Molecular and Biochemical Parasitology, 55 (1992) 105–114

MOLBIO 01809

Polymorphism in the circumsporozoite protein of the human malaria parasite *Plasmodium vivax*

Shoukat H. Qari[a], Ira F. Goldman[a], Marinete M. Povoa[b], Sylvia di Santi[c], Michael P. Alpers[d] and Altaf A. Lal[a]

[a]*Malaria Branch, Division of Parasitic Diseases, National Center for Infectious Diseases, Centers for Disease Control, Public Health Services, U.S. Department of Health and Human Services, Atlanta GA, USA;* [b]*Malaria Program, Instituto Evandra Chagas, Belem, Brazil;* [c]*Malaria Department, Sucen, Sã;o Paulo, Brazil and* [d]*Papua New Guinea Institute of Medical Research, Goroka, Papua New Guinea*

(Received 8 April 1992; accepted 12 June 1992)

The circumsporozoite (CS) protein that covers the surface of infectious sporozoites is a candidate antigen in malaria vaccine development. To determine the extent of B- and T-epitope polymorphism and to understand the mechanisms of antigenic variability, we have characterized the CS protein gene of *Plasmodium vivax* from field isolates representing geographically distant regions of Papua New Guinea (PNG) and Brazil. In the central repeat region of the CS protein, in addition to variation in the number of repeats, an array of mutations was observed which suggests that point mutations have led to the emergence of the variant CS repeat sequence ANGA(G/D)(N/D)QPG from GDRA(D/A)GQPA. Outside the repeat region of the protein, the nonsilent nucleotide substitutions of independent origin are localized in three domains of the protein that either harbor known T-cell determinants or are analogous to the *Plasmodium falciparum* immunodominant determinants, Th2R and Th3R. We have found that, with the exception of one CS clone sequence that was shared by one *P. vivax* isolate each from PNG and Brazil, the *P. vivax* CS protein types can be grouped into Papuan and Brazilian types. These results suggest that an in-depth study of parasite population dynamics is required before field trials for vaccine formulations based on polymorphic immunodominant determinants are conducted.

Key words: *Plasmodium vivax*; Circumsporozoite protein; Clonal typing; Epitope polymorphism

Introduction

In contrast to the progress towards vaccine development against *P. falciparum* malaria parasite, for which several stage-specific vaccine antigens have been characterized, studies of *P. vivax* parasite antigens have been limited, partly because of the lack of in vitro culture capability. Only two vaccine candidate antigens of *P. vivax* have been characterized: the circumsporozoite (CS) protein and the blood-stage antigen, PV 200 [1–7]. Immune responses against the CS protein have been studied, and vaccine formulations have been produced [8–16].

Vaccination with specific antigens or epitopes offers the possibility of inducing immunity against causative organisms of infectious diseases. However, natural polymorphism in the target determinants of the parasite proteins may compromise the efficacy of a subunit vaccine. Whether a vaccine is based on a critically important antigen present only in a single stage of the parasite or on a combination of important antigens present in several different stages, failure to deal with antigenic polymorphism could result in a vaccine that is unable to protect individuals exposed to a parasite with an antigenic variation. Therefore, whether the vaccine is targeted against single or multistage proteins, if a portion of a

Correspondence address: Altaf A. Lal, Malaria Branch, Mail Stop F 12, Centers for Disease Control, Atlanta, GA 30333, USA. Tel: (404) 488 4079, Fax: (404) 488 4427.

parasite population can escape the effects of a vaccine because of heritable polymorphism in vaccine target antigens, the natural population could rapidly evolve 'vaccine resistance' in a manner analogous to the development of drug resistance in malaria parasites. To ensure the efficacy of subunit vaccines based on polymorphic determinants, antigenic polymorphism of malaria parasites must be studied at the population level.

Polymorphism in the CS protein of *P. falciparum* has been shown to be restricted to the T-cell determinants [17–19]. However, polymorphism in the CS protein of *P. vivax*, which was originally observed in the repetitive B-cell determinant of the protein, has also been recently found outside the repeat region [5]. Variant *P. vivax* parasites from Thailand, Papua New Guinea (PNG) and Brazil have the repeat sequence ANGAGNQPG, which differs from the repetitive sequence GDRADGQPA of *P. vivax* CS proteins of parasites analyzed earlier [4,5]. We and others have subsequently shown that *P. vivax* parasites with the variant B-epitope repeat sequence have a wide geographic distribution [5,20]. In continuation of our earlier studies of the prevalence of CS proteins with the variant repeat sequence ANGAGNQPG in *P. vivax* field isolates from PNG and Brazil, we investigated: (1) nonsynonymous changes in the repeat region of *P. vivax* that result in the emergence of variant repeat sequences; (2) types of CS protein-bearing *P. vivax* parasites that predominate in PNG and Brazil; and (3) polymorphism outside the repeat region, in the T-cell determinants, of the *P. vivax* CS protein.

Materials and Methods

Origin of P. vivax parasites. Microscopically confirmed *P. vivax*-infected blood was collected directly from individuals living in a high malaria-endemic region of PNG (Madang) in March 1990 and from persons in relatively low malaria-endemic regions of Brazil (Paragaminos and around São Paulo) between 1986 and 1990.

Isolation and characterization of the circumsporozoite protein genes. As described earlier [5], blood collected from patients infected with *P. vivax* was passed through a two-layer column of acid-washed glass beads and CF 11 cellulose to remove leukocytes. Parasite genomic DNA was isolated from infected erythrocytes, and 100 ng was used in a polymerase chain reaction (PCR)-mediated CS protein gene amplification. Oligonucleotides AL60 (GTCGGAATTCATGAAGAACTTCATTCTC) and AL61 (CAGCGGATCCTTAATTGAATAATGCTAGG) corresponding to the terminal ends of the CS protein were used as amplification primers. The PCR-generated fragments were digested with *Eco*RI and *Bam*HI, cloned in Bluescript plasmid, and transformed into *Escherichia coli*. The repeat sequence of the recombinant CS clones was identified by DNA:oligonucleotide duplex hybridization analysis using oligonucleotide probes AL114 (ATCAACCAGGAGCAAATG), and AL116 (GGTGATAGAGCAGATGGA) complementary to a portion of the repeat sequences ANGAGNQPG and GDRAAGQPA respectively. AL54 (CCATGCAGTGTAACCTGTGGA) complementary to the genus conserved CS region, R II was also used. CS gene-specific primers were used to determine the nucleotide sequence of the CS gene by the dideoxynucleotide method [21].

Results

On PCR-mediated amplification of the *P. vivax* CS protein gene amplification, genomic DNA extracted from blood samples of the patients from PNG and Brazil yielded DNA fragments approximately 1.2 kb in length. We sequenced the CS protein genes of a total of 115 CS clones representing 15 isolates from PNG and 24 from Brazil. The entire sequence of a total of 13 CS clones representing 8 *P. vivax* isolates from Brazil and 3 clones representing 2 isolates from PNG was determined. The CS genes of the remaining 102 *P. vivax* clones were sequenced to reveal sequences outside the repeat regions and portions of

the central repeat sequences. Of the 115 CS clones, we have previously described the complete sequence of 4 CS genes (P19/D, P4/B, B7/4, and B19/2) and the partial sequence of another 20 CS clones [5]. For the purpose of presentation, we are designating the GDRA(D/A)GQPA and ANGA(G/D)(N/D)QPG CS protein repeats as type 1 and type 2 repeats, respectively.

Within-isolate polymorphism of the circumsporozoite proteins. To determine the presence of parasite polymorphs among patients from PNG, we sequenced 13 clones from the P19 isolate, 10 clones from the P4 isolate, and 2 clones from each of the isolates P6, P7, P9, P10, P18, P25, P26, P72, and P73. A distinct variability in sequences was noted among the clones of P19, P4, P6, and P10 isolates; these variations ranged from a change of a single amino acid residue to insertion of a stretch of amino acids (Fig. 1). The isolate P19 was a mixture of both type 1 and type 2 CS repeat-bearing parasites; two clones, P19/L and P19/M, had type 1 repeat sequences, while the remaining 11 clones had type 2 repeat sequences, indicating that the variants coexist within a single isolate.

Similarly, polymorphism was noted in the repeat and nonrepeat regions of the *P. vivax* CS protein from Brazil (Fig. 1). We sequenced 8 clones from the B19 isolate, seven from the B37, 6 each from the B5 and B10 isolates, 5 from the B7, 3 each from the B11, B15, B21, B31, B34, B40, and B43, and 2 clones from the B14, B20, B26, B30, B38, and B39. Like the mixed infection observed in the P19 isolate from PNG, isolates B7 and B19 had both type 1 and type 2 repeat sequence bearing parasites. However, no amino acid variation was noted among 2–7 clones sequenced from each of the 11 isolates from Brazil.

Evolution of the repeat region of the circumsporozoite protein. Based on the nature of the non-synonymous changes, we have grouped the repeat domain of the CS genes of *P. vivax* into four categories (Fig. 2). The repeat sequences in rows 1–3 encode the type 1 CS repeat sequences, and the sequence in the

							RI		Repeats						RII						
Isolate	Clones	13	38	52	82					294	295	298	310	311	316	341	350	355	359	364	374	388
P19	9	L	G	V	A		*	Type 2		G	D	A	T	N	K	V	A	E	L	T	G	L
	1	.	.	.	.		*		.	.	.	.	.	S	.	.	.	.	.	.	.	.
	1	.	.	.	.		*		.	.	.	.	P	.	.	.	.	.	.	.	.	.
	1	.	N	.	.			Type 1		E	N	+G	P	.	.	.	.	.	.	.	.	.
	1	.	N	.	.			Type 1		–	–	G	L	.	.	.	.	.	.	.	.	.
P4	8	.	.	.	.				.	.	.	.	P	.	I	.	.	.	.	.	.	.
	1	.	.	.	.				.	.	.	.	P	.	.	.	.	.	.	.	.	.
	1	.	.	.	.		*		.	.	.	.	P	.	.	.	.	.	.	.	.	.
P6	1	.	.	.	.				.	.	.	.	P	.	I	.	.	.	.	A	.	.
	1	.	.	.	.				.	.	.	.	P	.	I	.	.	.	.	.	.	.
P10	1	.	.	L	.				.	.	.	.	P	.	I	.	.	.	.	.	.	.
	1	.	.	.	.				.	.	.	.	P	.	I	.	.	G	.	.	.	.
B7	4	.	N	.	.			Type 1		–	–	G	P	.	.	.	.	.	.	.	.	.
	1	.	.	.	.				.	.	.	.	P	.	I	.	T	.	.	.	.	.
B10	4	.	N	.	.			Type 1		–	–	G	P	.	.	.	.	.	.	.	.	.
	1	.	N	.	.			Type 1		–	–	G	P	.	.	.	.	.	.	.	.	S
	1	.	N	.	.			Type 1		–	–	G	P	.	.	.	.	.	S	.	.	.
B11	2	.	N	.	.			Type 1		–	–	G	P	.	.	.	.	.	.	.	.	.
	1	.	N	.	.			Type 1		–	–	G	P	.	.	.	.	.	.	.	S	.
B15	2	.	N	.	.			Type 1		–	–	G	P	.	.	.	.	.	.	.	.	.
	1	.	N	.	.			Type 1		–	–	G	P	.	.	A	.	.	.	.	.	.
B19	7	.	N	.	.			Type 1		–	–	G	P	.	.	.	.	.	.	.	.	.
	1	.	.	.	.				.	.	.	.	P	.	I	.	.	.	.	.	.	.
B21	1	.	N	.	V			Type 1		–	–	.	.	.	.	.	.	.	.	.	.	.
	1	.	N	.	.			Type 1		–	–	G	P	.	.	.	.	.	.	.	.	.
	1	.	.	.	.			Type 1		–	–	.	.	.	.	.	.	.	.	.	.	.
B40	2	.	N	.	.			Type 1		–	–	G	P	.	.	.	.	.	.	.	.	.
	1	P	N	.	.			Type 1		–	–	G	P	.	.	.	.	.	.	.	.	.

Fig. 1. Parasite polymorphs within *Plasmodium vivax* isolates. P designates isolates from Papua New Guinea and B from Brazil. (.) indicates similarity with nine clones of isolate P19; (–) indicates deletion; (*) EDGAGNQP amino acid residues insertion; and (+) GNAGGNA insertion (like NK strain, between residue number 297 and 298). The nucleotide and amino acid residue numbers correspond to the sequence P19/D reported by us earlier [5].

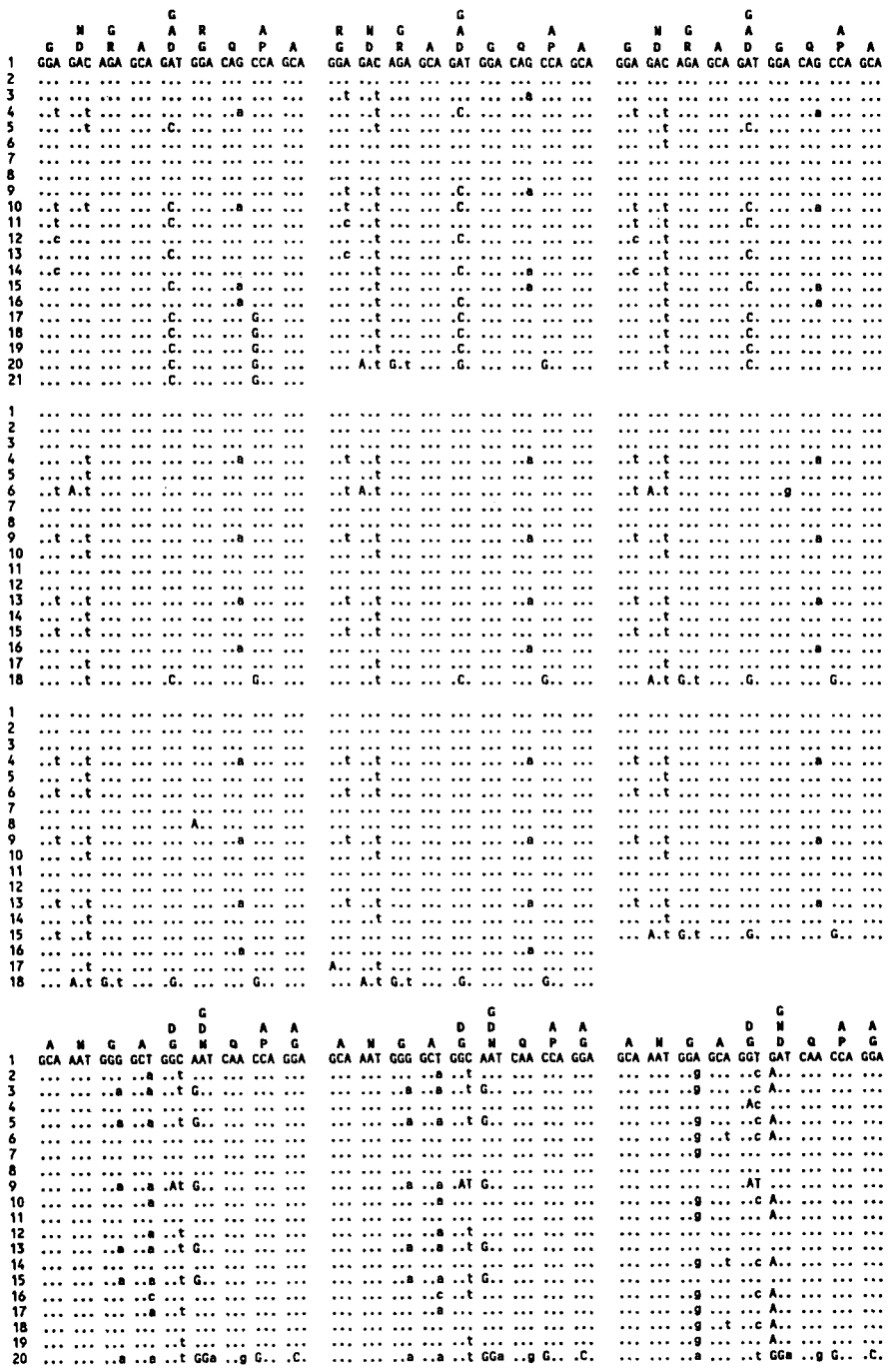

Fig. 2. Nucleotide sequence of the repetitive domains. Silent and non-silent mutations are shown by lower and upper case letters, respectively. Amino acid sequence and changes resulting from nonsilent nucleotides are shown in bold. Block 1, clones B7/5, B3/1 and B30/2 and B13/1; block 2: B5/1, B5/6 and B26/1; block 3: B38/2, B19/3 and B19/8 and B19/5; block 4: Repeat type-2 bearing clones P19/D, P19/B and P4/B, B7/4 and B19/2.

fourth row encodes the type 2 CS repeat sequence. Analysis of the nucleotide sequences of the diverging repeats reveals evolutionary relatedness between these sequences at both the nucleotide and amino acid level (Fig. 3). The CS protein repeat sequence from the parasites B7/5, B3/1, and B30/2, shown in the first row (Fig. 2), is similar to the previously identified sequence of parasites from Brazil [1], El Salvador [2], and North Korea [3], in that the 5th (D/A) and 8th (P/A) amino acid residues in the repeating unit are polymorphic. One clone, B13/1, had no variation in the 8th residue. The CS protein repeat sequences of the B5/1, B5/6, and B26/1 clones, shown in the second row, differ from the above described set of sequences in that no polymorphism was found in the 5th or 8th residue in the repeat unit, except in the degenerative last repeat unit. However, polymorphism was noted in the 2nd amino acid position of the repeat (D/N) in this group. The D/N transition was also found in the previous characterization of the CS gene from the North Korean isolate of *P. vivax*, but in the repeat units 4, 13, 14 and 20 [3]. The third group of the representative CS repeat sequences was seen in clones B38/2, B19/3, B19/8, and B19/5, which had no polymorphism except in the last degenerative repeat. In one clone, B38/2, there was a transversion from G to A at the 6th triplet codon of the 8th repeat.

The fourth category of repeat domain consists of the variant repeats identified in *P. vivax* parasites from PNG and Brazil. All clones from PNG (P19/D, P19/B, and P4/B) and from Brazil (B7/4 and B19/2) are polymorphic at the 5th and 6th amino acid residues, whereas the last degenerative repeat is polymorphic at the 8th and 9th positions as well. No polymorphism was noted in the repeat sequences reported from Thailand isolate [4].

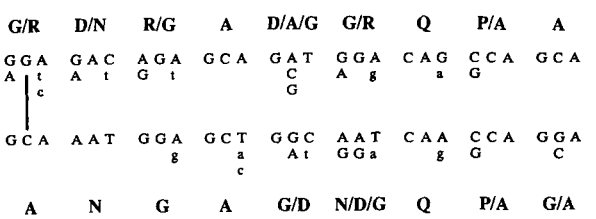

Fig. 3. Accumulation of silent (lower case) and nonsilent (upper case) mutations in a population of type-1 (top) and type-2 (bottom) repeat bearing clones. The deduced amino acid residues are shown in bold. A vertical line indicates the single point mutation that appears distinct in the differentiation of these two repeat types at the protein level.

Polymorphism in the nonrepeat region of CS protein. The comparison of the CS protein sequences outside the repeat region revealed that sequence polymorphism is restricted to three domains: amino to the conserved region, Region I (RI), and amino and carboxyl to the conserved region, Region II (RII) (Fig. 4). A 24-nucleotide insertion amino to the repeat region coding for EDGAGNQP was observed in 11 of the 13 clones in P19 and one of the 10 clones of P4 isolates from PNG. A similar insertion was originally seen in the CS gene of the VK247 *P. vivax* parasite [4].

Amino to the conserved region, RI, polymorphism is restricted to amino acid positions 11, 13, 38, 49, 52 and 82. We have found that CS proteins that bear type 1 CS repeats exhibit greater polymorphism in this domain. Among 73 type 1 CS clones analyzed, polymorphism was identified in 5 of 6 amino acid positions. In contrast, only one of 42 CS clones bearing type 2 repeats exhibited amino acid L instead of V at position 52.

The second and the most variable region is amino to RII. In this region, amino acid residues 294 (G/E), 295 (D/N), 298 (A/G), 310 (T/P/L/), 311 (N/S), and 316 (K/I) are polymorphic in both type 1 and type 2 CS repeat-bearing parasites. The CS protein genes from PNG (clones P19/M and P22/C), which are similar to the CS protein gene from the North Korean (NK) strain [3], have a 16-amino-acid deletion 3′ to the repeat sequence. This deletion is seen in all isolates from Brazil with type 1 repeats. The CS gene of PNG clones P19/L and P28/3 and the previously identified NK strain of *P. vivax* have a 7-amino acid insertion (GNAGGNA) at amino acid number 297 (Fig. 4). The *P. vivax* CS protein

	RI						Repeats		RII														
	11	13	38	49	52	82			294	295	298	310	311	316	341	350	355	359	364	374	388
1	S	L	G	A	V	A	*	Type 2	G	D	A	T	N	K	V	A	E	L	T	G	L
2	.	.	.	.	.		*	.	.	.	.	.	S	.	.	.	.	.	.	.	.
3	.	.	.	.	.	.	*	.	.	.	.	P	.	.	.	.	.	.	.	.	.
4	.	.	.	.	.	.		.	.	.	.	P	.	.	.	.	.	.	.	.	.
5	.	.	.	.	.	.		.	.	.	.	P	.	I	.	.	.	.	.	.	.
6	.	.	.	.	L	.		.	.	.	.	P	.	I	.	.	.	.	.	.	.
7	.	.	.	.	.	.		.	.	.	.	P	.	I	.	T	.	.	.	.	.
8	.	.	.	.	.	.		.	.	.	.	P	.	I	.	.	G	.	.	.	.
9	.	.	.	.	.	.		.	.	.	.	P	.	I	.	.	.	.	A	.	.
10	.	.	N	P	.	.		Type 1	.	.	+G	P	.	.	.	.	.	.	.	.	.
11	.	.	N	.	.	.		Type 1	E	N	+G	P	.	.	.	.	.	.	.	.	.
12	.	.	N	.	.	.		Type 1	-	-	G	L	.	.	.	.	.	.	.	.	.
13	P	.	N	.	.	.		Type 1	-	-	G	P	.	.	.	.	.	.	.	.	.
14	.	.	N	.	.	V		Type 1	-	-	G	P	.	.	.	.	.	.	.	.	.
15	.	.	N	.	.	.		Type 1	-	-	G	P	.	.	A	.	.	.	.	.	.
16	.	.	N	.	.	.		Type 1	-	-	G	P	.	.	.	.	.	S	.	.	.
17	.	.	N	.	.	.		Type 1	-	-	G	P	.	.	.	.	.	.	.	S	.
18	.	.	N	.	.	.		Type 1	-	-	G	P	.	.	.	.	.	.	.	.	S
19	.	.	N	.	.	.		Type 1	-	-	G	P	.	.	.	.	.	.	.	.	.
20	.	.	N	.	.	V		Type 1	-	-	.	.	.	.	.	.	.	.	.	.	.
21	.	.	.	.	.	.		Type 1	-	-	.	.	.	.	.	.	.	.	.	.	.
22	.	P	N	.	.	.		Type 1	-	-	G	P	.	.	.	.	.	.	.	.	.

Fig. 4. 'Clonal types' prevalent in Papua New Guinea (P) and Brazil (B). (•) and (+) are insertions EDGAGNQP and GNAGGNA respectively. 'Clonal-type': (1) 9 clones from single isolate (P19/A-F,H,I,K); (2) 1 clone (P19/G); (3) 1 clone (P19/J); (4) 3 clones from 2 isolates (P4/A, P25/C,D); (5) 24 clones from 11 isolates (P4/B-I, P3/C, P6/D, P7/C,D, P9/C,D, P15/C, P18/C,D, P26/C,D, P72/C,D, P72/C,D, P73/C,D, B19/2); (6) 1 clone (P10/C); (7) 1 clone (B7/4); (8) 1 clone (P10/D); (9) 1 clone (P6/C); (10) 1 clone (P28/C); (11) 1 clone (P19/L); (12) 1 clone (P19/M); (13) 1 clone (B13/1); (14) 1 clone (B3/1); (15) 1 clone (B15/3); (16) 1 clone (B10/3); (17) 1 clone (B11/2); (18) 1 clone (B10/2); (19) 61 clones from 23 isolates (P22/C, B4/1, B5/1-6, B7/2,3,5,6, B10/1,4-6, B11/1,3, B14/1,2, B15/1,2, B17/1, B19/1,4-9, B20/1,2, B21/2, B26/1,2, B28/1, B30/1,2, B31/1-3, B34/1-3, B37/1-7, B38/1,2, B39/1,2, B40/2,3, B42/1, B43/2,4); (20) 1 clone (B21/1); (21) 1 clone (B21/3); (22) 1 clone (B40/1).

sequence in this region, (T/P/L)(N/S)EKSV(K/I)EYLDKVRATVG, analogous to the T-helper site, Th2R, sequence of the *P. falciparum* CS protein, is also polymorphic. However, the immunogenic nature of this region in *P. vivax* needs to be established.

The third variable region is located carboxyl to the conserved region RII. In this region, polymorphism is restricted to amino acids 341 (V to A), 350 (A to T), 355 (E to G), 359 (L to S), 364 (T to A), 374 (G to S), and 388 (L to S). This polymorphic region contains the known T-cell determinants VTCG(V/A)GVRVRRR-VNA(A/T)NKKP, DLE(T/A)DVA-TNDKA, and DKAA(G/S)IFNVVSN and a cell adhesion site VTCG(V/A)GVRVR of the CS protein of *P. vivax* [9–11, 22]. In this region the *P. vivax* CS protein sequence (RVNA(A/T)NKKP(E/G)DLT(L/S)NDLE(T/A), which is analogous to the CTL containing determinant, Th3R, in *P. falciparum*, is also polymorphic.

Clonal diversity of the circumsporozoite protein bearing parasites. To determine the proportional prevalence of parasites with polymorphic CS epitopes in two geographically distant malaria-endemic regions of the world, PNG and Brazil, we have grouped the CS protein sequences of 115 clones analyzed here, based on both the repeat region type and outside repeat region polymorphism. Our comparison identified 22 different 'clonal types' of CS proteins and revealed that with one exception in each PNG and Brazil, the CS clonal types of *P. vivax* parasites from PNG were distinct from those from Brazil. For instance, CS clonal type 5 was found predominantly in Madang, PNG; 24 CS clones representing 11 of 15 *P. vivax* isolates from Madang, PNG were clonal type 5. In contrast, clonal type 19 was most prevalent in *P. vivax* parasites from Paragaminos, Brazil; 61 clones representing 23 of the 24 isolates from Brazil contained this form of polymorphic CS protein.

Discussion

We have undertaken a longitudinal, population level study of the CS protein of *P. vivax* from two distant malaria-endemic regions, PNG and Brazil, to elucidate the nature and extent of epitope variability in this vaccine candidate antigen. We have found that both repeat and nonrepeat regions of the *P. vivax* CS protein are polymorphic. Sequence comparison of the repeat of the CS proteins, aligned in groups of type 1 and type 2 CS repeats, shows a pattern of both silent and nonsynonymous mutations and indicates that point mutations have led to the emergence of variant repeat sequences (Figs. 2 and 3). A single point mutation, G to C (second nucleotide of first triplet) in the common genetic pool of the CS clones analyzed here, representing type 1 and type 2 CS repeats, makes these two repeat types different at the protein level (Fig. 3). The array of mutations and amino acid substitutions present in the population of type 1 repeats suggests that this type of CS repeat sequence may have mutated to type 2 variant repeats; however, the reverse cannot be ruled out. Examination of repeat diversity in the CS proteins of other human malaria parasites showed a similar point mutation based diversity in the repeat sequence. For example, the change of the repeat sequence from NAAG to NDAG in *Plasmodium malariae* [23] and from NANP to NVDP in *P. falciparum* [24] can be explained by point mutations and subsequent expansion of the repeating unit.

Genetic recombination between homologous single-copy CS alleles during meiosis would not result in progeny with variant repeat types in the event of a single or even double crossover. However, mutations that contribute to the polymorphic nature of the regions outside the repeat region can be redistributed between the progeny. A new repeat type can only emerge if there is crossover inside the repeat region between alleles with different repeat types, in which case the progeny will have a part of each type of repeat with a distinct junction to predict the exact position of genetic exchanges. Such a genetic event, which has been seen in the *P. falciparum* merozoite surface antigen gene [25], has not been observed in CS protein genes. However, this does not rule out the existence of *P. vivax* parasites with CS repeat sequences distinct from the known type 1 and type 2 sequences.

In the nonrepeat regions of the CS protein, nonsilent nucleotide substitutions are localized in the identified T-cell epitope and in regions analogous to the immunodominant Th2R and Th3R regions of the *P. falciparum* CS protein. We propose that these two polymorphic regions may harbor immunodominant determinants of the protein; the underlying rationale is that mutations in the parasite's proteins that interface with host immunity, and which are advantageous to the parasite, would be positively selected rapidly [26]. This school of thought is contrary to the other opinion that favours DNA based selection and maintainance of polymorphism in the CS protein of human malaria parasites [27,28]. However, irrespective of the origin and the mechanism that maintains polymorphic sequence bearing parasites studies of antigenic polymorphism of parasite candidate vaccine antigens constitutes a step towards identifying vaccine-essential immunodominant determinants of parasite proteins. It remains to be determined experimentally whether Th2R and Th3R analogous regions in the *P. vivax* CS protein are immunogenic in nature. Information on the sequence variation of these regions would allow us and others to test the vaccine related effects of polymorphism.

Comparison of the CS protein gene sequences of field isolates of *P. vivax* from distant geographic regions, Brazil and PNG, revealed that each of these regions had distinct types of polymorphic CS protein-bearing parasites. However, we also found that certain types of polymorphic CS proteins were present in both PNG and Brazil. Moreover, mutations leading to individual amino acid changes were shared by Papuan and Brazilian parasites, indicating, as has been suggested before, a common but independent origin of sequence polymorphism

[29]. Our results suggest that both immunologic and genetic events are operational in the selection and expansion of the repeat sequence.

To summarize, we have shown that polymorphism in the CS protein of *P. vivax* is resident in both repeat and nonrepeat regions, that point mutations cause repeat sequence-based antigenic diversity in the CS protein, and that different clonal types of *P. vivax* predominate in the malaria-endemic regions of PNG and Brazil investigated here. Similar cross-sectional and longitudinal studies of vaccine antigen diversity are essential before malaria vaccine programs can be implemented in malaria-endemic regions. While a cross-sectional geographic level study would reveal the extent and proportional prevalence of polymorphic epitope-bearing malaria parasites, a longitudinal study at the population level would help clarify the dynamics of parasite populations as they relate to parasite infectivity and disease.

Acknowledgements

This work was supported in part by USAID PASA Grant BST-0453-P-HC-2086-07, National Institute of Allergy and Infectious Diseases/Centers for Disease Control Intra-agency Agreement 1-Y02-AI00006-01, and United Nations Developmental Program/World Bank/World Health Organization Special Program in Research and Training in Tropical Diseases (Grants 890520 and 870284).

References

1 Arnot, D.E., Barnewell, J.W., Tam, J.P., Nussenzweig, V., Nussenzweig, R. and Enea, V. (1985) Circumsporozoite protein of *Plasmodium vivax*: gene cloning and characterization of the immunodominant epitope. Science 230, 815–816.

2 McCutchan, T.F., Lal, A.A., de la Cruz, V., Miller, L.H., Maloy, W.L., Charoenvit, R.L., Guerry, P., Wistar, R., Hoffman, S.L., Hockmeyer, W.T., Collins, W.E. and Wirth, D. (1985) Sequence of the immunodominant epitope for the surface protein on sporozoites of *Plasmodium vivax*. Science 230, 1381–1383.

3 Arnot, D.E., Barnwell, J.W. and Stewart, M.J. (1988) Does biased gene conversion influence polymorphism in the circumsporozoite protein-encoding gene of *Plasmodium vivax*? Proc. Natl. Acad. Sci. USA 85, 8102–8106.

4 Rosenberg, R., Wirtz, R.A., Lanar, D.E., Sattabongkot, J., Hall, T., Watters, A.P. and Prasittisuk, C. (1989) Circumsporozoite protein heterogeneity in the human malaria parasite *Plasmodium vivax*. Science 245, 973–976.

5 Qari, S.H., Goldman, I.F., Povoa, M.M., Oliveira, S., Alpers, M.P. and Lal, A.A. (1991) Wide distribution of the variant form of the human malaria parasite *Plasmodium vivax*. J. Biol. Chem. 266, 16297–16300.

6 Del Portillo, H.A., Gysin, J., Mattei, D., Khouri, E., Udagama, P., Mendis, K. and David, P. (1988) *Plasmodium vivax*: cloning and expression of a major blood stage surface antigen. Exp. Parasitol. 67, 346–353.

7 Del Portilo, H.A., Longacre, S., Khouri, E. and David, P. (1991) Primary structure of the merozoite surface antigen 1 of *Plasmodium vivax* reveals sequences conserved between different *Plasmodium* species. Proc. Natl. Acad. Sci. USA 88, 4030–4034.

8 Charoenvit, Y., Collins, W.E., Jones, T.R., Millet, P., Yuan, L., Campbell, G.H., Beaudoin, R.L., Broderson, J.R. and Hoffman, S.L. (1991) Inability of malaria vaccine to induce antibodies to a protective epitope within its sequence. Science 251, 668–671.

9 Nardin, E., Clavijo, P., Mons, B., Belkum, A.V., Ponnudurai, T. and Nussenzweig, R.S. (1991) T cell epitopes of the circumsporozoite protein of *Plasmodium vivax*: recognition by lymphocytes of a sporozoite-immunized chimpanzee. J. Immunol. 146, 1674–1678.

10 Rodrigues, M.M., Paiva, A.C., Dutra, A.P., Yoshida, N. and Nakaie, C. (1991) Identification of epitopes within the circumsporozoite protein of *Plasmodium vivax* recognized by murine T lymphocytes. Exp. Parasitol. 72, 271–277.

11 George IV, F.W., Law, J.L., Rich, K.A. and Martin, W.J. (1990) Identification of a T-cell epitope on the circumsporozoite protein of *Plasmodium vivax*. Infect. Immun. 58, 575–578.

12 Burkhot, T., Graves, P., Wirtz, R., Brabin, B., Battistutta, D., Cattani, J., Maizeles, R. and Alpers, M. (1989) Differential antibody responses to *Plasmodium falciparum* and *P. vivax* circumsporozoite proteins in a human population. J. Clin. Microbiol. 27, 1346–1351.

13 Fontes C.J., Bathurst, J. and Krettli, A.U. (1991) *Plasmodium vivax* sporozoite antibodies in individuals exposed during a single malaria outbreak in a non-endemic area. J. Trop. Med. Hyg. 44, 28–33.

14 Brown, A.E., Webster, H., Krinchai, K., Gordon, D., Wirtz, R. and Permpanich, B. (1991) Characteristics of natural antibody responses to the circumsporozoite protein of *Plasmodium vivax*. Am. J. Trop. Med. Hyg. 44, 21–27.

15 Wirtz, W.A., Rosenberg, R., Sattabongkot, J. and Webster, H. (1990) Prevalence of antibody to heterologous circumsporozoite protein of *Plasmodium vivax* in Thailand. Lancet 336, 593–595.

16 Cochrane, A.H., Nardin, E., Arruda, M., Maracis, M., Clavijo, P., Collins, W.E. and Nussenzweig, R. (1990) Wide spread reactivity of human sera with a variant repeat of the circumsporozoite protein of *Plasmodium vivax*. Am. J. Trop. Med. Hyg. 43, 446–451.

17 De la Cruz, V.F., Lal, A.A. and McCutchan, T.F. (1987) Sequence variation in putative functional

domains of the circumsporozoite protein of *Plasmodium falciparum* J. Biol. Chem. 262, 11935–11939.
18 Yoshida, N., Di Santi, S., Dutra, A., Nussenzweig, R., Nussenzweig, V. and Enea, V. (1990) *Plasmodium falciparum*: restricted polymorphism of T cell epitopes of the circumsporozoite protein in Brazil. Exp. Parasitol. 71, 386–392.
19 Shi, Ya Ping, Alpers, M.P., Povoa, M.M. and Lal, A.A. (In Press) Diversity in the immunodominant determinants of the circumsporozoite protein of *Plasmodium falciparum*. Am. J. Trop. Med. Hyg.
20 Kain, K.C., Keystone, J., Franke, E.D. and Lanar, D.E. (1991) Global distribution of a variant of the circumsporozoite gene of *Plasmodium vivax*. J. Infect. Dis. 164, 208–210.
21 Sanger, F., Nioklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463–5466.
22 Rich, K.A., George IV, F.W., Law, J.L. and Martin, W.J. (1990) Cell-adhesive motif in Region II of malaria circumsporozoite protein. Science 249, 1574–1577.
23 Lal, A.L., de la Cruz, V., Campbell, G.H., Procell, P.M., Collins, W.E. and McCutchan, T.F. (1988) Structure of the circumsporozoite gene of *Plasmodium malariae*. Mol. Biochem. Parasitol. 30, 291–294.
24 Dame, J.B., Williams, J., McCutchan, T., Weber, J., Wirtz, R., Hockmeyer, W., Maloy, W., Haynes, J., Schneider, I., Roberts, D., Sanders, G., Reddy, E., Diggs, C. and Miller, L.H. (1984) Structure of the gene encoding the immunodominant surface antigen on the sporozoite of the human malaria parasite *Plasmodium falciparum*. Science 225, 593–599.
25 Marshall, V.M., Coppel, R.L., Martin, R.K., Oduola, A.M., Anders, R.F. and Kemp, D.K. (1991) A *Plasmodium falciparum* MSA-2 gene apparently generated by intragenic recombination between the two allelic families. Mol. Biochem. Parasitol. 45, 349–352.
26 McCutchan, T.F., Good, M.F. and Miller, L.H. (1989) Polymorphism in the circumsporozoite (CS) protein of *Plasmodium falciparum*. Parasitol. Today 5, 143–146.
27 Arnot, D.E. (1989) Malaria and the major histocompatibility complex. Parasitol. Today 5, 138–142.
28 Arnot, D.E. (1991) Possible mechanism for the maintenance of polymorphisms in *Plasmodium* population. Acta Leidensia 60, 29–35.
29 McCutchan, T.F., Lal, A.A., Do Rosario, V. and Waters, A.P. (1992) Two types of sequence polymorphism in the circumsporozoite gene of *Plasmodium falciparum*. Mol. Biochem. Parasitol. 50, 37–46.