Characterization of variants of the gene encoding the p55 antigen in *Pneumocystis* from rats and mice

Liang Ma,† Geetha Kutty, Qiuyao Jia† and Joseph A. Kovacs

Critical Care Medicine Department, Warren Grant Magnuson Clinical Center, National Institutes of Health, Bethesda, MD 20892, USA

Variants of the p55 gene in rat-derived *Pneumocystis carinii* have been identified and its counterpart in mouse-derived *P. carinii* f. sp. muris has been cloned. By PCR amplification of *P. carinii* genomic DNA, five variants were identified that differed from each other in size and sequence, primarily in the number and size of encoded amino acid repeats. For *P. carinii* f. sp. *muris*, a single PCR fragment (471 bp) was obtained, which contained an incomplete ORF encoding a 157 aa protein that was most similar to a p55 variant in *P. carinii*, with nucleotide and amino acid sequence identity of 79 and 68 %, respectively. Southern blot analysis revealed the presence of more than one copy of the p55 gene in both *Pneumocystis* species. Thus, like other *Pneumocystis* antigens, p55 exhibits polymorphism that could potentially benefit the organism in host interactions.

**INTRODUCTION**

*Pneumocystis* is an important opportunistic pathogen that causes severe pneumonia in patients with AIDS and other immunocompromised patients. Two major groups of antigens have been identified in this organism by immunoblotting. One group, which is present as a broad band of 90–120 kDa, has been well-characterized as the major surface glycoprotein (MSG) in *Pneumocystis* organisms derived from rats, humans, mice, ferrets and rabbits (Graves et al., 1986; Walzer & Linke, 1987; Gigliotti et al., 1988; Kovacs et al., 1989, 1993; Bauer et al., 1993; Kutty et al., 2001). The second group, which appears as a broad band of 45–55 kDa in *P. carinii* and 35–45 kDa in human-derived *Pneumocystis jiroveci* (Stringer et al., 2002), is less well-studied (Walzer & Linke, 1987; Smulian et al., 1992, 1993, 2000). The gene that encodes a 45–55 kDa protein of *P. carinii*, designated p55, has been cloned (Smulian et al., 1992, 1993). This gene is characterized by the presence of 10 repeats of a heptapeptide motif that is rich in glutamic acids near the C-terminus. Southern blot analysis has shown that this gene is present as a single copy. Unlike MSG, which resides on the outer surface of *P. carinii*, p55 is located within the cell wall (Broomall et al., 1998). Currently, the role of p55 in the biology and pathogenesis of *Pneumocystis* infection remains unclear. Theus et al. (1994) demonstrated a cell-mediated immune response to p55 in animals exposed previously to *P. carinii*. A more recent study has shown that immunization with recombinant *P. carinii* p55 can afford partial protection against infection (Smulian et al., 2000).

Whilst screening a *P. carinii* cDNA library, we identified a DNA fragment that was highly homologous to the *P. carinii* p55 gene but lacked the heptapeptide repeats. We undertook further to characterize the p55 gene in *P. carinii* and to clone the p55 gene in mouse-derived *P. carinii* f. sp. *muris*.

**METHODS**

*Pneumocystis* organisms and DNA/RNA extraction. *P. carinii* organisms were isolated from the lungs of immunosuppressed rats, provided under contract by Indiana University (Indianapolis, IN, USA). *P. carinii* f. sp. *muris* organisms were isolated from the lungs of severe combined immune deficient (SCID) mice in our laboratory. Genomic DNA was extracted by treatment with proteinase K followed by phenol/chloroform extraction, as described previously (Ma et al., 1999). Total RNA was extracted by using a RNAzol B Isolation kit (Tel-Test). Experimentation guidelines of the US Department of Health and Human Services and the NIH were followed in the conduct of this research.

PCR and DNA sequencing. Based on regions conserved between the previously reported p55 sequence (Smulian et al., 1992, 1993) and a DNA fragment that we identified by sequencing randomly picked clones from a *P. carinii* cDNA library (Edman et al., 1989) (GenBank accession no. AF494450), we designed the primer pair PU947 and PD1637 (Table 1). PCR mixture (50 µl) contained 100 ng genomic DNA, 0.5 µM each primer, 0.2 mM dNTPs, 1× PCR buffer (10 mM Tris/HC1, pH 8.3; 50 mM KCl; 2.5 mM MgCl2) and 2.5 U AmpliTaq Gold DNA polymerase (PerkinElmer). Amplification was performed in a PTC-100 programmable thermal controller (MJ Research) with a touchdown cycling protocol as follows: 95 °C for 8 min, then 10 cycles of 1 min at 95 °C, 2 min at 65 °C (with a decrease by 1 °C every cycle to reach 50 °C in the last cycle) and 2 min at 72 °C, followed by 30 cycles of 1 min

**Abbreviation:** MSG, major surface glycoprotein.

The GenBank/EMBL/DDBJ accession numbers for the sequences of *Pneumocystis* p55 variant genes are AF494448–AF494452.
Table 1. Oligonucleotide primers and probes used in this study

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Sequence (5′→3′)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PU947</td>
<td>GCAAGACATTCTTAGCAGAATCGTG</td>
<td>Corresponding to positions 280–304 of p55 variant v3 (GenBank accession no. AF494450)</td>
</tr>
<tr>
<td>PD1637</td>
<td>ACTGCTAATGTCATTCTATAATTGCAC</td>
<td>Complementary to positions 769–796 of p55 variant v3</td>
</tr>
<tr>
<td>cp</td>
<td>GCAAGACATTCTTAGCAGAATCGTG</td>
<td>Corresponding to positions 280–316 of p55 variant v3</td>
</tr>
<tr>
<td>v0p</td>
<td>GATGAAAAGGAAAGCCTACTGTAGAGGGGGAAAG</td>
<td>Corresponding to positions 382–417 of p55 (M77242)</td>
</tr>
<tr>
<td>c1r</td>
<td>TCAGCTTCTTTTACAGATCTGCTAAAATGTCTTTC</td>
<td>Complementary to positions 14–62 of p55</td>
</tr>
<tr>
<td>v01r</td>
<td>TCAGGAGGAGGAGAAGTAGGACGACAGAAGACCTGACAC</td>
<td>Complementary to positions 525–577</td>
</tr>
<tr>
<td>v23r</td>
<td>TCAGGAGGAGGAGAAGTAGGACGACAGAAGACCTGACAC</td>
<td>Complementary to positions 525–577 of p55 variant v3</td>
</tr>
<tr>
<td>M39</td>
<td>GAGAGCTGCTCAAAAAATGTAGATTGTTCTACCTTTACA</td>
<td>Corresponding to positions 14–52 of p55 (AF494452)</td>
</tr>
<tr>
<td>M472</td>
<td>TCTACACGAACACCGTGACCTT</td>
<td>Complementary to 426–447 of p55 (AF494452)</td>
</tr>
</tbody>
</table>

at 95 °C, 2 min at 50 °C and 2 min at 72 °C. Each experiment included a negative control without template DNA. PCR products were sequenced either directly or after subcloning into vector pCR 2.1 as described previously (Ma et al., 1999). Sequence analysis was performed by using MacVector 7.0 software (Oxford Molecular). Nucleotide sequences obtained in this study are available from GenBank under accession numbers AF494448–AF494451 for P. carinii p55 variant genes and AF494452 for the P. carinii f. sp. muris p55 gene.

Southern blot analysis with P. carinii DNA. Based on the sequence alignment of the five p55 gene variants in P. carinii (Fig. 1), we designed conserved and variant oligonucleotide probes, cp and v0p (Table 1). Probes were labelled with digoxigenin-dUTP by using the DIG Oligonucleotide Tailing kit (Boehringer Mannheim). Genomic DNA samples (2·5 μg) were digested with SplI, XhoI or EcoRI. These enzymes recognize AT-rich sequences and have been found to cut Pneumocystis genomic DNA more frequently than other enzymes (Ma et al., 1999).

Fig. 1. Alignment of deduced amino acid sequences of p55 variants in P. carinii (rPC) and P. carinii f. sp. muris (mPC). Identical residues are shaded and boxed. The sequence of variant v0 is identical to that reported previously (Smulian et al., 1992). Lines above the sequences indicate heptapeptide repeats and lines below the sequences indicate hexapeptide repeats; the regions from which the sequences of probes cp and v0p were derived for use in Southern blotting are shown. GenBank accession numbers for the sequences of variants v1–v4 are AF494448–AF494451.
et al., 1999; Kutty et al., 2001). No restriction sites for these enzymes were present in the sequences of probes used in this study. Digested DNA was separated by 1% agarose gel electrophoresis and transferred to a Nytran membrane (Schleicher & Schuell). Blots were hybridized with probes overnight at 42 °C by using DIG Easy Hyb hybridization solution (Boehringer Mannheim), washed twice in 2× SSC and 0.1% SDS at room temperature for 5 min each, then twice in 1× SSC and 0.1% SDS at 42 °C for 15 min each. Hybridization signal was detected by using a DIG Luminescent Detection kit (Boehringer Mannheim). Blots were stripped by incubating twice for 30 min each at 80 °C in stripping buffer that contained 50% deionized formamide, 5% SDS and 50 mM Tris/HCl (pH 7.5) before rehybridizing with a different probe.

**Southern blot analysis with *P. carinii* f. sp. *muris* DNA.** A DNA fragment was amplified from *P. carinii* f. sp. *muris* genomic DNA by PCR with primers M39 and M472 (Table 1) and a PCR DIG Probe Synthesis kit (Boehringer Mannheim). Genomic DNA samples (9.4 μg) were digested with EcoRV, HindIII, HpaI or XbaI. Subsequent procedures were done under the same conditions as described above, except that high-stringency washing was performed in 0.16× SSC and 0.1% SDS at 65 °C for 15 min each.

**Northern blot analysis with *P. carinii* RNA.** Total RNA was separated by 1.2% agarose gel electrophoresis in the presence of formaldehyde, transferred to a Nytran membrane (Schleicher & Schuell) and hybridized at 42 °C overnight with probes c1r, v01r or v23r, which were designed from conserved regions or specific variants. Washing and detection conditions were the same as for Southern blotting with *P. carinii* DNA.

**RESULTS AND DISCUSSION**

**p55 in *P. carinii***

Whilst screening a *P. carinii* cDNA library, we identified a 906 bp fragment (GenBank accession no. AF494450), which contained a single incomplete ORF encoding a 269 aa protein that was homologous to the previously reported *P. carinii* p55 gene and protein (Smulian et al., 1992, 1993), with respectively 76 and 60% sequence identity in overlapping regions. Given that MSG and another *P. carinii* surface protein, PRT-1, are encoded by multi-copy gene families (Kovacs et al., 1993; Lugli et al., 1997; Russian et al., 1999), we wanted to determine whether p55-related proteins are encoded by multiple genes. We amplified *P. carinii* genomic DNA from a single rat isolate by PCR with primers (PU947 and PD1637; Table 1) that were based on regions conserved between these two sequences. PCR amplification resulted in two distinct bands. After subcloning of amplified products, we obtained 12 clones, which were subsequently sequenced. Alignment of these sequences identified closely related but unique p55 variants, designated v0–v4, which consisted of five, one, one, three and two clones, respectively. Replicate clones of each variant were identical, except that two of the five clones of variant v0 contained a single nucleotide change at unique positions, which may be PCR artefacts. Fig. 1 shows an alignment of deduced amino acid sequences of the five variants. The sequence of variant v0 is identical to the published p55 sequence (Smulian et al., 1992) and contains 10 repeats of a heptapeptide motif near the C-terminus. Compared to v0, the other four variants (v1–v4) differed primarily in the number and size of heptapeptide repeats, although sporadic single nucleotide polymorphisms were identified. Variant v1 differs from v0 by the presence of only two repeats of the same heptapeptide motif. Variants v2, v3 and v4 lack a 10 aa sequence near the N-terminus compared to v0 and contain one or five repeats of a heptapeptide, rather than heptapeptide, motif. Variants v2 and v3 differ by nine amino acids in the C-terminus; v3 is identical to the sequence we identified initially. Variant v4 is the shortest sequence and contains only one heptapeptide motif. Additional clones were obtained by subcloning the PCR product from another rat isolate; they were found to correspond to the five variants, with no additional sequences being identified.

Southern blotting with *P. carinii* genomic DNA and a probe (v0p) that was specific for variant v0 showed a single band with all four enzymes (Fig. 2), consistent with previous reports (Smulian et al., 1992). When the same blot was stripped and hybridized with a probe (cp) that corresponded to a region conserved among all five *p55* variants, two bands (one of which corresponded to the single band in Fig. 2a) were seen for all enzymes except for EcoRI, which showed a single, high-molecular-mass band. These findings show that *P. carinii* has at least two copies of the *p55* gene and may have additional copies or allelic variants.

To determine whether the different variants were transcribed, Northern blotting was performed by using *P. carinii* total RNA (Fig. 3). Hybridization with a probe (c1r) from a region conserved among all five variants showed two bands of ~1.5 and 1.7 kb in size (Fig. 3). When the same blot was stripped and hybridized with a probe that corresponded to a region conserved between variants v1 and v0 (probe v01r) or between variants v2 and v3 (probe v23r), only one band was observed. Probe v01r hybridized to the higher band and v23r hybridized to the lower band. As only partial sequences were available for variants v1–v4, precise size of the corresponding mRNA could not be predicted. It is unclear whether individual bands present on the blot represent single variants. Given that available sequences for variants v2 and v3 are identical in size (487 bp) but differ in nucleotide sequence, the single band that hybridized to probe v23r is probably a mixture of two transcripts of the same size. As the available DNA sequence of variant v1 is only 147 bp shorter than that of variant v0 in the overlapping region, they may not be separated well by the gel; thus, the single band detected with probe v01r may also represent a mixture of two transcripts. However, the possibility that RNA expression levels differ among variants cannot be ruled out.

**p55 in *P. carinii* f. sp. *muris***

PCR amplification of *P. carinii* f. sp. *muris* genomic DNA with primers PU947 and PD1637 (Table 1) yielded a single band. Direct sequencing revealed a 471 bp fragment (GenBank accession no. AF494452) with an incomplete ORF encoding a 157 aa protein, which was most similar to variant v3 of *p55* *P. carinii* (Fig. 1) with 79 and 68% identity in nucleotide and amino acid sequences, respectively. Southern
blotting with *P. carinii* f. sp. *muris* genomic DNA and a probe amplified by using primers M39 and M472 (Table 1) showed variable numbers of bands for four restriction enzymes (Fig. 4). Whilst we did not perform subcloning of the PCR product to obtain individual clones, variation of the number of bands in the Southern blot may reflect sequence variation or a complicated organization of multiple gene copies, which suggests that, like *P. carinii*, *P. carinii* f. sp. *muris* may have multiple copies or allelic variants of the p55 gene.

These studies expand our understanding of the molecular organization of p55, which is one of the few antigens of *Pneumocystis* that has been characterized. Although the function of p55 is unknown, several studies have shown that p55 is capable of eliciting strong cellular and humoral immune responses in animals naturally infected with *P. carinii* (Smulian et al., 1992, 1993). Active immunization studies with recombinant p55 antigen have demonstrated partial protection against subsequent *P. carinii* infection in immunosuppressed rats (Smulian et al., 2000). Hence, p55 appears to play an important role in host–organism interactions. In the present study, we identified multiple variants of the p55 gene, which are closely related but differ clearly. Whilst PCR or subcloning artefacts could have generated some of the variants that we have identified, identification of the same variants in organisms obtained from two separate animals strongly suggests that all variants are real. Whilst additional studies are needed to better characterize these variants, to try to identify additional variants and to localize the variants on chromosomes, it seems likely, based on variable sequence data as well as Southern and Northern blot studies, that p55 is encoded by multiple genes, similar to two other surface proteins, MSG and PRT-1 (Kovacs et al., 1993; Lugli et al., 1997; Russian et al., 1999). The presence of multiple p55 genes supports a hypothesis that was derived from studies of MSG (Angus et al., 1996; Kovacs et al., 1993), i.e. that *Pneumocystis* has the potential for antigenic variability, which could play a role in immune evasion. Previous studies of rats exposed naturally to *P. carinii* have demonstrated that the repeat domain in the carboxyl portion of the p55 molecule is the target of both cellular and humoral immune responses, whereas the amino portion of this molecule is immunologically silent (Theus et al., 1994). In the current study, we found that p55 variants differed from each other primarily in the number and size of repeats, suggesting that *Pneumocystis* uses variation in the repeat region to evade host immune defences.
In P. jiroveci, the antigen complex of 35–45 kDa is the most common antigen found in lungs and bronchoalveolar fluid of patients with *Pneumocystis* pneumonia and is recognized strongly by serum antibodies of infected patients (Walzer & Linke, 1987; Smulian et al., 1992, 1993). Thus, it would be of great interest to clone the p55 counterpart in *P. jiroveci*. Previous studies have demonstrated that affinity-purified antibodies to the recombinant rat-derived *P. carinii* p55 antigen react with the 35–45 kDa band of human-derived *P. jiroveci*, suggesting the presence of shared epitopes (Smulian et al., 1992). We attempted to amplify *P. jiroveci* genomic DNA by PCR with a number of different primer pairs that were designed from regions conserved among the six sequences shown in Fig. 1. However, no specific sequences were identified. In addition, no hybridization signal was observed in Southern blot analysis with *P. jiroveci* genomic DNA and a variety of probes (either oligonucleotides or PCR products), even under conditions of low stringency (data not shown). These results suggest that *P. jiroveci* does not have a homologue of p55 or that it has evolved substantially and does not have adequate homology to permit identification by these methods.

In conclusion, we have found five variants of the gene that encodes the p55 antigen in *P. carinii*. Our results suggest the presence of polymorphism in the p55 gene and that at least two copies of this gene are present in the *P. carinii* genome. The p55 gene in *P. carinii* f. sp. *muris* shows a high degree of homology to the *P. carinii* p55 gene, but lacks the heptapeptide repeats. Novel approaches are needed to characterize the p55 gene counterpart in *P. jiroveci*.

**ACKNOWLEDGEMENTS**

We would like to thank Rene Costello and Howard Mostowski for their assistance. Part of this work was presented at the 7th International Workshops on Opportunistic Protists, Cincinnati, Ohio, 13–16 June 2001, abstract no. PO16.

**REFERENCES**


