# Characterization of an Immuno-dominant Variable Surface Antigen from Pathogenic and Nonpathogenic Entamoeba histolytica

By Ursula Edman,<sup>\*</sup> Marco A. Meraz,<sup>‡</sup> Sloan Rausser,<sup>\*</sup> Nina Agabian,<sup>\*</sup> and Isaura Meza<sup>‡</sup>

From the \*Intercampus Program for Molecular Parasitology, University of California, Berkeley and San Francisco, California 94143; and the <sup>‡</sup> Department of Cell Biology, Centro de Investigacion y de Estudios Avanzados, Instituto Politecnico Nacional, Mexico City 07000, D.F. Mexico

## Summary

A 125-kD surface antigen of Entamoeba histolytica is recognized by 73% of immune sera from patients with amoebic liver abscesses. Using pooled human immune sera a cDNA clone ( $\lambda cM17$ ) encoding this antigen (M17) has been isolated from a  $\lambda$ gt11 expression library of the virulent stain E. histolytica HM1:IMSS. Monospecific antibodies, purified by binding to phage lysate of  $\lambda$ cM17, and mAb FA7 reacted exclusively with the 125-kD antigen by Western blot analysis. Surface binding and cap formation are observed with patient sera, purified monospecific antiserum, and mAb FA7. Corresponding genomic clones (pBSgM17-1/2/3) were isolated by hybridization with the cDNA clone. These contained an open-reading frame of 3345 bp, which is in good agreement with the mRNA size of  $\sim$ 3.0 kb as revealed by Northern hybridization with  $\lambda$ cM17. The inferred amino acid sequence predicts a 125,513 dalton protein that contains 17 potential N-linked glycosylation sites and is unusually rich in tyrosine and asparagine residues. A distinctly hydrophobic NH<sub>2</sub>-terminal region may serve as membrane anchor or signal sequence. In contrast to conservation of an immunodominant epitope recognized in pathogenic and nonpathogenic strains by monoclonal FA7 and human immune sera, amplification and sequence analysis of a 1,400-bp fragment of this gene from a fresh nonpathogenic isolate by use of the PCR demonstrate regions of significant sequence divergence in this antigen. A 1% sequence variability among different isolates of the pathogenic strain HM1:IMSS and a 12-13% variability between pathogenic and nonpathogenic strains are revealed by comparison to published partial amino acid sequences (Tannich, E., R.D. Horstmann, J. Knobloch, and H.H. Arnold. 1989. Proc. Natl. Acad. Sci. USA. 86:5118.). Some restriction enzymes were found that allowed PCR diagnosis of nonpathogenic and pathogenic isolates with the exclusion of E. histolytica-like Laredo, suggesting that a detailed study of nonpathogenic and pathogenic isolates in relation to the M17 antigen sequence will provide a basis of differentiating isolates.

Entamoeba histolytica is a common human pathogen that causes a spectrum of disease ranging from a commensal state in asymptomatic carriers to fulminant diarrhea or extraintestinal abscess formation. Virulent amoebae cause ulceration of the intestinal epithelium and may penetrate the bowel wall to form extra-intestinal abscesses, primarily in the liver. Several molecular activities thought to correlate with the virulent phenotype has been partially characterized. These include a sulfhydryl protease (1-3), a pore-forming protein (4-6), an N-acetyl-galactosamine-specific adherence lectin (7-13), a 220kD N-acetyl-glucosamine lectin (14, 15), and a 96-kD surface antigen (16-18); however, the role of each of these in pathogenesis remains ill defined. Most importantly, it is still unclear whether in a given strain invasiveness is a stable (19) or a variable (20-22) genotypic characteristic. Standard methods of differentiating between potentially virulent strains of *E. histolytica* include host symptomatology and serology and the pattern of a number of parasite isoenzymes which together constitute its zymodeme. It is this latter criterion that has been generally used in the current classification of *E. histolytica* isolates; pathogenic and nonpathogenic zymodemes are differentiated on the basis of polymorphisms in the electrophoretic mobility of the glycolytic enzymes phosphoglucomutase (PGI), <sup>1</sup> hexokinase (HK), and phosphoglucoisomerase (PGI) (19, 23-26). At least 18 zymodemes

<sup>&</sup>lt;sup>1</sup> Abbreviations used in this paper: HK, hexokinase; PGI, phosphoglucoisomerase; PGM, phosphoglucomutase.

 <sup>879</sup> J. Exp. Med. © The Rockefeller University Press • 0022-1007/90/09/0879/10 \$2.00
 Volume 172 September 1990 879-888

have been described in pathogenic and nonpathogenic amoebae, but the majority of clinical isolates fall into zymodemes I, II, and III.

More recently several other probes, based either on DNA sequences (27-29), or the detection of specific antigens (30), have been suggested as diagnostic reagents. However, in most cases, axenic cultivation (20) and cloning (31) of amoebae directly from fresh stool samples before assay with any of these probes have not been achieved and none of these probes have been validated by large scale screening of clinically defined isolates that have also been compared with extant criteria, such as zymodeme patterns. Since axenization and cloning of amoebae from patient isolates appear to favor outgrowth of the less fragile pathogenic strains and are known to result in reversible attenuation of virulence, it is especially important to develop probes that can directly discriminate virulent amoebae in fresh isolates. Clearly, the existence of such probes will be invaluable in clinically distinguishing infections caused by a mixture of strains or those that may result from phenotypic interconversion of pathogenic to nonpathogenic strains as has been suggested by Mirelman (22). Presented here is the isolation and characterization of the gene encoding a 125kD surface antigen from E. histolytica which is immunodominant in patients with invasive disease. The 125-kD antigen exhibits significant sequence variation among amoebal isolates and is potentially useful in differentiating nonpathogenic and pathogenic strains of E. histolytica.

#### Materials and Methods

Entamoeba Isolates and Cell Culture. Trophozoites of the axenized E. histolytica strains (HM1:IMSS, NIH:HK9) and E. histolytica-like Laredo were grown in TYI-S-33 media as described by Diamond et al. (32). Polyxenic isolates were grown in liquid Robinson's medium supplemented with 10% bovine serum and containing 5  $\mu$ g/ml of medium of each of the following antibiotics: kanamycin, erythromycin, and ampicillin. Amoebae were pelleted by centrifugation at 900 rpm and washed twice with PBS, pH 7.5. Polyxenic amoebae were further purified by centrifugation through a Percoll/PBS cushion at 3,000 rpm in a refrigerated Accuspin centrifuge. Isolates SD4 (pathogenic, zymodeme II) and REF 291 and SD116 (nonpathogenic, zymodemes III and I), were a generous gift of Dr. Sharon Reed from the University of California, San Diego. Nonpathogenic isolates Nos. 43 and 44 and pathogenic isolate No. 46, classified by zymodeme analysis using gradient PAGE (33), were isolated in Mexico City. They correspond to Sargeaunt zymodemes I, I, and II, respectively.

Human Immune Sera and Western Blot Analysis. Sera from 108 patients with amoebic liver abscesses were obtained from Drs. A. Isibasi and R. Landa at the Instituto Nacional de la Nutricion and La Raza-IMSS Hospitals, Mexico City. Diagnosis of hepatic abscess in patients was established by clinical symptoms, countercurrent immunoelectrophoresis, ELISA, and rectosigmoidoscopy. Human sera from donors without history of amoebiasis and negative for anti-amoebic antibodies as tested by immunoblot served as controls. Western blots of whole trophozoites were prepared by suspending washed cells in PBS containing 10 mM p-hydroxymercuribenzoate and Laemmli sample buffer, boiling for 5 min, fractionation by 10% or 5–15% gradient SDS-PAGE, and electrophoretic transfer to nitrocellulose filters. All sera were evaluated by Western blot analysis on extracts of whole amoebae. 29 sera with the highest titer were selected from the 108 samples and were pooled.

Antimembrane fraction serum. This serum was obtained by immunizing mice with 300  $\mu$ g of membrane fraction, prepared as described previously (34) and diluted 1:1 with PBS and CFA. Mice were injected intraperitoneally every 2 wk until titers reached 1:5,000 as assayed by Western blot.

mAb FA7. Whole amoebic extract from  $2 \times 10^6$  amoebae was fractionated by preparative 5-15% gradient SDS-PAGE. After electrophoretic transfer to nitrocellulose the 125-kD region was excised from the blot, ground to a powder, and suspended in PBS. 100  $\mu$ l of the suspension were diluted 1:9 with PBS and injected three times intraperitoneally into mice at 2-wk intervals with a final boost before the fusion. Hybridomas were selected by positive reaction with the 125-kD band in Western transfers of *E. histolytica* extracts. Harvest fluid from clone FA7 was used at a 1:1,000 dilution in Western blot analysis.

Antibody Capping by Live Trophozoites. Human immune serum, hybridoma harvest fluid from clone FA7, and purified monospecific antiserum, were added to live trophozoites at 1:500, 1:2,000 and undiluted, respectively. After formation of caps (10 min at 37°C), cells were fixed with 3.7% formaldehyde, washed with PBS, and stained with FITC-labeled anti-human or anti-mouse IgG. Undiluted harvest fluid from an anti-actin-producing clone was used as control for a nonsurface antigen (15, 35).

Preparation and Screening of Libraries. Genomic DNA and poly(Å)<sup>+</sup> RNA isolation and construction of the  $\lambda$ gt11 cDNA library from strain E. histolytica HM1: IMSS have been described previously (36). For construction of the genomic library from E. histolytica HM1:IMSS, 600 µl of NaI (GeneClean kit; Bio101) were added to 200  $\mu$ l (~20  $\mu$ g DNA) of agarose-embedded nuclei (36) in an Eppendorf tube and melted by incubation at 60°C for 5 min. 20  $\mu$ l of glassmilk were added, suspended well, and the mixture was incubated at room temperature for 5 min. The sample was vortexed for 1 min to shear the DNA and spun in a microfuge for 5 s. After removal of the supernatant the pellet was suspended in 1 ml wash buffer by vortexing for 30 s. The glassmilk was pelleted by a 5-s spin in the microfuge, and the supernatant was removed. The wash was repeated twice and the sheared and purified DNA was eluted into 100 µl 10 mM Tris-HCl (pH 8), 1 mM EDTA (TE) by incubation at 37°C for 5 min. Recovery and degree of shearing were assessed by agarose gel electrophoresis. All subsequent steps including addition of EcoRI linkers, methylation, ligation into the vector  $\lambda$ ZAPII, and packaging reaction were performed as described previously (37, 38). The  $\lambda$ gt11 cDNA library (3 × 10<sup>5</sup> phage) was screened with the pool of 29 patient sera at a 1:200 dilution (39, 40). The genomic library was screened with the  $\alpha$ -[<sup>32</sup>P]dCTPlabeled EcoRI fragment of  $\lambda cM17$ . Plasmids were rescued from genomic  $\lambda$ ZAPII clones as described previously (41). Phage DNA and plasmid DNA were purified by standard methods (42).

Sequence Analysis. With the exception of the first 207 bp, the entire sequence of gene M17 presented in Fig. 3 was determined on both strands in genomic clone pBSgM17-1 and on one strand in genomic clone pBSgM17-2. The internal EcoRI fragment representing the cDNA insert was also sequenced on both strands using nested deletion templates created with the Promega Biotech system (Madison, WI). Double-stranded sequence was also determined for two PCR fragments obtained by amplification of genomic DNA from isolate REF 291, Zymodeme III, derived from an asymptomatic Costa Rican refugee and kindly provided by Dr. S.L. Reed. Several oligonucleotides were used as primers in single-stranded DNA (M13mp18/19) and double-stranded DNA [pBSKS(+)] sequencing reactions with the Sequenase system (U.S. Biochemical Corp., Cleveland, OH) or ABI Sequencer (Applied Biosystems Inc., Foster City, CA).

Primer Extension Sequence Analysis. Primer extension sequence analysis was performed by reverse transcriptase-mediated extension of oligonucleotide primer SRO9 (5'AACTACTCCTGTGACTA-TTGCAGAAG3') annealed to 10  $\mu$ g poly(A)<sup>+</sup> enriched RNA in the presence of deoxyadenosine 5'-[ $\alpha$ -[<sup>35</sup>S]thio]triphosphate as described previously (36).

Polymerase Chain Reaction. The PCR was performed using a Cetus Corp./Perkin-Elmer DNA thermocycler. Reaction mixtures  $(50 \ \mu l)$  contained 25 pmol of each of the two oligonucleotide primer pairs SRO18 [5'GCAACTAGTGTTAGTTATAC3'] + SRO21 [5'GGTGGAATTTGGAATTCTGG3'] and SRO19 [5'GTATAA-CTAACACTAGT3'] + SRO22 [5'GCTGTTACACTTGAAAA-TAT3'], ~500 ng of genomic DNA, all four dNTPs each at 1.5 mM, 60 mM KCl, 25 mM Tris-HCl (pH 8), 0-20 mM MgCl<sub>2</sub>, 0.1% BSA, and 10% DMSO. The reaction mixture was overlaid with a drop of paraffin oil and denatured at 94°C for 10 min, and amplification was initiated by addition of 2.5 U of Thermus aquaticus DNA polymerase (Cetus Corp., Emeryville, CA). PCR parameters were 35 thermal cycles consisting of a 1-min denaturation of 94°C followed by a 3-min annealing period at 42°C, a 3-min ramp, and a 4-min extension period at 72°C. The amplification products were restricted with EcoRI and SpeI endonucleases and purified for subcloning into M13 by 2% low-melting-point agarose gel electrophoresis.

### Results

Western Blot Analysis. Sera from 108 different patients, diagnosed with amoebic liver abscess, were each reacted with Entamoeba whole cell extracts in Western blots; seven antigens (220, 190, 160, 125–129, 96, 75, 46 kD) were detected by >62% of the sera (43). Among these seven, a 125-kD antigen was immunodominant, reacting strongly and being recognized by >70% of the serum samples. We assume, based on their molecular weight and serological reactivity, that the 220-kD, the 160- and the 96-kD antigens represent the N-acetyl-glucosamine adherence lectin (14, 15), the N-acetyl-D-galactosamine adherence lectin (7-13), and the 96-kD integral membrane protein (16-18), respectively. Because it appeared that the 125-kD antigen had not been characterized, we chose to study this immunodominant antigen in more detail. A Western blot of whole cell extracts of axenically or polyxenically propagated pathogenic and polyxenically propagated nonpathogenic E. histolytica isolates was assayed with the pooled subset of 29 human immune sera (Fig. 1); the sera reacted strongly with a 125-kD antigen in all isolates regardless of source. Polyspecific antiserum prepared against amoebic plasma membrane (34) also reacted strongly with the 125-kD antigen (Fig. 1). The mAb FA7, prepared against partially purified 125-kD antigen, reacted specifically with an epitope of the 125-kD antigen; by Western analysis with FA7 this epitope was detected in different strains and species of Entamoeba (Fig. 1). In the Western blot with mAb FA7 additional bands of lower molecular weight and varying intensity are apparent in most of the isolates. Because potent proteases are present in whole amoebic extracts (1-3) we assume that these are degradation products of the 125-kD antigen, although processing intermediates of unknown origin can not be ruled out.

Localization of the 125-kD Antigen to the Surface of Amoebae. Live trophozoites will cap antibody-antigen complexes bound to their surface. Antibody-antigen caps were induced in HM1:IMSS trophozoites by incubation with above pooled patient serum, mAb FA7, or monospecific antibody recovered after specific binding and elution of pooled patient sera to phage lysates of cDNA clone  $\lambda$ cM17 (see below) (Fig. 2); a negative control antibody (anti-actin mAb) neither bound to trophozoite surfaces nor induced cap formation.

Isolation and Characterization of a cDNA Clone Encoding Part of the 125-kD Antigen. The pooled sera from amoebic abscess patients were used to screen a  $\lambda$ gt11 expression library



Figure 1. Western blot of whole Entamoeba extract fractionated by 5-15% SDS-PAGE, (lanes 2, 7, and 12) polyxenic pathogenic E. histolytica isolate SD-4, (lanes 3, 8, and 13) polyxenic nonpathogenic E. histolytica isolate SD116, (lanes 4, 9, and 14) E. histolytica-like Laredo, (lanes 5, 10, and 15) E. histolytica HK-9, (lanes 6, 11, and 16) E. histolytica HM1:IMSS probed with anti-membrane fraction serum (lanes 2-6), pooled human immune sera (lanes 7-11), and mAb FA7 (lanes 12-16); molecular masses are given in kilodaltons (molecular mass standards lane 1: 200, 97, 68, 43, 28 kD).

881 Edman et al.



А

Figure 2. Photographs ( $\times$ 800) of *E.* histolytica HM1:IMSS trophozoites labeled in vivo with primary antibodies. (*A*) Pool of human anti-*E.* histolytica immune sera at 1:500 dilution; (*B*) pool of human anti-*E.* histolytica immune sera purified by binding to  $\lambda$ cM17 phage lysates; (*C*) monoclonal FA7 harvest fluid at 1:1,000 dilution; (*D*) monoclonal anti-*E.* histolytica actin antibody at 1:1,000 dilution; secondary antibodies, FITC goat-anti-human and FITC goat-antimouse.

from E. histolytica HM1:IMSS; 46 reactive clones were each plaque purified and tested for recognition by each of the 29 patient sera included in the serum pool and by the antimembrane antibody. Clone  $\lambda cM17$  strongly reacted with 26 of 29 patient sera as well as with the anti-membrane serum. Monospecific antibody was selected from the pooled human sera by elution from filter-bound phage lysate of  $\lambda cM17$ . This eluate reacted with a single polypeptide of 125 kD by Western blot analysis of whole amoebic extracts; phage lysate of  $\lambda gt11$ , serving as negative control, did not bind antibodies reacting with amoebic antigens (data not shown). After nucleotide sequence analysis, the  $\lambda cM17$  1.9-kb insert revealed an ORF spanning the entire insert (Fig. 3). The lack of a 5' initiating methionine, the absence of a poly(A)-tail, and hybridization to a  $\sim$ 3 kb mRNA by Northern blot analysis (Fig. 4) indicated that NH2- and COOH-terminal sequences were lacking in  $\lambda cM17$ .

Isolation and Characterization of Genomic Clones Encoding the 125-kD Antigen. To isolate a genomic clone, a  $\lambda$ ZAPII library from *E. histolytica* HM1:IMSS was screened using the 1.9-kb insert of  $\lambda$ cM17 as a probe. Three genomic clones were identified, and two of these were sequenced using oligonucleotide primers derived from the cDNA sequence. The nucleotide sequence of the cDNA was identical in both genomic clones. An additional 556 bp of 5' and 870 bp of 3' sequence yielded an ORF of 3,345 bp which was also identical in both genomic clones (Fig. 3). The size of this ORF (gene M17) is in reasonable agreement with the mRNA size of  $\sim$ 3,000 bp determined by Northern blot analysis (Fig. 4). The inferred amino acid sequence predicts a 125-kD protein.

5' Flanking Sequence Comparison. The 5' flanking sequence of gene M17 shares striking similarities with the 5' flanking region of both actin and ferredoxin genes (Fig. 5), the only other genes of E. histolytica where sequence has been determined. The transcriptional start site of M17 was mapped to an adenine residue 17 bp 5' of the start codon by primer extension sequence analysis using oligonucleotide SRO9 (data not shown). 5' untranslated regions of actin (11 bp) (36, 44) and ferredoxin (9 bp) (45) genes were likewise very short as compared with other eukaryotic gene transcripts. A common sequence motif, 5'ATTCA3', is present at the transcriptional start site of both M17 and actin genes, the initiating nucleotide being an adenine residue as is most frequently found in other eukaryotes. While the same motif is also present in the flanking sequence of the ferredoxin gene, its cap site was mapped to the 3' thymidine rather than the 5' adenine residue (45) (Fig. 5). An additional sequence motif shared among these genes is YATTTAAA present at -29, -31, and -32 for the M17, actin, and ferredoxin gene flanking sequences, respectively. This sequence motif does not conform with the Goldberg-Hogness promoter consensus sequence TATAAATA, which in eukaryotic genes is located 25-30 bp upstream of the transcriptional start site. Nevertheless, this sequence is similar in the three E. histolytica genes, both in sequence and relative position, suggesting a consensus which in E. histolytica serves as the entry point for RNA polymerase.

Gene Copy Number. Southern blot and sequence analysis of the M17 gene and limited flanking regions indicate that this surface antigen is encoded by a single copy gene. When a Southern blot of genomic DNA from E. histolytica, restricted

	gaagetataaataagtistagaaatataaaagaatg		360
	ttaaaaatgaaaacaaacataaaaataagtgtatttaaagtgtttttaaasataactaatt*ATTCATAAATTAAAGT 20	1021	The Leu Asn Phe Asp Gin Arg Val Asp Ala Gly Ala Ala Val Ala Tyr Val Gly Arg Trp Aca TTA AAC TTT GAT CAA AGA GTT GAT GCA GGA GCT GCT GCT ATA GTA GGA GGA TGG ACA TTA AAC TTT GAT CAA AGA GTT GAT GCT GGA GCT GCT GTT GCA TIT GTA GGA AGA TGG
1	Met lau Gly Ser Lys Ser Lie He Ale Val Val Ale He Ale Ser Ale He He He Gly Arg tra GGT TCT AAR AGC ATT ATT GCT GTT GTG GCT ATA GCT TCT GCA ATA GTC ACA GGA 40	1001	Phe 380 Phe Thr Gin Asn Pro Ser Asp Trp Als Als Als Cys Val Gly Lys Asp Giy Leu He Asn men Fre Gin Asn Pro Ser Asp Trp Als Als Als Cys Val Gly Lys Asp Giy Leu He Asn
61	Val val val i le val val val val val i the Leu Ser val val Leu hir Arg Ser Ser val Lys GTA GTT GTT ATA GTT GTT GTT GTT GTA CA CTT TCT GTG GTT TTA ACA AGA AGT AGT GTT AA 60	1001	TTT ACT CAA GAC CCA TCC GAL TOG GCA GCI GCA TGI GII GGI AAG GAT GGA TTA ATA AAT TTT ACT CAA GAC CCA TCC GAC TGG GCC GCT GGA TGC GTT AAC AAA GAA GAA GAA TAA ATA His Ser Gly Asn Arg
121	ASD THE ASD SET HE HY VAL FLO ASD VAL LAD HE ASH ASD FLO GH AND HE ASH GAL GAC ACC AAC TCT ATT TAT GTT CCT GAT GTT ATT ACT AAC GAC CCA CAA ATG ACA ATG ACA ATG Not here The Low Glu Val The Sar Sar Sar Law Das Sar Glu The Law Dro. Law Glu	1141	400 Tyr Gly Asn Trp Gly Pro Leu His Glu Met Asn His His Met Gln Gly Thr Tyr Leu Lys TAT GGA AAT TGG GGA CCA TTA CAT GAA ATG AAT CAT CAT ATG CAA GGA ACT TAT TA AAA TTT GGA HAT TGG GGA CCA TTA CAT GAA ATG AAT CAT CAC ATG CAA GGA ACT TAT TA AAA
181	ATG GAT ACA TTA GAG GTT ATT TCT TCT AAA TTT AGT GGA ACA AHA CCA AAA GAA TGG 100 101 102 103 104 104 104 104 104 104 104 104		Ser Glu Glu Ann Tro Glu Tla Sar Ban Pro Glu Glu Glu Thr Ann An Val Mar Thr Sar 140
241	Inf met by fyt inf by fythe to the first inf of or or or or the first inf of of or or or or of the first inf of of or or of the first inf of of or or of the first inf of of or or of the first of the f	1201	GA GGA ANT TGG GGT ATT AGT AAT CCA GGA GAA GAA ACT AAT AAT GTT ATG ACA TCA ATT GGA GGA AAT TGG GGT ATT AGT AAT CCA GGA GAA GAA ACT AAT AAT GTT ATG ACA TCA ATT GGA GGA AAT GGG GGT AT <u>C</u> A <u>A</u> <u>A</u> <u>A</u> <u>A</u> <u>A</u> <u>A</u> <u>A</u> <u>C</u> <u>A</u>
301	by Son hon i to the hor of hon of hon and hon of hon the hon and the hon of the hon of the hon of hon the hon and hon the hon and hon the hon and hon	1261	440 Asn Tyr Ile Leu Tyr Thr Asn Ile Ala Gly His Arg Asn Gln Gly Leu Ser Gly Trp Asn AAT TAT ATT TTG TAT ACA AAT ATT GCT GGA CAT AGA AAT CAA GGA CTT AGT GGG TGG AAT
361	AAT GGA TCA TTA GGA AGA ATG CCT GAA AÀA TÀT GGT GGT GAT AÀA CAA TTT GAA GCT AAT 160 Gly Val Asn Trp Glu Ala Asp Arg Leu Glu Val Arg Tyr Gly Leu Phe Gly Arg Val Phe		ANT TAT ATT GTG TAT ACG ANT ATT GCA GGA CAT AGA AAA CAA GGG CTT AGT GGT TGG ANT Lys 460
421	GGÃ GTT ANT TGG GAN GCT GAT AGÃ TTA GAN GTT AGÃ TÀT GGT CTT TTC GGT AGÃ GTT TTT 180 Gly Gin Arg Ala Val Ala Trp Ala Phe Pro Gly Glu Ile Val Thr Ile Lys Phe Pro Lys	1321	Tye val Ser Asp Gly Tye Ser The Lie Tye Lys Lie Leu Lys Gly Glu Asn Asp Gln Pero Tat GTT CCT GAT GGT TAT TAT ACA TT TAT AAA ATT CTT AAA GGT GAA AAT GAT CAA CCT TAT GTT TCT GAT GGG TAT TCT ACA ATT ATT AAA ATT CTT AAA GGT GAA AAT GAT CAA CCT
481	GGT CAA AGA GCT GTT GCA TGG GCT TTT CCA GGA GAA ATC GTT ACA ATT AAA TTC CCT AAA 200 Gly Met Ser Tyr Lys Gly Ile Gln Val Gly Ile Gly Lys Cys Asn His Asn Pro Ser Asp		480 His Leu Arg Ser Tyr Val Asn Met Ala His Ala Phe Gly Thr Asp Thr Leu Ile Ala Leu
541	GGA ATG AGT TAT ANG GGA ATT CAG GTT GGT ATT GGT ANG TGT ANC CAT ANT CCT TCT GAT 220 Gin Trp Leu Asn Val Asn Asn Trp Ser Asn Asp Arg Met Pro Ile Asp Ser Ile Gly Phe	1381	CAT TTA AGG TCT TAT GTT AAT ATG GCA CAT GCA TTT GGA ACA GAC ACT TTA ATT GCT TTA CAT TTA AGG TCT TAT GTT AAT ATT GCA CAT GCA TTT GGA ACA GAT ACT TTA ATT GCT TTA 11e
601	CAA TGG TTA AAT GTT AAT AAC TGG TCA AAT GAT AGA ATG CCA ATC GAT TCA ATT GGA TTT 240 Asp Leu Gly Leu Asm Thr Thr Gln Pro Tyr Ile Ile Asm Asp Thr Phe Lys Ile Gly Ser	1441	500 Val Lys Ser Tyr Tyr Gly Leu Trp Tyr Glu Asn Asn Phe Glu Ser Lys Tyr Ser Ile Lys GTT AAA TCT TAT TAT GGA TAT TGG TAT GAA AAT AAT TTT GAA AGT AAA TAT TCA ATT AAA
661	GAT TTA GGA CTT AAT ACA ACG CAA CCA TAC ATT ATT AAT GAT ACA TTT AAA ATA GGA TCA 260 Pro Phe Gly Gly Met ILs Tyr Leu Arg Ser Asp Thr Thr Phe Thr Asn Ser Phe Tyr Val		GIT AAA TET TAT TAT GGL CTA TGG TAT GAA AAT AAT TAT GAA GGT GAC TAT TCA ATT AAG Tyr Gly Glu Sign and S
721	280 Thr Phe Ser Asn Val Gly Arg Ala Pro Ile Ile Asn Tyr Asn Ile Thr Thr Asn Glu Glu Dan more har come col act come col act controls and the thr Thr Asn Glu Glu	1501	AG GAT TCL ACT TCL GCT TTC TGT TTG TTA GCT GCA ATA AGA GAT ACT AGA AGA GAT TCL ACT TCL GCT TTC TTG TTA GCT GCA ATA GCT ACA AAA AGA GAT ACT AGA
841	And her was and out out out out and and the form and	1561	540 Tyr Leu Cys Ser Leu Phe Lys Tyr Asp Ile Gin Ser Asn Val Ser Giu Ala Ile Lys Asn TAC TTA TGT TCT CTA TTT AAA TAT GAT ATA CAA AT CA AM GTT TCA GAA GCA ATT AAA AAT
901	320 Arg Leu Val Leu Thr Ser Arg Asn 11e Arg Ser Leu Glu Asp Ala Gin Tyr 11e Ser Asp Aga CTT GTT CTT ACA AGA ANT ATT AGA AGT TTG GAA GAT GCA CAA TAT ATT AGT GAT		TAI TTA TGT TCT CTI TTT AAA TAC GAT ATA CAA CAA AAT GTT TCA GAA GCA ATT AAA AAC Gln 560
961	340 Phe Trp Leu Lys Ala Ile Ser Ile Ser Asn Tyr Ala Val Thr Leu Glu Asn Ile Pro Ile Trc TgG TTA ARA GCA ATT AGT ATT TCT AAT TAT GCT GTT ACA CTT GAA AAT ATT CCA ATT	1621	Met Asn Tyr Pro Thr Tyr Tyr Pro Phe Phe Asn Leu Tyr Ala Met Ser Tyr Asn Gly Asn Arg Ant TAT CCA ACT TAT TAT CCA TTC TTC AAC CTC TAT GCC ATG AGT TAT ATG GA AAT ARG AAT TAT CCA ACT TAT TAT CCA CT TC TTC AAC GTI TAT GCA TAT GAA AAT GGA AAT
	CCA ATE		Val
	580	2341	800 Phe lie Arg lie Gly Tyr Cys Tyr His Gly Thr Glu Glu Ser Ser Val Asp Val Ser Lys TIT ATT AGA ATA GGA TAT TGT TAT CAT GGA ACA GAA GAA TCA AGT GTT GAT GTT TCT AAA
1681	Tyr Tyr Gly Arg Fro Tyr Lys 11e Fro Tyr Gly Arg Thr Arg Leu Asn Phe Thr Ala Thr TAC TAT GGA AGA CC TAT ANA ATT CCA TAT GGA AGA ACT AGA TTG ANT TTC ACT GGA ACT TAT TAT GGA AGA <u>A</u> CA TAT ANA ATT CCA TAT GGI AGA ACT AGA TTG ANT TTI ACA GCA ACC		ITT ATT AGA ATA GGG TAT TGT TAT CAA GGA ACA GAA CAG TCT AGT GTT TGC AAA Gin Gin Asn 820
1741	Thr Thr 600 Thr Ala Ile Asp Pro Lys Ala Thr Ser Val Ser Tyr Thr Ile Lys Ser Gly Leu Thr Lys For Con Thr Ala Chi Con And Ser Con The Lys Thr Ise Type Thr Con Thr Lys	2401	Cys ser val ser asp lae gy ser ser met val deu asm giu lys val lys thr gly ala TGC AGT GTA TGC AGT ATT GGA AGC TCT ATG GTT CTT AAT GGA AAA GTT AAA ACA GGA GCA TGT AGT GGA TLA GAT ATT GGA AGC GJU Lau
1/41	ACT GCT ATA GAT CCA ADA GCA ACT AGT GTT AGT TAT ACT ATA AAT AAA ACT GCT ATA GAT CCA AGT GCA ACT AGT GTT AGT GTT AGT AAA Ser His	2461	840 Lys Glu Pro Glu Phe Gln Ile Pro Pro Ile Lys Tyr Ser Arg Pro Thr Arg Phe Leu Thr AAA GAA CGA GAA TTC CAA ATT CCA CCA ATT AAA TAC AGC AGA CCA ACA CCT TTC TTA ACT
1801	Gly Lys Leu Glu Arg Val Glu Asp Asn Val Tyr Asp Tyr Thr Pro Phe Phe Gly Ile Glu GGA AMA TTA GAA GGA GTT GAA GAC AAT GTT TAT GAC TAT ACA CCA TTC TTT GGA ATA GAA GGA AMG TTA GAA CAA GTT GAA GAA AMT GTT TAT GAT TAT ACA CCA AMC TTT GGA GGA GAT	2521	Asn Ala Tyr Arg Thr Ile Pro Lys Cys Leu Asn Gly Asp Asp Ala Cys Ser Ile Lys Cys AAT GCA TAT AGA ACT ATT CCA AAA TGT TTG AAT GGT GAC GAT GCT TGT TCT ATT AAA
	Gln Glu Asn Ala Asp 640 Glu Asn Asp Thr Phe Val Leu Asn Ile Asp Cys Val Val Asn Gly Glu Lys Val His Ile	2581	880 Leu Ser Leu Leu Pro Leu Lys His Asp Asp Ser Ser Lys Cys Ser Asn Met Phe Asp Asp CTC TCC TTA TTA CCA CTT AAA CAT GAT GAT TCA AGT AAA TGT TCT AAT ATG TTT GAT GAT
1861	GAN MAT GAT AGA TITI GITI TIA AAT ATT GAT GIT GITI GITI	2641	900 Asn Tyr Ser Thr Met Tyr His Ser Arg Trp Thr Gly Gln Gly Thr Thr Phe Pro Val Asn AAT TAT TCT ACT ATG TAT CAT TCA AGA TGG ACT GGA CAA GGA ACT ACT TTC CCA GTT AAT
1921	660 Glu Gln Glu Gly Thr Phe Glu Leu Asp Pro His Gln Vel Glu Tyr Glu Vel Tyr Lys Asp GAA CAA GAA GGA ACA TIT GAA TTA GAT CCA CAA CAA GTA GAA TAT GAA GTT TAT AAA GAT	2701	920 Tyr Thr Phe Glu Phe Ser Glu Asn Val Thr Phe Asn Asn Leu Tyr Val His His Arg Arg TAT ACA TTT GAA TTC TCA GAA AAT GTA ACA TTT AAT AAT CTT TAT GTT CAT CAT AGA AGA
	GAA CAA GAC GGA ACA TIT GAA CTA GAC CCA CAT CAA GTA GAC TAT GAA GIT TAT AAA GAT Asp 680	2761	940 Pro Glu Asp Ser Trp Gly Tyr Phe Glu Met Phe Val Lys Ser Pro Glu Thr Gly Glu Met CCT GAA GAT TCA TGG GGA TAC TTT GAA ATG TTT GTT AAA TCT CCA GAA ACA GGA GAA ATG GGA
1981	Val Gin Thr Arg Abp Met Alia Gin Ala lle Abn Ile lle Gin Abn Lya Thr Arg Abn Abp GTT CAA ACA AGA GAT ATG GCA CAA GCT ATT ANT ATT ATT CAG AAT AMA ACT CGT AAT GAT GTT AAA ACA AAA GAT ATG GAA CAA GCT CTT AAT ACT ATT CAG AAT AAA ACT ICT AAT TAT	2821	Glu Leu Leu Glu Lys Tyr Lys His Pro Lys Ser Thr Thr Thr Glu Leu Asn Phe Gln Lys GAG TTA TTA GAA AAA TAT AAG CAT CCA AAG TCT ACT ACA ACA GAA CTT AAT TTC CAA AAA 980
2041	Lys Lys Giu Leu Thr Ser Tyr 700 Thr Gly Arg Ala Ser Phe Phe Gly Ile Gly Thr Tyr Asn Asp Gly Ser Met Glon Ser Met Ard AGB AGC GGT TEA THE THE THE THE THE GAT THE GAT THE AST OF THE AST OF THE AST	2881	Leu Val Thr Thr Asp Arg Val Gin Phe Ile Val Tyr Asn Asn Ser Asn Gly Gly Asn Tyr TTA GTT ACA ACT GAT CGT GTC CAA TTT ATT GTC TAT AAT AAT TCA AAT GGT GGG AAT TAT 1000
2041	ACA GGT AGG TCT ACT TC TTT GGA ATT GGA ATT AGA ATT ATT GGA ACA ATG CA TCA ATG ACA GGT AGG TCT ACT TC TTT GGA ATT GA ATT ATT GAT ATG ATA GA ATG ATG	2941	Val Asn Val Val Glu Leu Ser Phe Asn Ile Lys Glu Thr Phe Lys Asn Tyr Thr Asn Ser GTC AAT GTT GTA GAA TTG TCT TTC AAT ATT AAG GAA ACC TTT AAG AAT TAT ACA AAT TCA
			1020
2101	Leu Val Glu Lys Gly Lys Leu Ile Val Pro Lys Ser Gly Tyr Tyr Thr Leu Phe Met Lys TTA GTA GAA AAA GGT AAA TTG ATA GTT CCA AAA TCT GGA TAT TAT ACA TTG TTT ATG AAA TTA GTA GAA AAA GGG AAA <u>T</u> G ATA GTT CCA ACA TCA GGA TAT TAT ACA TTG TTT ATG AAA	3001	Phe Gly Pio Lys Ile Lys Ser Thr Gly Phe Lys Lys Val Thr Thr Pro Gly Ala Ser Gly TTT GGA CCA AGG ATT AAA AGT ACT GGA TTT AAG AAA GTT ACT ACA CCA GGT GT TCA GGA Gly Twr An Ala Yah Ag Gly Ly Gly Gly Gly Gly G T
2101	Leu Val Glu Lys Gly Lys Leu Ile Val Pro Lys Ser Gly Tyr Tyr Thr Leu Phe Met Lys TTA GTA GAA AAA GGT AAA TTG ATA GTT CCA AAA TCT GGA TAT TAT ACA TTG TTT ATG AAA TTA GTA GAA AAA GGT AAA <u>C</u> TG ATA GTT CCA A <u>CA</u> TC <u>A</u> GGA TAT TAT ACA TTG TTT ATG AAA Leu Thr Ala Asp Asp Leu Gly Arg Leu Leu Leu Asn Ile Thr Gly Glu Tyr Glu Gln Leu Leu Aan	3001 3061	Phe Gly Pro Lys Ile Lys Ser Thr Gly Phe Lys Lys Val Thr Thr Pro Gly Ala Ser Gly TTT GGA CCA AGA ATT AAA AGT ACT GGA TTT AAG AAA GTT ACT ACA CCA GGT GCT TCA GGA 1040 Gly Tyr Leu Ala Val Asn Glu Lys Glu Gly Glu Gly Ser Leu Cys Phe Lys Ala Lys Val GGA TAT CTT GCA GTA AAT GAA AAG GAA GGA GGA GGA TCA CTT GT TTC AAA GCT AAA GTC Thr Lwa Phe Glu Leu Tyr Gly Tyr Arr Lws The The Ser Clu Lys Dhe Arr Will 1060
2101 2161	Leu Val Glu Lys Gly Lys Leu Ile Val Pro Lys Ser Gly Tyr Tyr Thr Leu Phe Met Lys TTA GTA GAA AAA GGT AAA TTG ATA GTT CCA AAA TCT GGA TAT TAT ACA TTG TTT ATG AAA TTA GTA GAA AAA GGT AAA GTG ATA GTT CCA ACA TCA GGA TAT TAT ACA TTG TTT ATG AAA Leu Thr Ala Asp Asp Leu Gly Arg Leu Leu Leu Asn Ile Thr Gly Glu Tyr Glu Gin Leu Leu Asp GCA GAT GAT TTA GGA AGG TTG TTA TG AAA TAT ACT GGA GAG TAT GAA CAA TTA TTA GAT GCA GAT GAT TTA GGA AGG TTG TTG TTA TAT GAA GGA GAG TAT GAA CAA TTA TTA AAT YA Asn	3001 3061 3121	Phe Gly Pto Lys lie Lys Ser Thr Gly Phe Lys Lys Val Thr Thr Pro Gly Ala Ser Gly TTT GGA CCA ANG ATT AMA AGT ACT GGA TTT ANG AMA GTT ACT ACA CCA GGT GCT TCA GGA GGA TAT CTT GCA GTA ANT GAA AGG GLA GGA GGA GGA GGA TCA CTT TGT TTC AMA GCT AMA GGA TAT CTT GCA GTA ANT GAA AAG GAA GGA GGA GGA GGA TCA CTT TGT TTC AMA GCT AMA GGA TAT CTT GCA GTA TAT GAA AAG GAA AGA AGA GGA AGA TCA CTT TGT TTC AMA GCT AMA GGA TAT CTT GCA GTA TAT GAA AAG GAA AGA GAA AGA TCA CTT TGT TTC AMA GCT AMA GCT AMA TTC GGT CTT TAT GGA TAT AGA AMA ACA ACA TCT GGA AAG TTT AGA GTT ACA ATT ACT AMA TTC GGT CTT TAT GGA TAT AGA AMA ACA ACA TCT GGA AAG TTT AGA GTT ACA AGT AMA TTC GGT CTT TAT GGA TAT AGA AMA ACA ACA TCT GGA AAG TTA AGA GTT ACA AGT AMA TTC GGT CTT TAT GGA TAT AGA AMA ACA ACA TCT GGA AAG TTA AGA TTA AGA TAT ACA ATT GGA TAT AGA TAT AGA AMA ACA ACA TCT AGA GTT ACA ATT AGT AMA TTC GGT CTT TAT GGA TAT AGA AMA ACA ACA TCT GGA AAG TTA AGA TAT AGA TA
2101 2161 2221	Leu Val Glu Lys Glu Lys Leu Ile Val Pro Lys Ser Gly Tyr Tyr Thr Leu Phe Met Lys TTA GTA GAA AAA GGT AAA TTG ATA GTT CCA AAA TCT GGA TAT TAT ACA TTG TTT ATG AAA TTA GTA GAA AAA GGT AAA CTG ATA GTT CCA ACA TCG GA TAT TAT ACA TTG TTT ATG AAA Leu Thr 740 Ala Asp Asp Leu Gly Arg Leu Leu Leu Asn Ile Thr Gly Glu Tyr Glu Gin Leu Leu Asp GCA GAT GAT TTA GGA AGG TTG TTG TTA TAT ATT ACT GGA GAG TAT GAA CAA TTA TTA GAT GCA GAT GAT TTA GGA AGG TTG TTG TTA ATA GGT AGG AGA GTA TAA CAA TTA TATA AAT Asn Val Asn Cat TTT TGGA GGT TAT TAT ATG GAA GAG TAT GAA CAA TTA TTA AAT Asn Val Lys Thr Tyr Leu Gly Gly Tyr Ser Lys Thr Leu Asn Gly Ser Tyr Ala Thr Val Lys GTT AAA ACA TAT CTT GGA GGT TAT CAA AAA ACT CTT AAT GGA AGT TAT GCA ACT GTA AAA	3001 3061 3121 3181	Dee Gly Pro Lys Ile Lys Ser Thr Gly Phe Lys Lys Val Thr Thr Pro Gly Ala Ser Gly TTT GGA CCA AMG ATT AMA AGT ACT GGA TTT AMG AMA GTT ACT ACA CCA GGA GGT TCA GGA 1040 Gly Tyr Leu Ala Val Asn Glu Lys Glu Gly Glu Gly Ser Leu Cys Phe Lys Ala Lys Val GGA TAT CTT GCA GTA ANT GAA AMG GAA GGA GAM GGA TCA CTT TOT TTC AMA GCT AMA GTC Thr Lys Phe Gly Leu Tyr Gly Tyr Arg Lys Thr Thr Ser Gly Lys Phe Arg Val Thr 1 ACT AMA TTC GGT CTT TAT GGA TAT AGA AMA ACA ACT TGT GGA GAA GTT AGA GTT ACA ATT 1080 Asp Ser Gln Pro Gly Glu Val Thr Ser Gln Ser Tyr Phe Ser Asp Ser Glu Arg Thr Leu GAC TCA CAA CCA GGT GGA GAT GTA ACT AGC CAMA GTT ATA TTC TGA CTA GAA CAT TT 1080 Phe Tyr Ala His Thr Phe Asp Glu Thr Glu Ala Asn Lys Val His Asn Ile Cva Mer Glu
2101 2161 2221	Lev Val Glu Lys Gly Lys Leu IIe Val Pro Lys Ser Gly Tyr Tyr Thr Leu Phe Met Lys TTA GTA GAA AAA GGT AAA TTG ATA GTT CCA AAA TCT GGA TAT TAT ACA TTG TTT ATG AAA Lu Thr ATA GTA GAA AAA GGT AAA CTG ATA GTT CCA AQA TCA GGA TAT TAT ACA TTG TTT ATG AAA Leu Thr Ata Asp Asp Leu Gly Arg Leu Leu Law Asn IIe Thr Gly Glu Tyr Glu Gln Leu Leu Asp GCA GAT GAT TTA GGA AGA TTG TTA TTG ATA ATT ACT GGA GAG TAT GAA CAA TTA TTA ACA GCA GAT GAT TTA GGA AGA TTG TTA TTG ATA ATT ACT GGA GAG TAT GAA CAA TTA TTA ATA Ma Asp Asp Leu Gly Arg Leu Leu Law Asn IIe Thr GGA GAG TAT GAA CAA TTA TTA ATT GCA GAT GAT TTA GGA AGA TTG TTG TTA TTG ATA ATT ACT GGA GAG TAT GAA CAA TTA TTA ATT Val Asn Asn 760 Val Lys Thr Tyr Leu Gly Gly Tyr Ser Lys Thr Leu Asn Gly Ser Tyr Ala Thr Val Lys GTT AAA ACA TAT CTT GGA GGT TAT TCA AAA ACT CTT AAT GGA AGT TAT GCA ACT GTA AAA IIe Thr 780	3001 3061 3121 3181 3241	Phe Gly Pio Lys lie Lys Ser Thr Gly Phe Lys Lys Val Thr Thr Pro Gly Ala Ser Gly TTT GGA CCA ANG ATT AMA AGT ACT GGA TTT AGA AMA GTT ACT ACA CCA GGT GCT TCA GGA GGA TAT CTT GCA GTA AMT GAA AGG GAA GGA GGA GGA GGA CCA TT GGT TC AAA GCT AMA GTC Lys Phe Gly Leu Yr Gly Tyr Arg Lys Thr Thr Ser Gly Lys Phe Arg Val Thr Lie ACT AMA TTC GGT CTT TAT GGA TAT AGA AMA ACA ACA TCT GGT AGA GTT AGA GTT ACA ATT CTT GCA GTA ATT GGA TAT AGA AMA ACA ACA TCT GGA TAG TT AGA GTT ACA LST AMA TTC GGT CTT TAT GGA TAT AGA AMA ACA ACA TCT GGA AGA TTT AGA GTT ACA ACT AMA TTC GGT CTT TAT GGA TAT AGA AMA ACA ACA TCT GGA AGA GTT AGA GTT ACA LST AMA TTC GGT CTT TAT GGA TAT AGA AMA ACA ACA TCT GGA AGA TTT AGA CTA ACT TG LST ACT AMA TTC GGT CTT TAT GGA TAT AGA AMA ACA ACA TCT GGA AGA CTT AGA CTA ACT TTG LST ACT ACA CCA GGT GAA GTT ACT ACC CCA AGT TAT TTC TCT GAC CT AGA CGA ACT TTG LIO Phe Tyr Ala His Thr Phe Asp Glu Thr Glu Ala Asn Lys Val His Asn Lle Cys Met Glu TCT CTA GCT CAT ACT TTT GAT GGA ACA GAA GTA ACT AMA CT ACA ATT LST Val Val Glu Gly Thr Val Asn Leu Asp IIE ILE GLY Ser Ser
2101 2161 2221 2281	Leu Val Glu Lys Gly Lys Leu IIe Val Pro Lys Ser Gly Tyr Tyr Thr Leu Phe Met Lys TTA GTA GAA AAA GGT AAA TG ATA GTA CCA AAA TCT GGA TAT TAT ACA TG TTT AG AAA TTA GTA GAA AAA GGT AAA GT ATA GTA CA ACA CCA GGA TAT TAT ACA TG TTT AG AAA Leu Thr 740 Ale Asp Asp Leu Gly Arg Leu Leu Leu Asn IIe Thr Gly Glu Tyr Glu Gin Leu Leu Asp GCA GAT GAT TTA GGA AGA TTG TAT ATG AAT ATT ACA GGA GA TAT ATA TA ATA GCA GAT GAT TTA GGA AGA TTG TTA TTA ATG GGA GAG TAT GAA CAA TTA TAT AAT GCA GAT GAT TTA GGA AGA TTG TTA TTA ATG GGA GAG TAT GAA CAA TTA TAT AAT GCA GAT GAT TTA GGA AGA TTG TTA TTA ATG GGA GAG TAT GAA CAA TTA TAT AAT GCA GAT GAT TTA GGA AGA TTG TTA TTA ATG GGA GAG TAT GAA CAA TTA TAT AAT GCA GAT GAT TAT GGA AGG TTG TTG TTA AAT GGA AGA TAT GAA CAA TTA TAT AAT Val Asn 760 Val Lys Thr Tyr Leu Gly Gly Tyr Ser Lys Thr Leu Asn Gly Ser Tyr Ala Thr Val Lys GTT AAA ACA TAT CTT GGA GGT TAT TCA AAA ACT CTT AAT GGA ACT TAT GCA ACT GTA AAA ILeu Glu Lys Asp Val Gly Tyr Pro Phe IIe Leu Tyr Asn Leu Asn Thr Gly Gly Gln Gly TTG GAA AAA GAT AGT GAA TAT CCA TTT ATT CTT TAT ATT TG GAA GAT ATA CGA GGA CAA GGA	3001 3061 3121 3181 3241 3301	Phe Gly Pro Lys Ile Lys Ser Thr Gly Phe Lys Lys Val Thr Thr Pro Gly Ala Ser Gly TTT GGA CCA AG ATT AAA AGT ACT GGA TIT AAG AAA GTT ACT ACA CCA GGT GCT TCA GGA 1040 Gly Tyr Leu Ala Val Asn Glu Lys Glu Gly Glu Gly Ser Leu Cys Phe Lys Ala Lys Val GGA TAT CTT GCA GTA AAT GAA AAG GAAA GGA GAA CAT TOT TTC AAA GCT AAA GT TAA GCT AAT CAT AGA AAG GAA GGA GAA GCA GAA GT Thr Lys Phe Gly Leu Tyr Gly Tyr Arg Lys Thr Thr Ser Gly Lys Phe Arg Val Thr 11e ACT AAA TTC GGT CTT TAT GGA TAT AGA AAA ACA ACT CTG GA AAG TTT AGA GTT ACA ATT 1080 Asp Ser Gln Pro Gly Glu Val Thr Ser Gln Ser Tyr Phe Ser Asp Ser Glu Arg Thr Leu GAC TCA CCA GGT GAA GTT ACT AGA CATA ACT GAA AGT CT GAA CAA ATT TGT ATC Phe Tyr Ala His Thr Phe Asp Glu Thr Glu Ala Asn Lys Val His Aan Ile Cys Met GIA Val Val Glu Gly Tr Val Asn Leu Asp Ile Ile Gly Ser Ser GTT GTT GAA GGA ACA GTT AAT CTT GAT ACT ATT GGT TCT TA AGGTLAACATT GGA 3444

Figure 3. Inferred amino acid sequence and nucleotide sequence of coding region and flanking region obtained from genomic clone pBSgM17-1. The sequence of the internal EcoRI fragment was identical in both genomic clones (pBSgM17-1/2) and the cDNA clone  $\lambda$ cM17. Shown below is the partial nucleotide sequence of PCR amplification products derived from nonpathogenic isolate REF291. Nucleotide substitutions are underlined and amino acid substitutions are indicated below the partial sequence derived from REF291.



Figure 4. Transfer blot of E. histolytica HM1:IMSS RNA probed with cDNA clone  $\lambda cM17$  indicates a single hybridizing band migrating at  $\sim$ 3 kb. Hybridization conditions were 50% formamide, 0.2× SSC, 42°C. Autoradiography shown required a 72-h exposure.

with BglII and EcoRV in single and double digests, was probed with BamHI-BglII, BglII-EcoRV, and EcoRV fragments of  $\lambda$ cM17, only unique restriction fragments hybridized with each probe (data not shown). Furthermore, the nucleotide sequence of both genomic clones and the cDNA clone is identical.

Detection of Sequences Related to Gene M17 in Nonpathogenic E. Histolytica and Mapping of RFLPs. Western blot analysis suggested that the 125-kD antigen or a closely related antigen that shared the epitope recognized by poly- and monoclonal antisera was found in both pathogenic and nonpathogenic E. histolytica isolates as well as E. histolytica-like Laredo. By Southern blot analysis, even under low stringency hybridization and wash conditions (25% formamide, 2 × SSC,

37°C) sequences related to M17 were difficult to detect in nonpathogenic E. histolytica isolates and E. histolytica-like Laredo (data not shown). To confirm the presence of a closely related gene in nonpathogenic amoebae, two fragments spanning most of the sequence contained within the cDNA clone  $\lambda$ cM17 were amplified in a PCR using oligonucleotide pairs SRO19/SRO22, and SRO18/SRO21 as primers on genomic template DNA derived from nonpathogenic isolate REF 291. By nucleotide sequence analysis of the two subcloned PCR amplification products, REF 291 had 145 nucleotide substitutions over 1410 residues (10.3%) as compared with the sequence of  $\lambda cM17$  (HM1:IMSS) (Fig. 3). These substitutions result in 57 amino acid differences per 470 residues (12.1%). A computer search of published protein sequences with the entire 3,345 bp M17 gene sequence revealed that the internal gene fragment represented by the  $\lambda cM17$  insert encoded a protein sequence similar to that deduced for a DNA fragment isolated from nonpathogenic and pathogenic strains of E. histolytica by Tannich et al. (29) and proposed by these authors as a potential diagnostic probe for strain differentiation. Specifically, when we compared the amino acid sequences of Tannich et al. with that of  $\lambda cM17$  we detected five substitutions between pathogenic HM1:IMSS isolates (1%) (Fig. 6). As the nucleotide sequence of the DNA fragment was not published by Tannich et al., we infer from the amino acid sequence that at least three of these five differences between the E. histolytica HM1: IMSS laboratory strains must have arisen from more than one nucleotide substitution and are therefore unlikely to represent cDNA synthesis or sequencing artifacts. When the 470 amino acid sequence derived from the PCR product of nonpathogenic isolate REF291 was compared with isolate SAW 1734 (29), six amino acid substitutions (1.3%) were detected (Fig. 6). Over the same 470 amino acids, 61 amino acid residues (12.9%) differ among the pathogenic HM1:IMSS (29) and nonpathogenic SAW 1734 (29) strains (Fig. 6). Overall there are 65 variable residues over a stretch of 470 amino acids (13.8%) when these four isolates were compared (Fig. 6).

As the partial M17 amino acid sequences of nonpathogenic strains SAW 1734 (29) and REF 291 were significantly more similar to one another than to their pathogenic counterparts, PCR amplification of the same gene fragments from six additional strains was undertaken to examine the possibility of defining RFLPs that could reliably differentiate pathogenic from nonpathogenic amoebal isolates. Using oligonucleotide primers SRO19, SRO22, SRO18, and SRO21, PCR products of the same size were amplified from genomic DNA of strains SD116, SD4, Nos. 43, 44, 46, and HK9. Based

respectively,

M17*	A G A A A T A T A A A A G A A T G T T A A A A	С А А А С А Т А А А А А А А Т А А G T G	Figure 5. Alignment of the 5' flanking se-
ACTIN		G T T A A C T C C A A A C A A A	ferredoxin (45); a likely Goldberg-Hogness con-
FERREDUXIN	>>>	>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>	sensus sequence at -29, -31, -32, respectively
M17*	ATTTAAAGTGTTTTTTAAAAAACTAATTATT	CATAAATTAAAGTTATGI	is boxed and in bold face. The 5' end of the
ACTIN*	с а с <u>т <b>а т т т а а а</b></u> д а с т д а с а а а а а с т д а а т т с а д <b>т а т т т а а а</b> д а т с а т а а т д а а с т д а а т т а а	A A T C A T T A A T T A A T A T G G	similarities around the cap site are boxed and
FERREDOXIN	G T <u>C A T T T A A A</u> T A C A A C A A A T T A A T C T T T T 7 << << << << << << << << << << << << <<	Met	shaded.

884 An Immuno-dominant Variable Surface Antigen from Entamoeba histolytica

on our nucleotide sequence of these fragments from HM1: IMSS and REF291, we predicted that restriction endonucleases EcoRV, SspI, PvuII, AccI, and HincII among others would cleave the PCR products into restriction fragments which might be expected to correlate with the pathogenic or nonpathogenic phenotype of the isolate. An example of such an analysis with EcoRV and SspI is presented in Fig. 7. A restriction site for EcoRV is absent in nonpathogenic No. 43, No. 44, and REF291 but present in nonpathogenic SD116 and Laredo as well as pathogenic HM1:IMSS, HK9, SD4, and No. 46 (Fig. 7). Digestion with restriction endonuclease SspI shows a distinct pattern for pathogenic (HM1:IMSS, HK9, SD4, 46) versus non-pathogenic (No. 43, No. 44, SD116, REF291) strains with the exception of E. histolytica-like Laredo, which would appear pathogenic by this criterion (Fig. 7). Similarly, restriction with AccI distinguishes pathogenic from nonpathogenic isolates with the exception of Laredo, which appears to have an additional restriction site for the enzyme. HincII digestion shows the same restriction fragments in nonpathogenic isolates No. 43, No. 44, and REF291 and pathogenic isolate SD4 but no restriction sites in commensal Laredo and pathogenic isolates No. 46, HM1:IMSS, and HK9.

### Discussion

It has long been known that amoebiasis is a spectral disease, asymptomatic infections with "nonpathogenic" amoebae and life-threatening infections with "pathogenic" amoebae defining opposite ends of the spectrum. With the availability of effective treatment regimens, early diagnosis is crucial for the prevention of disease and transmission. However, much controversy has centered on benefits and drawbacks of initiating therapy in asymptomatic infections. Thus, recent investigations have focussed on a molecular genetic analysis of virulence and the definition of marker molecules that have high predictive value and can be applied in a clinically feasible fashion.

We have identified a variable, immunodominant 125-kD surface antigen in *E. histolytica* HM1:IMSS. The amino acid sequence inferred from the nucleotide sequence of the coding region of the 125-kD antigen is unusual with respect to its high Asn (90 = 8.2%), Tyr (70 = 6.3%), and hydroxyl amino acid residue (Ser, 85 = 7.6%; Thr, 90 = 8.1%) content. While a total of 17 *N*-linked glycosylation sites suggests that the 125-kD antigen may be glycosylated, Western blot analysis shows that this antigen migrates as a compact band on SDS-PAGE. A distinctly hydrophobic NH<sub>2</sub>-ter-

HM1:IMSS* HM1:IMSS# REF 291* SAW 1734#	P 1 P 1 P 1 P 1	TL TL TL TL	N I N I N I			R V R V R V R V	D D D D	A ( A ( A (	G A G A G A G A	A A A A	v v v v	A ' A ' A 1 A 1	Y V Y V F V F V	G D G G	R R R R	W W W	F 1 F 1 F 1	000	N N H H	P 9 P 9 P 9	s D 5 D 5 D 5 D	W W W	A A A A	A		v v v v	G I G I N I N I	<pre>&lt; D &lt; D &lt; D &lt; C </pre>	G G R G	L L L L	N   N   N   N	YYS	G G G G	N N N N	W ( W ( W (	3   3   3   3   3		н н н	E M E M E M	AN AN AN
HM1:IMSS* HM1:IMSS# REF 291* SAW 1734#	нн нн нн нн	M Q M Q M Q M Q	G 1 G 1 G 1 G 1	ΓΥ Υ ΓΥ	L 1 L 1 L 1 L 1	K G K G R G R G	G G G	NV NV NO	V G G G V G		S S K K	N N E E	P G P G P G	EEEE	E E E		2222	: V : V : V : V	M M M	T 9 T 9 T 9 T 9	5 I 5 I 5 I 5 I	ZZZZ	Y Y Y Y		LY LY LY	T T T T	N N N N	A   A   A   A	C C C	H H H H	R N R N R F R N		G G G G	L L L L	S ( S ( S ( S (	3 V 3 V 3 V 3 V 3 V	/ N / N / N / N	Y Y Y Y		5 C 5 C 5 C 5 C
HM1:IMSS* HM1:IMSS# REF 291* SAW 1734#	G Y G Y G Y G Y	ST ST ST ST	1 ) 1 ) 1 ) 1 )	( K ( K ( K ( K	I I I I	L K L K L N	G G G G G	E M E M E M		0000	P P P	HL 1 HL 1 HL 1 HL 1	LR LR LR LR	5 5 5 5	Y Y Y Y			f A f A f A A	H H H H	A ] A ] A ] A ]	PG FG FG FG	T T T	D D D D	T   T   T   T	L I L 1 L 1 L 1	A A A A		/ K / K / K	s s s s	Y Y Y Y	Y ( Y ( Y ( Y (	) L ) L ] L ] L	W W W	Y Y Y Y			F F Y Y	E E E	S I S I G I	k Y K Y E Y E Y
HM1:IMSS* HM1:IMSS# REF 291* SAW 1734#	5 1 5 1 5 1 5 1	K R K R K R K R	D 5 D 5 D 5	T T T T	S S S S	AF AF AF AF	C C C C	L   L   L   L	LA LA LA	A A A A	L L I I	V V A	TK TK TK TK	R R R R	D D D D	T   T   T   T	RY RY RY RX	( L ( L ( L	C C C C	5   5   5   5	LF LF LF	K K K	Y   Y   Y   Y	D D D D		s s Q Q	N N N N N N	/ S / S / S	E E E	A A A A	1 ) 1 ) 1 ) 1 )		M M M	NNNN	Y   Y   Y   Y	(2 9 (2 9 (2 9 (2 9	Y Y Y Y	Y Y Y Y	P   P   P   P	F F F F F F F F
HM1:IMSS* HM1:IMSS# REF 291* SAW 1734#	N L N L N V N V	Y A Y A Y A Y A Y A	M 5 M 5 M 5 M 5	Y Y Y Y Y		G N G N G N G N	Y Y Y Y	Y ( Y ( Y ( Y (	G R G R G R O R	P P T	Y Y Y Y	K K K	IP IP IP IP	Y Y Y Y	G G G G	R R T T	TR TR TR	t L t E t L t L	N N N	F F F F	ΤΑ ΤΛ ΤΑ ΤΑ	, T , T , T , T	T C T T	A S A A		P P P P	K K S S	А Т А Т А Т А Т	S 5 5 5 5	V V V	9 ) 9 ) 5 F 5 )		I I I	K K K	5 ( 5 ( 5 (		. T . T . T	K K K	G ) G ) G ) G )	( ). ( ). ( ). ( ).
HM1:IMSS* HM1:IMSS# REF 291* SAW 1734#	E R E R E Q E Q	V E V E V E V E	D D E E		Y Y Y Y Y	DY DY DY DY	T T T T	P I P I P M P M	F F F F N F N F	G G G	I A A A	E E D	e n e n e n	D D D	T T T T	F F F F	V L V L V L V L	. N . N . N	I J I I			v v v v		G 1 G 1 G 1 G 1	e k e k e k	v v v v	H   H   H   H	I E I E I E	0000	E E D		5 T 5 T 5 T 5 T	F F F	E E E	L   L   L   L		H H H H	0 0 0 0 0	V 1 V 1 V 1 V 1	E Y E Y E Y E Y
HM1:IMSS* HM1:IMSS# REF 291* SAW 1734#	E V E V E V E V	Y K Y K Y K Y K		Q Q K K	T T T T	R D R D K D K D	M M M	A ( A ( E ( E (	2 A 2 A 2 A 2 A 2 A	I I L L	N N N	I I T T		N N N	K K K	TI TI T! T!	R N R N S N S N	D D Y Y	T T T T	G 1 G 1 G 1 C 1	R A R A T S T S	S S T T	F F F F			C G G C	T ) T ) N )	( N ( D ( D	D D D D	G G G	5 N 5 N F N F N		S S S S	M M M		/ 1 / 1 / 1	K K K	G G G	K I K I K I	
HM1:IMSS* HM1:IMSS# REF 291* SAW 1734#	V P V P V P V P	KS KS TS TS	G Y G Y G Y G Y	Y Y Y Y Y	TI TI TI TI	L F L F L F	M M M M	K / K / K /		D D D			{ L { L { L { L { L	L L L L	L L L L				E E E	Y 1 Y 1 Y 1 Y 1	E Q E Q E Q E Q	L L L L			/ K / K / K	T T T T	Y I Y I Y I	000	C G G G	Y Y Y Y Y Y	5 K 5 K 5 K	T T T T	L L I I	N N N	G : G : G 1 G 1	5 ) 5 ) 7 )	A A A A	T T T T	V I V I V I V I	( L ( I. ( L ( I.
HM1:IMSS* HM1:IMSS# REF 291* SAW 1734#	E K E K E K E K	D V D V D T D T	G) G) E) E)	P P P P	F F F F	I L I L I L I L	Y 1 Y 1 Y 1 Y 1	N I N I N I N I	. N . N . N	T T T T	G I G I G I		2 G 2 G 2 G 2 G	F F F F	1 1 1 1	R R R R	1 G 1 G 1 G 1 G	Y Y Y Y	C C C C	Y H Y H Y Q Y Q	H G H G Q G Q G	T T T T		E 5 2 5 2 5	1 S 1 S 1 S 5 S	v v v		/ S / S / S	K K K			S S L L	D D D D	1 I 1 I 1 I						

Figure 6. Alignment of the amino acid sequences inferred from nucleotide sequences of cDNA and genomic clones (HM1:IMSS\*) and of PCR amplification products (REF291\*) with those published by Tannich et al. (HM1:IMSS#, SAW 1734#) (29). Conserved amino acids are shaded, variable residues differentiating pathogenic from nonpathogenic isolates are in bold face, and additional variable residues are in plain text.

885 Edman et al.



Figure 7. Restriction endonuclease (EcoRV lower case letters, SspI capital letters) digests of PCR products generated by amplification of genomic DNA from *E. histolytica* isolates/strains using oligonucleotide primers SRO 18 + 21 and SRO19 + 22. (a/A) No. 43, (b/B) No. 44, (c/C) SD116, (d/D) REF291, (e/E) *E. histolytica*-like Laredo, (f/F) No. 46, (g/G) HK9, (h/H) = HM1:IMSS.

minal region of 35 amino acids may serve as anchor or signal sequence. Compared with known prokaryotic and eukaryotic signal sequences this region contains an unusually long (20 amino acids) NH2-terminal (n) region with a single positively charged residue, an 8 amino acid long hydrophobic core (h) region and a 7 amino acid long polar COOH-terminal (c) region with an amino acid composition similar to those seen in other signal sequences (46). By extrapolation this would imply that the antigen is either a peripheral membrane protein or it may be anchored in the membrane by other means such as a glycophospholipid anchor. Alternatively, the hydrophobic NH2-terminal itself may serve to anchor the antigen in the membrane with the COOH-terminal externally exposed as no additional trans membrane domains could be discerned. From searches of protein and nucleic acid sequence data banks, a small domain was identified sharing sequence similarity with the  $\beta$  chain of the human fibronectin receptor, the  $\beta$ -1 chain of the mouse integrin, and band 3 precursor of the chicken integrin (Fig. 8). However intriguing, the functional significance of this similarity will need to be assessed by generation of antibodies to this domain for use in attachment or invasion assays.

Tannich et al. (29) had reported a sequence of 1.9 kb that differed substantially between pathogenic and nonpathogenic E. histolytica and also suggested that RFLPs in this gene fragment would allow the differentiation of pathogenic from nonpathogenic E. histolytica. Examination of the Tannich sequence revealed that it was derived from an internal fragment of M17, the gene sequenced in its entirety in this paper. While human sera and mAb FA7 demonstrated the presence of the 125-kD antigen in a number of nonpathogenic and pathogenic amoebal isolates (Fig. 1), DNA hybridization data with the  $\lambda cM17$ probe suggested that there were regions of substantial sequence variability in other portions of this molecule. For this reason, fragments of this antigen within a variable sequence region were amplified by the PCR to search for RFLPs that correlated with phenotype. If in fact RFLPs could be used to differentiate pathogenic from nonpathogenic strains, then in combination with FA7, PCR could provide a potent diagnostic protocol.

Based on the nucleotide sequence differences of strains HM1:IMSS and REF291, presented in Fig. 3, five restriction endonucleases were chosen that should have yielded RFLPs which correlate with the pathogenic and nonpathogenic phenotype of the isolate from which the PCR fragments were derived. PCR fragments from four fresh polyxenic nonpathogenic isolates of zymodeme I (SD116, No. 43, No. 44) or III (REF291), from axenized Laredo, from two fresh polyxenic pathogenic isolates (SD4, No. 46) of zymodeme II and two axenized established pathogenic laboratory strains (HM1: IMSS, HK9) of zymodeme II were subjected to this analysis. Our results provide strong evidence that the M17 antigen is a highly variable protein and that distinct sets of amino acid substitutions exist in pathogenic versus nonpathogenic strains of Entamoeba. These differences provide the basis for RFLPs which, in the limited sampling of this study, are correlated with pathogenic and nonpathogenic isolates. The exception to this correlation is the E. histolytica-like Laredo strain that is often used as a prototype nonpathogen in laboratory studies. Although Laredo was first isolated as a human commensal it appears morphologically more like free-living amoebae and belongs to an unusual zymodeme that is rarely found in patient isolates. Thus, Laredo's RFLP pattern may not preclude the use of this criterion in the clinical context.

M17 nonpathogenic

M17 pathogenic

beta chain, human FNR (B27079) band 3 precursor chick integrin(IJCH3) beta chain mouse integrin (SO1659)



Figure 8. Similarity of a small region of the amino acid sequence inferred from the M17 nucleotide sequence to the  $\beta$  chain of the human fibronectin receptor, the band 3 precursor of the chick integrin, and the  $\beta$ -chain of the mouse integrin.

Given the rather limited sample size tested here, it is, however, noteworthy that none of the pathogenic isolates revealed a nonpathogenic RFLP pattern in this analysis.

While gene fragments from HM1:IMSS analyzed in these experiments were derived from the same original isolate, they have been propagated in different laboratories for some time. The extent of sequence differences indicates that a low degree of variation probably occurs within this gene family in the absence of selective pressure by the host immune system. Although the position and nature of most of the amino acid substitutions are conserved, it is clear that overall extensive interstrain variability and the modest intrastrain variability among strains with the same phenotype necessitate that RFLPs be validated on large numbers of amoebal isolates. Zymodeme characterization and the use of additional diagnostic markers will be required to develop a reliable set of criteria for the differentiation of pathogenic from nonpathogenic E. *histolytica*.

To further understand the significance of the 125-kD surface antigen sequence variation and its potential role in pathogenesis, sequence analysis of the entire gene from several nonpathogenic and pathogenic isolates has been undertaken. Evolutionary analysis of the small ribosomal subunit gene sequences from several nonpathogenic *E. histolytica* isolates and comparison to the known small ribosomal subunit gene sequences of *E. histolytica* HM1:IMSS, *E. histolytica*-like Laredo and *E. invadens* (manuscript in preparation) will improve our ability to interpret the significance of the variability in the 125-kD antigen in view of the degree of evolutionary divergence between these different *Entamoeba* strains/species.

This work was supported by grants from the John D. and Catherine T. MacArthur Foundation, the World Health Organization/Rockefeller Foundation Tropical Diseases Partnership, and by Conacyt of Mexico.

Address correspondence to Dr. Nina Agabian, University of California School of Medicine, 3333 California St., San Francisco, CA 94143-1204.

Received for publication 4 June 1990.

## References

- Keene, W.E., M.G. Pettit, S. Allen, and J.H. McKerrow. 1986. The major neutral proteinase of *Entamoeba histolytica*. J. Exp. Med. 163:536.
- Reed, S.L., W.E. Keene, J.H. McKerrow, and I. Gigli. 1989. Cleavage of C3 by a neutral cysteine protease of *Entamoeba* histolytica. J. Immunol. 143:189.
- 3. Otto, J., and E. Werries. 1989. Specificity of a cysteine proteinase of *Entamoeba histolytica* against various unblocked synthetic peptides. *Mol. Biochem. Parasitol.* 33:257.
- Lynch, E.C., I.M. Rosenberg, and C. Gitler. 1982. An ionchannel forming protein produced by *Entamoeba histolytica EMBO (Eur. Mol. Biol. Organ.) J.* 1:801.
- Young, J.D.-E., T.M. Young, L.P. Lu, J.C. Unkeless, and Z.A. Cohn. 1982. Characterization of a membrane pore-forming proteins from *Entamoeba histolytica*. J. Exp. Med. 156:1677.
- 6. Young, J.D.-E., and Z.A. Cohn. 1985. Molecular mechanisms of cytotoxicity mediated by *Entamoeba histolytica*: characterization of a pore-forming protein (PFP). J. Cell. Biochem. 29:299.
- Chadee, K., W.A. Petri, D.J. Innes, and J.I. Ravdin. 1987. Rat and human colonic mucins bind to and inhibit adherence lectin of *Entamoeba histolytica*. J. Clin. Invest. 80:1245.
- Petri, W.A., R.D. Smith, P.H. Schlesinger, C.F. Murphy, and J.I. Ravdin. 1987. Isolation of the galactose-binding lectin that mediates the in vitro adherence of *Entamoeba histolytica*. J. Clin. Invest. 80:12387.
- Petri, W.A., M.P. Joyce, J. Broman, R.D. Smith, C.F. Murphy, and J.I. Ravdin. 1987. Recognition of the galactose- or N-acetylgalactoseamine-binding lectin of *Entamoeba histolytica* by human immune sera. *Infect. Immun.* 55:2327.
- 10. Petri, W.A., and J.I. Ravdin. 1987. Cytopathogenicity of Entamoeba histolytica: the role of amebic adherence and contact-

dependent cytolysis in pathogenesis. Eur. J. Epidemiol. 3:123.

- 11. Chadee, K., M.L. Johnson, E. Orozco, W.A. Petri, and J.I. Ravdin. 1988. Binding and internalization of rat colonic mucins by the galactose/N-acetyl-D-galactoseamine adherence lectin of *Entamoeba histolytica*. J. Infect. Dis. 158:398.
- Petri, W.A., M.D. Chapman, T. Snodgrass, B.J. Mann, J. Broman, and J.I. Ravdin. 1989. Subunit structure of the galactose and N-acetyl-D-galactoseamine-inhibitable adherence lectin of *Entamoeba histolytica*. J. Biol. Chem. 264:3007.
- Petri, W.A., J. Broman, G. Healy, T. Quinn, and J.I. Ravdin. 1989. Antigenic stability and immunodominance of the Gal/ GalNac adherence lectin of *Entamoeba histolytica*. Am. J. Med. Sci. 297:163.
- Rosales-Encina, J.L., I. Meza, A. López-De-León, P. Talamás-Rohana, and M. Rojkind. 1987. Isolation of a 220-kilodalton protein with lectin properties from a virulent strain of *Ent*amoeba histolytica. J. Infect. Dis. 156:790.
- Meza, I., F. Cázares, J.L. Rosales-Encina, P. Talamás-Rohana, and M. Rojkind. 1987. Use of antibodies to characterize a 220kilodalton surface protein from *Entamoeba histolytica*. J. Infect. Dis. 156:798.
- Torian, B.E., S.L. Reed, B.M. Flores, C.M. Creely, J.E. Coward, K. Vial, and W.E. Stamm. 1990. The 96-kilodalton antigen as an integral membrane protein in pathogenic and nonpathogenic isolates. *Infect. Immun.* 58:753.
- Torian, B.E., S.L. Reed, B.M. Flores, J. Plorde, and W.E. Stamm. 1989. Serologic response to the 96,000-Da surface antigen of pathogenic *Entamoeba histolytica*. J. Infect. Dis. 159:794.
- Torian, B.E., S.A. Lukehart, and W.E. Stamm. 1987. Use of monoclonal antibodies to identify, characterize, and purify a 96,000-dalton surface antigen of pathogenic *Entamoeba histo*-

lytica. J. Infect. Dis. 156:334.

- Sargeaunt, P.G. 1987. The reliability of Entamoeba histolytica zymodemes in clinical diagnosis. Parasitol. Today. 3:40.
- Mirelman, D., R. Bracha, A. Chayen, A. Aust-Kettis, and L.S. Diamond. 1986. *Entamoeba histolytica*: effect of growth conditions and bacterial associates on isoenzyme patterns and virulence. *Exp. Parasitol.* 62:142.
- Mirelman, D. 1987. Ameba-bacterium relationship in amebiasis. *Microbiol. Rev.* 51:272.
- Mirelman, D. 1987. Effect of culture conditions and bacterial associates on the zymodemes of *Entamoeba histolytica*. Parasitol. Today. 3:37.
- Sargeaunt, P.G., J.E. Williams, and J.D. Grene. 1978. The differentiation of invasive and non-invasive *Entamoeba histolytica* by isoenzyme electrophoresis. *Trans. R. Soc. Trop. Med. Hyg.* 72:519.
- Sargeaunt, P.G., and J.E. Williams. 1978. Electrophoretic isoenzyme patterns of the pathogenic and non-pathogenic intestinal amoebae of man. *Trans. R. Soc. Trop. Med. Hyg.* 73:225.
- Sargeaunt, P.G., J.E. Williams, and R.A. Neal. 1980. A comparative study of *Entamoeba histolytica* (NIH:200, HK9, etc.) "E. histolytica-like" and other morphologically identical amoebae using isoenzyme electrophoresis. *Trans. R. Soc. Trop. Med. Hyg.* 74:469.
- Moss, D.M., and H.M. Mathews. 1987. A fast electrophoretic isoenzyme technique for the identification of invasive and noninvasive *Entamoeba histolytica* and "*E. histolytica*-like" Organisms. *J. Protozool.* 34:253.
- Garfinkel, L., M. Giladi, M. Huber, C. Gigler, D. Mirelman, M. Revel, and S. Rozenblatt. 1989. DNA probes specific for *Entamoeba histolytica* possessing pathogenic and non-pathogenic zymodemes. *Infect. Immun.* 57:926.
- Samuelson, J., R. Acuna-Soto, S. Reed, F. Biagi, and D. Wirth. 1989. DNA hybridization probe for clinical diagnosis of *Ent*amoeba histolytica. J. Clin. Microbiol. 27:671.
- Tannich, E., R.D. Horstmann, J. Knobloch and H.H. Arnold. 1989. Genomic DNA differences between pathogenic and nonpathogenic *Entamoeba histolytica*. Proc. Natl. Acad. Sci. USA. 86:5118.
- Strachnan, W.D., P.L. Chiodini, W.M. Spice, A.H. Moody, and J.P. Ackers. 1988. Immunological differentiation of pathogenic and non-pathogenic isolates of *Entamoeba histolytica*. Lancet. i:561.
- Gillin, F.D., and L.S. Diamond. 1978. Clonal growth of Entamoeba histolytica and other species of Entamoeba in agar. J. Protozool. 25:539.
- 32. Diamond, L.S., D.R. Harlow, and C.C. Cunnick. 1978. A

new medium for the axenic cultivation of Entamoeba histolytica and other Entamoeba. Trans. R. Soc. Trop. Med. Hyg. 72:431.

- 33. Meza, I., M. De La Garza, M.A. Meraz, B. Gallegos, M. De La Torre, M. Tanimoto, and A. Martinez-Palomo. 1986. Isoenzyme patterns of *Entamoeba histolytica* isolates from asymptomatic carriers: use of gradient acrylamide gels. Am. J. Trop. Med. Hyg. 35:1134.
- Aley, S.B., W.A. Scott, and Z.A. Cohn. 1980. Plasma membrane of *Entamoeba histolytica*. J. Exp. Med. 152:391.
- 35. Meza, I., M. Sabanero, F. Cazares, and J. Bryan. 1983. Isolation and characterization of actin from *Entamoeba histolytica*. J. Biol. Chem. 258:3936.
- Edman, U., I. Meza, and N.M. Agabian. 1987. Genomic and cDNA actin sequences from a virulent strain of *Entamoeba* histolytica. Proc. Natl. Acad. Sci. USA. 84:3024.
- Gubler, U., and B.J. Hoffman. 1983. A simple and very efficient method for generating cDNA libraries. Gene (Amst.). 25:263.
- Morgan, D.O., J.C. Edman, D.N. Standring, V.A. Fried, M.C. Smith, R.A. Roth, and W.J. Rutter. 1987. Insulin-like Growth Factor II Receptor as a Multifunctional Binding Protein. *Nature (Lond.)*. 329:301.
- Young, R.A., and R.W. Davis. 1983. Efficient isolation of genes using antibody probes. Proc. Natl. Acad. Sci. USA. 80:1194.
- Weinberger, C., S.M. Hollenberg, E.S. Ong, J.M. Harmon, S.T. Brower, J. Cidlowski, E.B. Thompson, M.G. Rosenfeld, et al. 1985. Identification of human glucocorticoid receptor complementary DNA clones by epitope selection. *Science (Wash.* DC). 228:740.
- Short, J.M., J.M. Fernandez, J.A. Sorge, and W.D. Huse. 1988.
   λZAP: a bacteriophage lambda expression vector with *in vivo* excision properties. *Nucleic. Acids Res.* 16:7583.
- Maniatis, T., E.F. Fritsch, and J. Sambrook. 1981. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Press, Cold Spring Harbor, New York. 545 pp.
- Meraz, M.A., U. Edman, N. Agabian, and I. Meza. 1989. Surface molecules of *Entamoeba histolytica* with immunodominant characteristics. J. Cell Biol. 107:746a. (Abstr.)
- Huber, M., L. Garfinkel, C. Gitler, D. Mirelman, M. Revel, and S. Rozenblatt. 1987. Entamoeba histolytica: cloning and characterization of actin cDNA. Mol. Biochem. Parasitol. 24:227.
- 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.
- von Heijne, G. 1985. Signal sequences, the limits of variation. J. Mol. Biol. 184:99.