

Circumsporozoite Protein of *Plasmodium vivax*: Gene Cloning and Characterization of the Immunodominant Epitope

Abstract. The gene encoding the circumsporozoite (CS) protein of the human malaria parasite *Plasmodium vivax* has been cloned. The deduced sequence of the protein consists of 373 amino acids with a central region of 19 tandem repeats of the nonapeptide Asp-Arg-Ala-Asp/Ala-Gly-Gln-Pro-Ala-Gly. A synthetic 18-amino acid peptide containing two tandem repeats binds to a monoclonal antibody directed to the CS protein of *Plasmodium vivax* and inhibits the interaction of this antibody with the native protein in sporozoite extracts. The portions of the CS gene that do not contain repeats are closely related to the corresponding regions of the CS genes of two simian malarias, *Plasmodium cynomolgi* and *Plasmodium knowlesi*. In contrast, the homology between the CS genes of *Plasmodium vivax* and *Plasmodium falciparum*, another malaria parasite of humans, is very limited.

DAVID E. ARNOT

JOHN W. BARNWELL

Department of Medical and Molecular Parasitology, New York University Medical Center, New York 10016

JAMES P. TAM

Rockefeller University, New York 10021

VICTOR NUSSENZWEIG

Department of Pathology, New York University Medical Center

RUTH S. NUSSENZWEIG

VINCENZO ENEA

Department of Medical and Molecular Parasitology, New York University Medical Center

Plasmodium vivax, the causative agent of benign tertian malaria, occurs throughout the world and is the dominant species of this parasite in China, Sri Lanka, and certain parts of Central America (1). Although causing much less mortality than *P. falciparum*, vivax malaria, with its characteristic relapsing fevers, is an incapacitating disease. We

report here the sequence of a gene we have cloned from *P. vivax*. This gene encodes the immunodominant surface antigen of the mosquito-borne infectious sporozoites, the circumsporozoite (CS) protein.

Protective immunity to malaria is conferred by immunization with irradiated sporozoites, and host antibody against sporozoites is directed against the CS protein (2-4). The gene encoding this protein has been cloned from two species of simian malaria parasites, *P. knowlesi* (5, 6, 7) and *P. cynomolgi* (8, 9) and from the human malaria parasite *P. falciparum* (10, 11). We have used a probe derived from the *P. cynomolgi* CS gene to isolate the homologous gene from a bacteriophage λ library of cloned *P. vivax* genomic DNA. The primary structure of the *P. vivax* CS protein is deduced here from the sequence of the cloned gene, and it has the same general features as have been reported for other CS proteins. In particular, it has a central domain of repetitive amino acid sequences constituting roughly half of the total protein. Sequence analysis reveals a close homology between the CS gene of *P. vivax* and that of the simian malaria *P. knowlesi* (6), but little homology to the CS gene of *P. falciparum* (11). These findings corroborate earlier hypotheses suggesting a close evolutionary relation between the simian malaria parasites and *P. vivax*, a parasite of humans (12).

Homology between a CS gene probe from *P. cynomolgi* (Gombak) (8) and genomic DNA from a bloodstream form of *P. vivax* was first detected by Southern blotting. As shown in Fig. 1, under medium stringency conditions, this *P. cynomolgi* CS gene probe hybridizes with single Acc I, Hpa II, and Bgl II fragments of *P. vivax* (Belem) DNA. Since the DNA fragment generated by Bgl II was of a suitable size (15 kilobases) for cloning in the bacteriophage λ EMBL vectors (13), a complete Bgl II digest of *P. vivax* bloodstream form

DNA was size-fractionated and ligated into the Bam HI sites of EMBL3 arms, and then packaged in vitro and plated (14).

Approximately 5000 plaques were screened with the *P. cynomolgi* complementary DNA (cDNA) clone being used as the probe, and three positive clones were identified. The restriction map of one of these clones, λ VX1B, is shown in Fig. 2. The structure of this clone is consistent with the sizes of the cross-hybridizing fragments obtained in the genomic Southern blots (Fig. 1). A series of plasmid subclones were therefore derived for further mapping and DNA sequence analysis (Fig. 2). The DNA sequence analysis was initially carried out from the Xba I and the Acc I sites which also occur in the *P. cynomolgi* CS gene. Homology with that gene was detected around these sites and a complete sequence analysis of *P. vivax* genomic DNA in the region shown in the fine

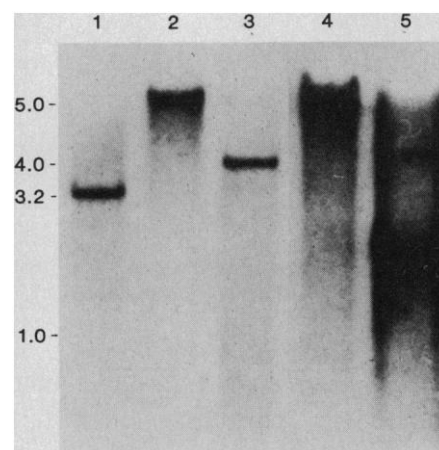


Fig. 1. Southern blot of genomic DNA from bloodstream forms of *P. vivax* (Belém strain). The DNA was bound to nitrocellulose and probed with a 700-bp Pst I fragment encoding the COOH-terminal domain and 3' untranslated region of a CS protein complementary DNA clone of *P. cynomolgi* (8, 9). The bloodstream parasites were obtained from an infected *Saimiri* monkey. Hybridization was carried out in 38 percent formamide, 6 \times standard saline citrate (SSC), 50 mM sodium phosphate buffer, pH 6.4, 10 percent dextran sulfate, and 100 μ g of salmon sperm DNA per milliliter, with the use of a nick-translated probe (3×10^6 count/min) (14). Lanes 1 to 4: 0.5 μ g of *P. vivax* DNA cut with Acc I, Sst I, Hpa II, and Bgl II, respectively. Lane 5: 0.5 μ g of *P. cynomolgi* (Gombak) DNA cut with Bgl II. After overnight hybridization of the DNA at 42°C, the filters were washed four times for 5 minutes each at room temperature in 3 \times SSC, 50 mM sodium phosphate buffer, pH 6.4, 0.1 percent sodium dodecyl sulfate, and then for 1 minute at 42°C in the same buffer. Autoradiography was carried out at -70°C with intensifying screens. Sst I does not cleave sufficiently close to the CS gene to release a resolved fragment on this 1 percent agarose gel.

Table 1. Inhibition by the synthetic peptide [Asp-Gly-Gln-Pro-Ala-Gly-Asp-Arg-Ala]₂ of the binding of ¹²⁵I-labeled monoclonal antibody 2F2 to immobilized *P. vivax* sporozoite extracts.

Concentration of synthetic peptide (μ g/ml)	Bound radioactivity* (count/min)
0	16,387 \pm 1,085†
200	842‡
20	1,091
2	2,130
0.2	3,146

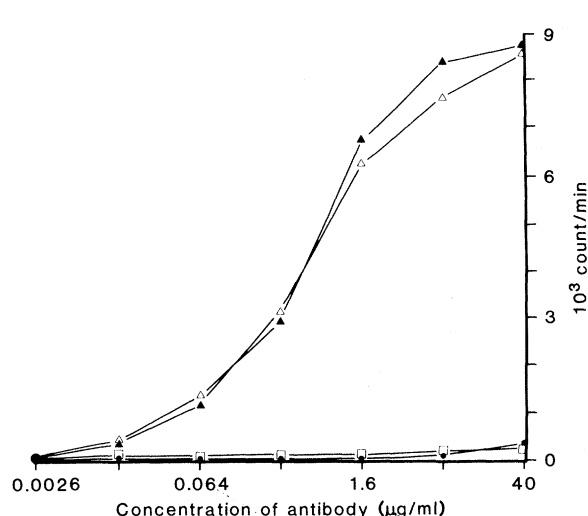
*The labeled 2F2 antibody (10⁵ count/min; specific activity about 2×10^6 cpm/ μ g) was diluted in phosphate-buffered saline (PBS) containing 1 percent bovine serum albumin (BSA) and the indicated amount of synthetic peptide and then added to microtiter wells coated with *P. vivax* sporozoite extracts. After 1 hour of incubation at room temperature, the wells were washed three times with PBS and 1 percent BSA and counted in a gamma counter. In preliminary experiments it was determined that the binding of the labeled 2F2 to the *P. vivax* extract was specific, since it was inhibited by diluting it with cold 2F2, but not with a cold nonrelevant monoclonal antibody of the same subclass.
†Mean \pm standard deviation of counts in four wells.
‡Mean of duplicates.

dental. Alternatively, the similarity may indicate that these sequences diverged from a common repeat by accumulation and spread of point mutations, insertions, and deletions. A similar proposal has been made to account for divergence in the repetitive amino acid sequences of the S-antigens of *P. falciparum* (16).

Sequence homology between the CS proteins of *P. vivax* and the other major human malaria *P. falciparum* is very low. The only region where there is any significant homology is a small region of the COOH-terminal domain (nucleotides 1086–1101) containing two cysteine residues previously noted to be conserved between *P. knowlesi* and *P. falciparum* (11, 18). However, outside their unrelated repeated sequences there is considerable homology between *P. vivax* and *P. knowlesi*. In the NH₂-terminal domain there is 69 percent amino acid and 72 percent nucleotide sequence homology; the COOH-terminal domain has 63 percent amino acid and 68 percent nucleotide sequence homology. In accordance with the Southern blotting results shown in Fig. 1, the CS protein of *P. cynomolgi* also has extensive sequence homology with *P. vivax* in the domains flanking the central repetitive region of the protein (9).

Antibodies to native CS proteins are largely, if not entirely, directed against the immunodominant repeated epitope (4, 19). To verify whether the repeat sequences deduced from the genomic clone λ VXIB contained this epitope, we synthesized an 18-amino acid peptide comprising the sequence (DGQPAGDRA)₂ (20) and compared its antigenic properties to those of sporozoite material. We found that the monoclonal antibody 2F2, which is specific for *P. vivax* sporozoites, bound to the synthetic peptide immobilized on wells of microtiter trays. Significant binding was observed at antibody concentrations below 1 μ g/ml. The binding was totally inhibited by the presence of homologous peptide (25 μ g/ml) in the fluid phase but not affected by an unrelated peptide (Fig. 4). The 18-amino acid peptide also inhibited, in a dose-dependent fashion, the interaction between the monoclonal antibody 2F2 and extracts of *P. vivax* sporozoites (Table 1). A strong inhibitory effect of the peptide could be observed at concentrations below 10⁻⁶M. Taken together these findings demonstrate that the immunodominant epitope is contained within the peptide.

For the experiments shown in Table 1 we used a synthetic 18-amino acid peptide that constitutes the CS epitope of a Brazilian isolate, a monoclonal antibody



incubation, the wells were washed with PBS-BSA and incubated with 50 μ l of ¹²⁵I-labeled affinity purified goat antiserum to mouse immunoglobulin G (10⁵ count/min; specific activity about 2 \times 10⁶ cpm/ μ g). After an additional hour of incubation the wells were washed with PBS-BSA and the radioactivity was counted. As shown, 2F2 but not 2A10 bound to the immobilized (DGQPAGDRA)₂ peptide. The binding of 2F2 was totally inhibited by (DGQPAGDRA)₂ but not affected by (NANP)₃. In other experiments (not shown) we found that the binding of 2A10 to the corresponding epitope (NANP)₃ was not inhibited by the synthetic peptide (DGQPAGDRA)₂. The points in the curves represent the mean of duplicate measurements. The variation between the replicates was less than 10 percent of the mean values.

prepared against sporozoites from Southeast Asia (3), and a sporozoite extract of the Chesson (New Guinea) strain of *P. vivax*. The results demonstrate that the repeated epitope must be present in all of these isolates and are in agreement with earlier data showing that the *P. vivax* CS epitope occurs in all of the many different isolates tested (21). Invariance of the repetitive epitope among different geographic isolates is also a characteristic of the CS protein of *P. falciparum* (21). This feature should facilitate the development of sporozoite vaccines based on the repeated epitope of CS proteins, produced either by genetic engineering or chemical synthesis.

The constancy of the CS protein epitopes in these human malaria parasites contrasts with the variability of the repetitive epitope sequences in different strains of *P. cynomolgi* (22) and *P. knowlesi* (7). A number of factors may be responsible for this paradox. The human parasites may have spread from their ancestral populations too recently compared to the geographically isolated and ecologically restricted simian parasites for CS epitope variation to have occurred. The dynamics of the host-parasite interaction, which differ in human and simian malaria, might affect the rate of generation and fixation of CS antigenic diversity. It is also possible that the genetic mechanisms governing the formation, maintenance, and evolution of the CS repeats may operate differently in different parasites. It should be possible to understand more about the evolution

Fig. 4. Monoclonal antibody 2F2 against *P. vivax* CS protein was serially diluted in phosphate-buffered saline-bovine serum albumin (PBS-BSA) (\blacktriangle); or in PBS-BSA containing 25 μ g of the synthetic peptide (DGQPAGDRA)₂ per milliliter (\bullet); or as a control in PBS-BSA containing 25 μ g of synthetic peptide (NANP)₃, per milliliter (\triangle). Another monoclonal antibody (2A10) to the *P. falciparum* CS protein (3) was serially diluted in PBS-BSA (\square). Portions (25 μ l) of the dilutions were delivered to the bottom of microtiter plates that had been coated with the synthetic peptide (DGQPAGDRA)₂, and saturated with 1 percent BSA. After 1 hour of

and function of antigenic diversity in sporozoites by further analysis of the CS genes of primate plasmodia.

References and Notes

1. P. C. C. Garnham, *Malaria Parasites and Other Haemosporidia* (Blackwell, Oxford, 1966).
2. R. S. Nussenzweig and V. Nussenzweig, *Philos. Trans. R. Soc. London Ser. B*, **307**, 117 (1984).
3. E. H. Nardin et al., *J. Exp. Med.*, **156**, 20 (1982).
4. F. Zavala et al., *ibid.*, **157**, 1947 (1983).
5. J. Ellis et al., *Nature (London)*, **302**, 536 (1983).
6. L. S. Ozaki et al., *Cell*, **34**, 815 (1983).
7. S. Sharma, P. Svec, G. H. Mitchell, G. N. Godson, *Science*, **229**, 779 (1985).
8. V. Enea et al., *Proc. Natl. Acad. Sci. U.S.A.*, **81**, 7520 (1984).
9. D. Arnot et al., in preparation.
10. V. Enea et al., *Science*, **225**, 628 (1984).
11. J. B. Dame et al., *ibid.*, p. 593.
12. L. J. Bruce-Chwatt, *Bull. WHO*, **32**, 363 (1965); G. W. Coatney, W. F. Collins, M. Warren, P. G. Contacos, *The Primate Malaria* (U.S. Department of Health, Education and Welfare, Washington, D.C., 1971); P. G. Sergiev and N. A. Tiburskaya, in *Progress in Parasitology* (International Congress Series) (1965), vol. 91, p. 165; T. F. McCutchan, J. B. Dame, L. H. Miller, J. Barnwell, *Science*, **225**, 808 (1984).
13. A. M. Frischauf et al., *J. Mol. Biol.*, **170**, 827 (1983).
14. These procedures are standard as described by T. Maniatis et al., *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1982).
15. R. L. Coppel et al., *Nature (London)*, **302**, 751 (1983); R. L. Coppel et al., *ibid.*, **310**, 789 (1984); R. F. Anders, G. V. Brown, A. Edwards, *Proc. Natl. Acad. Sci. U.S.A.*, **80**, 6652 (1983); M. Koenen et al., *Nature (London)*, **311**, 382 (1984); J. V. Ravetch et al., *ibid.*, **312**, 616 (1984); J. V. Ravetch, J. Kochan, M. Perkins, *Science*, **227**, 1593 (1985); H. D. Stahl et al., *Proc. Natl. Acad. Sci. U.S.A.*, **82**, 543 (1985).
16. A. F. Cowman et al., *Cell*, **40**, 775 (1985).
17. M. Galinski, D. Arnot, V. Enea, unpublished data.
18. U. Vergara et al., *Mol. Biochem. Parasitol.*, **14**, 283 (1985).
19. F. Zavala et al., *Science*, **228**, 1436 (1985).
20. An 18-residue peptide amide was synthesized with a structure [Asp-Gly-Gln-Pro-Ala-Gly-Asp-Arg-Ala]₂. The carboxyl terminus is an α -carboxamide and not a free carboxylic acid. The peptide was synthesized by the stepwise solid-phase method [R. B. Merrifield, *J. Am. Chem. Soc.*, **85**, 2149 (1963)] on a multidetachable benz-

hydramine resin (*p*-acyloxybenzhydramine-copolystyrene-1 percent divinylbenzene resin) [J. P. Tam, R. D. DiMarchi, R. B. Merrifield, *Tetrahedron Lett.* 2851 (1981)]. Boc-Ala-*p*-acyloxybenzhydramine resin (0.4 mmol per gram substitution of resin) was placed into the reaction vessel of a Beckman 990 synthesizer and a double-coupling protocol via dicyclohexylcarbodiimide was used to give a coupling efficiency of >99.85 percent completion per step. The benzyl-based side chain protecting groups and tertiary butoxycarbonyl (Boc) for the NH₂ α terminus were used. In the present synthesis the sequence Asp-Gly is prone to a cyclization reaction to form aspartimide during the synthesis and acid deprotection of the protected peptide-resin. To prevent base-catalyzed aspartimide formation during the synthesis and the strong-acid catalyzed deprotection step, a new protection group, Asp-(OCH₂Hex, aspartyl-β-cyclohexyl ester) was used [J. P. Tam *et al.*, *Tetrahedron Lett.* (1979), p. 4033]. The unpurified peptide, when examined in high-performance liquid chromatography (HPLC), gave a single symmetrical peak, accounting for >83 percent of all peptide content. The crude peptide was purified by preparative low-pressure liquid chromatography (100 to 120 pounds per square

inch) on a reversed-phase C-18 column (2.5 by 30 cm) using aqueous CF₃CO₂H (0.05 percent) and an acetonitrile gradient. The purified material gave a single symmetrical peak on analytical HPLC, and an amino acid analysis gave the correct theoretical values of amino acids. The overall yield based on the first alanine attached to the resin was 72 percent.

21. F. Zavala *et al.*, *J. Immunol.* 135, 2790 (1985).
22. A. H. Cochrane *et al.*, *Mol. Biochem. Parasitol.* 14, 111 (1985).
23. J. Viera and J. Messing, *Gene* 19, 259 (1982).
24. A. M. Maxam and W. Gilbert, *Proc. Natl. Acad. Sci. U.S.A.* 74, 560 (1977).
25. J. Messing and J. Viera, *Gene* 19, 169 (1982).
26. F. Sanger, S. Nicklen, A. R. Coulson, *Proc. Natl. Acad. Sci. U.S.A.* 74, 5436 (1977).
27. We thank D. Keeney and L. Caiati for technical assistance, W. E. Collins for providing *P. vivax* sporozoites, E. Nowoswiat and P. Lomedico for oligonucleotides, and B. Robles and C. Perry for manuscript preparation. We also acknowledge the Agency for International Development (No. DPE 0453-A-00-5012-00) and the McArthur Foundation for generous support of our research.

24 July 1985; accepted 19 September 1985

Novel Role for Phycoerythrin in a Marine Cyanobacterium, *Synechococcus* Strain DC2

Abstract. *Cyanobacterial picoplankton contribute substantially to oceanic primary productivity. The colored protein phycoerythrin is the major component of their light-harvesting apparatus. It was found that in Synechococcus strain DC2 a variable proportion of the light energy absorbed by phycoerythrin is lost as autofluorescence and therefore is not passed to a photoreaction center. Phycoerythrin may serve two functionally distinct roles in this organism: as a nitrogen reserve and as a collector of quanta for photosynthesis.*

M. WYMAN

Department of Biological Sciences,
University of Warwick,
Coventry CV4 7AL, Great Britain

R. P. F. GREGORY

Department of Biochemistry,
University of Manchester,
Manchester M13 9PL, Great Britain

N. G. CARR

Department of Biological Sciences,
University of Warwick

Cyanobacteria of the genus *Synechococcus* (diameter, <1.0 μm) make a substantial contribution to marine primary productivity in both tropical and temper-

ate waters (1, 2). Because the availability of nitrogen is often considered to limit productivity in the oceans (3), we investigated the growth of *Synechococcus* strain DC2 in light- or nitrogen (nitrate)-limited chemostat cultures.

Marine *Synechococcus* species isolated from the picoplankton community (4) of oceanic habitats owe their red color to the presence of the phycobiliprotein phycoerythrin (1, 5), a component of the photosynthetic light-harvesting apparatus in cyanobacteria and red algae (6). Several marine *Synechococcus* isolates, including strain DC2, synthesize phycoerythrins whose novel spectral proper-

ties (5, 7) bear closer resemblance to those found in red algae than those of freshwater or terrestrial cyanobacteria. Under optimal conditions, phycobiliproteins may account for up to 50 percent of the total protein in the cyanobacterial cell (8); however, light (9) and the availability of nitrogen (10) profoundly influence the phycobiliprotein content and composition of cyanobacteria. In all phycoerythrin-producing cyanobacteria examined thus far, light energy absorbed by this phycobiliprotein is transferred to reaction-center chlorophyll with high efficiency (90 to 95 percent) (6, 8).

One feature of marine *Synechococcus* species that prompted this study is the high degree of autofluorescence detected upon excitation of phycoerythrin (1, 5), a finding which suggests that, in these organisms, the phycoerythrin in the phycobilisome is not efficiently coupled to the photosynthetic apparatus. We propose that in *Synechococcus* strain DC2 the pigment not only collects quanta for photosynthesis but is accumulated when nitrogen is readily available, providing a dynamic pool of stored nitrogen. Under these conditions a substantial fraction of the light energy absorbed by the biliprotein is dissipated as autofluorescence and therefore not utilized in photosynthesis.

Table 1 shows the influence of increasing irradiance on the phycobiliprotein composition of *Synechococcus* strain DC2 and a typical phycoerythrin-rich cyanobacterium, *Oscillatoria rubescens*. In contrast to the response of the latter, increasing irradiance resulted in a progressive enrichment in the relative phycoerythrin content of *Synechococcus* strain DC2. At the highest irradiance (60 μE m⁻² sec⁻¹), the ratio of phycoerythrin to phycocyanin reached 21.6; considerably in excess of that reported in a wide range of other cyanobacteria (9) even under the most favorable conditions for phycoerythrin synthesis.

The efficiency with which light energy absorbed by the phycoerythrin of *Synechococcus* strain DC2 is utilized in photosynthesis was examined by two complementary approaches. In the first, glycerol was used to uncouple energy transfer between phycobiliproteins *in vivo* (11). The effect of increasing irradiance (and hence increasing rate of growth) on the relative light-harvesting efficiency of phycoerythrin was examined (Fig. 1). In nitrogen-sufficient cultures a greater proportion of absorbed light energy was lost in the form of phycoerythrin autofluorescence at high growth rates than in cells exposed to decreasing irradiance. In contrast, at all

Table 1. Influence of irradiance on the mean phycoerythrin (PE) and phycocyanin (PC) composition (± standard error, *n* = 3) of *Synechococcus* strain DC2 and *O. rubescens*.

Irradiance (μE m ⁻² sec ⁻¹)	PE (μg per unit biomass)	PC (μg per unit biomass)	PE:PC
<i>Synechococcus strain DC2 (low nitrogen)</i>			
60	16.3 ± 0.7	1.3 ± 0.1	12.5 ± 0.1
<i>Synechococcus strain DC2 (high nitrogen)</i>			
60	21.9 ± 1.9	1.0 ± 0.1	21.6 ± 0.5
48	22.4 ± 5.6	1.2 ± 0.1	18.2 ± 2.8
20	45.8 ± 2.2	2.8 ± 0.1	16.2 ± 0.4
6	36.0 ± 2.4	3.0 ± 0.5	12.0 ± 2.9
<i>Oscillatoria rubescens</i>			
45	20.1 ± 0.7	11.3 ± 1.5	1.9 ± 0.2
12	36.1 ± 5.1	13.8 ± 0.4	2.6 ± 0.3