organelles play a key role in erythrocyte invasion and have been studied as vaccine targets, with the aim to induce antibodies to block invasion. One erythrocyte-binding molecule in the rhoptry is a complex of high-molecular-mass proteins called the RhopH complex [4,5]. The RhopH complex is distributed throughout the erythrocyte and PV membrane (PVM) and has been detected in ring-stage parasites [6], suggesting an important role during PV establishment. The importance of the complex has further been emphasized by the failure of attempts to disrupt the *pfrhoph3* gene locus, suggesting its necessity for parasite survival [7].

The RhopH complex comprises three distinct components: RhopH1, RhopH2, and RhopH3 [8–12]. The genes encoding RhopH1 are members of the *rhoph1/clag* gene family, which was originally defined by the cytoadherence-linked asexual gene (*clag*) on chromosome 9 in *Plasmodium falciparum* (*clag9*) and consists of at least three members; *clag2*, *3.1*, and *9* [13–15]. Although not yet determined experimentally, molecules encoded by *clag3.2* and *8* are likely parts of the RhopH complex as judged by their similarity in amino acid sequence and transcription pattern with other members [15]. Because only one RhopH1/Clag participates to form a single RhopH complex [15,16], five types of *Pf*RhopH complex are expected to exist, each of which contains one *rhoph1/clag* gene product. In this report we employ 'RhopH1/Clag' (protein) and '*rhoph1/clag*' (gene) as the family name, and 'Clag' (protein) and '*clag*' (gene) for each member.

Erythrocyte-binding proteins discharged from *P. falciparum* merozoites are considered to be targets of host immune responses. Strong diversifying selections on microneme proteins have been detected (e.g., AMA-1 and EBA-175), suggesting that polymorphism of these proteins has been maintained to evade host immunity in parasite populations [17,18]. Antibodies against the *Pf*RhopH complex partially inhibit the growth of *P. falciparum* in vitro and in vivo, consistent with its potential as a vaccine target [19–21]. Although the RhopH complex has been shown to induce host protective immunity and is likely to be under host immune pressure, the genetic diversity and immunologic characteristics of this complex are not fully understood. Here, we analyzed sequence polymorphism in five *rhoph1/clag* members, *rhoph2*, and *rhoph3* and show that some of the *rhoph* genes are under positive/diversifying selection. In addition, we assessed a population genetic mechanism that might drive the evolution of the *rhoph1/clag* multigene family.

## 2. Materials and methods

### 2.1. Malaria parasites

All cloned lines of *P. falciparum* were maintained in vitro, essentially as described previously [22]. The parasite lines examined originated from Southeast Asia (Dd2, FVO, Camp, T9/96, T9/102, K1, and Thai838), Papua New Guinea (MAD20), Central and South America (HB3, 7G8, DIV17, DIV29, DIV30, PC49, PC54, Santa Lucia, and Haiti), and Africa (RO33, 123/5, 128/4, SL/D6, LF4/1, 102/1, M2, M5, Fab9, 713, P13, and KMWII) and have been previously described [23–25]. Their geographic origins have also been previously described [26].

### 2.2. DNA and RNA isolation

Genomic DNA was obtained as described previously [24]. Total RNA was isolated from schizont stage-enriched HB3 and Dd2 parasite lines using the RNeasy mini kit (Qiagen, Valencia, CA). Complementary DNA was synthesized using random hexamers and an Omniscript reverse transcription kit (Qiagen) after DNase treatment.

### 2.3. Polymerase chain reaction (PCR) amplification and sequencing

Nucleotide sequences corresponding to open reading frame (ORF) were determined for five *pfrhoph1/clag* genes, *rhoph2*, and *rhoph3* in four parasite lines: Dd2, HB3, 7G8, and FVO. DNA fragments were PCR amplified with KOD-Plus DNA polymerase (Toyobo, Japan) using a panel of oligonucleotides specific for the genes (Supplemental Table 1) and sequenced directly using an ABI PRISM® 310 genetic analyzer (Applied Biosystems, Foster City, CA) or sequenced after cloning into pGEM-T Easy® plasmid (multiple plasmid clones sequenced for each DNA fragment; Promega, Madison, WI). To PCR amplify DNA fragments including the entire ORF of *clag3.1* or *3.2*, LA *Taq* DNA polymerase (TaKaRa, Japan) was used with oligonucleotide primers 3.1F (5′-TGTGCAATATATCAAAGTGTACATGC-3′) and 3.1R (5′-TAGAAAATATTAGAATTGCTATTATGTAC-3′) or 3.2F (5′-AATAGTTGAGTACGCACTAATATGTC-3′) and 3.2R (5′-ACACAAATTCTTAATAATTATATAAAACC-3′), respectively. A highly polymorphic region identified in *clag2*, *3.1*, *3.2*, and *8* in this study was further analyzed by increasing the number of parasite lines ($n = 25$) from different geographic areas.

### 2.4. Plasmodium reichenowi sequences

A TBLASTN search was performed against the *P. reichenowi* preliminary genome shotgun database (Dennis strain; Sanger Centre, UK) using Clag2, 3.1, 3.2, and 8 amino acid sequences as queries. For *prclag2* and *prclag8*, sequences were assembled using SeqMan II accompanied with Lasergene software (DNASTAR Inc., Madison, WI) with manual corrections. Regions covered by at least two independent reads and showing identical sequences were selected and used for analysis (Supplemental Figs. S1 and S2). The generated sequences were 3273 bp long for *prclag2*, corresponding to nucleotide (nt) positions 193–870, 1021–1902, 2458–3432, and 3448–4185 of *pfclag2* (3D7), and 2175 bp long for *prclag8*, corresponding to nt positions 1459–4173 of *pfclag8* (3D7). For Clag3 orthologs in the *P. reichenowi* genome, only sequences possessing homology with the 5′ untranslated region (UTR) (reich908g11.plk) or 3′ UTR (reich1194c08.plk and reich289f06.qlk) were used.

### 2.5. Sequence alignment and analysis

The entire ORFs for the 7 *Pf*RhopH complex-related genes (5 *rhoph1/clag* genes, *rhoph2*, and *rhoph3*) in four culture-adapted

*P. falciparum* lines—Dd2 (Southeast Asia), 7G8 (Brazil), HB3 (Honduras), and FVO (Vietnam)—were aligned with those retrieved from a genome database (3D7 line, presumably African in origin) using a CLUSTAL W program [27] with manual corrections; nucleotide diversity ($\pi$) and its standard error (S.E.) were computed with the Jukes and Cantor method using MEGA 3.1 software [28] after excluding insertions/deletions (indels) and highly polymorphic regions. The mean numbers of synonymous substitutions per synonymous site ($d_S$) and nonsynonymous substitutions per nonsynonymous site ($d_N$) and their standard errors were computed using the Nei and Gojobori method [29] with the Jukes and Cantor correction, implemented in MEGA 3.1. The statistical difference between $d_S$ and $d_N$ was tested using a one-tailed Z-test with 500 bootstrap pseudosamples using MEGA 3.1. A value of $d_N$ significantly higher or lower than $d_S$ at the 95% confidence level was taken as evidence for positive or purifying selection, respectively. The $d_N$:$d_S$ ratio was evaluated using a sliding window method (50 bases with a step size of 10 bases) in DnaSP 4.0 [30]. Positive selection was also evaluated using the McDonald–Kreitman test [31]. Before estimating the time to the most recent common ancestor (TMRCA) for *P. falciparum clag2* and *8* polymorphism, the evolutionary rate constancy of *clag2* and *8* between *P. falciparum* and *P. reichenowi* was validated using a *Plasmodium yoelii* ortholog *Py*RhopH1A (accession number AB060734) as an outgroup using Tajima's relative rate test [32] implemented in MEGA 3.1. Mean and 95% confidence intervals (CI) for estimated TMRCA were computed based on the model assuming the distribution of the distance and the substitution rate were Gamma-distributed [33]. Gene conversion was evaluated for each exon using an algorithm by Betrán et al. [34] implemented in DnaSP 4.0.

Unrooted dendrograms of the *pfrhoph1/clag* members were constructed using the neighbor-joining and maximum parsimony methods in MEGA 3.1, and Tajima's relative rate test was used to evaluate the evolutionary rate among members. Indels and highly polymorphic regions could not be satisfactorily aligned and were therefore excluded from the analysis. The sequences (3D7 parasite line) used to construct trees and the evolutionary rate were as follows: nt positions 154–312, 331–573, 727–1122, 1207–1266, 1324–1560, 1609–2988, 3004–3288, and 3382–3924 for *clag3.2*; nt positions 160–318, 337–579, 733–1128, 1213–1272, 1330–1566, 1615–2994, 3010–3294, and 3388–3930 for *clag3.1*; nt positions 223–381, 400–642, 799–1194, 1279–1338, 1390–1626, 1696–3075, 3091–3375, and 3553–4095 for *clag2*; nt positions 130–288, 307–549, 706–1101, 1186–1245, 1300–1536, 1606–2985, 3001–3285, and 3415–3957 for *clag8*; and nt positions 82–240, 265–507, 652–1047, 1132–1191, 1276–1512, 1582–2961, 2977–3261, and 3394–3936 for *clag9*.

## 3. Results

### 3.1. Polymorphism of the PfRhopH complex-related genes

All seven *Pf*RhopH complex-related genes showed greater nucleotide diversity levels than the average (+2 S.E.) of 204 ORFs on *P. falciparum* chromosome (chr) 3 [35] (Table 1). Among the seven genes, *clag2*, *3.1*, *3.2*, and *8* are highly polymorphic with nucleotide diversity ($\pi = 0.0053$–$0.0164$) comparable to malaria vaccine candidate antigen protein genes such as *eba-175* ($\pi = 0.0030$) and *ama-1* ($\pi = 0.0166$) [17,18]. The observed nucleotide diversity levels of *clag2*, *3.1*, and *3.2* should be taken as minimum estimates, because indels and highly polymorphic regions were excluded from this analysis to obtain reliable alignments. The highly polymorphic nature of four *rhoph1/clag* genes at the nucleotide level extends to the amino acid level, which is represented by high $d_N$ values (Table 1). Thus, the genes encoding RhopH1/Clag are more polymorphic than RhopH2 and RhopH3.

Among the four RhopH1/Clag showing high polymorphism (Clag2, 3.1, 3.2, and 8), the majority of polymorphic sites are clustered in a region at amino acid (aa) positions 1000–1200

Table 1
Nucleotide diversity of the *Pf*RhopH complex genes[a]

| Gene | *n* | Indel | Sites | $\pi$ | $\pi$ (S.E.) | $d_N$ | $d_N$ (S.E.) | $d_S$ | $d_S$ (S.E.) | $d_N/d_S$ | *P*[b] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *clag2*[c] | 5 | (+) | 4,317 | 0.0053 | (0.0008) | 0.0032 | (0.0007) | 0.0133 | (0.0028) | 0.24 | (0.0003) |
| *clag3.1*[c] | 5 | (+) | 4,140 | 0.0164 | (0.0015) | 0.0062 | (0.0011) | 0.0582 | (0.0058) | 0.11 | ($<10^{-10}$) |
| *clag3.2*[c] | 5 | (+) | 4,134 | 0.0138 | (0.0011) | 0.0063 | (0.0011) | 0.0445 | (0.0050) | 0.14 | ($<10^{-10}$) |
| *clag8* | 5 | (−) | 4,182 | 0.0066 | (0.0007) | 0.0065 | (0.0011) | 0.0069 | (0.0020) | 0.94 | ns |
| *clag9* | 5 | (−) | 4,020 | 0.0009 | (0.0003) | 0.0011) | (0.0004) | 0.0000 | (0.0000) | ∞ | 0.002 |
| *rhoph2* | 5 | (−) | 4,134 | 0.0009 | (0.0003) | 0.0010 | (0.0004) | 0.0005 | (0.0005) | 2.00 | ns |
| *rhoph3* | 5 | (−) | 2,691 | 0.0013 | (0.0004) | 0.0012 | (0.0005) | 0.0015 | (0.0010) | 0.80 | ns |
| *clag2*[d] | 24 | (+) | 522 | 0.0131 | (0.0032) | 0.0114 | (0.0042) | 0.0192 | (0.0074) | 0.60 | ns |
| *clag8*[d] | 26 | (−) | 585 | 0.0267 | (0.0042) | 0.0305 | (0.0060) | 0.0132 | (0.0061) | 2.31 | 0.020 |
| *Chr 3*[e] | 5 | | 202,069 | 0.00044 | (0.00006) | 0.00039 | (0.0060) | 0.00068 | (0.00010) | 0.57 | |

[a] *n*, Number of sequences sampled; sites, sites analyzed excluding noncoding sequences and alignment gaps; $\pi$, pairwise nucleotide diversity; $d_N$, number of nonsynonymous substitutions over numbers of nonsynonymous sites; $d_S$, number of synonymous substitutions over numbers of synonymous sites; S.E., standard error computed using the Nei-Gojobori method with Jukes-Cantor correction. S.E. was estimated using the bootstrap method with 500 replication.

[b] *P*-value indicates that $d_N$ is significantly greater than $d_S$. Those shown in parenthesis indicate that $d_S$ are significantly greater than $d_N$. The statistical difference between $d_S$ and $d_N$ was tested using an one-tail Z-test with 500 bootstrap pseudosamples implemented in MEGA 3.1. ns indicate not significant ($P > 0.05$).

[c] For optimal sequence alignment, nt 3433–3435 was excluded from *clag2*, nt 3337–3447 from *clag3.1*, and nt 88–99 and 3343–3444 from *clag3.2* for the analysis. Nucleotide numbering are after the 3D7 line sequences.

[d] nt 3022–3606 of *clag8* and nt 3106–3420 and 3436–3642 of *clag2* were used.

[e] Data from 204 ORF on *P. falciparum* chr 3 using five parasite lines [35].

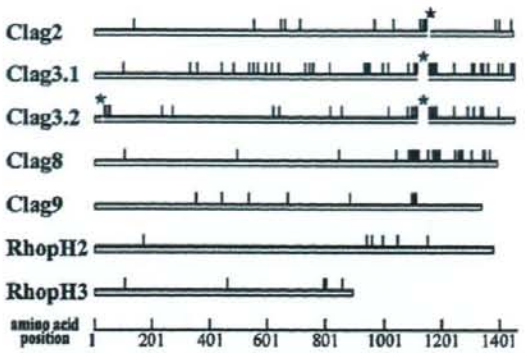


Fig. 1. Locations of amino acid polymorphism of seven components of the *Pf*RhopH complex among five *P. falciparum* parasite lines (3D7, HB3, Dd2, FVO, and 7G8). Indels are shown as gaps with asterisks (aa position 1145 for Clag2, aa positions 1113–1149 for Clag3.1, and aa positions 30–33 and 1115–1148 for Clag3.2). Numbers are those of 3D7 line sequences.

(Fig. 1). In addition, numerous polymorphic sites in this region have more than one amino acid substitution, whereas most polymorphisms in the other regions are dimorphic (at both nucleotide and amino acid levels). Most indels are also located in this region (Fig. 1, asterisks). Thus, the region at aa positions 1000–1200 of RhopH1/Clag is the most highly polymorphic region of the *Pf*RhopH complex.

### 3.2. Gene conversion between clag3.1 and 3.2

Of interest, *clag3.1* and *3.2* share some polymorphic sites. Because *clag3.1* and *3.2* have 96.7% nucleotide identity (3D7 parasite line) and are located on chr 3 and separated by only 10 kb harboring one putative ORF (PFC0115c) (Fig. 2A), we assessed gene conversion between these two loci. Using an algorithm by Betrán et al. [34], we identified multiple gene conversion tracts located at nt positions 1314–1353, 1447–1452, 1612–1659, 1702–1785, 1852–1983, and 2148–2208 in 3D7 *clag3.1*; nt positions 3824–4240 in HB3 *clag3.1*; nt positions 189–247 in 7G8 *clag3.1*; nt positions 813–817 and 3821–4182 in 3D7 *clag3.2*; nt positions 88–151 in HB3 *clag3.2*; and nt positions 3320–3755 in 7G8 *clag3.2* (Fig. 3). The detected conversion tracts had less than 5% informative nucleotides showing a mosaic origin, indicating that the probability of these tracts being involved in a recombination event more than once is negligible [34]. No gene conversion was detected between the other *rhoph1/clag* genes.

Because gene conversion potentially accelerates nucleotide diversity, we evaluated the evolutionary rates of *clag3.1* and *3.2*. Results showed that *clag2*, *3.1*, and *3.2* form a single clade and *clag8* another (Fig. 4); thus we performed Tajima's relative rate test using *clag8* as an outgroup and found that the evolutionary rates between *clag3.1* and *2* and between *clag3.2* and *2* were significantly different for all combinations of the sequences from five parasite lines. Because *clag3.1* and *3.2* were more diverse than *clag2*, *clag3.1* and *3.2* appear to have evolved more rapidly than *clag2*.

### 3.3. Amino acid polymorphism of the region around aa positions 1000–1200 of Clag2, 3.1, 3.2, and 8

Because extensive polymorphisms were observed around aa positions 1000–1200 in Clag2, 3.1, 3.2, and 8, we further analyzed polymorphism in this region with additional sequences from parasite lines originating worldwide. Alignment of Clag2 sequences showed multiple amino acid substitutions per site at multiple sites, e.g., five amino acids at aa position 1139 (K, R, S, G, and I). Indels were also observed (Supplemental Fig. S3). Clag8 has even higher levels of amino acid substitutions at between 1077 and 1136; five different amino acids (I, S, R, G, and N) at 1100, seven at 1101 (D, S, T, E, N, I, and K), six at 1104 (S, N, I, K, R, and T), and five at 1105 (G, D, T, S, and N) (Supplemental Fig. S4). Clag3.1 and 3.2 are also highly polymorphic (Fig. 5), which will be discussed later.

### 3.4. Copy number polymorphism of rhoph1/clag genes on chr 3

Notably, when PCR amplification was performed to obtain DNA fragments of the entire ORFs of *clag3.1* or *3.2*, 17 parasite

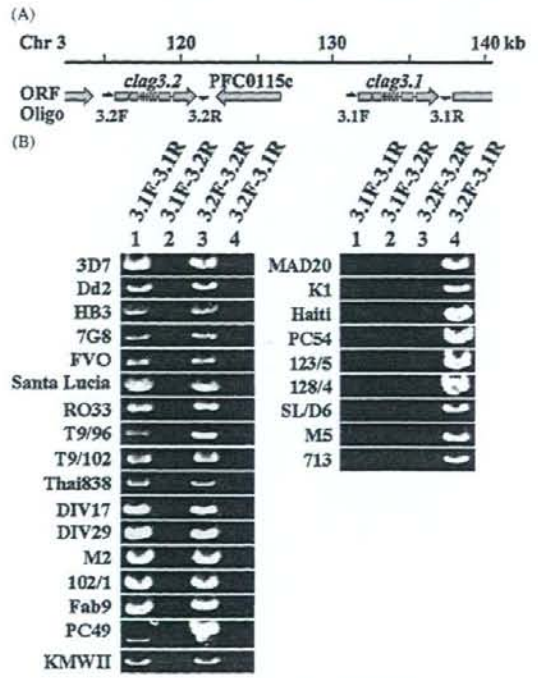


Fig. 2. Copy number polymorphism of *rhoph1/clag* genes on chr 3. (A) Genome organization around *clag3.2* and *3.1* gene loci on chr 3. The locations of the oligonucleotide primers are indicated. Oligonucleotide 3.2F and 3.1F were designed on the 5′ UTR of *clag3.2* and *3.1*, respectively. Oligonucleotide 3.2R and 3.1R were designed on 3′ UTR of *clag3.2* and *3.1*, respectively. (B) PCR-amplified DNA fragments of 26 *P. falciparum* lines with different combinations of oligonucleotides.

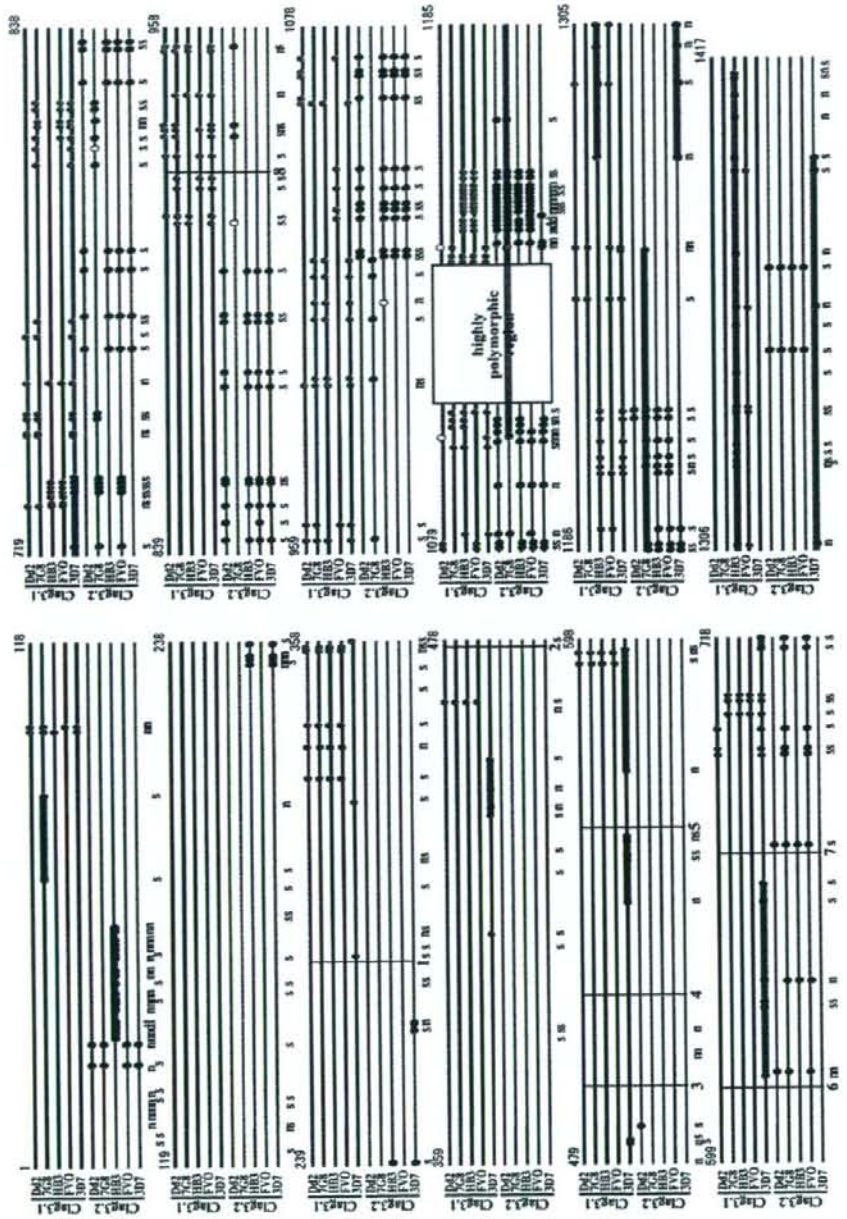


Fig. 3. Gene conversion tracts in *clag3.1* and *clag3.2*. Polymorphic codons (circles) in the coding sequences of Clag3.1 and 3.2 were compared in five *P. falciparum* lines. Clag3.1, black bar; Clag3.2, gray bar. Polymorphisms matching the paralogous sequence are shown in gray or black circles, respectively, and rare polymorphisms by an open circle. Exons are separated by vertical bar with the intron number at the bottom. Polymorphic sites that differ between consensus sequences are shown below the line classified as nonsynonymous (n), synonymous (s), and deletion (d). Gene conversion tracts identified using algorithm by Betrán et al. [34], wide gray bars.

lines showed the 2 expected positive bands with the primer sets 3.1F–3.1R and 3.2F–3.2R, whereas 9 parasite lines showed a positive band only with the primer set 3.2F–3.1R, which suggests that these 9 parasite lines possessed a hybrid gene with *clag3.2* sequence at the 5′ UTR and *clag3.1* sequence at the 3′ UTR (Fig. 2B). DNA fragments were not amplified with other primer combinations, indicating that artificial amplification due to primer mispairing was negligible. This is consistent with a

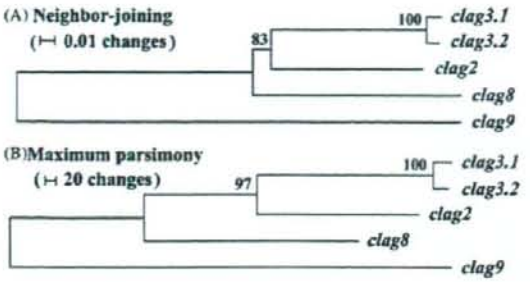


Fig. 4. Unrooted dendrograms of *pfrhoph1/clag* genes using nucleotide sequences from the 3D7 parasite line. The trees were constructed by the neighbor-joining and maximum parsimony methods using MEGA 3.1. Numbers on branches indicate bootstrap values (500 pseudoreplicates).

recent report by Chung et al. [36], who found that some parasite lines possess only a single *rhoph1/clag* on chr 3 by Southern blot hybridization. We here designate this *clag3* gene as *clag3h* (*clag3 hybrid*; Clag3H for protein). In addition, we obtained two distinct sequences for *clag3.1* from the KMWII parasite line (Fig. 5), using several cloned plasmids after experiencing difficulty in direct sequencing of PCR products. Sequences for *clag2*, *3.2*, and *8* were easily obtained from the KMWII line by direct sequencing of the PCR products, supporting the assumption that this line was a clone. Thus, the KMWII line appears to possess at least three *clag3*-related sequences in the genome. This data suggests that the number of *clag3*-related sequences in *P. falciparum* varies from one to at least three.

To deduce the direction of the one-gene to two-gene (or vice versa) change, we searched *P. reichenowi* orthologs in the genome database and found one sequence read (reich908g11.plk) showing high similarity with the sequence around the start codon of *clag3.1* and *3.2*. We also found two reads (reich289f06.qlk and reich1194c08.plk) showing strong homology with the sequence around the stop codons of *clag3.1* and *3.2*. Comparison of the nucleotide sequences at the UTR revealed that reich908g11.plk and reich289f06.qlk were similar to the *pfclag3.2* sequence and that reich1194c08.plk was similar to the *pfclag3.1* sequence (Fig. 6). Thus, duplication of *clag3.1* and *3.2* gene loci appears to predate the divergence of *P. falciparum* and *P. reichenowi*, suggesting that a single *rhoph1/clag* (*clag3h*) found in some *P. falciparum* lines is likely a result of an unequal crossover between two closely related genes. Notably, Clag3H had characteristic amino acids that were not observed in Clag3.1 and 3.2. For example, Ala at 1116 was found in three of nine Clag3H (30%). If Clag3H originated recently, for example during culture, the amino acid allele observed in Clag3H would also exist in Clag3.1 or 3.2; however, Ala at 1116 was not found in a total of 36 sequences of non-Clag3H protein sequences. Three in nine Clag3Hs is a significant excess compared to zero Ala at 1116 in 36 non-Clag3H sequences by Fisher's exact test ($P = 0.013$). This suggests that at least some Clag3H have accumulated some unique amino acid substitutions since their creation.

### 3.5. Selection on the PfRhopH complex

Positive selection was evaluated by comparing synonymous and nonsynonymous substitutions (Table 1). A significant excess of $d_N$ over $d_S$ was observed for *clag9* (entire ORF of five parasite lines) and for *clag8* (highly polymorphic region at nt positions 3022–3606 of 26 parasite lines), suggesting positive selection acting on these genes. A sliding window plot of $d_N$:$d_S$ ratios revealed that *clag2* and *8* had the highest peaks, around nt positions 3000–3600 (Fig. 7). It should be noted that the corresponding regions of *clag3.1* and *3.2* are the regions showing highly extensive polymorphism with indels (asterisks in Fig. 7), thereby preventing evaluation of $d_N$:$d_S$ ratios in this region. The peak at the N-terminus of *clag3.2* is due to introduction of part of the *clag3.1* sequence into the HB3 line *clag3.2* by gene conversion (see Fig. 3).

Positive selection was further evaluated by the McDonald–Kreitman test using *P. reichenowi* orthologs for *clag2* and *8*. Significant excess of intraspecific nonsynonymous substitutions over synonymous substitutions was observed in *clag8* as compared with interspecies fixed differences of nonsynonymous and synonymous changes, suggesting positive selection (Table 2).

### 3.6. Early origin of the clag2 and 8 polymorphism

We estimated the TMRCA for *clag2* and *8* polymorphism using aligned regions. Distances of synonymous single-nucleotide polymorphisms are $0.0139 \pm 0.0031$ for *clag2* and $0.0106 \pm 0.0030$ for *clag8*. Distances between *P. falciparum* and *P. reichenowi* are $0.0455 \pm 0.0082$ and $0.0748 \pm 0.0120$ for *clag2* and *8*, respectively. Assuming that the divergence time of *P. falciparum* and *P. reichenowi* was 6 million years ago (mya) [37,38], the estimated TMRCA of the polymorphism of *clag2*

Table 2
The McDonald–Kreitman test of selection for *Plasmodium falciparum clag2* and *8*

| Locus | $n^a$ | No. of sites | Fixed differences between species | | Polymorphic sites within *P. falciparum* | | |
|---|---|---|---|---|---|---|---|
| | | | Syn | Nsyn[b] | Syn | Nsyn | $P^c$ |
| *clag2* | 5 | 3273 | 23 | 47 | 18 | 12 | (0.011) |
| *clag8* | 5 | 2715 | 36 | 55 | 11 | 38 | 0.030 |

[a] *n*, Number of *P. falciparum* lines used.
[b] Syn, synonymous; Nsyn, nonsynonymous substitutions.
[c] Fisher's exact test (one-tailed) was used. *P*-value indicates that Nsyn are significantly greater than Syn. Value in parenthesis indicates that Syn are significantly greater than Nsyn.

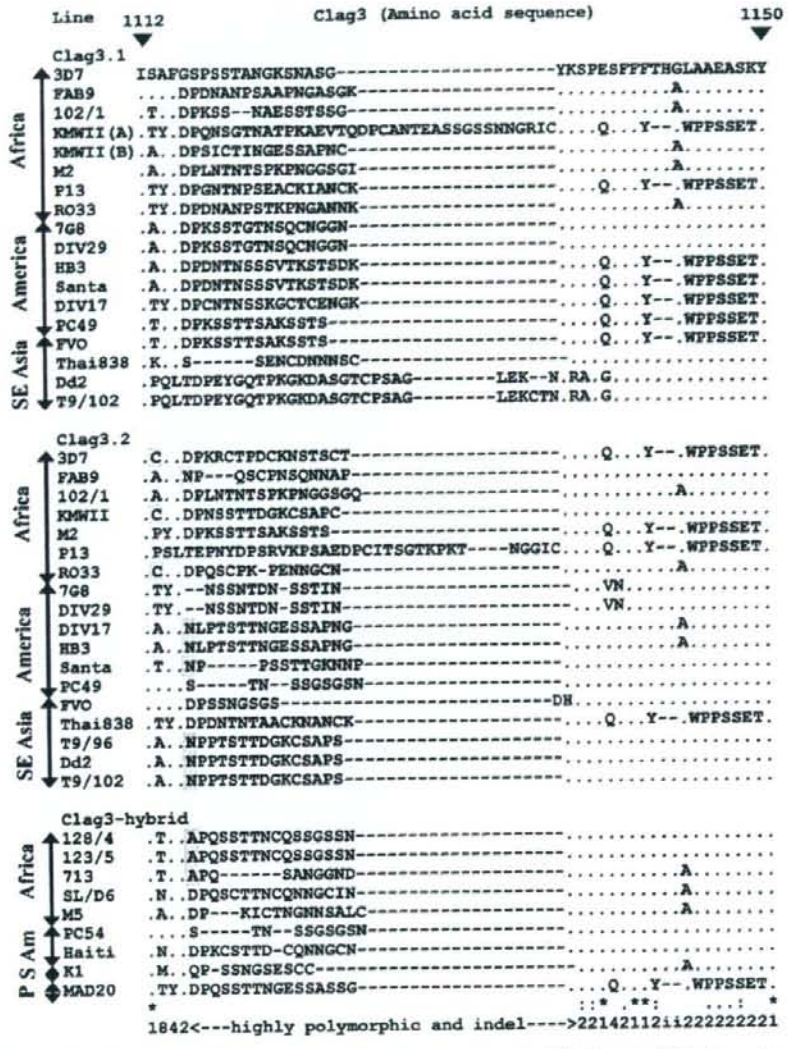


Fig. 5. Polymorphism of Clag3. An amino acid region 1112–1150 (after 3D7 line Clag3.1 sequence) of *P. falciparum* (27 lines) was aligned. Geographic origins are shown at left: SE Asia or S, Southeast Asia; Am, America; P, Papua New Guinea; Santa, Santa Lucia cloned line. Identical, conserved, or semiconserved residues in the alignment are indicated with asterisk, colon, or period, respectively. The number of amino acid replacements at each position and the region with indels are shown at the bottom. Cys residue at aa position 1113 and Asn and Ala residues at aa position 1116 are masked.

and *8* are 1.89 (95% CI, 1.02–3.18) and 0.87 (95% CI, 0.42–1.54) mya, respectively.

## 4. Discussion

### *4.1. Diversifying selection on the rhoph1/clag gene loci*

The present study revealed that the RhopH1/Clag-encoding genes *clag2*, *3.1*, *3.2*, and *8* contain a highly polymorphic region, particularly at nt positions 3000–3600. Diversifying selection increases nucleotide diversity ($\pi$), and an excess of $d_N$ to $d_S$ is indicative of positive selection favoring amino acid replacement [39]. Thus, the observed excess of $d_N$ to $d_S$ at nt positions 3000–3600 of *clag8* suggests that the polymorphism in *clag8* is positively maintained. An excess of $d_N$ to $d_S$ was also observed for *clag9*, indicating that this gene is also under positive selection. The most polymorphic region, in which positive selection was detected for *clag8*, was excluded from *clag2*, *3.1*, and *3.2* due to extensive sequence variation that made sequence alignment unreliable. Further analysis is required to evaluate positive selection on these three *rhoph1/clag* genes.


Fig. 6. Nucleotide sequence alignment of the 5′ and 3′ UTRs for *P. falciparum clag3.1* and *3.2* genes and *P. reichenowi* orthologous sequences. A *P. reichenowi* sequence possessing homology with *pfclag3.1* was not found in the current database. Nucleotide sequences corresponding to the ORF and the UTR are shown with upper case and lower case letters, respectively. Putative start and stop codons are boxed. Characteristic nucleotides are displayed in reverse (*clag3.1*) or masked with gray (*clag3.2*).

To date, observation of such high levels of polymorphism for *clag2*, *3.1*, *3.2*, and *8* ($\pi = 0.0053$–$0.0164$; $d_N = 0.0032$–$0.0065$) has not been reported for other known malaria rhoptry protein genes. The high polymorphism observed in *clag2*, *3.1*, and *3.2* is consistent with the observation by Kidgell et al. [40] based on the hybridization of genomic DNA from a panel of parasite lines to an oligonucleotide array for the *P. falciparum* genome. In addition, the polymorphism levels are comparable to those of the microneme proteins such as *eba-175* ($\pi = 0.0030$; $d_N = 0.0037$) and *ama-1* ($\pi = 0.0166$; $d_N = 0.0207$), which are exposed to host immune responses [17,18]. Rhoptry proteins are released into the PV and are considered to be minimally exposed to host immunity. If RhopH1/Clag polymorphism is generated by host immune pressure, the questions arises as to how RhopH1/Clag is exposed to host immunity. There are a few possible explanations. First, RhopH1/Clag may be released from the merozoites before attachment to the erythrocyte surface, thereby becoming a target of host immunity. Second, the RhopH complex, which is released into PVs, may be leaked to the surface of infected erythrocytes through the junction between invading parasite and the erythrocyte membrane. Leaked RhopH complex, and therefore parasite-infected erythrocytes, are then potential targets of host immunity. Indeed, the *Pf*RhopH complex and rhoptry-associated protein 2 (RAP-2, RSP-2), another malaria rhoptry protein, have been detected on the erythrocyte surface upon parasite attachment to erythrocytes [41,42]. The last possibility is that RhopH1/Clag, after release into the PV, may be distributed to the parasite-derived membranous network (i.e., Maurer's clefts) in the erythrocyte cytosol, where it is exposed to host immunity. RhopH2 and RhopH3 have recently been observed in materials deriving from Maurer's cleft by proteome analyses, consistent with this possibility [43,44].

There are no obvious associations between particular haplotypes and their geographic origins, and most haplotypes co-exist in different geographic areas, similar to other known polymorphic antigens such as MSP-1 [45,46]. RhopH1/Clag polymorphism might be maintained in natural parasite populations to evade host immunity. Using a T-cell epitope prediction algorithm (SYFPEITHI software) [47], we found that binding of a predicted T-cell epitope peptide to a particular HLA allotype was dramatically affected by the RhopH1/Clag polymorphism in silica. For example, aa positions 1094–1108 of 3D7 line Clag8 (KRISTSIDHISGGKW) was predicted as a T-cell epitope of HLA-DRB1*1101 with a score of 22, but the score for the corresponding region of Camp line Clag8 (MRISSTSTYISNNEW) was 0, emphasizing a potential involvement in immune evasion of RhopH1/Clag polymorphism. We consider the algorithm useful because the score of HLA-DRB1*0701 for the *Pf*CSP Th2R domain, a well characterized malaria polymorphic T-cell epitope peptide, is 22 for the K1 parasite line (KIQYSLSTEWSPCSV) but only 12 for that of the 3D7 line (KIQNSLSTEWSPCSV).

### 4.2. Evolution of *PfRhopH1/Clag* polymorphism of the extant *P. falciparum* population

Gene conversion has been reported for other *P. falciparum* loci, such as *falcipain 2* [48] and *var* [49], as a source of genetic diversity. In this study, we show that *clag3.1/3.2*, which interchange their sequences by gene conversion, evolved more rapidly than *clag2*. The precise function of the RhopH complex remains unknown, and thus whether the gene conversion observed was functionally advantageous or neutral is also unknown; however, gene conversion can be a mechanism for antigenic variation to evade host immunity. Some examples include the *vsg* gene *of Trypanosoma brucei*, causative agent of African sleeping sickness, the *ves* gene of a cattle parasite *Babesia bovis* [50], and *var* genes in *P. falciparum* [51].

Based on the shared hybridization pattern between *clag3.1* and *3.2*, Chung et al. [36] proposed that these genes are alleles of the same locus; however, because the origin of two *rhoph1/clag* loci on chr 3 appear to predate the *P. falciparum–P. reichenowi* divergence, these should be categorized as paralogous genes but not the same gene. Shared features between these loci detected by Southern blot hybridization by Chung et al. can be simply explained by the gene conversion identified in this study. *clag3h* could be generated by an unequal crossover between *clag3.1* and *3.2* of a set of chromosomes during meiosis. If such crossover had occurred, parasite lines possessing three *rhoph1/clag* on chr 3 would be expected, with a molecule having its 5′ end derived from *clag3.1* and its 3′ end from *clag3.2* (Supplemental Fig. S5, model 1); however, this type of *rhoph1/clag* was not detected in this study. KMWII line appears to possess 3 *rhoph1/clag*

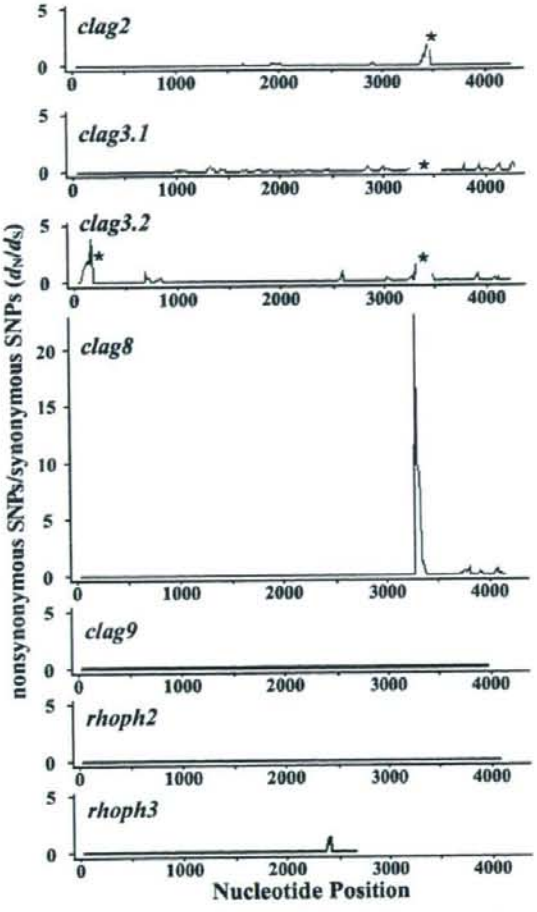


Fig. 7. Sliding window plot of $d_N/d_S$ ratio for seven genes of the *Pf*RhopH complex. For optimum sequence alignment, nt positions 3433–3435 was excluded from *clag2*, nt positions 3337–3447 from *clag3.1*, and nt positions 88–99 and 3343–3444 from *clag3.2* for the analysis (asterisks), because these regions were highly polymorphic with indels. Nucleotide numbers are those of 3D7 line sequences. Window length is 50 bp, and step size is 10 bp.

genes on chr 3, but the third *rhoph1/clag* on chr 3 appears not to be generated by the mechanism described above, because this *rhoph1/clag* was obviously a duplicated *clag3.1* gene amplified with the *clag3.1*-specific primer set. Thus, *clag3h* is more likely a product of a recombination event between *clag3.1* and *3.2* on the same chromosome (Supplemental Fig. S5, model 2). Because a unique amino acid of Clag3H (e.g., Ala at aa position 1116) suggests a relatively old origin of *clag3h*, recombination events between *clag3.1* and *3.2* might be rare in the natural population.

Four highly polymorphic *rhoph1/clag* genes contained unexpectedly large numbers of synonymous substitutions. Based on the ratio of interspecific distance to intraspecific distance, the TMRCA of the polymorphism of *P. falciparum clag2* and *8* were estimated to be 1.89 (95% CI, 1.02–3.18) and 0.87 (95% CI, 0.42–1.54) mya, respectively. Although there is still controversy surrounding its accuracy, TMRCA of the extant *P. falciparum* population was estimated to be approximately 0.1–0.2 mya based on the genetic distance in nuclear genome housekeeping genes between *P. falciparum* and *P. reichenowi* ([52], Tanabe, unpublished data). Thus, polymorphism of *clag2* and *8* appears to be generated between the divergence of *P. falciparum* and *P. reichenowi* and TMRCA of the extant *P. falciparum* populations. Early origins of the polymorphism have been suggested for merozoite surface proteins *Pf*MSP-1 and *Pf*MSP-2, for which the origin of the polymorphism was proposed to predate the *P. falciparum–P. reichenowi* divergence (thus termed 'ancient origin'), or TMRCA of the extant *P. falciparum* population, respectively [53,54]. Early origins of the polymorphism older than TMRCA of extant *P. falciparum* populations would suggest that *rhoph1/clag* polymorphisms confer an advantage to the parasite and were positively selected for during the recent evolution of *P. falciparum*.

In summary, four factors appear to affect current *rhoph1/clag* polymorphism; (i) older origin than TMRCA of the extant *P. falciparum* population; (ii) gene conversion and (iii) copy number polymorphism for *rhoph1/clag* on chr 3; and (iv) positive diversifying selection. Multigene families play important roles in many aspects of malaria biology, e.g., responsibility for redundancy of erythrocyte invasion or antigenic variation of parasite-infected erythrocytes. Given the abundance of multigene families in the *P. falciparum* genome [55], combination of the mechanisms described in this study can be a powerful driving force to generate high biologic redundancy for parasite survival.

## Acknowledgements

We thank NIAID intramural editor Brenda Rae Marshall for assistance. Sequence data of *P. reichenowi* were obtained from The Sanger Centre website (http://www.sanger.ac.uk/Projects/P_reichenowi/). This work was supported in part by Grants-in-Aid for Scientific Research 15406015 (to MT), 18390131 (to KT), and 17590372 and 17406009 (to OK) from the Ministry of Education, Culture, Sports, Science and Technology, Japan, and by the Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, USA (to XS).

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.molbiopara.2007.11.004.

## References

[1] Cowman AF, Crabb BS. Invasion of red blood cells by malaria parasites. Cell 2006;124:755–66.
[2] Kaneko O. Erythrocyte invasion: vocabulary and grammar of the *Plasmodium* rhoptry. Parasitol Int 2007;56:255–62.
[3] Bannister LH, Mitchell GH, Butcher GA, Dennis ED. Lamellar membranes associated with rhoptries in erythrocytic merozoites of *Plasmodium knowlesi*: a clue to the mechanism of invasion. Parasitology 1986;92:291–303.

[4] Sam-Yellowe TY, Perkins ME. Interaction of the 140/130/110 kDa rhoptry protein complex of *Plasmodium falciparum* with the erythrocyte membrane and liposomes. Exp Parasitol 1991;73:161–71.
[5] Rungruang T, Kaneko O, Murakami Y, et al. Erythrocyte surface glycosylphosphatidyl inositol anchored receptor for the malaria parasite. Mol Biochem Parasitol 2005;140:13–21.
[6] Lustigman S, Anders RF, Brown GV, Coppel RL. A component of an antigenic rhoptry complex of *Plasmodium falciparum* is modified after merozoite invasion. Mol Biochem Parasitol 1988;30:217–24.
[7] Cowman AF, Baldi DL, Duraisingh M, et al. Functional analysis of proteins involved in *Plasmodium falciparum* merozoite invasion of red blood cells. FEBS Lett 2000;476:84–8.
[8] Holder AA, Freeman RR, Uni S, Aikawa M. Isolation of a *Plasmodium falciparum* rhoptry protein. Mol Biochem Parasitol 1985;14:293–303.
[9] Brown HJ, Coppel RL. Primary structure of a *Plasmodium falciparum* rhoptry antigen. Mol Biochem Parasitol 1991;49:99–110.
[10] Shirano M, Tsuboi T, Kaneko O, Tachibana M, Adams JH, Torii M. Conserved regions of the *Plasmodium yoelii* rhoptry protein RhopH3 revealed by comparison with the *P. falciparum* homologue. Mol Biochem Parasitol 2001;112:297–9.
[11] Ling IT, Kaneko O, Narum DL, et al. Characterisation of the *rhoph2* gene of *Plasmodium falciparum* and *Plasmodium yoelii*. Mol Biochem Parasitol 2003;127:47–57.
[12] Kaneko O, Tsuboi T, Ling IT, et al. The high molecular mass rhoptry protein, RhopH1, is encoded by members of the *clag* multigene family in *Plasmodium falciparum* and *Plasmodium yoelii*. Mol Biochem Parasitol 2001;118:237–45.
[13] Holt DC, Gardiner DL, Thomas EA, et al. The cytoadherence linked asexual gene family of *Plasmodium falciparum*: are there roles other than cytoadherence? Int J Parasitol 1999;29:939–44.
[14] Ling IT, Florens L, Dluzewski AR, et al. The *Plasmodium falciparum clag9* gene encodes a rhoptry protein that is transferred to the host erythrocyte upon invasion. Mol Microbiol 2004;52:107–18.
[15] Kaneko O, Yim Lim BY, Iriko H, et al. Apical expression of three RhopH1/Clag proteins as components of the *Plasmodium falciparum* RhopH complex. Mol Biochem Parasitol 2005;143:20–8.
[16] Ghoneim A, Kaneko O, Tsuboi T, Torii M. The *Plasmodium falciparum* RhopH2 promoter and first 24 amino acids are sufficient to target proteins to the rhoptries. Parasitol Int 2007;56:31–43.
[17] Polley SD, Conway DJ. Strong diversifying selection on domains of the *Plasmodium falciparum* apical membrane antigen 1 gene. Genetics 2001;158:1505–12.
[18] Baum J, Thomas AW, Conway DJ. Evidence for diversifying selection on erythrocyte-binding antigens of *Plasmodium falciparum* and *P. vivax*. Genetics 2003;163:1327–36.
[19] Siddiqui WA, Tam LQ, Kramer KJ, et al. Merozoite surface coat precursor protein completely protects *Aotus* monkeys against *Plasmodium falciparum* malaria. Proc Natl Acad Sci USA 1987;84:3014–8.
[20] Cooper JA, Ingram LT, Bushell GR, et al. The 140/130/105 kilodalton protein complex in the rhoptries of *Plasmodium falciparum* consists of discrete polypeptides. Mol Biochem Parasitol 1988;29:251–60.
[21] Doury JC, Bonnefoy S, Roger N, Dubremetz JF, Mercereau-Puijalon O. Analysis of the \ weight rhoptry complex of *Plasmodium falciparum* using monoclonal antibodies. Parasitology 1994;108:269–80.
[22] Trager W, Jensen JB. Human malaria parasites in continuous culture. Science 1976;193:673–5.
[23] Jongwutiwes S, Tanabe K, Nakazawa S, Uemura H, Kanbara H. Coexistence of GP195 alleles of *Plasmodium falciparum* in a small endemic area. Am J Trop Med Hyg 1991;44:299–305.
[24] Kaneko O, Soubes SC, Miller LH. *Plasmodium falciparum*: invasion of *Aotus* monkey red blood cells and adaptation to *Aotus* monkeys. Exp Parasitol 1999;93:116–9.
[25] Su X, Ferdig MT, Huang Y, et al. A genetic map and recombination parameters of the human malaria parasite *Plasmodium falciparum*. Science 1999;286:1351–3.
[26] Mu J, Ferdig MT, Feng X, et al. Multiple transporters associated with malaria parasite responses to chloroquine and quinine. Mol Microbiol 2003;49:977–89.
[27] Thompson JD, Higgins DG, Gibson TJ, Clustal W. improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22:4673–80.
[28] Kumar S, Tamura K, Nei M. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform 2004;5:150–63.
[29] Nei M, Gojobori T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol 1986;3:418–26.
[30] Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 2003;19:2496–7.
[31] McDonald JH, Kreitman M. Adaptive protein evolution at the *Adh* locus in *Drosophila*. Nature 1991;351:652–4.
[32] Tajima F. Simple methods for testing the molecular evolutionary clock hypothesis. Genetics 1993;135:599–607.
[33] Haubold B, Wiehe T. Statistics of divergence times. Mol Biol Evol 2001;18:1157–60.
[34] Betrán E, Rozas J, Navarro A, Barbadilla A. The estimation of the number and the length distribution of gene conversion tracts from population DNA sequence data. Genetics 1997;146:89–99.
[35] Mu J, Duan J, Makova KD, et al. Chromosome-wide SNPs reveal an ancient origin for *Plasmodium falciparum*. Nature 2002;418:323–6.
[36] Chung WY, Gardiner DL, Anderson KA, Hyland CA, Kemp DJ, Trenholme KR. The CLAG/RhopH1 locus on chromosome 3 of *Plasmodium falciparum*: two genes or two alleles of the same gene? Mol Biochem Parasitol 2007;151:229–32.
[37] Escalante AA, Ayala FJ. Phylogeny of the malarial genus *Plasmodium*, derived from rRNA gene sequences. Proc Natl Acad Sci USA 1994;91:11373–7.
[38] Goodman M. The genomic record of Humankind's evolutionary roots. Am J Hum Genet 1999;64:31–9.
[39] Hughes AL, Nei M. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. Nature 1988;335:167–70.
[40] Kidgell C, Volkman SK, Daily J, et al. A systematic map of genetic variation in *Plasmodium falciparum*. PLoS Pathog 2006;2:e57.
[41] Sam-Yellowe TY, Shio H, Perkins ME. Secretion of *Plasmodium falciparum* rhoptry protein into the plasma membrane of host erythrocytes. J Cell Biol 1988;106:1507–13.
[42] Douki JB, Sterkers Y, Lepolard C, et al. Adhesion of normal and *Plasmodium falciparum* ring-infected erythrocytes to endothelial cells and the placenta involves the rhoptry-derived ring surface protein-2. Blood 2003;101:5025–32.
[43] Sam-Yellowe TY, Fujioka H, Aikawa M, Hall T, Drazba JA. A *Plasmodium falciparum* protein located in Maurer's clefts underneath knobs and protein localization in association with Rhop-3 and SERA in the intracellular network of infected erythrocytes. Parasitol Res 2001;87:173–85.
[44] Vincensini L, Richert S, Blisnick T, et al. Proteomic analysis identifies novel proteins of the Maurer's clefts, a secretory compartment delivering *Plasmodium falciparum* proteins to the surface of its host cell. Mol Cell Proteomics 2005;4:582–93.
[45] Ferreira MU, Kaneko O, Kimura M, Liu Q, Kawamoto F, Tanabe K. Allelic diversity at the merozoite surface protein-1 (MSP-1) locus in natural *Plasmodium falciparum* populations: a brief overview. Mem Inst Oswaldo Cruz 1998;93:631–8.
[46] Sakihama N, Ohmae H, Bakote B, Kawabata M, Hirayama K, Tanabe K. Limited allelic diversity of *Plasmodium falciparum* merozoite surface protein 1 gene from populations in the Solomon Islands. Am J Trop Med Hyg 2006;74:31–40.
[47] Rammensee H, Bachmann J, Emmerich NP, Bachor OA, Stevanovic S. SYFPEITHI: database for MHC ligands and peptide motifs. Immunogenetics 1999;50:213–9.
[48] Nielsen KM, Kasper J, Choi M, et al. Gene conversion as a source of nucleotide diversity in *Plasmodium falciparum*. Mol Biol Evol 2003;20:726–34.

[49] Freitas-Junior LH, Bottius E, Pirrit LA, et al. Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of *P. falciparum*. Nature 2000;407:1018–22.

[50] Dzikowski R, Deitsch K. Antigenic variation by protozoan parasites: insights from *Babesia bovis*. Mol Microbiol 2006;59:364–6.

[51] Kraemer SM, Kyes SA, Aggarwal G, et al. Patterns of gene recombination shape *var* gene repertoires in *Plasmodium falciparum*: comparisons of geographically diverse isolates. BMC Genomics 2007;8:45.

[52] Tanabe K, Sakihama N, Hattori T, et al. Genetic distance in housekeeping genes between *Plasmodium falciparum* and *Plasmodium reichenowi* and within *P. falciparum*. J Mol Evol 2004;59:687–94.

[53] Polley SD, Weedall GD, Thomas AW, Golightly LM, Conway DJ. Orthologous gene sequences of merozoite surface protein 1 (MSP-1) from *Plasmodium reichenowi* and *P. gallinaceum* confirm an ancient divergence of *P. falciparum* alleles. Mol Biochem Parasitol 2005;142:25–31.

[54] Roy SW, Ferreira MU, Hartl DL. Evolution of allelic dimorphism in malarial surface antigens. Heredity 2008;100:103–10.

[55] Michon P, Stevens JR, Kaneko O, Adams JH. Evolutionary relationships of conserved cysteine-rich motifs in adhesive molecules of malaria parasites. Mol Biol Evol 2002;19:1128–42.

INFECTION AND IMMUNITY, Apr. 2008, p. 1702–1708
0019-9567/08/$08.00+0 doi:10.1128/IAI.01539-07

Vol. 76, No. 4

# Wheat Germ Cell-Free System-Based Production of Malaria Proteins for Discovery of Novel Vaccine Candidates[▽]†

Takafumi Tsuboi,[1,2]* Satoru Takeo,[1] Hideyuki Iriko,[2]‡ Ling Jin,[2] Masateru Tsuchimochi,[1] Shusaku Matsuda,[1] Eun-Taek Han,[1]§ Hitoshi Otsuki,[3] Osamu Kaneko,[3]¶ Jetsumon Sattabongkot,[4] Rachanee Udomsangpetch,[5] Tatsuya Sawasaki,[1] Motomi Torii,[3] and Yaeta Endo[1]

*Cell-Free Science and Technology Research Center[1] and Venture Business Laboratory,[2] Ehime University, Matsuyama, Ehime 790-8577, Japan; Department of Molecular Parasitology, Ehime University Graduate School of Medicine, Toon, Ehime 791-0295, Japan[3]; Department of Entomology, Armed Forces Research Institute of Medical Sciences, Bangkok 10400, Thailand[4]; and Department of Pathobiology, Faculty of Science, Mahidol University, Bangkok 10400, Thailand[5]*

Received 20 November 2007/Returned for modification 22 December 2007/Accepted 1 February 2008

**One of the major bottlenecks in malaria research has been the difficulty in recombinant protein expression. Here, we report the application of the wheat germ cell-free system for the successful production of malaria proteins. For proof of principle, the Pfs25, PfCSP, and PfAMA1 proteins were chosen. These genes contain very high A/T sequences and are also difficult to express as recombinant proteins. In our wheat germ cell-free system, native and codon-optimized versions of the Pfs25 genes produced equal amounts of proteins. PfCSP and PfAMA1 genes without any codon optimization were also expressed. The products were soluble, with yields between 50 and 200 μg/ml of the translation mixture, indicating that the cell-free system can be used to produce malaria proteins without any prior optimization of their biased codon usage. Biochemical and immunocytochemical analyses of antibodies raised in mice against each protein revealed that every antibody retained its high specificity to the parasite protein in question. The development of parasites in mosquitoes fed patient blood carrying *Plasmodium falciparum* gametocytes and supplemented with our mouse anti-Pfs25 sera was strongly inhibited, indicating that both Pfs25-3D7/WG and Pfs25-TBV/WG retained their immunogenicity. Lastly, we carried out a parallel expression assay of proteins of blood-stage *P. falciparum*. The PCR products of 124 *P. falciparum* genes chosen from the available database were used directly in a small-scale format of transcription and translation reactions. Autoradiogram testing revealed the production of 93 proteins. The application of this new cell-free system-based protocol for the discovery of malaria vaccine candidates will be discussed.**

*Plasmodium falciparum* is the protozoan responsible for the widespread return of malaria to tropical countries, particularly in Africa. This reemergence is generally credited to two causes: the development of multidrug-resistant parasites and the development of insecticide-resistant mosquitoes (10). Through decades of work, scientists have learned that vaccination could be a potent curative, but efforts to develop a successful vaccine have not yet succeeded (25). One of the bottlenecks in vaccine development is at the malaria protein production step and is mainly due to the lack of a methodology to enable preparation of quality proteins in an efficient manner. *P. falciparum* genes have a very high A/T content (average, 76% per gene), and a number of them encode repeated stretches of amino acid sequences (8); these features have been proposed as the major factors limiting *P. falciparum* protein expression in cell-based systems. Moreover, the presence of glycosylation machinery in eukaryotic cell-based systems can produce inappropriately glycosylated recombinant malaria proteins, resulting in incorrect immune responses (9, 21, 26). In fact, the three pioneering genome-wide studies on the production of *P. falciparum* proteins in cell-based systems faced serious problems. For instance, Aguiar et al. (1) were able to obtain expression in *Escherichia coli* cells of only 39 of 292 malaria genes cloned into the glutathione *S*-transferase (GST) fusion vector. Mehlin et al. carried out an even more challenging trial in which 1,000 genes encoding relatively small (<450 amino acids) malaria cytosolic proteins were expressed in *E. coli* (24). In that study only 30% of the genes were expressed and only 6.3% of the proteins were soluble, yielding 0.9 mg to 406 mg of protein per liter of culture medium. The other approach used an engineered *E. coli* strain with tRNAs genetically supplemented to allow reading of the high number of A/U codons in malaria mRNA (31). A significant improvement in protein solubility, up to 20.9%, was observed (38 out of 182 proteins tested were soluble). However, although the *E. coli* translation system is known to support folding of prokaryotic and small eukary-

* Corresponding author. Mailing address: Cell-Free Science and Technology Research Center, Ehime University, 3 Bunkyo-cho, Matsuyama, Ehime 790-8577, Japan. Phone: 81-89-927-8277. Fax: 81-89-927-9941. E-mail: tsuboi@ccr.ehime-u.ac.jp.

† Supplemental material for this article may be found at http://iai.asm.org/.

‡ Present address: Division of Medical Zoology, Department of Microbiology and Immunology, Faculty of Medicine, Tottori University, Yonago, Tottori 683-8503, Japan.

§ Present address: Department of Parasitology, Kangwon National University College of Medicine, Chunchon 200-701, Korea.

¶ Present address: Department of Protozoology, Institute of Tropical Medicine, Nagasaki University, Nagasaki 852-8523, Japan.

▽ Published ahead of print on 11 February 2008.

otic proteins, the multidomain proteins common in eukaryotes tend to fold incorrectly in the *E. coli* system, resulting in the formation of inclusion bodies.

Through decades of laborious work, scientists have identified three leading vaccine candidates from the pool of *P. falciparum* proteins: Pfs25 (19), PfCSP (5, 12, 34), and PfAMA1 (6, 11). Pfs25, a zygote/ookinete surface protein, is a promising candidate as a transmission-blocking vaccine. This protein is composed of four tandem epidermal growth factor-like domains, containing three putative N-linked glycosylation sites beside a signal peptide for the attachment of a glycosylphosphatidylinositol moiety (GPI anchor) at the C terminus. These characteristics render Pfs25 very difficult to express (18, 20). PfCSP, with its biased codon usage and lopsided amino acid composition, allows for only a minute amount of protein to be expressed in *E. coli* cells (34). The other antigen candidate is the PfAMA1 gene, which codes for a type 1 integral membrane protein of merozoites and is also difficult to express. Only a synthetic and codon-optimized gene has produced a fairly large amount of PfAMA1 protein in *E. coli* cells. Furthermore, a series of labor-intensive and technically complex refolding processes of the aggregates were required to use the protein as an antigen (6). The fact that only a few vaccine candidates are currently available (23) is most likely the result of difficulties in expressing malarial antigens in high quantity with their correct conformation.

We previously developed a wheat germ cell-free protein synthesis system for practical use in protein production. The system is especially powerful when used for the production of eukaryotic proteins because of its eukaryotic nature. We established two wheat germ cell-free protein protocols for practical use. The first can be used to produce a small amount of protein from a large number of cDNAs, in parallel, for the examination of product qualities and for the genome-wide biochemical annotation of gene products. In this approach, the templates for transcription are constructed using the split-PCR approach (29). The solution resulting from transcription is then directly used as the mRNA source in the small-scale bilayer translation system (28). The second protocol enables the production of large quantities of proteins. In this case, suitable gene products are first selected using the small-scale parallel production method and subsequent functional screening. Genes of interest are then cloned into the pEU plasmid (29), and the mRNA is transcribed. In the translational step, the protein production employs either the bilayer or the discontinuous batch translation method. The bilayer method has acceptable performance for the production of hundreds of micrograms of protein. Since 150 mg of a control protein in a reaction volume of 50 ml was produced in 5 h with the latter reaction method, the cell-free method can be scaled up (27). The system has been acknowledged in the fields of structural and functional genomics of eukaryotes (7, 32) and has proved advantageous due to its capacity to yield high-quality proteins. Taken together, the system seems to be powerful when used for the production of malaria parasite proteins, as no glycosylation takes place during the standard reaction. However, to date, there is no Good Manufacturing Practice facility for production of recombinant proteins for clinical studies using the wheat germ cell-free system in the world. In the present study, we first tested the versatility of the wheat germ cell-free system using as control models the leading vaccine candidate genes from *P. falciparum*. In addition, a series of experiments was conducted to prove the value of the system for the parallel expression of malaria proteins. The results presented here suggest that the wheat germ cell-free system may be useful as an additional protein production method in the field of *P. falciparum* research.

## MATERIALS AND METHODS

**Genomic cloning and construction of genes encoding fragments of Pfs25, PfCSP, and PfAMA1.** The nucleotide sequences for the signal peptide and the GPI anchor were excluded from the expression constructs for genes encoding the PfCSP and Pfs25 proteins. The truncated versions of the PfCSP and Pfs25-3D7 genes were amplified by PCR from the genomic DNA of the *P. falciparum* 3D7 strain and subcloned into pEU3 (a vector carrying the C-terminal $His_6$ tag) (29) at the EcoRV site. The gene encoding Pfs25-TBV was a generous gift from Anthony W. Stowers (NIAID, NIH, Rockville, MD) (35). Pfs25-TBV, a synthetic version of the Pfs25 gene, was codon optimized for expression in the yeast *Saccharomyces cerevisiae*, and the replacement of Asn with Gln at three N-glycosylation sites was performed (20). DNA encoding full-length PfAMA1 protein was amplified from the genomic DNA of *P. falciparum* 3D7 and cloned into pEU-E01-GST (a vector with an N-terminal GST tag followed by a tobacco etch virus protease cleavage site) between the XhoI and BamHI sites. These pEU plasmid vectors are the expression vectors designed specifically for the wheat germ cell-free system (16). The inserted nucleotide sequences were confirmed using the ABI PRISM 310 Genetic Analyzer and the BigDye Terminator v1.1 Cycle Sequencing kit (Applied Biosystems, Foster City, CA).

**Parallel construction of the DNA template from the parasite RNA.** We selected 124 genes annotated as dominantly expressed in the blood stages of *P. falciparum* based on the microarray data integrated in the PlasmoDB database (http://www.plasmodb.org) (see Table S1 in the supplemental material). Extracted total RNA from *P. falciparum* 3D7 asexual blood-stage parasites was reverse transcribed into cDNA by using SuperScript III reverse transcriptase (Invitrogen, Carlsbad, CA), and PCR amplification was performed using *LA Taq* DNA polymerase (Takara Bio, Otsu, Japan). The 5′ primers were designed as 46-mers: 16-mer nucleotide sequences (5′-CCACCCACCACCACCA) as the S1 tag sequence followed by a 30-mer of unique sequence covering each 5′ open reading frame containing the start codon. For the 3′ primers, 30-mer nucleotide sequences covering each unique sequence upstream from the termination codon were prepared. The PCR products were then cloned into the pCR2.1 plasmid using a TOPO TA cloning kit (Invitrogen), and their sequences at both ends were confirmed. Translation templates were prepared by in vitro transcription from PCR products amplified by the split-primer PCR method described earlier (29).

**Production and purification of the Pfs25-3D7/WG, Pfs25-TBV/WG, PfCSP/WG, and PfAMA1/WG proteins.** We employed the wheat germ cell-free protein expression system for protein production using the bilayer translation reaction method described previously (28). Briefly, 250 μl of transcription mixture containing 25 μg of the plasmid DNA, 80 mM HEPES-KOH, pH 7.8, 16 mM magnesium acetate, 2 mM spermidine, 10 mM dithiothreitol, 2.5 mM each of nucleoside triphosphates, 250 U of SP6 RNA polymerase (Promega, Madison, WI), and 250 U of RNasin (Promega) was incubated for 6 h at 37°C. After the incubation, the transcription solution containing transcribed mRNA was mixed with 250 μl of wheat germ extract (60 $A_{260}$ units) supplemented with 2 μl of 20-mg/ml creatine kinase in a single well of a six-well plate. The 5.5-ml substrate mix (30 mM HEPES-KOH, pH 7.8, 100 mM potassium acetate, 2.7 mM magnesium acetate, 0.4 mM spermidine, 2.5 mM dithiothreitol, 0.3 mM amino acid mix, 1.2 mM ATP, 0.25 mM GTP, and 16 mM creatine phosphate) from the ENDEXT Wheat Germ Expression S kit (CFS Co., Ltd., Matsuyama, Japan) was then added on top of the translation mix and incubated at 26°C for 12 h. After incubation, the reaction mixture was centrifuged at 21,900 × *g* for 20 min. Recovered supernatants were passed through Amicon Ultra centrifugal filter units (10-kDa molecular mass cutoff) (Millipore, Billerica, MA) to replace the translation buffer with phosphate-buffered saline. The samples containing the synthesized Pfs25-3D7/WG, Pfs25-TBV/WG, and PfCSP/WG proteins were purified using the Ni-nitrilotriacetic acid agarose column (Qiagen, Valencia, CA). The PfAMA1/WG protein was purified by passing the supernatant through the glutathione-Sepharose 4B column (GE Healthcare Bio-Sciences, Piscataway, NJ), followed by tobacco etch virus protease (Invitrogen) cleavage to remove the

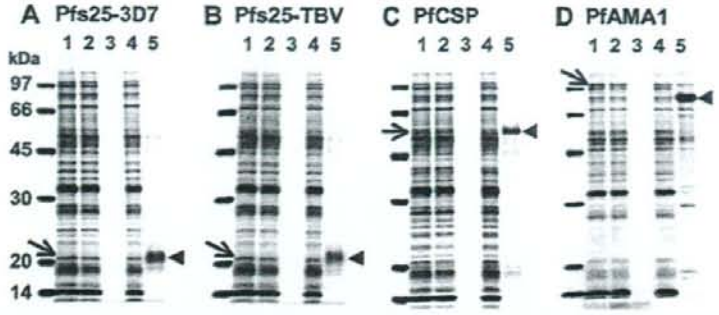


FIG. 1. SDS-PAGE analysis of the proteins expressed in the wheat germ cell-free system. Pfs25-3D7/WG (A), Pfs25-TBV/WG (B), PfCSP/WG (C), and PfAMA1/WG (D) were separated on SDS-12.5% polyacrylamide gels under reducing conditions and stained with Coomassie brilliant blue. The samples in each gel were as follows: total reaction mixture (lane 1), supernatant and precipitated fractions after brief centrifugation (lanes 2 and 3, respectively), and unbound and affinity-purified proteins (lanes 4 and 5, respectively). Products and purified proteins are indicated by arrows and arrowheads, respectively.

GST tag. Concentrations of affinity-purified proteins were determined using the Bradford protein assay kit (Bio-Rad Laboratories, Hercules, CA). Protein samples were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) under reducing conditions (22), and the bands were visualized with Coomassie brilliant blue. Purified protein samples were stored in aliquots at −80°C until further use. For parallel protein synthesis from 124 malaria genes, the transcription and translation reactions were performed by a method similar to that described above. The 125-μl substrate mixture was overlaid on top of the 25-μl translation mixture containing transcribed mRNA in the presence of U-[$^{14}$C]Leu (11.1 kBq; 15 GBq/mmol of Leu). The reaction was performed in 96-well plates. Proteins were separated by SDS-PAGE and identified by autoradiography using an imaging analyzer (BAS-2500; Fujifilm, Tokyo, Japan). The solubility of each product was expressed as the percentage of trichloroacetic acid-insoluble radioactivity (counted using a liquid scintillation counter [LSC-6100; Aloka, Mitaka, Japan]) in a supernatant fraction recovered from centrifugation at 21,900 × *g* for 20 min compared to that of the total reaction mixture. The amount of target protein was estimated using the following formula where count is the radioactivity of the protein produced; Leu is the number of Leu residues in the protein, used to estimate the moles of Leu incorporated; MW is molecular weight; and ratio is the ratio of intensity of a specific protein band to the total intensity of bands on the autoradiogram: protein concentration = count/Leu × MW × ratio.

**Preparation of antiserum.** Groups of female BALB/c mice (five mice in each group) were subcutaneously immunized three times in the 1st, 3rd, and 5th weeks with 10 μg of affinity-purified proteins emulsified in Freund's adjuvant. As the control, a group of mice was administered GST in Freund's adjuvant, using the same protocol as described above. Antiserum preparation was as described elsewhere (2).

**Preparation of *P. falciparum* asexual blood-stage parasites, ookinetes, and sporozoites.** A mature schizont-rich fraction was obtained from cultured *P. falciparum* strain 3D7 (30). Parasite pellets were kept at −80°C until extract preparation.

To obtain ookinetes and sporozoites of *P. falciparum*, we used parasites derived from patient blood. The use of all human materials in this study was reviewed and approved by the Institutional Ethics Committee of the Thai Ministry of Public Health and the Human Subjects Research Review Board of the United States Army. Peripheral blood was collected with heparinized syringes under written informed consent from patients who came to the malaria clinics in the Mae Sod district, Thailand. Infection with *P. falciparum* was confirmed by the microscopic observation of Giemsa-stained thick and thin blood smears. The gametocytemic patient blood was divided into two parts. One was used to grow zygotes/ookinetes in vitro for both Western blotting and immunocytochemical analyses, and the other half was subjected to propagation of sporozoites in mosquitoes for two further analyses, as described elsewhere (33). Western blot analysis and immunocytochemistry were performed as described previously (3, 17).

**Transmission-blocking assays.** We collected 20 ml of peripheral blood from a volunteer patient. Blood was divided into aliquots (300 μl/tube) and briefly centrifuged, and plasma was discarded. Mouse immune sera against both Pfs25-3D7/WG and Pfs25-TBV/WG were serially diluted with heat-inactivated normal human serum prepared from malaria-naïve donors. Next, 180 μl of each diluted solution was added to the *P. falciparum*-infected blood cells and incubated for 15 min at room temperature. The mixture was placed in a membrane feeding apparatus kept at 37°C to allow *Anopheles dirus* A mosquitoes to feed on the blood in each apparatus for 30 min. Fully engorged mosquitoes were maintained for a week in the insectary. Oocysts that developed within the midgut were counted from 20 randomly selected mosquitoes. The Kruskal-Wallis test was applied to examine the differences in oocyst counts per mosquito between immunized groups and the control group fed on mouse serum raised against GST. Probability values (*P*) of less than 0.05 were considered statistically significant.

## RESULTS AND DISCUSSION

We were able to successfully express the Pfs25/WGs, PfCSP/WG, and PfAMA1/WG proteins using the wheat germ cell-free system. Expression of the Pfs25 (Pfs25-3D7/WG) protein from a gene with a native nucleotide composition was shown by subsequent SDS-PAGE analysis (Fig. 1A) to be comparable in amount to that of Pfs25-TBV/WG (Fig. 1B) expressed from the codon-optimized synthetic gene. On the SDS-polyacrylamide gels, two protein bands appeared at 20 kDa, the expected mobility of the Pfs25 truncated form, lacking the signal peptide and the GPI anchor. Almost all of the Pfs25-3D7/WG protein from the biased DNA was recovered in the supernatant fraction (Fig. 1A, lane 2) and was easily purified as a single dominant band along with other nonspecific faint bands by affinity chromatography (Fig. 1A, lane 5). The amount of purified Pfs25-3D7/WG was 35 μg per 6.0 ml of the reaction mixture, while that obtained from the codon-optimized gene was comparable, at 30 μg protein per reaction mixture. These results demonstrate that the wheat germ cell-free system that we employed produced equal amounts of proteins with and without prior optimization of their biased codon usage in the DNA. Similarly, the amounts of the other two proteins, PfCSP/WG (Fig. 1C), and PfAMA1/WG (Fig. 1D), produced from a gene with a native nucleotide composition were 26 and 102 μg per reaction, respectively.

**Immunological characterization of the protein products.** To determine the creation of conformation-dependent epitopes in Pfs25 and AMA1, we examined and confirmed the reactivity of

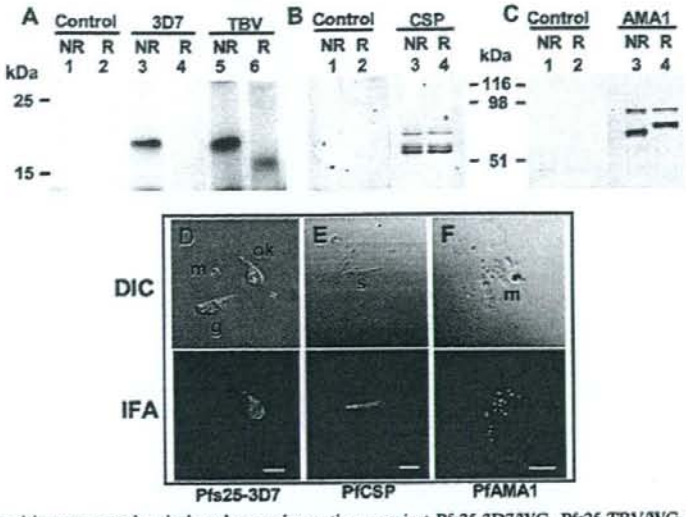


FIG. 2. Western blot and immunocytochemical analyses using antisera against Pfs25-3D7/WG, Pfs25-TBV/WG, PfCSP/WG, and PfAMA1/WG. Extracts prepared from *Plasmodium falciparum* zygotes/ookinetes (A), sporozoites (B), and schizonts (C) were separated on SDS-12.5% polyacrylamide gels under nonreducing (NR; lanes 1, 3, and 5) and reducing (R; lanes 2, 4, and 6) conditions. (A) Proteins on polyvinylidene fluoride membranes were immunostained either with mouse anti-Pfs25-3D7/WG serum (lanes 3 and 4) or mouse anti-Pfs25-TBV/WG serum (lanes 5 and 6) or with the negative-control serum (lanes 1 and 2). (B) The membrane was immunostained with either mouse anti-PfCSP/WG serum (lanes 3 and 4) or the control serum (lanes 1 and 2). (C) The membrane was immunostained with either mouse anti-PfAMA1/WG serum (lanes 3 and 4) or the control serum (lanes 1 and 2). (D to F) Samples prepared from *Plasmodium falciparum* immature ookinete (D), sporozoite (E), and schizont (F) were immunostained with the antiserum indicated at the bottom of the panel. Upper panels represent images obtained by differential interference contrast (DIC) microscopy, and lower panels represent immunostained images (immunofluorescence assay [IFA]) visualized with goat anti-mouse immunoglobulin G-fluorescein conjugate. These images have been taken by confocal scanning laser microscopy (LSM5 PASCAL; Carl Zeiss MicroImaging, Thornwood, NY). g, gametocyte; ok, immature ookinete; m, merozoite; s, sporozoite. Bars, 5 μm.

anti-Pfs25 conformation-specific monoclonal antibody 4B7 (a generous gift from Carole A. Long [NIAID, NIH, Rockville, MD]) against Pfs25/WGs and the reactivity of anti-PfAMA1 3D7 conformation-specific monoclonal antibody 1E9 (a generous gift from Carole A. Long) against PfAMA1/WG by Western blotting under nonreducing conditions (data not shown). To evaluate the immunogenicity of each protein prepared, we then raised mouse antisera and determined their reactivity to the parasite-derived native proteins. Extract from approximately $5 \times 10^5$ zygotes/ookinetes per lane was separated by SDS-PAGE, and Western blot analysis was performed. Specific bands with the expected mobility of native Pfs25 protein were detected under nonreducing conditions using antisera against Pfs25-3D7/WG and Pfs25-TBV/WG. Anti-Pfs25-3D7/WG serum did not show any reactivity under reducing conditions (Fig. 2A). These results suggest that the Pfs25-3D7/WG protein prepared here retained a conformation similar to that of the native protein. The identity of the faint band detected at the lower position with anti-Pfs25-TBV/WG under reducing conditions is unclear at present (Fig. 2A). Similar experiments were performed using anti-PfCSP/WG and anti-PfAMA1/WG sera to study extracts from respective stages of the parasite. The analyses clearly showed specific reactivity of each antiserum to PfCSP and PfAMA1 proteins (Fig. 2B and C). Anti-PfCSP serum reacted to three protein bands in the sporozoite extract under both reducing and nonreducing conditions (Fig. 2B). The upper and lower bands appeared to correspond to precursor and mature forms, respectively, as reported earlier by Coppi et al. (4). Anti-PfAMA1 serum gave two signals, with the upper and lower bands corresponding to mature and processed forms, respectively (15). The signal shift of the two bands upon introduction of a reducing reagent was most likely due to the high content of disulfide bonds within the protein (14). These results are consistent with previously reported findings (13).

Immunocytochemical staining was performed against immature ookinetes obtained by in vitro short-term culture using anti-Pfs25-3D7/WG. As shown in Fig. 2D (differential interference contrast and immunofluorescence assay), the antiserum specifically stained the surface of the immature ookinete but not the gametocyte and the merozoite. Antiserum against Pfs25-TBV/WG yielded similar results (data not shown). These findings were consistent with our previous report in which Pfs25-TBV prepared from yeast cells was used to raise antiserum (2). These findings also verified that Pfs25 prepared using our protocols from a gene with an A/T-rich native nucleotide composition can yield a protein of sufficient quality to raise a specific antibody. Experiments using anti-PfCSP/WG and anti-PfAMA1/WG on the target stages of the parasite showed typical staining patterns. The entire surface of the slender sporozoite was stained by anti-PfCSP/WG serum (Fig. 2E), and the anti-PfAMA1/WG serum clearly visualized punc-

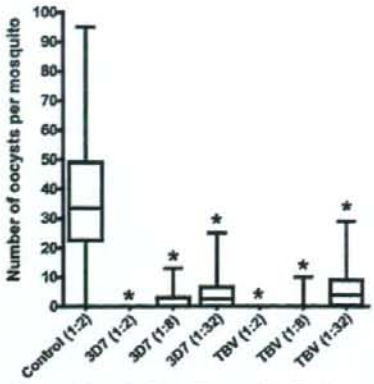


FIG. 3. Transmission-blocking efficacy of antibodies against *Plasmodium falciparum* parasites. The median numbers of oocysts per mosquito ($n = 20$) with quartiles (box plot) and ranges (lines on both top and bottom of the box) were compared among groups of mosquitoes fed on either anti-Pfs25 sera serially diluted or control mouse serum. Dilution of test immune serum is shown as 1:2 to 1:32. Statistical analysis was performed using the Kruskal-Wallis test for comparison of oocyst numbers between the test immune sera and control serum. Asterisks indicate statistically significant differences compared to the control group ($P < 0.05$).

tate localization of PfAMA1 at the apical end of merozoites (Fig. 2F).

**Efficacy evaluation of the proteins as vaccine antigens.** In view of a practical application of the system for discovery of malaria vaccine candidates, we evaluated the quality of antigens produced by performing a parasite growth inhibition assay using the antibodies raised against those antigens. We focused on Pfs25-3D7/WG and Pfs25-TBV/WG. Pfs25-TBV is currently the sole transmission-blocking vaccine candidate under clinical trial (23). A transmission-blocking assay was performed using both anti-Pf25-3D7/WG and anti-Pfs25-TBV/WG. A mixture containing *P. falciparum*-gametocyte infected erythrocytes and one of the antisera was fed to mosquitoes. The number of developed oocysts in the mosquitoes was then later counted. Both antisera at twofold dilution completely inhibited oocyst development, as we have seen no mosquito harboring oocysts (Fig. 3). The number of oocysts was inversely proportional to the concentration of antiserum added, findings consistent with previous experiments, in which Pfs25-TBV prepared from yeast was used to raise antiserum (2). It is important at this moment to stress the difference between this study and other studies: our proteins were produced from a non-codon-optimized gene in a cell-free system, while in other studies a codon-optimized engineered Pfs25-TBV gene was transformed into yeast cells (20). The results presented here strongly indicate the value of the wheat germ cell-free system for the production of malaria proteins that require complicated procedures in other systems.

**Parallel syntheses of *P. falciparum* proteins.** Although cell-based expression systems have been widely used in this field, they are limited mainly in their ability for efficient production of *P. falciparum* protein, primarily because of the complexity of the genome. In order to evaluate the capability of our cell-free system for parallel expression from the parasite genes, we selected 124 genes (see Table S1 in the supplemental material) encoding asexual blood-stage parasite proteins, based on the PlasmoDB database. Autoradiography demonstrated that 93 of the 124 genes yielded protein products. The average yield of expressed protein estimated for each full-size product was 1.9 μg per 150 μl of reaction mixture, an amount sufficient for preliminary antigen discovery studies using hyperimmune serum. Average protein solubility was 65% (see Table S1 in the supplemental material). There was significant inverse correlation between yield and molecular size of the protein; the greater the size, the lower the protein yield. There was also weak but significant inverse correlation between the protein yield and the relative frequency of low-complexity regions. In addition, solubility was inversely correlated with the pI value (Table 1). These observations have already been documented in earlier studies (24, 31). Surprisingly, we did not see any correlation between yield and A/T content, pI value, or the existence of a transmembrane domain (data not shown). We then analyzed the statistical difference in molecular weights, pI values, A/T contents, and relative frequencies of low-complexity regions between the expressed and nonexpressed groups of molecules, using the Mann-Whitney U test. The molecular weights in the nonexpressed group were significantly higher than those of the expressed group ($P < 0.0001$). In contrast, pI values, A/T contents, and the relative frequencies of low-complexity regions did not differ significantly (see Table S1 in the supplemental material). We currently have no explanation for why 25% of the tested genes failed to produce proteins in our system. One possible explanation is the sequence errors most likely present in the PCR products that were used as templates for transcription and subsequent translation. Such templates would cause mistranslation of the protein by frameshift.

In summary, the ability of the wheat germ cell-free protein synthesis system to produce *P. falciparum* proteins was examined. We found that (i) without the need for codon optimization, the cell-free system is able to produce a sufficient amount of high-quality proteins of the leading malaria vaccine candidates, Pfs25, PfCSP, and PfAMA1; (ii) biochemical, immunocytochemical, and biological analyses demonstrated that the prepared proteins could be directly used for immunization after a simple affinity purification step; and (iii) the system proved suitable for use as a parallel

TABLE 1. Correlation of expression or solubility and characteristics[a]

| Parameter | Correlation coefficient (*P* value) | |
|---|---|---|
| | Protein concn | % Solubility |
| Mol wt | −0.3177 (0.0019)[b] | −0.1221 (0.2436) |
| pI | −0.1214 (0.2464) | −0.3519 (0.0005)[b] |
| % A/T | −0.1505 (0.1498) | |
| Low complexity[c] | −0.2276 (0.0283)[b] | |
| % Solubility | −0.0494 (0.6385) | |

[a] Spearman's correlation coefficients by rank were calculated among the 93 proteins expressed. The probability values of the statistical significance are shown in parentheses.
[b] $P < 0.05$ was considered to indicate a statistically significant correlation.
[c] Relative frequency of low-complexity regions per molecular weight.

way to produce parasite proteins. We believe that the wheat germ cell-free protein synthesis system may be a key tool for decoding genetic information above and beyond malaria vaccine research.

## ACKNOWLEDGMENTS

We thank Jeeraphat Sirichaisinthop and the staff of the Vector Borne Disease Training Center, Pra Budhabat, Saraburi, Thailand, for assistance in setting up the field sites and the staff of the Department of Entomology, AFRIMS, Bangkok, Thailand, as well as Hiroko Suzuki, Limei Yin, Kana Kato, and Aya Tamai for their technical assistance. We are grateful to Ivona Kozieradzki for critical reading of the manuscript and valuable comments.

This work was supported in part by Grants-in-Aid for Scientific Research (18390129 and 19406009) and Scientific Research on Priority Areas (19041053) from the Ministry of Education, Culture, Sports, Science and Technology of Japan and by a Grant-in-Aid from the Ministry of Health, Labor and Welfare (H17-Sinkou-ippan-019) of Japan.

## REFERENCES

1. **Aguiar, J. C., J. LaBaer, P. L. Blair, V. Y. Shamailova, M. Koundinya, J. A. Russell, F. Huang, W. Mar, R. M. Anthony, A. Witney, S. R. Caruana, L. Brizuela, J. B. Sacci, Jr., S. L. Hoffman, and D. J. Carucci.** 2004. High-throughput generation of *P. falciparum* functional molecules by recombinational cloning. Genome Res. **14:**2076–2082.
2. **Arakawa, T., A. Komesu, H. Otsuki, J. Sattabongkot, R. Udomsangpetch, Y. Matsumoto, N. Tsuji, Y. Wu, M. Torii, and T. Tsuboi.** 2005. Nasal immunization with a malaria transmission-blocking vaccine candidate, Pfs25, induces complete protective immunity in mice against field isolates of *Plasmodium falciparum*. Infect. Immun. **73:**7375–7380.
3. **Arakawa, T., T. Tsuboi, A. Kishimoto, J. Sattabongkot, N. Suwanabun, T. Rungruang, Y. Matsumoto, N. Tsuji, H. Hisaeda, A. Stowers, L. Shimabukuro, Y. Sato, and M. Torii.** 2003. Serum antibodies induced by intranasal immunization of mice with *Plasmodium vivax* Pvs25 co-administered with cholera toxin completely block parasite transmission to mosquitoes. Vaccine **21:**3143–3148.
4. **Coppi, A., C. Pinzon-Ortiz, C. Hutter, and P. Sinnis.** 2005. The *Plasmodium* circumsporozoite protein is proteolytically processed during cell invasion. J. Exp. Med. **201:**27–33.
5. **Dame, J. B., J. L. Williams, T. F. McCutchan, J. L. Weber, R. A. Wirtz, W. T. Hockmeyer, W. L. Maloy, J. D. Haynes, I. Schneider, D. Roberts, et al.** 1984. Structure of the gene encoding the immunodominant surface antigen on the sporozoite of the human malaria parasite *Plasmodium falciparum*. Science **225:**593–599.
6. **Dutta, S., P. V. Lalitha, L. A. Ware, A. Barbosa, J. K. Moch, M. A. Vassell, B. B. Fileta, S. Kitov, N. Kolodny, D. G. Heppner, J. D. Haynes, and D. E. Lanar.** 2002. Purification, characterization, and immunogenicity of the refolded ectodomain of the *Plasmodium falciparum* apical membrane antigen 1 expressed in *Escherichia coli*. Infect. Immun. **70:**3101–3110.
7. **Endo, Y., and T. Sawasaki.** 2006. Cell-free expression systems for eukaryotic protein production. Curr. Opin. Biotechnol. **17:**373–380.
8. **Gardner, M. J., N. Hall, E. Fung, O. White, M. Berriman, R. W. Hyman, J. M. Carlton, A. Pain, K. E. Nelson, S. Bowman, I. T. Paulsen, K. James, J. A. Eisen, K. Rutherford, S. L. Salzberg, A. Craig, S. Kyes, M. S. Chan, V. Nene, S. J. Shallom, B. Suh, J. Peterson, S. Angiuoli, M. Pertea, J. Allen, J. Selengut, D. Haft, M. W. Mather, A. B. Vaidya, D. M. Martin, A. H. Fairlamb, M. J. Fraunholz, D. S. Roos, S. A. Ralph, G. I. McFadden, L. M. Cummings, G. M. Subramanian, C. Mungall, J. C. Venter, D. J. Carucci, S. L. Hoffman, C. Newbold, R. W. Davis, C. M. Fraser, and B. Barrell.** 2002. Genome sequence of the human malaria parasite *Plasmodium falciparum*. Nature **419:**498–511.
9. **Gowda, D. C., and E. A. Davidson.** 1999. Protein glycosylation in the malaria parasite. Parasitol. Today **15:**147–152.
10. **Greenwood, B., and T. Mutabingwa.** 2002. Malaria in 2002. Nature **415:**670–672.
11. **Gupta, A., T. Bai, V. Murphy, P. Strike, R. F. Anders, and A. H. Batchelor.** 2005. Refolding, purification, and crystallization of apical membrane antigen 1 from *Plasmodium falciparum*. Protein Expr. Purif. **41:**186–198.
12. **Heppner, D. G., Jr., K. E. Kester, C. F. Ockenhouse, N. Tornieporth, O. Ofori, J. A. Lyon, V. A. Stewart, P. Dubois, D. E. Lanar, U. Krzych, P. Moris, E. Angov, J. F. Cummings, A. Leach, B. T. Hall, S. Dutta, R. Schwenk, C. Hillier, A. Barbosa, L. A. Ware, L. Nair, C. A. Darko, M. R. Withers, B. Ogutu, M. E. Polhemus, M. Fukuda, S. Pichyangkul, M. Gettyacamin, C. Diggs, L. Soisson, J. Milman, M. C. Dubois, N. Garcon, K. Tucker, J. Wittes, C. V. Plowe, M. A. Thera, O. K. Doumbo, M. G. Pau, J. Goudsmit, W. R. Ballou, and J. Cohen.** 2005. Towards an RTS, S-based, multi-stage, multi-antigen vaccine against falciparum malaria: progress at the Walter Reed Army Institute of Research. Vaccine **23:**2243–2250.
13. **Hodder, A. N., P. E. Crewther, and R. F. Anders.** 2001. Specificity of the protective antibody response to apical membrane antigen 1. Infect. Immun. **69:**3286–3294.
14. **Hodder, A. N., P. E. Crewther, M. L. Matthew, G. E. Reid, R. L. Moritz, R. J. Simpson, and R. F. Anders.** 1996. The disulfide bond structure of *Plasmodium* apical membrane antigen-1. J. Biol. Chem. **271:**29446–29452.
15. **Howell, S. A., C. Withers-Martinez, C. H. Kocken, A. W. Thomas, and M. J. Blackman.** 2001. Proteolytic processing and primary structure of *Plasmodium falciparum* apical membrane antigen-1. J. Biol. Chem. **276:**31311–31320.
16. **Kamura, N., T. Sawasaki, Y. Kasahara, K. Takai, and Y. Endo.** 2005. Selection of 5′-untranslated sequences that enhance initiation of translation in a cell-free protein synthesis system from wheat embryos. Bioorg. Med. Chem. Lett. **15:**5402–5406.
17. **Kaneko, O., B. Y. Y. Lim, H. Iriko, I. T. Ling, H. Otsuki, M. Grainger, T. Tsuboi, J. H. Adams, D. Mattei, A. A. Holder, and M. Torii.** 2005. Apical expression of three RhopH1/Clag proteins as components of the *Plasmodium falciparum* RhopH complex. Mol. Biochem. Parasitol. **143:**20–28.
18. **Kaslow, D. C., I. C. Bathurst, T. Lensen, T. Ponnudurai, P. J. Barr, and D. B. Keister.** 1994. *Saccharomyces cerevisiae* recombinant Pfs25 adsorbed to alum elicits antibodies that block transmission of *Plasmodium falciparum*. Infect. Immun. **62:**5576–5580.
19. **Kaslow, D. C., I. A. Quakyi, C. Syin, M. G. Raum, D. B. Keister, J. E. Coligan, T. F. McCutchan, and L. H. Miller.** 1988. A vaccine candidate from the sexual stage of human malaria that contains EGF-like domains. Nature **333:**74–76.
20. **Kaslow, D. C., and J. Shiloach.** 1994. Production, purification and immunogenicity of a malaria transmission-blocking vaccine candidate: TBV25H expressed in yeast and purified using nickel-NTA agarose. Biotechnology (New York) **12:**494–499.
21. **Kedees, M. H., N. Azzouz, P. Gerold, H. Shams-Eldin, J. Iqbal, V. Eckert, and R. T. Schwarz.** 2002. *Plasmodium falciparum*: glycosylation status of *Plasmodium falciparum* circumsporozoite protein expressed in the baculovirus system. Exp. Parasitol. **101:**64–68.
22. **Laemmli, U. K.** 1970. Cleavage of structural proteins during assembly of the head of bacteriophage T4. Nature **227:**680–685.
23. **Malkin, E., F. Dubovsky, and M. Moree.** 2006. Progress towards the development of malaria vaccines. Trends Parasitol. **22:**292–295.
24. **Mehlin, C., E. Boni, F. S. Buckner, L. Engel, T. Feist, M. H. Gelb, L. Haji, D. Kim, C. Liu, N. Mueller, P. J. Myler, J. T. Reddy, J. N. Sampson, E. Subramanian, W. C. Van Voorhis, E. Worthey, F. Zucker, and W. G. Hol.** 2006. Heterologous expression of proteins from *Plasmodium falciparum*: results from 1000 genes. Mol. Biochem. Parasitol. **148:**144–160.
25. **Richie, T. L., and A. Saul.** 2002. Progress and challenges for malaria vaccines. Nature **415:**694–701.
26. **Samuelson, J., S. Banerjee, P. Magnelli, J. Cui, D. J. Kelleher, R. Gilmore, and P. W. Robbins.** 2005. The diversity of dolichol-linked precursors to Asn-linked glycans likely results from secondary loss of sets of glycosyltransferases. Proc. Natl. Acad. Sci. USA **102:**1548–1553.
27. **Sawasaki, T., and Y. Endo.** 2008. The wheat germ cell-free protein synthesis system, p. 111–139. *In* A. S. Spirin and J. R. Swartz (ed.), Cell-free protein synthesis, methods and protocols. Wiley-VCH Verlag GmbH and Co. KGaA, Weinheim, Germany.
28. **Sawasaki, T., Y. Hasegawa, M. Tsuchimochi, N. Kamura, T. Ogasawara, T. Kuroita, and Y. Endo.** 2002. A bilayer cell-free protein synthesis system for high-throughput screening of gene products. FEBS Lett. **514:**102–105.
29. **Sawasaki, T., T. Ogasawara, R. Morishita, and Y. Endo.** 2002. A cell-free protein synthesis system for high-throughput proteomics. Proc. Natl. Acad. Sci. USA **99:**14652–14657.
30. **Trager, W., and J. B. Jensen.** 1976. Human malaria parasites in continuous culture. Science **193:**673–675.
31. **Vedadi, M., J. Lew, J. Artz, M. Amani, Y. Zhao, A. Dong, G. A. Wasney, M. Gao, T. Hills, S. Brokx, W. Qiu, S. Sharma, A. Diassiti, Z. Alam, M. Melone, A. Mulichak, A. Wernimont, J. Bray, P. Loppnau, O. Plotnikova, K. Newberry, E. Sundararajan, S. Houston, J. Walker, W. Tempel, A. Bochkarev, I. Kozieradzki, A. Edwards, C. Arrowsmith, D. Roos, K. Kain, and R. Hui.** 2007. Genome-scale protein expression and structural biology of *Plasmodium falciparum* and related Apicomplexan organisms. Mol. Biochem. Parasitol. **151:**100–110.
32. **Vinarov, D. A., B. L. Lytle, F. C. Peterson, E. M. Tyler, B. F. Volkman, and J. L. Markley.** 2004. Cell-free protein production and labeling protocol for NMR-based structural proteomics. Nat. Methods **1:**149–153.

33. **Wirtz, R. A., J. Sattabongkot, T. Hall, T. R. Burkot, and R. Rosenberg.** 1992. Development and evaluation of an enzyme-linked immunosorbent assay for *Plasmodium vivax*-VK247 sporozoites. J. Med. Entomol. **29**:854–857.
34. **Young, J. F., W. T. Hockmeyer, M. Gross, W. R. Ballou, R. A. Wirtz, J. H. Trosper, R. L. Beaudoin, M. R. Hollingdale, L. H. Miller, C. L. Diggs, and M. Rosenberg.** 1985. Expression of *Plasmodium falciparum* circumsporozoite proteins in *Escherichia coli* for potential use in a human malaria vaccine. Science **228**:958–962.
35. **Zou, L., A. P. Miles, J. Wang, and A. W. Stowers.** 2003. Expression of malaria transmission-blocking vaccine antigen Pfs25 in Pichia pastoris for use in human clinical trials. Vaccine **21**:1650–1657.

*Editor:* W. A. Petri, Jr.

Molecular & Biochemical Parasitology 159 (2008) 142–145

Contents lists available at ScienceDirect

Molecular & Biochemical Parasitology

Short communication

# Disruption of the *Plasmodium berghei* 2-Cys peroxiredoxin TPx-1 gene hinders the sporozoite development in the vector mosquito

Kazuhiko Yano[a], Hitoshi Otsuki[b], Meiji Arai[c,1], Kanako Komaki-Yasuda[a,d], Takafumi Tsuboi[e], Motomi Torii[b], Shigeyuki Kano[a], Shin-Ichiro Kawazu[a,d,f,*]

[a] Research Institute, International Medical Center of Japan, 1-21-1 Toyama, Shinjuku-ku, Tokyo 162-8655, Japan
[b] Department of Molecular Parasitology, Ehime University School of Medicine, Toon, Ehime 791-0295, Japan
[c] Division of Medical Zoology, Department of Infection and Immunity, Jichi Medical University School of Medicine, Shimotsuke, Tochigi 329-0498, Japan
[d] Precursory Research for Embryonic Science and Technology, Japan Science and Technology Agency, 5, Sanbancho, Chiyoda-ku, Tokyo 102-0075, Japan
[e] Cell-Free Science and Technology Research Center, Ehime University, Matsuyama, Ehime 790-8577, Japan
[f] National Research Center for Protozoan Diseases, Obihiro University of Agriculture and Veterinary Medicine, 2-13 Inada-cho, Obihiro, Hokkaido 080-8555, Japan

ARTICLE INFO

*Article history:*
Received 19 December 2007
Received in revised form 5 March 2008
Accepted 5 March 2008
Available online 13 March 2008

*Keywords:*
Insect stage
Peroxiredoxin
*Plasmodium berghei*
Thioredoxin peroxidase

ABSTRACT

To investigate the physiologic role of cytosolic 2-Cys peroxiredoxin of *Plasmodium berghei* (PbTPx-1), we infected the vector mosquito *Anopheles stephensi* with a parasite carrying a targeted knockout of *pbtpx-1* (Prx-KO). The number of Prx-KO midgut oocysts at 14–15 days post-feeding (pf) was comparable to that of the parent strain (WT); however, the numbers of sporozoites that formed in midgut oocysts and accumulated in the salivary gland of Prx-KO-infected mosquitoes by 21 days pf were decreased to 10–20% and 3–10%, respectively, of those values in WT-infected mosquitoes. A higher frequency of DNA strand breaks was detected in Prx-KO oocysts than in WT oocysts. Sporozoites carrying the targeted disruption had reduced infectivity in mice; however, the knockout did not affect the ability of the sporozoite to reach the liver parenchyma and initiate exo-erythrocytic form (EEF) development. TPx-1 may be involved in development during exponentially multiplying stages, such as sporozoites and EEF.



As *Plasmodium* spp. actively proliferate in erythrocytes of their vertebrate hosts, the parasites are subjected to the toxic effects of reactive oxygen species (ROS) through their asexual development [1,2]. In contrast, the parasites fertilize and multiply in the digestive tract (midgut) of the vector mosquito and subsequently mature intracellularly in the salivary grand. In these environments, the parasites are also likely to be subject to oxidative stress [3,4]. Because *Plasmodium* spp. are highly susceptible to such oxidative stress, their antioxidant defenses are considered to play essential roles in survival throughout the lifecycle and thus may be potential targets for malaria chemotherapies [5].

Superoxide dismutase, catalase, glutathione (GSH) peroxidase, and peroxiredoxin (Prx) are the four major cellular antioxidant enzymes in aerobes [6,7]. An interesting feature of antioxidant system of malaria parasites is that the parasites do not possess genes that encode catalase or genuine GSH peroxidase, but they are equipped with a 1-Cys Prx, two typical 2-Cys Prxs, a 1-Cys antioxidant protein (AOP) and a GSH peroxidase-like thioredoxin peroxidase [1,2,8,9]. We recently reported that disruption of the gene encoding cytosolic 2-Cys Prx (PbTPx-1) of the rodent malaria parasite *Plasmodium berghei* did not affect its asexual proliferation in mouse erythrocytes but that the disruption caused a defect in gametocyte development [10]. In the present report, we examined the insect-stage phenotype of parasites carrying a targeted knockout (KO) of the Prx gene (*pbtpx-1*). We also examined the phenotype of Prx-KO sporozoites during the early stage of liver infection.

To investigate the effect of *pbtpx-1* disruption on the mosquito stage of the parasite, *Anopheles stephensi* were fed on BALB/c mice that showed high levels of gametocytemia (Table 1). Prx-WT oocysts developed similarly to those of WT in the midgut, and the final numbers of sporozoites that formed in midgut oocysts and that accumulated in the salivary gland by 21 days post-feeding (pf) were equivalent level to those of WT ($P = 0.4–0.8$ and $P = 0.4–0.5$, respectively). For Prx-KO1-3 populations, the number of midgut oocysts at 14–15 days pf was comparable to those of WT. The normal oocyst formation in Prx-KO suggests that the *pbtpx-1* disruption does not affect the gamete fertilization, ookinete formation or transformation of the ookinetes to oocysts, which requires ookinete invasion of

*Abbreviations:* DHFR-TS, dihydrofolate reductase-thymidylate synthase; GSH, glutathione; Prx, peroxiredoxin; ROS, reactive oxygen species.

* Corresponding author at: National Research Center for Protozoan Diseases, Obihiro University of Agriculture and Veterinary Medicine, 2-13 Inada-cho, Obihiro, Hokkaido 080-8555, Japan. Tel.: +81 155 495846; fax: +81 155 495643.
*E-mail address:* skawazu@obihiro.ac.jp (S.-I. Kawazu).
[1] Present address: Department of Immunology and Parasitology, School of Medicine, University of Occupational and Environmental Health, Japan, Yahatanishi-ku, Kitakyusyu 807-8555, Japan.

0166-6851/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.molbiopara.2008.03.002

**Table 1**
Development of Prx-KO *Plasmodium berghei* in *Anopheles stephensi*

| | Parasites | 14–15 days post-feeding (pf) | | | 21 days pf no. of sporozoites/mosquito[a] | | |
|---|---|---|---|---|---|---|---|
| | | Total oocyst number | No. of mosquitoes[b] | Matured oocyst (%) | Midgut | Salivary gland | No. of mosquitoes[b] |
| Experiment 1 | WT | 1760 | | 32.1 | 22000 ± 2740 | 8600 ± 990 | |
| | Prx WT | 1920 | 28 | 27.2 | 25200 ± 2850 | 7600 ± 610 | 20 |
| | KO1 | 1540 | | 4.2 | 2870 ± 120[c] | 270 ± 60[c] | |
| Experiment 2 | WT | 1330 | | 18.6 | 50400 ± 510 | 9950 ± 1700 | |
| | Prx WT | 1040 | 20 | 17.9 | 49100 ± 7700 | 9600 ± 600 | 30 |
| | KO1 | 870 | | 3.1 | 5470 ± 800[c] | 330 ± 70[c] | |
| Experiment 3 | WT | 1020 | | 14.5 | 30000 ± 3280 | 7070 ± 510 | |
| | Prx WT | 1460 | 20 | 11.9 | 28910 ± 3240 | 7920 ± 1180 | 25 |
| | KO1 | 1310 | | 3.7 | 4720 ± 1130[c] | 800 ± 160[c] | |
| Experiment 4 | WT | 700 | | 19.7 | 14100 ± 1370 | 6430 ± 740 | |
| | KO2 | 940 | 20 | 1.2 | 3070 ± 930[c] | 550 ± 150[c] | 30 |
| | KO3 | 1220 | | 1.8 | 1700 ± 360[c] | 270 ± 90[c] | |
| Experiment 5 | WT | 1590 | | 17.9 | 44400 ± 800 | 13800 ± 530 | |
| | KO4 | 1020 | 25 | 4.1 | 4280 ± 970[c] | 920 ± 460[c] | 25 |
| | KO5 | 760 | | 3.7 | 4940 ± 420[c] | 660 ± 160[c] | |

The *P. berghei* ANKA strain was obtained from the Armed Forces Research Institute of Medical Sciences, Thailand. The Prx knockout (Prx-KO) parasite, which carries a targeted disruption of *pbtpx-1* (PlasmoDB, PB000037.01.0), was established by double-crossover homologous recombination with a selectable marker, the dihydrofolate reductase-thymidylate synthase (DHFR-TS) gene with a pyrimethamine-resistance mutation [10]. Five Prx-KO populations (Prx-KO1-3 and Prx-KO4-5 were obtained by two independent electroporation experiments), one wild-type parasite population with pyrimethamine resistance (*dhfr-ts/mt* at *dhfr-ts* locus) and intact *pbtpx-1* (Prx-WT) and the parent strain (WT) were used to infect the animals. After parasite infection, the numbers of gametocytes in the peripheral blood were monitored, and *A. stephensi* mosquitoes were fed on five-week-old BALB/c mice (Clea Japan, Japan) when the number of gametocytes reached 20–30 per $1 \times 10^5$ erythrocytes. At 14–15 days pf, the mosquitoes were dissected, and oocyst numbers in the midgut were counted [17]. An oocyst filled with needle-shaped sporozoites was counted as a mature oocyst. One week later (21 days pf), the remaining mosquitoes were dissected, and sporozoite numbers in the midgut oocysts, hemolymph and salivary glands were examined [18]. The animal experiments in this study were carried out in compliance with the Guide for Animal Experimentation at either Ehime University School of Medicine or at the International Medical Center of Japan. Differences were evaluated with Student's *t*-test. $P < 0.05$ was considered statistically significant.

[a] Mean ± S.D.

[b] No. of mosquitoes dissected.

[c] The difference between WT and KO population was significant ($P < 0.01$).

epithelial cells and attachment to the basal lamina of the mosquito's midgut [11]. The lower number of oocysts in Prx-KO4 and 5 populations when compared to that of WT may be attributed to the mosquito's blood feeding, which varies between each experiment. *P. berghei* develops asynchronously in the infected mosquitoes, and the midgut contains both young and mature oocysts. In our experiments, in WT-infected mosquitoes, 15–32% of the oocysts were mature oocysts, which contained needle-shaped sporozoites, at 14–15 days pf. However, the percentages of mature oocysts in Prx-KO-infected mosquitoes at the same time points were lower (1–4%) than those of WT-infected mosquitoes. This phenotype was observed in all Prx-KO populations. Electron microscopic observation of the midguts at 14–15 days pf revealed that there was damaged oocyst with irregular sporogonic development (Fig. 1A), supports the idea that sporogony is abnormal in Prx-KO. This finding strongly suggests that the immature oocysts in Prx-KO populations failed to develop further. Consequently, the final numbers of sporozoites that formed in midgut oocysts and that accumulated in the salivary gland by 21 days pf were significantly decreased to 10–20% and 3–10%, respectively, in Prx-KO-infected mosquitoes than in WT-infected mosquitoes (Table 1). The ratio of salivary gland sporozoites to midgut sporozoites was significantly lower ($P < 0.05$) in Prx-KO-infected mosquitoes (0.06–0.21) than in WT-infected mosquitoes (0.20–0.46). Thus, the reduction in the number of salivary gland sporozoites in Prx-KO populations may not be due solely to the reduced number of sporozoites in the midgut oocysts. Because the ratio of hemolymph sporozoites to midgut sporozoites was comparable between Prx-KO-infected mosquitoes and WT-infected mosquitoes (data not shown), *pbtpx-1* disruption may not affect sporozoite entry to the salivary gland, but it may affect parasite survival in salivary gland after the entry. Sporozoite invasion of the salivary gland requires multiple steps [12], including entry into the secretory cavity, where the parasite has direct contact with saliva. NADPH oxidase activity which produces ROS was detected in anopheline salivary homogenate [13]. Disruption of *pbtpx-1* may affect the parasite survival in the mosquito saliva, but the phenotype requires further investigation to clarify this.

To investigate the role of TPx-1 in oocyst maturation, the Prx-KO and WT populations were subjected to *in vitro* oocyst culture (Fig. 1B). Ookinetes derived from WT population could efficiently transform into oocysts, and they conspicuously enlarged in size after the 5th day of culture. The number of oocysts, those that grew more than 6 μm in diameter with normal morphology, was 11146 ± 678 (28 ± 2% of the total number of ookinetes initially added to the culture) on the 7th day of culture. Ookinetes from the Prx-KO population transformed into oocysts like those of the WT population (data not shown). However, their growth after transformation was less pronounced than that observed by WT oocysts. The maximum number of oocysts observed on the 7th day of culture was 6740 ± 898 (17 ± 2% of the total number of ookinetes initially added to the culture), which was significantly lower than that of the WT population. These results indicated that the reduced oocyst development observed in Prx-KO-infected mosquitoes could be reproduced *in vitro*.

To investigate whether the gene disruption can promote DNA damage during the insect stage, oocysts harvested on the 8th day of culture were subjected to comet assay. We found that the comet-tail length of Prx-KO oocysts was significantly longer than that of WT oocysts ($P < 0.05$) (Fig. 1C). Although comet-tail length is considered to reflect breaks in the cellular DNA [14], it is not clear whether the DNA alterations are causal or just a consequence of degrading parasites. However, our findings from immunoelectron microscopic observation also suggested that there was an increase in formation of 8-hydroxy-2'-deoxyguanosine (8-OHdG), a marker of oxidative DNA damage, in Prx-KO oocyst nuclei during early developmental stage (Fig. 1D). Taken together, our data suggest that disruption of *pbtpx-1* induces DNA damage in oocyst nuclei. How malaria parasites deal with the accumulation of oxidative DNA damage during

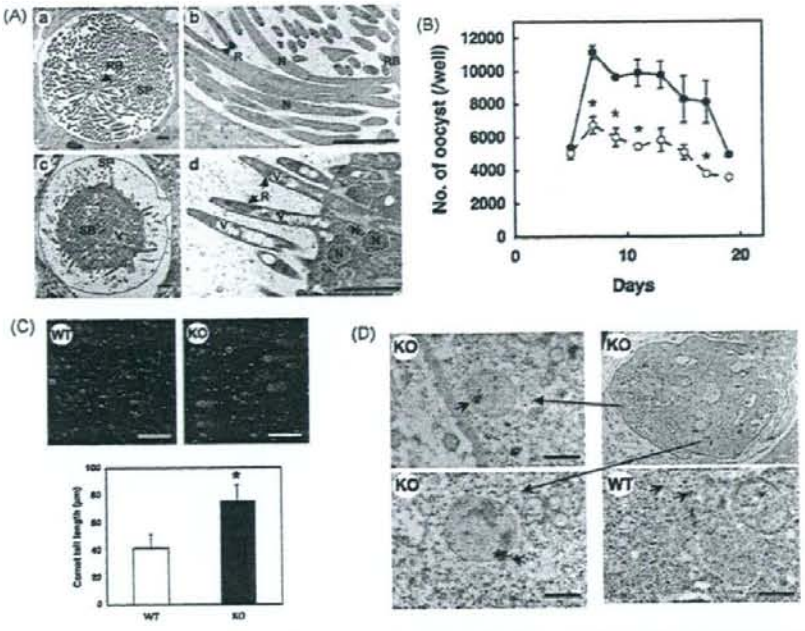


**Fig. 1.** (A) Electron microscopy image of oocyst and sporozoite in midguts of mosquitos infected with WT (a and b) and Prx-KO (c and d). Electron microscopy was performed as described previously [18] with a transmission electron microscope (Hitachi H-7000, Japan). The midguts were dissected at 15 days pf. Matured oocyst in WT is packed with fully formed sporozoites, and the sporoblast decreases in size to become one residual body (a). The sporozoites show normal structure with a firm nucleus and rhoptries (b). Oocyst in KO contains one large sporoblast, which does not form multiple islands (c). The sporoblast and the sporozoites have deformed nuclei, swollen rhoptries and abnormal cytoplasmic vacuoles (d). Eighty-two oocysts, 39 of which had begun sporozoite budding, from three midgut samples were observed for Prx-KO, and none of the oocysts contained normal sporozoites. Such abnormal sporogony was rarely seen in WT-infected mosquitoes. N, SB, SP, R, RB and V indicate nucleus, sporoblast, sporozoite, rhoptries, residual body and vacuole, respectively. Bars indicate 5 μm. (B) Effect of *pbtpx-1* disruption on oocyst development *in vitro*. Oocyst culture was basically carried out as described previously [19]. Ookinetes ($4 \times 10^4$) derived from WT (●) and Prx-KO (○) populations were initially added to the culture well. The number of oocysts (those that grew more than 6 μm in diameter with normal morphology) was counted on every other day from 5th to 19th day of culture. Data are mean ± S.D. of the oocyst number in triplicate cultures. *The difference between the WT and KO populations was significant ($P < 0.05$). (C) Effect of *pbtpx-1* disruption on DNA damage in oocysts. DNA strand breaks in oocysts harvested on the 8th day of culture were evaluated by comet assay. Comet assay was performed according to the manufacturer's instructions (CometAssay™, Trevigen, Inc., USA). Quantitative analysis was done by measuring the tail length of the degraded DNA from each oocyst cell spot in WT and Prx-KO (KO) populations with a confocal laser scanning microscope (LSM510, Carl Zeiss, Germany). Bars indicate 100 μm. Data are mean ± S.D. of the comet-tail length (μm) of 75 oocysts from WT (□) and Prx-KO populations (■). *$P < 0.05$. (D) *pbtpx-1* disruption promotes 8-OHdG formation in oocyst nuclei. 8-OHdG formation in nuclei of oocysts at an early developmental stage (young oocyst developed in the mosquito's midgut at 15 days pf) was detected by immunoelectron microscopy with anti-8-OHdG antibody (Japan Institute for the Control of Aging, Japan) and secondary antibody conjugated with gold particle. Immunoelectron microscopy was performed as described previously [18]. The presence of gold particles in the nuclei of Prx-KO oocysts (arrowheads in panels labeled KO) indicates formation of 8-OHdG, which was rarely observed in nuclei of WT oocysts. Arrowheads in panel labeled WT showed traceable deposition of these particles in the cytosol. Bars represent 1 μm.

the lifecycle is of interest. Because oocyst maturation was reduced in Prx-KO *in vitro*, host factors may not contribute to the accumulation of oxidative DNA damage during the mosquito stage. Recent data from microarray analyses of oocysts cultured in *P. berghei* revealed that expression of genes encoding thioredoxin and 2-Cys peroxiredoxin (PbTPx-1) are upregulated during the early stage of oocyst development [15]. This finding suggests that ROS are present during the early stage of oocyst development (young oocyst), when oxygen metabolism in the cell may be elevated. There is a consistent data indicating that genes encoding enzymes in the mitochondrial electron transport chain, which produces ROS, are upregulated in young oocysts [15].

To evaluate infectivity of Prx-KO populations in mice, salivary-gland sporozoites were injected intravenously into BALB/c mice, and the infection rate and the period required for the parasite to develop 0.5% parasitemia of erythrocytic stage (pre-patent period) were compared to those of WT. Parasitemia of the animals was monitored every 12 h. After inoculation of 100, 1000 and 10000 sporozoites from Prx-KO2, Prx-KO3 and WT populations, all animals developed erythrocytic-stage infection. Therefore, the infectivity of the Prx-KO population was assessed based on pre-patent period (**Supplementary Table 2**). The pre-patent period recorded for WT infections increased according to decreasing numbers of sporozoites in the inoculum, the mean pre-patent period for animals ($n = 5$) inoculated with 10000, 1000, and 100 sporozoites were 4.6, 5.3 and 6.2 days, respectively. The mean pre-patent period for animals ($n = 5$) inoculated with 1000 Prx-KO sporozoites were 6.1 (KO2) and 6.3 (KO3) days, which was equivalent to that recorded in animals inoculated with 100 WT sporozoites. These data indicate that the infectivity of Prx-KO populations in mice was reduced to 1/10 of WT.

To investigate the influence of Prx-KO on sporozoite invasion to the liver parenchyma, parasite burden in the liver shortly after the sporozoite inoculation, which represents the number of sporozoites that could cross the sinusoidal cell layer, was compared between WT, Prx-WT, and Prx-KO1 by TaqMan® fluorescent quantitative RT-PCR. For this purpose, mice were inoculated intravenously with 2000 sporozoites, and their livers were perfused with PBS 1 h after the inoculation and then removed. The parasite burden in the liver sample, which represents number of sporozoites in Kupffer cells, the space of Disse, and hepatocytes, was assessed as the ratio of the amount of parasite 18S rRNA to the amount