A SURVEY OF THE Th2R AND Th3R ALLELIC VARIANTS IN THE CIRCUMSPOROZOITE PROTEIN GENE OF *P. FALCIPARUM* PARASITES FROM WESTERN THAILAND

Chutima Kumkhaek[1], Kooruethai Phra-ek[2], Pratap Singhasivanon[3], Sornchai Looareesuwan[4], Chakrit Hirunpetcharat[5], Alan Brockman[2], Anne Charlotte Grüner[6], Nicolas Lebrun[7], Laurent Rénia[6], François Nosten[2], Georges Snounou[8] and Srisin Khusmith[1]

[1] Department of Microbiology and Immunology; [3]Department of Tropical Hygiene, [4]Department of Tropical Medicine, Faculty of Tropical Medicine; Mahidol University, Bangkok; [5]Department of Microbiology, Faculty of Public Health, Mahidol University, Bangkok; [2]Shoklo Malaria Research Unit, Mae Sot, Thailand; [6]Département d'Immunologie; [7]Service Commun de Séquençage, Institut Cochin, INSERM U567, CNRS UMR 8104, Université René Descartes, Paris; [8]Unité de Parasitologie Bio-Medicale, CNRS URA 2581, Institut Pasteur, Paris, France

Abstract. Allelic variation in the *Plasmodium falciparum* circumsporozoite protein (CS) gene has been determined by sequencing the immunodominant T-cell epitopes, Th2R and Th3R, from 95 isolates from two malaria-endemic areas in the west of Thailand. Comparison with a reference sequence revealed only non-synonymous point mutations in the two epitope regions. Point mutations were found outside these epitopes in a minority of samples, and all but four were also non-synonymous. A relatively high number of variants, 11 Th2R and 9 Th3R, were detected and comprised some that had not been previously observed. However, the Th2R*05 and the Th3R*01 allelic variants predominated, as they were found in more than 70% of the 101 sequences obtained.

INTRODUCTION

Circumsporozoite protein (CS) is an antigen found at the surface of the sporozoite (Yoshida *et al*, 1980), the form of the malaria parasite transmitted by the mosquito. CS has long been considered an important target of protective immune response against pre-erythrocytic parasites (sporozoites and hepatic stages). The *P. falciparum* CS gene, *pfcs*, encodes a protein consisting of relatively conserved regions flanking a highly-repetitive domain composed of tetrapeptide repeats; the sequence of the majority is NANP and that of the remainder is NVDP (Dame *et al*, 1984). The immunodominant B-cell epitopes are confined to this repeat region, while T-cell epitopes are found in the C-terminal domain (Zavala *et al*, 1983; Good *et al*, 1988). These epitopes provide T-cell help for the production of anti-sporozoite antibody that inhibits hepatocyte invasion by sporozoites, and induce $CD4^+$ and $CD8^+$ cytotoxic T-effector cells that are thought to kill infected hepatocytes (Good *et al*, 1988). The CS structure of other parasite species is similar to that described above, though the sequences of the repetitive elements differ substantially between species.

CS has been the focus of immunological studies since the 1970s and its gene was the first *Plasmodium* gene to be cloned and partially sequenced (Godson *et al*, 1983). Numerous experimental vaccines thus include, or are based on, CS (Richie and Saul, 2002). The most recent, RTS,S, consists of 19 NANP repeats as well as the carboxyl terminus of the CS fused to the hepatitis B surface antigen (Stoute *et al*, 1998). To date, this formulation associated to the potent adjuvant (SBAS2) proved the most efficacious in protecting vaccinated human volunteers against an experimental challenge with a homologous strain (Stoute *et al*, 1998). However, the protection observed in a Phase II trial of this vaccine in Gambian adults proved to be of short duration (Bojang *et al*, 2001).

Correspondence: Dr Srisin Khusmith, Department of Microbiology and Immunology, Faculty of Tropical Medicine, Mahidol University, 420/6 Rajvithi Rd, Bangkok 10400, Thailand.
Tel: +66 (0) 2354 9100-19 ext 1594; Fax: +66 (0) 2643 5583
E-mail: tmskm@mahidol.ac.th

Polymorphism in the vaccine candidate antigen sequence is a major concern, since it might adversely influence the induction of immunity as well as provide the parasite with a means to escape the protective responses induced. Sequencing of *pfcs* genes or gene fragments from laboratory and field isolates revealed that this antigen, like many others in *P. falciparum*, displays extensive genetic diversity (De La Cruz *et al*, 1987; Lockyer *et al*, 1989; Yoshida *et al*, 1990; Doolan *et al*, 1992; Qari *et al*, 1992; Shi *et al*, 1992; Jongwutiwes *et al*, 1994; De Stricker *et al*, 2000; Escalante *et al*, 2002). The number and arrangement of the two repeated tetrapeptides varied with the parasite line, though this was not found to alter binding of antibodies (Zavala *et al*, 1985). Variations were found in the pre- and post-repeat regions, as were occasional point mutations in the N-terminal region, and their immunological relevance has yet to be established. However, the most striking polymorphisms consisted of a series of non-synonymous point mutations centered on two important T-helper epitopes, Th2R and Th3R, found in the C-terminal domain of the molecule (Lockyer *et al*, 1989). Variation within the CS gene of *P. falciparum* is postulated to result from immune selection. Correlation between human T-cell proliferations to the Th3R epitope with protection from falciparum malaria infection, was observed for a limited number of patients from a malaria endemic area (Hoffman *et al*, 1989). Recent studies further showed that human $CD4^+$ T-cell clones specific for a *P. falciparum* Th/Tc epitope that overlaps the Th2R epitope, recognized a large number of variant peptides that correspond to polymorphisms detected in *P. falciparum* isolates from different geographical areas (Moreno *et al*, 1993).

In view of the concern that *pfcs* T-cell epitope polymorphisms might impinge on the efficacy of CS-based vaccines, field isolates from diverse geographic locations were analyzed for Th2R and Th3R diversity. Two trends emerged from these studies. The degree of variation observed, though large, did not fulfil the potential indicated by the number of positions where alternate residues were noted. The distribution of allelic variants varied with the geographical origin of the parasites. The lowest degree of variation was found in *P. falciparum* from Brazil, Papua New Guinea, and Thailand, whilst the highest was observed in parasites of African origin.

Since sequencing was used to gather the above data, it was often limited to a relatively small number of field samples. A hybridization-based method, developed for the analysis of samples from the Gambian RTS, S trial, provided a means to survey diversity in a larger number of samples (Alloueche *et al*, 2000). We wished to employ this method to analyze diversity in *P. falciparum* parasites obtained from two regions located on the Thai-Myanmar border. However, the oligonucleotides developed for PCR-SSOP did not include some of the Th2R and Th3R allelic variants described after its inception. In order to establish whether this set of PCR-SSOP oligonucleotides can be meaningfully used for Thai samples, we surveyed the diversity of the CS T-cell epitopes by sequencing the corresponding amplified region from a large number of samples obtained from patients infected with *P. falciparum* in two regions of western Thailand.

MATERIALS AND METHODS

Blood sample collection

A total of 95 blood samples diagnosed with *P. falciparum* infection were obtained from patients with informed consent. The ethical issue of this study has been approved by the Ethics Committee of the Faculty of Tropical Medicine. Forty samples were obtained from patients admitted to the Hospital for Tropical Diseases, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand from June to September, 2001 (n = 40) with infections acquired in Phetchaburi (n = 1), Ratchaburi (n = 31), Kanchanaburi (n = 4) or Tak (n = 4) provinces. The remaining 55 samples were collected from patients admitted to the clinic at the Mae La Karen refugee camp north of Mae Sot, Tak Province during 1996 (n = 19), 1999 (n = 17), or 2001 (n = 19). In all cases, admission blood samples were collected in ethylene diamine tetraacetic acid (EDTA) tubes. Two hundred microliters of blood were centrifuged and the red blood cell pellets were stored at -20ºC.

DNA amplification, cloning and sequencing

Amplification by polymerase chain reaction

(PCR) was carried out as previously described (Alloueche *et al*, 2000), with minor modifications in the preparation of the template. Briefly, 5 µl of lysed freeze-thawed infected red blood cell pellet were added to 150 µl phosphate-buffered saline solution and mixed by brief vortexing. The released parasites were pelleted by centrifugation for 5 minutes at 5,000*g*, and the supernatant carefully discarded. Fifty µl of amplification reaction mixture, containing 1x PCR buffer, 200 nM of each primer, 250 µM dNTPs, 2.0 mM $MgCl_2$ and 2.5 U Taq DNA polymerase (Applied Biosystems, UK), were added directly to the pellet before initiating the amplification cycle. The primers were designed to hybridize to conserved regions spanning the Th2R and Th3R region (forward primer; 5′-ACAATCAAGGTAATGGACAAGG-3′ and reverse primer; 5′-ACGACATTAAACACA CACTGGAAC-3′), and to result in the amplification of a 319 base-pair fragment. The cycling conditions were 95°C for 5 minutes, 58°C for 2 minutes and 72°C for 2 minutes for 1 cycle, then 94°C for 1 minute, 58°C for 2 minutes and 72°C for 1 minute for 32 cycles followed by 10 minutes extension at 72°C. The PCR amplifications were performed using a PTC 200 (MJ Research, USA). The amplified products were electrophoresed on 1.5% agarose gel in Tris-borate-EDTA buffer, and visualized under UV light.

The unique PCR product obtained for each sample was purified using the QIAquick gel extraction kits (QIAGEN, Germany), and cloned using the TOPO TA Cloning Kit (Invitrogen, USA). Plasmid DNA containing the CS fragment was purified from positive bacterial colonies using the QIAquick Miniprep Spin Kit (QIAGEN, Germany). Sequencing was performed by automated sequencer at the Hôpital Cochin, Paris, France. Sequence alignments were performed using the Gene Jockey II program (Biosoft, UK). The nomenclature for the Th2R and Th3R allelic variants and the designation for new allelic variants were proposed in reference to the scheme followed by Alloueche and colleagues (2000).

RESULTS

Following DNA purification from 95 isolates, a 319 bp fragment at the 3′-end of the *pfcs* gene, encoding the C-terminal domain where the T-cell epitopes are found, was amplified for each isolate. The PCR fragments were cloned and sequenced, and the sequences compared to the corresponding sequence of the previously published 7G8 *pfcs* gene (REF). The comparison was confined to 275 bp of the fragment, as the 5′- and 3′-end corresponding to the oligonucleotides used for the amplification were excluded from analysis. Given that a single isolate could contain a mixture of genetically different parasites, three colonies were picked for sequencing for 15 of the isolates. For most of these, the same sequence was obtained. However, for three of the isolates, one or more of the cloned fragments differed in sequence. In total, 101 of the cloned fragments were included for analysis of *pfcs* 3′-end domain polymorphisms.

None of the 101 sequences obtained exactly matched the reference sequence from the 7G8 *pfcs* gene (Dame *et al*, 1984), and all differences observed were point mutations (Fig 1). Mutations occurring outside the two T-helper epitopes were only observed in 12 of the sequences obtained (Fig 1). These mutations were found in 15 positions corresponding to 14 codons, and proved to be synonymous in only 4 cases. Mutations occurring in the Th2R and/or Th3R T-helper epitopes, however, accounted for the major proportion of the polymorphisms observed (Fig 1). As compared to the reference 7G8 sequence, mutations were found at 9 of the 17 residues of the Th2R epitope, and at six of the twelve residues of the Th3R epitope (Figs 2 and 3). All the point mutations observed in these T-helper epitopes led to an alteration in the amino acid encoded by the corresponding codon, *ie* non-synonymous. Overall, the non-synonymous mutations were mostly associated with mutations at the first or second bases of the codon, while for the four silent mutations the third base of the codon was altered.

A high diversity of Th2R and Th3R epitopes was observed in the samples analyzed, since 11 Th2R types and 9 Th3R types were detected in the 101 sequences analyzed. Most of the variants were, however, only observed at a low frequency, while the Th2R*05 and the Th3R*01 were found in 84 and 90% of the sequences, respectively. The

```
                         His Asn Met Pro Asn Asp Pro Asn Arg Asn Val Asp Glu Asn Ala Asn Ala Asn Asn Ala      27
1    ACAATCAAGGTAATGGACAAGG T CAC AAT ATG CCA AAT GAC CCA AAC CGA AAT GTA GAT GAA AAT GCT AAT GCC AAC AAT GCT

     Val Lys Asn Asn Asn Asn Glu Glu Pro Ser Asp Lys His Ile Glu Gln Tyr Leu Lys Lys Ile Gln Asn Ser Leu Ser Thr      54
84   GTA AAA AAT AAT AAT AAC GAA GAA CCA AGT GAT AAG CAC ATA GAA CAA TAT TTA AAG AAA ATA CAA AAT TCT CTT TCA ACT

     Glu Trp Ser Pro Cys Ser Val Thr Cys Gly Asn Gly Ile Gln Val Arg Ile Lys Pro Gly Ser Ala Asn Lys Pro Lys Asp      81
165  GAA TGG TCC CCA TGT AGT GTA ACT TGT GGA AAT GGT ATT CAA GTT AGA ATA AAG CCT GGC TCT GCT AAT AAA CCT AAA GAC

     Glu Leu Asp Tyr Glu Asn Asp Ile Glu Lys Lys Ile Cys Lys Met Glu Lys      105
246  GAA TTA GAT TAT GAA AAT GAT ATT GAA AAA AAA ATT TGT AAA ATG GAA AAA  TGTTCCAGTGTGTTTAATGTCGT
```

Fig 1–DNA and amino acid sequence of the amplified fragment. The sequence is based on the *pfcs* gene of the 7G8 line of *P. falciparum*, Accession number K02194 (Dame *et al*, 1984). The oligonucleotide used for PCR are in italic. The amino acid sequences of the Th2R and Th3R epitopes are highlighted. The bases where silent point mutations were observed are lightly highlighted, whereas those associated with non-synonymous mutations are highlighted in black.

Th2R*01	P CCA	S AGT	D GAT	K AAG	H CAC	I ATA	E GAA	Q CAA	Y TAT	L TTA	K AAG	K AAA	I ATA	K AAA	N AAT	S TCT	I ATT
Th2R*05														Q C--			L C--
Th2R*22						M --T								Q C--			L C--
Th2R*15								E G--						Q C--			L C--
Th2R*12							T AC-	E G--						Q C--			L C--
Th2R*20								K A--				E G--		Q C--			L C--
Th2R*17												R -G-		Q C--			L C--
Th2R*18												N --T		Q C--			L C--
Th2R*21													M --T	Q C--			L C--
Th2R*16												R -G-	M --T	Q C--			L C--
Th2R*13														Q C--	Y T--		L C--
Th2R*19		G G--												Q C--	Y T--		L C--

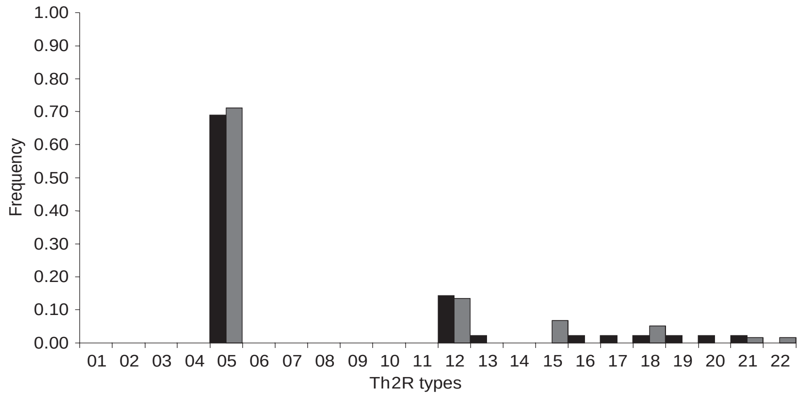

Fig 2–The sequence and frequency of the Th2R allelic variants observed in Thai *P. falciparum* isolates. The sequences obtained are compared to the DNA and amino acid sequences of the Th2R*01 type present in the *pfcs* gene of the 7G8 parasite line (Dame *et al*, 1984), that was not observed for any of the sequences obtained in this study. For each of the 11 types observed in the Thai isolates, the mutated base in each codon is shown below the corresponding altered amino acid, and dashes represent identity. The remaining codons are left blank to indicate no sequence variation. Black bars represent samples collected mainly in Ratchaburi Province and stippled bars samples collected in Tak Province.

pattern of frequency distribution was not significantly different for the samples obtained from the two different geographical regions, or collected in different years.

Among the variants observed in Thai isolates, six Th2R allelic types (Th2R*16 to Th2R*22), and four Th3R allelic types (Th3R*14to Th3R*17) were not included in the selection used to design SSOP oligonucleotides (Allloueche *et al*, 2000).

DISCUSSION

In this study, we surveyed the allelic variations of the immunodominant T-cell epitope regions, Th2R and Th3R, in the *P. falciparum* CS gene from a large number of field isolates collected from distinct malaria-endemic areas close to the Thai-Myanmar border. This was achieved by sequencing

cloned fragments derived by DNA amplification. Despite the fact that numerous allelic variants for both epitope regions were detected, the overall diversity could be considered restricted, since the majority of the sequences belonged to a single allelic variant.

When compared with the list of variants, (Th2R*01 to Th2R*15) and (Th3R*01 to Th3R*13) considered when the PCR-SSOP methodology was developed (Alloueche *et al*, 2000), we observed 7 novel variants for Th2R and 4 novel variants for Th3R. However, a review of the published sequences for these epitopes revealed that a total of 28 variants for Th2R, and 17 variants for Th3R, had been observed. Of the novel variants observed in the Thai isolates used in this study, only one (Th2R*20) had been observed previously (Accession number AF181833). Mutations leading to amino acid substitutions at the 2nd, 6th and 13th residues of the Th2R epitopes (Th2R*19, Th2R*22, and Th2R*21 plus Th2R*16, respectively) had not been previously observed, nor had those at the 2nd residue of the Th3R epitopes (Th3R*15 and Th3R*16). Since 4 of the novel Th2R and the 4 novel Th3R allelic types were observed only in a single sequence, it is possible that they are the result of amplification artefact. It is felt that this is unlikely for a number of reasons: a) the error rate observed in this study appeared to be quite low, since in 75 of the 101 cloned fragments (a total of 20,625 bp), the sequences were identical, b) all but four of the mutations observed were non-synonymous, and c) most mutations observed for the two epitope regions were found at residues previously shown to be variable.

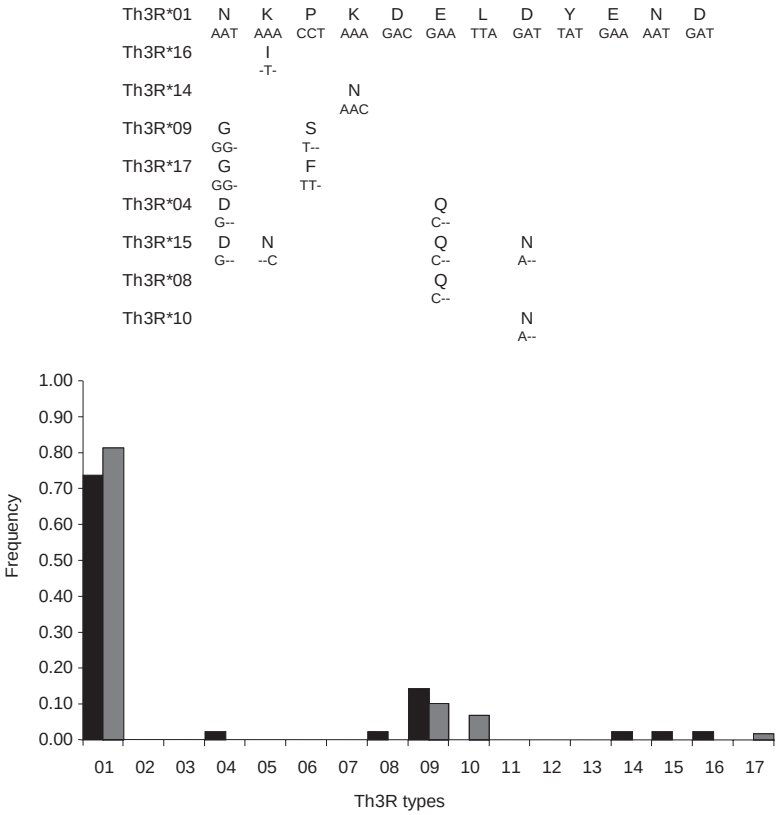

Fig 3–The sequence and frequency of the Th3R allelic variants observed in Thai *P. falciparum* isolates. The sequences obtained are compared to the DNA and amino acid sequences of the Th2R*01 type present in the *pfcs* gene of the 7G8 parasite line (Dame *et al*, 1984), that was found in the majority of the sequences obtained in this study. For each of the eight other types observed in the Thai isolates, the mutated base in each codon is shown below the corresponding altered amino acid, and dashes represent identity. The remaining codons are left blank to indicate no sequence variation. Black bars represent samples collected mainly in Ratchaburi Province and stippled bars samples collected in Tak Province.

The frequency distribution of the Th2R and Th3R allelic types can be interpreted as evidence for a relatively low diversity of these *pfcs* epitopes in Thailand. These results concur similar studies conducted in the same area of Thailand (Mae Sot district) on samples collected in 1988-1989, where the Th2R*05 and Th3R*01 allelic variants were also found to be predominant (Jongwutiwes *et al*, 1994). Similar levels of diversity, where one allelic variant predominates, have been found in Brazil (Yoshida *et al*, 1990; Shi *et al*, 1992) and Papua New Guinea (Doolan *et al*, 1992; Shi *et al*, 1992).

This contrasts with higher levels of diversity observed in Africa (Lockyer *et al*, 1989; Alloueche *et al*, 2000; Escalante *et al*, 2002). Such a pattern is consistent with the notion that parasite populations in areas of high transmission intensity maintain a higher effective population size, recombine more frequently and are consequently more subject to diversifying positive natural selection, thus sustaining higher levels of parasite polymorphism (Escalante *et al*, 2002). Transmission intensities in Thailand are 1 to 2 orders of magnitudes lower than those calculated for hyperendemic areas in sub-Saharan Africa.

Allelic variants of both the Th2R and Th3R regions have been shown to affect specific $CD4^+$ and $CD8^+$ T-cell responses (Guttinger *et al*, 1988; Plebanski *et al*, 1997). Murine and human T-cells primed with one sequence usually fail to respond to other Th2R and Th3R variants (De La Cruz *et al*, 1988; Guttinger *et al*, 1988), though a study of human T-cell proliferation in immune Gambians suggested some T-cell cross-reactivity to variant Th2R and Th3R peptides (Good *et al*, 1988). The non-synonymous nature of the T-cell epitopes analyzed sustains the hypothesis that these mutations arose, and are maintained, as a result of immune selection. Such an origin for antigenic diversity in the *pfcs* epitopes would have implications for the development of protective immunity and its induction by vaccination.

Since the PCR-SSOP methodology (Alloueche *et al*, 2000) would be suitable to type a majority of the circulating *P. falciparum* populations, longitudinal epidemiological studies of *pfcs* diversity can now be envisaged. Furthermore, the restricted diversity of the Th2R and Th3R *pfcs* epitopes in Thai parasites could be exploited to assess the functional capacity of these variants in T-cell activation. These types of study can be instrumental in determining whether CS antigen polymorphism represents in fact a major obstacle to developing an anti-sporozoite vaccine against the malaria parasite.

ACKNOWLEDGEMENTS

The authors wish to thank the patients for their kind participation in this study and the staff of the Hospital for Tropical Diseases (Faculty of Tropical Medicine, Mahidol University), and of the clinical center at the Mae La Karen refugee camp, for their assistance in blood collection and slide reading. We also wish to thank Dr Porntip Petmitr (Department of Protozoology, Mahidol University) for her kindness in providing *P. falciparum* K1 strains. This study was supported by a grant from the VIHPAL Program (Ministère de la Recherche, France), the Royal Golden Jubilee PhD Program, the Basic Research Grant for the Golden Jubilee PhD of The Thailand Research Fund and Mahidol University Grant.

REFERENCES

Alloueche A, Silveira H, Conway DJ, *et al*. High-throughput sequence typing of T-cell epitope polymorphisms in *Plasmodium falciparum* circumsporozoite protein. *Mol Biochem Parasitol* 2000; 106: 273-82.

Bojang KA, Milligan PJ, Pinder M, *et al*. Efficacy of RTS,S/AS02 malaria vaccine against *Plasmodium falciparum* infection in semi-immune adult men in The Gambia: a randomised trial. *Lancet* 2001; 358: 1927-34.

Dame JB, Williams JL, McCutchan TF, *et al*. Structure of the gene encoding the immunodominant surface antigen on the sporozoite of the human malaria parasite *Plasmodium falciparum*. *Science* 1984; 225: 593-9.

De La Cruz VF, Lal AA, McCutchan TF. Sequence variation in putative functional domains of the circumsporozoite protein of *Plasmodium falciparum*. Implications for vaccine development. *J Biol Chem* 1987; 262: 11935-9.

De La Cruz VF, Maloy WL, Miller LH, Lal AA, Good MF, McCutchan TF. Lack of cross-reactivity between variant T cell determinants from malaria circumsporozoite protein. *J Immunol* 1988; 141: 2456-60.

De Stricker K, Vuust J, Jepsen S, Oeuvray C, Theisen M. Conservation and heterogeneity of the glutamate-rich protein (GLURP) among field isolates and laboratory lines of *Plasmodium falciparum*. *Mol Biochem Parasitol* 2000; 111: 123-30.

Doolan DL, Saul AJ, Good MF. Geographically restricted heterogeneity of the *Plasmodium falciparum* circumsporozoite protein: relevance for vaccine development. *Infect Immun* 1992; 60: 675-82.

Escalante AA, Grebert HM, Isea R, *et al*. A study of genetic diversity in the gene encoding the circumsporozoite protein (CSP) of *Plasmodium falciparum* from different transmission areas - XVI. Asembo Bay Cohort Project. *Mol Biochem Parasitol* 2002; 125: 83-90.

Godson GN, Ellis J, Svec P, Schlesinger DH, Nussenzweig V. Identification and chemical synthesis of a tandemly repeated immunogenic region of *Plasmodium knowlesi* circumsporozoite protein. *Nature* 1983; 305: 29-33.

Good MF, Berzofsky JA, Miller LH. The T cell response to the malaria circumsporozoite protein: An immunological approach to vaccine development. *Ann Rev Immunol* 1988; 6: 663-88.

Good MF, Pombo DJ, Quakyi IA, *et al*. Human T-cell recognition of the circumsporozoite protein of *Plasmodium falciparum*: immunodominant T-cell domains map to the polymorphic regions of the molecule. *Proc Natl Acad Sci USA* 1988; 85: 1199-203.

Guttinger M, Caspers P, Takács B, *et al*. Human T cells recognize polymorphic and non-polymorphic regions of the *Plasmodium falciparum* circumsporozoite protein. *EMBO J* 1988; 7: 2555-8.

Hoffman SL, Oster CN, Mason C, *et al*. Human lymphocyte proliferative response to a sporozoite T cell epitope correlates with resistance to falciparum malaria. *J Immunol* 1989; 142: 1299-303.

Jongwutiwes S, Tanabe K, Hughes MK, Kanbara H, Hughes AL. Allelic variation in the circumsporozoite protein of *Plasmodium falciparum* from Thai field isolates. *Am J Trop Med Hyg* 1994; 51: 659-68.

Lockyer MJ, Marsh K, Newbold CI. Wild isolates of *Plasmodium falciparum* show extensive polymorphism in T cell epitopes of the circumsporozoite protein. *Mol Biochem Parasitol* 1989; 37: 275-80.

Moreno A, Clavijo P, Edelman R, *et al*. $CD4^+$ T cell clones obtained from *Plasmodium falciparum* sporozoite-immunized volunteers recognize polymorphic sequences of the circumsporozoite protein. *J Immunol* 1993; 151: 489-99.

Plebanski M, Lee EAM, Hill AVS. Immune evasion in malaria: altered peptide ligands of the circumsporozoite protein. *Parasitol* 1997; 115: S55-S66.

Qari SH, Goldman IF, Póvoa MM, Di Santi SM, Alpers MP, Lal AA. Polymorphism in the circumsporozoite protein of the human malaria parasite *Plasmodium vivax*. *Mol Biochem Parasitol* 1992; 55: 105-13.

Richie TL, Saul A. Progress and challenges for malaria vaccines. *Nature* 2002; 415: 694-701.

Shi Y-P, Alpers MP, Póvoa MM, Lal AA. Diversity in the immunodominant determinants of the circumsporozoite protein of *Plasmodium falciparum* parasites from malaria-endemic regions of Papua New Guinea and Brazil. *Am J Trop Med Hyg* 1992; 47: 844-51.

Stoute JA, Kester KE, Krzych U, *et al*. Long-term efficacy and immune responses following immunization with the RTS,S malaria vaccine. *J Infect Dis* 1998; 178: 1139-44.

Yoshida N, Di Santi SM, Dutra AP, Nussenzweig RS, Nussenzweig V, Enea V. *Plasmodium falciparum*: restricted polymorphism of T cell epitopes of the circumsporozoite protein in Brazil. *Exp Parasitol* 1990; 71: 386-92.

Yoshida N, Nussenzweig RS, Potocnjak P, Nussenzweig V, Aikawa M. Hybridoma produces protective antibodies directed against the sporozoite stage of malaria parasite. *Science* 1980; 207: 71-3.

Zavala F, Cochrane AH, Nardin EH, Nussenzweig RS, Nussenzweig V. Circumsporozoite proteins of malaria parasites contain a single immunodominant region with two or more identical epitopes. *J Exp Med* 1983; 157: 1947-57.

Zavala F, Masuda A, Graves PM, Nussenzweig V, Nussenzweig RS. Ubiquity of the repetitive epitope of the CS protein in different isolates of human malaria parasites. *J Immunol* 1985; 135: 2790-3.