IMP Dehydrogenase from *Pneumocystis carinii* as a Potential Drug Target

MARY JEANNE O'GARA,¹ CHAO-HUNG LEE,² GEOFFREY A. WEINBERG,³[†] JENIECE M. NOTT,¹ AND SHERRY F. QUEENER¹*

Departments of Pharmacology & Toxicology,¹ Pathology,² and Pediatrics,³ Indiana University School of Medicine, Indianapolis, Indiana 46202

Received 31 May 1996/Returned for modification 6 September 1996/Accepted 17 October 1996

Mycophenolic acid, a specific inhibitor of IMP dehydrogenase (IMPDH; EC 1.1.1.205), is a potent inhibitor of *Pneumocystis carinii* growth in culture, suggesting that IMPDH may be a sensitive target for chemotherapy in this organism. The IMPDH gene was cloned as a first step to characterizing the enzyme and developing selective inhibitors. A 1.3-kb fragment containing a portion of the *P. carinii* IMPDH gene was amplified by PCR with two degenerate oligonucleotides based on conserved sequences in IMPDH from humans and four different microorganisms. Northern hybridization analysis showed the *P. carinii* IMPDH mRNA to be approximately 1.6 kb. The entire cDNA encoding *P. carinii* IMPDH was isolated and cloned. The deduced amino acid sequence of *P. carinii* IMPDH shared homology with bacterial (31 to 38%), protozoal (48 to 59%), mammalian (60 to 62%), and fungal (62%) IMPDH enzymes. The IMPDH cDNA was expressed by using a T7 expression system in an IMPDH-deficient strain of *Escherichia coli* (strain S ϕ 1101). *E. coli* S ϕ 1101 cells containing the *P. carinii* IMPDH gene were able to grow on medium lacking guanine, implying that the protein expressed in vivo was functional. Extracts of these *E. coli* cells contained IMPDH activity that had an apparent K_m for IMP of 21.7 \pm 0.3 μ M and an apparent K_m for NAD of 314 \pm 84 μ M (mean \pm standard error of the mean; n = 3), and the activity was inhibited by mycophenolic acid (50% inhibitory concentration, 24 μ M; n = 2).

Pneumocystis carinii is an opportunistic pathogen that has become a serious threat to immunodeficient patients. *P. carinii* pneumonia was the leading cause of morbidity and mortality in AIDS patients in the United States and remains common even with widespread use of prophylaxis (30). Chemotherapy for this pneumonia is limited by side effects associated with the compounds considered drugs of choice. Since the biology of *P. carinii* is poorly understood, chemotherapy designed specifically for this organism has lagged behind drug development for other pathogens.

Because of the critical need for new therapy, some of the genes that have been cloned from *P. carinii* are being evaluated as potential targets for new or improved anti-*Pneumocystis* drugs. The targets being studied and the compounds that effectively inhibit these gene products include folic acid synthase (sulfonamides), dihydrofolate reductase (antifolates such as trimethoprim, trimetrexate, and piritrexim), β -tubulin (benzimidazoles), and ATP-dependent topoisomerase (pentamidine and its analogs) (30). A self-splicing group I intron (35) that was found within the 16S and the 26S rRNA genes of *P. carinii* but that was found to be absent from those of humans is also being pursued as a therapeutic target. The splicing of this intron is inhibited by some aminoglycosides, tetracycline, pentamidine, and L-arginine (26), but only pentamidine has in vivo activity against the organism.

IMP dehydrogenase (IMPDH; EC 1.1.1.205) is an exploitable target for chemotherapy in several organisms (3, 41). IMPDH catalyzes the NAD⁺-dependent conversion of IMP to XMP (Fig. 1). In many organisms this reaction is the ratelimiting step for de novo synthesis of guanine nucleotides (18). If the pathway is blocked by an IMPDH inhibitor, DNA and RNA synthesis will be inhibited, resulting in death of the cell or organism. The studies reported here were designed to evaluate IMPDH as a potentially novel drug target in *P. carinii*.

MATERIALS AND METHODS

P. carinii culture screen. Culture of *P. carinii* over 7 days has been used extensively to screen compounds with activity against the organism (30). The method used for these studies has been described previously (4, 8), with inoculation of 24-well plates with approximately 7×10^5 organisms per ml of medium in wells containing established monolayers of HEL cells (human embryonic lung fibroblasts; HEL-299; American Type Culture Collection). *P. carinii* were evaluated by direct counting; host cell sloughing was monitored as an index of toxicity.

PCR. The oligonucleotide primers used for PCR were based on conserved amino acid residues in known IMPDH proteins. The primer sequences were 5'-ATGGA(T/C)ACNGTNACNGA-3' (sense primer based on MDTVTE) and 5'-GA(A/T)CCCAT(A/T)CC(A/T)CG (A/G)TA-3' (antisense primer based on YRGMGS). The primers were designed to minimize degeneracy in the third position of codons based on the codon biases of nucleotide sequences of P. carinii dihydrofolate reductase (14), thymidylate synthase (15), and β -tubulin (13). Genomic DNA was isolated from P. carinii grown in spinner flasks (25). For DNA amplification, 0.5 µg of P. carinii genomic DNA served as template for PCR (32). Samples were denatured for 10 min in boiling water and chilled on ice, and the remaining reaction components were added. The 100-µl reaction mixture contained 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 0.2 mM (each) dATP, dCTP, dGTP, and dTTP, 0.2 µM (each) primers MDTVTE and YRGMGS, 2.5 U of AmpliTaq DNA polymerase, and 0.5 to 3.0 mM MgCl₂. The negative control was identical except that the DNA template was replaced with water. Samples were overlaid with 50 to 75 µl of mineral oil to prevent evaporation during thermal cycling. PCR conditions for DNA amplification were (i) an initial denaturation at 94°C for 10 min, (ii) 35 cycles consisting of denaturation at 94°C for 1 min, followed by primer annealing at 41°C for 1 min and extension for 3 min at 72°C, and (iii) a final 10-min extension at 72°C. Fifteen microliters of each PCR sample was analyzed by electrophoresis on 1% agarose. Amplified fragments were purified by using Wizard PCR Preps (Promega, Madison, Wis.) and were subcloned into pCRII (Invitrogen, LaJolla, Calif.) for sequencing.

PFGE, **Southern analysis, and DNA sequencing**. Pulsed-field gel electrophoresis (PFGE) was performed as described previously (43). The PFGE-resolved *P. carinii* chromosomes were transferred to a Nytran Maximum Strength Plus membrane (Schleicher & Schuell, Keene, N.H.) by the method of Southern (36). The blot was probed with a genomic fragment of the *P. carinii* IMPDH gene amplified by PCR. Probes were radiolabeled with $[\alpha$ -³²P]dCTP according to the manufac-

^{*} Corresponding author. Mailing address: Indiana University School of Medicine, 635 Barnhill Dr., MS A519, Indianapolis, IN 46202. Phone: (317) 274-1563. Fax: (317) 274-1560. E-mail: queenes@ indyunix.iupui.edu.

[†] Present address: University of Rochester, Rochester, NY 14627.



FIG. 1. Metabolic pathway illustrating role of IMPDH. Enzymes shown in the pathway are indicated as follows: 1, purine nucleoside phosphorylase; 2, hypoxanthine-guanine phosphoribosyl transferase; 3, GMP synthase; 4, adenylosuccinate synthase and adenylosuccinate lyase, respectively (two-step reaction); 5, AMP deaminase; 6, adenine phosphoribosyl transferase; 7, adenine deaminase; 8, adenosine kinase; 9, adenosine deaminase. IMPDH is also shown.

turer's protocol for the Random Primed Labeling Kit (United States Biochemical Corp., Cleveland, Ohio). DNA sequencing was performed with the Sequenase, version 2.0, kit (U.S. Biochemical Corp.). Two oligonucleotides were synthesized for use as internal sequencing primers. These primers were referred to as RACE 53 (5'-CCGTTTGCTCCTCTATCG-3') and RACE 33 (5'-GCAT TCATTACAGGATATCGG-3').

Isolation of total RNA from *P. carinii* and Northern blot analysis. *P. carinii* organisms were cultured for 7 to 8 days in spinner flasks on monolayers of HEL cells (25); organisms in the supernatant were harvested by centrifugation and washed twice in phosphate-buffered saline. After suspension in medium to protect RNA from degradation (10), RNA was isolated from the pelleted organisms and total RNA was electrophoresed on a formaldehyde–1.2% agarose gel (31). The RNA was blotted onto a Nytran Maximum Strength Plus membrane by Northern transfer (1). The blot was hybridized with the same probe used for Southern analysis.

5' and 3' RACE. Kits for 5' and 3' rapid amplification of cDNA ends (RACE) were used according to the manufacturer's instructions (Gibco BRL, Gaithersburg, Md.). Oligonucleotides for 5' and 3' RACE included RACE-32 (5'-GCA GGAACTACTGAATCTCC-3'; nucleotides 1193 to 1212 for 3' RACE) and RACE-51 (5'-CTGGTAATTTAGATGCTAAAGG-3'; nucleotides 719 to 740) for cDNA synthesis and RACE-52 (5'-AAGTGCTGTTAAGTTTCCTTCT-3'; nucleotides 661 to 682), a nested primer for PCR in the final step of 5' RACE.

Reverse transcriptase PCR (RT-PCR). Total RNA (0.5 to 1.0 μ g) from *P. carinii* was aliquoted into 0.5-ml RNase-free tubes, and the volume was brought up to 15.5 μ l with diethyl pyrocarbonate-treated H₂O. The YRGMGS primer described above was used in a 25- μ l reaction mixture to reverse transcribe the *P. carinii* IMPDH mRNA (24). The reverse-transcribed product was amplified by PCR with primers YRGMGS and IMPDH-RP2 (5'-TCGTTCATGTAAGAAG GG-3') under the same PCR conditions described above.

Overlap extension of RACE and RT-PCR products. RT-PCR and RACE products purified on GlassMAX spin cartridges (Gibco BRL) served as templates for overlap extension (Gibco BRL Technical Services). Initially, the RT-PCR product and the 3' RACE fragment were extended. The extended product was purified, and overlap extension was performed again with the 5' RACE product. The full-length *P. carinii* IMPDH cDNA thus generated was cloned into pCRII for DNA sequencing. The newly constructed plasmid was designated pPCI-19.

Construction of the IMPDH-T7 expression vector. Two oligonucleotides, PCI-F (5'-GCGTGATATCATGAGTTCACCGATGGA-3') and PCI-R (5'-CG GGATCCCGTTAACTCCAAAGCTTTTTTTG-3'), were designed with *Bsp*HI and *Bam*HI restriction endonuclease sites at their 5' ends, respectively, in order to amplify by PCR IMPDH cDNA from pPCI-19 (pCRII-IMPDH recombinant clone). The PCR product was digested with the enzymes described above and was cloned into the *NcoI* and *Bam*HI sites of the T7 expression vector pET-8c, generating pET/19-3C and pET/19-3H. Recombinant clones were identified by digesting plasmid DNA with *Bg*II and *Bam*HI, since the *NcoI* and *Bsp*HI sites were lost when they were ligated.

Expression of \vec{P} . *carinii* **IMPDH in** *Escherichia coli*. A single colony of BL21(DE3) transformed with pET/19-3H was inoculated into 3 ml of LB medium with 100 µg of ampicillin per ml, incubated overnight at 37°C, diluted 1:40 in LB medium containing 100 µg of ampicillin per ml, and incubated for 4 h at 37°C. Expression of T7 RNA polymerase was induced by adding isopropyl- β -D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. The culture was incubated at 37°C for 30 min. *E. coli* RNA polymerase was inhibited by adding 200 µg of rifampin per ml. The culture was incubated at 37°C for an additional

2.5 h and was harvested at 10,000 rpm for 15 min. The pellet was resuspended in 1.5 ml of buffer A (10 mM sodium phosphate buffer [pH 7.0] containing 9 μ g of phenylmethylsulfonyl fluoride per ml, 20 μ g of leupeptin per ml, 50 μ g of aprotinin per ml). Cells were disrupted by sonication. The supernatant from centrifugation at 100,000 × g for 15 min at 4°C was assayed immediately for IMPDH activity.

IMPDH enzyme assay. *E. coli* lysates were assayed spectrophotometrically for IMPDH catalytic activity (27). A 1-ml enzyme reaction mixture consisting of 100 mM Tris-HCl (pH 8.0), 100 mM KCl, 3 mM EDTA, 0.1 mM allopurinol, and 275 to 575 μ M NAD was incubated at 37°C for 5 min with monitoring of the A_{340} to detect the formation of NADH. The concentration of IMP in the reaction mixture was 180 to 360 μ M; the control, however, had no IMP. Reactions were done in duplicate and were initiated by the addition of enzyme.

Isolation of IMPDH from inclusion bodies. Recombinant protein was isolated from inclusion bodies by a modification of the procedure of Landman et al. (23). The crude cell extract was centrifuged at $10,000 \times g$ for 15 min at 4°C. The pellet was resuspended in 1 ml of 50 mM Tris-HCl (pH 8.0)–1 mM dithiothreitol (DTT)–0.01 mM EDTA–10% glycerol; after sonication the preparation was centrifuged at 10,000 × g for 15 min. The pellet was stored overnight at -70° C and was resuspended in 5 ml of 50 mM Tris-HCl (pH 8.0)–0.1 M DTT–5 M urea–1 MM EDTA, and the mixture was incubated at room temperature for 2 to 3 hrs. Urea was removed with a 50-ml column of Sephadex G-25 (fine) equilibrated with 50 mM Tris-HCl (pH 8.0)–1 mM DTT–6. Fractions were assayed for protein and enzymatic activities. Samples of the fractions were loaded onto a sodium dodecyl sulfate (SDS)–12% polyacrylamide gel (22) for visualization of proteins by Coomassie blue staining.

Analysis of newly synthesized proteins. Incorporation of ³⁵S-Met was used to identify newly synthesized proteins (38). Three-milliliter cultures were grown to an optical density at 600 nm of 0.4 to 0.5, and 1 ml was removed and centrifuged for 30 s. The cell pellet was washed once with 1 ml of M9 medium and was resuspended in 1 ml of M9 medium with 0.02% (each) an 18-amino-acid mixture (minus Met and Cys). The resuspended cells were incubated for 1 h at 37°C and were then split into two 0.5-ml samples. One of the 0.5-ml samples was induced with 1 mM IPTG (final concentration). Both samples were incubated at 37°C for 30 min, at which time rifampin (200 μ g/ml) was added to both tubes. The cultures were incubated for 30 min at 37°C. Translation-grade ³⁵S-Met (10 μ Ci; NEN, Boston, Mass.) was added to the samples, and the mixtures were incubated for 5 min at 37°C. The reaction was stopped by centrifuging the culture for 30 s. The medium was discarded, and the cells were resuspended in 100 µl of 60 mM Tris-HCl (pH 6.8) containing 3% SDS, 1% β-mercaptoethanol, 10% glycerol, and 0.01% bromophenol blue. The samples were boiled for 5 min, and 20 µl of each sample was resolved by SDS-polyacrylamide gel electrophoresis (PAGE) on a 12% polyacrylamide gel. ¹⁴C-methylated protein markers (Amersham, Arlington Heights, Ill.) were used as molecular weight standards. The gel was fixed for 1 h in 30% methanol-10% acetic acid, soaked in En³Hance (NEN) for 1 h, and washed with cold tap water for 30 min. The gel was dried overnight before being placed in an X-ray cassette for fluorography with XAR-5 film and an intensifying screen. The film was developed after a 24 h of exposure by using an X-OMAT rapid process developer.

Experiments with IMPDH-deficient E. coli. E. coli Sol101 (supplied by Sabrina van Ginkel, Southern Research Institute, with permission of Per Nygaard, Department of Biological Chemistry, Institute of Molecular Biology, University of Copenhagen) requires methionine, thiamine, and guanine for growth. The strain lacks IMPDH, which confers the growth requirement for guanine. E. coli Sol1101 was sequentially transformed with plasmid pET-8c containing an ampicillin resistance marker and the putative P. carinii gene for IMPDH driven by a T7 promoter and with plasmid pGP1-2 containing a gene for T7 RNA polymerase and a kanamycin resistance marker. The T7 RNA polymerase gene was under the control of the temperature-sensitive repressor λ cI857, which was expressed at 30 to 37°C but inactivated at 42°C; therefore, a temperature shift to 42°C could be used to turn on the synthesis of the T7 RNA polymerase and, in turn, transcription of the P. carinii IMPDH gene. In mid-logarithmic phase, T7 RNA polymerase expression was induced by a temperature shift to 42°C for 30 min, and the cultures were allowed to grow at 37°C for an additional 4 to 6.5 h. The cultures were harvested by centrifugation, and the cells were washed, resuspended in buffer A, and sonicated as described above.

Nucleotide sequence accession number. The nucleotide sequence of the entire *P. carinii* IMPDH cDNA was assigned GenBank accession number U42442.

RESULTS

Susceptibility of *P. carinii* **to known IMPDH inhibitors.** Growth of *P. carinii* in short-term culture is inhibited by agents such as trimethoprim-sulfamethoxazole, primaquine-clindamycin, and pentamidine, which are also known to be effective in animals and humans (40). The culture system has also been used to predict the activities of experimental compounds. For example, albendazole was active against *P. carinii* in culture (4) and in animals (5), whereas the imidazole antifungal agents



FIG. 2. Effect of IMPDH inhibitor mycophenolic acid (MPA) on the growth of *P. carinii* in culture. Cultures were established by published protocols by using an inoculum from *P. carinii*-infected rat lung to infect monolayers of HEL cells (8). Control cultures received only drug diluent, whereas experimental cultures received mycophenolic acid to achieve final concentrations of 0.01, 0.1, or 1 μ M.

were not active in culture or animal models (7). A known inhibitor of IMPDH, mycophenolic acid, was tested against *P. carinii* in this culture system and was a highly effective inhibitor of the growth of *P. carinii* (Fig. 2); the monolayer host cells showed no sign of toxicity during this short-term culture. The sensitivity of *P. carinii* to low concentrations of mycophenolic acid implied that the purine metabolic pathway was present in *P. carinii* and that inhibition of that pathway might be detrimental to the organism. Tests were attempted in a rat model of *P. carinii* pneumonia (6), but the toxicity of the drug precluded adequate testing in animals (data not shown), suggesting that more specific inhibitors of *P. carinii* IMPDH would be needed for successful therapy. Therefore, the cloning of IMPDH from *P. carinii* was undertaken as a first step in this process.

Cloning of the cDNA encoding P. carinii IMPDH. The amino acid sequences of IMPDH enzymes from E. coli (39), Bacillus subtilis (20), Leishmania donovani (44), mouse (40), Chinese hamster (11), and human (11, 28) cells, which have two isozymes, were compared and aligned for homology by using the CLUSTAL V program in PC/Gene, version 6.8 (Fig. 3). The alignment revealed three conserved regions referred to as MDTVTE, GSICIT, and YRGMGS, respectively, on the basis their single-letter amino acid abbreviations. MDTVTE fell in the N-terminal portion of the protein, whereas YRGMGS was near the carboxy terminus, with GSICIT near the middle of the protein. The three hexapeptide sequences were converted from the amino acids into nucleotide sequences, taking into account the codon bias observed in other P. carinii genes. Oligonucleotide PCR primers, named for the amino acid sequences, were 5'-ATG GA(T/C) ACN GTN ACN GA-3' (MDTVTE), 5'-GA(A/T) CCC AT(A/T) CC(A/T) CG(A/G) TA-3' (YRGMGS), and 5'-GGN TCN AT(A/T) TG(T/C) AT(A/T) AC-3' (GSICIT). These degenerate primers were used in various combinations to amplify a portion of the IMPDH gene from P. carinii. A 1,369-bp fragment was amplified in the presence of 2.5 to 3.0 mM MgCl₂ with primers MDTVTE and YRGMGS, but no product was obtained in the presence of lower magnesium concentrations. An amplified DNA product was not obtained with the GSICIT primer at any concentration of magnesium.

Use of the 1.3-kb product amplified from *P. carinii* genomic DNA to screen a cDNA library from *P. carinii* produced clones that shared sequences with the 1.3-kb product, but no clones contained the entire gene. Since the 1.3-kb product represented the middle of the IMPDH gene, RT-PCR and 5' RACE and 3' RACE reactions seemed ideal techniques for obtaining the missing 5' and 3' ends of the gene.

Oligonucleotides directed at internal sequences of the 1.3-kb genomic PCR product were designed and used in RT-PCR and 5' RACE and 3' RACE reactions as described in Materials and Methods (Fig. 4). The primer for 3'-RACE (5'-GCA GGA ACT ACT GAA TCT CC-3') was near the 3' end of the 1.3-kb fragment. The product obtained from this procedure contained a poly(A) tail approximately 60 bp long; the entire product was 546 bp. RT-PCR was performed with the primer YRGMGS near the 3' end of the 1.3-kb product and another oligonucleotide (IMPDH-RP2; 5'-TCG TTC ATG TAA GAA GGG-3') near the middle of the 1.3-kb fragment. These primers yielded a product of 648 bp. The 5' RACE procedure used two primers near the middle of the 1.3-kb fragment. The first primer (5'-CTG GTA ATT TAG ATG CTA AAG G-3') was used to generate cDNA with a homopolymeric dCMP tail to serve as an anchor for the second round of PCR with a second primer (5'-AAG TGC TGT TAA GTT TCC TTC T-3') just to the 5' side of the first primer. This procedure gave a product of 682 bp.

In order to obtain the full-length cDNA encoding P. carinii IMPDH, it was necessary to join the three smaller fragments into one large DNA fragment. It was possible to obtain the full-length cDNA for the P. carinii IMPDH gene because the primers for RACE were designed so that the 3' end of the 5' RACE product overlapped the 5' end of the RT-PCR product and the 3' end of the RT-PCR product overlapped the 5' end of the 3' RACE product (Fig. 4). These overlapping regions were complementary to each other, and by annealing to each other, they acted as primers for AmpliTaq DNA polymerase to extend in the first PCR cycle. The initial overlap extension reaction joined the cDNA from the RT-PCR (682 bp) and the 3' RACE reaction (546 bp) to produce a 1.1-kb fragment. This fragment included the portion of the cDNA for P. carinii IM-PDH beginning with oligonucleotide IMPDH-RP2, a primer used to sequence the 1.3-kb genomic PCR product, at the 5' end through the 3' poly(A) tail. This fragment was used in the second overlap extension reaction with the 5' RACE product to generate the full-length cDNA, which was a 1.7-kb fragment (Fig. 5). The entire P. carinii IMPDH cDNA was cloned into pCRII for sequencing.

Characterization of the *P. carinii* **IMPDH gene.** The sequence of the 1.3-kb genomic fragment contained five putative introns of 144, 48, 52, 48, and 45 nucleotides on the basis of the presence of the characteristic 5'-GT and 3'-AG splicing signal. These putative introns were very A/T-rich, like introns identified in other published *P. carinii* genes (13–15). Primer GSICIT, which failed to amplify DNA, was found to span an intron-splicing site, as determined from the sequence.

When the putative introns were removed from the sequence, an open reading frame of 1,032 bp was generated. The DNA sequence of the 648-bp RT-PCR product was identical to that of the same region of the 1.3-kb genomic PCR product without the putative intron sequences. This result confirmed that the regions identified as introns in the 1.3-kb genomic PCR product were authentic introns on the basis of the absence of these sequences in the RT-PCR product. The DNA sequence of the 546-bp 3' RACE product revealed a polyadenylation signal,

205 277 273

305

377 373 373

355

427 422

416

454

523

514

514

PC	M	1	PC	SDLMKNLHFPLASKLPDSKOLICAOAVGTRPDDRTRLKHLVEAGLDTVVL
SC	*AAIRDYKTALDFTKSLPRPDGLSVOELMDS	31	SC	T*LMKNONY*LASKSANTKO*LCGASTG*MDa*KE*LEL*V/KA*L*V///T*
CH	*ADVI.TSCCTSVVPDDCLTACOLENC	26	CH	T*LKKNEDV*LASKDAK_KO*LCGAATG*HED*KV*LDI*AIA*V*MAG*
MII	*aDVLISCCTSVVP	26	MIT	T MAGANDI MASADAK-NO DOGARIG HED KI "DDD"ALA" V VVV
10	* DVI ISCOUCYUB.	26	10	T DIGUNDI DASKDAK-KQ"DCGAAIG"HED"KI"DDL"ALA"V"VVV"
1101		20	HOZ HC	T "LIXANRDI" LASKDAR" KQ "LCGAAIG" HED "KI "LDL" AQA" V "VVV"
HSI	*ADYLISGGIGIVPBUITAQULFAS	20	HS	T^LKKNRDY*LASKDSQ-KQ*LCGAAVG*RED*KY*LDL*TQA*V*VIV*
1.64	*ESTTKVKVNGFVESTSSSAAPAIQTKSTIGFDAELQDGLSCKELFQN	48	DM	T*LKKARSY*NASKDSN-KQ*LVGAAIG*RSE*KA*LAL*VAN*V*VII*
LD	*ATNNRDGCTAEELFRG	25	LD	R*AVRARDY*HSTLD-KSGR*ICAAATS*RPE*KR*VAA*ADV*V*VLV*
TB	*ENTNLRTKTLRDGTTAEELFSQ	23	TE	R*AVRARDY*NSSLD-RNGH*LCAAATS*REA*KG*VAA*SEA*I*VLV*
PC	\$\$PMብጣም	q	50	
50		01	PC	DSSQGNSIYQINMIKWIKKEFPNLEVIAGNVVTREQAANLISAGADALRV
50		01	SC	******SIFELNMLKWV*ESF*G*E*IA****RE**A***A**A**A*GL*I
ĻН	G-DGLTYNDFLILFGYIDFTADQVDLTSALTKKITLKTPLV******	74	CH	******SIFQINMIKYM*EKY*N*Q*IG*****AA**K***D**V*AL*V
MŲ	GDGLTYNDFLILPGY1DFTADQVDLTSALTKK1TLKTPLV*******	74	MU	******SIFQINMIKYI*EKY*S*Q*IG*****AA**K***D**V*AL*V
HS2	GDGLTYNDFLILPGYIDFTADQVDLTSALTKKITLKTPLV*******	74	HS2	*****SIFQINMIKYI*DKY*N*Q*IG*****AA**K***D**V*AL*V
H\$1	ADDLTYNDFLILPGFIDFIADEVDLTSALTRKITLKTPLV*******	74	HS1	******SVYQIAMVHYI*QKY*N*Q*IG*****AA**K***D**V*AL*V
DM	GEGLTYNDFLILPGYIDFTAEEVDLSSPLTKSLTLRAPLV*******	96	DM	******SVYQVEMIKYI*ETY*E*O*IG*****RA**K***D**V*GL*V
LD	DGLTYNDFIILPGFIDFGAADVNISGQFTKRIRLHIPIV*****I*	72	LD	******TIYQIAFIKWV*STY*H*E*VA*****0D**K***D**A*GI*I
TB	DGLSFNDFIILPGFIDFDS5KVNVSGQFTKNILLHLPLV*******	70	TB	******TIYQVSFIRWV*KTY*H*E*VA*****QD**K***D**A*SL*I
	^			
PĈ	ESDMAINLALLGGIGVIHHNCTIEEQTEMVRKVKKFENGFITSPIVLSLN	59	\mathbf{PC}	GMG8GSICITOEIMAVGRPOATAVYAVSEFASKEGVPTTADGGTENTCHT
SC	*SE**TFM**L**I*FI*H***P*D*ADM*RR**NYEN**INNPIVISPT	131	SC	**************************************
CH	*AG**TAM**T*T*FT*H***P*F*ANE*RK**KYEO**TTDPWULSPK	124	СH	***C**********************************
MIT	*AG**TAM**T*T*FT*H***P*F*ANE*EK**KVEO**TTD*\7.SPK	124	MIT	THE A VER DELARATION AND A VER DELARATIVA AND A VER DELARATIVA AND A VER DELARATIVA AND A VER DELARATIVA AND A VER
HC2	*AC**TAM**T*T*ET*U***P*E*ANE*RK*KVEO**TTDBV/J.SPK	124	10	***C**********************************
1102	*AD**TAM**M**T*FT*U***D*F*ANE*D**TEPO**TEPOPULI CCC	124	132	The second visit of the second s
TGN 1	*AD**IAM**M**I*FI*A**F*F*ANE*KK**NFEQ**IIDFVVLSFS	144	HSI	C***C*****VM*C*R**G**V*K*AEYARRFG**II****IQTV*HV
LAM	"SE" "IAM" "C" "I"II"H" " "P"I "ALE"HK" "KIKH" "MRDPSVMSPT	140	DM	***C*******VM*C*C**A**V*Q*STYARQFG**VI****IQSI*HI
Ш	*NE**KIM**M**V*VL*H***V*R*VEM*KS**AYRN**ISK*KSVPPN	122	LD	***S********VL*C*R**G**V*K*AQYCASRR**CT****LRQV*DI
TB	*SS**RAM**M**I*VI*N***V*Q*ARM*RS**LYRN**IMK*KSVSPD	120	TB	***S********VL*C*R**A**I*K*ARYAASRG**CV****LRNV*DV
PC	HRVRDVRRIKEELGFSGIPITDTGQLNGKLLGIVTSRDIQFHNNDESF	107	PC	TKALALGASAVMMGNLLAGTTESPGQYYYRDGQRLKS <u>YRGMGS</u> IDAMEHL
SC	TTVGEAKSMKEKY*FA*FPV*TDGKRNAK*V*VITSR**QFVEDNSLL	179	SC	T**L*L*SSTV*M*GMLAG*T*S**E*FYOD*K***A****ID**OKT
ĊН	DRVRDVFEAKARH*FC*IPI*DTGRMGSR*V*IISSR**DFLKEEEHDRF	174	CH	A**L*L*ASTV*M*SLLAA*T*A**E*FFSD*I***K******JD**DK-
MU	DRVRDVFEAKARH*FC*IPI*DTGRMGSR*V*IISSR**DFLKEEEHDRF	174	MU	A**L*L*ASTV*M*SLLAA*T*A**E*FESD*T***K*****LD**DK-
HS2	DRVRDVFEAKARH*FC*IPI*DTGRMGSR*V*IISSR**DFLKEEEHDCF	174	HS2	A**I.*I.*ASTV*M*SLLAA*T*A**F*FFSD*T**************
HŞ1	HTVRDVFEAKMRH*FS*IPI*ETGTMGSK*V*IVTSR**DFLAEKDHTTL	174	HS1	V**L*L*ACTV*M*CLLAA*T*A**F*FFCD*V**********
DM	NTVGDVLEARRKN*FT*YPV*ENGKLGGK*L*NVTSR**DF-RENOPEVL	195	DM	VISIANTALANCAUTACETACECETATESCONTANTALANTALANT
LD	TPISNITRIKEEK*IS*ILV*ENROPHGK*L*IVCTK**DYVKNKOTP	170	ID	CALL ADAV MODELAG STATE FFSD VAN KANANALEN ER-
TB	VPVSTIPNIKSEK*IS*ILV*ECGKVDGK*L*IVCTK**DEVKDASAP	168	100	CARL ANCA HAGHLSGATATA BARANA ANA ANA ANA ANA ANA ANA ANA ANA AN
12		100	IB	CALLANDALLASMIAGASATA BAFFRDAMAAAGAAAAAA AA
PC	LSEVMTKDLVTGSEGTR*EEANETLRSCKKGKLPTVDKEGNLTALLSP	155	DC	COVNECTION CONTRACTOR CONTRACTOR
80	VODU**KNDUTGAGGTT*SEGNET*KKTKK*P*LIADEKGNU/SMLS*	227	PC.	SGRWRGDNAASSRIF-GEADTIRVAQGVSGSVIDKGSIKKFIPYLRTGLQ
CH	LEET ** ADDDI MAN DACTO * ADDALE ANALA *********************************	221	50	GTKGNASTS**F-S*SDSVLV****S*AVV****IKKFIP*LYN*L*
	LEEL + * UDEDI UUD DI CUETURENDE VERMENDE VILLENDE VILLE VILLENDE VILLENDE VILLENDE VILLENDE	224	CH	HLSSQN**F-S*ADKIKV****S*AVQ****IHKFVP*LIA*I*
MO	LEEL ** KREDLIV VAPAGVI * KBANEI * OKSKK * K * PIVNENDELVAIIA*	224	MU	HLSSQN**F-S*ADKIKV****S*AVQ****IHKFVP*LIA*I*
HSZ	LEET**RREDLVVAPAGIT*REANET*QRSRR*R*PIVNENDELVATIA*	444	HS2	HLSSQN**F-S*ADKIKV****S*AVQ****IHKFVP*LIA*I*
HSI	LSEV**PRIELVVAPAGVT*REANEI*QRSKK*K*PIVNDCDELVAIIA*	224	HS1	SSSSQK**F-S*GDKVKI****S*SIQ****IQKFVP*LIA*I*
DM	LADI**TELVTAPNGIN*PTANAI*EKSKK*K*PIVNQAGELVAMIA*	243	DM	GDAKGAAMS**YHN*MDKMKV****S*SIV****VLRYLP*LEC*L*
LD	VSAV**RREKMTVERAPIQ*EEAMDV*NRSRY*Y*PIVNENDEVVNLCS*	220	LD	GKESGK**L-S*NEAVOV****S*NVV****AAKLIA*VSK*L*
TB	VSQY**RRENMTVERYPIK*EEAMDV*NRSRH*Y*PVLNDKDEVVCLCS*	218	TB	GRESGK**L-S*NETLQV****A*AVL****VLKLLA*IHK*L*
FIG.	3. Amino acid homology among eukaryotic IMPDH proteins.	The	PC	HSLQDIGVONLTELRKOVKEKNIRFEFRTVASOLEGNVHGLDSYOKKLWS
CLUST	AI V program was used to align and show homology among nine	e11-	SC	HSC****CRSLTLLKNNVORGKVR*EF*TASA*L**G**N*HS*EKRLHN
CL0317	IMPOLI and the line and the insert a	i	CH	HSC****AKSLTOVRAMMYSGELK*EK*TSSA*V**G**S*HS*EKRLF-
кагуонс	inition proteins. Hypnens are used to insert gaps to optimize al	ngn-	MU	HSC****AKSLTOVRAMTYSGELK*EK*TSSA*V**G**S*HS*EKRLF-

karyotic IMPDH proteins. Hyphens are used to insert gaps to optimize alignment of amino acids. Amino acids identical in all proteins are denoted with an asterisk, and conserved residues are marked with a caret. The IMPDH signature site is in boldface letters. Underlined amino acids represent amino acids used to design degenerate PCR primers. PC, P. carinii; SC, S. cerevisiae; CH, Chinese hamster; MU, mouse; HS1, human type I; HS2, human type II; DM, Drosophila melanogaster; LD, L. donovani; TB, T. brucei.

AATAAA (16), within the 3' untranslated region, as well as the poly(A) tail. The translated protein had a 63% amino acid homology with mammalian IMPDHs, whereas the homology among mammalian IMPDHs was 83 to 99% (Fig. 3).

PFGE and Southern analysis of PFGE-resolved P. carinii chromosomes were performed to verify that the amplified DNA product did not originate from the human feeder cells that contaminate preparations of DNA isolated from cultured P. carinii (Fig. 6A and B). The 1.3-kb fragment, labeled with $[\alpha^{-32}P]dCTP$, hybridized specifically to a single chromosome from rat-derived (430-kb) and mouse-derived (410-kb) P. carinii (Fig. 6B). No hybridization signal was observed with yeast or bacteriophage lambda markers or the P. carinii isolated from infected ferret lung. The absence of a hybridization signal from the 1.3-kb probe with ferret P. carinii chromosomes may be due to differences among P. carinii organisms infecting different hosts (34, 37, 42, 43). The nucleotide sequence of the amplified product was not identical to that of the human cDNA, and no hybridization band was observed in the upper

HS2	HSC****AKSLTOVRAMMYSGELK*EK*TSSA*V**G**S*HS*EKRLF-	514
HS1	HSC****ARSLSVLRSMMYSGELK*EK*TMSA*I**G**S*HS*EKRLY-	514
DM	HSC****ANSINKLRDMIYNGOLR*MK*THSA*L**N**G*FS*EKLRF-	537
LD	OSA****EISFDAIREKMYAGOVL*SR*SPTA*G**G**S*HS*EKKLFA	511
TB	OSA****EVSFDAIREKVYEGOVL*NR*TLTA*S**A**S*HH*ERKLFA	509
PC		
SC		
CH		
MU		
HS2		
HS1		
DM		
LD	AKM	514
TB	SKL	512

one-fourth of the gel, which is where human chromosomes run in this system. These results provide additional evidence that the 1.3-kb fragment did not originate from the feeder cells.

The same probe used for Southern hybridization studies was used to probe a Northern blot of total RNA from P. carinii in order to determine the size of the mRNA encoding IMPDH. A single band of 1.6 kb was observed on the autoradiogram (Fig. 7). This result was consistent with the size of the full-length transcripts for human (1,654 nucleotides [28]) and Chinese hamster (1,614 nucleotides [11]) IMPDHs. The size of transcript led us to believe that we had isolated the full-length cDNA.



FIG. 4. Strategy for obtaining the full-length cDNA for *P. carinii* IMPDH. The original 1.3-kb genomic PCR product was used to design primers for 3' RACE, 5' RACE, and RT-PCR, as shown. Assembly of the three products yielded a cDNA copy of the entire gene. Sequences of all primers are given in Materials and Methods.

Expression of the *P. carinii* **IMPDH gene.** The cDNA encoding *P. carinii* **IMPDH** was cloned into the T7 expression vector pET-8c as described in Materials and Methods for highlevel expression in *E. coli*. The T7 promoter, in the presence of T7 RNA polymerase, is so strong that nearly all protein synthesis in the cell is directed at making recombinant protein.

E. coli BL21(DE3) cells that have a chromosomal copy of T7 RNA polymerase were transformed with pET/19-3H contain-





FIG. 6. PFGE resolution of *P. carinii* chromosomes and Southern hybridization analysis. (A) Chromosomes resolved in 1% agarose. (B) Southern hybridization of chromosomes with IMPDH probe. Lanes 1, DNA size markers from *S. cerevisiae*; lanes 2, DNA size markers of bacteriophage lambda concatamers; lanes 3, rat *P. carinii* from spinner flask culture; lanes 4 and 5, *P. carinii* isolated from lungs of two different rats; lanes 6, *P. carinii*-infected mouse lung; lanes 7, *P. carinii*-infected ferret lung; lanes 8, *P. carinii*-infected mouse lung.

ing the putative gene for *P. carinii* IMPDH. T7 RNA polymerase was expressed by inducing the culture with IPTG. *E. coli* cells transformed with the recombinant IMPDH gene and induced with IPTG were shown by electron microscopy to contain inclusion bodies that occupied 25 to 33% of the cytoplasm,





FIG. 5. Overlap extension reactions for full-length *P. carinii* IMPDH cDNA. Lanes 1 and 5, $\lambda BStE$ II and Boehringer Mannheim DNA marker VI, respectively; lane 2, a control to which no template was added; lanes 3 and 4, different amounts of the reaction products with RACE and RT-PCR. In lane 3, the overlapping templates are visible, as is the full-length IMPDH cDNA (1.7 kb).

FIG. 7. *P. carinii* RNA and Northern hybridization analysis. Lanes 1 and 3, RNA ladders of 0.16 to 1.77 and 0.24 to 9.5 kb, respectively (BRL, Gaithersburg, Md.); lanes 2, total RNA prepared from cultured *P. carinii* and isolated by cesium chloride density gradient centrifugation. The 16S and 26S rRNAs (lower two bands; lane 2) were visualized by ethidium bromide staining of formalde-hyde–1.2% agarose gel after electrophoresis (A). (B) Hybridization of IMPDH probe with 1.6-kb mRNA on Northern blot.



FIG. 8. New protein synthesis determined by SDS-PAGE resolution of [35 S]methionine-labeled proteins and detected by fluorography. Lane M, molecular weight standards (radiolabeled); lanes 1 to 4, extracts from uninduced cultures; lanes 5 to 8, extracts from IPTG-induced cultures; lanes 1 and 5, untransformed bacteria; lanes 2 and 6, extracts from *E. coli* transformed which lacked the coding sequence for IMPDH; lanes 3, 4, 7, and 8, two different transformants, both of which received pGP1-2 and plasmid pET19-3 containing the IMPDH coding sequence.

but inclusion bodies were not seen in untransformed BL21 (DE3) cells (data not shown).

When induced cultures were grown in the presence of [³⁵S]methionine, a 50-kDa protein was the only newly synthesized protein detected (Fig. 8).

Characterization of the recombinant *P. carinii* **IMPDH protein.** Inclusion bodies from recombinant organisms were solubilized with urea, and the solubilized protein was subjected to Sephadex G-25 column chromatography and SDS-PAGE. The molecular size of the predominant protein was approximately 50 kDa as predicted, but IMPDH catalytic activity was not detected by using several different conditions of assay.

The 50,000-molecular-weight protein, isolated by urea solubilization and G-25 column chromatography, was subjected to amino acid sequencing. The sample was digested with trypsin because the protein appeared to be N-terminally blocked. Trypsin cleavage products were separated by reverse-phase high-pressure liquid chromatography, and three peptides were selected at random for sequencing. The sequences of the three fragments were DIQFHNNDESFLSEV, EFPNLEVIAGN VVT, and KFENGFITSPIVLS. The sequences of all 43 amino acids sequenced exactly matched the deduced amino acid sequence of the cDNA for P. carinii IMPDH. These three peptides showed only 25% homology with those regions of E. coli IMPDH (Fig. 9) and overall showed 32% homology with the complete sequence of the E. coli IMPDH protein (Table 1). Therefore, it was unlikely that the three peptides originated from E. coli IMPDH.

E. coli S ϕ 1101 lacks IMPDH activity and requires methionine, thiamine, and guanine for growth (28a). Experiments were performed to determine if IMPDH synthesized from the *P. carinii* IMPDH gene could function catalytically in *E. coli* S ϕ 1101 and thus confer the ability to grow in the absence of guanine. Methionine and thiamine were supplied to all transformants and controls. Untransformed *E. coli* S ϕ 1101 cells failed to grow in the absence of guanine, as did transformants containing only pET-8c, pGP1-2, or pET/19-3C (data not shown). Doubly transformed cells (S ϕ 1101 with pET/19-3C and pGP1-2, which supplies the T7 RNA polymerase gene) gained the ability to grow in the presence of both ampicillin

\mathbf{PC}	44	KFENGFITSPIVLSL	58
EĊ	88	KHESGVVTDPQTVLP	102
PC	97	DIQFHNNDESFLSEV	111
EC	138	 DVRFVTDL	145
PÇ	225	EFPNLEVIAGNVVT	238
EC	267	KYPDLOI IGGNVAT	280

FIG. 9. Alignment of amino acid sequences of three peptides from *P. carinii* IMPDH with the sequence of corresponding peptides from *E. coli*. *P. carinii* IMPDH expressed in *E. coli* was partially purified and exposed to trypsin digestion, and the resulting peptides were purified by high-pressure liquid chromatography. Three purified peptides available in adequate yields were chosen for amino acid sequencing, as described in Materials and Methods.

and kanamycin, as well as the ability to grow in the absence of guanine. These results implied that catalytically active *P. carinii* IMPDH is produced in the doubly transformed *E. coli*.

E. coli S ϕ 1101 doubly transformed with pET/19-3C and pGP1-2 was grown in the presence of ampicillin, kanamycin, methionine, and thiamine. Control cultures of untransformed *E. coli* S ϕ 1101 cells were grown in the presence of methionine, thiamine, and guanine. The supernatants obtained by centrifugation of sonicated extracts of each culture at 100,000 × g were assayed for IMPDH activity. IMPDH activity was found only in the doubly transformed *E. coli* cells containing the *P. carinii* IMPDH gene, again implying that catalytically active *P. carinii* IMPDH is synthesized in this organism. IMPDH activity was defined as a reduction of NAD that was dependent on the presence of IMP; background rates of NAD reduction were low under the standard conditions of the assay.

The catalytic activity was characterized in these supernatants obtained by centrifugation at 100,000 \times g. The reaction was linear with enzyme concentration over 7 min under the standard conditions of the assay. The apparent K_m for IMP was 21.7 \pm 0.3 μ M in three separate preparations, whereas the K_m for NAD was 314 \pm 84 μ M (mean \pm standard error of the mean; n = 3). The activity was inhibited by mycophenolic acid, with 50% inhibitory concentrations of 18 and 29 μ M in two experiments.

DISCUSSION

By taking advantage of strong homology in selected regions of IMPDH, we were able to amplify a fragment of the IMPDH gene from genomic *P. carinii* DNA. The full gene for this putative *P. carinii* IMPDH was obtained by combining the techniques of 5' RACE, 3' RACE, and RT-PCR with overlap extension. The resulting product with a 3' poly(A) tail and a 5' start codon was 1.7 kb.

Several lines of evidence confirmed that the gene isolated by this procedure was from *P. carinii*. The nucleotide sequence for the entire gene showed strong homology to the sequences of IMPDH genes from several sources, but it was not identical to any of them. Homology at the nucleotide level with the two human forms of IMPDH was 50 to 52%; it is important to examine these forms of IMPDH because the *P. carinii* isolate used to produce genomic DNA for the original isolation was grown on human embryonic lung cells. Additional evidence supporting the identification of the material as a gene from *P. carinii* came from Southern and Northern blot analyses, both

D. ()	% Homology												
Protein source	AC	BS	СН	DM	EC	HS1	HS2	LD	MU	PC	SC	TB	TF
AC		53.9	36.3	37.1	63.1	38.3	36.3	33.0	28.9	31.3	34.0	32.6	28.9
BS	53.9		38.2	34.9	52.4	40.5	38.2	30.0	38.4	38.3	34.7	31.8	31.0
CH	36.3	38.2		67.3	35.2	83.5	98.2	52.5	99.0	62.3	61.3	52.9	23.5
DM	37.1	34.9	67.3		35.2	67.9	67.5	49.8	28.2	58.6	57.0	51.6	28.2
EC	63.1	52.4	35.2	35.2		33.9	35.2	29.6	35.2	31.7	37.1	31.4	31.8
HS1	38.3	40.5	83.5	67.9	33.9		83.7	52.3	83.5	59.9	59.9	52.9	27.8
HS2	36.3	38.2	98.2	67.5	35.2	83.7		52.9	98.6	62.3	61.1	53.1	23.9
LD	33.0	30.0	52.5	49.8	29.6	52.3	52.9		52.3	47.8	48.4	75.6	17.1
MU	28.9	38.4	99.0	28.2	35.2	83.5	98.6	52.3		62.3	61.1	52.9	23.9
PC	31.3	38.3	62.3	58.6	31.7	59.9	62.3	47.8	62.3		61.9	50.0	22.3
SC	34.0	34.7	61.3	57.0	37.1	59.9	61.1	48.4	61.1	61.9		50.2	24.3
TB	32.6	31.8	52.9	51.6	31.4	52.9	53.1	75.6	23.7	50.0	50.2		23.7
TF	28.9	31.0	23.5	28.2	31.8	27.8	23.9	17.1	23.9	22.3	24.3	23.7	

TABLE 1. Amino acid homology between IMPDH proteins^a

^a AC, Acinetobacter colcoaceticus (2); BS, Bacillus subtilis (20); CH, Chinese hamster (11); DM, Drosophila melanogaster (33); EC, Escherichia coli (39); HS1, human type I (27); HS2, human type II (27); LD, Leishmania donovani (44); MU, mouse (40); PC, Pneumocystis carinii; SC, Saccharomyces cerevisiae (19); TB, Trypanosoma brucei (45); TF, Tritrichomonas foetus (9).

of which showed that the 1.3-kb PCR product amplified from *P. carinii* genomic DNA reacted with *P. carinii* nucleic acid but not with material from the mammalian host. PFGE also supported the identification of this product as the gene from *P. carinii*, in that the probe hybridized with only a single chromosome in rat-derived *P. carinii* and failed to react with mammalian chromosomes. This probe was used in separate experiments that demonstrated remarkable chromosomal polymorphism among different forms of *P. carinii* (12, 43).

The amino acid sequence deduced from the cDNA encoding *P. carinii* IMPDH was confirmed by sequencing three peptides from a tryptic digest of the recombinant protein expressed from a T7 promoter and purified after solubilizing the *E. coli* inclusion bodies with urea. The degree of identical amino acids found between *P. carinii* IMPDH and mammalian enzymes (60 to 62%) and with the only known fungal IMPDH enzyme, *Saccharomyces cerevisiae* (62%) (19), provided further support that the cDNA that we cloned encodes the *P. carinii* enzyme. Moreover, codon bias fit the pattern known from other *P. carinii* genes.

The cDNA for the cloned IMPDH gene was determined to be 1,739 bp. Although the exact size of the genomic copy of the gene is not known because the genomic gene was not cloned, it must be equal to or exceed 2,078 bp (the number of nucleotides in the introns identified in the genomic PCR product plus the size of the cDNA). The five introns are located between nucleotides 467 and 468, 1005 and 1006, 1023 and 1024, 1051 and 1052, and 1134 and 1135 of the cDNA, and the polyadenylation signal includes nucleotides 1651 to 1656. The *P. carinii* IMPDH nucleotide sequence of the cDNA and the introns within the genomic fragment are very A/T-rich, which is consistent with the sequences of other published *P. carinii* genes (13–15).

The open reading frame for *P. carinii* IMPDH encodes a protein with 454 amino acids and a molecular weight of 50,000, making it the smallest IMPDH protein identified to date (Fig. 3). Comparisons of our deduced *P. carinii* IMPDH amino acid sequence with all the published enzyme sequences revealed the greatest identity with fungal IMPDH and mammalian enzymes (Table 1), as expected. Protein homology is greatest within the C-terminal half of the enzyme. Less homology was found between *P. carinii* IMPDH and protozoal enzymes or between the *P. carinii* protein and bacterial IMPDHs (Table 1) (2, 9, 11, 19, 20, 28, 33, 39, 40, 44, 45).

The enzyme was shorter than the enzymes from higher eukaryotes, with which it shares the greatest sequence homology, but was similar to these enzymes in other ways. For example, *P. carinii* and another fungus (*S. cerevisiae*) have IMPDH proteins with pIs of 8.54. These values were calculated using the CHARGPRO program in PC/Gene, version 6.8, which plots the charge on the protein as a function of pH. The pI for *P. carinii* IMPDH was strikingly different from that for mammalian (6.13 to 6.91) or bacterial (5.84 to 6.18) IMPDH. Protozoal IMPDH is more diverse; *L. donovani* and *Trypanosoma brucei* have basic pI values (8.86 and 8.51, respectively), but *Tritrichomonas foetus* has a pI of 6.17.

The amino acid composition of P. carinii IMPDH was analyzed with the PROSITE program in PC/Gene in order to identify sites and signatures with potential biological significance. Five potentially significant sites were detected. Possible sites for four different types of posttranslational modifications included six protein kinase C and 11 casein kinase II phosphorylation areas, three N-glycosylation signatures, and 12 potential N-myristoylation sites. The fifth signature sequence identified was a single IMPDH-GMP reductase site. This region of the protein spans amino acids 253 to 265 (Fig. 3). Gly-X-Gly-X-Gly is the motif considered to be responsible for dehydrogenase activity within proteins possessing this sequence. The Gly-X-Gly-X-Gly motif is present in all known IMPDHs in the C-terminal portion of the enzymes. Tiedeman and Smith (40) observed IMPDH catalytic activity in both the smaller E. coli protein and a truncated mouse enzyme produced as a result of expressing a mouse IMPDH clone which did not contain the full-length cDNA but which did contain this motif.

Growth of the doubly transformed S ϕ 1101 cells and IMPDH activity detected in supernatants obtained by centrifugation of extracts of these cells at 100,000 × g demonstrates that the protein expressed from the cloned *P. carinii* IMPDH gene is catalytically active. Failure to observe activity in urea-solubilized inclusion bodies may reflect an altered tertiary structure of the protein. IMPDH proteins in bacteria and mammals are homotetramers (46). Conditions for assembling homotetrameric enzyme from urea-solubilized inclusion bodies have not been determined, but optimization could improve the yield of catalytically active IMPDH. The tetramer may not be the only active form of the enzyme. Two groups have reported that a dimer is the smallest functional unit that has catalytic activity (21, 29). These possibilities remain to be explored.

Preliminary characterization of the kinetic properties of recombinant *P. carinii* IMPDH assayed in extracts of the *E. coli* strain lacking IMPDH showed patterns typical for IMPDH. The apparent K_m for IMP was $21.7 \pm 0.3 \mu$ M and that for NAD was $314 \pm 84 \mu$ M. Most forms of mammalian IMPDH have similar K_m values for IMP, ranging from 14 to 30 μ M (18). K_m values for NAD vary more widely. For example, the K_m values for mammalian forms of IMPDH for NAD range between 24 and 46 μ M, but the value for *Eimeria tenella* is 150 μ M (18) and the value for *T. foetus* is 340 μ M (17). The large difference between NAD K_m values for mammalian IMPDH and *P. carinii* IMPDH again supports the conclusion that the cDNA cloned and expressed is not from mammalian host cells that could have contaminated the original DNA preparations.

The properties of the IMPDH protein from *P. carinii* suggest that the enzyme is worthy of consideration as a drug target for this organism. This conclusion is based on the potent action of mycophenolic acid in culture and the degree of amino acid homology between the *P. carinii* enzyme and the two human IMPDH isozymes. The two human enzymes are 59.9 to 62.3% homologous in amino acid sequence to the *P. carinii* enzyme. The differences between the mammalian forms of the enzyme and the *P. carinii* IMPDH confer a substantial difference in pI. Whether this difference in homology can also serve as the basis for selective binding of inhibitors that would have more activity against the organism and less toxicity for the human host is the subject for future studies.

ACKNOWLEDGMENTS

Mycophenolic acid was tested in culture against *P. carinii* under contract NO1-AI-72647 (J. W. Smith, principal investigator and S. F. Queener and Marilyn S. Bartlett, coinvestigators; used with permission).

REFERENCES

- Alwine, J. C., D. J. Kemp, and G. R. Stark. 1977. Method for detection of specific mRNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. Proc. Natl. Acad. Sci. USA 74:5350– 5354.
- Anderegge, U., W. H. Schnuck, O. Asperger, and H. P. Kleber. 1992. EMBL accession no. P31002. EMBL Data Library Submissions, Heidelberg, Germany.
- Balzarini, J., A. Karlsson, L. Wang, C. Bohman, K. Horska, I. Votruba, A. Fridland, A. V. Aerschot, P. Herdewijn, and E. DeClercq. 1993. Eicar (5ethynyl-1-beta-D-ribofuranosylimidazole-4-carboxamide). A novel potent inhibitor of inosinate dehydrogenase activity and guanylate biosynthesis. J. Biol. Chem. 268:24591–24598.
- Bartlett, M. S., T. D. Edlind, M. M. Durkin, M. M. Shaw, S. F. Queener, and J. W. Smith. 1992. Antimicrotubule benzimidazoles inhibit in vitro growth of *Pneumocystis carinii*. Antimicrob. Agents Chemother. 36:779–782.
- Bartlett, M. S., T. D. Edlind, C. H. Lee, R. Dean, S. F. Queener, M. M. Shaw, and J. W. Smith. 1994. Albendazole inhibits *Pneumocystis carinii* proliferation in inoculated immunosuppressed mice. Antimicrob. Agents Chemother. 38:1834–1837.
- Bartlett, M. S., J. A. Fishman, S. F. Queener, M. M. Durkin, M. A. Jay, and J. W. Smith. 1988. New rat model of *Pneumocystis carinii* infection. J. Clin. Microbiol. 26:1100–1102.
- Bartlett, M. S., S. F. Queener, M. M. Shaw, J. D. Richardson, and J. W. Smith. 1994. *Pneumocystis carinii* is resistant to imidazole antifungal agents. Antimicrob. Agents Chemother. 38:1859–1861.
- Bartlett, M. S., M. M. Shaw, P. Navaran, J. W. Smith, and S. F. Queener. 1995. Evaluation of potent inhibitors of dihydrofolate reductase in a culture model for growth of *Pneumocystis carinii*. Antimicrob. Agents Chemother. 39:2436–2441.
- Beck, J. T., S. Zhao, and C. C. Wang. 1994. Cloning, sequencing, and structural analysis of the DNA encoding inosine monophosphate dehydrogenase (EC 1.1.1.205) from *Tritrichomonas foetus*. Exp. Pathol. 78:101–112.
- Chomczynski, P., and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidium thiocyanate-phenol-chloroform extraction. Anal. Biochem. 162:156–159.

- Collart, F. R., and E. Huberman. 1988. Cloning and sequence analysis of the human and chinese hamster inosine-5'-monophosphate dehydrogenase cDNAs. J. Biol. Chem. 263:15769–15772.
- Cushion, M. T., J. Zhang, M. Kaselis, D. Guintoli, S. L. Stringer, and J. R. Stringer. 1993. Evidence for two genetic variants of *Pneumocystis carinii* coinfecting laboratory rats. J. Clin. Microbiol. 31:1217–1223.
- Edlind, T. D., M. S. Bartlett, G. A. Weinberg, G. N. Prah, and J. W. Smith. 1992. The beta-tubulin gene from rat and human isolates of *Pneumocystis carinii*. Mol. Microbiol. 6:3365–3373.
- Edman, J. C., U. Edman, M. Cao, B. Lundgren, J. A. Kovacs, and D. V. Santi. 1989. Isolation and expression of the *Pneumocystis carinii* dihydrofolate reductase gene. Proc. Natl. Acad. Sci. USA 86:8625–8629.
- Edman, U., J. C. Edman, B. Lundgren, and D. V. Santi. 1989. Isolation and expression of the *Pneumocystis carinii* thymidylate synthase gene. Proc. Natl. Acad. Sci. USA 86:6503–6507.
- Freifelder, D. M. 1987. Molecular biology, 2nd ed., p. 353. Jones and Bartlett Publishers, Inc., Boston, Mass.
- Hedstrom, L., and C. C. Wang. 1990. Mycophenolic acid and thiazole adenine dinucleotide inhibition of *Tritrichomonas foetus* inosine 5'-monophosphate dehydrogenase: implications on enzyme mechanism. Biochemistry 29: 849–854.
- Hupe, D. J., B. A. Azzolina, and N. D. Behrens. 1986. IMP dehydrogenase from the intracellular parasitic protozoan *Eimeria tenella* and its inhibition by mycophenolic acid. J. Biol. Chem. 261:8363–8369.
- Johnston, M. 1994. GenBank accession no. U00029. GenBank, National Center for Biotechnology Information, Bethesda, Md.
- Kanzaki, N., and K. Miyagawa. 1990. Nucleotide sequence of the Bacillus subtilis IMP dehydrogenase gene. Nucleic Acids Res. 18:6710.
- Krishnaiah, K. V. 1975. Inosinic acid 5'-monophosphate dehydrogenase of Escherichia coli: purification by affinity chromatography and some properties. Arch. Biochem. Biophys. 170:567–575.
- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–685.
- Landman, O., D. Shiffman, Y. Av-Gay, Y. Aharaonowitz, and G. Cohen. 1991. High level expression in *Escherichia coli* of isopenicillin N synthase genes from *Flavobacterium* and *Streptomyces*, and recovery of active enzyme from inclusion bodies. FEMS Microbiol. Lett. 84:239–244.
- Lee, C. H., C. Cheng, J. Wang, and L. Lumeng. 1992. Identification of hepatitis C viruses with a conserved sequence of the 5'-untranslated region. J. Clin. Microbiol. 30:1602–1604.
- Lee, C. H., J. J. Lu, M. S. Bartlett, M. M. Durkin, T. H. Liu, J. Wang, B. Jiang, and J. W. Smith. 1993. Nucleotide sequence variation in *Pneumocystis carinii* strains that infect humans. J. Clin. Microbiol. 31:754–757.
- Liu, Y., R. R. Tidwell, and M. J. Leibowitz. 1994. Inhibition of in vitro splicing of a group I intron of *Pneumocystis carinii*. J. Eukaryot. Microbiol. 41:31–38.
- Natsumeda, Y., and S. F. Carr. 1993. Human type I and II IMP dehydrogenases as drug targets. Ann. N.Y. Acad. Sci. 696:88–93.
- Natsumeda, Y., S. Ohno, H. Kawasaki, Y. Konno, G. Weber, and K. Suzuki. 1990. Two distinct cDNAs for human IMP dehydrogenase. J. Biol. Chem. 265:5292–5295.
- 28a.Nygaard, P. (Department of Biological Chemistry, Institute of Molecular Biology, University of Copenhagen). Personal communication.
- Okada, M., K. Shimura, H. Shiraki, and H. Nakagawa. 1983. IMP dehydrogenase. II. Purification and properties of the enzyme from Yoshida sarcoma ascites tumor cells. J. Biochem. 94:1605–1613.
- Queener, S. F. 1995. New drug developments for opportunistic infections in immunosuppressed patients: *Pneumocystis carinii*. J. Med. Chem. 38:4739– 4759.
- Rave, N., R. Crkvenjakov, and H. Boedtker. 1979. Identification of procollagen mRNAs transferred to diazobenzyloxymethyl paper from formaldehyde agarose gels. Nucleic Acids Res. 6:3559–3567.
- Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higu-chi, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermo-stable DNA polymerase. Science 239:487–491.
- Sifri, C. D., K. Wilson, S. Smolik, M. Forte, and B. Ullman. 1994. Cloning and sequence analysis of a *Drosophila melanogaster* cDNA encoding IMP dehydrogenase. Biochim. Biophys. Acta Gene Struct. Expression 1217:103– 106.
- Sinclair, K., A. E. Wakefield, S. Banerji, and J. M. Hopkin. 1991. Pneumocystis carinii organisms derived from rat and human hosts are genetically distinct. Mol. Biochem. Parasitol. 45:183–184.
- Sogin, M. L., and J. C. Edman. 1989. A self-splicing intron in the small subunit rRNA gene of *Pneumocystis carinii*. Nucleic Acids Res. 17:5349– 5359.
- Southern, E. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98:503–517.
- Stringer, J. R., S. L. Stringer, J. Zhang, R. Baughman, A. G. Smulian, and M. T. Cushion. 1993. Molecular genetic distinction of *Pneumocystis carinii* from rats and humans. J. Eukaryot. Microbiol. 40:733–741.

48 O'GARA ET AL.

- Tabor, S. 1990. Expression using the T7 RNA polymerase/promoter system, p. 16.2.1–16.2.11. *In* F. A. Ausubel, R. Brent, R. E. Kingston, et al. (ed.), Current protocols in molecular biology, John Wiley & Sons, Inc., New York, N.Y.
- Tiedeman, A. A., and J. M. Smith. 1985. Nucleotide sequence of the guaB locus encoding IMP dehydrogenase of *Escherichia coli* K12. Nucleic Acids Res. 13:1303–1316.
- Tiedeman, A. A., and J. M. Smith. 1991. Isolation and sequence of a cDNA encoding mouse IMP dehydrogenase. Gene 97:289–293.
- 41. Van Berg, A. A., P. A. W. Mooyer, H. Van Lenthe, E. H. Stet, R. A. De Abreu, A. B. P. Van Kuilenburg, and A. H. Van Gennip. 1995. The IMP dehydrogenase inhibitor mycophenolic acid antagonizes the CTP synthetase inhibitor 3-deazauridine in MOLT-3 human leukemia cells: a central role for phosphoribosyl pyrophosphate. Biochem. Pharmacol. 50:1095–1098.
- 42. Weinberg, G. A., and P. J. Durant. 1994. Genetic diversity of Pneumocystis

carinii derived from infected rats, mice, ferrets, and cell cultures. J. Eukaryot. Microbiol. **41**:219–224.

- Weinberg, G. A., M. J. O'Gara, and M. T. Cushion. 1994. Coinfection of rats with genetically diverse forms of *Pneumocystis carinii* demonstrated by *P. carinii* inosine monophosphate dehydrogenase gene polymorphism. J. Eukaryot. Microbiol. 41:119S.
- 44. Wilson, K., F. R. Collart, E. Huberman, J. R. Stringer, and B. Ullman. 1991. Amplification and molecular cloning of the IMP dehydrogenase gene of *Leishmania donovani*. J. Biol. Chem. 266:1665–1671.
- Wilson, K. E., R. L. Berens, C. D. Sifri, and B. Ullman. 1993. GenBank accession no. M97794. GenBank, National Center for Biotechnology Information, Bethesda, Md.
- Yamada, Y., Y. Natsumeda, and G. Weber. 1988. Action of the active metabolites of tiazofurin on purified IMP dehydrogenase. Biochemistry 27: 2193–2196.