Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 7085–7089, August 1992
Microbiology

A family of erythrocyte binding proteins of malaria parasites

(*Plasmodium*/invasion/adhesins/Duffy blood group antigen/sialic acid)

JOHN H. ADAMS*†, B. KIM LEE SIM‡§, STEPHEN A. DOLAN*, XIANGDONG FANG*, DAVID C. KASLOW*, AND LOUIS H. MILLER*

*Laboratory of Malaria Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892; †Department of Biological Sciences, University of Notre Dame, Notre Dame, IN 46556; ‡Walter Reed Army Institute of Research, Washington, DC 20307; and §Johns Hopkins University, Baltimore, MD 21205

Contributed by Louis H. Miller, March 27, 1992

ABSTRACT **Malaria erythrocyte binding proteins use the Duffy blood group antigen (*Plasmodium vivax* and *Plasmodium knowlesi*) and sialic acid (*Plasmodium falciparum*) on the erythrocyte surface as receptors. We had previously cloned the one *P. vivax* gene, the one *P. falciparum* gene, and part of one of the three *P. knowlesi* genes encoding these erythrocyte binding proteins and described the homology between the *P. knowlesi* and *P. vivax* genes. We have completed the cloning and sequencing of the three *P. knowlesi* genes and identified introns in the *P. vivax* and *P. falciparum* genes that correct the previously published deduced amino acid sequences. All have similar structures, with one or two exons encoding the signal sequence and the erythrocyte binding domain, an exon encoding the transmembrane domain, and two exons encoding the cytoplasmic domain with the exception of the *P. knowlesi* β gene. The regions of amino acid sequence homology among all the genes are the 5′ and 3′ cysteine-rich regions of the erythrocyte binding domain. On the basis of gene structure and amino acid homology, we propose that the Duffy binding proteins and the sialic acid binding protein are members of a gene family. The level of conservation (≈70%) of the deduced amino acid sequences in the 5′ cysteine-rich region between the *P. vivax* protein and the three *P. knowlesi* proteins is as great as between the three *P. knowlesi* proteins themselves; the *P. knowlesi* β protein just 3′ to this cysteine-rich region is homologous to the *P. vivax* protein but not to the other *P. knowlesi* proteins. Conservation of amino acid sequences among these organisms, separated in evolution, may indicate the regions where the adhesin function resides.**

Invasion of erythrocytes by malaria parasites is dependent on binding of parasite proteins to receptors on the erythrocyte surface (1). *Plasmodium knowlesi* and *Plasmodium vivax* require interaction with the Duffy blood group antigen (Duffy-positive human erythrocytes) and cannot invade Duffy-negative human erythrocytes. *Plasmodium falciparum*, the major human malaria, can invade Duffy-negative and -positive erythrocytes equally well and, instead, binds specifically to neuraminidase-sensitive sialic acids on erythrocyte glycophorin. Neuraminidase treatment of erythrocytes does not reduce invasion by *P. knowlesi*. Thus, different parasite species use different receptors for invasion of erythrocytes (1). The parasite proteins that bind to the Duffy blood group antigen in *P. knowlesi* (2) and *P. vivax* (3) or to sialic acid in *P. falciparum* (4) were identified, and the genes encoding them were partially cloned and sequenced (5–7). The *P. falciparum* protein that binds to erythrocytes has a molecular mass of 175 kDa and was called erythrocyte binding antigen 175 (EBA175). Because it has now been shown that EBA175 binds specifically to sialic acid (8, 9), we have renamed it the *P. falciparum* sialic acid binding protein. The Duffy binding proteins of *P. knowlesi* are a family of genes with differing specificities (5); only a portion of one gene of this family had been cloned. The homologous *P. vivax* gene was cloned by cross-hybridization with the *P. knowlesi* gene (7). The common characteristics of the Duffy binding proteins of *P. knowlesi* and *P. vivax* are 5′ and 3′ cysteine-rich regions separated by a nonhomologous hydrophilic region, a transmembrane domain, and a 45- to 46-amino acid conserved cytoplasmic domain (7). In the present paper, we complete the cloning of the Duffy binding protein gene family of *P. knowlesi*, identify an intron after the signal sequence of *P. vivax*, and identify three exons that encode the transmembrane and cytoplasmic domains of the sialic acid binding protein of *P. falciparum*.¶ On the basis of gene structure and amino acid sequence homology, the Duffy binding proteins and the sialic acid binding protein are members of the same gene family. Because of the evolutionary distance between *P. falciparum* and *P. vivax*/*P. knowlesi* (10), the homologous regions may be conserved for receptor binding function.

MATERIALS AND METHODS

Parasites. The Malayan H strain of *P. knowlesi* was grown in rhesus monkeys. The Dd2/NM clone of *P. falciparum* was grown in culture.

cDNA and Genomic Clones of the Duffy Binding Protein Gene Family. Genomic DNA and total RNA were extracted and purified from *P. knowlesi* schizont-infected erythrocytes by described methods (5). Poly(A)-enriched *P. knowlesi* RNA was used to synthesize a cDNA library (p31) primed with an oligonucleotide (5′-TCCATATTAGTTCCCAT-3′) from the 5′ region of the α gene following the manufacturer's protocols (RiboClone cDNA synthesis system, Promega). The cDNA was treated with T4 DNA polymerase I, ligated into *Sma* I-cut, bacterial alkaline phosphatase-treated pUC18 plasmid DNA (Pharmacia), and used to transform XL1-Blue *Escherichia coli* (Stratagene). Complete and size-selected (3–6 kilobase) genomic DNA libraries were made from *Eco*RI-digested DNA as described (5). Colony lifts were done on GeneScreen*Plus* (DuPont) by using the manufacturer's recommended procedure.

An oligonucleotide (5′-TCCCTCATGCCCAGTCACA-3′) from the 5′ end of the *P. vivax* Duffy binding protein gene was used to screen the p31 cDNA library as described (5). Colony lifts were hybridized at 45°C and washed at 51°C in 6× standard saline citrate (SSC)/0.5% SDS. This selected a 681-base-pair fragment from the *P. knowlesi* β gene, clone p31 8A (BcDNA 8a). To select additional Duffy binding

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

Abbreviations: EBA175, 175-kDa erythrocyte binding antigen; ORF, open reading frame.

¶The sequences reported in this paper have been deposited in the GenBank data base (accession nos. M90466, M90694, M90695, and M93396–93398).

protein gene fragments from the p31 cDNA library and genomic fragments from the *Eco*RI genomic library, a 581-base-pair fragment of this clone was amplified by using PCR [primers: antisense oligonucleotide, 5′-TCCATATTAGT-TCCCAT-3′; sense oligonucleotide, 5′-GCAAGGAATATG-GTCATGTTGG-3′ (AmpliTaq, Cetus)], radiolabeled (11), hybridized at 65°C, and washed in 2× SSC/0.5% SDS at 55°C. This selected an additional cDNA clone of 1532 base pairs from the β gene (bcDNA 9A) and three genomic DNA clones (pPkβ, pPkα5′, and pPkγ5′). The 3′ end of the *P. knowlesi* γ gene (pPkγ3′) was cloned from the 3- to 6-kilobase size-selected *Eco*RI library by using *P. knowlesi* β exon 2 as a probe. Mature asexual stage parasites from clone Dd2/NM2 (12) were developed in culture and used to prepare a cDNA library in the phagemid vector pcDNAII. The library was screened with a genomic clone from the *P. falciparum* sialic acid binding protein/EBA175 (pEBA1.8/Y1090). Both strands of each plasmid DNA were sequenced by the dideoxy nucleotide termination method (Sequenase, United States Biochemical).

Sequence Analysis. The *P. falciparum* sialic acid binding protein sequence was retrieved from Swiss-Prot release no. 18 (accession number P192145; GenBank/EMBL accession no. X52524). The partial *P. knowlesi* and *P. vivax* Duffy binding protein gene sequences are in the GenBank data base [*P. knowlesi* α, accession nos. M37512 (cDNA p1C1) and M37513 (pECO6); *P. vivax*, accession no. M37514].

Sequence analysis was done by using computer-based algorithms in PCGENE and GENEWORKS 2.0 (both programs from IntelliGenetics). The sequence alignments were done with CLUSTAL (in PCGENE) and ALIGNMENT (in GENEWORKS). The final alignments were manually adjusted to give the best fit.

RESULTS AND DISCUSSION

The Open Reading Frames (ORFs) of the Duffy Binding Proteins. The genomic sequences and intron/exon boundaries of *P. knowlesi* α, *P. knowlesi* β, *P. knowlesi* γ, and *P. vivax* are in the GenBank data base (accession nos. M90466, M90694, M90695, and M37514, respectively). *P. knowlesi* α and β genes were shown to be expressed by comparison of the sequence of genomic and cDNA clones (ref. 5; J.H.A., unpublished data). Proof of expression of the *P. knowlesi* γ and *P. vivax* genes was determined by analysis of the PCR products of cDNA and genomic DNA (J.H.A., unpublished data).

Intron 1 of *P. knowlesi* β and *P. vivax* was delineated by the sequence of the cDNA in the 5′ region of these genes. Intron 1 in the *P. knowlesi* α and γ genes was inferred from the genomic DNA sequence by analogy with the *P. knowlesi* β gene. There were stop codons between the signal sequences of the α and γ genes and the presumed exon 2. Intron 1 of the β gene was homologous to the DNA sequence of the α and γ genes, including donor and acceptor sites for splicing (deposited in the data base). Even though intron 1 of the *P. vivax* gene maintained an ORF, splicing was determined to be the usual event by analysis of cDNA (J.H.A., unpublished data).

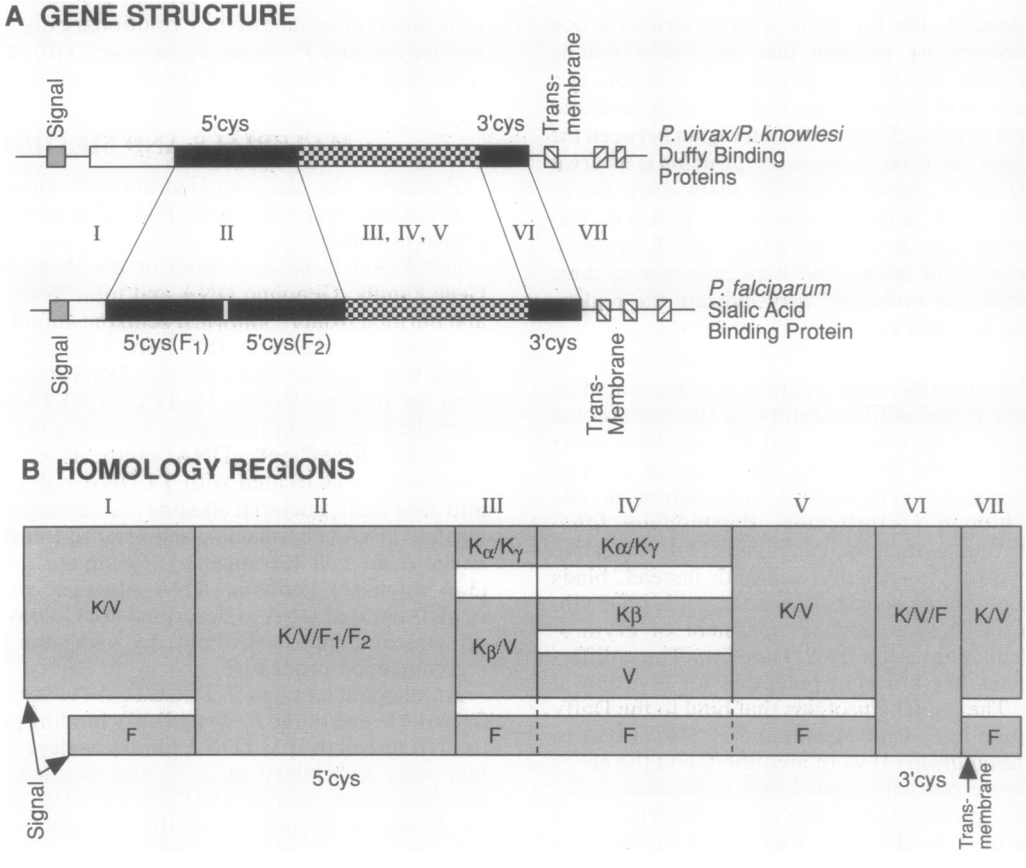

FIG. 1. Comparison of gene structure (*A*) and regions of amino acid sequence homology (*B*) of *P. vivax*, *P. knowlesi*, and *P. falciparum*. (*A*) Coding sequences are blocked and drawn to scale. Black areas are cysteine-rich regions (5′ cys and 3′ cys), and checked areas are the regions of low homology. Roman numerals designate different regions. The boundaries for regions I–VII can be determined from the sequence of *P. vivax* (7) (amino acids 250, 575, 732, 844, 945, and 1051) and the *P. knowlesi* data base. (*B*) Homology regions (I–VII) are delineated by using vertical lines. Horizontal lines within regions group the genes by amino acid sequence similarity (see Table 1). Dissimilar amino acid sequences are separated. K, *P. knowlesi*; V, *P. vivax*; F, *P. falciparum*; F_1 and F_2, duplicate copies of the *P. falciparum* 5′ cysteine-rich motifs.

Table 1. Amino acid sequence identity of erythrocyte binding proteins divided into regions I–VII*

Gene product	Region						
	I	II	III	IV	V	VI	VII
Pkγ	95	71	88	96	96	97	92
Pkβ	96	71	8†	16‡	93	95	94
Pv	55	70	12†	6‡	38	68	65
Pf	8	27 (F_1) 19 (F_2)	5	5	4	41	14

The values are the percent identity in each region relative to the *P. knowlesi* α gene product. Pk, *P. knowlesi*; Pv, *P. vivax*; Pf, *P. falciparum*.

*See Fig. 1.

†*P. vivax* and *P. knowlesi* β have 45% identity to each other in this region.

‡*P. vivax* and *P. knowlesi* β have only 8% identity to each other in this region.

The exon/intron boundaries for the 3′ part of the *P. vivax* and the *P. knowlesi* β and γ genes were inferred from cDNA for the *P. knowlesi* α gene (5). The splice sites for the other genes were determined by homology with the amino acid sequence after splicing out introns that had consensus splice sites (deposited in the data base).

Exons Are Functional Domains for the Duffy Binding Proteins. Each of the *P. knowlesi* and the *P. vivax* Duffy binding protein genes has a similar structure and sequence (see Fig. 1). The ORF of each gene comes from five exons except for the *P. knowlesi* β gene, which lacks intron 4 (deposited in the data base). Each exon represents a functional domain of the binding protein, as is true for other eukaryotic genes (13). Exon 1 encodes a signal sequence. The methionine start codon is followed by several positively charged residues and then 12 hydrophobic residues. The start codon is continuous with the ORF found in the cDNA and is immediately preceded by a sequence (ATACA) that fits the Kozak consensus sequence for initiation of translation (14–16). The *P. knowlesi* β cDNA fragment had 584 bases of untranslated region 5′ to this start codon; no ATG present in this 5′ untranslated region is followed by a long ORF (data not shown).

Exon 2 encodes an ORF of 987, 1065, 984, and 986 amino acids for *P. knowlesi* α, β, and γ and *P. vivax*, respectively, and contains the erythrocyte binding domain. The designation of exon 2, or a part of it, as the erythrocyte binding domain derives from the study of Adams *et al.* (5) in which the soluble erythrocyte binding proteins were 135 kDa and greater. As exon 2 encodes a large protein and is followed by a transmembrane domain encoded by exon 3, exon 2 must encode this large, soluble protein. Furthermore, the soluble erythrocyte binding proteins were not immunoprecipitated by the antiserum to the C-terminal peptide, which is located after the transmembrane domain. This antiserum immunoprecipitated higher molecular mass proteins from detergent-extracted parasites. Exon 3 encodes a transmembrane domain of 22, 21, 22, and 18 amino acids for *P. knowlesi* α, β, and γ and *P. vivax*, respectively. Exons 4 and 5 encode a cytoplasmic domain of 45 amino acids. Exon 4 of the *P. knowlesi* β gene encodes the entire cytoplasmic domain; intron 4 is absent from this gene.

Regions of the Erythrocyte Binding Domain in Exon 2 in *P. knowlesi* and *P. vivax*. Exon 2 contains five regions as defined by amino acid sequence identity (Fig. 1 and Table 1). Region I, which contains exon 1 and the beginning of exon 2, is highly conserved among the *P. knowlesi* genes (≈95%) and, to a lesser degree, is similar in *P. vivax* (55%; Table 1). The region is highly charged, with negative charges predominating (deposited in the data base).

The 5′ cysteine-rich region (region II; Figs. 1 and 2) has 12 cysteine residues, all of which are conserved in position, requiring only the insertion of one space in the *P. knowlesi* γ gene. The aromatic amino acids are highly conserved in this region: all 10 tryptophan residues are conserved; 7 phenylalanine residues and 10 tyrosine residues are conserved; between 2 and 5 phenylalanine or tyrosine residues for each gene are nonconserved (Fig. 2). Interestingly, in region II, the three genes from *P. knowlesi* have as much variability among themselves as between any one of them and *P. vivax*—that is, they all show ≈70% homology to the *P. knowlesi* α gene.

The 3′ cysteine-rich region (region VI; Figs. 1 and 2) spans the final 104 residues of exon 2 and immediately precedes the transmembrane domain. The 3′ cysteine-rich region is nearly identical among the *P. knowlesi* genes (Table 1). The homology between the *P. knowlesi* α gene and the *P. vivax* gene is 68%. All cysteine residues and 11 of 13–15 aromatic residues are conserved among the *P. knowlesi* and *P. vivax* genes (Fig. 2).

FIG. 2. Conserved 5′ and 3′ cysteine-rich regions of the erythrocyte binding proteins. Identical residues are shaded and cysteine residues are boxed. K, *P. knowlesi* consensus sequence encoded by the α, β, and γ genes (with a lowercase letter signifying a substitution in one of the proteins); V, *P. vivax* sequence; F, *P. falciparum* sequence; F_1 and F_2, duplicate copies of the *P. falciparum* 5′ cysteine-rich motifs; x_n, the number (n) of intervening residues. Dashes indicate the insertion of spaces into the sequence for best alignment.

Separating the cysteine-rich regions are regions III, IV, and V. *P. knowlesi* α is similar to *P. knowlesi* γ in all these regions (Table 1). In region III, *P. knowlesi* β and *P. vivax* have 45% homology to each other, including a stretch of 45 amino acids that is missing from *P. knowlesi* α and γ. The homology in region III between *P. knowlesi* α/γ and *P. knowlesi* β/*P. vivax* is low. With the exception of *P. knowlesi* α and γ, all genes are dissimilar in region IV. Region V has high homology (93–96%) between the *P. knowlesi* genes and low homology (38%) to *P. vivax*. Region V contains multiple copies of a slightly variant pentapeptide repeat SS(D/N)(Q/H)T in each *P. knowlesi* protein but only a single copy of SDQT in the *P. vivax* sequence.

As exon 2 is the erythrocyte binding domain, the sequences may give clues to the location of the erythrocyte binding regions. Using antisera to fusion proteins from the *P. knowlesi* α gene that recognize all *P. knowlesi* Duffy binding proteins, we have shown that the erythrocyte binding proteins have different specificities, only one of which binds to Duffy-positive human erythrocytes (5). Another protein recognized by these antisera binds specifically to erythrocytes from Old World monkeys, the natural host of *P. knowlesi*. The Duffy binding protein of *P. vivax* binds specifically to the Duffy blood group antigen of human erythrocytes, but not to monkey erythrocytes (3). The fact that the 5′ cysteine-rich region (region II) is conserved to the same degree between all four genes (Table 1) suggests that this region may be conserved for binding to the erythrocyte receptor. The similar degree of variability in region II may reflect differences in binding specificity. The homology of region III between *P. knowlesi* β and *P. vivax* may also be for similar receptor function for binding to the Duffy blood group antigen on human erythrocytes. Exploration of these questions must await studies with proteins recombinant between regions of the various genes.

Homology Between the Duffy Binding Protein and *P. falciparum* EBA175. The inclusion of the sialic acid binding protein of *P. falciparum* (EBA175) in the same gene family as the Duffy binding proteins is based on the similarity in gene structure (Fig. 1) and the amino acid homology in the 5′ and 3′ cysteine-rich regions (Fig. 2). The *P. falciparum* protein differs from the Duffy binding proteins in that it has two copies (F_1 and F_2) of the 5′ cysteine-rich region (17). The homology of the two 5′ cysteine-rich regions of *P. falciparum* and the equivalent regions of *P. vivax* and *P. knowlesi* are evident (Fig. 2). There are blocks of homology, including the cysteine residues and aromatic amino acids that encompass 325 amino acids in the 5′ cysteine-rich region. Twelve cysteine residues are present in roughly the same location in the repeat of *P. falciparum* motifs (F_1 and F_2) and the *P. vivax* and *P. knowlesi* proteins. The F_1 motif has one extra cysteine residue, and the F_2 motif has two extra cysteine residues that are not seen in the *P. vivax* and *P. knowlesi* proteins. Of nine tryptophan residues encoded by the *P. vivax* and *P. knowlesi* proteins, seven are aligned in F_1 and six are aligned in F_2. Only one nonaligned tryptophan is present in F_2, and none are present in F_1. If the tyrosine and phenylalanine are considered equivalent, they also align. In addition to the alignment of the dispersed cysteine residues and aromatic amino acids, there are blocks of homology for other amino acids (Fig. 2). In general, F_1 is more similar to the *P. vivax* and *P. knowlesi* proteins than it is to F_2. The 3′ cysteine-rich region of homology encompasses ≈104 amino acids (Fig. 2). The eight cysteine residues in the region align with only one space inserted in the *P. knowlesi* and *P. vivax* amino acid sequences. In addition to other amino acids, eight tyrosine/phenylalanine residues align.

An intron was not present between the signal sequence and the rest of the extracellular domain of the *P. falciparum* gene. Comparison of cDNA and genomic DNA sequences from the 3′ region of the *P. falciparum* gene demonstrated that the transmembrane region is encoded by one exon and that the cytoplasmic domain is encoded by two exons (Fig. 3*A*). The presence of an exon/intron division similar to the Duffy binding proteins suggests a common origin of the genes, because introns in malarial genes are uncommon. Many malarial genes contain no introns [e.g., circumsporozoite protein (18) and malaria surface antigen 1 (19)] or contain one intron between the signal sequence and the rest of the gene (reviewed in ref. 20). Some have more introns, but none have the same distribution as described for this group of genes. For example, a gene, AMA-1, encodes a protein that also has a signal sequence, a transmembrane region, and a 55-amino acid cytoplasmic tail, but has no introns (21, 22). As in the case of the erythrocyte binding proteins, AMA-1 localizes to the apical region of merozoites, but the organelle in which it is found has not yet been identified.

A third point in favor of placing the *P. falciparum* sialic acid binding protein in the same family as the Duffy binding proteins was the exon boundaries for the codons (Fig. 3*B*). The three exon boundaries at the 3′ region of the genes (before the transmembrane domain, after the transmembrane domain, and in the middle of the cytoplasmic domain) used the same distribution of codons on each side of the exons. If the codons are distributed randomly across the three exons, then the identical distribution in *P. falciparum* to the Duffy binding proteins could occur by chance alone only 1 time in 27 (⅓ × ⅓ × ⅓).

The Duffy binding protein and the sialic acid binding protein are both found in micronemes, an organelle at the invasive end of the parasite (5, 17). As the Duffy binding protein is not found on the surface of invasive merozoites (5), the protein is presumably released after contact with another erythrocyte receptor that mediates initial attachment and apical reorientation.

A

Exon 1	MKCNI...x_{1371}...TCFRKEAFS	Camp
Exon 2	SMPYYAGAGVLFIILVILGASQAKYQ	Camp, Dd2/NM
Exon 3	SSEGVMNENNENNNFLFEVTDNLDKLSNM	Camp, Dd2/NM
Exon 4	FNQQVQETNINDFSEYHEDINDIKFKI	Camp
	.....N..K	Dd2/NM

B

	1/2	2/3	3/4	4/5
Kα	AAG/GCT	G/GC	GA/T	G/AT
Kβ	AAG/GCG	G/GC	GA/T	*
Kγ	AAG/GCT	G/GC	GA/T	G/AT
V	AAG/GTA	G/AC	GA/T	A/AT
F	*	A/GT	AG/T	G/TC

FIG. 3. (*A*) Amino acid sequence of the sialic acid binding protein encoded by exons 2, 3, and 4 from the Camp strain and the Dd2/NM clone of *P. falciparum*. The amino acid sequence was deduced from the ORF of *P. falciparum* cDNA (Dd2/NM) and genomic DNA (Camp strain). The first five residues starting with the initiation methionine (M) and the final nine residues (amino acids 1386–1394) of exon 1 are from the previously published sequence (6). The sequence after these residues in Sim *et al.* (6) is now corrected by the sequence shown. Residues in the predicted transmembrane domain are underlined. The deduced amino acid sequence from the cDNA of Dd2/NM differed in that the last four amino acids of exon 4 were NFKK. The genomic and cDNA nucleotide sequences from which the splice sites for the exons were determined have been deposited in the Genbank data base (M93396–M93398). (*B*) Codon usage at the exon boundaries. The slash mark (/) represents the exon boundary with the codons presented for each side of the boundary. The numbers in each column refer to the exons of *P. knowlesi* α (Kα) gene and the analogous exons of *P. knowlesi* β (Kβ), *P. knowlesi* γ (Kγ), *P. vivax* (V), and *P. falciparum* (F). An asterisk (*) indicates that the analogous intron is missing (e.g., the *P. falciparum* intron between the signal sequence and the erythrocyte binding domain).

The structure of the binding domains may be of interest for vaccines and receptor-blocking therapy. Such studies require a knowledge of the binding domains in the parasite proteins. Identification of the binding domains may require erythrocyte binding studies with proteins recombinant between regions of the various genes. Certain difficulties must be overcome, however. First, transformation of malaria parasites has not been possible to date. Second, normal conformation of the cysteine-rich domain will probably require expression in a eukaryotic system (23). There has been no problem in eukaryotic expression systems obtaining localization to the cell surface or secretion with malarial genes that reach the surface without trafficking through invasion organelles of the parasite. The soluble antigen (24) and Pfs 25 (23) both come to the surface of eukaryotic cells. In contrast, the circumsporozoite protein that traffics to the surface through micronemes is not expressed on the surface of eukaryotic cells (25, 26) but remains localized in internal organelles. The Duffy binding proteins that also localize in micronemes may present similar problems.

An additional problem in expression in nonparasitic eukaryotic systems may be the requirement for proteolytic processing of the protein to create the active domains. For example, the *in vitro* translation product of the *P. knowlesi* Duffy binding proteins is larger than the membrane-bound proteins extracted from parasites, and the soluble protein that binds is of still lower molecular weight (5). If these are to be made in eukaryotic cells, then some ideas on potential binding domains may be critical for planning constructs. The homology regions between the *P. falciparum* sialic acid binding protein and the *P. vivax* and *P. knowlesi* Duffy binding proteins may assist in this design.

Regions of homology among *P. falciparum*, *P. vivax*, and *P. knowlesi* have been of great interest because of their wide evolutionary separation (10). This is especially true for molecules that may be under immune pressure, because those areas that are conserved may be motifs conserved for function. For example, the circumsporozoite proteins (CSPs) of *P. knowlesi* and *P. falciparum* are conserved in two regions, region I and region II (27). Peptides from CSP region I bind specifically to liver cells (28). Peptides from CSP region II also have cell adhesive properties (29) and sequence identity to a region in vertebrate adhesion proteins such as thrombospondin (30) and properdin (31). Thus, the original assumption that the two regions were conserved for function is consistent with the subsequent findings. It now remains to be tested if the regions of homology of the erythrocyte binding molecules between *P. falciparum* and *P. vivax*/*P. knowlesi* are conserved for binding function.

We thank Drs. David S. Peterson and David J. Bzik for discussions on the sequence analysis. Funds were provided in part by the University of Notre Dame.

1. Hadley, T. J., Klotz, F. W. & Miller, L. H. (1986) *Annu. Rev. Microbiol.* **40,** 451–477.
2. Haynes, J. D., Dalton, J. P., Klotz, F. W., McGinniss, M. H., Hadley, T. J., Hudson, D. E. & Miller, L. H. (1988) *J. Exp. Med.* **167,** 1873–1881.
3. Wertheimer, S. P. & Barnwell, J. W. (1989) *Exp. Parasitol.* **69,** 340–350.
4. Camus, D. & Hadley, T. J. (1985) *Science* **230,** 553–556.
5. Adams, J. H., Hudson, D. E., Torii, M., Ward, G. E., Wellems, T. E., Aikawa, M. & Miller, L. H. (1990) *Cell* **63,** 141–153.
6. Sim, B. K., Orlandi, P. A., Haynes, J. D., Klotz, F. W., Carter, J. M., Camus, D., Zegans, M. E. & Chulay, J. D. (1990) *J. Cell Biol.* **111,** 1877–1884.
7. Fang, X., Kaslow, D. C., Adams, J. H. & Miller, L. H. (1991) *Mol. Biochem. Parasitol.* **44,** 125–132.
8. Klotz, F. W., Orlandi, P. A., Reuter, G., Cohen, S. J., Haynes, J. D., Schauer, R., Howard, R. J., Palese, P. & Miller, L. H. (1992) *Mol. Biochem. Parasitol.* **51,** 49–54.
9. Orlandi, P. A., Klotz, F. W. & Haynes, J. D. (1992) *J. Cell Biol.* **116,** 901–909.
10. Waters, A. P., Higgins, D. G. & McCutchan, T. F. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 3140–3144.
11. Feinberg, A. P. & Vogelstein, B. (1984) *Anal. Biochem.* **137,** 266–267.
12. Dolan, S. A., Miller, L. H. & Wellems, T. E. (1990) *J. Clin. Invest.* **86,** 618–624.
13. Gilbert, W., Marchionni, M. & McKnight, G. (1986) *Cell* **46,** 151–153.
14. Kozak, M. (1987) *Nucleic Acids Res.* **12,** 857–872.
15. Saul, A. & Battistutta, D. (1990) *Mol. Biochem. Parasitol.* **42,** 55–62.
16. Robson, K. J. & Jennings, M. W. (1991) *Mol. Biochem. Parasitol.* **46,** 19–34.
17. Sim, B. K. L., Toyoshima, T., Haynes, J. D. & Aikawa, M. (1992) *Mol. Biochem. Parasitol.* **51,** 157–160.
18. Osaki, L. S., Svec, P., Nussenzweig, R. S., Nussenzweig, V. & Godson, G. N. (1983) *Cell* **34,** 815–822.
19. Holder, A. A., Lockyer, M. J., Odink, K. G., Sandhu, J. S., Moreno-Riveros, V., Nicholls, S. C., Hillman, Y., Davey, L. S., Tizard, M. L. V., Schwarz, R. T. & Freeman, R. R. (1985) *Nature (London)* **317,** 270–274.
20. Brown, H. & Coppel, R. L. (1991) *Mol. Biochem. Parasitol.* **49,** 99–110.
21. Peterson, M. G., Marshall, V. M., Smythe, J. A., Crewther, P. E., Lew, A., Silva, A., Anders, R. F. & Kemp, D. J. (1989) *Mol. Cell. Biol.* **9,** 3151–3155.
22. Waters, A. P., Thomas, A. W., Deans, J. A., Mitchell, G. H., Hudson, D. E., Miller, L. H., McCutchan, T. F. & Cohen, S. (1990) *J. Biol. Chem.* **265,** 17974–17979.
23. Kaslow, D. C., Isaacs, S. N., Quakyi, I. A., Gwadz, R. W., Moss, B. & Keister, D. B. (1991) *Science* **252,** 1310–1313.
24. Langford, C. J., Edwards, S. J., Smith, G. L., Mitchell, G. F., Moss, B., Kemp, D. J. & Anders, R. F. (1986) *Mol. Cell. Biochem.* **6,** 3191–3199.
25. Smith, G. L., Godson, G. N., Nussenzweig, V., Nussenzweig, R. S., Barnwell, J. & Moss, B. (1984) *Science* **224,** 397–399.
26. Kumar, S., Miller, L. H., Quakyi, I. A., Keister, D. B., Houghten, R. A., Maloy, W. L., Moss, B., Berzofski, J. A. & Good, M. F. (1988) *Nature (London)* **334,** 258–260.
27. 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., Roberts, D., Sanders, G. S., Reddy, E. P., Diggs, C. L. & Miller, L. H. (1984) *Science* **225,** 593–599.
28. Aley, S. B., Bates, M. D., Tam, J. P. & Hollingdale, M. R. (1986) *J. Exp. Med.* **164,** 1915–1922.
29. Rich, K. A., George, F. W., IV, Law, J. L. & Martin, W. J. (1990) *Science* **249,** 1574–1577.
30. Lawler, J. & Hynes, R. O. (1986) *J. Cell Biol.* **103,** 1635–1648.
31. Groundis, D. & Reid, K. B. M. (1986) *Nature (London)* **335,** 82–85.