Antibodies to Intercellular Adhesion Molecule 1-Binding Plasmodium falciparum Erythrocyte Membrane Protein 1-DBL Are Biomarkers of Protective Immunity to Malaria in a Cohort of Young Children from Papua New Guinea

Tessema, Sofonias K; Utama, Digjaya; Chesnokov, Olga; Hodder, Anthony N; Lin, Clara S; Harrison, G. L. Abby; Jespersen, Jakob S; Petersen, Bent; Tavul, Livingstone; Siba, Peter; Kwiatkowski, Dominic; Lavstsen, Thomas; Hansen, Diana S; Oleinikov, Andrew V; Mueller, Ivo; Barry, Alyssa E

Published in:
Infection and Immunity

Link to article, DOI:
10.1128/IAI.00485-17

Publication date:
2018

Document Version
Peer reviewed version

Link back to DTU Orbit

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Antibodies to ICAM1-binding PfEMP1-DBLβ are biomarkers of protective immunity to malaria in a cohort of young children from Papua New Guinea

Sofonias K Tessema1,2, Digiaya Utama1,2, Olga Chesnokov3, Anthony N Hodder1,2, Clara S Lin1,2, G.L. Abby Harrison1,2, Jakob S Jespersen4, Bent Petersen5,6, Livingstone Tavul7, Peter Siba7, Dominic Kwiatkowski8,9, Thomas Lavstsen4, Diana S Hansen1,2, Andrew V Oleinikov3, Ivo Mueller1,2,10 and Alyssa E Barry1,2,11

The Walter and Eliza Hall Institute of Medical Research, Victoria, Australia1; The University of Melbourne, Department of Medical Biology, Victoria, Australia2; Charles E. Schmidt College of Medicine, Florida Atlantic University, Boca Raton, Florida, USA3; Centre for Medical Parasitology, Department of Immunology and Microbiology, University of Copenhagen and Department of Infectious Diseases, Copenhagen University Hospital, Denmark4; Center for Biological Sequence Analysis, Technical University of Denmark, Kgs. Lyngby, Denmark5; Centre of Excellence for Omics-Driven Computational Biodiscovery (COMBio), Faculty of Applied Sciences, AIMST University, Kedah, Malaysia; ‡Vector Borne Diseases Unit, Papua New Guinea Institute of Medical Research, Goroka, Papua New Guinea; †Wellcome Trust Sanger Institute, Hinxton, United Kingdom; ‡MRC Centre for Genomics and Global Health, University of Oxford, Oxford, United Kingdom; Institut Pasteur, Paris, France10

Running Head: Naturally acquired immunity to ICAM1-binding PfEMP1-DBLβ

#Address correspondence to Alyssa E Barry, barry@wehi.edu.au
Plasmodium falciparum Erythrocyte Membrane Protein 1 (PfEMP1) mediates parasite sequestration to the cerebral microvasculature via binding of DBLβ domains to Intercellular Adhesion Molecule 1 (ICAM1) and is associated with severe cerebral malaria. In a cohort of 187 young children from Papua New Guinea (PNG), we examined baseline antibody levels to the ICAM1-binding PfEMP1 domain, DBLβ3PF11_0521, in comparison to four control antigens including NTS-DBLα and CIDR1 domains from another group A variant and a group B/C variant. Antibody levels for the group A antigens were strongly associated with age and exposure. Antibody responses to DBLβ3PF11_0521 were associated with a 37% reduced risk of high-density clinical malaria in the follow up period (adjusted incidence risk ratio, aIRR = 0.63 [95% CI: 0.45-0.88; p = 0.007]) and a 25% reduction in risk of low-density clinical malaria (aIRR = 0.75 [95% CI: 0.55-1.01; p = 0.06]), whilst there was no such association for other variants. Children who experienced severe malaria also had significantly lower antibody levels to DBLβ3PF11_0521 and the other group A domains than other children. Furthermore, a subset of PNG DBLβ sequences had ICAM1-binding motifs, formed a distinct phylogenetic cluster and were similar to sequences from other endemic areas. PfEMP1 variants associated with these DBLβ were enriched for DC4 and DC13 head-structures implicated in EPCR-binding and severe malaria, suggesting conservation of dual binding specificity. These results provide further support for the development of specific classes of PfEMP1 as vaccine candidates, and as biomarkers for protective immunity against clinical P. falciparum malaria.

Key words. Malaria; var genes; PfEMP1; immunity; DBLβ; ICAM1
BACKGROUND

Malaria due to infection with *Plasmodium falciparum* remains a major global public health issue, with more than 400,000 deaths and 215 million symptomatic episodes each year (1). Children with limited prior exposure to malaria bear the majority of the disease burden, however naturally acquired immunity eventually develops with age and exposure, and is associated with the acquisition of a diverse repertoire of antibodies to parasite-encoded variant antigens on the infected erythrocyte surface (2). The major target of this immunity is *P. falciparum* Erythrocyte Membrane Protein 1 (PfEMP1) (3, 4), which is differentially encoded by up to 60 highly polymorphic var genes per parasite genome (5-7). Expression of diverse PfEMP1/var gene variants allows clonal antigenic variation (8, 9) and cytoadhesion to a wide variety of host molecules including chondroitin sulphate A (10), CD36 (11), endothelial protein C receptor (EPCR) (12) and Intercellular Adhesion Molecule 1 (ICAM1) (13). Adhesion occurs via specialized PfEMP1 domains, known as Duffy Binding Like (DBL) and Cysteine-rich Interdomain Region (CIDR) (7). Whilst antibodies to PfEMP1 in general have been shown to be important mediators of protection against symptomatic malaria, the specific PfEMP1 variants targeted by protective immune responses are poorly understood.

Var genes have been classified into three major groups (A, B, and C) based on chromosome orientation and conserved structural and sequence features, and in addition there exists a group of chimeric genes (B/A), also known as domain cassette (DC) 8 (14). Group A and B/A var genes are expressed in parasites isolated from children with severe disease and are up-regulated in cytoadherent parasites linked to pathogenesis (reviewed by (15, 16)). PF11_0521 and PFD1235w are group A var genes that contain ICAM1-binding DBLβ domains (17-19). Infected erythrocytes co-localize with ICAM1 expression in the brain blood vessels suggesting that ICAM1 mediates parasite sequestration in cerebral malaria (20). These genes also belong to the subclass of group A PfEMP1 variants that have adjacent CIDR domains that bind EPCR (21), another important host-parasite interaction implicated in severe malaria (12). Dual binding to these host receptors has been linked to cerebral malaria (21). To our knowledge, only one study, conducted in Tanzanian children, has found an association between high levels of antibodies against the ICAM1-binding DBLβ domain of PF11_0521 (DBLβ3PF11_0521) and a reduced risk of severe malaria (22). The role of antibodies against ICAM1-binding DBLβ in protection against
clinical and severe malaria is thus not well understood, and has not been studied outside of sub-
Saharan Africa.

Considering the diverse functional roles of different PfEMP1 variants and domains, protective immu-
unity would be expected to vary considerably among different PfEMP1 subgroups and cytoadherent do-
mas. Therefore, this study aimed to investigate whether antibodies against DBL\textsubscript{3PF11\_0521} are associ-
ated with protection against clinical and severe malaria in comparison to domains from other PfEMP1 vari-
ants not associated with ICAM1-binding or severe malaria (23, 24). The study was conducted in a lon-
gitudinal cohort of very young (aged 1-3 years) children from Papua New Guinea (PNG) that are ac-
tively acquiring immunity to malaria (25), and to minimize the background of diverse PfEMP1 antibo-
dies that are acquired with high malaria exposure (25-27). Plasma antibody levels were measured at ba-
line and associated with prospective risk of uncomplicated (clinical) and severe malaria. To explore the PfEMP1 landscape of PNG, we also investigated the presence of ICAM1-binding motifs and the associated domain architecture of \textit{var} genes among 125 \textit{P. falciparum} isolates from three distinct geographic areas. The results support a role for PfEMP1 variants containing ICAM1-binding DBL\textsubscript{3} as targets for protective antimalarial immunity.
RESULTS

Group A PfEMP1 domains are serodominant among young PNG children

Baseline plasma samples from a longitudinal cohort of 187 1-3 year old children from the Maprik area of East Sepik Province, Papua New Guinea followed for 16 months (25) were screened for antibodies (IgG) to five PfEMP1 domains including the ICAM1-binding DBLβ3_PF11_0521, and four control antigens. These included two domains from a group A variant PF13_0003 which has a DBLβ3 domain but does not have an ICAM1 binding motif (see below): NTS-DBLα1.6_PF13_0003 and CIDRδ_PF13_0003, and two domains from a group B/C variant PFL1955w: NTS-DBLα0.16_PFL1955w and CIDRα3.4_PFL1955w. Amongst the children, seroprevalence was 2.5 to 4-fold higher for the three type A PfEMP1 domains (40.1% for DBLβ3_PF11_0521, 27.8% for CIDRδ_PF13_0003 and 24.1% for NTS-DBLα1.6_PF13_0003) than the type B/C domains (11.2% for NTS-DBLα0.16_PFL1955w and 10.1% for CIDRα3.4_PFL1955w domains) (Figure 1A). Pairwise comparisons of antibody responses to the five domains showed that seropositivity to DBLβ3_PF11_0521 is significantly higher than all the other domains. In addition, seropositivity to group A domains is significantly higher than the group B domains (Table S1). Similarly, a combined analysis of seropositivity to any one of the group A PfEMP1 (55%) was significantly higher than the seropositivity to any one of the group B/C domains (18.7%) after correcting for multiple comparisons (p = 0.003, Bonferroni adjusted pairwise t-test).

There was a low but significant correlation between antibody responses to DBLβ3_PF11_0521 and NTS-DBLα1.6_PF13_0003 (Spearman’s rho (rs) = 0.36, p < 0.0001). This was also the case for the two type A PF13_0003 domains (rs = 0.36, p < 0.0001). In contrast, there was no significant correlation between the two group B PFL1955w domains (rs = 0.11, p = 0.135), nor between domains from the different var gene subgroups (Figure 1B). These low but significant correlations among the group A domains are explained by the predominant expression of and exposure to type A PfEMP1 variants in early childhood infections (23, 24, 28, 29).
Antibodies to five PfEMP1 domains are differentially associated with age and infection status.

To investigate whether past and current exposure to malaria influence responses to the five PfEMP1 domains, we investigated associations of antibody levels for each of the domains with age and infection status respectively. Children were split into two groups on the basis of the median age (1.7yrs) and median antibody responses were compared. Antibody levels were significantly higher in the older children for group A but not group B/C domains (Figure 2A). Therefore, the older children had more past exposure to group A antigens than younger children, whereas group B/C domains were similarly recognized irrespective of age. Concurrent microscopic parasitemia (median = 3349 parasites/μL) at the time of antibody measurement was associated with significantly higher antibody levels as compared to the non-infected individuals for all domains except group B NTS-DBLα0.16_PFL1955w (p = 0.94, Wilcoxon rank-sum test) (Figure 2B), which was poorly reactive overall. Children who were infected at enrollment were 3.43 times more likely to be seropositive to at least one of the five PfEMP1 domains than non-infected children [Range: 1.6, 7.8, p = 0.002]. Therefore, current infection appears to boost antibody levels for all antigens across the cohort.

Antibodies to DBLβ3_PF11_0521 but not other domains are associated with reduced risk of high density clinical malaria.

To examine whether antibody responses against any of the domains were associated with protection against clinical malaria, we conducted a prospective analysis of risk of febrile episodes and antibody responses. Associations between plasma antibody levels and subsequent risk of symptomatic malaria (all clinical episodes (fever and ≥2500 parasites/μl) and high density clinical malaria (fever and ≥10000 parasites/μl)), were determined by grouping individuals into tertiles, and comparing high and low antibody groups. An important feature of the analysis is the adjustment for confounding variables at the individual level such as the molecular force of blood stage infection (molFOB, see Materials and Methods). Children with high levels of antibodies to DBLβ3_PF11_0521 had a 37% reduction in risk of high-density clinical malaria that was highly significant (febrile illness with ≥10000 parasites/μl: adjusted incidence rate ratio (aIRR) = 0.63 [95% CI: 0.45-0.88; p = 0.007]) and a 25% reduced risk of clinical malaria that was borderline significant (febrile illness with ≥2500 parasites/μl: (aIRR) = 0.75 [95% CI: 0.55-1.01; p = 0.06]).
However, there was no significant reduction in risk in either presentation of clinical malaria for the other four domains tested (Figure 3).

Children that developed severe malaria had significantly lower antibodies to DBLβ3\textsubscript{PF11\_0521}.

Having determined that antibody responses to DBLβ3\textsubscript{PF11\_0521} were associated with protection against clinical malaria, and in particular high-density clinical malaria, a biomarker for severe disease, we then wanted to examine antibody responses in the children that experienced severe disease in the follow-up period. According to WHO criteria (31), of the 187 children, 18 experienced severe \textit{P. falciparum} malaria during the follow-up period (25), Table S2. On average, these children were similar to those that did not develop severe malaria with respect to age (Severe cases: 1.64 yrs vs Non-severe cases: 1.89 yrs, \( p = 0.12 \)) and exposure, experiencing a similar number of distinct \textit{P. falciparum} infections during the follow-up period (Severe: 5.19 and non-severe: 5.17, \( p = 0.98 \)). However, children who experienced severe malaria had significantly lower antibodies to DBLβ3\textsubscript{PF11\_0521} at baseline than those that did not develop severe malaria (Figure 4, \( p = 0.004 \)). Children who developed severe malaria also had significantly lower antibodies to the other type A PfEMP1 domains (\( p < 0.01 \)) and the type B/C domain CIDRα3.4\textsubscript{PFL1955w} (\( p = 0.024 \)). For the NTS-DBLα0.16\textsubscript{PFL1955w} domain with low reactivity, there was no significant difference (\( p = 0.95 \), Figure 4).

PNG PfEMP1 with ICAM1-binding DBLβ domains are enriched for DC4 and 13.

The PF11\_0521 PfEMP1 variant is from 3D7, an isolate with possible African origin and dual-binding activity with ICAM1 (via DBLβ) and EPCR (via DC13, (21)). Previous studies investigating dual ICAM1-EPCR binding have focused on collections of reference strains or field isolates from different geographic areas (21, 32). To determine whether PNG isolates possess ICAM-binding motifs and to study the associated PfEMP1 domain architecture, we conducted a detailed analysis of \textit{var} genes extracted from whole genome sequence data of 125 \textit{P. falciparum} isolates from PNG. Among the genomes we identified 4044 full or partial ORF’s that were classified as \textit{var} genes using BLAST against a database of classified DBL and CIDR domains (mean no. distinct PfEMP1 per genome =32, range = 1-60, mean coverage =33 reads, range = 8-124 reads, Supporting File: Table S3). Of those, 117 genomes contained 1505 DBLβ domains with a mean coverage of 35 reads distributed among 1420 \textit{var} genes (i.e. many \textit{var} genes had...
multiple DBLβ domains). For the 8 PNG genomes without DBLβ domains, 6 had low coverage resulting in poor sampling of var genes (n=1-6), whilst two isolates had higher coverage and contained 12 and 21 var genes. Among 1505 PNG DBLβ sequences (Supporting File: Table S3, Dataset 1), 81 contained the ICAM1-binding motif (21). They included 47 DBLβ1 (58%), 33 DBLβ3 (41%) and 1 DBLβ7 (1%) that were distributed among 61 genomes, and 80 var genes (1 var gene had two ICAM1 binding motifs: DBLβ1 and DBLβ7) (Supporting File: Table S4). To examine PNG DBLβ diversity in context with parasite populations of other malaria endemic countries, 279 DBLβ sequences from other countries and reference isolates including DBLβ3PF11_0521 (linked to DC13) and DBLβ3PF1235w (another ICAM1 binding variant linked to DC4, (18, 19)) were included in the analysis (Supporting File: Dataset 2). Of these, 22 contained the ICAM1-binding motif and they were found within 11 DBLβ1 (50%) and 11 DBLβ3 (50%) domains (Supporting File: Table S5). All DBLβ1 and β3 from both datasets (406 PNG + 178 other, Datasets 1 and 2) were then combined for multiple alignment. Truncated sequences were removed resulting in 455 sequences, including 102 with the ICAM1 motif. Phylogenetic analysis of the SD3 region encompassing the ICAM1-binding motif, revealed an “ICAM1-binding” cluster populated entirely by DBLβ with the ICAM1-binding motif (Figure 5A). The average evolutionary distance of the DBLβ predicted to bind ICAM1 was 0.532 compared to 1.011 for the remaining DBLβ. PNG DBLβ sequences were distributed throughout the tree, and there was no evidence of population structure. Of note, PF13_0003 contains a DBLβ sequence, however it did not contain an ICAM1-binding motif and was divergent to sequences in the ICAM1-binding clade (Figure 5A). These results suggest that the majority of PNG parasites carry at least one PfEMP1 with predicted ICAM1-binding.

To determine the PfEMP1 context of DBLβ with predicted ICAM1 binding, we investigated the domain architecture of full-length var gene assemblies from the PNG genomic sequence data with respect to adjacent domains and specific DC classes. All PNG PfEMP1 sequences with DBLβ1 and β3 domains contained CIDRα1 domains (EPCR-binding). However, PfEMP1 containing DBLβ domains with the ICAM1 motif (n=80, Tables S3,4) were significantly enriched for adjacent DC13 structures (46.3% compared to 20.5% among PfEMP1 with DBLβ domains with no motif, p<0.0001, Binomial Exact Test) and DC4 (13.8% compared to 6.0%, p=0.008, Binomial Exact Test) (Figure 5B, Supporting File: Table S6), which are strongly
associated with severe malaria and dual EPCR/ICAM1-binding (21, 32). In the phylogenetic analysis, DBL\(\beta\) with adjacent DC4 and DC13 were divergent, although several DBL\(\beta\) associated with DC4 were identical, suggesting conservation of a common var gene (Figure S2). Other DC structures including DC8, DC16 and DC1-var1 were not found amongst the PfEMP1 with the ICAM1 binding motif (Figure 5B). This suggests that specific classes of PfEMP1 with dual binding specificity to ICAM1 and EPCR are maintained in PNG isolates, confirming previous observations in African isolates (21).
DISCUSSION

PF11_0521 belongs to a class of group A PFEMP1 with DBLβ domains shown to bind ICAM1, and found exclusively in PfEMP1 with EPCR-binding CIDRα1 (16, 35). These domains are adjacent to each other and together they may produce a binding phenotype often associated with sequestration leading to cerebral malaria (13, 18-21, 32). Our results support a role for these PfEMP1 in clinical and severe malaria in young children by demonstrating that (i) antibodies against DBLβ<sub>3PF11_0521</sub> are significantly associated with protection against high-density clinical malaria, (ii) children that developed severe malaria had lower levels of antibodies to DBLβ<sub>3PF11_0521</sub> prior to the disease episode (albeit antibodies to other antigens were also lacking in these children), (iii) there are ICAM1-binding motifs present in DBLβ sequences from PNG isolates, and (iv) PfEMP1/var genes with predicted ICAM1-binding are also predicted to bind EPCR (21, 32).

The cohort of 1-3 year old PNG children was specifically chosen to explore early antibody responses to PfEMP1, because their immunity to malaria was incomplete (25) and to limit the complex background antibody responses that are observed in older children and adults (27). Antibodies to PfEMP1 domains were associated with age and current infection in the cohort, consistent with exposure driving the acquisition and maintenance of immunity to malaria (2). In addition to the measurement of exposure to new infections in the follow up period (molFOB), these results provided a basis for exploring associations with the risk of disease adjusted for important confounding factors (33). Because the number of severe disease cases was small (n=18), we initially focused the prospective risk analysis on clinical infections, which occurred at a high rate in these children (25). High density clinical malaria (fever plus ≥10,000 parasites/μL) is considered a surrogate marker for severe disease since these children have intense infections, yet are not classified into any of the severe disease syndromes by WHO criteria (31). The significant reduction in risk of these high density clinical infections if children had high levels of DBLβ<sub>3PF11_0521</sub> antibodies, suggests that inhibition of ICAM1-binding or other binding phenotype by antibodies against these PfEMP1 may limit parasite burden and progression to severe malaria. Antibodies to DBLβ<sub>3PF11_0521</sub> have previously been associated with a reduced risk of hospitalization with severe or moderately severe malaria in Tanzanian children (22), clearly demonstrating potentially protective immune response associated with exposure to PfEMP1 with...
DBLβ3PF11_0521-like variants. Tanzanian children with high antibody reactivity to CIDRα1 from the other confirmed ICAM1-binding protein, PFD1235w, also had a lower risk of anemia (hemoglobin <11 g/dL) and clinical malaria (34). Furthermore, ICAM1-binding inhibitory antibodies are common in hyper-immune adults living in endemic areas (19) suggesting that they may play an important role in the maintenance of clinical immunity. The mechanism of protection may be through direct inhibition of ICAM1-binding as shown in vitro experiments (21, 35), indirectly, through prevention of binding of other domains, such as EPCR-binding CIDRα1 (12), or synergistic antibody responses targeting multiple domains.

Whilst a broad repertoire of PfEMP1 antibodies was acquired in these young children, only antibodies against the ICAM1-binding full-length DBLβ3PF11_0521 domain were associated with protection against both clinical and severe malaria. The results are strengthened by fact that children also acquired antibodies to domains from the other group A PfEMP1 variants tested in this study (NTS-DBLα1.6PF13_0003 and CIDRδPF13_0003), yet they were not associated with a reduced risk of clinical malaria. The higher recognition of group A antigens in the children overall, and the significantly higher recognition in older children for group A, but not group B/C antigens, suggests shared epitopes within this group of PfEMP1 variants and confirms a hierarchy of PfEMP1 exposure with age (27-30). One caveat to mention is the use of different assays for DBLβ3PF11_0521 and the control antigens, which prevented the direct comparison of antibody units between antigens. However, the prospective risk analyses and comparison between groups were done for each antigen independently, thus limiting potential biases of the different assays. Another limitation of this study is the small number of PfEMP1 proteins and the lack of a direct comparison of DBLβ3PF11_0521 with other non-ICAM1 binding DBLβ domains. Parallel analyses of antibody responses to a large panel and variety of PfEMP1 domains will be a valuable extension of this study.

Dissecting the association of antibodies to different parasite antigens with protection and exposure is important in understanding naturally acquired immunity to malaria (33). In the same cohort, high levels of antibodies to merozoite antigens were predictive of an increased risk of developing clinical malaria (33). In older PNG children aged 5-14 years however, antibodies against merozoite antigens were found to be associated with protection against clinical malaria (33). Therefore, merozoite antigens are biomarkers of accumulated malaria exposure in the
younger age group, however with increasing exposure and responses of higher magnitude, antibodies to merozoite antigens become biomarkers of protective immunity (33). We used the molFOB as a marker of exposure at an individual level to adjust for the confounding effects of exposure in the prospective risk analysis (26, 33). We observed significant protection against clinical malaria in children with high levels of DBLβ3PF11_0521 antibodies despite these differing exposures to malaria. That is, children who had antibodies to DBLβ3PF11_0521 had lower rates of clinical malaria than those who had low levels of these antibodies, even after adjusting for individual differences in the rate of new malaria infections.

We also found an association between the lack of antibodies to DBLβ3PF11_0521 and other group A PfEMP1 domains and the prospective risk of severe malaria, however in this analysis we also found associations with other antigens. Children who experienced only uncomplicated or asymptomatic malaria had significantly higher levels of antibodies to all three group A antigens and one of the group B antigens than children who developed severe malaria in the follow up period. Taken together the observed association of antibodies to DBLβ3PF11_0521 with reduced risk of high density clinical malaria and severe malaria suggests epitopes in DBLβ3PF11_0521–like sequences or adjacent PfEMP1 domains may be important targets of protective immunity.

The association with other domains in severe malaria may be due to the fact that we did not adjust for confounders and is also consistent with the early acquisition (and potential protective effects) of antibody responses to group A antigens compared to B antigens (27,28,30). However, we cannot rule out that other domains tested are important targets of protection against severe malaria.

The importance of host-parasite interactions via ICAM1 to the PNG parasite population is indicated by the maintenance of a class of relatively conserved DBLβ sequences with predicted ICAM1-binding (18, 21). Phylogenetic analysis of the C-terminal DBL subdomain 3 (SD3) domain of DBLβ sequences extracted from 125 parasite genomes of PNG together with those of geographically diverse isolates identified a cluster of sequences previously shown (19, 21) or predicted to bind ICAM1 as distinguished by a common sequence motif (21). Our data suggests that similar subsets of group A PfEMP1 with ICAM1-binding DBLβ are found in
PNG parasites. The lower diversity of this group of sequences compared to other non-ICAM1 binding DBLβ1/3 suggests positive selection due to functional specialization for binding to ICAM1. We also found that DBLβ domains with ICAM1-binding motifs are located adjacent to domain cassettes associated with EPCR binding and severe disease (DC13 and DC4) (18, 20, 21). This assemblage is predicted to confer a dual binding phenotype that has been associated with severe malaria (12, 21). These results are consistent with the conservation of these specialized classes of PfEMP1 across large geographic distances since previous studies have focused on African parasites or reference isolates from diverse locations (18, 21, 32). The high reactivity of children’s sera to the full-length DBLβ3PF11_0521 supports the notion that conserved epitopes exist that can be targeted by cross-reactive and protective antibodies (18, 21, 35). However, the ICAM1-binding motif is located in C-terminal part of the protein (SD3 region), and no sequence traits in the N-terminal part of the domain have been linked to ICAM1 binding. The protective association may therefore be accounted for by epitopes outside the SD3 region or as a result of its concurrence with other virulence-associated domains in the same PfEMP1.

Our analysis of antibodies to functionally diverse PfEMP1 domains extends previous insights into early exposure to PfEMP1 (22, 27, 28, 30), revealing that young children from PNG are highly exposed to group A antigens whilst having limited exposure to group B/C antigens. The finding that high levels of antibodies against DBLβ3PF11_0521 are associated with a reduced risk of high-density clinical and severe malaria supports a role for PfEMP1 in malaria pathogenesis via ICAM1-binding domains or adhesion of adjacent domains. The demonstration that PNG P. falciparum isolates contain PfEMP1/var genes with predicted ICAM1-binding closely linked to predicted EPCR-binding CIDR domains suggests positive selection and functional specialization of a subclass of dual binding PfEMP1 implicated in severe malaria syndromes. Studies investigating whether antibodies in clinically immune children interrupt binding interactions between ICAM1 and this class of DBLβ sequences would establish a more direct link to protection against malaria. Antibodies to the EPCR-binding CIDRα1 domains co-occurring with ICAM1-binding DBLβ may have synergistic protective effects, however this is yet to be established. This study adds to the growing body of evidence supporting the development of
specific classes of PfEMP1 as vaccine candidates. Furthermore, it suggests that this class of DBLβ domain could be used as diagnostic antigens to track population immunity during malaria elimination.
MATERIALS AND METHODS

Cohort study design

Plasma samples were collected during a longitudinal cohort survey conducted in the East Sepik Province of PNG. A detailed description of the study is published elsewhere (25, 26). Briefly, 190 children aged 1–3 yrs were enrolled at the start of the study in March 2006 and 74 additional children were enrolled over the following six months. Children were followed for 69 weeks with active and passive follow-up (25). Children were visited fortnightly with collection of 2 blood samples 24 hours apart for active detection of malaria infection every 8–9 weeks. The demographic and clinical characteristics of the study population and incidence of clinical malaria and severe malaria in each 8–9 week follow-up interval have been described in detail elsewhere (25, 26). Antibody assays were performed on plasma samples collected from 187 (of the 190) children enrolled in March 2006. Of the 187 children, 48 (25.6%) were microscopy positive for *P. falciparum*. The average molFOB in this subset of the cohort was 5.2 and the average number of clinical episodes was 2 per child per year at risk (25, 26). *P. falciparum* clinical episodes were defined as febrile illness (axillary temperature ≥37.5°C or history of fever in the preceding 48hrs) and >2500 parasites/µl. A high-density clinical episode was defined as febrile illness and >10,000 parasites/µl. Characteristics of children with severe malaria are summarized in Table S2.

Written informed consent was obtained before enrolment of each child. Ethical approval for the study was granted from the PNG Institute of Medical Research (10.21), the Medical Research Advisory Council of PNG (10.55) and the Walter and Eliza Hall Institute of Medical Research (11.03).

Protein expression, purification and refolding

*DBLβ3PF11_0521* (also known as *DBL2β*PF11_0521) was expressed, purified and refolded as described previously (17). As control *PfEMP1* proteins without ICAM1 binding activity, we selected NTS-DBLα and CIDR domains of two *var* genes of 3D7: *PF13_0003* is a group A *PfEMP1* that has been associated with the formation of rosettes, a phenotype linked with severe malaria (36). In contrast, *PFL1955w* is a group B/C *PfEMP1* with limited antibodies acquired in young children (36). The sequences of all five domains are available in the Supporting Materials (Text S1).

NTS-DBLα1.6*PF13_0003*, CIDRδ*PF13_0003*, NTS-DBLα0.16*PF1L1955w* and CIDRα3.4*PF1L1955w* codon-
optimized sequences were synthesized for E. coli expression (GeneArt). The GenBank accession numbers are PF13_0003 (XM_001349704) and PFL1955w (XM_001350761). Sequences were excised from the supplier’s vector using BamHI and XhoI and ligated into the pProExHTb expression vector (Invitrogen), which incorporates an N-terminal hexahistidine fusion tag. The vectors were then transformed into E. coli strain BL21 (DE3) for expression as described previously (37). Briefly, the transformed cultures were grown in super broth and expression was induced with 1mm isopropyl 1-thio-β-D-galactopyranoside, and cultures were grown for a further 3 h at 37 °C. The cells were harvested by centrifugation and lysed by sonication and processed either as insoluble inclusion bodies (NTS-DBLα1.6PF13_0003, CIDRδPF13_0003 and NTS-DBLα0.16PFL1955w) or soluble proteins (CIDRα3.4PFL1955w).

NTS-DBLα1.6PF13_0003, CIDRδPF13_0003 and NTS-DBLα0.16PFL1955w proteins were deposited as insoluble inclusion bodies. Cells were lysed by sonication, and the inclusion bodies solubilized by the addition of 6M guanidine HCl, pH 8.0. The solubilized proteins were purified by metal-chelating chromatography using a nickel-nitrilotriacetic acid (Ni-NTA) column (Qiagen) under reducing conditions. Optimum refolding conditions were determined for each protein. The NiNTA eluted CIDRα3.4PFL1955w and the refolded CIDRδPF13_0003 proteins were purified using strong anion-exchange chromatography. Proteins were eluted from Hitrap Q column (GE Healthcare) and the relevant fractions were pooled and concentrated. For CIDRα3.4PFL1955w, the protein was further purified by size exclusion chromatography. The refolded NTS-DBLα1.6PF13_0003 and NTS-DBLα0.16PFL1955w domains were further purified using cation-exchange chromatography. Bound proteins were eluted from a Hitrap SP column (GE Healthcare) and the relevant fractions were pooled, concentrated and further purified by size exclusion chromatography. The purity of each protein was assessed on SDS-PAGE gels and via western immunoblots using standard conditions. Briefly, proteins were run on a 4–12% Bis-Tris SDS-PAGE (Invitrogen). Standard Western blotting procedures were performed for non-reduced and reduced (by addition of β-mercaptoethanol) samples using nitrocellulose and the immunoblots were processed with enhanced chemiluminescence (ECL) substrates (GE Healthcare). For all Western blots, recombinant proteins were detected with pooled hyperimmune sera from highly exposed PNG adults (see Supplementary figure 1). A single batch of each protein was used for all serological screening. The purified proteins were assessed using a pool of
hyperimmune plasma (see Supplementary Figure 1) and a single batch of each protein was used for all serological screening.

Measurement of antibody responses

Plasma samples collected at enrolment (n=187) were tested for antibodies comprising total Immunoglobulin G (IgG) to DBLβ3PF11_0521 using a standard ELISA assay. For the other four domains, IgG levels were measured using the cytometric bead array (CBA) as described previously (38). The details are described below.

ELISAs were performed to measure total IgG using standard methods. Ninety-six well plates (Nunc, Denmark) were coated with 1μg/ml of DBLβ3PF11_0521 recombinant protein in PBS and incubated overnight at 4°C. PBS with 5% skim milk was used for blocking and PBS with 1% skim milk and 0.05% tween for diluting the plasma samples and antibodies. Plasma was added at 1 in 100 dilutions. For measurement of total IgG, horseradish peroxidase-conjugated mouse anti-human IgG (SouthernBiotech, USA) was used at a dilution of 1 in 1000. Finally, TMB microwell peroxidase substrate (KPL, Inc., Australia) was added, and the reaction was stopped using 1M H₃PO₄ and the optical density (OD) was measured at 450 nm. All samples were tested in duplicate. Background (determined from the wells with no plasma) was deducted and the threshold for a seropositive response was determined using reactivities of 1:100 diluted plasma samples from anonymous malaria-naïve Australian adults (n=12). The mean value among these negative control plasma samples plus 3SD was used as a cut-off value to define seropositivity.

CBAs were carried out using four micro-beads (BD Bioscience, San Diego, CA, USA) of distinct and non-overlapping fluorescence intensities covalently coupled to NTS-DBLa1PF13_0003, CIDRδPF13_0003, NTS-DBLa0.16PFL1955w and CIDRα3.4PFL1955w recombinant proteins according to the manufacturer's protocol. Briefly, 150μl of selected micro-beads were sonicated for 1 min and incubated with 3.8μl of 1M Dithiothreitol (DTT) for 1h at room temperature with agitation. The beads were washed 3 times and resuspended in 40 μl of coupling buffer (BD Bioscience). Recombinant proteins (1mg/ml concentration) were activated by incubating with 4μl of sulfosuccinimidyl 4-N-maleimidomethyl cyclohexane 1-carboxylate (2 mg/ml) for 1h. The protein mixture was then run through a buffer exchange spin column (Bio-Rad) pre-equilibrated...
with the coupling buffer (BD Bioscience). The activated protein was added to the washed microbeads and allowed to conjugate for 1h at room temperature with agitation. Four µl of N-Ethylmaleimide (2 mg/ml) was added and the mixture was incubated for another 15 min. The conjugated micro-beads were then washed, resuspended in 1 ml of storage buffer (BD Bioscience) and kept at 4°C in the dark. For assessment of antibody responses, 1µl of conjugated micro-beads was diluted in 50 µl of washing buffer (BD Bioscience) containing 1 in 100 dilutions of plasma samples. Duplicate samples were then incubated for 1hr at room temperature in the dark, washed and further incubated with mouse anti-human IgG phycoerythrin (PE) conjugated antibody (BD, Bioscience) for 1hr at room temperature. After washing, the samples were re-suspended in diluent buffer containing PE (BD Bioscience) and acquired using an LSR Fortessa analyzer (Becton Dickinson, New Jersey, USA). Analysis was performed using FlowJo software and the median fluorescence intensity (MFI) for each bead (recombinant protein) was calculated. Background (determined from the unconjugated beads with plasma samples and conjugated beads with no plasma samples) was deducted from the mean of each sample. The threshold for a seropositive response was determined using reactivities of 1:100 diluted plasma samples from anonymous malaria-naïve Australian adults (n=12). The mean value among these negative control plasma samples plus 3SD was considered seropositive. The analysis was done independently for each antigen.

A serial dilution of plasma samples from a pool of hyperimmune PNG adults were included in each plate to determine standard curves, which was later fitted using a 5-parameter logistic regression model (Giraldo J, Vivas, NM et al. 2002 Pharmacol Ther) to transform antibody measured by the two assays into relative antibody units and correct plate-to-plate variations within an assay.

**Statistical Analysis**

Statistical analyses were performed using STATA version 12.1 software (Stata Corporation, USA). Differences in median antibody levels by age and *P. falciparum* infection status were compared using the Wilcoxon rank-sum test. The proportions of seropositive children to different domains were compared using Chi-square tests. Correlation coefficients for antibody levels were determined using Pearson’s correlation.
Analyses of the cohort data showed significant over-dispersion in the number of clinical episodes per child (25), as a result a negative binomial model with generalized estimating equations (GEE) (based on an XTNBREG procedure) with an exchangeable correlation structure and a semi-robust variance estimator was used for the analyses of association of antibody levels and incidence of clinical (fever with ≥ 2500 parasites/µl) and high-density clinical episode (febrile illness and ≥ 10,000 parasites/µl) during the follow-up period. Antibody levels were grouped into tertiles (low, medium, and high responses) and their association with clinical and high-density clinical episode was assessed by univariate analyses adjusted for seasonal variation, village of residence, age at the time of enrollment, *P. falciparum* infection status and individual exposure as measured by the *mol*FOB. The *mol*FOB is the number of genetically distinct *P. falciparum* clones (based on *msp2* genotyping) each child acquired per year at risk (26).
Sequence Analyses

To identify ICAM1-binding DBL\(\beta\) domains and corresponding full length PfEMP1 sequences in the PNG parasite population, we extracted var gene sequences from the genomes of 125 clinical \textit{P. falciparum} isolates collected from three distinct geographic areas of PNG. These isolates were sequenced as part of the MalariaGEN Community Project (European Nucleotide Archive (ENA) accession numbers listed in the Supporting File: Table S3). Illumina\textsuperscript{®} short read data sequences were assembled using Velvet version 1.2.03 (39). Open reading frames were extracted using Virtual Ribosome ORF finder (40) and \textit{var} gene-encoded PfEMP1 domain sequences were extracted from these by BLAST using a library of previously annotated PfEMP1 domain sequences from 7 reference genomes \textit{MalariaGEN} Community Project (42) and seven reference genomes (41), including confirmed ICAM1-binding DBL\(\beta\) domains of PF11_0521 and PFD1235w (18) (n=279, Supporting File: Dataset 2). DBL\(\beta\) sequences were screened for a relaxed version of the ICAM1 motif: N-G-G-[PA]-x-Y-x(27)-G-P-P-x(3)-H (21) using the web-based server Scan Prosite (43). All sequences from the DBL\(\beta\) classes with ICAM1-binding motifs (DBL\(\beta\)1 and DBL\(\beta\)3) were aligned using Muscle with default settings in MEGA version 7.0 (37). After removing truncated sequences (n=40), and focusing on the S3 region of DBL\(\beta\), which contains the ICAM1 binding motif (21), we then conducted a phylogenetic analysis using a total of 367 PNG DBL\(\beta\)1 and DBL\(\beta\)3 sequences originating from PNG isolates together with 89 DBL\(\beta\)1 and DBL\(\beta\)3 sequences from the supplementary dataset. A maximum likelihood tree was estimated using the JTT substitution model with 1000 bootstrap repetitions in MEGA version 7.0 (44). We measured the average evolutionary divergence among sequences using the Dayhoff model in MEGA version 7.0 (44).

In addition, we characterized the domain architecture of PfEMP1 with DBL\(\beta\)1 and DBL\(\beta\)3 domains by extracting the domain classifications from the BLAST output (Supporting File: Table S6). We then assessed whether the frequency of each domain cassette amongst the ICAM1-motif containing sequences varied significantly from the expected frequency (among a subset of 331 PfEMP1 with DBL\(\beta\)1/3 domains) using an Exact Binomial test using R software (45).
Acknowledgments

We thank all participants in the study and their parents or guardians, staff involved in the study at the Papua New Guinea Institute of Medical Research. This research was supported by the National Health and Medical Research Council of Australia (NHMRC grant numbers GNT1005653 and GNT1027109 to AEB) and Danish Council for Independent Research (DFF-4004-00624B). Production of DBLβ3PF11_0521 protein was supported by a US National Institutes of Health Grant R01AI092120 to AVO. IM is supported by an NHMRC Senior Research Fellowship. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.
REFERENCES


45. FIGURE LEGENDS

Figure 1. Antibody responses to five PfEMP1 domains in 187 young Papua New Guinean children. A. Domain composition and seroprevalence of the five tested PfEMP1 domains. Seroprevalence is indicated in percentage above the relevant domains. B. Correlation coefficients for seropositivity to five PfEMP1 domains. Significant correlations (p < 0.001) are indicated by the asterisk.
Figure 2. Relationship between antibody responses to five PfEMP1 domains, age and infection status. A. IgG levels stratified by age (age groups determined by the median age (1.7 years)). B. IgG levels stratified by *P. falciparum* infection status. In both panel, box and whisker plots for the relative antibody units (in log10) are shown for the five PfEMP1 domains. Boxes show the interquartile range, midline is the median, whiskers the 95% confidence interval and dots are the outliers (95–99%). *P*-values for the differences were determined using the Wilcoxon rank-sum test between the groups.

Figure 3. Antibody responses to five PfEMP1 domains and prospective risk of symptomatic malaria. Antibody levels were grouped into three equal groups (High, Medium and Low). The incidence rate of clinical malaria and high-density clinical malaria were compared for high and low responders for each tested domain using negative binomial regression. The incidence rate ratios were adjusted for villages of residence, seasonal variation, age (continuous), infection status at the time of antibody measurement and differences in individual exposure (*molFOB*). Adjusted incidence rate ratio (aIRR) for the comparison of high and low responders and the 95% confidence intervals are shown. The *p*-values are indicated only when it is significant (*p*<0.05).

Figure 4. Antibody response to five PfEMP1 domains and development of severe malaria. Mean and standard error of the relative antibody units are shown for children who experienced severe malaria (red, *n*=18) and those who did not (black, *n*=169). *P*-values for *t*-test comparisons of the means are indicated for each domain.

Figure 5. Conservation of ICAM1 binding motifs and dual EPCR-ICAM1 binding cassettes in Papua New Guinea. *Var* gene sequences were assembled for 125 *P. falciparum* isolates of Papua New Guinea and domains classified as previously described (41). (A) Maximum Likelihood tree of 473 DBLβ1 and DBLβ3 sequences including 406 from PNG (pink = DBLβ1, red = DBLβ3) and 67 from isolates from diverse geographic locations (cyan = DBLβ1, blue = DBLβ3). Sequences containing the minimal ICAM1-binding motif are indicated by solid circles. Gene Ids described in this study are indicated against the DBLβ1/3 variant position in the tree. (B) Domain architecture of PNG *var* genes containing DBLβ1/3 domains with ICAM1 motifs. The presence of domain cassettes (DCs) among the 80 *var* genes containing DBLβ1/3 with the...
ICAM1 binding motif, relative to that among a subset of 331 PNG var genes containing any DBLβ1/3, is indicated on the right. Significance was determined by Binomial Exact test.