INFECTION AND IMMUNITY, May 1994, p. 1513–1519
0019-9567/94/$04.00+0

Vol. 62, No. 5

Conserved Sequence Homology of Cysteine-Rich Regions in Genes Encoding Glycoprotein A in *Pneumocystis carinii* Derived from Different Host Species

TERRY W. WRIGHT,[1] PATRICIA J. SIMPSON-HAIDARIS,[1,2,3]* FRANCIS GIGLIOTTI,[1,4] ALLEN G. HARMSEN,[5] AND CONSTANTINE G. HAIDARIS[1,6]

Departments of Microbiology and Immunology,[1] Medicine/Hematology Unit,[2] Pathology and Laboratory Medicine,[3] Pediatrics,[4] and Dental Research,[6] University of Rochester School of Medicine and Dentistry, Rochester, New York 14642, and Trudeau Institute, Saranac Lake, New York 12983[5]

Received 26 October 1993/Returned for modification 7 January 1994/Accepted 29 January 1994

***Pneumocystis carinii* surface glycoprotein A (gpA) exhibits host species-specific phenotypic and genotypic variation. Despite this heterogeneity, the gpAs of *P. carinii* isolated from different host species appear to be homologous molecules sharing certain biochemical and antigenic characteristics. Using two degenerate oligodeoxyribonucleotide primers corresponding to conserved cysteine regions from ferret and rat *P. carinii* gpAs, a PCR product of approximately 300 bp was amplified from ferret, rat, and SCID mouse *P. carinii*-infected lung genomic DNA. Northern (RNA) hybridization revealed a transcript of 3,450 nucleotides in *P. carinii*-infected SCID mouse lung mRNA, which is similar in size to the transcripts for ferret and rat *P. carinii* gpAs. Nucleotide sequence analysis of SCID mouse *P. carinii* gpA subclones derived from the PCR products identified two isoforms, which were 89% identical to each other in the amplified region and 73 and 54% identical to the rat- and ferret-derived *P. carinii* gpA genes, respectively. Comparison of the deduced amino acid sequences of mouse, ferret, and rat *P. carinii* gpAs revealed striking similarity in residues adjacent to and including the conserved cysteines. Furthermore, the spacing of two proline residues is invariant, and a potential N-linked glycosylation site is found at a similar position in all of the gpAs. Despite the heterogeneity observed in *P. carinii* gpAs, the conservation of cysteine residues and adjacent sequences implies similar secondary structure and, most likely, similar function for the gpAs of *P. carinii* isolated from different host species.**

Pneumocystis carinii is a major opportunistic pathogen of AIDS patients, as well as individuals suffering from a variety of other immunodeficiencies (16, 21). Infection of the susceptible host, if left uncontrolled, results in proliferation of organisms in lung alveoli, causing a life-threatening pneumonitis (PCP) (16). *P. carinii* has also been shown to infect a variety of mammalian hosts, including ferrets, rats, and mice. Although morphologically similar, immunologic and genetic analyses have indicated that *P. carinii* organisms isolated from different host species display antigenic (6, 11, 12, 17, 36, 37), karyotypic (15, 39), and genotypic (8, 29, 34) differences. Furthermore, recent work has demonstrated that *P. carinii* isolated from one host species is unable to survive and replicate after inoculation into a different immunodeficient host species (9).

A number of studies on *P. carinii* have focused on an abundant, immunodominant, mannosylated surface glycoprotein we refer to as gpA (6). gpA appears to play an important role in the *P. carinii*-host interaction. gpA may serve as an attachment ligand to lung cells by using either fibronectin (25) or lung surfactant protein A (41) as a bridge or by binding directly to the mannose-binding proteins present on the surface of alveolar macrophages (4). *P. carinii* gpA is also a target of both the humoral and cellular immune systems of the host. Antibody responses to gpA are observed after recovery from natural infection (38) and after immunization with recombinant gpA (14) or synthetic peptides deduced from the gpA cDNA sequence (18). Studies using purified gpA have shown a T-cell response to this molecule after both immunization and natural infection (5). Finally, passive immunization with a gpA-specific monoclonal antibody partially protects ferrets and rats against the development of PCP (10).

The gpAs of *P. carinii* isolated from different host species display a considerable degree of heterogeneity, including distinct monoclonal antibody-binding patterns and apparent molecular masses ranging from 95 to 140 kDa depending upon the host of origin (6, 7, 19, 20, 26, 30). Isolation of the bulk of the cDNA encoding ferret *P. carinii* gpA (14) allowed us to determine that this host species-specific heterogeneity extends to the genotypic level. Probes derived from the ferret *P. carinii* gpA cDNA sequence were unable to hybridize to the gpA genes of *P. carinii* isolated from other host species, even under conditions of low stringency (6). Recently, portions of the cDNAs encoding several rat *P. carinii* gpAs were isolated and their nucleotide sequences were determined (18, 33). Comparison of the cDNA sequences from ferret- and rat-derived *P. carinii* gpAs confirmed that these genes were not identical. The deduced amino acid sequences of ferret and rat *P. carinii* gpAs are only 33% identical and 54% similar when allowing for conservative amino acid substitutions. However, these gpA cDNAs each encode glycoproteins which are cysteine rich, contain several N-linked glycosylation sites, and are derived from mRNA species of similar size (14, 18, 33). Furthermore, we have observed amino acid homology concentrated around several conserved cysteine residues (see Fig. 1), indicating similarities in structure, and possibly function, between these two gpA homologs. Therefore, we hypothesized that the DNA sequences encoding the cysteine residues and surrounding conserved regions of ferret and rat *P. carinii* gpAs may also be conserved among *P. carinii* organisms isolated from other host species. In the present study, we utilized the PCR to demon-

* Corresponding author. Mailing address: Department of Medicine/Hematology Unit, P.O. Box 610, 601 Elmwood Ave., Rochester, NY 14642. Phone: (716) 275-8267. Fax: (716) 473-4314.

strate that degenerate oligodeoxyribonucleotide primers derived from the ferret (14) and rat (18) *P. carinii* gpA cDNAs were able to amplify gpA gene sequences from *P. carinii*-infected ferret, rat, and severe combined immunodeficient (SCID) mouse lung genomic DNA.

MATERIALS AND METHODS

Isolation of genomic DNA from *P. carinii*-infected and healthy lungs. Male ferrets (500 to 750 g) were purchased from Marshall Farms (North Rose, N.Y.). Ferrets were given drinking water containing 4 mg of dexamethasone (Lypho Med Inc., Rosemont, Calif.) per liter and 500 mg of tetracycline (Butler, Columbus, Ohio) per liter ad libitum to induce PCP. The 4- to 11-week-old C.B-17 *scid/scid* mice were obtained from the Trudeau Animal Breeding Facility (Saranac Lake, N.Y.). *P. carinii*-infected and *P. carinii*-free SCID mouse colonies were derived as described previously (9). *P. carinii*-infected and healthy lungs from ferrets, rats (kindly provided by W. T. Hughes, St. Jude Children's Research Hospital, Memphis, Tenn.), and SCID mice were homogenized in 5 volumes of phosphate-buffered saline as previously described (6). The tissue homogenates were digested overnight with proteinase K–10% sodium dodecyl sulfate (SDS) at 50°C. After standard phenol-chloroform extraction, the DNA was precipitated in ethanol, dried, and resuspended in 10 mM Tris-HCl (pH 8.0)–1 mM EDTA (14). The DNA concentration was measured by UV spectroscopy at 260 nm.

PCR. Three pairs of oligonucleotide primers were used to analyze each of the healthy and *P. carinii*-infected lung genomic DNA preparations. Primers Cys-sense [01C-S; 5′-AGAG(T/C)AGC(G/C)TG(C/T)TA(T/C)AA(A/G)AA(A/G)GG-3′] and Cys-antisense [02C-AS; 5′-ACA(C/T)T(T/G)CTCCTTCAA(C/T)TCAACACA-3′] were derived from the sequences of two highly conserved cysteine-rich regions of the ferret (see reference 14, Fig. 1, positions 1438 to 1460 and 1743 to 1720, respectively) and rat (see reference 18, Fig. 2, positions 1424 to 1446 and 1723 to 1700, respectively) *P. carinii* gpA cDNAs. The expected PCR reaction products from the amplification of the ferret and rat *P. carinii* gpA genes with the Cys primer pair were 306 and 300 bp, respectively. Primers pAZ102-E (5′-GATGGCTGTTTCCAAGCCCA-3′) and pAZ102-H (5′-GTGTACGTTGCAAAGTACTC-3′) amplify a 350-bp reaction product from the mitochondrial rRNA genes of *P. carinii* derived from each host species used in this study (8, 9, 34, 35). The β-actin-sense (5′-GCTGTGCTATGTTGCCCTAGACTTCGAGC-3′) and β-actin-antisense (5′-CGTACTCCTGCTTGCTGATCCACATGTGC-3′) primer sequences were based on the DNA sequence of the rat β-actin gene (23) and were used as a control for the integrity of the healthy lung DNA samples. This primer pair generates a PCR product of approximately 433 bp. All primers were purchased from the Department of Pediatrics Synthetic Oligomer Facility at the University of Rochester.

PCR was performed using the hot start method and Gene Amp kit reagents (Perkin Elmer Cetus, Norwalk, Conn.). The PCR reaction conditions were 0.2 mM deoxynucleoside triphosphate (dNTPs), 50 mM KCl, 3 mM $MgCl_2$, 10 mM Tris-HCl (pH 8.3), 1 μM each primer, and 2.5 U of Amplitaq DNA polymerase. For each PCR reaction, 0.02 to 0.08 μg of template DNA was used. Samples were denatured at 94°C for 90 s, annealed at 60°C for 90 s, and extended at 72°C for 120 s for a total of 35 cycles. Amplified DNA products were analyzed by 1.8% agarose gel electrophoresis in TAE buffer (14).

Restriction digests. The Cys primer pair amplification products from *P. carinii*-infected ferret, rat, and SCID mouse lung genomic DNAs were purified by using the Magic PCR Preps kit (Promega, Madison, Wis.). Each purified DNA was digested with *Pst*I, *Bcl*I, and *Hin*fI (Life Technologies Industries, Gaithersburg, Md.). One microgram of DNA was digested with 10 U of enzyme for 2.5 h at 37°C. The samples were then analyzed by 2.0% agarose gel electrophoresis in TAE buffer.

Northern (RNA) blot analysis. Total RNA was isolated from healthy C.B-17 and *P. carinii*-infected C.B-17 SCID mouse lungs by using the RNA isolation kit from Stratagene (La Jolla, Calif.). Poly(A)$^+$ mRNA was selected by using oligo(dT)-cellulose column chromatography (Collaborative Research, Bedford, Mass.) and analyzed by Northern blotting as described previously (28). The SCID mouse *P. carinii* gpA PCR product was labeled with [α-^{32}P]dCTP (NEN, Boston, Mass.), using the random primer labeling kit (Life Technologies Industries). Hybridization was performed at 60°C with 0.5 M Na_2HPO_4 (pH 7.2)–1 mM EDTA–7% SDS according to Bio-Rad's instructions for hybridization to nucleic acids on Zetaprobe membranes (Bio-Rad Laboratories, Centreville, N.Y.). High-stringency washes were done at 60°C using two changes of 40 mM Na_2HPO_4 (pH 7.2)–1 mM EDTA–5% SDS and then three changes of 40 mM Na_2HPO_4 (pH 7.2)–1 mM EDTA–1% SDS. Following exposure to film, the blot was stripped of hybridized radioactive probe by three washes for 20 min each in 40 mM Na_2HPO_4 (pH 7.2)–1 mM EDTA–1% SDS at 95°C. The blot was then reprobed in identical fashion with an [α-^{32}P]dCTP-labeled γ-actin probe (13).

Nucleic acid analyses. Nucleotide sequencing of the ferret and rat *P. carinii* Cys primer pair PCR products was performed by using the *fmol* PCR sequencing kit (Promega) with the 01C-S and 02C-AS primers after purification of the DNA with Magic PCR Preps kits. The SCID mouse *P. carinii* Cys primer pair PCR products were filled in to form blunt ends, using dNTPs and T4 DNA polymerase (Life Technologies Industries), and subcloned into the *Sma*I site of the plasmid vector pGEM7Zf(−) (Promega). Several subclones containing the expected 300-bp insert were then sequenced using the M13 universal and reverse sequencing primers with the Sequenase DNA sequencing kit (United States Biochemicals, Cleveland, Ohio). Sequenase data were analyzed using MacVector from IBI-A Kodak Company (New Haven, Conn.) and the University of Wisconsin Genetics Computer Group (UWGCG) software package (3). Comparison of the DNA and deduced amino acid sequences of the SCID mouse *P. carinii* gpA gene fragments to known sequences was performed with the FASTA program (24).

Nucleotide sequence accession number. The gene sequences for pMgpA-4 (accession no. U05262) and pMgpA-10 (accession no. U05263) have been deposited with GenBank.

RESULTS

PCR amplification of a conserved cysteine region in *P. carinii* gpA genes. Comparison of the published deduced amino acid sequences of ferret and rat *P. carinii* gpAs (14, 18) revealed that there were several conserved cysteine residues, some with highly conserved adjacent amino acid residues (Fig. 1). Degenerate oligodeoxyribonucleotide primers were designed based on the DNA sequences of several of these conserved regions and were used in the PCR to determine whether the gpA genes of *P. carinii* derived from different host species could be amplified with a single primer pair. Using the conserved Cys primer pair, a DNA fragment of approximately 300 bp was amplified from *P. carinii*-infected ferret, rat, and SCID mouse lung genomic DNAs. The expected reaction

```
                  *                                   *       *
FPc gpA     RVACYKKGLEAAAMSLLESRMKGMLKPGP-KNDYEECQKELLKQCKEV 526
RPc gpA     •A••••••QDRMLNKFFQKEL••K•GHVRFYS•PKD•K•YVVEN•TKL 474

                        *       *                           *
FPc gpA     RNVSAEVFEMCLYPKDTCKSLSDDVSDKAHDLMFTSFQKRDHPSQEDC 574
RPc gpA     DK---KYLPR•••••EL•YG••N•IFLQSKE•SALLDDQ••F•LKK•• 519

                      *               *       *
FPc gpA     VELKEQCEALEADADWLRSPCETLRTHCKFLYLSEGLKHHLLDEGKGK 622
RPc gpA     •••••K•DE•SS•SLLNLEK•I••KRR•EYFRV•••FRNVF•EKKDDS 567

                       *       *               *       *       *
FPc gpA     LSNNETCIKELEEKCHSWHRKKNETYAFPCALRNESCELMVWRVEGHC 670
RPc gpA     •MTQDN•T•A•H••••QLY•RRKNSFTVS•••PE•T•SY••FHTSQD• 615
```

FIG. 1. Predicted amino acid sequence homology in a cysteine-rich region of ferret and rat *P. carinii* gpAs. Amino acids are numbered as in reference 14 for ferret *P. carinii* gpA and as in reference 18 for rat *P. carinii* gpA. Amino acid residues encoding conserved Cys primer pairs are underlined. Identical amino acid residues are denoted with dots; positions at which rat *P. carinii* gpA differs from ferret *P. carinii* gpA are shown; conserved cysteine residues are denoted by asterisks; and gaps in the sequence alignment are indicated by hyphens.

product of 300 bp was not amplified from the uninfected lung genomic DNAs of the three host species (Fig. 2A). However, a smaller fragment of approximately 250 bp was amplified only from the healthy ferret lung DNA sample (Fig. 2A, lane 2) with the conserved Cys primer pair. Restriction enzyme and Southern blot analyses demonstrated that the 250-bp fragment amplified from the healthy lung was nonspecific. It was not related to the ferret *P. carinii* gpA PCR product in that it was not digested with *Pst*I, nor did it hybridize to the ferret *P. carinii* gpA-specific cDNA clone, pPoPc3 (14), from which the Cys primer pair sequences were derived (data not shown). The conserved *P. carinii* mitochondrial rRNA DNA primers amplified the expected 350-bp reaction product from each of the *P. carinii*-infected lung DNA samples, confirming that *P. carinii* DNA was present (Fig. 2B). The β-actin primer pair amplified the expected 430-bp reaction product from each of the template DNA samples, demonstrating that DNA was present in the healthy lung DNA samples as well as in the *P. carinii*-infected lung DNA samples (Fig. 2C).

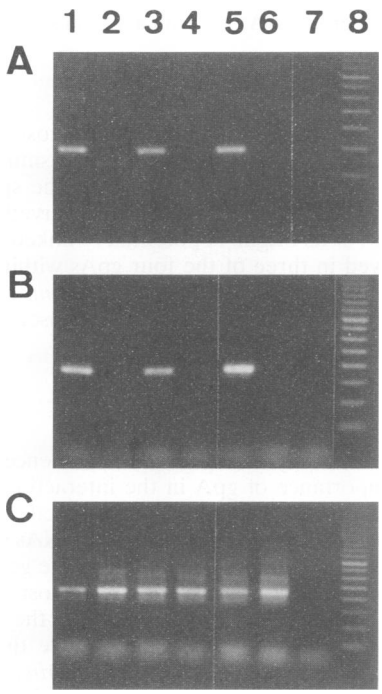

FIG. 2. PCR amplification of *P. carinii* DNA from different host species. (A) Reaction products were resolved on a 1.8% agarose gel. *P. carinii*-infected and healthy lung DNAs from the ferret (lanes 1 and 2), rat (lanes 3 and 4), and mouse (lanes 5 and 6) were amplified with the Cys primer pair designed from a conserved cysteine-rich region of the ferret and rat *P. carinii* gpAs. The expected product of approximately 300 bp was amplified from each of the *P. carinii*-infected lung DNAs but not from the healthy lung DNAs. Lane 7 shows the control in which no template was added. Molecular size standards are in 100-bp increments (lane 8). The identical DNA samples were also amplified with the pAZ102-E and -H (B) and the β-actin sense and antisense (C) primer pairs. The pAZ102-E and -H primers are conserved among all *P. carinii* organisms and amplified the expected product of approximately 350 bp from each of the *P. carinii*-infected lung DNAs. The β-actin primer pair amplified the predicted product of approximately 433 bp from each of the healthy and *P. carinii*-infected lung DNA samples.

Restriction enzyme mapping of the PCR products. Restriction digests of the PCR products were performed to examine the relationship of the gpA genes of *P. carinii* isolated from different host species. From the published DNA sequences, restriction sites were selected which distinguished between the ferret and rat *P. carinii* gpA PCR products. *Pst*I digested only ferret *P. carinii* gpA in the PCR-amplified region, generating the expected 270- and 36-bp DNA fragments (Fig. 3A, lane 1), while *Bcl*I digested only the rat *P. carinii* gpA PCR product, generating the expected 246- and 54-bp DNA fragments (Fig. 3B, lane 2). A battery of restriction enzymes was then tested, resulting in the identification of a third enzyme, *Hin*fI, which digested the SCID mouse *P. carinii* PCR product, resulting in unique DNA fragments of 200 and 100 bp (Fig. 3C, lane 3). These results suggested that the internal sequences of the PCR products amplified from each host species were distinct and did not result from PCR cross-contamination. Interestingly, none of the gpA PCR products were digested completely with the restriction enzymes used. This observation is consistent with published data indicating that more than one distinct gene encodes the *gpA* of *P. carinii* organisms which infect a single host species (18, 40).

Northern hybridization. To determine the size of the mouse *P. carinii* gpA transcript, Northern blot analysis was performed on healthy C.B-17 mouse lung and *P. carinii*-infected SCID C.B-17 mouse lung poly(A)$^+$ mRNA, using the SCID mouse *P. carinii* gpA PCR product as a probe. The labeled probe hybridized to a single mRNA species of approximately 3,450 nucleotides in length present only in the *P. carinii*-infected SCID mouse lung poly(A)$^+$ mRNA (Fig. 4A). Stripping and reprobing of the same blot with a γ-actin-specific probe

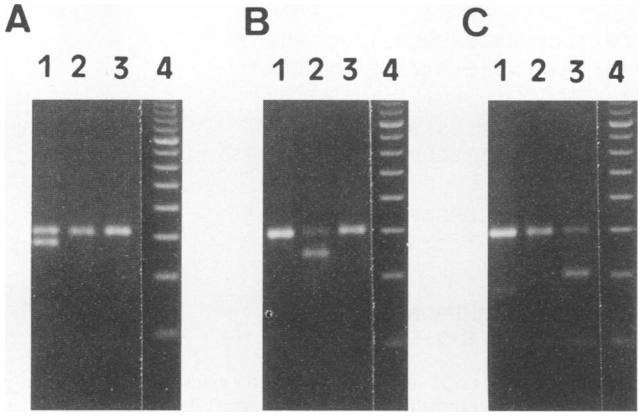

FIG. 3. Restriction enzyme analysis of the Cys primer pair amplification products of *P. carinii* DNA derived from different host species. The Cys primer pair amplification products from ferret (lanes 1), rat (lanes 2), and mouse (lanes 3) *P. carinii* were digested with *Pst*I (A), *Bcl*I (B), and *Hin*fI (C), demonstrating the nucleotide sequence diversity of these products. Molecular size standards are in 100-bp increments (lanes 4). The *Pst*I restriction site in the ferret *P. carinii* gpA PCR product and the *Bcl*I site in the rat *P. carinii* gpA PCR product were identified from the published cDNA sequences of these genes (14, 18). The *Hin*fI site in the mouse *P. carinii* PCR product was identified after testing a battery of restriction enzymes. Incomplete digestion of the amplification products appears to be the result of more than one gene encoding the gpA of *P. carinii* isolated from a single host species.

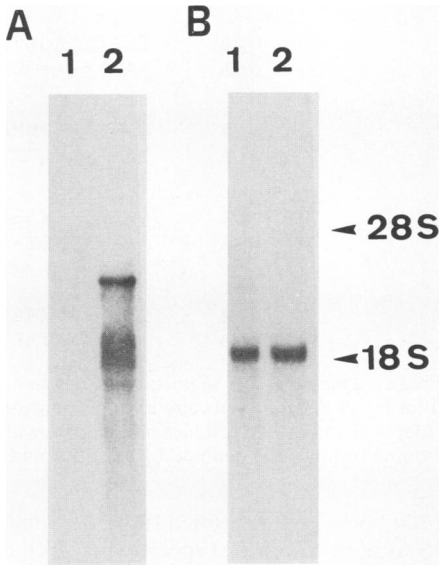

FIG. 4. Northern blot analysis of healthy and *P. carinii*-infected C.B-17 SCID mouse lung mRNAs. Lanes 1, 5 μg of healthy lung poly(A)$^+$ mRNA; lanes 2, 5 μg of infected lung poly(A)$^+$ mRNA. The mRNAs were probed with the SCID mouse-derived *P. carinii* Cys primer pair PCR product (A) and with a γ-actin-specific DNA probe (B).

demonstrated that an equal amount of mRNA was loaded in both lanes (Fig. 4B).

Sequence analysis of amplified *P. carinii* gpA genes. Direct nucleotide sequencing of the Cys primer pair PCR products from *P. carinii*-infected ferret and rat lung genomic DNAs confirmed that the sequences of these products corresponded to the respective published gpA sequences (data not shown). The gpA PCR product from the *P. carinii*-infected SCID mouse lung genomic DNA was subcloned into a plasmid vector for subsequent nucleotide sequence analysis. From the sequences of several clones, it was determined that at least two isoforms of the *P. carinii* gpA gene, represented by clones pMgpA-4 and pMgpA-10, were amplified from the infected SCID mouse lung DNA. Interestingly, only one of the clones, pMgpA-4, contained a *Hin*fI restriction site (Fig. 5). The presence of two gpA isoforms represented in the single PCR-amplified band, one with a *Hin*fI site and one without, explains, at least in part, why the PCR product was not completely digested with this enzyme. The DNA sequences of these two isoforms were 89% identical to each other, 73% identical to rat *P. carinii* gpA, and 54% identical to ferret *P. carinii* gpA in the PCR-amplified region.

Comparison of the putative amino acid sequences of the two mouse *P. carinii* gpA clones with the published deduced amino acid sequences of ferret and rat *P. carinii* gpA in the PCR-amplified region revealed that, although these sequences varied considerably, there were notable similarities (Fig. 6). The two mouse *P. carinii* gpA clones were approximately 80% identical to each other, 58% identical to rat *P. carinii* gpA, and 35% identical to ferret *P. carinii* gpA over the PCR-amplified region. When allowance was made for conservative amino acid substitutions, the mouse *P. carinii* gpAs were 85% similar in amino acid sequence to each other and 74% similar to the amino acid sequence of the rat *P. carinii* gpA. When the mouse *P. carinii* gpAs were compared with the ferret *P. carinii* gpA, a 53% similarity in sequence was observed across the amplified region (Table 1). Furthermore, several other similarities in the gpA amino acid sequences were notable. The spacing of the cysteines within the amplified region is conserved for *P. carinii* gpA regardless of its origin. A potential N-linked glycosylation site is conserved in three of the four gpAs within this region, the exception being one of the mouse *P. carinii* gpAs, and the spacing of two proline residues is exactly conserved in all gpA protein sequences (Fig. 6).

DISCUSSION

A considerable body of experimental evidence has demonstrated the importance of gpA in the interaction of *P. carinii* with its host. Despite the cloning of gpA-specific cDNAs from ferret (14) and rat (18, 33) *P. carinii*, the cDNAs when used as probes have not been useful in identifying the genes encoding *P. carinii* gpA homologs from different hosts, presumably because of the heterogeneity observed in the gpA-specific cDNAs identified thus far. However, despite the differences seen in the DNA sequences of the *P. carinii* gpA-specific cDNAs (8, 14, 18, 33) and the antigenic polymorphism observed in the molecules isolated from different hosts (5–7, 11, 12, 17, 19, 20, 26, 30, 36), the gpA homologs derived from different hosts share several attributes. These include the surface location of gpA, its relative abundance in relationship to other *P. carinii* proteins, a similar overall size, and the presence of a conserved carbohydrate epitope (6, 12). These similarities suggest a conserved function for gpA in *P. carinii* biology and in the interaction of the organism with its host. The apparent conserved function of gpA prompted us to search for structural similarities in the deduced primary amino acid sequences of ferret and rat *P. carinii*-derived gpAs. Other proteins, such as vaccinia growth factor and epidermal growth factor, have considerable differences at the DNA and primary amino acid sequence levels but have maintained similar sec-

FIG. 5. Nucleotide sequence alignment of two mouse-derived *P. carinii* gpA clones. The sequences of the pMgpA-4 and pMgpA-10 clones were aligned by using the UWGCG gap program (3). The *Hin*fI restriction site present at nucleotides 99 to 103 in pMgpA-4, but not in pMgpA-10, is bracketed. The degenerate Cys primer pair sequences are underlined.

ondary structure, and subsequently similar function, through conserved cysteine residues (2, 27, 31). We proposed that this conservation of structure and function may also be seen in the gpAs of *P. carinii* derived from different hosts.

Examination of the published deduced amino acid sequences indicated a particularly striking conservation of the positioning of cysteine residues and a region of limited homology concentrated around them between ferret and rat *P. carinii* gpAs (Fig. 1). Using degenerate oligodeoxyribonucleotide primers based on the conserved cysteine regions of *P. carinii* gpAs derived from ferret and rat, we amplified the gpA genes of *P. carinii* isolated from various host species by PCR. The predicted 300-bp reaction product was amplified from the *P. carinii*-infected genomic lung DNAs of ferrets, rats, and SCID mice, using the Cys primer pair; however, the predicted amplification product was not observed in parallel reactions performed using genomic DNA purified from uninfected lungs obtained from these animals. Interestingly, the expected PCR reaction product was not observed when three different archival *P. carinii*-infected human lung DNA samples were amplified with the Cys primers under either high (60°C) or low (45°C) primer annealing temperatures (40). Kovacs et al. (18) also were unable to amplify human *P. carinii* gpA by using primers based on the rat *P. carinii* gpA DNA sequence. Cloning and sequencing of human *P. carinii* gpA DNA will be necessary in order to draw definitive conclusions regarding its interrelationship with gpA gene sequences of animal-derived *P. carinii.*

We also report initial nucleotide sequence data for mouse *P. carinii* gpA and provide further evidence for the presence of multiple genes encoding the gpA of *P. carinii* organisms which infect a single host species. Northern blot analysis demon-

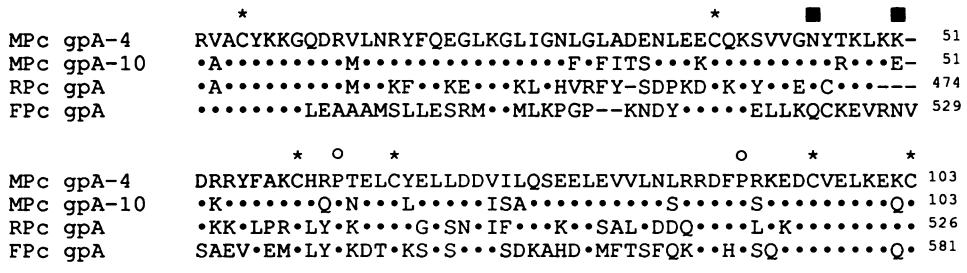

FIG. 6. Alignment of the predicted amino acid sequences of the gpAs of *P. carinii* isolated from three different host species. The deduced amino acid sequences of ferret-, rat-, and mouse-derived *P. carinii* gpAs in the cysteine-rich PCR-amplified region were aligned by using the UWGCG PileUp program (3). Identical amino acid residues are denoted by dots; positions in mouse gpA-10 (MPc gpA-10), rat gpA (RPc gpA), and ferret gpA (FPc gpA) which differ from mouse gpA-4 (MPc gpA-4) are shown; and gaps in the sequence alignment are indicated by hyphens. Conserved amino acid residues are highlighted above the sequence as follows: cysteines (asterisks), prolines (open circles), and potential N-linked glycosylation sites (closed squares).

TABLE 1. Comparison of similarity in the amino acid sequences corresponding to a conserved cysteine region of *P. carinii* gpAs from the ferret (F), mouse (M), and rat (R)

Sequence	% similarity[a]		
	FPc gpA	MPc gpA-4	MPc gpA-10
MPc gpA-4	52		
MPc gpA-10	53	85	
RPc gpA	52	79	74

[a] Percent similarity based on the total number of conservative amino acid substitutions plus the number of identical amino acids divided by the total number of residues.

strated that the mouse *P. carinii* gpA PCR product hybridized to a single mRNA band of approximately 3,450 nucleotides in length, which was similar in size to the mRNA transcripts encoding ferret and rat *P. carinii* gpAs (14, 18). Restriction enzyme digests of the PCR products obtained by amplification using the Cys primers demonstrated, as expected, that the gpA genes of *P. carinii* derived from different host species were distinct. Interestingly, none of the PCR products were completely digested by the enzymes. To examine the possibility that we had amplified more than one gpA gene from mouse *P. carinii* by using the Cys primers, the PCR products were subcloned for further analysis. DNA sequence analysis of several subclones revealed that at least two isoforms of the gpA gene existed in SCID mouse *P. carinii* (Fig. 5). One isoform contained a *Hin*fI restriction site, while the other isoform did not, explaining why the PCR product was not completely digested with *Hin*fI. This finding is consistent with data obtained previously using the ferret and rat models of PCP, which demonstrated that more than one isoform of the gpA gene exists in *P. carinii* organisms which infect a single host species (18, 33, 40). Nonetheless, comparison of the two mouse *P. carinii* gpA clones revealed a significant degree of identity in their deduced amino acid sequences which was comparable to the degree of identity observed among the gpA isoforms found in ferret and rat *P. carinii* (18, 33, 40). Thus, the Cys primer sequences used in the PCR were not only conserved among the gpA genes of *P. carinii* from different host species but also among the gpA genes of the *P. carinii* organisms which infect a single host species.

Alignment of the deduced amino acid sequences of the two mouse *P. carinii* gpA clones with the published amino acid sequences of ferret and rat *P. carinii* gpAs illustrates their interrelationship (Fig. 6). The overall similarity in the deduced amino acid sequences in the PCR-amplified region of *P. carinii* gpAs isolated from different hosts (Table 1) indicates a stronger homology than one would detect from DNA sequence comparison alone. Several residues are conserved in all of the gpA amino acid sequences examined, while numerous conservative amino acid substitutions are also present. In addition to the conserved cysteines found within the primer sequences (positions 4, 96, and 103), the cysteine residues at positions 38, 59, and 66 are conserved as well (Fig. 6, mouse *P. carinii* gpA numbering). Furthermore, the spacing of two proline residues at positions 62 and 91 (Fig. 6, mouse *P. carinii* gpA numbering) is invariant in all of the gpAs, and the potential N-linked glycosylation site at position 470 in the rat gpA and 45 in the mouse *P. carinii* gpA is found at position 528 in the ferret *P. carinii* gpA (Fig. 6). Conservation of these residues implies a similar secondary structure, and possibly a similar function, for the gpA molecules of *P. carinii* organisms which infect different host species. It was also evident that the amino acid sequence of mouse *P. carinii* gpA was more similar to rat *P. carinii* gpA than to ferret *P. carinii* gpA (Fig. 6 and Table 1). These data are consistent with the report that rat and mouse *P. carinii* are antigenically similar (37) and suggests that rat and mouse *P. carinii* are more closely related to each other than to ferret *P. carinii*.

The existence of multiple genes encoding a surface antigen has been noted for other eukaryotic microorganisms, including members of the protozoan genera *Trypanosoma* and *Giardia* (1, 22). Interestingly, although *P. carinii* has been grouped taxonomically with the fungi on the basis of the homology of *P. carinii* housekeeping genes with those found in fungi (34), the presence of multiple genes encoding a polymorphic surface antigen seems to be more characteristic of protozoa than fungi. Although it appears that a population of *P. carinii* organisms infecting a single animal host possesses multiple gpA genes, the question of whether a single *P. carinii* organism is capable of simultaneously expressing multiple genes for gpA remains unanswered. Alternatively, multiple strains of *P. carinii* each expressing only a single gpA gene at a given time may infect a single host. Furthermore, the role of multiple gpA genes in the biology of *P. carinii* and their function in the pathogenesis of PCP is also unknown. It has been speculated that the polymorphism observed in different *P. carinii* gpA isoforms within the same host allows the organism to subvert the immune response (18, 33) or, perhaps, produce a more pathogenic strain (32). However, if either mechanism was a successful means of escaping immune recognition, one would expect *P. carinii* to be more overt in its pathogenicity rather than an opportunist capable only of infecting severely immunocompromised hosts. More needs to be understood about the biology of *P. carinii* and the basis of acquired resistance to *P. carinii* infection before the functional basis of gpA polymorphism is to be fully appreciated. The development of a continuous culture system for clonal populations of *P. carinii*, coupled with the ability to generate and analyze mutants, will be required to achieve an understanding of the functional need for structural polymorphism in *P. carinii* gpA.

ACKNOWLEDGMENTS

This work was supported in part by grants HL-49610, HL30616, AI-23302, AI-28354, and AI-07362 from the National Institutes of Health, Bethesda, Md.

REFERENCES

1. **Borst, P.** 1991. Molecular genetics of antigenic variation. Immunol. Today **12**:A29–A33.
2. **Brown, J. P., D. R. Twardzik, H. Marquardt, and G. J. Todaro.** 1985. Vaccinia virus encodes a polypeptide homologous to epidermal growth factor and transforming growth factor. Nature (London) **313**:491–492.
3. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
4. **Ezekowitz, R. A. B., D. J. Williams, H. Koziel, M. Y. K. Armstrong, A. Warner, F. F. Richards, and R. M. Rose.** 1991. Uptake of *Pneumocystis carinii* mediated by the macrophage mannose receptor. Nature (London) **351**:155–158.
5. **Fisher, D. J., F. Gigliotti, M. Zauderer, and A. G. Harmsen.** 1991. Specific T-cell response to a *Pneumocystis carinii* surface glycoprotein (gp120) after immunization and natural infection. Infect. Immun. **59**:3372–3376.
6. **Gigliotti, F.** 1992. Host species-specific antigenic variation of a mannosylated surface glycoprotein of *Pneumocystis carinii*. J. Infect. Dis. **165**:329–336.
7. **Gigliotti, F., L. R. Ballou, W. T. Hughes, and B. D. Mosley.** 1988. Purification and initial characterization of a ferret *Pneumocystis*

carinii surface antigen. J. Infect. Dis. **158:**848–854.
8. **Gigliotti, F., P. J. Haidaris, C. G. Haidaris, T. W. Wright, and K. R. Van Der Meid.** 1993. Further evidence of host species specific variation in antigens of *Pneumocystis carinii* using the polymerase chain reaction. J. Infect. Dis. **168:**191–194.
9. **Gigliotti, F., A. G. Harmsen, C. G. Haidaris, and P. J. Haidaris.** 1993. *Pneumocystis carinii* is not universally transmissible between mammalian species. Infect. Immun. **61:**2886–2890.
10. **Gigliotti, F., and W. T. Hughes.** 1988. Passive immunoprophylaxis with specific monoclonal antibody confers partial protection against *Pneumocystis carinii* pneumonitis in animal models. J. Clin. Invest. **81:**1666–1668.
11. **Gigliotti, F., D. C. Stokes, A. B. Cheatham, D. S. Davis, and W. T. Hughes.** 1986. Development of murine monoclonal antibodies to *Pneumocystis carinii*. J. Infect. Dis. **154:**315–322.
12. **Graves, D. C., S. J. N. McNabb, M. A. Worley, T. D. Downs, and M. H. Ivey.** 1986. Analyses of rat *Pneumocystis carinii* antigens recognized by human and rat antibodies by using Western immunoblotting. Infect. Immun. **54:**96–103.
13. **Gunning, P., P. Ponte, H. Okayama, J. Engel, H. Blau, and L. Kedes.** 1983. Isolation and characterization of full-length cDNA clones for human α-, β-, and γ-actin mRNAs: skeletal but not cytoplasmic actins have an amino-terminal cysteine that is subsequently removed. Mol. Cell. Biol. **3:**787–795.
14. **Haidaris, P. J., T. W. Wright, F. Gigliotti, and C. G. Haidaris.** 1992. Expression and characterization of a cDNA clone encoding an immunodominant surface glycoprotein of *Pneumocystis carinii*. J. Infect. Dis. **166:**1113–1123.
15. **Hong, S. T., P. E. Steele, M. T. Cushion, P. D. Walzer, S. L. Stringer, and J. R. Stringer.** 1990. *Pneumocystis carinii* karyotypes. J. Clin. Microbiol. **28:**1785–1795.
16. **Hughes, W. T.** 1987. *Pneumocystis carinii* pneumonitis, vol. 1 and 2. CRC Press, Inc., Boca Raton, Fla.
17. **Kovacs, J. A., J. L. Halpern, B. Lundgren, J. C. Swan, J. E. Parrillo, and H. Masur.** 1989. Monoclonal antibodies to *Pneumocystis carinii*: identification of specific antigens and characterization of antigenic differences between rat and human isolates. J. Infect. Dis. **159:**60–70.
18. **Kovacs, J. A., F. Powell, J. C. Edman, B. Lundgren, A. Martinez, B. Drew, and C. W. Angus.** 1993. Multiple genes encode the major surface glycoprotein of *Pneumocystis carinii*. J. Biol. Chem. **268:**6034–6040.
19. **Linke, M. J., M. T. Cushion, and P. D. Walzer.** 1989. Properties of the major antigens of rat and human *Pneumocystis carinii*. Infect. Immun. **57:**1547–1555.
20. **Lundgren, B., G. Y. Lipschik, and J. A. Kovacs.** 1991. Purification and characterization of a major human *Pneumocystis carinii* surface antigen. J. Clin. Invest. **87:**163–170.
21. **Murray, J. F., S. M. Garay, P. C. Hopewell, J. Mills, G. L. Snider, and D. E. Stover.** 1987. Pulmonary complications of the acquired immunodeficiency syndrome: an update. Am. Rev. Respir. Dis. **135:**504–509.
22. **Nash, T. E., and M. R. Mowatt.** 1992. Characterization of a *Giardia lamblia* variant-specific surface protein (VSP) gene from isolate GS/M and estimation of the VSP gene repertoire size. Mol. Biochem. Parasitol. **51:**219–227.
23. **Nundel, U., R. Zakut, M. Shanie, S. Newman, Z. Levy, and D. Yaffe.** 1983. The nucleotide sequence of the rat cytoplasmic β-actin gene. Nucleic Acids Res. **11:**1759–1771.
24. **Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA **85:**2444–2448.
25. **Pottratz, S. T., J. Paulsrud, J. S. Smith, and W. J. Martin II.** 1991. *Pneumocystis carinii* attachment to cultured lung cells by *Pneumocystis* gp120, a fibronectin binding protein. J. Clin. Invest. **88:**403–407.
26. **Radding, J. A., M. Y. K. Armstrong, E. Ullu, and F. F. Richards.** 1989. Identification and isolation of a major cell surface glycoprotein of *Pneumocystis carinii*. Infect. Immun. **57:**2149–2157.
27. **Reisner, A. H.** 1985. Similarity between the vaccinia virus 19K early protein and epidermal growth factor. Nature (London) **313:**801–803.
28. **Simpson, P. J.** 1987. An improved method for messenger RNA isolation and characterization of in vitro translation products by Western blotting. Gene **56:**161–171.
29. **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.
30. **Tanabe, K., S. Takasaki, J. Watanabe, A. Kobata, K. Egawa, and Y. Nakamura.** 1989. Glycoproteins composed of major surface immunodeterminants of *Pneumocystis carinii*. Infect. Immun. **57:**1363–1368.
31. **Twardzik, D. R., J. P. Brown, J. E. Ranchalis, G. J. Todaro, and B. Moss.** 1985. Vaccinia virus-infected cells release a novel polypeptide functionally related to transforming and epidermal growth factors. Proc. Natl. Acad. Sci. USA **82:**5300–5304.
32. **Vogel, P., C. J. Miller, L. L. Lowenstine, and A. A. Lackner.** 1993. Evidence of horizontal transmission of *Pneumocystis carinii* pneumonia in simian virus-infected rhesus macaques. J. Infect. Dis. **168:**836–843.
33. **Wada, M., K. Kitada, M. Saito, K. Egawa, and Y. Nakamura.** 1993. cDNA sequence diversity and genomic clusters of major surface glycoprotein of *Pneumocystis carinii*. J. Infect. Dis. **168:**979–985.
34. **Wakefield, A. E., S. E. Peters, S. Banerji, P. D. Bridge, G. S. Hall, D. L. Hawksworth, L. A. Guiver, A. G. Allen, and J. M. Hopkin.** 1992. *Pneumocystis carinii* shows DNA homology with the ustomycetous red yeast fungi. Mol. Microbiol. **6:**1903–1911.
35. **Wakefield, A. E., F. J. Pixley, S. Banerji, K. Sinclair, R. F. Miller, E. R. Moxon, and J. M. Hopkin.** 1990. Detection of *Pneumocystis carinii* with DNA amplification. Lancet **336:**451–453.
36. **Walzer, P. D., and M. J. Linke.** 1987. A comparison of the antigenic characteristics of rat and human *Pneumocystis carinii* by immunoblotting. J. Immunol. **138:**2257–2265.
37. **Walzer, P. D., and M. E. Rutledge.** 1980. Comparison of rat, mouse, and human *Pneumocystis carinii* by immunofluorescence. J. Infect. Dis. **142:**449.
38. **Walzer, P. D., D. Stanforth, M. J. Linke, and M. T. Cushion.** 1987. *Pneumocystis carinii*: immunoblotting and immunofluorescent analyses of serum antibodies during experimental rat infection and recovery. Exp. Parasitol. **63:**319–328.
39. **Weinberg, G. A., and M. S. Bartlett.** 1991. Comparison of pulsed field gel electrophoresis karyotypes of *Pneumocystis carinii* derived from rat lung, cell culture, and ferret lung. J. Protozool. **38:**64S–65S.
40. **Wright, T. W., T. Y. Bissoondial, F. Gigliotti, C. G. Haidaris, and P. J. Simpson-Haidaris.** Unpublished observation.
41. **Zimmerman, P. E., D. R. Voelker, F. X. McCormack, J. R. Paulsrud, and W. J. Martin II.** 1992. 120-kD surface glycoprotein of *Pneumocystis carinii* is a ligand for surfactant protein A. J. Clin. Invest. **89:**143–149.