The structure of *Plasmodium yoelii* merozoite surface protein 1\textsubscript{19}, antibody specificity and implications for malaria vaccine design

Rachel D. Curd\textsuperscript{1}, Berry Birdsall\textsuperscript{2}, Madhusudan Kadekoppala\textsuperscript{1}, Solabomi A. Ogun\textsuperscript{1}, Geoffrey Kelly\textsuperscript{3} and Anthony A. Holder\textsuperscript{1}

\textsuperscript{1}Divisions of Parasitology, \textsuperscript{2}Molecular Structure, and \textsuperscript{3}NMR Centre, MRC National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

1. Summary

Merozoite surface protein 1 (MSP1) has been identified as a target antigen for protective immune responses against asexual blood stage malaria, but effective vaccines based on MSP1 have not been developed so far. We have modified the sequence of *Plasmodium yoelii* MSP1\textsubscript{19} (the C-terminal region of the molecule) and examined the ability of the variant proteins to bind protective monoclonal antibodies and to induce protection by immunization. In parallel, we examined the structure of the protein and the consequences of the amino acid changes. Naturally occurring sequence polymorphisms reduced the binding of individual protective antibodies, indicating that they contribute to immune evasion, but immunization with these variant proteins still provided protective immunity. One variant that resulted in the localized distortion of a loop close to the N-terminus of MSP1\textsubscript{19} almost completely ablated protection by immunization, indicating the importance of this region of MSP1\textsubscript{19} as a target for protective immunity and in vaccine development.

2. Introduction

Malaria is a global disease with hundreds of millions of clinical infections and at least a million deaths annually [1]. One strand of the strategy to control and finally eradicate malaria is to develop cheap and effective vaccines [2], but progress has been slow in part, because we understand poorly what is required to induce protective immunity. Individuals that survive continuous exposure to infection do eventually develop clinical immunity, suggesting that a vaccine against the asexual blood stage of the parasite is achievable.

Polymorphism in antigens such as merozoite surface protein 1 (MSP1) is an important mechanism of the malaria parasite to evade a protective immune response [3]. Variable sequences result in antigenic differences important in acquisition of immunity, for example affecting antigen processing and presentation to T cells or as epitopes for antibodies that provide protection against infection. They are an important consideration in vaccine design because efficacy against different parasite populations should be as broad as possible. MSP1\textsubscript{19}, the C-terminal region of MSP1 that is composed of two epidermal growth factor (EGF)-like domains, has been implicated as a target of protective immune mechanisms in...
a large number of studies, using molecular-, in vitro-, in vivo- and population-based methods (reviewed in [4]). Rodent parasites in laboratory mice have been used to examine immune responses to MSP119, for example, immunization with recombinant Plasmodium yoelii MSP119 may provide protection against a homologous [5,6], but not heterologous [7] parasite challenge. In natural isolates of P. yoelii, MSP119 sequence polymorphism is extensive [8], whereas in Plasmodium falciparum, the most important human malaria parasite, the MSP119 sequence is relatively conserved [9].

Monoclonal antibodies (mAbs) specific for P. yoelii MSP119 that protect by passive immunization against the homologous line have been identified [10,11]. Two of these antibodies, B6 and F5, can bind to the first EGF-like domain alone, whereas a third, B10, requires both EGF domains for binding [11]. These antibodies bind to epitopes constrained by disulfide bonds [11], and to some but not all of the natural sequence variants [8]. However, the contribution of such sequence differences to immune evasion following infection is unclear.

Crystallographic and NMR-based approaches have been used to examine the three-dimensional structure of MSP119 from a number of Plasmodium species [12–15]. Some epitopes of P. falciparum MSP119-specific mAbs that do or do not inhibit erythrocyte invasion and parasite growth have been mapped using such structural methods or by introducing amino acid substitutions to identify the key target regions for functional antibodies [16–19]; this is particularly important because some so-called blocking antibodies compete with the binding of inhibitory antibodies, rendering them ineffective [20,21].

For vaccine development, attention has been focused on the larger 42 kDa fragment of MSP119 (MSP142), but the result of phase II clinical trials of these vaccines has been disappointing [22]: despite the induction of antibody, there was no obvious clinical benefit. Therefore, we need to re-evaluate the parameters that are important in the design of candidates for vaccine development.

The antigen may be ‘engineered’ to improve desirable characteristics or remove undesirable ones [21,23], and it is important to understand the structural basis and constraints of these approaches. Amino acid differences may ablate the binding of antibodies by affecting either the local or global structure of the protein, and affect the ability of the protein to induce protection following vaccination. However, the consequence of amino acid substitutions on the structure of the protein is not predictable; the effects may be localized if the residue is on the surface, or profound, particularly if the substitution is a radical change in the side chain. Such changes may also affect B-cell epitopes, and antigen processing and presentation to T cells [24]. There have been no experimental studies to explore the relationship between amino acid sequence, protein three-dimensional structure and antigenic polymorphism in protection against malaria.

To examine some of these aspects experimentally, we used the P. yoelii model. We produced recombinant MSP119 protein with single amino acid sequence differences either at positions where variation is found naturally [8] or within a conserved region. We examined the effects of these changes on the binding of protective mAbs and on the three-dimensional structure of the protein. Then, we examined the immunogenicity of the modified proteins and their ability to provide protection against parasite challenge.

### 3. Material and methods

#### 3.1. Ethical statement

All animal work protocols were reviewed and approved by the Ethical Review Panel of the MRC-NIMR, and approved and licensed by the UK Home Office as governed by law under the Animals (Scientific Procedures) Act 1986 (project licence no. 80/1832). The experimental procedures were designed to minimize the extent and duration of any harm, and included predefined clinical and parasitological endpoints to avoid unnecessary suffering.

#### 3.2. Preparation of wild-type and variant MSP119 proteins

Production in Escherichia coli and purification of a P. yoelii YM WT GST fusion protein (GST-MSP119) has been described previously [6]. Four variant proteins with single amino acid changes—R12L, K16E, N17H and E28K (residues numbered according to the N-terminus of MSP119)—were produced following site-directed mutagenesis. The wild-type (WT) and variant proteins with an N-terminal hexa-His tag were also produced in Pichia pastoris using a synthetic gene with the N-glycosylation site removed, and inserted into the pPICK9K vector (Invitrogen), with transformation and selection of expressing cells.

#### 3.3. Antigenic analysis of wild-type and variant GST-MSP119 proteins

Binding of the mAbs B6, B10 and F5 [11] to GST-MSP119 variants was analysed by Western blotting, ELISA and surface plasmon resonance (SPR). For Western blotting, the proteins were electrophoresed on NuPAGE Bis–Tris 12% polyacrylamide gels under non-reducing conditions and transferred onto nitrocellulose. These conditions were used because cystine reduction or reduction and alkylation either severely reduces or abolishes reactivity with the mAbs [11]. Following blotting with 5% (w/v) skimmed milk in PBS, blots were blocked with 5% (w/v) skimmed milk in PBS, blots were incubated with mAbs (B6, 2 μg ml<sup>−1</sup>; B10, 2 μg ml<sup>−1</sup>; F5, 10 μg ml<sup>−1</sup>), and binding was visualized using the ECL kit (Amersham). For ELISA analysis of mAb binding, 96-well plates (Nunc Maxisorp) were coated with anti-GST antibody followed by 1 μg ml<sup>−1</sup> of WT or variant GST-MSP119 and PBS- and GST-alone controls. Doubling dilutions of primary antibody were added, followed by goat anti-mouse IgG-HRP and detection with O-phenylenediamine dihydrochloride at 490 nm. Surface plasmon resonance measurements were made on a BIAcore 2000 instrument with anti-GST antibody bound to the surface of a CMS sensor chip and GST-MSP119 proteins bound at 10 μg ml<sup>−1</sup>. mAb binding assays were performed at a constant flow rate of 5 μl min<sup>−1</sup> at 25°C, using B6 (1.29 mg ml<sup>−1</sup>), B10 (1.18 mg ml<sup>−1</sup>) and F5 (1.18 mg ml<sup>−1</sup>), recording the steady-state binding (within 2 min) in triplicate repeated binding assays.

#### 3.4. Immunization studies with wild-type and variant GST-MSP119 proteins

Groups of six eight-week-old BALB/c mice were immunized intraperitoneally (i.p.) with 10 μg protein in Freund’s
complete adjuvant, and boosted by two further injections of 40 μg of protein in Freund’s incomplete adjuvant 21 and 42 days later. Serum samples were taken and pooled for each group 14 days after the final immunization, and the titre of MSP1₁₉-specific antibodies was measured by ELISA, each group 14 days after the final immunization, and the 42 days later. Serum samples were taken and pooled for

4.1. The effects of single amino acid substitutions on mAb binding to MSP1₁₉

We produced recombinant proteins containing either single natural polymorphisms (R12L, K16E and N17H) or a substitution of a conserved residue, changing the charge (E28K; figure 1a). In order to test whether the amino acids changed in the MSP1₁₉ variants affected the binding of the protective B6, F5 and B10 mAbs [11], Western blotting, ELISA and SPR analyses were carried out using the WT and variant proteins (figure 1). For the F5 antibody, K16E and E28K abolished, R12L reduced and N17H had little effect on binding in all assays. For B6, K16E largely abolished, N17H reduced and R12L had no effect on binding. For the antibody B10, R12L and E28K reduced binding in the ELISA and SPR assays, but the differences were less pronounced on the Western blot. Overall, each variant gave a different pattern of reactivity, but no single amino acid change resulted in the loss of binding of all three mAbs, indicating that their contribution to the evasion of antibody-mediated protection is likely to be cumulative.

4.2. The effects of amino acid substitutions on the structure of MSP1₁₉

To understand the basis of the reduced binding of the mAbs, the structural consequences of these amino acid substitutions

3.5. NMR assignments and structure calculations

NMR spectra were recorded at 25 °C or 35 °C on Varian Inova (600 and 800 MHz) and Bruker Avance (600 and 700 MHz) spectrometers equipped with triple resonance z-gradient probes. All spectrometers except for the Varian 600 MHz were equipped with cryogenic probes. Water suppression was carried out using the WATERgate sequence [25]. ¹H-[¹⁵N]-HSQC spectra were acquired for ¹⁵N-labelled WT MSP1₁₉ and all variants. For the WT and the E28K variant, three-dimensional NMR experiments (HNCO, HNCACB and CBCA(CO)NH) were used for making the sequential assignments. Aliphatic and aromatic side chain assignments were derived from three-dimensional HCC-TOCSY, HCCCONH and ¹³C-edited NOESY-HSQC spectra recorded separately for the aliphatic and aromatic regions. For the WT MSP1₁₉, chemical shift assignments were obtained for more than 97% of the ¹H, ¹³C and ¹⁵N atoms of the protein backbone, and for more than 78% of the side chain atoms; similar levels were achieved for the E28K MSP1₁₉. The first 10 residues (His tag and linker region) are not included. All spectra were processed with NMRPipe software, visualized with NMRDRAW [26] and analysed using XEASY or Sparky software (T. D. Goddard and D. F. Kneller, Sparky 3, UCSF, USA).

Amide proton exchange rates were measured in a series of SOFAST ¹H-[¹⁵N]-HMQC spectra recorded every 5 min in a pseudo-three-dimensional experiment. To evaluate quantitatively the extent of protection of the amide protons, the H/D exchange rates were measured from the time course of the NH peak intensities over 2–3 h.

Distance restraints for calculating the families of structures for WT and E28K proteins were obtained from nuclear Overhauser effects (NOEs) in four spectra: an ¹⁵N-edited NOESY-HSQC, separate ¹³C-edited NOESY-HSQC spectra for the aliphatic and aromatic regions, and a two-dimensional ¹H NOESY acquired for a sample in 99% ²H₂O; all NOESY experiments used a mixing time of 100 ms. Peaks from the NOESY spectra were picked manually using Sparky. The spectra were transferred to XEASY and the peak volumes determined by integration [27]. The NOE peak lists for the four spectra together with the chemical shift lists for all assigned nuclei and a list of dihedral angle restraints from TALOS [28] were submitted to ARIA 1.2 and 2.2 software packages [28–32] for structure calculation using default settings. NOEs were assigned via an iterative procedure using ARIA, and violated restraints were checked manually in Sparky. Potential hydrogen bonds in preliminary structures were calculated using MOLMOL [33]. H bonds calculated to be in eight or more of the 10 structures, and confirmed by the persistence of an NH crosspeak in the ²H₂O exchanged ¹H-[¹⁵N]-HSQC spectrum, were included in subsequent ARIA calculations. The structures were also examined to identify the disulphide bonds [34], which were added to the ARIA structure calculations initially as restraints and then as covalent bonds. For the final ARIA runs in iteration 8, the number of structures was increased to 100, and the 20 best energy structures were used in the water refinement. The quality of the final structures was assessed using PROCHECK NMR [35], and MOLMOL was used to calculate the surface electrostatic potential. Structure visualization, analysis and comparisons were carried out using InsightII (Accelrys Software Inc.).

The total number of NOE-derived distance restraints used in the final calculations was 3391 for the WT and 2927 for the E28K variant. A breakdown of the NOEs is given in the electronic supplementary material, table S1. The ¹H, ¹³C and ¹⁵N chemical shifts have been deposited in the BioMagResbank database (http://www.bmrb.wisc.edu) under the accession numbers 19232 and 19234. The structural data and experimental restraints used in calculations have been submitted to the Protein Data Bank with PDB ID codes 2mgp and 2mgr.
were monitored using NMR spectroscopy. Comparison of the $^1$H/$^{15}$N-HSQC spectra of the variants with the WT MSP1$_{19}$ spectrum (see electronic supplementary material, figure S1) and quantification of chemical shift changes identified regions where changes in the structure had occurred. For the R12L, K16E and N17H variants, the $^1$H/$^{15}$N-HSQC spectra were very similar to the WT spectrum (comparison of WT and R12L shown in figure 2a; comparison of WT with K16E and with N17H shown in the electronic supplementary material, figure S2). Figure 2b shows histograms of the weighted $^1$H/$^{15}$N chemical shift difference for each residue of these three variants and WT MSP1$_{19}$. The chemical shift changes were small, associated with residues at or near the substitution, and the spectra of the variants could be assigned by comparison with the WT spectrum (see below). However, for the E28K variant, the two spectra were very different and a large number of NH peaks had shifted (figure 2c), necessitating a full, independent assignment for this variant. Comparison of the E28K variant with the WT spectrum is shown in figure 2d.

Figure 1. Single amino acid changes reduce the binding of protective monoclonal antibodies to *P. yoelii* MSP1$_{19}$. (a) The location and nature of the four amino acid substitutions examined in this study. Antibody binding was examined by (b) Western blotting, (c) ELISA and (d) surface plasmon resonance. The mAbs B6, B10 and F5 were used to probe either GST or GST-MSP1$_{19}$ with the WT sequence or the four variants R12L, E28K, N17H and K16E. (b) Anti-GST antiserum was used as a control and the location of GST (25 kDa) and the GST-MSP1$_{19}$ fusion proteins (37 kDa) is indicated. (c) ELISA plates coated with each antigen were probed with increasing dilutions of the individual mAbs. (d) Chips were coated with the different proteins and probed with the individual antibodies.
and WT chemical shifts (figure 2d) showed that 18 of the 48 residues in the first EGF domain had amide shifts of greater than 0.1 ppm, and two of them were greater than 1.0 ppm. By contrast, the amides of only the three terminal residues of the second EGF domain shifted by more than 0.1 ppm. Clearly, the E28K variant is having a larger effect on the first rather than on the second EGF domain; interestingly, the largest shift changes were for residues 9–14 and not at the site of the changed residue, position 28.

The direction and magnitude of the chemical shift changes seen for residues 9 and 10 suggested that H bond interactions may have been lost in the E28K variant. 2H2O exchange experiments to determine the rates of exchange of individual amide residues were performed (see electronic supplementary material, figure S3); rates for residues in the second domain were almost identical in the WT and E28K proteins (although two residues could not be quantified owing to overlap). By contrast, 14 residues in the first domain exhibited reduced protection from exchange in the variant, and residues 7–10 protected in the WT protein showed no protection in the variant. These data, together with the chemical shift changes, support the view that residues 9 and 10 are not H-bonded in E28K.

4.3. Comparison of the three-dimensional structure of wild-type and E28K MSP119

In the light of the substantial chemical shift differences between WT and E28K MSP119, complete structure determinations of both were undertaken to allow full comparison. The numbers of NOE distance restraints used in the final iteration of the ARIA structure calculations are shown in the electronic supplementary material, table S1; the Ramachandran plot quality is typical of that found for other EGF structures [36]. Electronic supplementary material, figure S4 shows the backbone traces of the 20 lowest-energy structures for WT and E28K MSP119; the traces for both families showed good convergence especially for the interior regions, although the ends of some loops (23–26, 40–42 and 73–75) exhibit more variability. Figure 3a shows the lowest-energy structures of the two proteins superimposed using residues 8–85 (the less well-defined end residues are not useful for super-positioning) to identify differences between them. The major difference between the two proteins is the conformation of the residue 9–14 loop. In the orientation shown,
the WT loop appears in the plane of the paper, whereas, in
the E28K variant, it appears perpendicular to the page.
A manual analysis of the NOE connectivities observed for
this loop region indicated a subtle change in the local pack-
ing. Changes in NOE patterns include the observation of
strong NOEs from T11 to W30 in the WT protein. These are
absent in the E28K variant, which instead exhibits strong con-
nectivities between T11 and C20/F21. The backbone amide
proton exchange data (see electronic supplementary material,
figure S3) also support a change in packing interactions.

There are also some differences in the conformation of the
last residues, 89–99, where the two EGF domains are in con-
tact. RMSD values calculated for the superimposition
indicate that the structures are broadly similar, with the
main changes in the first domain (figure 3a). Antibody binding
to this protein will be sensitive to changes in the conformation of the side chains of the surface
residues. Although surface residue side chains are not as well
deﬁned as buried residues in these families of structures, the
surface side chains do not have complete freedom and are
concentrated in a narrow region of conformational space
(figure 3b). Therefore, it is possible to say that the side
chains of certain surface residues (for example the charged
residues, D10, R12 and D13) have changed their position in
response to the change at residue 28. In the WT protein, the
side chain of E28 is less than 5 Å from the D10 side chain,

Figure 3. The structure of WT and E28K MSP1
(a) Superimposition of WT and E28K MSP1 in a ribbon representation; the first EGF domain (residues 1–48) is
coloured red and blue, and the second EGF domain (residues 49–99) is coloured orange and cyan, respectively. (b) Ribbon representation with charged residues that
have altered position highlighted, and (c) electrostatic surface potential with selected charged residues indicated. (b,c) Four views of the molecules with a 90°
rotation between each view are presented; Arg and Lys residues are depicted in blue and Asp and Glu residues are depicted in red. Data that support these structures
include D2O exchange rates (see electronic supplementary material, figure S3); the 20 lowest-energy structures for each protein are presented in the electronic
supplementary material, figure S4.
but very far from the R12 and D13 side chains. However, in the variant, the K28 side chain is near to the side chain of both D10 and D13. R12 and D13 appear to have swapped places, with D13 moving towards K28 and R12 moving away. The electrostatic surface potential for the WT and E28K proteins was mapped onto the two lowest-energy structures using MOLMOL, revealing two distinct clusters of positively and negatively charged residues on the two proteins (figure 3c).

4.4. Effect of single amino acid substitutions on the ability of MSP119 to protect against parasite challenge in vivo

Immunization studies were designed to look at the effect of the sequence changes on the ability of MSP119 to provide protection against challenge infection with the lethal P. yoelii YM parasite by vaccination. Following immunization and prior to challenge, a serum sample was taken from each mouse and pooled for each group, and the antibody titre against WT MSP119 was measured (figure 4a). Mice immunized with each of the variants had produced antibodies binding to MSP119, and overall there was little difference in the antibody titres, indicating that the structural differences between the variants had no major effect on the induction of polyclonal antibodies reactive with the WT protein. Following challenge, the parasitaemia was followed daily from day 3 on Giemsa-stained blood films (figure 4b). None of the mice immunized with GST or E28K was protected. In the GST-alone group, parasites were patent on day 4, and by day 6 all the animals had a high parasitaemia and were sacrificed. In the E28K group, parasitaemia rose rapidly to day 7 and all mice had to be sacrificed due to poor clinical condition or high parasitaemia. By contrast, the overall patterns of parasitaemia for the WT, R12L, K16E and N17H groups were very similar, with low parasitaemia and final resolution of the infection, although some mice died throughout the experiment. A significant difference in peak parasitaemia was observed between WT and E28K groups (ANOVA, \( p = 0.003 \)), and WT and GST groups (ANOVA, \( p < 0.0001 \)). The geometrical means of parasitaemia during the course of infection also differed significantly for E28K (Wilcoxon matched-pairs signed-rank test, two-tailed \( p < 0.005 \)) and for K16E (two-tailed \( p < 0.0001 \)) groups from WT. Although the K16E group mean parasitaemia was higher than that of the WT group, the mice resolved parasitaemia, revealing significant protection from infection. The Kaplan–Meier estimates of survival probability are shown in the electronic supplementary material, figure S5. In comparison with the mice that received WT MSP119, the mice immunized with the E28K variant showed decreased survival rates following parasite challenge (\( p = 2.48 \times 10^{-5} \)). On the other hand, the mice that received the R12L, K16E and N17H variants did not exhibit a significant difference in survival rate when compared with the WT protein-immunized mice, with \( p \)-values of 0.634, 0.998 and 0.317, respectively. Overall, the results indicate that proteins containing individual natural variants still induced protective antibodies. The variant with the greatest effect on protein structure still induced antibodies reactive with the WT protein, but these antibodies were not protective.

5. Discussion

MSP1 remains a candidate for inclusion in a malaria vaccine. Much attention has been focused on the C-terminal region of
this molecule, particularly MSP1_{19}, the last approximately 100 amino acids preceding the GPI addition site [4]. This region is the target of antibodies that inhibit parasite growth by a variety of mechanisms [37]. The fine specificity of the antibody is crucial, and the binding sites of antibodies with different properties have been defined using a variety of techniques (reviewed in [4]). For example, amino acid substitutions in *P. falciparum* MSP1_{19} have been used to map the binding of ‘inhibitory’, ‘neutral’ and ‘blocking’ antibodies [4,19]. Few studies have been carried out to determine the protective capacity of antibodies induced by immunization with such antigens [38–40]. Passive immunization with *P. yoelii* MSP1_{19}-mAbs can suppress a blood stage infection [10–11], and these antibodies react with some, but not all recombinant MSP1_{19} of other *P. yoelii* lines, suggesting that the natural polymorphisms they contain represent antigenic diversity important in strain-specific immunity [8]. Immunization with MSP1_{19} provides protection against homologous blood stage challenge. MSP1_{19} sequence is conserved across species, but a number of amino acid substitutions are either species-specific or polymorphic within a single species [8,12].

Here, we have made several variant MSP1_{19} proteins, and examined the consequences of these changes on the structure of the protein, its ability to bind to different antibodies and its ability to induce protective immune responses to infection. We produced four variant proteins. Two (K16E and N17H) are natural variants in *P. yoelii*, and both these changes correlate with lack of binding of the mAbs B6 and F5. Across *Plasmodium* spp. K16 is not conserved, but residue 17 is almost always Asn. A third variant (R12L) is at a residue that is largely conserved within *P. yoelii* but may differ in different *Plasmodium* species [8]. These changes produced only very local structural perturbations; nevertheless, they abolished binding of one or more specific mAbs. The modified proteins induced antibodies that react with the WT protein, and immunized animals were partially protected against parasite challenge by a mechanism that is probably antibody-mediated [6]. This suggests that these epitopes, although the target of protective mAbs, are not individually sufficient as targets for protective immunity, and therefore it is likely that immune selection has driven cumulative polymorphisms in which incremental structural changes have occurred to evade protective antibody.

The final variant (E28K) is at a residue invariant across all species of malaria parasite. This Glu residue sits on the surface of the molecule, but changing it had profound consequences for a small region of the three-dimensional structure of the protein. Substitution of Glu by the positively charged Lys had profound effects on its ability to provide protection in immunization experiments. When used to immunize mice no protection was provided against parasite challenge, even though antibodies were produced that reacted at a similar but somewhat reduced level with the WT protein, and consistent with the perturbation being restricted to a small part of the molecule. The induction of these antibodies rules out the loss of protection due to this protein being degraded more rapidly following immunization.

To understand the extent of the perturbation, we solved the solution structure of both the WT and E28K variant proteins. The substitution had resulted in a reorganization of the charged residues within a region of the first EGF domain to accommodate the change, and other regions of the molecule remained unperturbed. These results identify a small region of the first EGF domain of MSP1_{19} for which the structural integrity is very important in the induction of a protective immune response. This region is highlighted in the sequence alignments and three-dimensional structure comparisons for MSP1_{19} from a number of *Plasmodium* species shown in the electronic supplementary material, figure S6.

It is unclear whether the E28K substitution in the parasite MSP1 would reduce viability. To test this possibility, the specific change could be introduced into the parasite by transfection and the effect on growth assessed. However, it seems unlikely that the change would be detrimental, because the entire MSP1_{19} sequence can be replaced by the two EGF domains of MSP8 [41].

This study highlights the importance of relating the structural consequences of changes to an antigen that might result from ‘engineering’ to the functional outcome. ‘Antigen engineering’ seeks to improve immunogenicity by focusing the immune response on the ‘right’ epitopes, a process that has been called ‘immunofocusing’ [23]. In the case of MSP1_{19}, different antibody specificities have been defined, including a class that blocks the binding of neutralizing antibody [21]. Changes in primary sequence may also result in changes in antigen processing and presentation to T cells by dendritic cells or the presence of T-cell epitopes, both of which may either suppress or enhance the immunogenicity of the protein [24,42]. However, the structural consequence of engineered amino acid changes is poorly understood. With respect to vaccines for human malaria that contain MSP1, some progress has been made in the identification of amino acid substitutions that can be made with a view to improving the immunogenicity of the protein. However, the current inability to easily carry out *in vivo* challenge studies with *P. falciparum* restricts progress. ‘T-celiparuzmized’ rodent parasites in which the MSP1 EGF domains have been replaced by the equivalent region from other malaria species [43,44] have been used [45], but parasites that contain at least the full-length *P. falciparum* MSP1_{19} [46] and longer chimeric or full-length MSP1 are required to be able to easily test the outcome of modifications to this molecule.

We conclude that individual amino acid changes that occur naturally in MSP1_{19} can abolish the binding of individual protective mAbs, an observation made previously with other malarial antigens, such as apical membrane antigen 1 [47]. However, such amino acid changes may have a small localized effect on the structure of the protein and do not substantially reduce the protection provided by immunization with the protein. On the other hand, at least one amino acid substitution did not affect the binding of all of the protective mAbs, but had a profound effect on the structure of part of the first EGF domain. Although the protein-induced antibodies reacted with the WT protein, they were unable to provide protection against parasite challenge. This result identifies an important structural target for the development of protective antibodies, and with a fuller understanding of the importance of structural differences for immunogenicity and protection it should be possible to design new and more effective antigens for inclusion in malaria vaccines. For example, antigens that induce a strong antibody response to the first loop of the first EGF domain would be of particular interest. Furthermore, in a complementary approach, the design of antigens to quantify the level of antibody response to this region of the molecule is possible and could facilitate studies to obtain *in vitro* surrogate markers of protection.
Acknowledgements. This work was supported by the UK Medical Research Council (U117532087, U117584256 and U117533887), and the European Union through the BioMalPar Network of Excellence. R.D.C. held an MRC Research Studentship.

References

30. Dayringer HE, Tramontano A, Sprang SR, Fletterick RJ. 1986 Interactive program for visualization and


