Structure of the Gene Encoding the Immunodominant Surface Antigen on the Sporozoite of the Human Malaria Parasite Plasmodium falciparum

John B. Dame, Jackie L. Williams, Thomas F. McCutchan
James L. Weber, Robert A. Wirtz, Wayne T. Hockmeyer
W. Lee Maloy, J. David Haynes, Imogene Schneider
Donald Roberts, Greg S. Sanders, E. Premkumar Reddy
Carter L. Diggs, Louis H. Miller

The need for vaccines to relieve the current global resurgence of malaria is clear. Vaccines are being developed against each stage in the life-cycle of the malaria parasite because immunity to malaria is specific for each stage. The stages include sporozoites, which develop in mosquitoes and are injected by them into humans; asexual erythrocytic parasites, which cause the disease; and sexual stages, which develop in humans and transmit the infection to mosquitoes (1). A vaccine against sporozoites would, if effective, prime the human immune system to kill sporozoites injected by the mosquito and thus prevent the subsequent stages responsible for the disease and transmission of the infection to others.

Humans and other animals have been protected from malaria by immunization with irradiated sporozoites, but this method is impractical because of the limited supply and instability of sporozoites. The circumsporozoite (CS) protein that covers the surface of the sporozoite was first identified in Plasmodium berghei, a parasite of rodents (2). Monoclonal antibodies to this protein completely protected mice from challenge by infected mosquitoes (3). Analogous CS proteins have been identified for species of Plasmodium infecting monkeys and humans (4), including P. falciparum, the major cause of malaria in humans. The gene for the CS protein of P. knowlesi, which infects the Old World monkey Macaca irus, was cloned first because large numbers of sporozoites were available in infected mosquitoes for preparation of a complementary DNA (cDNA) library (5). This gene encoded a protein with a repeating amino acid sequence (12 amino acids repeated 12 times) which contained the epitope that bound the protective monoclonal antibodies. This repeating epitope was the major immunogen on the protein, because monoclonal antibodies blocked access of polyclonal antiserum to Triton X-100-solubilized sporozoite protein in the immunoradiometric assay (6). The presence of a repeating amino acid sequence is not a unique feature of the CS proteins. Such sequences are also found in other malarial proteins such as the S antigen (7).

We describe in this article the cloning of the gene for the CS protein of P. falciparum, its nucleotide sequence, and the amino acid sequence of the protein as deduced from the nucleotide sequence. Previously, the gene for the CS protein of P. knowlesi was cloned by using cDNA (5). The development of the technique of McCutchan et al. (8) made it possible to clone the intact gene for the CS protein of P. falciparum from genomic DNA of asexual erythrocytic parasites grown in continuous culture. We compare its structure with that of the previously described analogous protein of P. knowlesi.

Clones from the genomic DNA expression library. The P. falciparum genomic DNA library in the expression vector λgt11 (9) was produced with the use of mung bean nuclease as described (8). The library was made from the DNA of the 7G8 clone of the IMTM22 isolate of P. falciparum from Brazil (10). The library was plated at a density of 25,000 plaques per 150-mm plate on 27 plates and immunologically screened (11). A pool of five monoclonal antibodies (12) against the P. falciparum 7G8 CS protein (Table 1) was used at a dilution of 1:10,000 for screening. Thirty-five positive clones were obtained in the initial screening after 48 hours of autoradiography. Seventeen were rescanned at a...
density of 100 to 800 plaques per 85-mm plate. Eleven of the clones gave positive plaques on the second screening. These were cloned without screening from 85-mm plates containing fewer than 50 plaques; 10 of the 11 clones were immunoreactive when screened.

Inserts in the ten clones were of the following sizes: \( \lambda M P F I \), \( \lambda M P F 3 \), and \( \lambda M P F 1 1 \) were 2.3 kilobases (kb); \( \lambda M P F 5 \), \( \lambda M P F 8 \), and \( \lambda M P F 1 3 \) were 1.3 kb; \( \lambda M P F 1 5 \) was 1.35 kb; \( \lambda M P F 6 \) was 1.0 kb; and \( \lambda M P F 9 \) was 0.4 kb. Clone \( \lambda M P F 1 8 \) contained two inserts and was not studied further. The inserts of clones \( \lambda M P F 1 \), -3, -5, -8, -11, -13, and -15 cross-hybridized. \( \lambda M P F 6 \) and \( \lambda M P F 9 \) did not cross-hybridize, indicating that the two smaller inserts, although selected by the mixture of five monoclonal antibodies, came from a part of the genome outside the 2.3-kb fragment.

Clone \( \lambda M P F 5 \) was nick-translated and used to probe a Southern blot (13) of Hind III digests of human and \( P. falciparum \) genomic DNA. A single band of hybridization was present at 14 kb in the \( P. falciparum \) lane (data not shown). The probe did not hybridize to human DNA.

Expression of the CS protein in Escherichia coli. The clones in \( \lambda g t 1 1 \) were introduced as lysogens into \( E. coli \) strain \( R Y 1 0 8 9 \) (ATCC 37196). The phages were induced at 44°C and then isopropylthio-
galactoside (IPTG) was added to the medium to enhance expression of \( \beta \)-galactosidase and possible fusion proteins (14). Lysates of the induced bacteria were analyzed for reactivity with each of the five monoclonal antibodies by Western blotting. The results indicated that the gene for the CS protein encoded a protein of 412 amino acids as deduced from the nucleotide sequence. The protein contains 41 tandem repeats of a tetrapeptide, 37 of which are Asn-Ala-Asn-Pro and four of which are Asn-Val-Asp-Pro. Monoclonal antibodies against the CS protein of Plasmodium falciparum were inhibited from binding to the protein by synthetic peptides of the repeat sequence. The CS protein of Plasmodium falciparum and the CS protein of a simian malaria parasite, Plasmodium knowlesi, have two regions of homology, one of which is present on either side of the repeat. One region contains 12 of 13 identical amino acids. Within the nucleotide sequence of this region, 25 of 27 nucleotides are conserved. The conservation of these regions in parasites widely separated in evolution suggests that they may have a function such as binding to liver cells and may represent an invariant target for immunity.

The lysates used for ELISA were subjected to electrophoresis on a sodium dodecyl sulfate-polyacrylamide gel (SDS-PAGE) and electroblotted onto nitrocellulose (17). The proteins on the nitrocellulose paper were identified by monoclonal antibodies to \( P. falciparum \) sporozoites. These antibodies bound to two doublets at molecular ratios of \( M_r \) of 60,000 and 57,000 and 53,000 and 51,000 on the protein blots of \( \lambda M P F 1 \), -3, -5, -8, -11, and -13 (Fig. 1), although the intensity for \( \lambda M P F 1 3 \) was greatly reduced. No binding occurred to the \( \lambda g t 1 1 \) vector without an insert. Monoclonal antibodies bound to proteins from sporozoites at \( M_r \) of 60,000, 53,000, and 51,000 (Fig. 1). Thus, all the sporozoite genes for CS protein in \( \lambda g t 1 1 \) produced a protein of similar mobility to the highest molecular weight form of the CS protein present in the sporozoite preparation (\( \sim M_r 60,000 \)).

On induction with IPTG, a marked increase in expression of \( \beta \)-galactosidase (\( M_r 1 1 6,000 \)) was noted for \( \lambda g t 1 1 \), and a fusion protein at \( M_r 1 3 1,000 \) with \( \beta \)-galactosidase was noted for \( \lambda M P F 9 \) (data not shown). The clones with the CS protein gene gave only weak \( \beta \)-galacto-
sidase bands on induction; no fusion proteins were seen (data not shown). In addition, antibody to \( \beta \)-galactosidase did not bind to the \( M_r 60,000 \) CS protein, suggesting that this protein did not contain fragments of \( \beta \)-galactosidase (data not shown).

The \( \lambda g t 1 1 \) vector is designed to express inserts as \( \beta \)-galactosidase fusion proteins on induction with IPTG (9). Thus the apparent lack of fusion proteins among clones expressing the CS protein was unexpected (Fig. 1). This is explained for \( \lambda M P F 1 \), -5, -8, and -15 because the inserts are oriented so that their direction of transcription is opposite to that of \( \beta \)-galactosidase. Restriction mapping of the phage DNA with \( S t u I \), \( K p n I \), and \( S t u I + K p n I \) indicates that the asymmetric \( S t u I \) site in the insert (Fig. 2) is located in each case about 1.1 kb from the \( K p n I \) site in \( \lambda g t 1 1 \) at 18.58 kb on the map (9). It is not known whether the \( P. falciparum \) DNA 5' to the coding sequence in clones \( \lambda M P F 1 \) and \( \lambda M P F 1 5 \) contains sequences that can be used as promoters by the \( E. coli \) RNA polymerase, but there are no obvious binding sites for bacterial ribosomes (Fig. 3). Clones \( \lambda M P F 5 \) and \( \lambda M P F 8 \) begin only 11 base pairs (bp) before the gene. Thus, expression of the CS protein in these clones is probably from a late \( \lambda \) promoter. A similar phenomenon was observed in the \( \lambda g t 1 1 \) system with a yeast DNA insert (18).

Restriction mapping indicates that the insert in \( \lambda M P F 1 3 \) is in the correct orientation with the \( \beta \)-galactosidase gene. However, it is one base out of frame to produce a fusion protein with \( \beta \)-galactosidase (Fig. 3). The low levels of CS protein produced by this clone as detected on Western blotting (Fig. 1) and in the ELISA (Table 1) may be understood in light of this construction. The bias among the clones for either reversed orientation or out-of-frame inserts sug-

---

Abstract. The gene for the circumsporozoite (CS) protein of Plasmodium falciparum has been cloned and its nucleotide sequence determined. The gene encodes a protein of 412 amino acids as deduced from the nucleotide sequence. The protein contains 41 tandem repeats of a tetrapeptide, 37 of which are Asn-Ala-Asn-Pro and four of which are Asn-Val-Asp-Pro. Monoclonal antibodies against the CS protein of Plasmodium falciparum were inhibited from binding to the protein by synthetic peptides of the repeat sequence. The CS protein of Plasmodium falciparum and the CS protein of a simian malaria parasite, Plasmodium knowlesi, have two regions of homology, one of which is present on either side of the repeat. One region contains 12 of 13 identical amino acids. Within the nucleotide sequence of this region, 25 of 27 nucleotides are conserved. The conservation of these regions in parasites widely separated in evolution suggests that they may have a function such as binding to liver cells and may represent an invariant target for immunity.
gests that there was selection against the expected fusion proteins (for example, λmpF5 and λmpPB in the correct orientation), perhaps because a higher level of synthesis of the CS protein was toxic to E. coli.

Structure of the P. falciparum gene for the CS protein. In Fig. 3, the deduced amino acid sequence for the CS protein of \textit{P. falciparum} is shown below the nucleotide sequence of \textit{λ}mpF1. The CS protein gene sequence contains a large open reading frame which commences with an ATG initiation codon at position 78 and terminates with a TAG codon at position 1316. Multiple terminator codons were observed in the other five reading frames. The open reading frame shown in Fig. 3 could code for a polypeptide of 412 amino acids with an approximate M, of 44,000. As was observed for the CS protein of \textit{P. knowlesi} (19), the molecular weight of the CS protein of \textit{P. falciparum} by SDS-PAGE (~60,000) differed from the deduced molecular weight (44,000).

Examination of the amino acid sequence for the CS protein of \textit{P. falciparum} reveals several interesting structural features that are of importance not only for our understanding of the role of this protein for sporozoite function but also for developing vaccines that could effectively act against sporozoites.

An important structural feature of this protein is the presence of 41 tandem repeats of tetrapeptides. The primary repeating unit is Asn-Ala-Asn-Pro, which occurs 37 times; an alternative form is Asn-Val-Asp-Pro, which occurs at units 2, 4, 6, and 22. The change from Ala-Asn to Val-Asp results from point mutations where cysteine is replaced by thymine in the second position of the alanine codon and where adenine is replaced by guanine in the first position of the asparagine codon.

The nucleic acid sequence coding for the repeats is not as well conserved as the amino acid sequence. The repeated region, which has 41 units, is composed of 11 different combinations of nucleotide sequence. Eighteen of the units are of one type (AATGCAAACCCCA). Seven of the repeats differ in only one position from this sequence, 12 differ in two positions, two differ in three positions, one differs in four positions, and one differs in five positions. The alteration in the sequence may stabilize the repeat within the genomic DNA and prevent it from being eliminated or reshuffled by recombination (20).

At the amino terminal end of the protein, a stretch of 16 amino acids constitutes a probable signal sequence (Figs. 3 and 4). Between this signal sequence and the repeated region occurs a highly charged region that is characterized by the presence of both basic and acidic amino acids. Thus, 27 of 53 amino acids from amino acids 66 to 118, are charged. Following the repeat region, two other segments of the protein contain a high proportion of charged amino acids. These regions occur between amino acids 324 and 339 and between amino acids 360 and 388; they contain 50 and 48 percent charged amino acids, respectively. At the carboxyl terminal end, the protein has a sequence of 21 hydrophobic amino acids, which represents an anchor sequence for an integral membrane protein.

Immunoreactivity of synthetic peptides with antibodies to the repeat sequence. To prove conclusively that the repeating nucleotide unit of the \textit{P. falciparum} sporozoite gene was correct, we synthesized peptides (21) of the repeat (Fig. 5). These peptides were then used in a modification (22) of the ELISA (15) to determine if they would inhibit binding of the monoclonal antibody 2F1.1 to \textit{λ}mpF1. The results (Fig. 5) demonstrate that the 7-, 11-, and 15-residue peptides significantly inhibit binding of 2F1.1 to \textit{λ}mpF1. Inhibition of binding was evident at $5 \times 10^{-7}$M with the 15-residue peptide. The 7-residue peptide also inhibited binding of monoclonal antibody 2F1.1 to the sporozoite antigen substituted for \textit{λ}mpF1 (data not shown). Furthermore, the synthetic peptide inhibited binding of the other four monoclonal antibodies to \textit{λ}mpF1. These data indicate that the sequence of the repeating unit is correct.

The increased inhibition of binding seen with the 11- and 15-residue peptides may reflect secondary conformational changes. The data do not suggest that they contain two epitopes since neither could be detected in a double-sided assay with 2F1.1 (data not shown).

Regions of homology between CS proteins of \textit{P. falciparum} and \textit{P. knowlesi}. The CS protein of \textit{P. falciparum} and the CS protein of a simian malaria parasite, \textit{P. knowlesi} (19), have a similar overall structure (Fig. 4), but have only two short regions of homology. Both proteins appear to contain the same major features in that they have a repeated region in the middle of the protein, multiple regions with a high density of charged amino acids, a signal sequence at the amino terminal end, and a hydrophobic anchor sequence at the carboxyl terminal end (Fig. 4). However, computer analysis for the amino acid sequence homology [K tuple size of 1, window size

<table>
<thead>
<tr>
<th>Antigen</th>
<th>2E6.4</th>
<th>2F1.1</th>
<th>4D9.1</th>
<th>4D11.6</th>
<th>5G5.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>λmpF1 (2.3 kb)</td>
<td>1.2</td>
<td>1.3</td>
<td>1.7</td>
<td>1.5</td>
<td>1.9</td>
</tr>
<tr>
<td>λmpF3 (2.3 kb)</td>
<td>1.3</td>
<td>1.1</td>
<td>1.9</td>
<td>1.4</td>
<td>1.8</td>
</tr>
<tr>
<td>λmpF5 (1.3 kb)</td>
<td>1.1</td>
<td>0.9</td>
<td>1.5</td>
<td>1.2</td>
<td>1.4</td>
</tr>
<tr>
<td>λmpP8 (1.3 kb)</td>
<td>1.2</td>
<td>0.9</td>
<td>1.6</td>
<td>1.2</td>
<td>1.4</td>
</tr>
<tr>
<td>λmpF1 (2.3 kb)</td>
<td>1.1</td>
<td>0.9</td>
<td>1.6</td>
<td>1.3</td>
<td>1.4</td>
</tr>
<tr>
<td>λmpF3 (1.3 kb)</td>
<td>0.6</td>
<td>0.5</td>
<td>0.5</td>
<td>0.7</td>
<td>0.6</td>
</tr>
<tr>
<td>λmpF5 (1.35 kb)</td>
<td>1.1</td>
<td>1.1</td>
<td>2.0</td>
<td>1.7</td>
<td>2.0</td>
</tr>
<tr>
<td>Agtll, IPTG induced</td>
<td>0.5</td>
<td>0.5</td>
<td>0.4</td>
<td>0.6</td>
<td>0.4</td>
</tr>
<tr>
<td>Agtll, not induced</td>
<td>0.5</td>
<td>0.4</td>
<td>0.4</td>
<td>0.6</td>
<td>0.4</td>
</tr>
<tr>
<td>RY1089</td>
<td>0.4</td>
<td>0.5</td>
<td>0.3</td>
<td>0.6</td>
<td>0.3</td>
</tr>
</tbody>
</table>
of 20, gap penalty of 1 (23) revealed limited sequence homology over most of the protein. The average homology between the two proteins in the segment before the repeat (as divided in Fig. 4) is 37 percent; 37 of a possible 102 amino acids match. The repeats are 16 percent identical since one proline and one alanine align every 12 amino acids. The average homology between the segments of the two proteins after the repeats is 42 percent; 50 of a possible 119 amino acids match. As the secondary and tertiary structure of these proteins are unknown, they may have structural and functional

\[ \text{Fig. 3. The nucleotide sequence of the CS protein gene in \( \lambda \text{mpF1}. \) The Eco RI insert in \( \lambda \text{mpF1} \) was subcloned in pUC8 (27), and sequenced by the method of Maxam and Gilbert (28). The sequence of both DNA strands was determined for 100 percent of the CS protein coding region and greater than 70 percent of the flanking regions. The inserts of clones \( \lambda \text{mpF5}, \lambda \text{mpF10}, \lambda \text{mpF15}, \lambda \text{mpF15/B} \) were also subcloned into pUC8 and the ends sequenced. The first base of each clone 5' to the CS protein coding region is located to the right of the arrows. The Eco RI linkers (GGAAATTTC) ligated at both ends of the inserts are not shown as part of the sequence. The deduced amino acid sequence of the CS protein is given beneath the nucleotide sequence. Two regions of the protein nearly identical to the \textit{F. novicida} CS protein (19) are marked region I and region II. The repeat units are underlined, and the variant amino acids in the repeat are boxed. The amber terminator codon in the sequence is indicated with stars. The symbols for the nucleotides are: A, deoxyadenosine; G, deoxyguanosine; C, deoxycytidine; T, thymidine.} \]
similarities despite the difference in primary sequence. For example, repeats in CS proteins are immunodominant after vaccination with sporozoites (6).

The two regions of greatest sequence homology are seen on either side of the repeated region. When the two peptides are aligned a region of homology is apparent where three prolines are aligned and there is a perfect match of five contiguous amino acids (Lys-Leu-Lys-Gln-Pro) (region I, Figs. 3, 4, and 6).

The second region of homology (region II, Figs. 3, 4, and 6) contains 13 amino acids, 12 of which are conserved. The only amino acid that is not identical is the fourth residue, which is Ser in P. falciparum, but is Thr in P. knowlesi. This region contains two cysteine residues which were implicated earlier by Ozaki et al. (19) in the formation of an intramolecular loop.

The nucleic acid sequence encoding the CS protein of P. falciparum is nearly identical to that of the P. knowlesi gene in region II (Figs. 3 and 6). A 27-base sequence is present which differs from the comparable sequence in P. knowlesi at only two positions. The next longest region of homology between the two sequences was only eight nucleotides in length. The larger conserved sequence may be useful as a probe to clone the genes encoding the CS proteins of other Plasmodium species.

The two regions of homology of amino acid sequence between P. falciparum and P. knowlesi indicate conservation of sequence for organisms widely separated in evolution. It was originally assumed that the primate malarial parasites had evolved in parallel with the evolution of primates. Recently, however, McCutchan et al. (24) demonstrated that the DNA composition and genome arrangement of P. falciparum is different from that of the primate malaria parasites P. knowlesi, P. cynomolgi, P. fragile, and P. vivax. These regions, which are nearly identical between P. falciparum and P. knowlesi, may be conserved for important functions of the protein such as receptor for cell invasion. Both P. falciparum and P. knowlesi sporozoites can infect the human liver.

Discussion

The method used to clone the CS protein gene depends on the finding that mung bean nuclease, under controlled conditions of formamid concentration and temperature, preferentially cuts the 5' and 3' ends of genes (8). Such DNA
fragments containing genes can be cloned into the expression vector λgt11 (9) and the clones can be screened for expression with antibody. The CS protein gene was expressed in λgt11 despite the fact that the insert was not in frame with the β-galactosidase gene (λMPF13) or that the insert was in the opposite orientation to the β-galactosidase gene (λMPF1, -3, -11, and -15). The transcription was from a k promoter rather than from the lac promoter. Expression in λgt11 of genes not in frame (9) and of genes inserted in the opposite direction (18) have been described.

The method has the added feature that genes for proteins with cross-reactive epitopes may also be identified. Clone AmP9 in the present study produced a protein that cross-reacted with one of the monoclonal antibodies but did not hybridize with the CS protein gene. The question of whether this gene codes for the asexual protein cross-reactive with sporozoites (16) can now be addressed.

Despite the potential problem of shared epitopes among proteins, the evidence that we have cloned the sporozoite gene is as follows. First, the similarity between this protein of P. falciparum and the CS protein of P. knowlesi is striking (Fig. 4). They are of similar size with calculated molecular weights of 44,426 and 36,792 for P. falciparum and P. knowlesi, respectively. They have analogous regions that include a signal sequence, charged regions, a region of repeating peptides in the middle of the protein, and an anchor sequence. Second, there are two regions of amino acid homology between the two proteins (Fig. 6). Third, five monoclonal antibodies known to react with the surface of P. falciparum sporozoites (12) recognized the protein synthesized in E. coli (Table 1). The cross-reactive protein from clone λmP9 only reacted with one of these monoclonal antibodies. Fourth, the protein synthesized in E. coli was of similar size in SDS-PAGE to the protein from P. falciparum sporozoites (Fig. 1). Fifth, synthetic peptides of the repeat blocked binding of a monoclonal antibody to the sporozoite protein in an ELISA (Fig. 5).

The sporozoite gene encodes a protein of 412 amino acids which consists of a signal sequence, a charged region, a central region of 41 four-amino acid units (repeats), two other charged regions, a probable cysteine loop, and an anchor sequence (Fig. 4). Thirty-seven of the repeats in the central region are identical (Asn-Ala-Asn-Pro); four have an alternative sequence (Asn-Val-Asp-Pro).

An analogous set of CS proteins are found on sporozoites of all Plasmodium species studied to date (4). Monoclonal antibodies to CS proteins confer protection in vivo (3) or neutralize sporozoite infectivity in vitro (25, 26). Although monoclonal antibodies may cross-react among Plasmodium species, antibody-mediated immunity appears to be species-specific (25), and in the case of P. knowlesi protective monoclonal antibodies are directed against the repeating epitope (5). These data (5) and the finding that monoclonal antibodies to sporozoites react in an assay that requires two or more epitopes led Zavala et al. (6) to propose that monoclonal antibodies to CS proteins react with an immunodominant region that has a repeating epitope. We have confirmed this hypothesis in

![Graph](image)

**Fig. 5.** Inhibition of binding of the monoclonal antibody 2F1.1 to CS protein by synthetic peptides of the predominant repeating amino acid sequence. Synthetic peptides (21) containing increasing lengths of the predominant repeat sequence were prepared and used to inhibit binding of 2F1.1 to a lysate of AmP91 growing in RY1089 (22). The data are given as the mean ± standard error of three replicates. The synthetic sequences tested were as follows: A, Asn-Pro-Asn-Ala; C, Pro-Asn-Ala-Asn-Pro-Asn-Ala; O, Pro-Asn-Ala-Asn-Pro-Asn-Ala; X, Pro-Asn-Ala-Asn-Pro-Asn-Ala; and 1, an unrelated decapeptide.

**Fig. 6.** (A and B) Regions of sequence identity between the CS proteins of P. falciparum and P. knowlesi. Amino acid sequences of the P. falciparum and P. knowlesi CS proteins were aligned by the method of Wilbur and Lipman (23) with a window size of 20, and gap penalty of 2 (23). The two regions of most extended homology are shown with amino acid numbers indicated. Region I identifies two amino acids from the repeat portion of the protein in P. falciparum. In P. knowlesi the last three amino acids of this region are part of the repeating portion of the protein. (C) The sole region of extended nucleotide sequence homology between the P. falciparum and P. knowlesi CS protein genes. The nucleotide sequences of the coding region of the P. falciparum and P. knowlesi CS protein genes were aligned by the method of Wilbur and Lipman (23) with a window size of 20, and gap penalty of 2 (23). The location of this sequence is indicated by the nucleotide numbers from Fig. 3 for P. falciparum and from (19) for P. knowlesi.
that the five monoclonal antibodies to the CS protein of P. falciparum are directed against repeating units (Asn-Ala-Asn-Pro) in the protein. It thus seems reasonable that the repeating epitope could form the basis for a vaccine. The immunogen for experimental vaccines against P. falciparum sporozoites could be produced in a bacterial fermentation system or through polymerization of a synthetic peptide.

In addition to the repeating epitope, other targets for sporozoite vaccines may include the conserved regions (Fig. 6) which were defined because of their homology to the P. knowlesi CS protein. The striking homology of region II between the evolutionarily divergent malaria parasites P. falciparum and P. knowlesi (24) suggests conservation for a sporozoite function such as reception for liver invasion. If this region is conserved in other human malarial and is exposed to the immune system, immunization with this region from P. falciparum may give protection against other species of human malaria. Furthermore, if this homologous region is involved in reception for liver invasion, then the malarial parasite may be unable to vary the sequence in this region.

References and Notes

11. The library was screened essentially as described [R. A. Young and R. W. Davis, Science 222, 778 (1983)]. Pooled ascitic fluids of five hybridomas were diluted 1:10,000 in 0.15M NaCl, 0.05M tris, pH 7.5 (tris-buffered saline (TBS)) containing 0.05 percent Tween-20 and 3 percent bovine serum albumin (BSA) and absorbed multiple times with a concentrated lysate of Agt11-infected RY1090 cells (air-dried onto nitrocellulose filters) to remove antibodies to E. coli and A. Antigen produced by the plaque formation was transferred in situ on to IPTG-saturated nitrocellulose filters. The filters were then washed in 500 ml of TBS containing 0.3 percent Tween-20, 3 percent BSA, 5 mM MgCl2, and 3 units of deoxyribonuclease I at room temperature for 30 minutes and incubated with the absorbed pool of monoclonal antibodies overnight at 4°C. All further manipulations were done at room temperature. After this and each of the next two steps, the filters were washed successively in TBS containing 0.05 percent Tween-20, TBS containing 1 percent Triton X-100, and TBS containing 0.05 percent Tween for 30 minutes in each solution. The signal of the mouse monoclonal antibodies was amplified by incubating the filters for 1 hour in rabbit antiserum to mouse immunoglobulin G (IgG) (Cappel) which had been diluted 1:500 in TBS containing 0.05 percent Tween-20 and 3 percent BSA and preabsorbed as described above for the ascitic fluid. Antibodies bound to the filters were detected by incubating up to five filters in 30 ml of TBS containing 0.05 percent Tween-20 and 1 lgM of [3H]-labeled protein A (Amersham) and then washing them and subjecting them to autoradiography.
12. J. L. Williams et al., unpublished data.
14. Lysogens were grown in media containing 50 pg of ampicillin per milliliter at 32°C until the absorbance at 550 nm was 0.4 to 0.8. The cultures were grown at 44°C for 20 minutes. IPTG was then added to a final concentration of 2 mM, and the culture was shaken at an additional 18°C. Lysis took 20 min for conditional lethal. Three hours after the IPTG was added, the cultures were plated on or 100 ml portions were mixed with 30 ml of monoclonal antibody 2E11 conjugated to horseradish peroxidase in 1.5-ml microcentrifuge tubes and held for 1 hour at room temperature. Wells of the microtiter plate were emptied and 30 ml of the monoclonal-antibody mixture was placed in each well and held for 1 hour at room temperature. The wells were again washed as above and 150 ml of the peroxidase substrate was added as described [P. K. Nakane and A. Kawai, J. Histochem. Cytochem. 22, 1084 (1974)].
21. We thank I. Quak and for the cultured axenial stage parasites used to prepare the DNA for the expression library and W. Davis for typing the manuscript. We thank L. A. McNicol and I. Green for comments on the manuscript, and F. A. Neva and P. K. Russell for their support in this work.
25 June 1984; accepted 29 June 1984.