Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 8625–8629, November 1989
Biochemistry

Isolation and expression of the *Pneumocystis carinii* dihydrofolate reductase gene

JEFFREY C. EDMAN*†, URSULA EDMAN‡, MING CAO*, BETTINA LUNDGREN§, JOSEPH A. KOVACS§, AND DANIEL V. SANTI¶‖

*Hormone Research Institute and Departments of †Laboratory Medicine, ¶Biochemistry and Biophysics, and ‖Pharmaceutical Chemistry and ‡Intercampus Program for Molecular Parasitology, University of California, San Francisco, CA 94143-0534; and §Critical Care Medicine Department, National Institutes of Health, Bethesda, MD 20892

Communicated by William J. Rutter, July 17, 1989 (received for review June 8, 1989)

ABSTRACT ***Pneumocystis carinii* dihydrofolate reductase (DHFR; 5,6,7,8-tetrahydrofolate: NADP$^+$ oxidoreductase, EC 1.5.1.3) cDNA sequences have been isolated by their ability to confer trimethoprim resistance to *Escherichia coli*. Consistent with the recent conclusion that *P. carinii* is a member of the Fungi, sequence analysis and chromosomal localization show that DHFR is neither physically nor genetically linked to thymidylate synthase. Expression of recombinant *P. carinii* DHFR in heterologous hosts provides an abundant source of the enzyme that may form a basis for the development of new therapies for this enigmatic pathogen. Studies with the recombinant enzyme show that trimethoprim is a very poor inhibitor of *P. carinii* DHFR and, in fact, is a more potent inhibitor of human DHFR.**

Pneumocystis carinii pneumonia is a leading cause of morbidity and mortality in AIDS (1, 2). Since the onset of the AIDS epidemic, the incidence of *P. carinii* pneumonia has risen from ≈200 to >25,000 cases per year in the United States. Due to the lack of a continuous *in vitro* culture system and the cumbersome nature of the rat model of *P. carinii* pneumonia, the development of anti-*P. carinii* therapy has rested largely on the assumption that antiprotozoan agents were likely to be effective (3–5). Despite the fact that *P. carinii* has recently been shown to be a member of the Fungi (6), the two principal therapeutic modalities, trimethoprim/sulfamethoxazole and pentamidine, were developed on this basis. Before the AIDS epidemic, these agents were sufficient for treatment of the rare cases of *P. carinii* pneumonia. However, in the human immunodeficiency virus-positive patient, therapy and prophylaxis with the standard anti-*P. carinii* agents are complicated by frequent toxic and allergic side effects (7). New compounds active against *P. carinii* are clearly needed.

The inability to propagate *P. carinii* reliably *in vitro* and the limited quantities of *P. carinii* enzymes that can be purified from infected rat lungs have hindered the search for anti-*P. carinii* agents. Purification and characterization of the intracellular targets for such agents would enable the development of new therapies for *P. carinii* pneumonia. Of the known anti-*P. carinii* agents, the dihydrofolate reductase (DHFR; 5,6,7,8-tetrahydrofolate: NADP$^+$oxidoreductase, EC 1.5.1.-3) inhibitors are the most thoroughly characterized. DHFR plays a central role in the *de novo* synthesis of nucleic acid precursors. DHFR inhibitors (methotrexate, trimethoprim, pyrimethamine) are effective antineoplastic, antibacterial, and antiprotozoal agents. *P. carinii* pneumonia clearly responds to the combination of a DHFR inhibitor (trimethoprim or pyrimethamine) and a sulfonamide. Despite their obvious efficacy when used in conjunction with a sulfonamide, available information suggests that trimethoprim and pyrimethamine, are in themselves poor inhibitors of *P. carinii* DHFR (IC_{50} values of 39,600 and 2400 nM, respectively, compared with 8 and 2500 nM for *Escherichia coli* DHFR at similar substrate concentrations; refs. 8 and 9). Other antifolates have been shown to be more effective inhibitors of *P. carinii* DHFR (8, 10) but require concomitant administration of leucovorin to prevent host toxicity. If pure *P. carinii* DHFR were available for study, then it should not be difficult to find inhibitors that surpass the efficacy of the known antifolates.

To facilitate the development of new anti-*P. carinii* agents and to further extend our knowledge of the basic biology of this unusual organism, DNA sequences encoding *P. carinii* DHFR have been isolated.** Expression of *P. carinii* DHFR in a heterologous host (*E. coli*) demonstrates the applicability of recombinant DNA technology in defining approaches toward new therapies for *P. carinii* pneumonia.

MATERIALS AND METHODS

Bacterial Strains and Media. *E. coli* strains XL1B and HB101 were used as bacteriophage λ and plasmid hosts. The DHFR-deficient strain D3-157 used for expression experiments was obtained from Sara Singer (11). *E. coli* JM101 and M13 vectors were used for DNA sequencing. Medium used was Luria broth plus appropriate antibiotics (ampicillin at 100 μg/ml and/or trimethoprim at 100 μg/ml).

***P. carinii* Preparations.** *P. carinii* from steroid-treated rats and human *P. carinii* from autopsy samples were prepared as described (12).

***P. carinii* DNA Libraries and Phagemid Rescue.** The construction of a rat *P. carinii* cDNA library in the bacteriophage vector λ-ZAP (Stratagene) has been described (13). Phagemids were rescued from the library by coinfection of *E. coli* XL1B with 8×10^7 plaque-forming units of λ phage and 2×10^9 plaque-forming units of R408 helper phage in 45 ml of Luria broth (14). After 4 hr at 37°C, the culture was heated to 70°C for 30 min, clarified by centrifugation, and the supernatant containing rescued phagemid was saved. The supernatant had a titer of 4×10^4 ampicillin-resistant colony-forming units per μl.

Transverse-Alternating Field Electrophoresis of *P. carinii* Chromosomes. *P. carinii* chromosomes from infected rat lung were prepared for transverse-alternating field-effect electrophoresis (ref. 15; B.L. and J.A.K., unpublished work). Electrophoresis was performed in three stages: stage 1, 30 min at 170 mA with a pulse time of 4 sec; stage 2, 15 hr at 150 mA with a pulse time of 25 sec; stage 3, 7 hr at 170 mA with a pulse time of 35 sec in 0.8% agarose with 10 mM Tris/0.5

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: DHFR, dihydrofolate reductase; nt, nucleotide(s).
**The sequences reported in this paper have been deposited in the GenBank data base (accession nos. M26495 and M26496).

mM EDTA/4.3 mM acetic acid at 16°C. After electrophoresis, the gel was photographed, exposed to a 302-nm UV transilluminator for 7 min, denatured in 0.5 M NaOH/1.5 M NaCl for 1 hr, neutralized in 1 M Tris·HCl, pH 8/1.5 M NaCl and blotted to a Hybond-N nylon filter in 10× SSC (1× SSC is 0.15 M sodium chloride/0.015 M sodium citrate, pH 7) for 18 hr. The membrane was baked for 2 hr at 80°C and prehybridized in 50% (wt/vol) formamide, 5× Denhardt's solution (1× Denhardt's solution is 0.02% polyvinylpyrrolidone/0.02% Ficoll/0.02% bovine serum albumin), and 5× SSC at 42°C. Hybridization was initiated by the addition of ^{32}P-labeled 710-base-pair (bp) *Sal* I–*Ssp* I fragment at 10^5 cpm/ml containing the entire coding region of *P. carinii* DHFR and continuing incubation at 42°C for 16 hr. The membrane was washed twice for 15 min in 2× SSC at room temperature, three times for 15 min in 0.1× SSC at 50°C, and exposed to x-ray film.

Expression Plasmid Construction and Characterization of Expression Products. The coding region from a cDNA clone of *P. carinii* DHFR was amplified using the polymerase-chain-reaction technique (16). The 5′ primer was dGGGATCCATATGAATCAGCAAAAGTCTTT, which contains nucleotides (nt) 1–20 and creates *Bam*HI and *Nde* I sites proximal to the initiation codon. The 3′ primer was dGGTCGACAAGCTTTAATATTGAATAAATAGAATAA (complementary to nt 669–689) and creates *Sal* I and *Hin*dIII sites. The amplified product was digested with *Bam*HI and *Hin*dIII and subcloned into Bluescript to yield pDHFR1. pDHFR1 was sequenced to ensure that the *P. carinii* DHFR-coding region was unaltered by the amplification process. A 712-bp *Sma* I and *Nde* I fragment of pDHFR1 containing the entire coding region of *P. carinii* DHFR was subcloned into the *Nde* I–*Sma* I-digested expression vector pDLTS-Nde (17) to generate pDLDHFR. This plasmid was introduced into the DHFR-deficient *E. coli* strain D3-157 (11). Cells containing pDLDHFR were grown to saturation in LB medium plus ampicillin at 100 μg/ml, collected by centrifugation, resuspended in 10 mM Tris·HCl, pH 8.0/1 mM EDTA, and lysed by sonication. Soluble protein extracts were then prepared by centrifugation. Purification of *P. carinii* DHFR was achieved by the passage of crude extracts over methotrexate-Sepharose and elution with 2 mM dihydrofolate (18).

RESULTS AND DISCUSSION

Isolation of *P. carinii* DHFR cDNA Sequences. Attempts to isolate *P. carinii* DHFR gene sequences by low-stringency hybridization with a variety of cloned DHFR sequences as probes were unsuccessful, presumably due to the poor conservation of DHFR sequences (19). As trimethoprim had been shown to be a weak inhibitor of partially purified *P. carinii* DHFR (8), we attempted to isolate *P. carinii* DHFR sequences by their ability to confer trimethoprim resistance to *E. coli*. Such a direct expression procedure allowed the isolation of the *Saccharomyces cerevisiae* DHFR gene (20, 21). A plasmid equivalent of a rat *P. carinii* cDNA library (13) cloned into the λ ZAP vector was generated by coinfection of *E. coli* XL-1B cells with the λ phage library and the replication-defective single-stranded phage R408 (14). This technique results in the excision of inserts and adjoining plasmid sequences from ZAP recombinants and packaging of their single-stranded equivalents into phagemids. The supernatant from this infection was heated to inactivate λ phage and contained 4×10^4 ampicillin-resistant colony-forming units per μl. Infection of XL-1B with 10^6 colony-forming units and plating to medium containing trimethoprim at 100 μg/ml allowed the isolation of 30 trimethoprim-resistant colonies. Plasmid DNA was isolated from each and used to transform *E. coli* HB101. Two plasmids were found capable of transferring trimethoprim resistance. The remainder of the original trimethoprim-resistant clones were presumed to represent chromosomal DHFR mutations.

Sequence of *P. carinii* DHFR cDNA. Restriction analysis of the two plasmids showed substantial overlap, and one plasmid was chosen for DNA sequence analysis. The sequence of the entire 848-bp insert of ppcDHFRcD was determined (Fig. 1) and was shown to contain a 618-bp open reading frame. The predicted amino acid sequence from this open reading frame demonstrated significant similarity to known DHFR sequences (Fig. 2). To ensure that the DHFR sequence obtained by the direct selection procedure was not altered due to the presence of trimethoprim, independent clones were isolated by nucleic acid hybridization. The coding region of one of these isolates was identical to clones originally obtained by trimethoprim selection.

The predicted start codon of the cDNA conforms poorly to the favorable environment determined for *S. cerevisiae* translational initiation (RNNATGRNNT; ref. 29). The sequences obtained are A + T rich (68% in the coding region; 71% overall). Codon usage is strongly biased toward adenosine or thymidine in the third position. Of the 194 codons where a choice between AT or GC in the third position is possible, 168 (87%) are found to contain an adenosine or a thymidine residue.

The *P. carinii* DHFR Gene Has an Intron. To assess the natural variability of DHFR sequences from different sources of *P. carinii*, specific primers were used in conjunction with the polymerase chain reaction to amplify DHFR coding sequences from DNA isolated from five rat and two human *P. carinii* preparations. The expected amplification product was 712 bp. However, amplification of *P. carinii* genomic DNA revealed a 750-bp product from all sources. This product was subcloned and shown to contain a 43-bp insertion at nt 267 (Fig. 1). This sequence disrupts the reading frame and demonstrates features consistent with a small intron. The 5′- and 3′- splice junctions, GCA and TAG, correspond to the consensus sequences for these regions determined in other fungal introns (30). In addition, the intron size is similar to those seen in the *P. carinii* thymidylate synthase gene (45–55 nt; ref. 13). Other than the intron, the only change in the sequence of the 750-bp product was an A → G transition in the 3′-untranslated region (position 654).

***P. carinii* DHFR Is Not Bifunctional.** In all protozoa where DHFR has been characterized, it is found as a bifunctional protein with thymidylate synthase (TS-DHFR; refs. 31 and 32). In contrast, all other organisms for which these genes have been characterized contain separate thymidylate synthase and DHFR genes and proteins. The *P. carinii* DHFR gene does not encode a bifunctional protein. Furthermore, by chromosomal localization (see below), the thymidylate synthase and DHFR genes are not linked. The absence of a bifunctional TS-DHFR further supports the conclusion that *P. carinii* is not a member of the Protozoa (6).

Comparison of *P. carinii* DHFR with Other Known DHFR Sequences. *P. carinii* DHFR genomic sequences predict a protein of 206 amino acids and a M_r of 23,868. This size is in accord with the value of 20,000–30,000 previously determined by gel filtration of partially purified *P. carinii* DHFR (33) and, as in *S. cerevisiae*, is slightly larger than mammalian and bacterial DHFRs. In Fig. 2, the protein sequence is aligned to the DHFR sequences from *Lactobacillus casei* (22), human (23), *S. cerevisiae* (20, 21, 24), and *Leishmania major* (25, 26). Due to the low degree of similarity between DHFRs, alignments in the absence of three-dimensional structures must be somewhat arbitrary. Nonetheless, both primary structure and three-dimensional structure were considered in making these alignments (19, 27). *P. carinii* DHFR contains those residues that are conserved in all DHFRs. The larger size of *P. carinii* DHFR is accounted for primarily by insertions in regions that are loops between known secondary

```
                                    1                                  10                                  20
                                   Met Asn Gln Gln Lys Ser Leu Thr Leu Ile Val Ala Leu Thr Thr Ser Tyr Gly Ile Gly Arg Ser
-30   TTTTGTCGACGGACAGTCCGTTTGATTTTG  ATG AAT CAG CAA AAG TCT TTA ACA TTG ATT GTT GCA CTT ACA ACT TCT TAT GGA ATT GGC CGA TCA

                                   30                                  40                                  50
      Asn Ser Leu Pro Trp Lys Leu Lys Lys Glu Ile Ser Tyr Phe Lys Arg Val Thr Ser Phe Val Pro Thr Phe Asp Ser Phe Glu Ser Met
+67   AAC TCT CTT CCA TGG AAA TTA AAG AAA GAA ATA AGT TAT TTT AAA CGA GTA ACC TCT TTT GTA CCA ACT TTT GAT TCA TTT GAA TCG ATG

                                   60                                  70                                  80
      Asn Val Val Leu Met Gly Arg Lys Thr Trp Glu Ser Ile Pro Leu Gln Phe Arg Pro Leu Lys Gly Arg Ile Asn Val Val Ile Thr Arg
+157  AAT GTT GTA TTG ATG GGT CGA AAA ACA TGG GAA AGT ATT CCT TTG CAA TTT CGG CCC CTT AAA GGT CGT ATT AAT GTT GTT ATC ACT CGA

                                   90                                  100                                 110
      Asn Glu Ser Leu Asp Leu Gly Asn Gly Ile His Ser Ala Lys Ser Leu Asp His Ala Leu Glu Leu Leu Tyr Arg Thr Tyr Gly Ser Glu
+247  AAT GAA TCT CTG GAT CTA GGA AAT GGA ATT CAT TCT GCA AAA TCC TTG GAT CAT GCT TTG GAA TTG TTA TAT CGT ACA TAT GGT TCT GAA

          GCATGTTAAAAAAATTAATCTGAAATTCGATTGACAATATTAG

                                   120                                 130                                 140
      Ser Ser Val Gln Ile Asn Arg Ile Phe Val Ile Gly Gly Ala Gln Leu Tyr Lys Ala Ala Met Asp His Pro Lys Leu Asp Arg Ile Met
+337  AGT TCG GTT CAA ATT AAT CGA ATT TTC GTT ATA GGT GGT GCA CAG CTA TAT AAA GCA GCT ATG GAT CAT CCT AAA TTA GAT AGA ATT ATG

                                   150                                 160                                 170
      Ala Thr Ile Ile Tyr Lys Asp Ile His Cys Asp Val Phe Phe Pro Leu Lys Phe Arg Asp Lys Glu Trp Ser Ser Val Trp Lys Lys Glu
+427  GCT ACA ATA ATA TAC AAG GAT ATT CAT TGT GAT GTA TTT TTT CCA CTT AAA TTT AGG GAT AAA GAA TGG TCT TCT GTA TGG AAA AAA GAA

                                   180                                 190                                 200
      Lys His Ser Asp Leu Glu Ser Trp Val Gly Thr Lys Val Pro His Gly Lys Ile Asn Glu Asp Gly Phe Asp Tyr Glu Phe Glu Met Trp
+517  AAA CAT TCA GAT TTA GAA TCT TGG GTT GGT ACT AAA GTT CCT CAT GGT AAA ATA AAT GAA GAC GGT TTT GAT TAT GAA TTC GAA ATG TGG

                  206
      Thr Arg Asp Leu OC
+607  ACA AGA GAT TTA TAA ATCCCTTTCAAATCTTTTTATGGCTTTTTAATACTACTAATTCTCTTGTTATTCTATTTATTCAATATTCTAAATTCTTTTATTTCGAAAACCTTTCGA
                                                        G

+721  TCTACCAATTCAACTCTCTTCATACTTTTGCTATCTCATAACTATAAAATCCTACATATTAATTACAAAAAAAATATCTATATTCAATTAAACAAAAA
```

FIG. 1. *P. carinii* DHFR cDNA sequence. The DNA sequence of the 848-bp insert of ppcDHFRcD is presented. The first residue of the predicted start codon is designated nt 1. Nucleotide numbers are at left, and amino acid numbers are above the sequence. The 43-bp intron in the genomic sequence is shown below the cDNA sequence. The single base change in the genomic clones is shown below the sequence at position 654.

structure elements (19). However, similar to *S. cerevisiae* DHFR, a particularly large insertion is present near the carboxyl terminus (amino acids 181–193). The most dramatic structural differences between bacterial and vertebrate DHFR also occur in this region (27). A β-sheet (βG in Fig. 2) in the bacterial enzyme is changed to a β bulge due to a small insertion in the vertebrate enzyme. Whether or not the insertion seen in *P. carinii* DHFR results in a similar structural change or even more radical change relative to the bacterial enzyme must await three-dimensional structure analysis.

Chromosomal Localization of the *P. carinii* DHFR Gene. It has recently been shown that *P. carinii* contains 14 chromosomes ranging in size from 295–710 kilobase pairs (kbp; ref.

```
                                  |-βA--|                  |----αB-----|           |--βB-| |-αC-|
L.casei                             TAFLWAQDRDGLIGKDGHLPW-HLPDDLHYFRAQTV------------GKIMVVGRRTYESFP   50
H.sapiens                         VGSLNCIVAVSQNMGIGKNGDLPWPPLRNEFRYFQRMTTTSSVEGK-----QNLVIMGKKTWFSIP   63

P.carinii                       MNQQKSLTLIVALTTSYGIGRSNSLPW-KLKKEISYFKRVTSFVPTFDSFES--MNVVLMGRKTWESIP   69

S.cerevisiae                  MAGGKIPIVGIVACLQPEMGIGFRGGLPW-RLPSEMKYFRQVTSLTKDPNK-----KNALIMGRKTWESIP   71
L.major       MSRAAARFKIPMPETKADFAFPSLRAFSIVVALDMQHGIGDGESIPW-RVPEDMTFFKNQTTLLRNKKPPTEKKRNAVVMGRKTWESVP   93


                        |-βC-|           |βD|        |---αE---|             |-βE-| |----αF----|
L.casei         --KRPLPERTNVVLTHQEDYQAQG-AVVVH----------DVAAVFAYAKQHPDQ--------ELVIAGGAQIFTAFK-----DDVD   111
H.sapiens       EKNRPLKGRINLVLSRELKEPPQGAHFLSR---------S-LDDALKLTEQPELANKVD----MVWIVGGSSVYKEAMNH---PGHL   130

P.carinii       LQFRPLKGRINVVITRNESLDLGNGIHSAK---------S-LDHALELLYRTYGSESSVQIN-RIFVIGGAQLYKAAMDH---PKLD   140

S.cerevisiae    PKFRPLPNRMNVIISRSFKDDFVHDKERSI----------VQSNSLANAIMNLESNFKEHLE-RIYVIGGGEVYSQIF-----SITD   136
L.major         VKFRPLKGRLNIVLSSKATVEELLAPLPEGQRAAAAQDVVVVNGGLAEALRLLARPLYCSSIETAYCVGGAQVYADAMLSPCIEKLQ   175


                  |---βF--|                  |--βG                                  βG--|       |--βH--|
L.casei         TLLVTRLAGSFEGDT-KMIPLNWDDFTKVSS----------------------------RTVEDTNPALTHTYEVWQKKA          162
H.sapiens       KLFVTRIMQDFESDT-FFPEIDLEKYKL--------------------LPEYPG--VLSDVQEEKG--IKYKFEVYEKND          188

P.carinii       RIMATIIYKDIHCDV-FFPLKFRDKEWSSVWKKEKHSD----------LESWVGTKVPHGKINEDG--FDYEFEMWTRDL          206

S.cerevisiae    HWLITKINPLDKNAT-PAMDTFLDAKKLEEVFSEQDPAQLKEFLPKVELPETDC--DQRYSLEEKG--YCFEFTLYNRK          211
L.major         EVYLTRIYATAPACTRFFPFPPENAAT--------------------AWDLASS--QGRRKSEAEG--LEFEICKYVPRN          230
```

FIG. 2. Alignment of *P. carinii* DHFR with selected DHFRs. The sequences from *Lactobacillus casei* (22), human (23), *P. carinii*, *S. cerevisiae* (20, 21, 24), and *Leishmania major* (25, 26) DHFRs are aligned. Residues that are conserved in all known DHFRs are in boldface. Alignments were achieved by taking into account both the obvious primary structure similarities and three-dimensional structure equivalence (27). Above the alignments are the secondary-structure regions as defined for the *L. casei* enzyme (28).

34, and B.L. and J.A.K., unpublished work). In addition, there is a variation of pattern of chromosomes in different isolates (Fig. 3*b*, lanes 1, 3, and 4). A *P. carinii* DHFR probe was used to probe a filter-replicate of *P. carinii* chromosomes separated by transverse-alternating-field electrophoresis. Two isolates were shown to have DHFR sequences on a single chromosome of 520 kbp (Fig. 3*a*, lanes 3 and 4). A third isolate had DHFR sequences on a chromosome of slightly larger molecular weight (Fig. 3*a*, lane 1). The significance of the size variation is unclear but may reflect strain differences in *P. carinii*. No hybridization to *S. cerevisiae* chromosomes was detected (Fig. 3*a*, lane 2). As *P. carinii* thymidylate synthase is found on a chromosome of 330 kbp (13), there is no genetic linkage of these proteins.

Expression of *P. carinii* DHFR in *E. coli*. Trimethoprim resistance in the original isolates indicated that functional *P. carinii* DHFR was being synthesized. The total DHFR activity in these isolates was equal to *E. coli* containing vector alone (data not shown). This small amount of *P. carinii* DHFR is apparently sufficient to overcome the level of trimethoprim used to isolate these clones. However, further work on *P. carinii* DHFR depends on the ability to generate sufficient quantities of the enzyme to perform detailed structural and kinetic characterization. To enhance the level of expression, the coding region of *P. carinii* DHFR was placed in the expression vector pDLTS-Nde (17) and used to transform the DHFR-deficient *E. coli* strain D3-157. The expression construct could complement the DHFR deficiency in this strain and significantly increase its growth rate in rich medium. Lysates of cells with this construct contain an ≈25-kDa protein not present in cells containing vector alone (Fig. 4, lanes a and b). *P. carinii* DHFR purified from these cells by methotrexate-Sepharose affinity chromatography comigrates with the ≈25-kDa protein (Fig. 4, lane c). Enzymatic activity measurements and densitometric scanning of stained gels showed that the level of *P. carinii* DHFR was

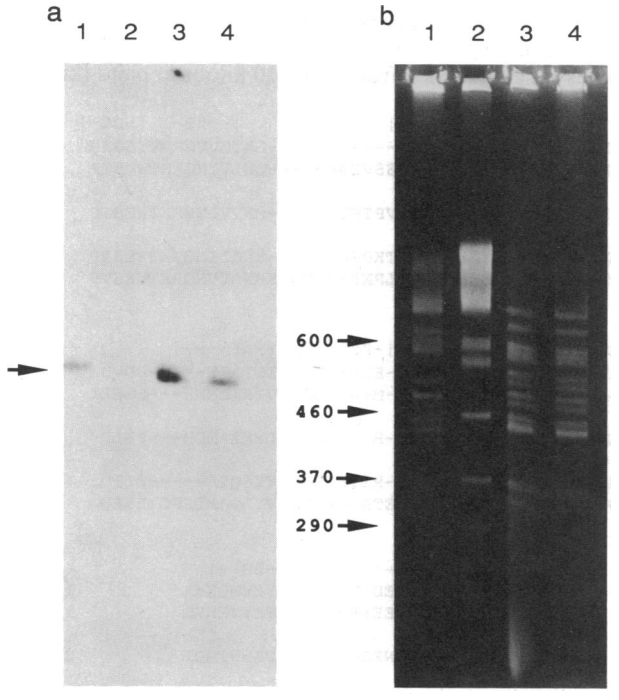

FIG. 3. Chromosomal localization of the *P. carinii* DHFR gene. (*a*) Autoradiogram of the filter replicate of the gel in *b* hybridized to a *P. carinii* DHFR probe. (*b*) Ethidium bromide-stained gel. Lanes: 1, 3, and 4; *P. carinii* trophozoites from three different infected rats; 2, *S. cerevisiae* chromosomes (Clontech). Note that all 14 chromosomes are not resolved under these conditions.

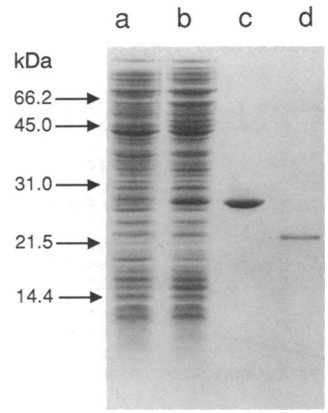

FIG. 4. Expression of *P. carinii* DHFR in *E. coli*. Extracts from cells containing pDLDHFR (lane b) and cells containing vector alone (lane a) were separated on a 15% polyacrylamide gel and stained with Coomassie brilliant blue. Lanes c and d are affinity-purified preparations of *P. carinii* DHFR from cells containing pDLDHFR and commercially obtained *E. coli* DHFR (Sigma), respectively. The migration of size standards is indicated at left.

≈5% of soluble protein in *E. coli* containing the *P. carinii* DHFR expression construct.

Trimethoprim Is a Poor Inhibitor of Recombinant *P. carinii* DHFR. Trimethoprim has been reported to be a weak inhibitor of *P. carinii* DHFR activity in crude homogenates of organisms from rat lung (8). With the purified recombinant enzyme, trimethoprim is a very poor inhibitor of the enzyme. At a dihydrofolate concentration of 25 μM, the IC_{50} value for trimethoprim was 20,000 nM for the *P. carinii* enzyme. For comparison, the IC_{50} for trimethoprim with human DHFR was 2000 nM and with *E. coli* DHFR was 2 nM. A nonselective inhibitor of DHFR, methotrexate, shows IC_{50} values of ≈0.1 nM for all three DHFRs.

SUMMARY

P. carinii DHFR cDNA sequences have been isolated by direct expression in *E. coli*. The open reading frame of this gene directs the synthesis of a 23,868-Da protein with identifiable similarities to known DHFRs. The *P. carinii* DHFR gene contains an intron of 43 bp. The DNA sequences are A + T rich and show a marked bias toward codons containing these residues in the third position. *P. carinii* DHFR is not found as a bifunctional enzyme with thymidylate synthase. Chromosomal localization demonstrates that DHFR is not genetically linked to thymidylate synthase. *P. carinii* DHFR has been expressed in large quantities in *E. coli*, which provides a source of pure enzyme for future studies. Furthermore, DNA sequence analysis shows that human *P. carinii* DHFR is identical to the enzyme from rat *P. carinii*.

The marked sequence differences between microbial and human DHFRs have made this enzyme an important target for antimicrobial chemotherapy (35), as exemplified by the widely used antimicrobials, trimethoprim and pyrimethamine. Trimethoprim in combination with sulfamethoxazole is one of the major therapeutic and prophylactic regimens for *P. carinii* pneumonia (4, 7). Surprisingly, the trimethoprim component is a very poor inhibitor of *P. carinii* DHFR, showing an IC_{50} value 10,000-fold higher than with the *E. coli* enzyme. Moreover, the drug inhibits the host DHFR more potently than the target *P. carinii* enzyme. Several speculations arise from this finding. (*i*) The sulfa component of the trimethoprim/sulfamethoxazole combination clearly plays a major role in the efficacy towards *P. carinii*. This fact is supported by the lack of efficacy of trimethoprim alone in the treatment of the rat model of *P. carinii* infection (36) and the

known synergism of trimethoprim and the sulfa drugs (35). (*ii*) Trimethoprim is probably not the optimal DHFR inhibitor to be used in the treatment of *P. carinii*. (*iii*) With the availability of large amounts of recombinant *P. carinii* DHFR, surpassing the efficacy of trimethoprim should not be difficult. This could be accomplished in the short term simply by screening the available collections of antifolates and avoiding classes of compounds that inhibit the human enzyme. In the longer term, *P. carinii* DHFR represents a worthy and promising target for drug design based upon molecular structure.

We acknowledge the help of Patricia Nolan in performing the DHFR assays and Dr. W. Sirawaraporn for purification of the enzyme. This work was performed with the support of funds provided by the State of California and allocated on the recommendation of the University-wide Task Force on AIDS to J.C.E. U.E. was supported by the MacArthur Foundation Consortium on the Biology of Parasitic Diseases. B.L. was supported, in part, by the Danish Medical Research Council. M.C. is a scholar of the Cheng Scholars Program.

1. Mills, J. (1986) *Rev. Infect. Dis.* **8,** 1001–1011.
2. Kovacs, J. A., Hiemenz, J. W., Macher, A. M., Stover, D., Murray, H. W., Shelhammer, J., Lane, H. C., Urmacher, C., Honig, C., Longo, D. L., Parker, M. M., Natanson, C., Parillo, J. E., Fauci, A. S., Pizzo, P. A. & Masur, H. (1984) *Ann. Intern. Med.* **100,** 663–671.
3. Ivády, G., Páldy, L. & Unger, G. (1958) *Mschr. Kinderheilk.* **106,** 10–14.
4. Hughes, W. T., Feldman, S., Chaudhary, S. C., Ossi, M. J., Cox, F. & Sanyal, S. K. (1978) *J. Pediatr.* **92,** 285–291.
5. Frenkel, J. K., Good, J. T. & Schults, J. A. (1966) *Lab. Invest.* **15,** 1559–1577.
6. Edman, J. C., Kovacs, J. A., Masur, H., Santi, D. V., Elwood, H. J. & Sogin, M. L. (1988) *Nature (London)* **334,** 519–522.
7. Kovacs, J. A. & Masur, H. (1988) *J. Infect. Dis.* **158,** 254–259.
8. Allegra, C. J., Kovacs, J. A., Drake, J. C., Swan, J. C., Chabner, B. A. & Masur, H. (1987) *J. Exp. Med.* **165,** 926–931.
9. Burchall, J. J. (1983) in *Inhibition of Folate Metabolism in Chemotherapy*, ed. Hitchings, G. H. (Springer, Berlin), pp. 55–74.
10. Kovacs, J. A., Allegra, C. J., Swan, J. C., Drake, J. C., Parrillo, J. E., Chabner, B. A. & Masur, H. (1988) *Antimicrob. Agents Chemother.* **32,** 430–433.
11. Singer, S., Ferone, R., Walton, L. & Elwell, L. (1985) *J. Bacteriol.* **164,** 470–472.
12. Kovacs, J. A., Halpern, J. L., Swan, J. C., Moss, J., Parillo, J. E. & Masur, H. (1988) *J. Immunol.* **140,** 2023–2031.
13. Edman, U., Edman, J. C., Lundgren, B. & Santi, D. V. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 6503–6507.
14. Short, J. M., Fernandez, J. M., Sorge, J. A. & Huse, W. D. (1988) *Nucleic Acids Res.* **16,** 7583–7600.
15. Schwartz, D. C. & Cantor, C. R. (1984) *Cell* **37,** 67–75.
16. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988) *Science* **239,** 487–491.
17. Davisson, V. J., Sirawaraporn, W. & Santi, D. V. (1989) *J. Biol. Chem.* **264,** 9145–9148.
18. Grumont, R., Sirawaraporn, W. & Santi, D. V. (1988) *Biochemistry* **27,** 3776–3784.
19. Blakeley, R. L. (1984) in *Folates and Pterins*, eds. Blakeley, R. L. & Benkovic, S. J. (Wiley, Toronto), Vol. 1, pp. 191–253.
20. Barclay, B. J., Huang, T., Nagel, M. G., Misener, V. L., Game, J. C. & Wahl, G. M. (1988) *Gene* **63,** 175–185.
21. Lagosky, P. A., Taylor, G. A. & Haynes, R. H. (1987) *Nucleic Acids Res.* **15,** 10355–10371.
22. Andrews, J., Clore, G. M., Davies, R. W., Gronenborn, A. M., Gronenborn, B., Kalderon, D., Papadopoulos, P. C., Schafer, S., Sims, P. F. G. & Stancombe, R. (1985) *Gene* **35,** 217–222.
23. Masters, J. N. & Attardi, G. (1983) *Gene* **21,** 59–63.
24. Fling, M. E., Kopf, J. & Richards, C. A. (1988) *Gene* **63,** 165–174.
25. Grumont, R., Washtien, W. L., Caput, D. & Santi, D. V. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 5387–5391.
26. Beverley, S. M., Ellenberger, T. E. & Cordingley, J. S. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 2584–2588.
27. Volz, K. W., Matthews, D. A., Alden, R. A., Freer, S. T., Hansch, C., Kaufman, B. T. & Kraut, J. (1982) *J. Biol. Chem.* **257,** 2528–2536.
28. Matthews, D. A., Alden, R. A., Bolin, J. T., Filman, D. J., Freer, S. T., Hamlin, R., Hol, W. G. J., Kisliuk, R. L., Pastore, E. J., Plante, L. T., Xuong, N. & Kraut, J. (1978) *J. Biol. Chem.* **253,** 6946–6954.
29. Dobson, M. J., Tuite, M. F., Roberts, N. A., Kingsman, A. J., Kingsman, S. M., Perkins, R. E., Conroy, S. C., Dunbar, B. & Fothergill, L. A. (1982) *Nucleic Acids Res.* **10,** 2625–2637.
30. Parker, R. & Patterson, B. (1987) in *Molecular Biology of RNA: New Perspectives*, eds. Inouye, M. & Dudock, B. S. (Academic, New York), pp. 133–148.
31. Garret, C. E., Coderre, J. A., Meek, T. D., Garvey, E. P., Claman, D. M., Beverley, S. M. & Santi, D. V. (1984) *Mol. Biochem. Parasitol.* **11,** 257–265.
32. Hughes, D. E., Shonekan, O. A. & Simpson, L. (1989) *Mol. Biochem. Parasitol.* **34,** 155–166.
33. Kovacs, J. A. & Masur, H. (1988) in *Parasitic Infections*, eds. Leech, J. H., Sande, M. A. & Root, R. K. (Churchill Livingstone, New York), pp. 177–193.
34. Fishman, J. A., Ullu, E., Armstrong, M. Y. K. & Richards, F. F. (1989) *J. Protozool.* **36,** 4S–5S.
35. Wormser, G. P. & Keusch, G. T. (1983) in *Inhibition of Folate Metabolism in Chemotherapy*, ed. Hitchings, G. H. (Springer, Berlin), pp. 1–8.
36. Hughes, W. T. & Smith, B. L. (1984) *Antimicrob. Agents Chemother.* **26,** 436–440.