THE NOVEL \textit{PLASMODIUM GALLINACEUM} SPOROZOITE PROTEIN, \textit{PG93}, IS PREFERENTIALLY EXPRESSED IN THE NUCLEUS OF OOCYST SPOROZOITES

ALEXIS N. LACRUE, ANTHONY A. JAMES, AND BRENDA T. BEERNTSEN*

\textit{Department of Veterinary Pathobiology, University of Missouri, Columbia, Missouri; Microbiology & Molecular Genetics, Molecular Biology & Biochemistry, University of California, Irvine, California}

Abstract. To study gene expression differences between oocyst and salivary gland sporozoites, cDNA libraries previously constructed from the two sporozoite populations of the avian malaria parasite, \textit{Plasmodium gallinaceum}, were used in a subtractive hybridization protocol to isolate \textit{Pg93}, a novel oocyst sporozoite gene. \textit{Pg93} encodes a putative \(\sim 76 \text{kDa} \) translated protein that was predicted to localize to the nucleus. Transcriptional analysis indicates that \textit{Pg93} is preferentially expressed in oocyst sporozoites versus salivary gland sporozoites. Immuno-localization assays confirm both the nuclear prediction and transcriptional analysis, suggesting that \textit{Pg93} is a nuclear protein. BLAST sequence analysis indicates that \textit{Pg93} represents a novel gene that has significant homology with a \textit{Plasmodium falciparum} hypothetical protein and translated \textit{Plasmodium knowlesi} and \textit{Plasmodium vivax} nucleotide sequences. This is the first characterization of a \textit{Plasmodium} nuclear protein that shows preferential expression in one sporozoite population as compared with the other population.

INTRODUCTION

Malaria is a mosquito-borne disease caused by protozoan parasites of the genus \textit{Plasmodium}. According to the World Health Organization (WHO), 300–500 million people are infected with malaria with more than 1 million fatal cases annually, most of whom are children in Africa. Malaria is exhibiting a resurgence due in part to the increase in drug-resistant parasites and insecticide-resistant mosquito vectors. Despite a century of research, efforts to eradicate malaria have been ineffective, and currently the disease is endemic in more than 90 countries throughout the world. However, with the advent of new technologies (e.g., microarrays and MudPIT) and new information garnered from genomics and proteomics advances, novel ways of controlling malaria now may be discovered.

Malaria sporozoites develop within the mosquito and travel from the oocyst, located on the basal lamina side of the midgut epithelium, to the salivary glands where they are transmitted to the vertebrate host when an infected mosquito takes a blood meal. This dual infectivity for host tissues makes oocyst and salivary gland sporozoites excellent targets for malaria control and prevention efforts. Consequently, molecular, cellular, and biochemical investigations of \textit{Plasmodium} sporozoites and particularly mosquito-sporozoite interactions are critical.

Although studies have shown that the infectivity of \textit{Plasmodium} sporozoites is developmentally regulated, efforts to determine the genes/proteins involved in sporozoite infectivity have focused primarily on sporozoites isolated from the salivary glands with currently little information available on gene expression specifically in the sporozoites located in the oocyst. Therefore, a subtractive hybridization protocol was undertaken to identify differentially or preferentially expressed oocyst sporozoite genes. A novel sporozoite gene, \textit{Pg93}, which is expressed preferentially in oocyst sporozoites, was isolated using a subtraction technique, and its initial characterization is described herein.

MATERIALS AND METHODS

Mosquito maintenance and the parasite transmission cycle. \textit{Aedes aegypti} (Red strain) mosquitoes used for the transmission cycle were 3–10 days old. The mosquitoes were maintained at 27°C and 80% relative humidity with a 16-hour light and 8-hour dark cycle. Sucrose (0.3 M) was fed to the mosquitoes ad libitum. \textit{Plasmodium gallinaceum} strain 8A was maintained in White Leghorn chickens by exposure to \textit{P. gallinaceum} sporozoite-infected mosquitoes. All animal care and the experiments described herein were approved by and performed in accordance with guidelines of the National Institutes of Health and the University of Missouri-Columbia Animal Care and Use Committee.

cDNA libraries. The oocyst sporozoite and salivary gland sporozoite cDNA libraries that were used in this study were provided by Dr. Louis Miller. These libraries were commercially prepared by Stratagene (La Jolla, CA) in the Lambda Zap II Custom cDNA library vector. The A. aegypti midgut cDNA library, which was used as a driver library in the subtraction protocol, was constructed using 1.8 \( \mu g \) of midgut poly A+ RNA and made using the Lambda Zap Express cDNA synthesis and Gigapack kit (Stratagene) following the manufacturer’s instructions.

Subtractive hybridization protocol. A subtractive hybridization protocol was conducted as previously described. Using this technique, single-stranded (ss), circular DNA plasmids were generated by an excision protocol from the oocyst sporozoite cDNA library (i.e., target library) and the salivary gland sporozoite and midgut cDNA libraries (i.e., driver libraries). An asymmetric polymerase chain reaction (PCR) with biotinylated dUTP was performed using the driver libraries as templates, and the resulting PCR product was hybridized with the ss plasmid DNA from the target library. Common sequences were subtracted away from the mix via the addition of streptavidin-coated magnetic beads, leaving behind nonbiotinylated target DNA sequences that were unique to or preferentially expressed in oocyst sporozoites. This target DNA was transformed into \textit{Escherichia coli} cells, and plasmid DNA was isolated from each colony. After DNA sequence determination, the \textit{Pg93} cDNA clone was identified.

Isolation of the \textit{Pg93} genomic clone. A \textit{P. gallinaceum} genomic library, kindly provided by Dr. David Kaslow, was...
screened to isolate a Pg93 genomic clone. The genomic library, maintained in the pBlueScript plasmid vector (Stratagene), was streaked to a density of 20,000 colonies per LB plate. The colonies then were transferred to nylon filters following a previously outlined protocol. The Pg93 cDNA (25 ng) was radiolabeled with 32P-dATP using the Megaprime DNA labeling kit (Amersham Biosciences, Piscataway, NJ) and hybridized to the filters overnight at 60°C. After hybridization, the filters were washed at 60°C and exposed to x-ray film as previously described. Those colonies that hybridized with the radiolabeled probe were isolated, plasmid DNA obtained, and the DNA sequence determined. The sequences then were aligned with the Pg93 cDNA clone, using the Vector NTI sequence analysis program (Informax, Inc., Bethesda, MD) to confirm that an appropriate genomic clone had been isolated.

**Sporozoite and sporozoite-infected tissue isolation.** For the confocal microscopy experiments, *P. gallinaceum* mature oocyst sporozoites were isolated from *A. aegypti* mosquitoes 8–10 days postexposure (PE) to an infective blood meal, and sporozoites were collected from salivary glands 12–14 days PE. The exact time for dissecting sporozoites was determined by examining a sample of midguts and salivary glands from the same infected cohort. Oocyst sporozoites were purified by placing dissected, infected midguts on a microscope slide and using the pressure of a coverslip to expel sporozoites from the oocyst. Sporozoites were collected in a drop of *Aedes* saline, transferred to a 1.5 mL microcentrifuge tube, and washed in *Aedes* saline. Infected salivary glands were disrupted gently in *Aedes* saline (125 µL) by pipetting to release the sporozoites. The homogenate was centrifuged at 18,000 × g and washed several times in *Aedes* saline (50 µL). After the final wash, sporozoites were resuspended in 50 µL of *Aedes* saline, and an aliquot was placed on a hemocytometer and sporozoites counted using compound microscopy.

For the rapid amplification of cDNA ends (RACE) protocol, sporozoite-infected midguts were dissected from mosquitoes 8–10 days PE and examined microscopically to determine oocyst maturity. The infected midguts selected for use had a mature oocyst burden of at least 20 oocysts and often had a considerably higher oocyst burden (> 50 oocysts/gut).

For the reverse transcription PCR (RT-PCR) protocol, abdomens, containing midguts with mature sporozoites in oocysts (8–10 days PE), were obtained by separating the abdominal region from the thorax with a razor blade. Thoraces, containing sporozoite-infected salivary glands, were obtained by slicing the thorax from the abdomen on Days 12–14 PE. Control blood meal abdomens and thoraces were obtained as described above. To prevent contamination, a new razor blade was used with each group. Whole naive mosquitoes also were used as controls.

**RNA isolation.** Total RNA needed for the RACE and RT-PCR protocols was isolated using the TRizol reagent as directed by the manufacturer (Invitrogen/Life Technologies, Carlsbad, CA) or the Qiagen RNAeasy mini kit (Qiagen, Valencia, CA). To remove residual genomic DNA, the purified total RNA was treated with amplification grade DNase I using 1 U DNase/µg RNA per the manufacturer’s instructions (Promega, Madison, WI).

**Rapid amplification of cDNA ends.** To obtain the 5′ end of the *Pg93* cDNA, a 5′ RACE PCR was performed using the First Choice RLM-RACE Kit (Ambion, Austin, TX) following the protocol supplied by the manufacturer. In brief, 30 µg of total RNA isolated from 60–90 *P. gallinaceum*-infected *A. aegypti* midguts, containing mature oocyst sporozoites, were treated with calf intestinal phosphatase (CIP; 10 U) and then incubated at 37°C for 1 hour. After CIP treatment and a phenol-chloroform extraction, the RNA was treated with tobacco acid pyrophosphatase (TAP-0.5 U) and incubated at 37°C for 1 hour. A kit adapter was ligated to the 5′ end of full-length decapped mRNA (10 µL final volume), and first-strand cDNA (20 µL) was synthesized using 2–4 µL of the decapped mRNA, a gene-specific anti-sense oligonucleotide primer, Superscript II reverse transcriptase (50 U), and the supplied buffers. The first-strand cDNA (2 µL) was used as a template in a 50 µL outer RACE PCR amplification containing final concentrations of the following: 1.25 U of Super Taq Plus polymerase (Ambion), 0.2 mM each dNTP, 1X reaction buffer, 2.0 mM MgSO4, and 0.4 µM each of the RACE kit and *Pg93* primers. The kit outer RACE primer was 5′-GC-TGATGCGGTGAATGAAACTG-3′ and the *Pg93* outer primer was 5′-CTATCATTTTTATGTTATCTTTT-3′. The outer RACE PCR product (2 µL) then was used in a nested inner RACE PCR reaction (50 µL) that consisted of the kit inner RACE primer: 5′-CGGCGATTCGAACTGATTTGCTGTTTATG-3′ and a *Pg93* inner primer 5′-TCATCATTTTTTGTGGTTC-3′. The following thermocycling conditions were used for the PCR reactions: 94°C for 3 minutes, followed by 35 cycles of 94°C for 30 seconds, 56°C for 30 seconds, 62°C for 3 minutes with a final extension at 62°C (the optimal annealing temperature for *Plasmodium* genes) for 10 minutes. PCR products were analyzed using agarose gel electrophoresis. Negative and positive control PCR reactions, using either distilled water (negative control) or *Pg93* DNA (positive control), also were done. RACE products (2–4 µL) then were ligated into the pPCR-Script Amp SK(+) cloning vector (10 ng) and transformed into XL10-Gold Kan ultracompetent cells (40 ng) and transformed into XL10-Gold Kan ultracompetent cells (40 ng) using the PCR-Script Amp Cloning Kit per the manufacturer’s instructions (Stratagene). Plasmid DNA was isolated from overnight cultures and sequenced by the campus sequencing facility. The resulting sequences were aligned and analyzed using previously described methods.

**Transcriptional expression analysis by RT-PCR.** To examine the expression of *Pg93* in sporozoites, RT-PCR was carried out using the Superscript First-Strand Synthesis System for RT-PCR (Invitrogen/Life Technologies). Total RNA (5 µg), isolated from sporozoite-infected *A. aegypti* abdomens and thoraces, was primed initially with oligo (dT) (0.5 µg per 20 µL reaction) following the manufacturer’s instructions. PCR amplification was performed with gene-specific forward (F) and reverse (R) oligonucleotide primers: *Pg93* (F, 5′-GA-AAAAATACAAACTTATGATG-3′; R, 5′-CTTT-ATATCTTAACTTATATGCGG-3′); *P. gallinaceum* circumsporozoite gene (CS) 16F (5′-TGATGCCAACAACCTGGAGGAGG-3′; R, 5′-AAAGAAACATCTTTTATGCTCGG-3′); and *A. aegypti* muscle actin gene (Aaeact-1) 16F (5′-GGTATCCACGAACATTGCTCAC-3′; R, 5′-TTTGGACAACATCTCTAC-3′). Each reaction (50 µL) had a final concentration of 1X PCR buffer without magnesium, 3.0 mM MgCl2, 0.2 mM each dNTP, 2.5 U of Taq polymerase (Promega), and 0.4 µM of each gene-specific primer. Initial denaturation of the template occurred at 94°C for 5 minutes. Amplification was performed for 30 cycles at 94°C for
30 seconds, 54°C for 30 seconds, and 62°C for 1 minute with a final extension at 62°C for 10 minutes. Aect-1 was used to show that RNA was present in all samples, and amplification of CS indicated the presence of sporozoite RNA in the parasite-infected samples. RNA isolated from nonfed mosquitoes and mosquitoes exposed to a noninfected blood meal served as controls. The amplification products were separated using agarose gel electrophoresis.

For the time-line RT-PCR analysis, *P. gallinaceum*–infected *A. aegypti* abdomens and thoraces were collected on Days 1–9 PE and Day 12 PE, respectively. Total RNA then was isolated and DNase treated as described above. This RNA (3 μg/sample) was used to make first-strand cDNA that then was used in an amplification reaction with gene-specific forward (F) and reverse (R) oligonucleotide primers: Pg93 (F, 5′-TAACTACAACACACAAATTCAAAAGG-3′; R, 5′-CCTAAAGATTCATCATCTTTCTTT-3′) and Aect-1 (see above). Each reaction was prepared as previously described and the initial denaturation of the template occurred at 94°C for 2 minutes. The subsequent amplification cycles were performed as previously described. This time-line experiment was performed twice, and the same results were observed each time.

To determine if Pg93 transcripts were present in asexual blood stages, total RNA was isolated from *P. gallinaceum*–infected chicken blood (0.2 mL) on Days 7–8 postinfection and noninfected chicken blood (0.2 mL) using TRI Reagent BD (Sigma, St. Louis, MO) per the manufacturer’s instructions. Total RNA (8 μg) was primed with random hexamers to generate first-strand cDNA, and PCR amplification then was performed with gene-specific forward (F) and reverse (R) oligonucleotide primers: Pg93 (same as the time-line RT-PCR primers); *P. gallinaceum* asexually expressed subunit ribosomal RNA (SSU rRNA)36 (F, 5′-GGCTTAGTTACCATTAAATAGGAGG-3′; R, 5′-CTACCGCTAATTAGCAGTTAAGG-3′); and chicken mRNA carbonic anhydrase II (CA II)39 (F, 5′-GAGTCTACAGGTTGGTGGTGCAGG-3′; R, 5′-CTGCGATTGGGCAGG-3′). Each reaction (50 μL) was prepared as described above. Initial denaturation of the template occurred at 94°C for 2 minutes and further amplification cycles were performed as previously described. CA II was used to show the presence of RNA in all samples, and the *P. gallinaceum* asexually expressed SSU rRNA gene indicated the presence of asexual blood stage RNA in the parasite-infected samples. The asexual blood stage transcriptional analysis was performed twice with the same results observed each time.

**DNA isolation and sequence determination of Pg93.** DNA, representing Pg93 genomic and cDNA clones and RACE products, was isolated from overnight cultures using Wizard mini-prep (Promega) and Qiagen midi-prep kits. The sequencing primers used in this study were obtained from Integrated DNA Technologies Inc. (Coralville, IA). Sequence of the isolated DNA was determined by the campus sequencing facility. DNA and protein sequence alignments were done using Vector NTI (Explorer or Contig Express). DNA and translated protein sequences were subjected to BLAST analysis to determine if they had similarity to sequences contained within the National Center for Biotechnology Information (NCBI) databases3 and the PlasmoDB database.34 To find homologues to the Pg93 protein, a PlasmoDB BLAST search was performed using the NCBI BLASTP program and “All *Plasmodium ORFs*” as the target database. PSORT II, a protein localization program, was used to predict protein localization sites. The ExPaSy Proteomics tools also were used for additional protein analysis.3 DNA sequences contained within the Wellcome Trust Sanger Institute Pathogen Sequencing Unit *P. gallinaceum* database were used to complete the 5′ end of the Pg93 genomic sequence. The sequence data have been submitted to GenBank (accession numbers AY775165–AY775167).

**Degenerate PCR reactions.** BLAST analysis of the Pg93 sequence against the PlasmoDB database indicated that *Plasmodium berghei* and *Plasmodium yoelii* did not contain a homologue to Pg93. To confirm these findings, degenerate PCR primers were designed from the conserved unknown domain 2, representing amino acids 551–558, and unknown domain 3, representing amino acids 625–631, of the Pg93 homologues. The primers were used in an amplification reaction containing 400 ng of genomic DNA isolated from *P. gallinaceum*, *P. falciparum*, *P. berghei*, and *P. yoelii*. The *P. falciparum* genomic DNA was provided by Dr. M. Ferdig (University of Notre Dame), the *P. gallinaceum* DNA provided by Dr. J. Vinetz (UCSD), and the *P. berghei* and *P. yoelii* DNA by Dr. A. Waters (Nijmegen). In brief, the templates were used in a 50 μL PCR amplification containing final concentrations of the following: 2.50 U of Taq polymerase (Promega), 0.4 mM dNTPs, 1X reaction buffer, 6.0 mM MgCl2, and 1.4 μM of each degenerate primer. The forward degenerate primer was 5′-YAYTAYGARAMIGTIGAIG-AAT-3′ and the reverse was 5′-AAYRAACCIARIGGIII-1CC-3′. Amplifications were performed at 94°C for 5 minutes, followed by 34 cycles of 94°C for 1 minute 30 seconds, 40°C for 2 minutes 30 seconds, and 50°C for 5 minutes. PCR products were analyzed by agarose gel electrophoresis. Negative and positive control PCR reactions, using either distilled water (negative control) or Pg93 cDNA (5 ng), also were done. To confirm that the *P. berghei* and *P. yoelii* genomic DNA was amplifiable, positive control reactions containing *P. berghei* and *P. yoelii* genomic DNA and gene-specific primers for the *P. berghei* thrombospondin related adhesive protein (TRAP)28 and Pyo1, a novel sporozoite gene (Beerntsen ET and others, unpublished data), respectively, were performed using a standard gene amplification protocol.

To confirm that the correct PCR products were amplified, the *P*93 degenerate PCR product and the positive control product were gel-purified using the Wizard SV gel and PCR clean-up system (Promega). The purified PCR products (12 ng) then were ligated into the pGEM-T Easy Vector (Promega) and transformed into JM109 competent cells (50 μL) per the manufacturer’s instructions (Promega). Plasmid DNA was isolated from overnight cultures and sequenced by the campus sequencing facility. The resulting sequences were aligned and analyzed using previously described methods.

**Recombinant protein expression and purification.** To produce polyclonal antibodies for immunoblotting studies and localization assays, recombinant Pg93 fusion protein (rPg93) was generated. A fragment, representing the carboxy region of the Pg93 open reading frame (ORF), was selected, and NeoI and Smal cloning sites were incorporated into the primers used to amplify this region. The primers, forward, 5′-GGGATCCATGGTTAGATATAGTAGACTATAAACAAAGG-3′, and reverse, 5′-CATCCCGCCGTCAGAGGTTAACCTAAAGATTGC-3′, were used to amplify a sequence
that encoded a region of the Pg93 protein defined by amino acids 432–630. The PCR amplifications were performed using the Pg93 genomic clone (25 ng) as a template and included reaction components as previously described. The amplification consisted of 30 cycles of 94°C for 30 seconds, 53°C for 45 seconds, and 62°C for 1 minute with a final extension temperature of 62°C for 10 minutes. The amplification product was digested with NcoI and Smal following the manufacturer’s protocol (Promega) and ligated into the NcoI/Smal double-digested dual-affinity expression vector, pAKSS. The Pg93 fragment then was expressed as a glutathione-S-transferase (GST)-Pg93-6xHis fusion protein. Purification of the rPg93-fusion protein was done via a His-Plus Kit protocol (Novagen, Madison, WI) modified by the inclusion of imidazole (final concentration of 50 mM) in the wash buffer, followed by a final wash without imidazole. Recombinant Pg93 protein was eluted with 50 mM EDTA, and the eluted fractions were analyzed for the presence of recombinant protein by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (10% gel) and Western blot analysis as described in the immunoblotting section.

Antibody production and purification. Two sterile wiffle balls (1.5 in diameter each) were surgically implanted into the lateral flank area caudal to the last rib of a female New Zealand White rabbit. Initially, pre-immune serum was obtained from the rabbit and then to eliminate cross-reactivity with E. coli, GST, and mosquito proteins, the pre-immune serum and the subsequently obtained anti-Pg93 serum were both depleted of anti-E. coli and anti-GST antibodies by chromatography on an E. coli–GST lyase-bound nickel column. To further purify these sera, they were preabsorbed to Optitran nitrocellulose strips (Schleicher & Schuell, Keene, NH) containing GST–E. coli lyase proteins, mosquito midgut proteins, and salivary gland proteins. Briefly, 10 A. aegypti midguts were resuspended in 10 μL of distilled water, and 10 pairs of salivary glands were resuspended in 5 μL of distilled water. Sample buffer was added to a final concentration of 0.5% bromophenol blue, 4% SDS, and 10% 2-mercaptoethanol (1X). The samples were homogenized using a pestle homogenizer (Kimble/Kontes, Vineland, NJ) and denatured at 95°C for 10 minutes. Following denaturation, samples were placed on ice for 2 min and then separated using SDS-PAGE (10% gel). Proteins were transferred onto an Optitran nitrocellulose membrane (Schleicher and Schuell) using the protocol described in the immunoblotting section. The membrane, containing lanes of midgut, salivary gland, or GST/E. coli lyse proteins, was cut into strips, and one strip was added to a 1.5 mL microfuge tube containing a 1:200 dilution of either pre-immune or anti-Pg93 serum and incubated at 4°C for 10 minutes on a rotator for a minimum of 1 hour. This procedure then was repeated for the other two strips.

After a 4-week recovery period, 100 μg of purified rPg93 fusion protein was injected, once every 2 weeks, into the wiffle balls. Fourteen days after the final injection (third injection), high titer serum was obtained from the rabbit and the anti-Pg93 serum was IgG-purified on a protein A column (Pierce Biotechnology, Inc., Rockford, IL). To eliminate any potential cross-reactivity, the anti-Pg93 serum was treated in the same manner as described for the pre-immune serum.

Immunoblotting. Western blot analysis was performed to confirm that the rPg93 fusion protein size was correct and that the rabbit anti-Pg93 polyclonal antibodies recognized the rPg93 protein. Briefly, the purified rPg93 protein and an aliquot of a GST–E. coli lysate were mixed 1:1 with sample buffer, denatured for 10 minutes at 95°C, and fractionated using SDS-PAGE (10% gel). The proteins were then transferred onto an Optitran nitrocellulose membrane (Schleicher and Schuell) at 25 volts for 50 min using a Trans-blot, SD-semi-dry transfer cell (BioRad, Hercules, CA). After overnight blocking with 5% non-fat dry milk/0.01% NaN₃ in Tris-buffered saline pH 7.5 (TBS) (0.15 M NaCl, 50 mM Tris), the blot was washed twice with TBS for 10 minutes and incubated with one of the following: 1) anti-Pg93 IgG (1:10,000), 2) rabbit pre-immune serum (1:10,000), or 3) mouse anti-GST mAb (Sigma) (1:2000) diluted in TBS/0.05% Tween 20 (TBS-T) for 1 hour at room temperature. After three 10-minute washes with TBS-T, the blot was incubated with horseradish peroxidase–conjugated goat anti-rabbit IgG (Sigma) at 1:20,000 dilution or anti-mouse IgG F(ab’)₂ (Jackson ImmunoResearch, West Grove, PA) 1:10,000 in TBS-T, and the electrochemical luminescence (ECL) kit (Amersham Biosciences) was used to detect positive signals. GST/E. coli–depleted rabbit pre-immune serum was used as a negative control to show that there was no cross-reactivity with bacterial proteins. The GST/E. coli lysate was used to show that the GST/E. coli–depleted anti-Pg93 antibodies did not recognize the GST or any E. coli proteins in the Western blot analysis.

Confocal microscopy. Sporozoites isolated from oocysts or salivary glands were resuspended in methanol-free 4% paraformaldehyde (EM Sciences, Washington, PA) in PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, and 2 mM KH₂PO₄)/0.075% NP40 and incubated for 30 minutes. After the incubation, 6,000–10,000 sporozoites were placed onto ringed slides (10 μL per ring) coated with 3-amino-propyltriethoxysilane (Sigma) to increase sporozoite adherence to the slides and allowed to air-dry. After air-drying, slides were spray washed four times with 1.0 mL of PBS pH 7.0/0.075% NP40 and blocked in PBS containing 1% non-fat dry milk/0.075% NP40 overnight at 4°C. Sporozoites then were incubated in primary antibody (anti-Pg93 IgG, rabbit preimmune serum, or anti-CS mAb) diluted 1:200 in blocking solution, overnight at 4°C. After three washes with PBS pH 7.0/0.075% NP40, sporozoites were incubated in the dark with an Alexa-488–labeled goat anti-rabbit or anti-mouse antibody (1:200) (Molecular Probes, Eugene, OR) and propidium iodide (nuclear stain) (Molecular Probes) at 1:50 dilution in blocking solution for 3 hours at room temperature. The slides then were washed 3–4 times with PBS pH 7.0/0.075% NP40 and one drop of Mowiol (Calbiochem, La Jolla, CA), a water-based mounting media, was added to each slide followed by a coverslip. The red and green fluorescence were visualized simultaneously using 488-nm light from an argon laser and 568-nm light from a krypton/argon laser on a Radiance 2000 confocal system (BioRad) coupled to an Olympus IX70 inverted microscope (Tokyo, Japan). The individual red images representing the nucleus of the sporozoite and green images representing the sporozoite were superimposed and analyzed using Laser Sharp 2000 software (BioRad). This confocal microscopy experiment was repeated three times with an estimated 100 sporozoites examined in each of the three groups per experiment.

RESULTS

DNA isolation and sequence determination and analysis of Pg93. After the subtractive hybridization protocol, an 871-bp
Figure 1. Nucleotide and deduced amino acid sequence of the novel *Plasmodium gallinaceum* gene, *Pg93*. The nucleotide sequence for the 2016 base pair consensus sequence is provided. Amino acids are represented by their single letter code, and amino acids in bold represent putative nuclear localization sequences. The asterisk (*) denotes the translational stop codon. The arrowheads denote the first and final nucleotide of the intron. The numbering begins at the first nucleotide encoding the open reading frame. Genbank accession numbers are AY775165 for the *Pg93* cDNA clone, AY775166 for the *Pg93* genomic clone, and AY775167 for a representative *Pg93* RACE product.

Figure 2. Alignment and PCR analysis of *Pg93* and its homologues. A. Alignment of the *Pg93* amino acid sequence with its homologues. *Plasmodium falciparum* (Pf93 orf /H11505 PFE1055c in PlasmoDB), *Plasmodium knowlesi* (Pk93 orf /H11505 Genomic Pk_195d11p1c in PlasmoDB), and *Plasmodium vivax* (Pv93 orf /H11505 Genomic Pv_402361 in PlasmoDB) all have distinct regions of identity with the *Pg93* protein as determined by aligning the translated proteins using the Vector NTI Explorer program. Amino acids highlighted in black are conserved in all proteins, and those highlighted in gray are conserved in a majority of the proteins. The three unknown domains (UKND) have a line above them. Numbers in parenthesis indicate the beginning and end of each amino acid sequence. B, A PCR amplification, using degenerate primers based on conserved amino acid sequences of the translated *Pg93* gene, suggests that *Plasmodium berghei* and *Plasmodium yoelii* do not contain homologues to *Pg93*. The amplifications were performed with degenerate primers designed from conserved domains (UKND) 2 and 3 of the *Pg93* protein and its homologues. The templates for each reaction consisted of *P. gallinaceum* (PG), *P. falciparum* (PF), *P. berghei* (PB), and *P. yoelii* (PY) genomic DNA with *Pg93* cDNA (5 ng) used as a positive control (+) and the PCR cocktail without DNA as a negative control. The arrow denotes the size (∼250 bp) of the degenerate PCR product. A basepair ladder is shown to the left of the figure.

---

LACRUE AND OTHERS 638

---
cDNA, designated *Pg*93, was isolated. This cDNA had a stop codon at the 3'-end corresponding to the region encoding the presumed carboxy terminus of the protein, but was incomplete at the 5' end. Therefore, a RACE protocol was performed to obtain the complete 5' end. Four RACE products were sequenced and aligned to generate a consensus sequence that resulted in a complete cDNA sequence (1,902 bp). Starting at the initiating methionine, the translated consensus cDNA sequence encodes a putative 634 amino acid (aa) protein with an approximate molecular mass of 76 kDa.

A.

---

B.
To obtain a full-length Pg93 genomic sequence, the Pg93 cDNA and a partial Pg93 genomic clone (1946 bp that was incomplete at the 5′ end) were used to identify the necessary 5′ genomic sequence from a *P. gallinaceum* whole-genome shotgun sequence database produced by the Pathogen Sequencing Unit at the Wellcome Trust Sanger Institute. The resulting consensus genomic sequence was 2016 bp in length with a single intron at base pairs 162–275 (Figure 1). The intron contains consensus splice-junction acceptor and donor sequences with 5′-end GT and 3′-end AG dinucleotides. Like the consensus cDNA sequence, the genomic sequence appears to contain a probable initiating methionine that follows Kozak’s rules, which state that a methionine is compatible with the start of translation if it contains an A nucleotide at the −3 position relative to the AUG start codon (data not shown).

Using Vector NTI analysis tools, the predicted protein is basic with a calculated net charge of 30.42 at a pH of 7.0 and a pI of 9.29. The PSORT II analysis program predicted a number of nuclear localization signals (NLS) including PYLKSSK (aa 51–57) and RKKKR (aa 231–235). It also predicted bipartite NLS that included KKKRSIEDKIKD-KKKKR (aa 70–87), KKOAHIGNKDDKHRT (aa 268–284), KDDYKATKKEEEEK (aa 345–362), and KKEKGKEKIEE (aa 374–390) (Figure 1). The ScanProsite program detected multiple putative sites for N-myristolation, N-glycosylation, and phosphorylation (data not shown). No signal peptide sequence was identified as determined by the SignalP program. Transcription factors are found in the nucleus, and although Pg93 was predicted and later confirmed to encode a nuclear protein, Pg93 does not appear to encode any known transcription factors. The Pg93 protein lacks domains and motifs commonly associated with transcription factors like DNA binding domains, zinc fingers, helix-turn-helices, leucine zippers, and high mobility group consensus sequences. The Pg93 protein also does not appear to have sequence homology with other known nuclear factors such as the splicing proteins involved in RNA splicing.

When compared with *Plasmodium* sequences in the PlasmoDB database via BLAST analysis, the Pg93 protein shared 29% identical amino acids and an additional 19% conserved residues with a *P. falciparum* hypothetical protein (PFE1055c in PlasmoDB) and 27% identical amino acids and 20% additional conserved residues with a translated *P. knowlesi* nucleotide sequence (Genomic Pk_195d1l1p1c in PlasmoDB) (Figure 2A). BLAST analysis of the Pg93 protein with the translated *P. vivax* sequences in the PlasmoDB database revealed that it shared 25% identity and had an additional 21% conserved residues with a translated *P. vivax* nucleotide sequence (Genomic Pv_402361 in PlasmoDB). Regions of high identity/similarity are referred to as unknown domains (UKND) 1, 2, and 3. BLAST analysis of these unknown domains, using the short, nearly exact match NCBI database, revealed that they have no significant similarity with other known proteins.

Using a degenerate PCR protocol, an ~250-bp PCR product, representing the region between the second and third conserved domains, was amplified from 400 ng of *P. gallinaceum* and *P. falciparum* genomic DNA, as well as the positive control DNA (5 ng Pg93 cDNA) (Figure 2B). After sequence analysis and alignment of the degenerate PCR products, it was determined that the flanking sequences matched the conserved regions from which the primers were designed, and the region between the primers corresponded with the sequence between the conserved regions. Equal amounts of *P. berghei* and *P. yoelii* genomic DNA failed to generate a product using the degenerate primers (Figure 2B), but when gene-specific primers were used, the appropriately sized products were amplified (data not shown). These results suggest that the homologue to Pg93 is not present in either *P. berghei* or *P. yoelii*. This experiment was performed twice and the same results were observed each time.

**Transcriptional analysis of Pg93 via RT-PCR.** The Pg93 gene transcript was present in both sporozoite-infected thoraces, containing mature salivary gland sporozoites, and in abdomens, containing mature oocyst sporozoites (Figure 3A). However, Pg93 transcripts appear to be more abundant within the sporozoite-infected abdomens as compared with sporozoite-infected salivary glands. Detection of the CS transcript indicates that there was sporozoite RNA in both of the sporozoite-infected tissue samples. Although there were only minor differences in CS abundance between the two sporozoite populations, the levels of Pg93 transcripts appear to be considerably higher in the oocyst sporozoites versus the salivary gland sporozoites even though the same starting amounts of total RNA and cDNA templates were used for each gene examined. This experiment was performed three times with independent RNA samples, and the same preferential expression pattern for Pg93 was seen on each occasion.

A more in-depth time-line RT-PCR analysis of Pg93 transcripts in the mosquito vector showed that there is an abundance of Pg93 transcripts present in the *P. gallinaceum*-infected *A. aegypti* mosquito midgut on Days 1–2 PE (Figure 3B). Although transcript abundance declines on Day 3 PE, it then increases beginning on Day 6 PE until Day 8 PE, during which time sporozoites are forming within the oocysts.

Transcriptional studies assessing *P. gallinaceum*-infected chicken blood show that Pg93 transcripts also are present during the erythrocytic stages (Figure 3C). Detection of the *P. gallinaceum* asexually expressed SSU rRNA gene indicates the presence of parasite RNA in the infected chicken blood, and detection of CA II shows that chicken RNA was present in all of the samples.

**Recombinant Pg93 protein expression and purification.** The recombinant Pg93-GST fusion protein (rPg93) was produced in *E. coli*, purified, and its size (55 kDa) was confirmed by SDS-PAGE (Figure 4A). Anti-GST monoclonal antibodies recognized the GST portion of the rPg93 protein (Figure 4B), indicating that the appropriate protein was produced. The recombinant protein then was injected into a rabbit for polyclonal antibody production. The GST/E. coli-depleted anti-Pg93 serum was IgG-purified on a protein A column, preabsorbed to noninfected mosquito tissues, and then used to detect rPg93 (Figure 4C). As expected, the purified anti-Pg93 antibodies recognized only rPg93 and did not recognize the GST protein. A negative control reaction, using preabsorbed rabbit pre-immune serum as a primary antibody, did not show cross-reactivity with rPg93 or any *E. coli* proteins (data not shown).

**Confocal microscopy.** In all oocyst sporozoite samples examined, the Pg93 protein localized to the nucleus as indicated by the yellow color where the red (nuclear stain) and green (secondary antibody) overlap (Figure 5) with limited cyto-
4. Western blot analysis using recombinant Pg93-exposed to a noninfected blood meal. A, A Coomassie blue–stained sodium dodecyl sulfate-polyacrylamide gel (10%) shows that the purified rPg93 protein was produced and is the appropriate size (~55 kDa). For all figures, the top right arrow denotes the size (~55 kDa) of the induced rPg93 protein (Lane 2) and the bottom right arrow denotes the ~29-kDa GST control protein in the lysate (Lane 1). B, Western blot analysis shows that the rPg93 fusion protein is recognized by anti-GST monoclonal antibodies, indicating that a protein of the correct size was produced (Lane 2). As expected, the GST control protein in the lysate also was recognized (Lane 1). C, The IgG-purified/preabsorbed anti-Pg93 polyclonal antibodies specifically recognize the purified rPg93 fusion protein. Western blot analysis shows that the induced rPg93 protein (Lane 2) is recognized by the anti-Pg93 polyclonal antibodies, indicating that the appropriate antibodies were produced in the rabbit. As expected, the ~29-kDa GST control protein is not recognized because the anti-Pg93 antibodies were preabsorbed against GST/E. coli proteins (Lane 1). Molecular weight markers in kilodaltons (kDa) are shown to the left of the figures.

**DISCUSSION**

In this study, we completed the initial molecular and cellular characterization of a novel oocyst sporozoite gene, *Pg93*, which was isolated via a subtractive hybridization protocol and appears to be preferentially expressed in oocyst sporozoites. Initially, the *Pg93* cDNA sequence was incomplete, and when salivary gland sporozoite and oocyst sporozoite cDNA libraries were screened, we were not able to obtain a full-length cDNA clone. When a RACE protocol using RNA isolated from salivary gland sporozoites was attempted, *Pg93* still could not be amplified. Only when RNA isolated from infected abdomens containing mature oocyst sporozoites was used for the RACE protocol was the cDNA sequence then completed, such that it contained an initiating methionine.

The consensus *Pg93* genomic sequence, like the consensus cDNA sequence, appears to contain a probable initiating methionine. A 114-bp intron is present in the consensus genomic sequence, and when the consensus *Pg93* sequence was subjected to BLAST analysis against the PlasmoDB database, homologous sequences were identified in *P. falciparum*, *P. knowlesi*, and *P. vivax*. These three homologous sequences represent open reading frames (ORFs) that were present in the PlasmoDB database, and each was predicted by PSORT II analysis to localize to the nucleus. The ScanProsite pro-
were generated and immunolocalization studies were performed. Not surprisingly, the immunolocalization assays indicated that the Pg93 gene product localizes primarily to the nucleus in oocyst sporozoites but is virtually undetectable in salivary gland sporozoites. These data confirm the PSORT II prediction that Pg93 encodes a nuclear protein. These data also support the transcriptional analysis, which suggests that Pg93 is preferentially expressed in oocyst sporozoites.

RT-PCR analysis of presporozoite stages in P. gallinaceum–infected mosquitoes and analysis of asexual blood-stages showed the presence of Pg93 transcripts in all stages examined. Pg93 transcripts were observed when ookinetes would be present in the infected mosquito (Day 1 PE) and during initial oocyst formation (Days 2–3 PE), as well as sporozoite development (Days 6–8 PE). It is possible, however, that the transcripts observed at Days 1–2 PE were from the asexual blood stage parasites ingested during the blood meal and did not represent new gene expression. Although Pg93 transcripts were detected in the asexual blood stages, preliminary immuno-fluorescence assays using Pg93 antibodies did not detect Pg93 protein expression in these particular stages (data not shown). These data are consistent with PlasmoDB developmental array analysis of P93 showing low levels of expression during the salivary gland sporozoite stage and expression during the asexual blood stages. Currently, there is neither oocyst sporozoite transcriptional data available for P93 nor is there mass spectrometry data for the P93 protein.

Many nuclear proteins are transcription factors; however, sequence analysis of the Pg93 protein suggests that it does not contain known transcription factor features, such as a helix-loop-helix, coiled-coil region, zinc finger, 30,32 or high mobility group sequences.27 Pg93 also does not appear to have known RNA recognition domains or sequence homology with other known nuclear factors such as the serine/arginine-rich (SR) and SR-like proteins that are involved in RNA splicing in the nucleus.12 Although the function of the Pg93 protein is currently unknown, as it does not encode a known transcription or splicing factor, it represents the first example of a sporozoite nuclear protein that is preferentially expressed in the oocyst sporozoite stage.

Received November 19, 2004. Accepted for publication March 18, 2005.

Acknowledgments: The authors thank Dr. Louis Miller for the sporozoite cDNA libraries and Dr. David Kaslow for the P. gallinaceum genome library. The whole genome shotgun sequence data for P. gallinaceum were produced by the Pathogen Sequencing Unit at the Wellcome Trust Sanger Institute and can be obtained from http://www.sanger.ac.uk/Projects/P_gallinaceum.

Financial support: This study was supported by the National Institutes of Health (NIH) grants AI 01657, AI 53156, and F32 AI 09731 to B.T.B and a Minority Biomedical Research Support (MBRTI)/NIH grant (R25 GM56901) and Missouri Alliance for Graduate Education and the Professoriate (MAGEP)/NSF Fellowship to A.N.L. A.A.J. was supported by a Burroughs-Wellcome Fund Molecular Parasitology Award.

Authors’ addresses: Alexis N. LaCrue and Brenda T. Beerntsen, Department of Veterinary Pathobiology, University of Missouri, 201 Connaway Hall, Columbia, MO 65211, Telephone: 573-882-5033, Fax: 573-884-5414, E-mail: anelbh@missouri.edu and BeerntsenB@missouri.edu. Anthony A. James, Microbiology & Molecular Genetics and Molecular Biology & Biochemistry, University of California,
REFERENCES


37. Wattam AR, Christensen BM, 1992. Induced polypeptides asso-