

Two *Plasmodium falciparum* genes express merozoite proteins that are related to *Plasmodium vivax* and *Plasmodium yoelii* adhesive proteins involved in host cell selection and invasion

Julian C. Rayner*, Mary R. Galinski†, Paul Ingravallo‡, and John W. Barnwell*§

*Division of Parasitic Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention, Chamblee, GA 30341; †Emory Vaccine Center, Emory University, Atlanta, GA 30329; and ‡Schering-Plough Research Institute, Kenilworth, NJ 07033

Edited by Louis H. Miller, National Institutes of Health, Bethesda, MD, and approved June 15, 2000 (received for review October 29, 1999)

Two related *Plasmodium falciparum* genes and their encoded proteins have been identified by comparative analyses with *Plasmodium vivax* reticulocyte binding protein 2 (PvRBP-2). The *P. falciparum* genes have a structure which suggests that they may be the result of an evolutionary duplication event, as they share more than 8 kb of closely related nucleotide sequence but then have quite divergent unique 3' ends. Between these shared and unique regions is a complex set of repeats, the nature and number of which differs between the two genes, as well as between different *P. falciparum* strains. Both genes encode large hydrophilic proteins, which are concentrated at the invasive apical end of the merozoite and are predicted to be more than 350 kDa, with an N-terminal signal sequence and a single transmembrane domain near their C termini. Importantly, they also share gene structure and amino acid homology with the *Plasmodium yoelii* 235-kDa rhoptry protein family, which is also related to PvRBP-2. Together these *Plasmodium* proteins define an extended family of proteins that appear to function in erythrocyte selection and invasion. As such, they may prove to be essential components of malaria vaccine preparations.

Parasites of the genus *Plasmodium* are estimated to cause between 300 and 500 million cases of malaria, the majority of which are caused by *Plasmodium vivax* and *Plasmodium falciparum* (1). *Plasmodium* parasites have a complex life cycle involving a series of developmental stages in both mosquitoes and mammals, but the clinical manifestations of malaria are all caused by the asexual blood stage. Merozoites, ovoid cells with an apical prominence at one end, invade red blood cells (RBCs), wherein they undergo a growth and multiplication phase (schizogony). The resulting schizont eventually ruptures the RBC, releasing newly formed merozoites for subsequent rounds of invasion.

How merozoites identify and invade RBCs has long been a focus of research (2, 3). The merozoite first attaches to a RBC at any point on its surface, and then reorients to bring its apical end into contact with the RBC. The initial attachment stages are reversible, and merozoites can disassociate and attach to a new potential target cell. The subsequent steps are irreversible, and involve the formation of an electron-dense adhesion zone between the apical end of the merozoite and the RBC. This zone then moves around the merozoite toward its posterior end, with a concurrent invagination of the RBC membrane and entry of the merozoite. This cascade of molecular events also involves release of proteins from the rhoptries and micronemes, specialized apical organelles central to the invasion process.

The molecular adhesion details behind this tantalizing outline are sketchy. The merozoite surface proteins (MSPs), several of which have been described in a number of species of *Plasmodium*, together make up a structurally complex coat around the outer membrane of the merozoite and may have a role in the initial reversible adhesive interaction between the merozoite and the RBC (4). The *P. vivax* and *Plasmodium knowlesi* Duffy

binding proteins (DBPs) and their *P. falciparum* orthologue, EBA-175 (erythrocyte binding antigen 175) have been shown to bind well-defined glycoprotein motifs on the RBC membrane (5–7). Studies of the *P. knowlesi* DBP have implicated these related microneme proteins in the formation of the adhesion zone that characterizes the beginning of the entry phase of invasion (8). Apical membrane antigen 1 (AMA-1) and MAEBL, a chimeric protein containing features of both the DBP/EBA-175 proteins and AMA-1, may also play a role in RBC invasion, but their precise functions are as yet unclear (9, 10).

Further complexity is due to the fact that some proteins recognize subpopulations of RBCs. In *P. vivax*, the reticulocyte binding proteins (PvRBP-1 and PvRBP-2) were identified through their ability to bind reticulocytes, the immature subpopulation of the RBC pool that *P. vivax* merozoites preferentially invade (11, 12). PvRBP-2 is distantly related to a 235-kDa rhoptry protein of the rodent parasite *Plasmodium yoelii*, which adheres to mouse erythrocytes, and is encoded by a multigene family (13–15). Passive transfer of specific monoclonal antibodies (16) or active immunization with purified protein (17) restricts parasites of a normally virulent strain to a self-limiting infection of reticulocytes, suggesting that, like the PvRBPs, the *P. yoelii* 235-kDa proteins play a role in host cell selection.

Here we present two *P. falciparum* orthologues of PvRBP-2, which we have termed *Plasmodium falciparum* RBP-2 homologues *a* and *b* (PvRBP2-Ha and -Hb). We propose that together with the RBP-2 of *P. vivax* and the *P. yoelii* 235-kDa proteins, they constitute an important *Plasmodium* family that plays a pivotal role in the early phases of merozoite invasion.

Materials and Methods

Parasites. The FVO strain of *P. falciparum* was obtained from a line adapted to *Aotus* monkeys and then established in *in vitro* culture (18, 19). Material from the *P. falciparum* 3D7 strain was obtained from *in vitro* culture in human O⁺ erythrocytes.

DNA Cloning and Analysis. *P. falciparum* FVO or 3D7 genomic DNA (gDNA) was prepared as described (11). DNA probes were

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: PvRBP, *Plasmodium vivax* reticulocyte binding protein; PvRBP2-H, PvRBP2 homologue; TMD, transmembrane domain; gDNA, genomic DNA.

Data deposition: Partial sequences for PvRBP2-Ha and -Hb have been deposited in the GenBank database (accession nos. AF196347 and AF196348).

§To whom reprint requests should be addressed at: Division of Parasitic Diseases, Centers for Disease Control and Prevention, MS F-13, 4770 Buford Highway, Chamblee, GA 30341. E-mail: wzb3@cdc.gov.

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.

Article published online before print: *Proc. Natl. Acad. Sci. USA*, 10.1073/pnas.160469097. Article and publication date are at www.pnas.org/cgi/doi/10.1073/pnas.160469097

labeled with [α - 32 P]dATP by random priming (Prime-it Kit; Stratagene). A λ ZAPII (Stratagene) gDNA library was constructed from *P. falciparum* FVO strain DNA by using mung bean nuclease methodology (20) with a range of formamide concentrations (30–45%) to allow digestion around (\approx 30%) as well as within (\approx 35–45%) genes. This library was screened with a PCR fragment amplified from *P. falciparum* FVO gDNA by using primers 5'-TCTAGGATTGTCGATGAAGCT-3' and 5'-ATTAGTATGATCTGTTCTCCTGA-3'. DNA sequencing was performed by the dideoxynucleotide method, both with Sequenase (USB) and with automated dye termination chemistry through commercial sequencing firms.

Long-range PCR used the Expand PCR system (Roche). Forward primers were SF7 (5'-GAAGACCACACTATTTGTAGC-3'), SF5 (5'-GCAGACAAAATGAAGAAAG-3'), and SF3 (5'-AGTGAAGTTGAGCAAGA-3'). Reverse primers were AR2 (5'-GAACATCATCATTCGGTTC-3') and BR2 (5'-TACTAGCATCACGTTGGTC-3'). Southern blot probes were as follows: Sh1 (a 1.2-kb *Xba*I fragment from clone 20.2b), Sh2 (an 850-bp PCR fragment amplified from FVO gDNA by using PCR primers 5'-TTGATTATGTAGATGTGGACG-3' and 5'-TACCCCTTTTAATGAACGG-3'), A (a 600-bp PCR fragment amplified from FVO gDNA by using primers 5'-CACACATAAAGCTACACTTCC-3' and 5'-CGAGTTAGCAGAAAGAGAACA-3'), and B (an 800-bp *Nco*I/*Eco*RV fragment from clone 18.2). PCR was performed with recombinant *Taq* DNA polymerase (GIBCO/BRL).

Databank BLAST-based searches were conducted by using MACVECTOR through Entrez (National Center for Biotechnology Information, National Institutes of Health) and directly from the Sanger Centre Internet site at <http://www.sanger.ac.uk>. Sequence data for *P. falciparum* chromosome 13 were obtained from http://www.sanger.ac.uk/Projects/P_falciparum/. Sequencing of *P. falciparum* chromosome 13 was accomplished as part of the Malaria Genome Project with support by The Wellcome Trust. Contig names have changed since the submission of this paper. The contigs discussed here can be retrieved from the GnuZIP archive MAL13.Contigs.4.4gz, available at ftp://ftp.sanger.ac.uk/pub/databases/P.falciparum_sequences/UNFINISHED_SEQUENCES/MAL13.

RNA Analysis. *P. falciparum* FVO strain RNA was prepared from mature schizonts by using ToTALLY RNA Isolation Kit (Ambion) and treated with RNase-free DNase (Ambion). cDNA was synthesized and purified with a 5' RACE kit (GIBCO/BRL) and the gene-specific primers AR2 and BR2. Nested PCR was carried out using forward primer SF1 (5'-CGAAAAGAATATCAAACACG-3') and reverse primer AR2 or BR2 in the primary reactions, then forward primer SF2 (5'-AGGATCTTGGTGATGA-3') and reverse primer AR1 (5'-AGGTTTATATCGACGAGTC-3') or BR1 (5'-CGCTTCTGTAAATCTACTG-3') in the secondary reactions.

Protein Analysis. For immunoblots, protein samples made by boiling partially purified late and segmented schizonts in reducing SDS/PAGE sample buffer were run on 6% polyacrylamide gels (37.5:1 acrylamide to bisacrylamide), transferred to nitrocellulose membranes (Schleicher and Schuell), probed with primary antisera, and developed with an ECL kit (Amersham Pharmacia Biotech). Immunofluorescence assays were performed on air-dried thin smears of *P. falciparum* (FVO strain)-infected erythrocytes, taken from culture and fixed with ice-cold acetone, with secondary antibodies [fluorescein isothiocyanate (FITC)-linked goat anti-rabbit IgG and tetramethylrhodamine isothiocyanate (TRITC)-linked goat anti-rat IgG] from Zymed Laboratories. Rabbit antisera were raised against four glutathione *S*-transferase (GST) fusion proteins by using the pGEX system (Pharmacia). All fusions were made by amplifying

fragments from *P. falciparum* FVO gDNA, with primers 5'-ATTGATTATGTAGATGTGGACG-3' and 5'-TACCCCTTTAATGAACGG-3' (S1); 5'-ATCCGTTCAATAAAGGGG-3' and 5'-AGGATTTTCTGGAGTTATTCCG-3' (S3); 5'-AGAAAGAAAATCGAGTTAGCAG-3' and 5'-CTATTAGTATGATCTGTTCTCCTG-3' (A); and 5'-ATCCAAGAATTAGAAGAGCAA-3' and 5'-TTTGTTATGGTTGAATACCT-3' (B). Fusion proteins were purified from *Escherichia coli* by using standard methods on glutathione-Sepharose 4B (Pharmacia) and injected into New Zealand White rabbits in Freund's complete and incomplete adjuvant. The antisera were checked for cross-reactivity, and all reacted only with the fusion protein against which they were raised.

Gene and deduced protein sequences were analyzed and compared by using the MACVECTOR 6.5 DNA/Protein analysis module (Oxford Molecular). Potential signal peptide sequences were analyzed by using the weight matrix method of von Heijne (21) and a neural network method (22) accessible at <http://www.cbs.dtu.dk/services/SignalP>.

Results

Identification of *P. falciparum* PvRBP-2 homologues. The 3' portion of the *PvRBP-2* gene, clone 16-6 (12), hybridizes to *P. falciparum* (FVO strain) gDNA on Southern blots under reduced stringency (data not shown). Subsequent BLASTp searching of the GenBank database with 16-6 translated sequence revealed a 2.4-kb *P. falciparum* (3D7 strain) DNA sequence that includes an open reading frame (ORF) of 500 amino acids designated *P. falciparum* unknown membrane protein (PfUMP, accession no. L04159). Reprobing gDNA Southern blots with a corresponding *P. falciparum* FVO-strain PCR-amplified fragment revealed the same pattern of hybridization bands as was obtained with 16-6 and also suggested the presence of two homologous genes (data not shown). This PCR fragment was therefore used to screen a *P. falciparum* (FVO strain) gDNA library in search of possible *PvRBP-2* homologues.

Two initial positive clones, 6A1 and 4D1, and subsequent related clones (1.1 and 18.2) were sequenced in their entirety (Fig. 1A). The latter clones were isolated by using the 5'-most 500 bp of 6A1 as a probe. 6A1 and 1.1 include a long ORF spanning all of the sequence, whereas 18.2 and 4D1 both have stop codons near their 3' ends. The clones fell into two groups, encoding two related genes with an unusual relationship to each other, being almost identical at their 5' ends (subsequently called the shared region), but divergent at their 3' ends (subsequently called the unique regions). In 1.8 kb of sequence shared by clones 1.1 and 18.2, there is only a single nucleotide difference, causing a V to A substitution in the predicted amino acid sequence. Intriguingly, where the sequences diverge, both genes encode a mixture of amino acid repeat motifs, which are discussed further below. The C termini after the repeats are clearly divergent (Fig. 1B). However, CLUSTAL alignment suggests that they are still related and both have homology to the *PvRBP-2* protein and the *P. yoelii* 235-kDa family member E-3 (see below). The hybridization data, sequence homologies, and further supporting data described below have led us to term these two genes *Plasmodium falciparum* reticulocyte binding protein 2 homologues *a* and *b* (*PfRBP2-Ha* and *-Hb*), corresponding to the 6A1/1.1 and 4D1/18.2 sequences, respectively.

The *PfRBP2-Ha* and *PfRBP2-Hb* Genes Share >8 kb of Closely Related Sequence Before Diverging. The presence of an extensive shared region made it difficult to identify the full *PfRBP2-Ha* and *PfRBP2-Hb* gene sequences by conventional library screening, as any clone that did not include some 3' unique sequence was impossible to assign to one gene or the other. An example is clone 20.2b (Fig. 1A). BLAST searching of the Sanger Centre database facilitated this process, revealing several contigs from *P. falciparum* (3D7 strain) chromosome 13 that contained

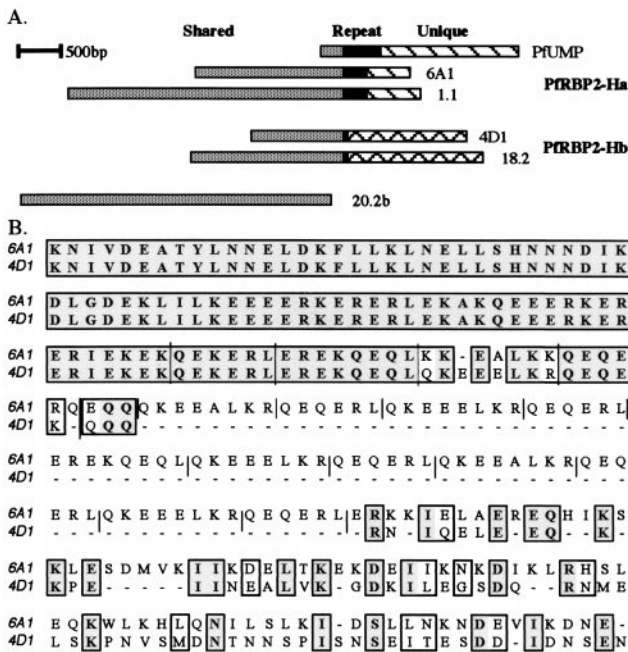


Fig. 1. *PfrBP2-H* genes are composed of shared, repetitive, and unique domains. (A) Schematic diagram of *PfrBP2-H* DNA clones and the PfUMP GenBank sequence. Sequences are grouped according to their relationship to *PfrBP2-Ha* (PFUMP, 6A1, 1.1) or *-Hb* (4D1, 18.2). Clone 20.2b could be a fragment of either. Regions shared by all sequences (gray), repeat regions (black), and regions unique to *PfrBP2-Ha* or *-Hb* (hatched) are noted. (B) CLUSTAL alignment of predicted amino acid sequence from clones 6A1 and 4D1. Areas of identity and similarity are boxed, with identical residues highlighted in gray. Not all of the 6A1 or 4D1 sequences are shown. Solid lines indicate the boundaries of repeated motifs.

PfrBP2-H sequences. Contig 08813 (2.2 kb) contains the *PfrBP2-Ha* unique region, whereas contig 04278 (11.5 kb) contains the *PfrBP2-Hb* unique region (Fig. 2A). Comparisons of these 3D7 contigs with our FVO gDNA clones showed only one nucleotide difference in the unique region of *PfrBP2-Hb* and none in the unique region of *PfrBP2-Ha*.

As discussed above, our FVO sequence data revealed only one difference between *PfrBP2-Ha* and *-Hb* in 1.8 kb of closely related sequence 5' to the repeat region and divergent 3' unique domains. Comparing the shared region from our largest FVO *PfrBP2-Ha* clone, 1.1, with the corresponding region of the 3D7 *PfrBP2-Hb* contig 04278 revealed only seven nucleotide differences in over 3.2 kb. It remains to be determined whether these are true allelic differences or sequencing errors not yet detected in the preliminary genome project data. One such difference introduces a stop codon into the ORF of *PfrBP2-Hb* (contig 04278) before the repeat region, and we have determined by cloning and sequencing this region from 3D7 gDNA that this particular base difference indeed represents a sequence error in the genome database (data not shown). What is clear, however, is that *PfrBP2-Ha* and *-Hb* share at least 3.2 kb of extremely highly related sequence.

To determine how far this striking region of shared sequence extends, long-range PCR was performed on FVO and 3D7 gDNA by using a series of forward primers (SF3, SF5, and SF7) designed from the 5' end of the 04278 sequence paired with a reverse primer (AR2 or BR2) from either the *PfrBP2-Ha* or *-Hb* unique region (Fig. 2A). Similar-sized PCR fragments, between 5.5 kb and 9 kb, were amplified with the constant forward primers paired with either reverse primer (Fig. 2B and data not shown). These could be digested with restriction enzymes spe-

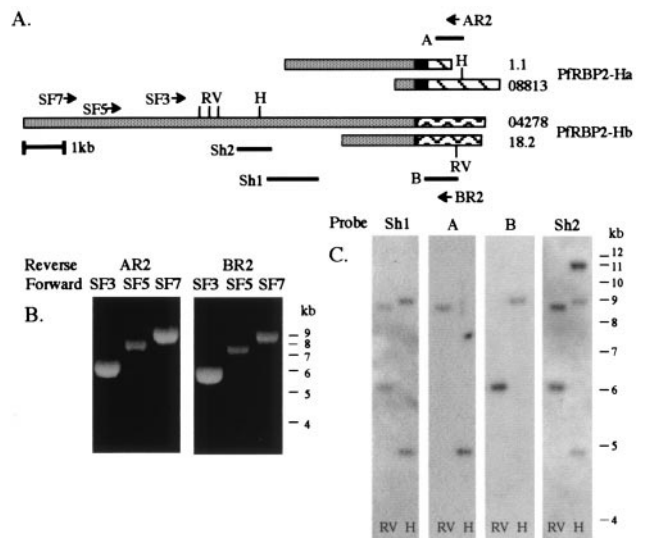


Fig. 2. *PfrBP2-Ha* and *-Hb* share >8 kb of common sequence. (A) *PfrBP2-Ha* (1.1 and 08813) and *-Hb* (18.2 and 04278) schematic. Gray, black, and hatched boxes define shared, repeat, and unique regions, respectively. Primers used in long-range PCR are marked (SF7, SF5, SF3, AR2, and BR2), as are *EcoRV* (RV) and *HindIII* (H) restriction sites and fragments used as probes in Southern blots (Sh1, Sh2, A, and B). (B) Long-range PCR performed on FVO DNA by using reverse primers from the unique regions (AR2 or BR2) paired with forward primers designed from contig 04278 (SF3, SF5, and SF7). (C) Southern blots of *EcoRV* (RV) and *HindIII* (H)-digested FVO DNA, probed with fragments Sh1, A, B, and Sh2. DNA 1-kb ladder markers from GIBCO/BRL are indicated.

cific to either the *PfrBP2-Ha* or *-Hb* unique regions, confirming that they were specific products and not the result of mispriming.

Two *PfrBP2-H* genes with an extended shared sequence were also evident in Southern blot experiments. *HindIII* cuts in the unique region of *PfrBP2-Ha* but not *-Hb*, and *EcoRV* cuts in the unique region of *PfrBP2-Hb* but not *-Ha*. As predicted, a probe in the shared region (probe Sh1) hybridizes to two bands in gDNA digested with either *EcoRV* or *HindIII* (Fig. 2C), whereas probes from the unique regions (probes A and B) detected only one or the other of these bands. In contrast, probe Sh2, which anneals slightly upstream of Sh1 and overlaps the shared *HindIII* site (Fig. 2A), hybridizes to three fragments in *HindIII*-digested DNA. The two smaller bands are clearly the 3' fragments of the two genes detected with probe Sh1, whereas the larger band of 11 kb represents two similarly migrating fragments containing the 5' end of both genes. The fact that its intensity is twice as strong as either of the smaller bands and that no further bands were visible even after prolonged exposure supports this conclusion.

The Southern blot data suggest that the region shared by *PfrBP2-Ha* and *-Hb* extends past the 5' end of contig 04278 for 5.5 kb and is at least 14.5 kb in total, while the long range PCR confirms at least 8 kb of overlap. The precise sequence of this shared region will be clarified further as the genome project is finalized.

***PfrBP2-Ha* and *-Hb* Are Homologous to *PvrBP2* and the *P. yoelii* p235 Proteins.** The 04278 contig contains a major ORF of over 9 kb (once the database sequencing error noted above is corrected). Upstream of this ORF we noticed a potential intron of 214 bp and initial exon of 57 bp (Fig. 3B). The presence of the intron and transcription of the short exon 1 were confirmed by using reverse transcription-PCR (data not shown). The SF7 primer used in long-range PCR (Fig. 2) is situated at the start of exon 1, suggesting that the sequence shared by the *PfrBP2-H* genes includes their translation start sites.

The *PfrBP2-H* gene structures are thus predicted to consist of

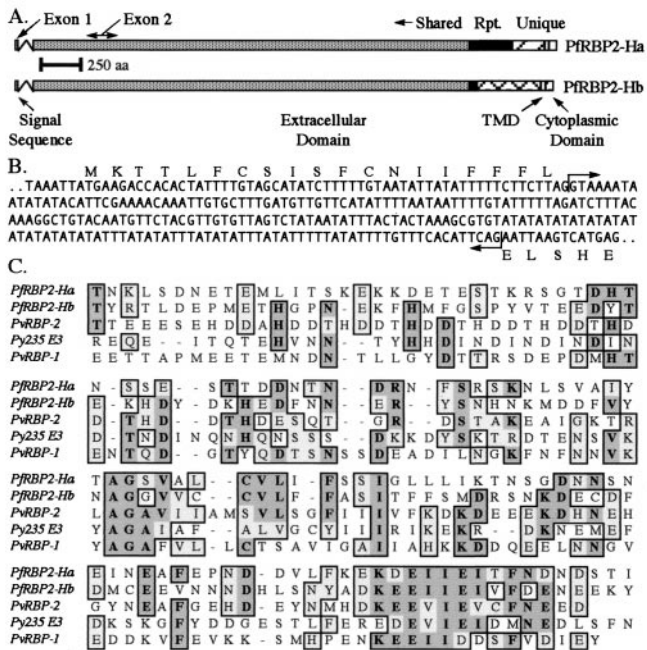


Fig. 3. Structure of the *PfrBP2-H* genes and their relationship with each other and with other *Plasmodium* proteins. (A) Schematic of the *PfrBP2-H* genes, highlighting structural features as marked. Shared regions (gray), repeats (black), and unique regions (hatched) are also noted. (B) Sequence of the 5' end of the *PfrBP2-H* genes, showing the boundaries of the intron (arrows) and sequence of the short exon 1 that encodes a signal peptide. (C) CLUSTAL alignment of the C terminus of the *PfrBP2-H* proteins, the *P. vivax* reticulocyte-binding proteins, and the E-3 member of the *P. yoelii* p235 family. Identical residues are highlighted in gray and similar residues in light gray.

a short exon 1, a short intron, and then a large exon 2 encoding the majority of the protein (Fig. 3A). This is highly reminiscent of the *PvRBP-1* and *-2* genes (11, 12). Like the *PvRBP* genes, the *PfrBP2-Hb* (04278) exon 1 encodes a hydrophobic core flanked by charged amino acids, which yields high probability scores for signal peptide assessment (21, 22). The remaining large portion of the protein is highly hydrophilic, with a putative TMD and cytoplasmic domain at the C terminus. This deduced model also predicts that both of the *PfrBP2-H* genes encode similar-sized large proteins, with estimated molecular masses of 370 kDa and 383 kDa for *PfrBP2-Ha* and *-Hb*, respectively.

There is low but significant amino acid homology between the *PfrBP2-Hb* (04278) ORF deduced sequence and both the *PvRBPs* and the *P. yoelii* 235-kDa rhoptry proteins, with CLUSTAL analysis showing homology scores (identity plus similarity) of 38% for *PvRBP-1*, 44% for *PvRBP-2*, and 42.5% for the E3 and E8 *P. yoelii* 235-kDa rhoptry proteins. However, individual regions show a much stronger homology ($\approx 50\%$), in particular at the C terminus (Fig. 3C) and over a 500 amino acid region previously reported to show homology between *PvRBP-2* and *Py 235* proteins (14). Using the 04278 translated sequence in tBLASTN searches suggests that it is more closely related to *PvRBP-2* and the *P. yoelii* 235-kDa proteins than it is to *PvRBP-1* (data not shown).

The Repeated Region of *PfrBP2-Ha* and *-Hb* Differs Between These Two Proteins and in Their Counterparts in Various *P. falciparum* Strains. The repeated region that marks the boundary between the highly conserved 5' shared domain and the divergent 3' unique domains was the only aspect of the sequencing project contigs that differed markedly from our gDNA clones. When a code is used for the consensus repeat motifs (Fig. 4), all sequences began with a

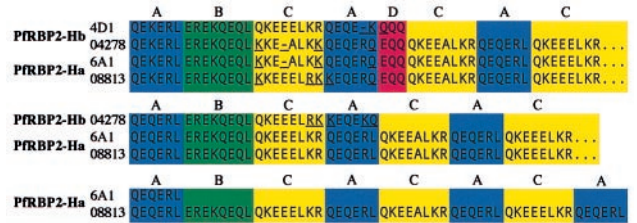


Fig. 4. Alignment of *PfrBP2-H* repeated motifs. Clones 4D1 and 6A1 were derived from *P. falciparum* strain FVO, whereas 04278 and 08813 sequences originate from strain 3D7. One-letter code refers to sequences as follows: A, QE(K/Q)ERL (blue); B, EREKQEQL (red); C, QKEE(E/A)LKR (black); and D, EQQ (green). Divergences from these consensus sequences are noted with underlines.

conserved pattern of repeats, ABCAD. The FVO *PfrBP2-Hb* sequence has no further repeats, whereas the 3D7 *PfrBP2-Hb* sequence has an extension reading CACABCA. The FVO *PfrBP2-Ha* also has a shorter repeat pattern than its 3D7 counterpart, reading CACABCACACA after the core consensus repeats, where the 3D7 sequence has the arrangement of CACAB-CACACABCACACA. We cloned and sequenced the repeated regions from 3D7 gDNA and confirmed these differences (data not shown). Preliminary analysis of other strains of *P. falciparum* suggests that such variation is widespread, including variation in repeat number between strains in a single infected individual (J.C.R. and J.W.B., unpublished observations). This repeat region may vary between strains in a manner similar to that of other well defined repeat regions in the *P. falciparum* genome, such as in the *MSP1* gene (23).

***PfrBP2-Ha* and *-Hb* Are Both Transcribed, Translated, and Expressed at the Apical Pole of Merozoites.** To show that both the *PfrBP2-Ha* and *-Hb* genes are transcribed, cDNA specific to each was generated by reverse transcription of total schizont stage RNA from FVO parasites by using reverse transcription primers designed from their unique regions. Nested PCR was then performed with these cDNA as templates and reverse primers specific to one or the other unique region (Fig. 5A). A PCR product could be generated with *PfrBP2-Ha* unique region reverse primers on cDNA synthesized with a primer specific to *PfrBP2-Ha* (Fig. 5B, lane 3). Similarly, a PCR product was generated with *PfrBP2-Hb* unique region reverse primers on cDNA specific to *PfrBP2-Hb* (lane 6). Crucially, no PCR product was produced with *PfrBP2-Ha* reverse primers on *PfrBP2-Hb* cDNA (lane 4), or with *PfrBP2-Hb* primers on *PfrBP2-Ha* cDNA (lane 5), indicating the absence of gDNA contamination. No PCR product was produced if reverse transcriptase was omitted at the cDNA synthesis step (data not shown). Both *PfrBP2-Ha* and *PfrBP2-Hb* are therefore clearly transcribed in schizonts.

Antisera were raised against several regions of *PfrBP2-Ha* and *-Hb* expressed as glutathione *S*-transferase (GST) fusion proteins. Two GST fusion proteins, S1 and S3, contained fragments of the shared region, while two others, A and B, contained the unique regions of *PfrBP2-Ha* and *-Hb*, respectively (Fig. 5A). Immunoblot analyses using a number of parasite protein preparations suggested that the *PfrBP2-H* proteins are subject to extensive and rapid proteolytic degradation, as several bands were detected that were not present in uninfected erythrocyte controls (Fig. 5C and data not shown). The sizes of these bands were not consistent between preparations, but antisera against both unique regions and the shared S3 sequence recognize high molecular mass bands of approximately 370 kDa in all samples (arrow in Fig. 5C). This size corresponds with the molecular mass predicted for each of the *PfrBP2-H* proteins.

To localize the *PfrBP2-H* proteins, immunofluorescence

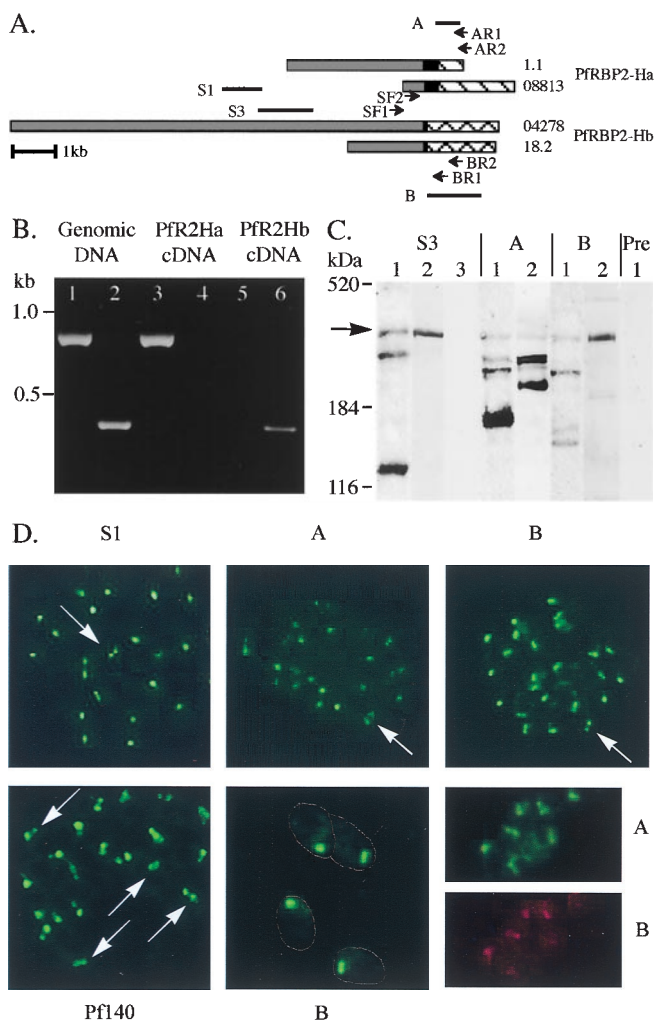


Fig. 5. *PfrBP2-Ha* and *-Hb* are expressed at the apical end of merozoites. (A) *PfrBP2-Ha* (1.1 and 08813) and *-Hb* (18.2 and 04278) schematic. Gray, black, and hatched boxes define shared, repeated, and unique regions, respectively. Primers used in reverse transcription-PCR are marked (SF1, SF2, AR1, AR2, BR1, and BR2), as are the fragments of the genes expressed to raise antisera (S1, S3, A, and B). (B) Nested PCR was performed by using primers specific to *PfrBP2-Ha* (SF1/AR2 followed by SF2/AR1, lanes 1, 3, and 5) or *-Hb* (SF1/BR2 followed by SF2/BR1, lanes 2, 4, and 6), with either FVO DNA (lanes 1 and 2) or cDNA as templates. The cDNA was generated from schizont stage total RNA and primers specific to *PfrBP2-Ha* (lanes 3 and 4) or *PfrBP2-Hb* (lanes 5 and 6) for reverse transcription reactions. (C) Western immunoblots prepared with anti-S3 (S3), anti-*PfrBP2-Ha* (A), anti-*PfrBP2-Hb* (B), or preimmune (Pre) antiserum as detection agent. Lanes contain either *P. falciparum* mature schizont SDS extracts (1 and 2) or uninfected erythrocyte ghost preparations (3). Protein size standards from GIBCO/BRL and Coomassie-stained apolipoprotein B (520 kDa) are noted. (D) Immunofluorescence assays performed with antisera against fragments of the shared domain (S1), fragments of the unique regions, anti-*PfrBP2-Ha* (A) and anti-*PfrBP2-Hb* (B), and the rhoptry antigen Pf-140, on air-dried, acetone-fixed, *P. falciparum*-infected erythrocytes. The lower center panel (B) is at a higher magnification ($\times 3,750$; others are $\times 1,250$), with the outline of the merozoites marked. The lower right shows costaining of single merozoites with rabbit anti-A serum (A) and rat anti-B serum (B).

assays were performed on early to late *P. falciparum* schizonts and free merozoites (Fig. 5D and data not shown). Each of the *PfrBP2-H* antisera, whether raised against fragments of the shared domain, anti-S1 and anti-S3, or against the unique regions, anti-A and anti-B, yielded a punctate staining pattern, generally with a single dot of fluorescence, although occasionally

two foci were visible in a single merozoite (marked by arrows in Fig. 5D Upper). This double-dot pattern is reminiscent of proteins such as the Pf140 protein (24), which are localized to the paired rhoptry organelles (Fig. 5D Lower Left). It should be noted, however, that the appearance of paired foci predominates with antiserum against the Pf140 protein, but is not so typical with the *PfrBP2-H* antisera. Higher magnification (Fig. 5D Lower Center) reveals that the staining is concentrated at the apical end of the merozoites. Costaining infected cells with rabbit anti-A serum and rat anti-B serum suggests that single merozoites express both *PfrBP2Ha* and *-Hb*, because individual merozoites were labeled with antisera against both unique domains (Fig. 5D Lower Right).

Discussion

We have identified and characterized *P. falciparum* homologues (*PfrBP2-Ha* and *-Hb*) of the *P. vivax* *RBP-2* gene. Like the *PvRBP-2* gene, the *PfrBP2-H* genes are expressed in the merozoite stage of the parasite and are predicted to encode large type I membrane-anchored proteins that localize to the apical pole.

The structure of the *PfrBP2-Ha* and *-Hb* genes, with 5' shared sequence, junctional repeated motifs, and 3' unique regions, is unusual and suggests that the two genes may have arisen as a result of a duplication event, followed by evolutionary drift of the 3' ends. The nature and extent of this duplication will become more evident with the completion of the *P. falciparum* genome project, but it would appear that it is an evolutionarily distant event, because both genes are present in *P. falciparum* strains from distinct geographical locations (J.C.R. and J.W.B., unpublished observations). Regions of almost identical nucleotide sequence have been observed previously in *Plasmodium*—for example, several *Plasmodium* species contain two identical copies of the EF-1 α gene (25). Repeated sequences within genes are also commonplace in *Plasmodium*, although those observed here have a unique cadence of alternating consensus motif units. What is unusual is that these repeats are strikingly placed between the strongly conserved predominant part of these genes and the more loosely maintained unique region that includes the TMD and cytoplasmic domains.

The general structure of the *PfrBP2-H* proteins, with a signal sequence and an extensive extracellular domain followed by a putative TMD and short cytosolic domain is the same as that of *PvRBP-2* and the two *P. yoelii* 235-kDa rhoptry proteins so far characterized (11–13, 26). Their two-exon gene structures are also similar. Although the overall amino acid homology between these proteins is not high, there are blocks of outstanding homology between the *PvRBP-2*, *Py235*, and *PfrBP2-H* proteins that are suggestive of a familial relationship. The identification here of domains with notable significant homology amidst divergent sequence mimics the familial relationship of the *P. knowlesi* and *P. vivax* DBPs and the *P. falciparum* EBA-175 (5). The rodent and each of the human malaria parasite species form evolutionarily distant clades, so considerable differences between members of the family could be expected, especially if proteins are under immune pressure (27). However, the similarities in gene structure, amino acid homology, and biology noted here together support the premise that the *PfrBP2-Hs*, *PvRBP-2*, and the *Py235* family members are related. We have not included *PvRBP-1* in this grouping, although it does show some similarity to both *PfrBP2-Hs*, as we have detected and characterized a distinct *PvRBP-1* homologue in *P. falciparum* (J.C.R., M.R.G., and J.W.B., unpublished observations).

There are clearly some features of this gene family that differ between the human parasites *P. falciparum* and *P. vivax* and the rodent parasite *P. yoelii*. There are up to 50 members of the 235-kDa protein gene family in *P. yoelii* (14), and the two so far reported are in fact very similar (26). Cross-hybridization suggests a large family of related genes is also present in the rodent

malaria *Plasmodium berghei* (14). Furthermore, each merozoite originating from a single *P. yoelii* schizont apparently expresses a distinct member of the family (28), a mechanism that may help the parasite evade the host immune response or facilitate the use of different host cell niches (29). Does a similar mechanism operate in *P. vivax* and *P. falciparum*? It would appear not, because both *PvRBP-1* and *-2* hybridize as single-copy genes and only two *PfRBP2-H* genes have been detected. We certainly do not rule out the possibility that the *P. vivax* and *P. falciparum* genomes may contain other more distantly related homologues, but such a scenario would clearly be different from the large closely related families found in *P. yoelii* and *P. berghei*. Given the apparent lack of a large multigene family in *P. falciparum*, the mechanism of transcriptional control described for the *P. yoelii* 235-kDa proteins would be of less relevance. Besides, our immunofluorescence assay data indicate that a single merozoite expresses both PfRBP2-Ha and -Hb.

What then is the function of PvRBP-2 and its orthologues? The PvRBPs were identified in *P. vivax* by their ability to bind to reticulocytes and are thus predicted to function in selecting reticulocytes for invasion by *P. vivax* merozoites (11). The *P. yoelii* 235-kDa protein family has also been shown to play a role in host cell selection (16, 17). In this case, however, the 235-kDa ligands appear to select mature normal erythrocytes, because neutralizing antibodies confine *P. yoelii* merozoite invasion to young RBCs (16). The PvRBPs have also been implicated in regulating invasion (30). Merozoites can disassociate from potential host cells up to the point that an electron-dense junction forms between the apical pole and the RBC. The case of *P. vivax* and its preferential invasion of reticulocytes implies the existence of a control step between initial random attachments and the irreversible cascade of events leading to merozoite entry. In this way junction formation and the later stages of invasion would be triggered only after a correct host cell has been selected, and unproductive invasion attempts into inappropriate host cells would be avoided (11, 30). The PvRBPs are prime candidates for such a triggering role, given their location at the apical end of the merozoite, their extracellular structures that are predicted to bind target cells, and their short intracellular domains that could undergo conformational change upon binding and thus transmit a signal to the merozoite cytoplasm. It follows that the PfRBP2-Hs may participate in a similar scheme in *P. falciparum*

invasion of its primary host cell, the mature normocyte population of erythrocytes. The presence of two closely related PfRBP2-H proteins with almost identical extracellular domains but differences in their TMD and cytoplasmic domains raises the possibility that they may bind the same ligand but communicate different biological signals to the merozoite cytoplasm.

Given this hypothesis, and the localization of the PfRBP2-Hs at the invasive end of merozoites, we carried out extensive erythrocyte binding assays (EBAs), using labeled *P. falciparum* culture supernatants to establish whether the PfRBP2-Hs binds to RBCs, but with inconclusive results (data not shown). This work however, did corroborate our immunoblot data, that the PfRBP2-Hs are extensively and rapidly degraded, as full-length PfRBP2-H product was not detectable in any of the culture supernatants we obtained, although fragments of various sizes were detected. Such cleavage could easily be removing adhesive domains or disrupting tertiary structure, so our lack of consistent results in no way rules out the possibility that the PfRBP2-Hs bind to RBCs.

In summary, we have identified two related *P. falciparum* proteins (PfRBP2-Ha and -Hb), which are expressed at the invasive apical end of the merozoite, and which, on the basis of gene structure and amino acid homology, are related to the *P. vivax* RBP-2 and *P. yoelii* 235-kDa proteins. We propose that these proteins, as previously shown for their orthologues, play an important adhesion function and possibly also a signaling role in the early stages of merozoite invasion. Comparative studies across species will continue to provide insights toward assessing the essential features of these and other critical components of the invasion cascade, which should be considered for inclusion in malaria vaccines or as possible drug targets.

We thank Dr. Basima Al-Khedery for comments on the manuscript and Dr. Anthony Holder for providing the Pf140 antiserum. J.C.R. was supported by the Human Frontier Science Program. This research was supported by Grant AI 24710-12 from the National Institutes of Health and by the National Center for Infectious Diseases, Centers for Disease Control and Prevention. Sequence data for *P. falciparum* chromosome 13 were obtained from The Sanger Centre Web site at <http://www.sanger.ac.uk/Projects/P.falciparum/>. Sequencing of *P. falciparum* chromosome 13 was accomplished as part of the Malaria Genome Project with support by The Wellcome Trust.

- World Health Organization (1997) *Wkly. Epidemiol. Rec.* **72**, 269.
- Dvorak, J. A., Miller, L. H., Whitehouse, W. C. & Shiroishi, T. (1975) *Science* **187**, 748-750.
- Bannister, L. H., Butcher, G. A., Dennis, E. D. & Mitchell, G. H. (1975) *Parasitology* **71**, 483-491.
- Barnwell, J. W. & Galinski, M. R. (1998) in *Malaria: Parasite Biology, Pathogenesis and Protection*, ed. Sherman, I. W. (Am. Soc. Microbiol., Washington, DC), pp. 93-120.
- Adams, J. H., Sim, B. K. L., Dolan, S. A., Fang, X., Kaslow, D. C. & Miller, L. H. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 7085-7089.
- Adams, J. H., Hudson, D. E., Torii, M., Ward, G. E., Wellems, T. E., Aikawa, M. & Miller, L. H. (1990) *Cell* **63**, 141-153.
- Sim, B. K., Toyoshima, T., Haynes, J. D. & Aikawa, M. (1992) *Mol. Biochem. Parasitol.* **51**, 157-159.
- Miller, L. H., Aikawa, M., Johnson, J. G. & Shiroishi, T. (1979) *J. Exp. Med.* **146**, 277-281.
- Waters, A. P., Thomas, A. W., Deans, J. A., Mitchell, G. H., Hudson, D. E., Miller, L. H., McCutchan, T. G. & Cohen, S. (1990) *J. Biol. Chem.* **265**, 17974-17979.
- Kappe, S. H. I., Noe, A. R., Fraser, T. S., Blair, P. L. & Adams, J. H. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 1230-1235.
- Galinski, M. R., Corredor Medina, C., Ingravallo, P. & Barnwell, J. W. (1992) *Cell* **69**, 1213-1226.
- Galinski, M. R., Xu, M. & Barnwell, J. W. (2000) *Mol. Biochem. Parasitol.* **108**, 257-262.
- Keen, J., Sinha, K. A., Brown, K. N. & Holder, A. A. (1994) *Mol. Biochem. Parasitol.* **65**, 171-177.
- Owen, C. A., Sinha, K. A., Keen, J. K., Ogun, S. A. & Holder, A. A. (1999) *Mol. Biochem. Parasitol.* **99**, 183-192.
- Ogun, S. A. & Holder, A. A. (1996) *Mol. Biochem. Parasitol.* **76**, 321-324.
- Freeman, R. R., Trejdosiewicz, A. J. & Cross, G. A. M. (1980) *Nature (London)* **284**, 366-368.
- Holder, A. A. & Freeman, R. R. (1981) *Nature (London)* **294**, 361-364.
- Trager, W. & Jensen, J. B. (1976) *Science* **193**, 673-676.
- Barnwell, J. W., Nichols, M. E. & Rubinstein, P. (1989) *J. Exp. Med.* **169**, 1795-1802.
- Vernick, K. D., Imberski, R. B. & McCutchan, T. F. (1988) *Nucleic Acids Res.* **16**, 6883-6896.
- von Heijne, G. (1983) *Eur. J. Biochem.* **133**, 17-21.
- Nielsen, H., Engelbrecht, J., Brunak, S. & von Heijne, G. (1997) *Protein Eng.* **10**, 1-6.
- Miller, L. H., Roberts, T. R., Shahabuddin, M. & McCutchan, T. F. (1993) *Mol. Biochem. Parasitol.* **59**, 1-14.
- Holder, A. A., Freeman, R. R., Uni, S. & Aikawa, M. (1985) *Mol. Biochem. Parasitol.* **14**, 293-303.
- Vinkenog, R., Aparecida Sperança, M., van Breemen, O., Ramesar, J., Williamson, D. H., Ross-MacDonald, P. B., Thomas, A. W., Janse, C. J., del Portillo, H. A. & Waters, A. P. (1998) *Mol. Biochem. Parasitol.* **94**, 1-12.
- Sinha, K. A., Keen, J. K., Ogun, S. A. & Holder, A. A. (1996) *Mol. Biochem. Parasitol.* **76**, 329-332.
- McCutchan, T. F., Kissinger, J. C., Touray, M. G., Rogers, M. J., Li, J., Sullivan, M., Braga, E. M., Kretteli, A. U. & Miller, L. H. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 11889-11894.
- Preiser, P. R., Jarra, W., Capiod, T. & Snounou, G. (1999) *Nature (London)* **398**, 618-622.
- Barnwell, J. W. (1999) *Nature (London)* **398**, 562-563.
- Galinski, M. R. & Barnwell, J. W. (1996) *Parasitol. Today* **12**, 20-29.