with the *C. acetobutylicum* ADH1 (36%), the anaerobic bacterium *Zymomonas mobilis* ADH2 (35%), the *E. coli* 1, 2-propanediol oxidoreductase (31%), the gram-negative anaerobe *Bacillus methanolicus* methanol dehydrogenase (28%), the *Saccharomyces cerevisiae* ADH4 (28%), and the *C. kluyveri* 4-hydroxy-butanol dehydrogenase (encoded by a gene that is on the same operon as *sucD*) (19%) (Fig. 5) (54, 61). The carboxy half of the *G. lamblia* ADHE showed less extensive positional identities with the *E. histolytica* ADH3, *C. acetobutylicum* butanol dehydrogenase, and an *E. coli* ORF derived from sequencing its genome (GenBank accession no. U28377) (2, 45, 51). Similarities between the class 3 ADH were in two sequences identified by PROSITE, including one with three His residues, which have been implicated in binding ferrous ions (5). The connector regions between ALDH and ADH portions of the parasite and eubacterial ADHE were short. The *G. lamblia* connector sequence (QVMHVGEGA LAKAAADLISRG) showed no resemblance to that of *E. histolytica* (PKIFFEPHSIRYLRELKE) (Fig. 5).

**Phylogeny of class 3 ADH.** ML, P, and NJ trees were drawn from an alignment of 6 eukaryotic and 10 eubacterial class 3 ADH, including the carboxy half of the *G. lamblia*, *E. histolytica*, *E. coli*, and *C. acetobutylicum* ADHE but excluding any regions lacking easy alignment (Fig. 6) (14–16, 21). Each method identified a cluster comprising the carboxy half of the *E. histolytica*, *E. coli*, and *C. acetobutylicum* ADHE. P, but not NJ, analysis identified significant support for the placement of *G. lamblia* ADHE as a sister to the aforementioned clade (88% versus 38% bootstrap support) and placed as a sister to this the clade comprising *C. kluyveri* 4HBD and *E. coli* ADHC (63.5% versus 46% bootstrap support). The node linking the members of this larger clade was collapsed and marked with an asterisk, because the relationship between these ADH was unresolved. Rather than distinguishing the ADHE of eukaryotes from those of the bacteria, each method instead identifies similarity between the ADHE of *E. histolytica* and those of gram-negative (*E. coli*) and gram-positive (*C. acetobutylicum*) eubacteria.

Distal from this portion of the unrooted class 3 ADH tree, *Z. mobilis* ADH2, *S. cerevisiae* ADH4, and *Schizosaccharomyces pombe* ADH6 formed a third cluster, suggesting horizontal transfer of these genes between bacteria and yeasts. Although other pairs of ADH were clearly related (those of *Citrobacter freundii* and *K. pneumoniae* as well as those of *Rhodococcus* strain NI86/21 and *Alcaligenes eutrophus*), it was impossible to relate these ADH to each other or to the yeast and *Z. mobilis* ADH. Finally, when the *E. histolytica* ADH3, *C. acetobutylicum* BDH, and *E. coli* ORF were added to the alignment of the 16 ADH shown in Fig. 6, these enzymes formed a fourth clade (data not shown) (45, 51). This result suggested that the *E. histolytica adh3* gene was also transferred horizontally from bacteria.

**Phylogeny of *E. histolytica*, *G. lamblia*, and *T. vaginalis* FD.** In an effort to better determine which extant bacterium might best represent the lineage that served as the most likely source for the parasite genes encoding the fermentation enzymes, phylogenetic trees were constructed by using the *E. histolytica*, *G. lamblia*, and *T. vaginalis* FD (27, 28, 56). Each of these three FD aligned with a different groups of putative FD homologs. Because considerable differences between these groups precluded their reliable alignment, each was taken independently for subsequent phylogenetic analyses.

The *E. histolytica* FD is a 59-amino-acid peptide that contains two similar halves, each with four Cys residues arranged to form a 2[4Fe-4S] center that is like those of the POR (Fig. 4 and 7) (5). The *E. histolytica* FD is half the size of the 12-kDa FD present in mitochondria, enterobacteria, and the *T. vaginalis* hydrogenosome and somewhat smaller than the 8-kDa FD present in plant chloroplasts and cyanobacteria (28). Thirty-four 6-kDa FD were identified by using BLAST with FD of either *E. histolytica*, *G. lamblia*, or *Clostridium thermosaccharolyticum*. The *E. histolytica* FD was identical in 21 positions to the FD of the archaebacteria *Methanosarcina thermophila* and *Methanosarcina barkeri* (37%). The *E. histolytica* FD was different from the archaebacterial FD in a 15-amino-acid stretch between the iron-binding sites.

The *E. histolytica* FD was inferred by phylogenetic analyses as intermediate between a cluster composed of the *M. thermophila* and *M. barkeri* FD and a cluster of the gram-positive bacteria *C. thermosaccharolyticum* and *Peptostreptococcus elsdenii* (Fig. 8). A distal cluster contained numerous FD of clostridia, as well as an FD of the α-purple bacterium *R. rubrum*.

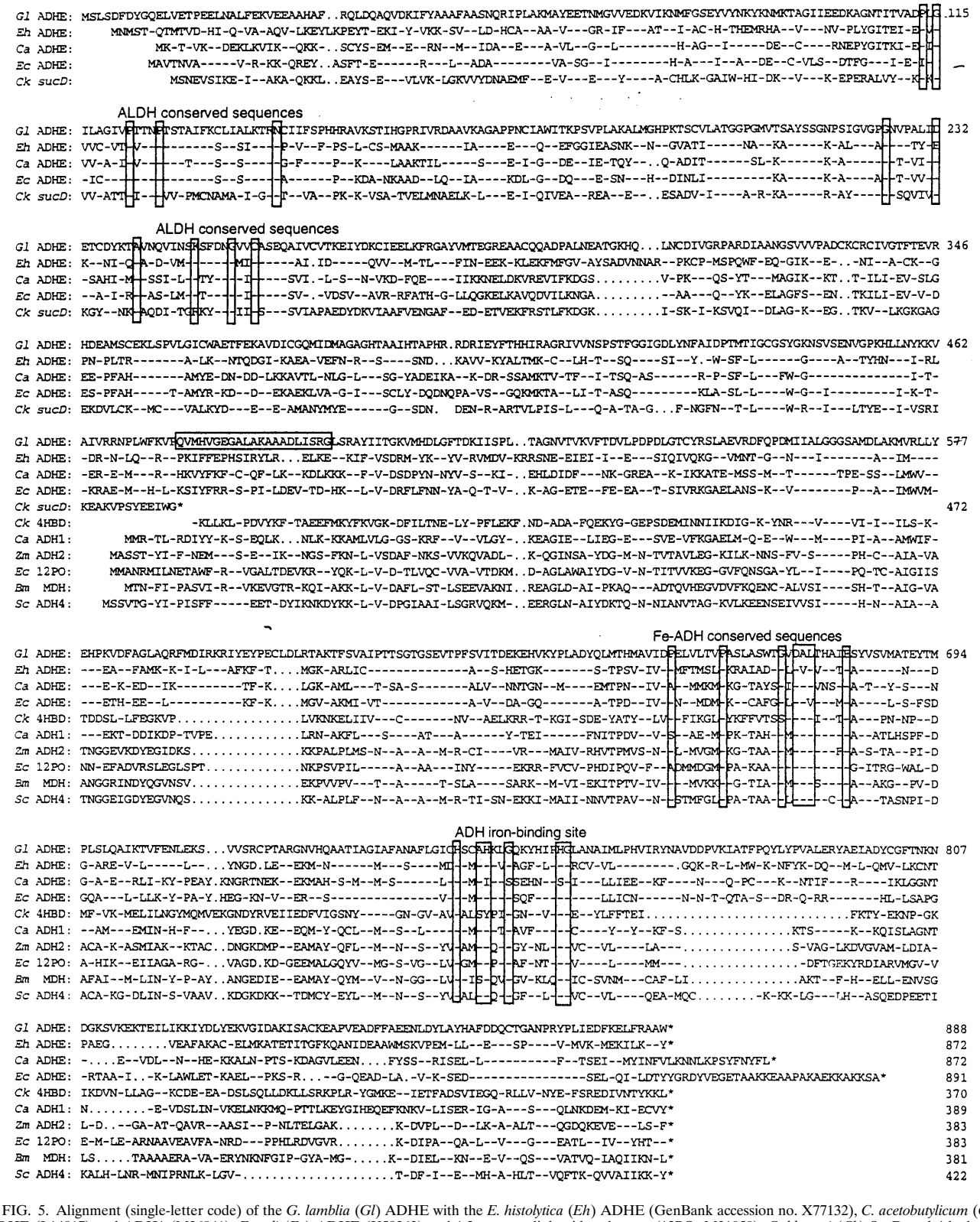



FIG. 5. Alignment (single-letter code) of the *G. lamblia* (*Gl*) ADHE with the *E. histolytica* (*Eh*) ADHE (GenBank accession no. X77132), *C. acetobutylicum* (*Ca*) ADHE (L14817) and ADH1 (M26941), *E. coli* (*Ec*) ADHE (X59263) and 1,2-propanediol oxidoreductase (12PO; M31059), *C. kluyveri* (*Ck*) SucD and 4-hydroxybuterate dehydrogenase (4HBD) (both L21902), *Z. mobilis* (*Zm*) ADH2 (M15394), *B. methanolicus* (*Bm*) methanol dehydrogenase (MDH; M65004), and *S. cerevisiae* (*Sc*) ADH4 (X05992) (9, 41, 61, 64). Dashes mark amino acid identities with *E. histolytica* POR, and periods mark gaps. The horizontal box marks unique connector regions in *G. lamblia* between amino-half ALDH sequences and carboxy-half class 3 ADH sequences. Vertical boxes mark conserved sequences found in CoA-dependent ALDH (shown here) and CoA-independent ALDH (not shown) (33, 69). Vertical boxes also mark conserved sequences found in class 3 ADH.

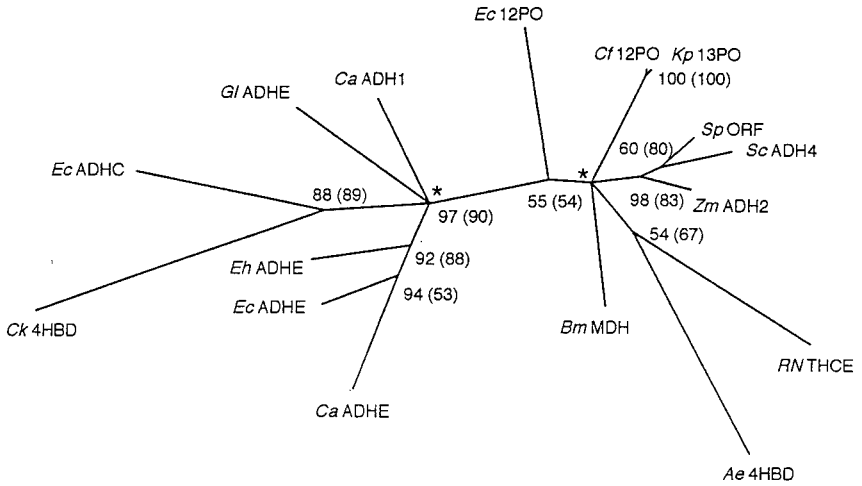



FIG. 6. Fifty percent phenogram of the 10 class 3 ADH shown in Fig. 5 as well as the *E. coli* (*Ec*) ADHC (GenBank accession no. X59544), *C. freundii* (*Cf*) 1,2-propanediol dehydrogenase (12PO; U09771), *K. pneumoniae* (*Kp*) 1,3-propanediol oxidoreductase (13PO; U30903), *Rhodococcus* strain NI86/21 (*RN*) THCE (U21071), *A. eutrophus* (*Ae*) 4-hydroxybuterate dehydrogenase (4HBD; L36817), and *S. pombe* (*Sp*) ORF (Z49811). *E. histolytica* (*Eh*) ADH3 (D49910), *C. acetobutylicum* (*Ca*) butanol dehydrogenase BDH1 (M96945), and an *E. coli* ORF (U28377) were part of a fourth clade class 3 ADH clade, which is not shown. Branch lengths in this unrooted tree were derived from the ML analysis, while bootstrap values at each node came from NJ trees and P trees (in parentheses). Asterisks mark nodes where either NJ or P analysis failed to identify the same bifurcating branch in over 50% of the bootstrap replicates.

The topology relating the various, highly similar clostridia could not be further resolved. These somewhat surprising results mean that if the *E. histolytica* ferredoxin gene indeed shares more recent common ancestry with the FD of clostridia, as has been previously suggested (27), this history cannot be readily inferred from sequence comparison.

The *G. lamblia* FD, which is also 6 kDa and for which a 27-amino-acid amino-half sequence is available, was intermediate between the clostridial cluster and a cluster composed of sulfur bacteria *Desulfovibrio desulfuricans*, *Desulfovibrio gigas*, *Desulfovibrio vulgaris*, and *Desulfovibrio africanus*, archaebacteria *Methanococcus thermolithotrophicus* and *Methanococcus jannaschii*, and the thermophilic eubacterium *Thermotoga maritima* (data not shown) (51, 56). The *G. lamblia* FD showed a 48% positional identity with the nearest *D. africanus* FD and only 37% positional identity with the *E. histolytica* FD. Interestingly, an Ala residue substitutes for the second of four Cys residues in the iron-sulfur center of the *G. lamblia* FD (5). These results suggested that the *G. lamblia* FD gene did not share close common ancestry with the FD gene of *E. histolytica*.

The FD of the *T. vaginalis* hydrogenosome, which contains a conserved Cys-$X_5$-Cys-$X_2$-Cys iron-sulfur-binding sequence, is part of a completely different group of 12-kDa [2Fe-2S] FD, which the includes γ-purple bacterium *Pseudomonas putida*, the photosynthetic bacterium *Rhodobacter capsulatus*, and *Caulobacter crescentus* (Fig. 9) (5, 28). Interestingly, this clade did not include the 12-kDa FD of mitochondria (data not shown). These results suggest that the *T. vaginalis* FD gene did not share a close ancestor with the amebic and giardia FD genes.

## DISCUSSION

The apparent absence of homologs in more eukaryotes limits the certainty with which we may identify the lineages that gave rise to the fermentation enzymes now present in various eukaryotic and eubacterial organisms. More definitive resolution of these gene trees would be possible if candidate protozoan representatives could be compared to two or more significantly differentiated clades. Instead, the available comparisons permit the less definitive conclusion that the fermentation genes from these amitochondriate protozoa do not themselves comprise groups whose common history can be verified by phenetic or cladistic means. Genes which could serve as appropriate outgroups, while not needed to reconstruct unrooted trees by the means that we have used, would therefore considerably aid in identifying the evolutionary history of these enzymes.

The argument for the bacterial origins of the *E. histolytica* fermentation genes may be summarized as follows.

(i) The sequences of the *E. histolytica* POR, FD, ADHE, ADH1, and ADH3 are all more similar to those of bacteria than to those of other amitochondriate protozoan parasites, suggesting that these genes derive by horizontal transfer from

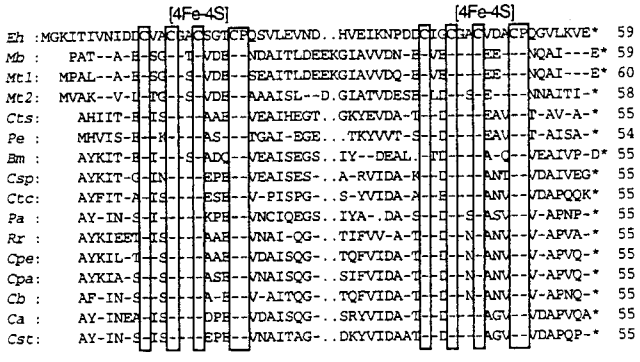



FIG. 7. Alignment (single-letter code) of the *E. histolytica* (*Eh*) FD (GenBank accession no. J03970) with the 6-kDa FD of *Methanosarcina barkeri* (*Mb*; PIR FEMZB), *M. thermophila* 1 and 2 (*Mt1* and *Mt2*; PIR A42960 and GenBank accession no. M83188), *C. thermosaccharolyticum* (*Cts*; PIR FECLCT), *P. elsdenii* (PIR FEME), *Butyribacterium methylotrophicum* (*Bm*; PIR JU01216), *Clostridium* sp. (*Csp*; PIR FECLCE), *R. rubrum* (*Rr*; PIR FEQFR), *Clostridium perfringens* (*Cpe*; PIR JX0144), *Clostridium pasteurianum* (*Cpa*; PIR FECLCP), *Clostridium butyricum* (*Cb*; PIR FECLCB), *Peptostreptococcus asaccharolyticus* (*Pa*; PIR FEPE), *Clostridium sticklandii* (*Cst*; PIR S36791), *Clostridium acidiurici* (*Ca*; PIR FECLCU), and *Clostridium thermocellum* (*Ctc*; PIR A24932). Dashes mark amino acid identities with *E. histolytica* FD, and periods mark gaps. Vertical boxes mark 2[4Fe-4S]-binding sites.

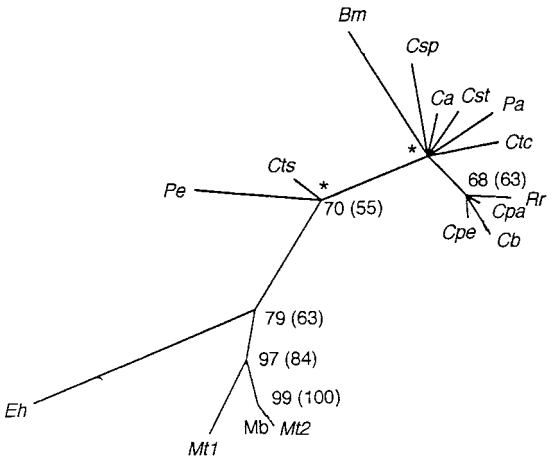



FIG. 8. Fifty percent phenogram of 6-kDa FD shown in Fig. 7. Branch lengths in this unrooted tree were derived from the ML analysis, while bootstrap values at nodes came from NJ trees and P trees (in parentheses). Asterisks mark nodes where either NJ or P analysis failed to identify the same bifurcating branch in over 50% of the bootstrap replicates. For abbreviations, see the legend to Fig. 7.

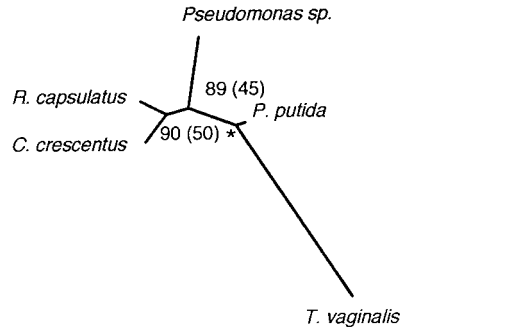



FIG. 9. Fifty percent phenogram of the 12-kDa FD of *T. vaginalis* (GenBank accession no. A36003), *P. putida* (J05406), *Pseudomonas* sp. (E42971), *C. crescentus* (S32573), and *R. capsulatus* (S45612). Branch lengths in this unrooted tree were derived from the ML analysis, while bootstrap values at nodes (in parentheses) came from NJ trees and P trees.

a bacterium rather than from a common eukaryotic ancestor (9, 18, 19, 27, 31, 45, 64).

(ii) Genes homologous to those that encode the *E. histolytica* fermentation enzymes are absent from the genome of *S. cerevisiae*, which has been sequenced in its entirety, and have not been identified in any other eukaryotes, including humans, where ADH, FD, and aldehyde dehydrogenases have been intensively studied (33, 66).

(iii) *E. histolytica* appears to have numerous other bacterial genes, including those that encode two $PP_i$-dependent glycolytic enzymes, HSP60, NNT, SOD, disulfide oxidoreductases, and possibly glyceraldehyde-3-phosphate dehydrogenase (8, 10, 13, 26, 50, 55, 68).

The *G. lamblia* genes encoding ADHE and FD are also likely derived from bacteria by horizontal transfer, as they are more similar to those of bacteria than to those of the other protozoan parasites. Other *G. lamblia* genes that appear to be derived from bacteria include the NADP-dependent glutamate dehydrogenase and adenylate kinase, which show 57 and 36% identities with *E. coli* and *Zea mays* chloroplast enzymes, respectively (47, 65). Because there was no close bacterial match for the *G. lamblia* POR, the possibility that its gene derives from a eukaryotic ancestor, as has been previously suggested, cannot be ruled out (1, 39, 57).

Whether bacterium-like POR genes of *E. histolytica* came from a phagocytosed bacterium (endosymbiont hypothesis), from direct transfer of plasmid DNA, or from some other method of transfer of genetic material cannot be determined (e.g., *Toxoplasma gondii* may have acquired a plastid by secondary endosymbiosis of a green alga) (18–20, 30, 59, 63). From which bacterial lineage these *E. histolytica* genes may have come is also unclear. Although an α-purple bacterium likely donated HSP60 to the eukaryotes prior to the divergence of the *E. histolytica* lineage, the *E. histolytica* POR, ADHE, ADH1, and ADH3 were no closer to those of gram-positive bacteria than to those of gram-negative bacteria, while the *E. histolytica* FD was closest to that of the archaebacterium *M. barkeri* (9, 13, 31, 45, 64). This is despite the fact that most of the *E. histolytica* fermentation enzymes are about as similar to their bacterial homologs as the *E. histolytica* HSP60 is to its nearest mitochondrial homolog (49% similarity with the *Cucurbita* sp. protein) (Fig. 2). It is thus likely that the fermentation genes have not yet been sequenced of the bacterium with which *E. histolytica* shares common ancestors for POR, FD, ADHE, ADH1, and ADH3 genes.

Although *E. histolytica* fermentation pathways show a strong superficial resemblance to those of clostridia (particularly with their multiple ADH), variation in the methods of energy metabolism within a particular bacterial group are great (67). This variation makes it impossible to use biochemical similarity to resolve bacterial phylogeny (62). Further, it does not appear likely that the *E. histolytica*, *G. lamblia*, and *T. vaginalis* FD came from the same bacterium, unless that bacterium had multiple FD genes as has been described for *Rhodobacter capsulatus* (six) (42). Nor does it appear likely that the *E. histolytica* POR, ADHE, and PPFK genes came from the same bacterium as those of *G. lamblia*, unless the bacterium had multiple POR, ADHE, and PPFK genes, and different fermentation genes were lost from *E. histolytica* and *G. lamblia* (3, 26, 48).

The bacterium-like genes of *E. histolytica* and *G. lamblia* are remarkable in that they encode cytosolic proteins lacking any sequences that might target them to an organelle (49). Indeed, no mitochondrion- or hydrogenosome-like organelles have ever been found in *E. histolytica* by electron microscopy (37). Further, we were unable to locate any mitochondrion-like structures by using the hydrophobic and cationic dye ruthenium red 123 (unpublished data and reference 12). Other examples of cytosolic proteins encoded by bacterium-like genes include the *S. cerevisiae* ADH4 and type II aldolase and the glyceraldehyde-3-phosphate dehydrogenase of *T. vaginalis* and *Euglena gracilis* (36, 50, 61). Conversely, there are numerous examples of horizontal transfer of eukaryotic genes to prokaryotes (e.g., glucose-6-phosphate isomerase and fibronectin III domains) (34, 50).

Previous investigators have suggested that the similarity of amebic fermentation enzymes to those of bacteria define *E. histolytica* as an "ancient" eukaryote (6, 9, 27, 39, 44, 45, 64). In contrast, the phylogenetic analyses performed here for POR, ADHE, ADH1, and FD (and elsewhere for SOD, PPFK, and ADH1) suggest that the *E. histolytica* genes encoding fermentation enzymes may have been transferred from bacteria and therefore do not represent those of ancient anaerobic eukaryotes (3, 31, 50). Possible candidates for the ancient eukaryotic fermentation enzymes include (i) pyruvate decarboxylase and ADH, which convert pyruvate to acetaldehyde and ethanol, respectively, and (ii) lactate dehydrogenase, which

converts pyruvate to lactate (38, 58, 66). Consistent with the idea that the genes encoding these fermentation enzymes derive from an ancestral eukaryote, pyruvate decarboxylase, lactate dehydrogenase, and ADH are all cytosolic enzymes, the sequences of which are easily distinguished from their bacterial homologs. Remarkably, both pyruvate decarboxylase and lactate dehydrogenase are absent from *E. histolytica* (39, 44).

## ACKNOWLEDGMENTS

We thank Ted Nash for the *G. lamblia* gDNA library, Peter Upcroft for submitting the unpublished *G. lamblia por* sequence to GenBank. Egbert Tannich for the *E. histolytica* cDNA library and for sequencing numerous *E. histolytica* genes encoding bacterium-like enzymes, Miklos Muller for helpful conversations and for sequencing numerous *G. lamblia* and *T. vaginalis* genes, Jun Adachi and Masami Hasegawa for MOLPHY 2.2, Joseph Felsenstein for PHYLIP 3.57c, and David Roos, Stephen Rich, and Richard Lewontin for useful comments on the manuscript.

This work was supported in part by the NIH and the MacArthur Foundation.

## REFERENCES

1. **Adam, R. D.** 1991. The biology of *Giardia* sp. Microbiol. Rev. **42:**465–488.
2. **Altschul, S. F., W. Gish, W. Miller, E. W., Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
3. **Alves, A. M. C. R., W. G. Meijer, J. W. Vrijbloed, and L. Dijkhuizen.** 1996. Characterization and phylogeny of the *pfp* gene of *Amycolatopsis methanolica* encoding $PP_i$-dependent phosphofructokinase. J. Bacteriol. **178:**149–155.
4. **Arnold, W., A. Rump, W. Klipp, U. B. Priefer, and A. Puhler.** 1988. Nucleotide sequence of a 24,206-base-pair DNA fragment carrying the entire nitrogen fixation gene cluster of *Klebsiella pneumoniae*. J. Mol. Biol. **203:** 715–738.
5. **Bairoch, A.** 1991. PROSITE: a dictionary of sites and patterns in proteins. Nucleic Acids Res. **19:**2241–2245.
6. **Bakker-Grunwald, T., and C. Wostmann.** 1993. *Entamoeba histolytica* as a model for the primitive eukaryotic cell. Parasitol. Today **9:**27–31.
7. **Bruchhaus, I., M. Leippe, M. Lioutas, and E. Tannich.** 1993. Unusual gene organization in the protozoan parasite *Entamoeba histolytica*. DNA Cell Biol. **12:**925–933.
8. **Bruchhaus, I., and E. Tannich.** 1993. Primary structure of the pyruvate phosphate dikinase in *Entamoeba histolytica*. Mol. Biochem. Parasitol. **62:** 153–156.
9. **Bruchhaus, I., and E. Tannich.** 1994. Purification and molecular characterization of the $NAD^+$-dependent acetaldehyde/alcohol dehydrogenase from *Entamoeba histolytica*. Biochem. J. **303:**743–748.
10. **Bruchhaus, I., and E. Tannich.** 1995. Identification of an *Entamoeba histolytica* gene encoding a protein homologous to prokaryotic disulphide oxidoreductases. Mol. Biochem. Parasitol. **70:**187–195.
11. **Bui, E. T. N., P. J. Bradley, and P. J. Johnson.** 1996. A common evolutionary origin for mitochondria and hydrogenosomes. Proc. Natl. Acad. Sci. USA **93:**9651–9656.
12. **Chen, L. B.** 1989. Fluorescent labeling of mitochondria. Methods Cell Biol. **29:**103–123.
13. **Clark, C. G., and A. J. Rogers.** 1995. Direct evidence for secondary loss of mitochondria in *Entamoeba histolytica*. Proc. Natl. Acad. Sci. USA **92:**6518–6521.
14. **Dayhoff, M. O., R. M. Schwartz, and B. C. Orcutt.** 1978. A model of evolutionary change in proteins, p. 345–352. *In* M. O. Dayhoff (ed.) Atlas of protein sequence structure, vol. 5, suppl. 3. National Biomedical Research Foundation, Washington, D.C.
15. **Faulkner, D. V., and J. Jurka.** 1988. Multiple aligned sequence editor (MASE). Trends Biochem. Sci. **13:**321–322.
16. **Felsenstein, J.** 1989. PHYLIP—phylogeny inference package (version 3.2). Cladistics **5:**164–166.
17. **Goodlove, P. E., P. R. Cunningham, J. Parker, and D. P. Clark.** 1989. Cloning and sequence analysis of the fermentative alcohol-dehydrogenase-encoding gene of *Escherichia coli*. Gene **85:**209–214.
18. **Gray, M. W.** 1989. The evolutionary origins of organelles. Trends Genet. **5:**294–299.
19. **Gray, M. W., and W. F. Doolittle.** 1982. Has the endosymbiont hypothesis been proven? Microbiol. Rev. **46:**1–42.
20. **Gupta, R. S.** 1995. Evolution of the chaperonin families (Hsp60, Hsp10 and Tcp-1) of proteins and the origin of eukaryotic cells. Mol. Microbiol. **15:**1–11.
21. **Hasegawa, M., and M. Fujiwara.** 1993. Relative efficiencies of maximum likelihood, maximum parsimony, and neighbor-joining methods for estimating protein phylogeny. Mol. Phylogenet. Evol. **2:**1–5.
22. **Hashimoto, T., Y. Nakamura, T. Kamaishi, F. Nakamura, J. Adachi, K.-I. Okamoto, and M. Hasegawa.** 1995. Phylogenetic place of mitochondrion-lacking protozoan, *Giardia lamblia*, inferred from amino acid sequences of elongation factor 2. Mol. Biol. Evol. **12:**782–793.
23. **Hawkins, C. F., A. Borges, and R. N. Perham.** 1989. A common structural motif in thiamin pyrophosphate-binding enzymes. FEBS Lett. **255:**77–82.
24. **Horner, D. S., R. P. Hirt, S. Kilvington, D. Lloyd, and T. M. Embley.** 1996. Molecular data suggest an early acquisition of the mitochondrion endosymbiont. Proc. R. Soc. Lond. Ser. B **263:**1053–1059.
25. **Hrdy, I., and M. Muller.** 1995. Primary structure and eubacterial relationship of the pyruvate:ferredoxin oxidoreductase of the amitochondriate eukaryote, *Trichomonas vaginalis*. J. Mol. Evol. **41:**388–396.
26. **Huang, M., R. A. Albach, K. P. Chang, R. L. Tripathi, and R. G. Kemp.** 1995. Cloning and sequencing of a putative pyrophosphate-dependent phosphofructokinase gene from *Entamoeba histolytica*. Biochim. Biophys. Acta **1260:** 215–217.
27. **Huber, M., L. Garfinkel, C. Gitler, D. Mirelman, M. Revel, and S. Rozenblatt.** 1988. Nucleotide sequence analysis of an *Entamoeba histolytica* ferredoxin gene. Mol. Biochem. Parasitol. **31:**27–34.
28. **Johnson, P. J., C. E. D'Oliveira, T. E. Gorrell, and M. Muller.** 1990. Molecular analysis of the hydrogenosomal ferredoxin of the anaerobic protist *Trichomonas vaginalis*. Proc. Natl. Acad. Sci. USA **87:**6097–6101.
29. **Kletzin, A., and M. W. W. Adams.** 1996. Molecular and phylogenetic characterization of pyruvate and 2-ketoisovalerate ferredoxin oxidoreductases from *Pyrococcus furiosus* and pyruvate ferredoxin oxidoreductase from *Thermotoga maritima*. J. Bacteriol. **178:**248–257.
30. **Kohler, S., C. F. Delwiche, P. W. Denny, L. G. Tilney, P. Webster, R. J. M. Wilson, J. D. Palmer, and D. S. Roos.** 1997. A plastid of probable green algal origin in Apicomplexan parasites. Science **275:**1485–1489.
31. **Kumar, A., P.-S. Shen, S. Descoteaux, J. Pohl, G. Bailey, and J. Samuelson.** 1992. Cloning and expression of an $NADP^+$-dependent alcohol dehydrogenase gene of *Entamoeba histolytica*. Proc. Natl. Acad. Sci. USA **85:**1782–1786.
32. **Lee, C. C., X. Wu, R. A. Gibbs, R. G. Cook, D. M. Muzny, and C. T. Caskey.** 1988. Generation of cDNA probes directed by amino acid sequence: cloning of urate oxidase. Science **239:**1288–1291.
33. **Lindahl, R., and J. Hempel.** 1991. Aldehyde dehydrogenases: what can be learned from a baker's dozen sequences? Adv. Exp. Med. Biol. **284:**1–8.
34. **Little, E., P. Bork, and R. F. Doolittle.** 1994. Tracing the spread of fibronectin type III domains in bacterial glycohydolases. J. Mol. Evol. **39:**631–643.
35. **Margulis, L.** 1981. Symbiosis in cell evolution. W. H. Freeman & Company, New York, N.Y.
36. **Markos, A., A. Miretsky, and M. Muller.** 1993. A glyceraldehyde-3-phosphate dehydrogenase with eubacterial features in the amitochondriate eukaryote, *Trichomonas vaginalis*. J. Mol. Evol. **37:**631–643.
37. **Martinez-Palomo, A.** 1982. Cell biology, p. 5–59. *In* K. N. Brown (ed.), The biology of *Entamoeba histolytica*. John Wiley, Research Studies Press, New York, N.Y.
38. **Mucke, U., T. Wohlfarth, U. Fiedler, H. Baumlein, K. P. Rucknagel, and S. Konig.** 1996. Pyruvate decarboxylase from *Pisum sativum*. Eur. J. Biochem. **237:**373–382.
39. **Muller, M.** 1992. Energy metabolism of ancestral eukaryotes: a hypothesis based on the biochemistry of amitochondriate parasitic proteins. BioSystems **28:**33–40.
40. **Muller, M.** 1993. The hydrogenosome. J. Gen. Microbiol. **139:**2879–2889.
41. **Nair, R. V., G. N. Bennett, and E. Papoutsakis.** 1994. Molecular characterization of an aldehyde/alcohol dehydrogenase gene from *Clostridium acetobutylicum* ATCC 824. J. Bacteriol. **176:**871–885.
42. **Naud, I., M. Vincon, J. G. F. Arin, E. Forest, and Y. Jouanneau.** 1994. Purification of a sixth ferredoxin from *Rhodobacter capsulatus*. Eur. J. Biochem. **222:**933–939.
43. **Ravdin, J. I.** 1995. Amebiasis. State-of-the-art clinical article. Clin. Infect. Dis. **20:**1453.
44. **Reeves, R. E.** 1984. Metabolism of *Entamoeba histolytica* Schaudinn, 1903. Adv. Parasitol. **23:**105–142.
45. **Rodriguez, M. A., M. Baez-Camargo, D. M. Delgadillo, and E. Orozco.** 1996. Cloning and expression of an *Entamoeba histolytica* $NADP^+$-dependent alcohol dehydrogenase gene. Biochim. Biophys. Acta **1306:**23–26.
46. **Rodriguez, M. A., M. E. Hidalgo, T. Sanchez, and E. Orozco.** 1996. Cloning and characterization of the *Entamoeba histolytica* pyruvate:ferredoxin oxidoreductase gene. Mol. Biochem. Parasitol. **78:**273–277.
47. **Rozario, C., and M. Muller.** 1995. Primary structure of a putative adenylate kinase gene of *Giardia lamblia*. Mol. Biochem. Parasitol. **71:**279–283.
48. **Rozario, C., M. W. Smith, and M. Muller.** 1995. Primary sequence of a putative pyrophosphate-linked phosphofructokinase gene of *Giardia lamblia*. Biochim. Biophys. Acta **1260:**218–222.
49. **Schatz, G.** 1987. Signals guiding proteins to their correct locations in mitochondria. Eur. J. Biochem. **165:**1–6.
50. **Smith, M. W., D.-F. Feng, and R. F. Doolittle.** 1992. Evolution by acquisition: the case for horizontal gene transfers. Trends Biochem. Sci. **17:**489–493.
51. **Smith, R. F., and T. F. Smith.** 1992. Pattern-induced multi-sequence align-

ment (PIMA) algorithm employing secondary structure-dependent gap penalties for use in comparative protein modeling. Protein Eng. **5:**35–41.
52. **Sogin, M. L.** 1991. Early evolution and the origin of eukaryoties. Curr. Opin. Genet. Dev. **1:**457–463.
53. **Sogin, M. L., J. H. Gunderson, H. J. Elwood, and D. A. Peattie.** 1989. Phylogenetic meaning of a kingdom concept: an unusual ribosomal RNA for *Giardia lamblia*. Science **243:**75–77.
54. **Sohling, B., and G. Gottschalk.** 1996. Molecular analysis of the anaerobic succinate degradation pathway in *Clostridium kluyveri*. J. Bacteriol. **178:**871–880.
55. **Tannich, E., I. Bruchhaus, R. D. Walter, and R. D. Horstmann.** 1992. Pathogenic and nonpathogenic *Entamoeba histolytica*: identification and molecular cloning of an iron-containing superoxide dismutase. Mol. Biochem. Parasitol. **49:**61–72.
56. **Townson, S. M., G. R. Hanson, J. A. Upcroft, and P. Upcroft.** 1994. A purified ferredoxin from *Giardia duodenalis*. Eur. J. Biochem. **220:**439–446.
57. **Townson, S. M., J. A. Upcroft, and P. Upcroft.** 1996. Characterization and purification of pyruvate:ferredoxin oxidoreductase from *Giardia duodenalis*. Mol. Biochem. Parasitol. **79:**183–193.
58. **Tsuji, S., M. A. Qureshi, E. W. Hou, W. M. Fitch, and S. S.-L. Li.** 1994. Evolutionary relationships of lactate dehydrogenases (LDHs) from mammals, birds, an amphibian, fish, barley, and bacteria: LDH cDNA sequences from Xenopus, pig, and rat. Proc. Natl. Acad. Sci. USA **91:**9392–9396.
59. **Vines, R. R., J. E. Purdy, B. D. Ragland, J. Samuelson, B. J. Mann, W. A. Petri, Jr.** 1995. Stable transfection of *Entamoeba histolytica*. Mol. Biochem. Parasitol. **71:**265.
60. **Wieland, O. H.** 1983. The mammalian pyruvate dehydrogenase complex: structure, and regulation. Rev. Biochem. Physiol. Pharmacol. **96:**123–170.
61. **Williamson, V. M., and C. E. Paquin.** 1987. Homology of *Saccharomyces cerevisiae* ADH4 to an iron-activated alcohol dehydrogenase from *Zymomonas mobilis*. Mol. Gen. Genet. **209:**374–381.
62. **Woese, C. R.** 1987. Bacterial evolution. Microbiol. Rev. **51:**221–271.
63. **Yang, D., Y. Oyaizu, H. Oyaizu, G. J. Olsen, and C. R. Woese.** 1985. Mitochondrial origins. Proc. Natl. Acad. Sci. USA **82:**4443–4447.
64. **Yang, W., E. Li, T. Kairong, and S. L. Stanley, Jr.** 1994. *Entamoeba histolytica* has an alcohol dehydrogenase homologous to the multifunctional *adhE* gene product of *Escherichia coli*. Mol. Biochem. Parasitol. **64:**253–260.
65. **Yee, J., and P. F. Dennis.** 1992. Isolation and characterization of a NADP-dependent glutamate dehydrogenase from the primitive eukaryote *Giardia lamblia*. J. Biol. Chem. **267:**7539–7544.
66. **Yokoyama, S., R. Yokoyama, C. S. Kinlaw, and D. E. Harry.** 1990. Molecular evolution of zinc-containing long-chain alcohol genes. Mol. Biol. Evol. **7:**143–154.
67. **Young, M., N. P. Minoton, and W. L. Staudenbauer.** 1989. Recent advances in the genetics of the clostridia. FEMS Microbiol. Rev. **63:**301–326.
68. **Yu, Y., and J. Samuelson.** 1994. Primary structure of an *Entamoeba histolytica* nicotinamide nucleotide transhydrogenase. Mol. Biochem. Parasitol. **68:**323–328.
69. **Zhang, W.-W., and J. Samuelson.** 1994. Cloning and expression of the gene for an $NADP^+$-dependent aldehyde dehydrogenase of *Entamoeba histolytica*. Mol. Biochem. Parasitol. **63:**157–161.