

UNIVERSITY OF WESTMINSTER



**WestminsterResearch**

<http://www.wmin.ac.uk/westminsterresearch>

**Stage-dependent localization of a novel gene product of the malaria parasite, Plasmodium falciparum.**

**Thanh Nguyen<sup>1</sup>, Hisashi Fujioka<sup>2</sup>, Angray Kang<sup>\*,3</sup>, William Rogers<sup>4</sup>, David Fidock<sup>5</sup>, Anthony James<sup>1</sup>**

\* Angray Kang now works in the School of Biosciences, University of Westminster

<sup>1</sup> Department of Molecular Biology and Biochemistry, University of California, Irvine, California 92697

<sup>2</sup> Institute of Pathology, Case Western Reserve University, Cleveland, Ohio 44106

<sup>3</sup> Diversa Corp., San Diego, California 92121

<sup>4</sup> Naval Medical Research Center, Silver Springs, Maryland 20910

<sup>5</sup> Department of Microbiology and Immunology, Albert Einstein College of Medicine, Bronx, New York 10461

This is an electronic version of an article published in Journal of Biological Chemistry, 276 (28). pp. 26724-26731, July 2001. Journal of Biological Chemistry is available online at:

<http://www.jbc.org/cgi/content/full/276/28/26724>

---

The WestminsterResearch online digital archive at the University of Westminster aims to make the research output of the University available to a wider audience. Copyright and Moral Rights remain with the authors and/or copyright owners. Users are permitted to download and/or print one copy for non-commercial private study or research. Further distribution and any use of material from within this archive for profit-making enterprises or for commercial gain is strictly forbidden.

---

Whilst further distribution of specific materials from within this archive is forbidden, you may freely distribute the URL of WestminsterResearch. (<http://www.wmin.ac.uk/westminsterresearch>).

In case of abuse or copyright appearing without permission e-mail [wattsn@wmin.ac.uk](mailto:wattsn@wmin.ac.uk).

## Stage-dependent Localization of a Novel Gene Product of the Malaria Parasite, *Plasmodium falciparum*\*

Received for publication, April 16, 2001, and in revised form, May 11, 2001  
Published, JBC Papers in Press, May 22, 2001, DOI 10.1074/jbc.M103375200

Thanh V. Nguyen‡, Hisashi Fujioka§, Angray S. Kang¶, William O. Rogers||, David A. Fidock\*\*, and Anthony A. James‡ ††

From the ‡Department of Molecular Biology and Biochemistry, University of California, Irvine, California 92697, the §Institute of Pathology, Case Western Reserve University, Cleveland, Ohio 44106, the ¶Diversa Corp., San Diego, California 92121, the ||Naval Medical Research Center, Silver Springs, Maryland 20910, and the \*\*Department of Microbiology and Immunology, Albert Einstein College of Medicine, Bronx, New York 10461

A novel *Plasmodium falciparum* gene, *MB2*, was identified by screening a sporozoite cDNA library with the serum of a human volunteer protected experimentally by the bites of *P. falciparum*-infected and irradiated mosquitoes. The single-exon, single-copy *MB2* gene is predicted to encode a protein with an  $M_r$  of 187,000. The *MB2* protein has an amino-terminal basic domain, a central acidic domain, and a carboxyl-terminal domain with similarity to the GTP-binding domain of the prokaryotic translation initiation factor 2. *MB2* is expressed in sporozoites, the liver, and blood-stage parasites and gametocytes. The *MB2* protein is distributed as a ~120-kDa moiety on the surface of sporozoites and is imported into the nucleus of blood-stage parasites as a ~66-kDa species. Proteolytic processing is favored as the mechanism regulating the distinct subcellular localization of the *MB2* protein. This differential localization provides multiple opportunities to exploit the *MB2* gene product as a vaccine or therapeutic target.

*Plasmodium falciparum* is the most virulent etiological agent of human malaria, responsible for over 90% of mortality due to the disease. Each year 300–500 million people are infected by malaria parasites, and this results in 1.5–3 million deaths (1). Efforts to eradicate malaria generally have failed, and currently the disease is endemic in more than 90 countries throughout the tropics. Widespread and increasing drug and insecticide resistance have exacerbated the situation, undermining the effectiveness of existing malaria control methods that depend on chemotherapy and vector control, respectively. Novel means to fight the disease are needed urgently, and a vaccine is predicted to have the greatest impact in addition to being the most cost-effective control measure (2).

Experimental support for the development of a vaccine for human malaria was provided first by the use of radiation-attenuated sporozoites as immunogens (3). The success of this experimental vaccination provided the impetus for the search for mechanisms of protective immune responses and the target antigens involved. The circumsporozoite (CS) protein was iden-

tified as the major surface antigen of *Plasmodium* sporozoites (4–6). The CS protein has been a leading vaccine candidate antigen because irradiated sporozoite-induced, protected human volunteers have high titers of anti-CS<sup>1</sup> antibodies (7), and CS-specific monoclonal antibodies and cytotoxic T-lymphocytes could adoptively transfer protection in a rodent malaria model system (8). However, attempts to induce protection in humans using *P. falciparum* CS-based vaccines, despite recent improvement in their immunogenicity, have repeatedly yielded only partial success (9–13).

The inability to develop a vaccine based on the CS protein was interpreted to indicate that additional antigens play a role in irradiated sporozoite-mediated protection against infection (14). It then becomes important to identify antigens that may act independently, additively or synergistically with the CS protein in the development of a multicomponent vaccine. Here we report the characterization of *MB2*, a novel gene encoding a *P. falciparum* sporozoite surface antigen identified by screening a CS-depleted sporozoite expression cDNA library with serum from a human volunteer protected by the bites of *P. falciparum*-infected and irradiated mosquitoes. The *MB2* gene is expressed in sporozoites, the exoerythrocytic stages, the asexual blood stages, and gametocytes, but the gene product is localized differentially in these developmental stages. This differential localization provides multiple opportunities to exploit the *MB2* gene product as a vaccine and drug target.

### EXPERIMENTAL PROCEDURES

**Library Construction**—A CS-depleted sporozoite cDNA library was constructed from a *P. falciparum* salivary gland sporozoite cDNA library (strain NF54; Ref. 15) using a hydroxyapatite column-based subtractive hybridization technique (16). Briefly, to prepare the target cDNA sense-strands, DNA from the unsubtracted library was linearized with *NotI* and used as a template to transcribe antisense cRNAs with T7 RNA polymerase (Megascript, Ambion). Template DNA was removed by DNase treatment and the antisense cRNA strands were used to generate cDNA sense strands in a reaction using SuperScript (Life Technologies, Inc.) reverse transcriptase. To prepare the driver cRNA, a CS clone, G89 (17), was linearized with *NotI*. Digestion products were used to generate antisense CS cRNA with T7 RNA polymerase (Megascript, Ambion).

\* This work was supported by grants from the Burroughs-Wellcome Fund and the John D. and Catherine T. MacArthur Foundation (to A. A. J.) and by National Institutes of Health Grant AI35827 (to H. F.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF378132–AF378138.

‡† To whom correspondence should be addressed: Dept. of Molecular Biology and Biochemistry, 3205 Bio Sci II, University of California, Irvine, CA 92697-3900. Tel.: 949-824-5930; Fax: 949-824-2814; E-mail: aajames@uci.edu.

<sup>1</sup> The abbreviations used are: CS, circumsporozoite; CRS, cell-surface retention signal; IEM, immunoelectron microscopy; GST, glutathione S-transferase; NLS, nuclear localization signal; ORF, open reading frame; PV, parasitophorous vacuole; aa, amino acid(s); UTR, untranslated region; bp, base pair(s); MSP, merozoite surface protein; TRAP, thrombospondin related anonymous protein.

The target cDNA sense strands were allowed to reassociate with a 50-fold excess of the driver cRNA antisense strands. The reassociation mix was loaded onto a hydroxyapatite column and non-duplex, single-stranded target cDNA was separated from duplex cDNA/cRNA by elution with a high molarity phosphate buffer. Primers specific for the UniZap  $\lambda$  phage vector (Stratagene) were used to amplify the subtracted cDNA, and the amplification products were subcloned into the phage arms of the UniZap vector and packaged.

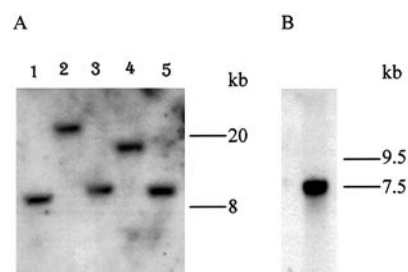
**Library Screening**—Phage were plated and lifted onto nitrocellulose membranes that were soaked in 10 mM isopropylthio- $\beta$ -D-galactoside and air-dried prior to use. Membranes were incubated in the serum of human volunteer 5<sup>2</sup> at a 1:100 dilution, and horseradish peroxidase-conjugated anti-human IgG+IgM were used to detect positive antibody reactions by the ECL system (Amersham Pharmacia Biotech). Positive phage were screened a second time to isolate single phage clones. Additional cDNA and genomic clones (strain ITO) were recovered using MB2-derived <sup>32</sup>P-labeled probes and standard library-screening techniques (18).

**DNA Sequencing of MB2**—The primary nucleotide sequences of all clones were determined by the dideoxynucleotide chain termination method (19) using a <sup>32</sup>P nucleotide terminator kit (Amersham Pharmacia Biotech). Specific oligonucleotide primers for sequencing were made by Heligen Laboratories (Huntington Beach, CA). Contiguity of clones was verified by gene amplification of genomic DNA using the following primers: a, 5'-GGTGATGACATTGAAGATATGAATG-3'; b, 5'-CAATAGAATAGATATAATACC; c, 5'-CTGGGTCATCATATGGAAAAGTG-3'; and d, 5'-CAATACACCCTGCAACCTTTCC-3'.

**Southern and Northern Analyses**—*P. falciparum* genomic DNA was isolated using a phenol/chloroform-based procedure (18) from blood-stage parasites (strain FCR3) cultured *in vitro*. The DNA was digested with various restriction endonucleases, and Southern blots were prepared as described (18). The probe for Southern blot analyses was prepared by labeling the sporozoite cDNA clone, *spz-MB2*, with radioactive [<sup>32</sup>P]ATP using the Megaprime DNA system (Amersham Pharmacia Biotech). Total RNA was isolated from blood-stage parasites of the same strain cultured *in vitro* using the Trizol<sup>®</sup> reagent (Life Technologies, Inc.). 15–20  $\mu$ g of total RNA were electrophoresed and Northern blots were prepared as described (18). Two <sup>32</sup>P-labeled probes consisting of nucleotides 1–580 and 2393–2836 of the coding sequence of *MB2* were used separately on filters to which RNA from blood-stage parasites had been transferred.

**Recombinant Protein Expression and Purification**—Segments of the MB2 open reading frame (ORF) were expressed in bacteria as GST-MB2-6xHis fusion proteins from the dual-affinity expression vector, pAK1-6H (20). *Nco*I and *Sma*I cloning sites were created for each insert by amplifying NF54 strain genomic DNA. The primer pair, 5'-GATGCCATGGAATATAATAGAATATGCTCA-3' and 5'-GATCCCGGGTTTTTATTATTAGAAGAATCA-3', was used to amplify a sequence that encodes a peptide, designated MB2-B, that overlaps amino acids (aa) 95–206. The primer pair, 5'-GATGCCATGGATTCTTCTAATAATAAATAAAT-3' and 5'-ATGCATCCCCGGGTCATTTTTATTGGAAGAAATTCCTC-3', was used to amplify a sequence that encodes a second peptide, designated MB2-C, that overlaps aa 200–316. The primer pair, 5'-GTATGCCATGGTCCACGAAAATAAAGAATATAATTCAG-3' and 5'-GATCCCGGGTCATCGAGCATTCTTTGGTC-3', was used to amplify a sequence encoding the peptide, MB2-FA, that overlaps aa 764–945. Finally, the primer pair, 5'-GATGCCATGGATGG TAATAGAACAAATAATGAC-3' and 5'-GATCCCGGGTACGCTTCGATTATATCGTTTGGCTC-3', was used to amplify a sequence that encodes the peptide, MB2-IF2, that overlaps aa 1337–1606. The amplification products were digested with *Nco*I and *Sma*I and ligated into pAK1-6H. The ligation mixture was used to transform *Escherichia coli* DH10B, and transformants were selected. Bacterial cells were grown at 37 °C in SuperBroth (Life Technologies, Inc.) to an  $A_{600} = 0.6$  and induced in 1 mM final concentration of isopropylthio- $\beta$ -D-galactoside for 3–4 h to express recombinant proteins. Purification of recombinant proteins was done using the ProBond resin (Invitrogen) modified by the inclusion of imidazole at 85 mM final concentration in the washing buffer. Eluted fractions were analyzed for the presence of recombinant proteins by SDS-polyacrylamide gel electrophoresis and immunoblotting.

**Rabbit Immunization**—400  $\mu$ g of purified recombinant protein were injected subcutaneously into a rabbit four times at 2-week intervals. Ten days following the last injection, high titer sera were obtained from the rabbit. The sera were depleted of anti-GST antibodies by chromatography on GST-bound nickel columns.



**FIG. 1. Southern and Northern analyses of the MB2 gene and transcription product.** A, Southern blot of *P. falciparum* genomic DNA, strain FCR3, digested with various restriction enzymes and hybridized with the *spz-MB2* cDNA clone. Lane 1, *Eco*RI/*Hind*III; lane 2, *Pst*I/*Hind*III; lane 3, *Pst*I/*Eco*RI; lane 4, *Pst*I/*Eco*RV; lane 5, *Pst*I/*Nde*I. B, Northern blot of *P. falciparum* blood-stage mRNA hybridized with a probe derived from nucleotides 1–580 of the MB2 ORF. The lane contained 20  $\mu$ g of total RNA. The approximate locations of molecular size markers in kilobases (kb) are indicated to the right of each of the panels.

**Immunoelectron Microscopy**—*P. falciparum* parasites and parasite-infected cells or tissues were fixed for 30 min at 4 °C with 1% formaldehyde, 0.1% glutaraldehyde in a 0.1 M phosphate buffer, pH = 7.4. Fixed samples were washed, dehydrated and embedded in LR White resin (Polysciences, Inc.). Thin sections (70–80 nm) were blocked in a phosphate buffer containing 5% w/v nonfat dry milk and 0.01% v/v Tween 20 (21). Grids were incubated at 4 °C overnight in solutions containing variable concentrations of rabbit antiserum reactive to domain-specific recombinant proteins diluted in the blocking buffer. Pre-immune sera were used as negative controls. After washing, grids were incubated for 1 h in 15 nm gold-conjugated goat anti-rabbit IgG (Amersham Pharmacia Biotech) diluted 1:40 in phosphate buffer containing 1% bovine serum albumin and 0.01% Tween 20. Following the 1-h incubation, grids were rinsed with phosphate buffer containing 1% bovine serum albumin and 0.01% Tween 20 and fixed with glutaraldehyde to stabilize the gold particles. Samples were stained with uranyl acetate and lead citrate and examined by electron microscopy.

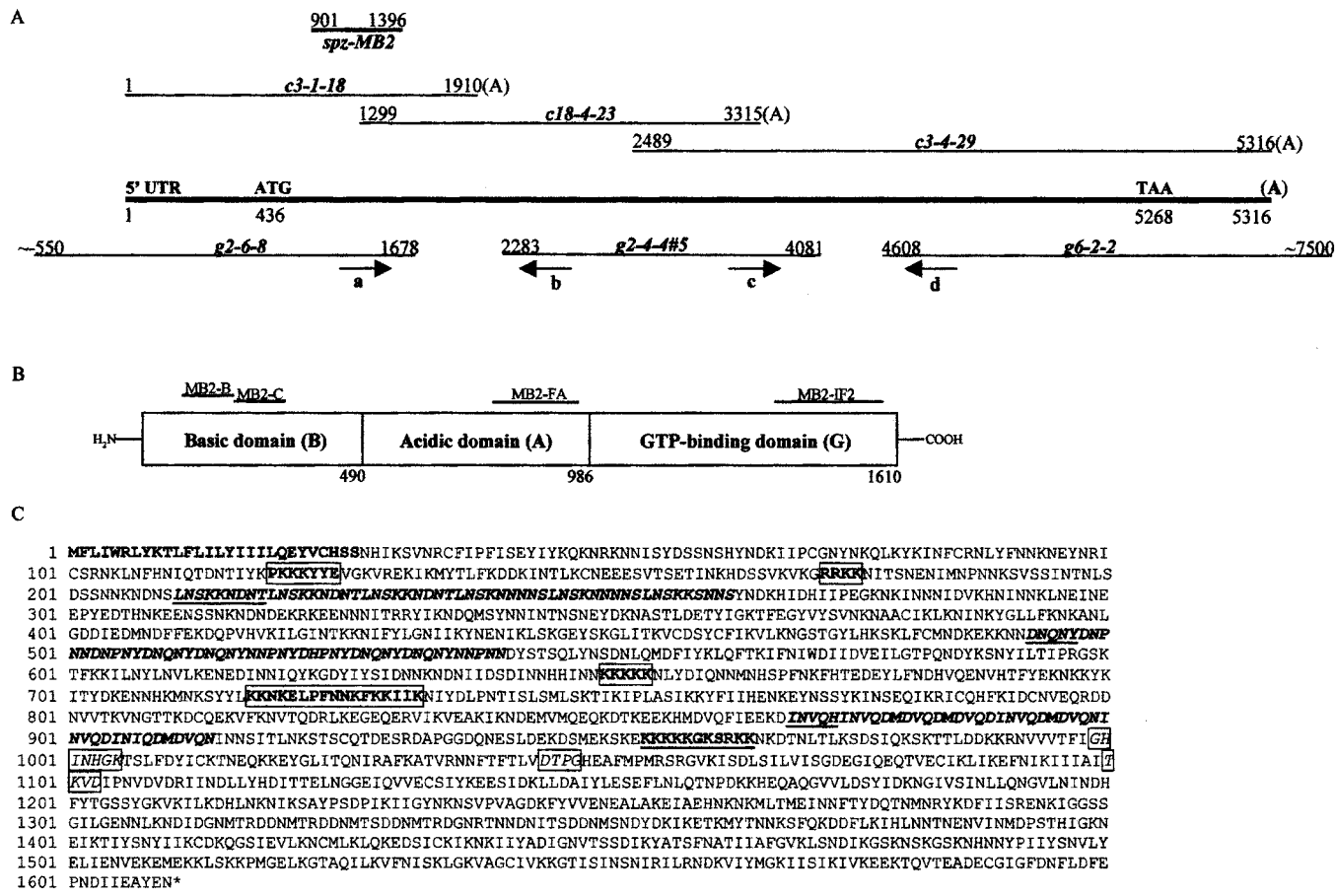
**Immunoblot Analysis**—Protein extracts from parasites were prepared by boiling them in sample buffer for 10 min (22). For the sporozoite stage, parasites were isolated from dissected salivary glands of infected mosquitoes. For the asexual blood stages, parasites were obtained from a saponin lysis of infected red blood cells grown in culture (23). Protein extracts were fractionated on 8% SDS-polyacrylamide gel electrophoresis and transferred onto nitrocellulose membranes. The membranes were incubated in rabbit antiserum diluted 1:500 for 1 h. Horseradish peroxidase-conjugated anti-rabbit IgG was used to detect positive signals using the ECL kit (Amersham Pharmacia Biotech). Pre-immune sera and lysates from uninfected human red blood cells were used as negative controls.

## RESULTS

**Library Construction**—To evaluate the efficiency of CS depletion in the subtracted cDNA library, 100 ng of the recovered single-stranded cDNA were amplified with oligonucleotide primers complementary to the cloning vector to generate a heterogeneous mixture of fragments of 200–1000 bp in size. The products were amplified again with gene-specific primers. Although two marker genes, *TRAP* (24, 25) and *P2* (26), were detected, no signal corresponding to the CS gene was amplified, indicating that the subtraction technique was highly efficient (data not shown). The amplification products were subcloned into the UniZap vector and packaged in  $\lambda$  phage, producing a library of  $1.45 \times 10^6$  primary phage.

**Library Screening, Southern and Northern Analyses of MB2**—A screening of  $1 \times 10^4$  primary phage with the human volunteer serum led to the selection of 18 candidate phage clones. These were rescreened, and 12 phage again were reactive for antibodies. Southern analyses showed that 1 of the 12 secondary clones hybridized specifically to *P. falciparum* genomic DNA and showed patterns of hybridization consistent with a single-copy gene (Fig. 1A). This 496-bp sporozoite cDNA clone was designated *spz-MB2* and was selected for further characterization. Northern analyses of RNA isolated from

<sup>2</sup> W. O. Rogers, unpublished data.



**FIG. 2. Structure of the MB2 gene and expression products.** A, schematic representation of cDNA and genomic clones used to identify and assemble a cDNA containing the complete ORF of the MB2 gene. The cDNAs, *spz-MB2*, *c3-1-18*, *c18-4-23*, and *c3-4-29*, are represented as horizontal lines above a linear representation of the full-length MB2 cDNA. The numbers above each cDNA refer to the terminal nucleotide positions in the completed cDNA. The As in parentheses in the cDNA clones represent the internal and terminal priming poly(A) sites of the oligo(dT) primers. The full-length cDNA is represented as a horizontal line numbered with the positions of the translation initiation (ATG) and translation termination (TAA) codons, and the beginning and end of the sequence. The 5' end untranslated region (5'-UTR) and polyadenylation sequences (A) also are indicated. The three horizontal lines at the bottom denote the MB2 genomic clones, *g2-6-8*, *g2-4-4#5*, and *g6-2-2*. The locations of the terminal nucleotides with respect to the cDNA are indicated above each line. Four horizontal arrows (a-d) represent the orientation and approximate location of gene amplification primers used to verify the contiguity of the sequence in the parasite genome. B, schematic representation of the MB2 protein sequence. The three domains, basic (B), acidic (A), and GTP-binding (G), are indicated as blocks with the junctions of the domains numbered below. The amino (H<sub>2</sub>N) and carboxyl (COOH) ends are labeled. The four short horizontal lines represent the approximate extents of the polypeptides, MB2-B, MB2-C, MB2-FA, and MB2-IF2, used to generate antibodies. C, primary amino acid sequence of the conceptual translation of the MB2 gene. Amino acids in bold represent the putative signal peptide; bold and boxed, putative nuclear localization sequences; bold and italicized, repeat regions with a single repeat unit underlined; bold and underlined, cell-surface retention sequence; italicized and boxed, motifs conserved in the G domain.

blood-stage parasites cultured *in vitro* and hybridized with probes derived from both the 5' and 3' ends of the complete ORF (described below) produced a single positive signal at ~7.5 kilobases (Fig. 1B).

**Sequence Analysis of the MB2 cDNA and Gene**—A comparison of the size of the *spz-MB2* cDNA with the mRNA detected in the Northern analyses indicates that it is not a full-length cDNA. Furthermore, primary sequencing of *spz-MB2* showed that it lacked a translation termination codon and represented an incomplete ORF. Sequence complementary to MB2 was detected in an asexual blood-stage cDNA library using specific gene amplification primers, and therefore the library was screened with the *spz-MB2* cDNA. Two overlapping blood-stage cDNAs, *c3-1-18* and *c18-4-23*, were identified (Fig. 2A). Nucleotide sequence analysis revealed that the reading frame of *spz-MB2* was contained entirely within a contig formed by these two cDNAs. The *c3-1-18* clone contained a putative translation initiation codon and a 435-bp 5' end untranslated region (UTR). The 3' end termini of the *c3-1-18* and *c18-4-23* cDNAs each have what appear to be polyadenylation (poly(A))

sequences characteristic of the 3' end termini of processed mRNAs. However, there were no translation termination codons located to the 5' end of the poly (A) tracks in either of the cDNAs, and the overlap of *c3-1-18* with *c18-4-23* revealed that the 17 terminal A nucleotides in *c3-1-18* comprise an internal A-rich nucleotide stretch in *c18-4-23*. Therefore, we concluded that the oligo(dT) primed the mRNA for cDNA synthesis from within the coding region.

To obtain additional 3' end sequence of MB2, a *Sau3AI* genomic library (strain ITO) was screened using as a probe the 400 nucleotides at the 3' end of *c18-4-23*. A genomic clone, *g2-4-4#5*, was identified having overlapping and contiguous sequence with *c18-4-23*. The sequence of *g2-4-4#5* confirmed that the 17-A region at the 3' end of *c18-4-23* is an internal A-rich nucleotide track, supporting the conclusion that these A-rich internal nucleotide tracks were primed by oligo(dT). To obtain additional 3' end cDNA sequence, the 600 nucleotides at the 3' end of *g2-4-4#5* were used to screen the blood-stage cDNA library, resulting in the identification of the cDNA clone, *c3-4-29*. Sequencing of *c3-4-29* revealed that it was contiguous

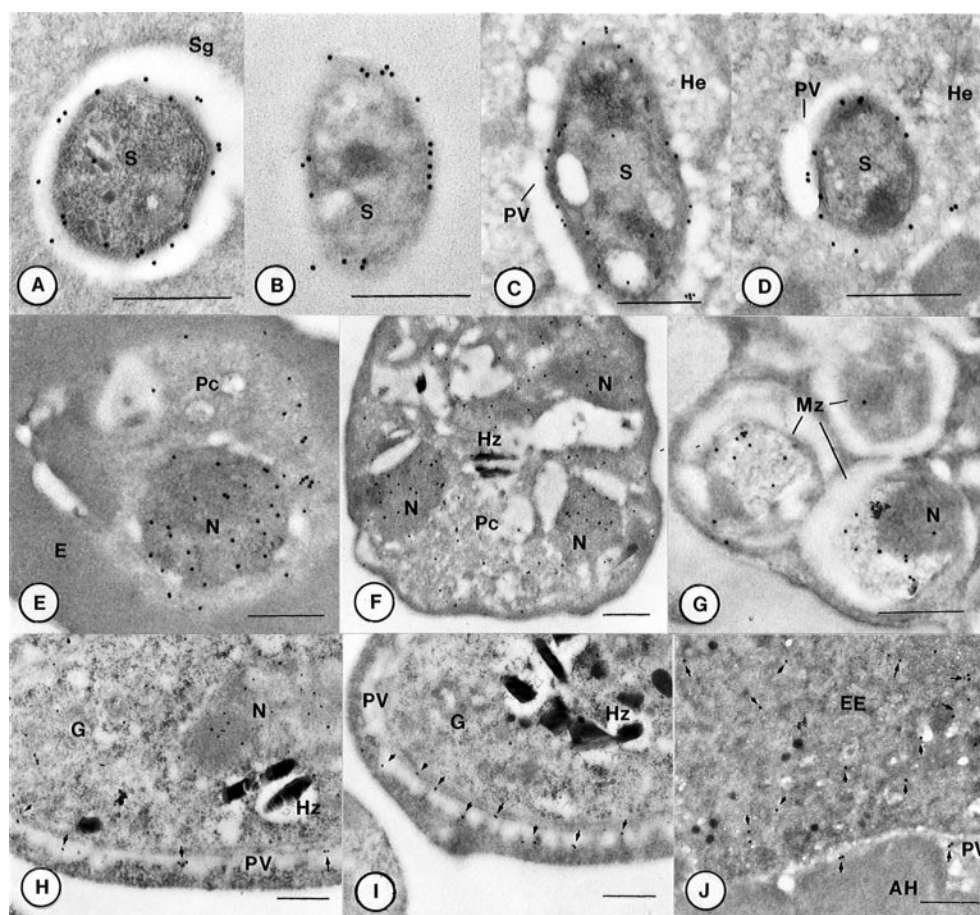


FIG. 3. Immunolocalization of the MB2 protein in different developmental stages of *P. falciparum*. A–D, sporozoite preparations. A shows a cross-section of a sporozoite (S) in the mosquito salivary gland (Sg) reacted with anti-B domain antiserum. B is a cross-section of a free sporozoite reacted with anti-A domain antiserum. C and D are partial-oblique and cross sections (respectively) of sporozoites in HepG2-A16 cells (He) reacted with anti-B and anti-A domain antisera, respectively. PV is the parasitophorous vacuole space. E–G, asexual stage parasite preparations. E is a cross-section of a trophozoite in an erythrocyte (E), showing localization principally to the parasite nucleus (N) and some in the parasite cytoplasm (Pc). F is a section of schizonts showing MB2 localization to the nucleus and some cytoplasm. Hemazoin (Hz) also is visible. Both E and F were reacted with anti-B domain antiserum. G shows sections of parasites at the merozoite (Mz) stage reacted with anti-A domain antiserum, showing only cytoplasmic localization. H and I, localization of MB2 in gametocytes (labeled G) reacted with anti-B and anti-A domain antisera, respectively. MB2 can be detected in the nucleus, cytoplasm, and the PV space. Arrows indicate the location of gold particles. J, localization of MB2 in the exoerythrocytic (EE) stages of an *Aotus* monkey hepatocyte (AH) reacted with anti-B domain antiserum. All bars are 0.5  $\mu$ m in length.

with *c18-4-23*. In addition, there are three stop codons at the 3' end of *c-3-4-29*, commencing at nucleotides 5266, 5272, and 5293, and there is a putative poly(A) region near the 3' end of the last stop codon. The positions of the stop codons and the authenticity of the poly(A) of the MB2 cDNA were supported by the genomic clone, *g6-2-2*, identified by screening the genomic library with a probe derived from the 3' end of *c3-4-29*. Similarly, the 5' end UTR and the start codon of MB2 also were verified by the clone *g2-6-8*, isolated from the genomic library using a probe derived from the 5' end of *c3-1-18*. The overlapping primary sequences of the three blood-stage cDNA clones and the contiguity of their reading frames allowed us to assemble a complete ORF of MB2 that is 4830 nucleotides in length, of which 77% of the bases are A-T pairs (data not shown). No nucleotide polymorphisms were observed among the cDNA and genomic sequences, indicating that there is a single allele of the MB2 gene encoded and expressed in the parasite strains used in our analyses.

Because the nucleotide sequences of the three genomic clones did not overlap, we designed gene amplification primers a, b, c, and d (Fig. 2A) to assess the contiguity of the MB2 gene in the parasite genome. Amplification products produced by the primer pairs a+b and c+d with parasite genomic DNA as the template gave the predicted product sizes, ~700 and ~1000 bp,

respectively (data not shown), indicating that MB2 is organized as a contiguous, single-exon gene in the parasite genome.

**Sequence Analysis of the MB2 Putative Translation Product**—MB2 encodes a putative translation product that is 1610 aa in length with an approximate molecular mass of 187 kDa (Fig. 2, B and C). The predicted protein is rich in asparagine (15%) and lysine (13%) and is strongly basic with a calculated net charge of +20 at pH 7 and a pI of 8.3. The primary amino acid sequence can be separated into three distinct, linear domains, the first of which is an amino-terminal basic domain of 490 residues (aa 1–490) with a calculated net charge of +30 and a pI of 9.4. We have designated this the “B” domain. This domain contains a region of six 9-aa imperfect repeats (aa 211–264) with the consensus sequence L, N, S, K, K, N, D/N, N, T/S. The central acidic domain, designated “A,” encompasses 496 residues (aa 491–986) with a calculated net charge of –6.2 and a pI of 6.1. The boundary between the B and A domains was selected to maximize the basic and acidic properties of the respective domains. The A domain contains two regions of imperfect repeats of 5 amino acids. The first region (aa 493–542) contains 10 repeats with a consensus sequence of D, N, Q/P, N, Y. The second region (aa 870–914) contains nine repeats with a consensus of I/M, N/D, V, Q, D. No similarities to any sequences of known function deposited in the data bases

were detected for either the B or A domains. Finally, a 624-residue carboxyl-terminal domain (aa 987–1610) with sequence similarity to the GTP-binding domain of the prokaryotic translation initiation factor 2, IF2, as revealed by the BLAST search program (27) has been designated "G." The boundary between the A and G domains was selected based on the start of the regions of similarity of the *MB2* protein with known IF2 molecules. In contrast to its overall hydrophilic nature, the *MB2* polypeptide contains at the amino terminus a strongly hydrophobic region (aa 1–25) mapped by a Kyte-Doolittle hydrophobicity plot. The PSORT computer program (28) predicted an uncleavable signal peptide in the hydrophobic amino-terminal region of *MB2*. However, the SignalP program (29) predicted that the signal peptide could be cleaved between a pair of S-S residues at aa 27–28 (Fig. 2C). Currently, we have no experimental data that support one alternative over the other. The PSORT program also predicted a number of nuclear localization signals (NLS), PKKK (aa 120–123), RRKK (aa 173–176), KKKKK (aa 652–656), and a bipartite NLS, KKNKELPFNN-KFKKIIK (aa 718–734), within the B and A domains. Multiple putative sites for *N*-glycosylation, *N*-myristoylation, and phosphorylation were detected by the ScanProsite program (Ref. 30; data not shown). There is a polybasic motif, KKKKKGKSRKK (aa 956–966), just before the start of the G domain, that could function as a plasma membrane localization signal as well as a cell-surface retention sequence (CRS) (31). This sequence also could be a putative NLS, although PSORT failed to identify it as such. The similarity of the G domain to the GTP-binding domains of the prokaryotic IF2 proteins includes the conservation of sequence and spacing of three motifs, GX<sub>4</sub>GK (aa 999–1005), DX<sub>2</sub>K (aa 1046–1049), and NKXD (aa 1100–1104), common to this family of proteins (32). There is a small variation in the third motif, TKXD, in *MB2* as compared with the consensus seen in other G proteins (Fig. 2C).

**Ultrastructural Localization of *MB2* Protein**—Immunoelectron microscopy (IEM) was used to study the subcellular localization of the *MB2* antigen. All rabbit antisera prepared against recombinant peptides derived from the B and A domains (Fig. 2A), and reacted with sectioned material containing sporozoites, showed that *MB2* protein was localized predominantly to the surface (Fig. 3, A–D). This was true of sporozoites in salivary glands (Fig. 3, A and B), as well as those that invaded *in vitro* cells of the human liver cell line, HepG2-A16 (Fig. 3, C and D). No antibody reaction was detected with sporozoites using the anti-G domain antibody (data not shown). Preimmune control sera for all reagents were negative (data not shown).

In contrast, the majority of the *MB2* protein detected in blood-stage parasites using both antisera against the B domain was localized in the nucleus, with some antibody reactivity detected in the cytoplasm (Fig. 3, E and F) and data not shown). Rabbit antisera against the A and G domains detected protein only in the cytoplasm of these parasites (Fig. 3G and data not shown). Furthermore, the numbers of gold particles observed in sections of parasites exposed to antibodies against the A and G domains were low when compared with the signal produced by the anti-B domain antisera, indicating probably that the majority of *MB2* protein present at the blood stages does not contain the A and G domains.

*MB2* protein was detected in the cytoplasm, nucleus, and parasitophorous vacuole (PV) space of gametocyte-stage parasites using the anti-B domain antiserum (Fig. 3H). *MB2* protein detected by the anti-A domain antiserum was localized only in the PV space (Fig. 3I), indicating that the protein detected in the nucleus and cytoplasm with the anti-B domain antiserum does not contain the A domain. The anti-G domain

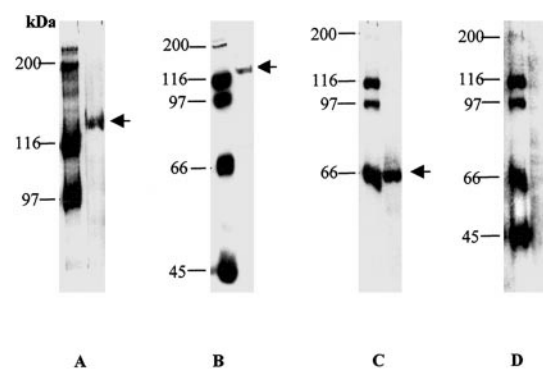


FIG. 4. Immunoblot analysis of protein extracts of *P. falciparum* sporozoite and blood stages. A and B, protein extracts prepared from sporozoites recovered from salivary glands of infected mosquitoes. C and D, proteins extracts prepared from asexual blood-stage parasites. A and C were probed with anti-B domain (MB2-B) antiserum; B and D were probed with anti-A domain (MB2-FA) antiserum. The molecular size markers (in kDa) are indicated to the left of each figure, and arrows to the right mark the locations of the *MB2* polypeptides.

antiserum produced a high background signal, making it difficult to interpret any specific localization pattern.

Finally, we attempted to use IEM to look at the localization of *MB2* protein in the exoerythrocytic stages of the parasite. A section of the liver of an infected *Aotus* monkey was reacted with the anti-B domain antiserum. It is hard to locate the parasites in these sections, but we were able to find some that revealed that the protein is localized mostly in the cytoplasm with some in the PV space (Fig. 3J). Sections reacted with the anti-A domain antiserum had high backgrounds obscuring any evidence of a specific localization pattern.

**Immunoblot Analyses**—A series of immunoblotting experiments were performed with parasite protein extracts prepared from the sporozoite and blood stages to determine the relative size of the *MB2* protein (Fig. 4). Both anti-B and anti-A domain antisera detected a single polypeptide of ~120 kDa at the sporozoite stage (Fig. 4, A and B). No immunoblot analyses were done on sporozoite preparations with anti-G domain antiserum because of the negative results obtained in the IEM analyses. Immunoblotting using anti-B domain antibody detected a single polypeptide of ~66 kDa at the blood stages (Fig. 4C). Furthermore, the 66-kDa polypeptide was not detected with either the anti-A or anti-G domain antibodies (Fig. 4D and data not shown). These data are consistent with the IEM study and indicate that the *MB2* polypeptide located at the surface of sporozoites consists of only the B and A domains, and the polypeptide translocated into the nucleus of parasites at the blood stages consists primarily of the B domain.

#### DISCUSSION

A major difficulty in immunoscreening of expression libraries to identify *P. falciparum* sporozoite antigens is the abundant and immunodominant characteristics of the CS protein. The abundant expression of the CS gene results in its representation in high frequency in cDNA libraries (15), and its immunodominant repeat domain induces a high antibody titer in the host providing the screening antiserum (7). Thus, to identify novel, non-CS antigens, we depleted the CS sequence from a sporozoite cDNA expression library prior to screening. The effectiveness of this approach was demonstrated by the isolation of *MB2*, a gene encoding a novel antigen with stage-dependent localization.

The Southern analyses revealed that *MB2* is present in the parasite genome most likely as a single-copy gene. The nucleotide sequence data indicated that it consists of a single exon and is represented most likely in our parasite samples by a

single allele. The Northern analyses with mRNA obtained from the blood-stage parasites indicated that *MB2* is expressed as a single large transcript. We do not know if this size is common to the other developmental stages of the parasite because of the difficulty in obtaining sufficient mRNA for blotting experiments. However, because the gene is single-copy and contains no intron, it is unlikely that multiple transcripts are produced, resulting either from expression of different genes or alternative splicing of a single gene. Thus, it is likely that the single-transcript expression of *MB2* is common to other developmental stages.

The overall length of the reconstructed *MB2* cDNA is 2.2 kilobases smaller than the RNA species detected in the Northern analyses. This difference results most likely from large 5' end, and perhaps 3' end, untranslated regions. Although we have neither a single genomic nor cDNA clone that spans the entire ORF of *MB2*, based on the overlapping primary sequence of the cDNA clones, the contiguity of their reading frames, and the gene amplification analyses of the genomic clones, we are confident that we have the complete expressed sequence of the *MB2* gene.

The complete ORF of *MB2* predicts a full-length protein of 187 kDa. However, there are many predicted sites for post-translational modification by myristoylation, glycosylation and phosphorylation. Therefore, it is likely that the actual molecular weight of the primary protein structure is increased by processing of individual amino acids. The predicted *MB2* protein is rich in asparagine (15%) and lysine (13%), and therefore is strongly basic. Asparagine is the most commonly used (~12%) amino acid in *P. falciparum*, followed by lysine and glutamic acid (~10%) (33, 34). Two other sporozoite surface proteins, CS and the sporozoite-threonine-and-asparagine-rich protein (STARP) (35), contain 29% and 25% asparagine, respectively. It has been speculated that asparagine-rich motifs in the amino acid sequence might be targets of opsonizing antibodies, promoting parasite phagocytosis by immune cells (36–38). Whether the *MB2* protein is a target of opsonizing antibodies is not known, but it is recognized by the immune serum of a human volunteer protected by the irradiated sporozoite vaccine.

Unlike a number of characterized *Plasmodium* genes that are active only in certain stages, the *MB2* gene is expressed in many developmental stages of the parasite life cycle. However, the *MB2* gene product has differential localization throughout development. The stage-dependent differential localization of the *MB2* protein suggests strongly that it has a multifunctional role during development of the parasites. It is conceivable that it functions as a signal recognition molecule while it is on the surface of the sporozoites. It then may transmit a signal to the nucleus by migrating there during the blood stages. Once inside the nucleus, it may function in the regulation of gene expression, participating in the process of turning off genes that are not required and activating genes that are required for blood stage infection. Examples of genes that are known to be inactivated as the parasite develops to the blood-stage are *CS* (39–40) and *TRAP* (25), and genes that are activated are merozoite surface protein genes, *MSPs* (41).

In gametocytes, the protein product is localized in the nucleus, cytoplasm, and the PV space. This differential localization may indicate that the *MB2* gene product is in a transitional phase from its functional role in the nucleus to the cell surface or it may have a role in the development of the sexual stages of the parasite. In the exoerythrocytic stage, the *MB2* protein detectable by anti-B domain antisera is localized mainly in the cytoplasm, although some can be detected in the PV space. As with the gametocytes, this expression may be a

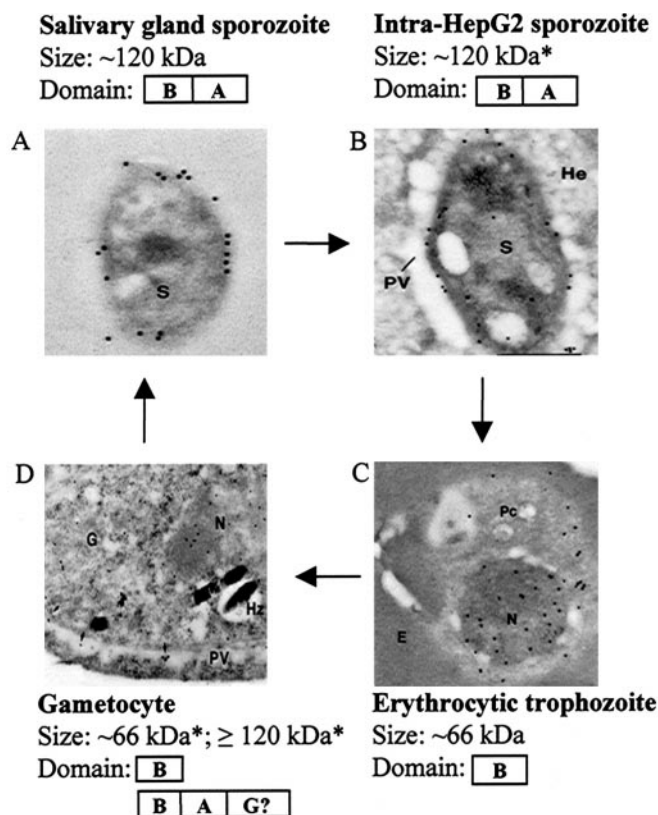


Fig. 5. Summary of expression and localization data for the *MB2* protein. Each panel (A–D) lists the stage of the parasite (first line) and the molecular size determination based on immunoblotting (second line). The third line indicates which domains were detected in either the immunoblotting or immune electron microscopy experiments. The immune electron micrographs are excerpted from Fig. 4. The asterisks in B and D indicate that the molecular size is not confirmed by immunoblotting analyses. The question mark (?) in D indicates that the presence of the G domain could not be unequivocally confirmed. All abbreviations are as in Figs. 3 and 4.

transitional phase in the specific localization as the parasite develops in the liver. The expression of *MB2* in this stage is important potentially as a vaccine target since the hepatocyte expresses major histocompatibility complex molecules that can be recognized by T cells. Research in the last 10 years has indicated that the infected hepatocyte can be an important target for immune attack (11).

Although the *MB2* protein is localized on the surface membrane of sporozoites, the primary amino acid sequence contains no apparent transmembrane domain or glycosylphosphatidylinositol anchor signal. However, the amino acid sequence does contain a polybasic motif that was shown to function as plasma membrane localization signal as well as CRS motif (31, 42). Many cytokines are retained on the membrane surface of the producer cell in a process mediated by the CRS. Studies have shown that, if the basic amino acids are deleted or mutated to acidic or neutral amino acids, then the membrane localization of the protein is affected (43, 44). Therefore, we predict that the polybasic motif is the most likely domain used to localize the *MB2* product to the membrane surface of sporozoites.

We provide evidence supporting the conclusion that protein processing is the mechanism by which *MB2* regulates its differential cellular localization (Fig. 5). The *MB2* protein is localized mostly to the surface membrane of sporozoites as a ~120-kDa species consisting of the B and A domains (Figs. 3 (A and B) and 5A). It is important to note that the predicted molecular mass of the B and A domains combined is ~116 kDa. This similarity in molecular mass of the actual protein and the

predicted domains is remarkably close and does not take into consideration the effects of post-translational changes to specific amino acids or the arbitrary boundaries assigned to the domains. As the sporozoites invade hepatocytes represented by the human cultured liver cells, the B and A domains of the MB2 protein are still detectable on the membrane surface (Figs. 3 (C and D) and 5B). We do not have immunoblot data for this stage, but we infer from the IEM study that the size of the protein is most likely ~120 kDa, similar to the size detected in free sporozoites. As the parasite develops to the blood stages, the majority of MB2 protein detected inside the parasite nucleus consists only of the B domain and is represented in immunoblots by a ~66-kDa species (Figs. 3 (E and F) and 5C). The predicted molecular mass, ~57 kDa, of the B domain selected by our analysis of the amino acid primary structure, is consistent with this smaller size polypeptide. As the parasite differentiates to gametocytes, the MB2 protein is found in the PV space as well as the nucleus and cytoplasm (Figs. 3 (H and I) and 5D). Based on the different labeling patterns seen with the anti-B and anti-A domain antisera, it is likely that the signal in the cytoplasm and nucleus originates from the ~66-kDa moiety. The protein in the PV space contains at least the B and A domains, and may contain the G domain. However, as noted in the results, the IEM study using the anti-G domain antibody is inconclusive, and we do not have immunoblot data that would provide the size of MB2 for the gametocyte stage.

We propose that the MB2 protein detected weakly in the cytoplasm of blood-stage parasites by the anti-A and anti-G domain antisera most likely represents the full-length, newly synthesized MB2 protein that has not been processed proteolytically into the ~66-kDa polypeptide. We propose further that the full-length MB2 protein is processed specifically at the sporozoite stage to the ~120-kDa polypeptide during synthesis and/or cellular trafficking. The ~120-kDa species contains the polybasic CRS-like motif, allowing it to be preferentially retained on the surface of the sporozoite. The secondary or higher-order structure of the ~120-kDa protein may conceal the NLS in the B and A domains. At the blood stage, the full-length MB2 protein is processed specifically into the ~66-kDa polypeptide as supported by the absence of the ~120-kDa species. The processing of the MB2 protein into the ~66-kDa polypeptide would remove the polybasic motif, thus removing the membrane targeting signal, and perhaps this processing exposes the nuclear localization signals allowing the ~66-kDa polypeptide to translocate to the nucleus. Finally, MB2 protein in the gametocyte stage may be processed into at least two forms, one of which consists of at least the B and A domains and is exported to the PV space. The other form, consisting most likely of only the B domain, is transported to the nucleus.

Another interesting feature of the MB2 protein is the G domain, which has significant sequence similarity to the prokaryotic IF2. Our data indicated that the G domain is not present in the MB2 protein detected on the sporozoite surface, nor is it present in the nucleus at the blood stages. It is conceivable that the cleavage of the MB2 protein requires energy, and this requirement is fulfilled by the G domain since it can bind to GTP. The cleavage process most likely includes removal of the G domain as evidenced by the inability to detect it with specific antiserum in most stages of the parasite. Alternatively, because MB2 can bind potentially to GTP, it is possible that there are conformational differences between the GTP-bound, GDP-bound, and unbound states that can regulate the distinct proteolytic processing of the MB2 protein.

In the last two decades, research efforts to develop recombinant vaccines against malaria have yielded largely limited successes. Although we have made significant progress in un-

derstanding immune responses and identifying a number of antigens, we are still not certain as to which host immune mechanisms and target antigens are relevant to protective immunity. In particular, the parasite antigenic composition at the sporozoite stage is still largely unknown. Because of the complexity of the *Plasmodium* species, we do not know if we have reached the heart of its antigenic repertoire required for identifying antigens that are relevant to protection or we are still wandering at the periphery, being misled by a limited number of identified antigens. However, vaccine development remains a viable option because it is clear that sterile immunity can be experimentally induced in humans (3). We believe that an optimal malaria vaccine would need to target multiple stages of the parasite life cycle, including the sexual stages developed inside the invertebrate host. As more novel antigens are being characterized, it is our expectation that the knowledge acquired from them can bridge the gap between recombinant and sporozoite-attenuated vaccines.

**Acknowledgments**—We are grateful to Jacques Prudhomme and Irwin Sherman (University of California, Riverside, CA) for providing blood stage parasites, Michael Hollingdale (Leeds University, Leeds, United Kingdom) for the infected *Aotus* and HepG2 preparations, and John Sacci (Naval Medical Research Center, Silver Springs, MD) for sporozoites. The study protocol was approved by institutional review boards at Naval Medical Research Center and Walter Reed Army Institute of Research, 1989–1994 (Protocol HURRAO Log A-4839) in compliance with all federal regulations governing the protection of human subjects. The experiments reported herein were conducted according to the principles set forth in the “Guide for the Care and Use of Laboratory Animals.”

#### REFERENCES

- World Health Organization (1998) *Fact Sheet N94*, World Health Organization, Geneva, Switzerland
- Miller, L. H., and Hoffman, S. L. (1998) *Nat. Med.* **4**, 520–524
- Clyde, D. F., Most, H., McCarthy, V. C., and Vanderberg, J. P. (1973) *Am. J. Med. Sci.* **266**, 169–177
- Yoshida, N., Nussenzweig, R. S., Potocnjak, P., Nussenzweig, V., and Aikawa, M. (1980) *Science* **207**, 71–73
- Zavala, F., Cochrane, A. H., Nardin, E. H., Nussenzweig, R. S., and Nussenzweig, V. (1983) *J. Exp. Med.* **157**, 1947–1957
- Dame, J. B., Williams, J. L., McCutchan, T. F., Weber, J. L., Wirtz, R. A., Hockmeyer, W. T., Maloy, W. L., Haynes, J. D., Schneider, I., and Roberts, D. (1984) *Science* **225**, 593–599
- Herrington, D., Davis, J., Nardin, E. H., Beier, M., Cortese, J., Eddy, H., Losonsky, G., Hollingdale, M. R., Szein, M., Levine, M., Nussenzweig, R. S., Clyde, D. F., and Edelman, R. (1991) *Am. J. Trop. Med. Hyg.* **5**, 539–547
- Weiss, W. R., Berzovsky, J. A., Houghten, R. A., Sedegah, M., Hollingdale, M. R., and Hoffman, S. L. (1992) *J. Immunol.* **149**, 2103–2109
- Ballou, W. R., Hoffman, S. L., Sherwood, J. A., Hollingdale, M. R., Neva, F. A., Hockmeyer, W. T., Gordon, D. M., Schneider, I., Wirtz, R. A., Young, J. F., Wasserman, G. F., Reeve, P., Diggs, C. L., and Chulay, J. D. (1987) *Lancet* **1**, 1277–1281
- Herrington, D., Clyde, D. F., Losonsky, G., Cortesia, M., Murphy, F. R., Davis, J., Baqar, A., Felix, A. M., Heimer, E. P., Gillesen, D., Nardin, E. H., Nussenzweig, R. S., Nussenzweig, V., Hollingdale, M. R., and Levine, M. M. (1987) *Nature* **328**, 257–259
- Hoffman, S. L., Franke, E. D., Rogers, W. O., and Mellouk, S. (1993) in *Molecular Immunologic Considerations in Malaria Vaccine Development*, (Good, M. F., and Saul, A. J., eds) pp. 149–167, CRC Press, London
- Stoute, J. A., Slaoui, M., Heppner, D. G., Momin, P., Kester, K. E., Desmons, P., Wellde, B. T., Garcon, N., Krzych, U., and Marchand, M. (1997) *N. Engl. J. Med.* **336**, 86–91
- Stoute, J. A., Kester, K. E., Krzych, U., Wellde, B. T., Hall, T., White, K., Glenn, G., Ockenhouse, C. F., Garcon, N., Schwenk, R., Lanar, D. E., Sun, P., Momin, P., Wirtz, R. A., Golenda, C., Slaoui, M., Wortmann, G., Holland, C., Dowler, M., Cohen, J., and Ballou, W. R. (1998) *J. Infect. Dis.* **178**, 1139–1144
- Galey, B., Druilhe, P., Ploton, I., Desgranges, C., Asavanich, A., Harinasuta, T., Marchand, C., Brahimi, K., Charoenvit, Y., Paul, C., Young, J., Gross, M., and Beaudoin, R. L. (1990) *Infect. Immun.* **9**, 2995–3001
- Fidock, D. A., Nguyen, T. V., Ribeiro, J. M., Valenzuela, J. G., and James, A. A. (2000) *Exp. Parasitol.* **95**, 220–225
- Usui, H., Falk, J. D., Dopazo, A., Lecea, L. D., Erlander, M. G., and Sutcliffe, J. G. (1994) *J. Neurosci.* **14**, 4915–4926
- McCutchan, T. F., Hansen, J. L., Dame, J. B., and Mullins, J. A. (1984) *Science* **225**, 625–628
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: a Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463–5467



20. Stratmann, T., Schmid, S. R., Harper, J., and Kang, A. S. (1997) *Protein Expression Purif.* **11**, 72–78
21. Aikawa, M., and Atkinson, C. T. (1990) *Adv. Parasitol.* **29**, 151–214
22. Laemmli, U. K. (1970) *Nature* **227**, 680–685
23. Hyde, J. E., and Read, M. (1993) *Methods Mol. Biol.* **21**, 133–143
24. Robson, K. J., Hall, J. R., Jennings, M. W., Harris, T. J., Marsh, K., Newbold, C. I., Tate, V. E., and Weatherall, D. J. (1988) *Nature* **335**, 79–82
25. Rogers, W. O., Malik, A., Mellouk, S., Nakamura, K., Rogers, M. D., Szarfman, A., Gordon, D. M., Nussler, A., Aikawa, M., and Hoffman, S. L. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 9176–9180
26. Fidock, D. A., Nguyen, T. V., Dodemont, H. J., Eling, W. M., and James, A. A. (1998) *Exp. Parasitol.* **89**, 125–128
27. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z. Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res.* **25**, 3389–3402
28. Nakai, K., and Kanehisa, M. (1992) *Genomics* **14**, 897–911
29. Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997) *Protein Eng.* **10**, 1–6
30. Appel, R. D., Bairoch, A., and Hochstrasser, D. F. (1994) *Trends Biochem. Sci.* **19**, 258–260
31. Hancock, J. F., Paterson, H., and Marshall, C. J. (1990) *Cell* **63**, 133–139
32. Dever, T. E., Glynnias, M. V., and Merrick, W. C. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 1814–1818
33. Hyde, J. E., and Sims, P. F. (1987) *Gene* **61**, 177–187
34. Weber, J. L. (1987) *Gene (Amst.)* **52**, 103–109
35. Fidock, D. A., Bottius, E., Brahimi, K., Moelans, I. I., Aikawa, M., Konings, R. N., Certa, U., Olafsson, P., Kaidoh, T., Asavanich, A., Guerin-Marchand, C., and Druilhe, P. (1994) *Mol. Biochem. Parasitol.* **64**, 219–232
36. Gysin, J., Gavaille, S., Mattei, D., Scherf, A., Bonnefoy, S., Mercereau-Puijalon, O., Feldmann, T., Kun, J., Muller-Hill, B., and Pereira da Silva, L. (1993) *J. Immunol. Methods* **159**, 209–219
37. Barale, J. C., Attal-Bonnefoy, G., Brahimi, K., Pereira da Silva, L., and Langsley, G. (1997) *Mol. Biochem. Parasitol.* **87**, 169–181
38. Barale, J. C., Candelle, D., Attal-Bonnefoy, G., Dehoux, P., Bonnefoy, S., Ridley, R., Pereira da Silva, L., and Langsley, G. (1997) *Infect. Immun.* **65**, 3003–3010
39. Suhrbier, A., Hamilton, A. J., Nicholas, J., and Sinden, R. E. (1988) *Eur. J. Cell Biol.* **46**, 25–30
40. Atkinson, C. T., Aikawa, M., Aley, S. B., and Hollingdale, M. R. (1989) *Am. J. Trop. Med. Hyg.* **41**, 9–17
41. Smythe, J. A., Coppel, R. L., Brown, G. V., Ramasamy, R., Kemp, D. J., and Anders, R. F. (1988) *Immunology* **85**, 5195–5199
42. Lokeshwar, V. B., Huang, S. S., and Huang, J. S. (1990) *J. Biol. Chem.* **265**, 1665–1675
43. Ostman, A., Anderson, M., Betsholtz, C., Westermark, B., and Heldin, C. H. (1991) *Cell Regul.* **2**, 503–512
44. Cadwallader, K. A., Paterson, H., Macdonald, S. G., and Hancock, J. F. (1994) *Mol. Cell. Biol.* **14**, 4722–4730