

One-step design of a stable variant of the malaria invasion protein RH5 for use as a vaccine immunogen

Ivan Campeotto^{a,1}, Adi Goldenzweig^{b,1}, Jack Davey^a, Lea Barfod^c, Jennifer M. Marshall^c, Sarah E. Silk^c, Katherine E. Wright^a, Simon J. Draper^c, Matthew K. Higgins^{a,2}, and Sarel J. Fleishman^{b,2}

^aDepartment of Biochemistry, University of Oxford, Oxford OX1 3QU, United Kingdom; ^bDepartment of Biomolecular Sciences, Weizmann Institute of Science, Rehovot 7610001, Israel; and ^cThe Jenner Institute, University of Oxford, Oxford OX3 7DQ, United Kingdom

Edited by David Baker, University of Washington, Seattle, WA, and approved December 16, 2016 (received for review October 11, 2016)

Many promising vaccine candidates from pathogenic viruses, bacteria, and parasites are unstable and cannot be produced cheaply for clinical use. For instance, *Plasmodium falciparum* reticulocyte-binding protein homolog 5 (PfRH5) is essential for erythrocyte invasion, is highly conserved among field isolates, and elicits antibodies that neutralize in vitro and protect in an animal model, making it a leading malaria vaccine candidate. However, functional RH5 is only expressible in eukaryotic systems and exhibits moderate temperature tolerance, limiting its usefulness in hot and low-income countries where malaria prevails. Current approaches to immunogen stabilization involve iterative application of rational or semirational design, random mutagenesis, and biochemical characterization. Typically, each round of optimization yields minor improvement in stability, and multiple rounds are required. In contrast, we developed a one-step design strategy using phylogenetic analysis and Rosetta atomistic calculations to design PfRH5 variants with improved packing and surface polarity. To demonstrate the robustness of this approach, we tested three PfRH5 designs, all of which showed improved stability relative to wild type. The best, bearing 18 mutations relative to PfRH5, expressed in a folded form in bacteria at >1 mg of protein per L of culture, and had 10–15 °C higher thermal tolerance than wild type, while also retaining ligand binding and immunogenic properties indistinguishable from wild type, proving its value as an immunogen for a future generation of vaccines against the malaria blood stage. We envision that this efficient computational stability design methodology will also be used to enhance the biophysical properties of other recalcitrant vaccine candidates from emerging pathogens.

malaria | immunogen design | Rosetta | PROSS

Malaria places the gravest public-health burden of all parasitic diseases, leading to ~215 million human clinical cases and ~440,000 deaths annually (1). The most virulent parasite species, *Plasmodium falciparum*, is endemic in large parts of sub-Saharan Africa and Southeast Asia. It causes the majority of malaria deaths and is a major target for vaccine development. There are currently no licensed malaria vaccines on the market, and the leading anti-sporozoite subunit vaccine, RTS,S/AS01, has achieved only modest efficacy of 30–50% in phase II/III clinical trials (2, 3). There is therefore an urgent need to develop additional vaccines, including those that target various stages of the parasite life cycle (4).

One of the leading malaria vaccine candidates is *Plasmodium falciparum* reticulocyte-binding protein homolog 5 (PfRH5), a protein required for the establishment of blood stage infection. PfRH5 is released onto the surface of infective *Plasmodium falciparum* merozoites, binding to human basigin in an interaction that is essential for erythrocyte invasion (4–7). Compared with other *Plasmodium* surface antigens, it is remarkably conserved across field isolates (7–11) and antibodies that bind either PfRH5 or basigin show robust growth-inhibitory effects in vitro against all tested strains of *Plasmodium falciparum* (5, 7–9, 11–13). Additionally, in a challenge trial, immunization with PfRH5-based vaccines protected *Aotus* monkeys against heterologous challenge with a virulent *Plasmodium falciparum* strain (12). PfRH5 is therefore the best-performing antigen against the blood stage of the parasite, and

clinical trials are already underway to test its safety, immunogenicity, and efficacy in immunized human volunteers (4).

Despite this promise, PfRH5 suffers from two significant shortcomings as a subunit vaccine candidate. First, the protein has limited stability at high temperatures, and second, despite extensive protein engineering (11), correctly folded, soluble, and functional PfRH5 has not been produced in microbial expression hosts. Instead, production has relied on more expensive eukaryotic expression systems, such as transiently transfected HEK293 cells (7) or stable insect cell lines (11, 14). Because the most likely use for PfRH5-based vaccines would require infant immunization in hot and underdeveloped regions, where a cold chain for transporting vaccine formulations is very challenging, a stabilized and lyophilized variant that can be cheaply produced in microbial cells, and that will retain efficacy when stored at elevated temperatures, is highly desirable. We therefore aimed to design versions of PfRH5 with improved expression levels and thermal stability, without compromising their effectiveness as immunogens.

Many potential vaccine immunogens are only marginally stable. To address this problem, approaches for immunogen stabilization or grafting of immunogenic epitopes onto stable scaffolds have been implemented (15–21). However, key vaccine immunogens frequently have complex folds with significant flexibility and low stability. Together with the strict requirement to maintain neutralizing immunological responses, this means that current efforts for

Significance

Malaria is one of the world's most devastating infectious diseases, affecting hundreds of millions of people and resulting in nearly half a million deaths each year. The parasites that cause malaria must invade the red blood cells of an infected patient, while blocking this process prevents the disease. The PfRH5 protein is an exciting vaccine candidate required for red blood cell invasion by *Plasmodium falciparum*, the most deadly malaria parasite. Here, we describe our use of a streamlined computational methodology to design variants of PfRH5 that can be produced more simply and cheaply and that show greater thermal stability. This method has broad potential to help the design of vaccines that target many of the world's most deadly diseases.

Author contributions: I.C., A.G., S.J.D., M.K.H., and S.J.F. designed research; I.C., A.G., J.D., L.B., J.M.M., S.E.S., K.E.W., and M.K.H. performed research; K.E.W. contributed new reagents/analytic tools; I.C., A.G., J.D., L.B., J.M.M., S.E.S., S.J.D., M.K.H., and S.J.F. analyzed data; and I.C., A.G., M.K.H., and S.J.F. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 5MIO).

¹I.C. and A.G. contributed equally to this work.

²To whom correspondence may be addressed. Email: matthew.higgins@bioch.ox.ac.uk or sarel@weizmann.ac.il.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1616903114/-DCSupplemental.

immunogen stabilization often require time-consuming and labor-intensive cycles. For instance, in the design of superior HIV and respiratory syncytial virus immunogen variants, multiple rounds of rational design, random mutagenesis, and biochemical, immunological, and structural characterization were applied (15–21). Although effective, such iterative strategies limit the ability to respond quickly to emerging pathogens.

We recently described a stability-design algorithm, called PROSS (22), and demonstrated its effectiveness in designing variants of challenging human enzymes with much improved thermal stability and increased bacterial expression levels, without affecting protein function. Guided by the recent structures of PfrH5, alone and in complex with its human receptor basigin and with neutralizing antibodies (11, 23), we embarked on a one-step, structure-based computational design process to generate PfrH5 variants, aiming to improve expression levels and thermal stability while retaining structure, ligand binding, and immunogenic properties.

Results

Extending Sequence Analysis Below the “Twilight Zone.” The PROSS workflow comprises three stages. It first analyzes homologs of the target protein to identify, at each amino acid position, mutations that are most likely to occur through natural diversity of the protein family. Second, starting from a molecular structure of the target protein, Rosetta atomistic design simulations suggest a subset of these mutations, which are individually predicted to stabilize the wild-type protein. At the last step, Rosetta combinatorial sequence optimization is used to design several optimized variants, typically comprising >10 mutations each, with improved native-state energy. In all steps, amino acid conformations in the active site are held fixed to preserve function. In this workflow, high natural sequence diversity provides an essential source of information on tolerated variation, and we have so far applied PROSS only to targets with dozens and even hundreds of unique homologs (22).

PfrH5 presented an unusual challenge for the phylogenetic sequence analysis implemented in PROSS. As of May 2015, sequences of PfrH5 from *Plasmodium falciparum* field isolates showed 99% sequence identity to one another, and only one ortholog, from *Plasmodium reichenowi* (with 66% sequence identity to PfrH5) was available. This extremely low sequence diversity made application of the PROSS methodology “as is” impractical. To increase diversity, we extended the search below the twilight zone of sequence identity (<25%) and took special measures to validate the sequence alignment. Specifically, context-specific iterated BLAST (CSI-BLAST) identified 72 unique homologs at 15–25% sequence identity to PfrH5, more than 90% of which showed only 15–18% sequence identity to PfrH5. From this set, we excluded sequences from genera other than *Plasmodium*, and any sequences that showed more than 1% gaps in the aligned segment. Because the RH5 fold is considered unique and highly conserved in *Plasmodium* invasion proteins (11), these restrictions increase the likelihood that the sequences in the alignment belong to the same fold. As a further safety measure, we generated two alignments, a “permissive” alignment containing all the homologs described above and a “strict” one at >18% sequence identity. These alignments had 14 and 8 sequences, respectively (Datasets S1 and S2). As a template structure, we used a version of RH5 with both the flexible N terminus (residues 1–140) and a disordered loop (residues 248–296) removed, as this PfrH5 Δ NL variant contains the structured region of RH5 and retains the capacity to bind to basigin and induce production of inhibitory antibodies (11). The two alignments and the PfrH5 Δ NL structure were then provided to the PROSS algorithm in two independent runs. To preserve the function and immunological efficacy of the designed PfrH5 variants, we did not allow mutations at amino acid positions within 5 Å of the contact sites of either basigin or two anti-PfrH5 inhibitory antibodies, 9AD4 and QA1 (11, 13) (Table S1). We then visually inspected the designs and selected three for experimental testing,

one based on the strict alignment bearing 18 mutations relative to PfrH5 (PfrH5 Δ NL_{HS1}) and two using the permissive alignment bearing 25 and 15 mutations (PfrH5 Δ NL_{HS2} and PfrH5 Δ NL_{HS3}, respectively) (Fig. 1A, Table S1, and Datasets S1 and S2).

Generation of a Stabilized, Functional, and Bacterially Expressed PfrH5.

Genes encoding each of these three variants were designed and tested in *Escherichia coli* expression strain Rosetta-gami under different growth conditions. Although no detectable soluble expression was observed for PfrH5 Δ NL, each of the stabilized variants

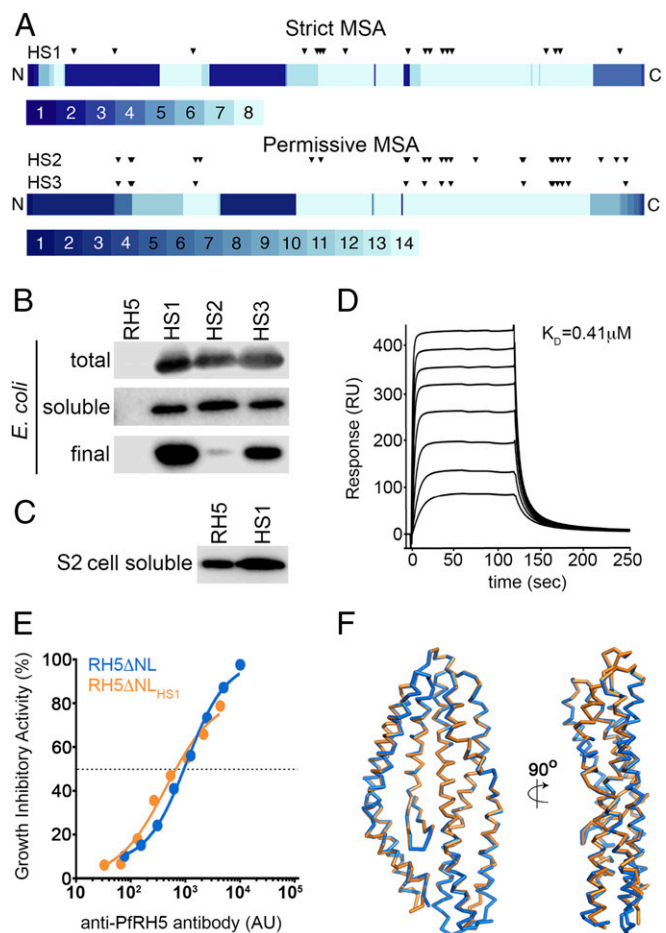


Fig. 1. Design of an *E. coli*-expressible PfrH5 variant. (A) Eighty percent of mutations in all three designed variants are located in the C-terminal half of PfrH5, where most of the aligned sequences contribute information on sequence diversity. The PfrH5 sequence is schematically shown from N to C terminus, and each position is colored according to the number of unique sequences contributing to the strict and permissive alignments, ranging from 1 to 8 and from 1 to 14, respectively. The locations of the mutations in each designed variant relative to wild-type PfrH5 Δ NL are indicated by triangles. (B) Expression levels of PfrH5 Δ NL (RH5), PfrH5 Δ NL_{HS1} (HS1), PfrH5 Δ NL_{HS2} (HS2), and PfrH5 Δ NL_{HS3} (HS3) from *E. coli*. “Total” is whole lysed cells, “soluble” is material after cell lysis and clarification, and “final” is after immobilized metal ion affinity chromatography and size exclusion chromatography. (C) Expression levels of PfrH5 Δ NL (RH5) and PfrH5 Δ NL_{HS1} (HS1) secreted in the cell supernatants from a stable *Drosophila melanogaster* Schneider 2 (S2) cell line. (D) Surface plasmon resonance analysis of the binding of PfrH5 Δ NL_{HS1} to basigin, with twofold dilutions of PfrH5 Δ NL_{HS1} from a maximal concentration of 8 μM . (E) In vitro GIA of purified IgG against 3D7 clone parasites from mice immunized with either PfrH5 Δ NL or PfrH5 Δ NL_{HS1}. The anti-PfrH5 antibody response was measured for each sample of purified IgG and is plotted against the measured level of GIA. The dashed line indicates 50% GIA, and each GIA data point represents the mean of each sample tested in triplicate. (F) The structure of PfrH5 Δ NL_{HS1}:9AD4 (orange) overlaid on the structure of PfrH5 Δ NL:9AD4 (blue).

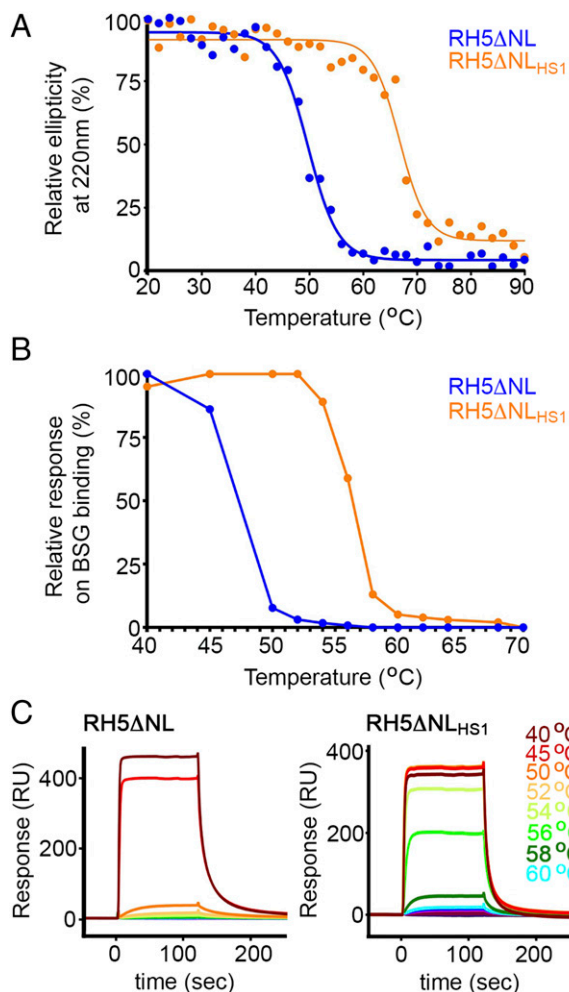


Fig. 2. Increased thermal stability of variant PfrH5ΔNL-HS1. (A) Determination of the effect of temperature on the ellipticity of PfrH5ΔNL and PfrH5ΔNL-HS1 at a wavelength of 220 nm as measured by circular dichroism. (B and C) Determination of the effect of temperature on the binding of PfrH5ΔNL and PfrH5ΔNL-HS1 to basigin (BSG), measured by surface plasmon resonance. Protein, at 16 μ M, was incubated for 60 min at the specified temperature before analysis at 8 μ M.

expressed at similar levels (Fig. 1B). Protein from each variant was purified by immobilized metal ion chromatography, followed by size exclusion chromatography (Fig. S1). This showed PfrH5ΔNL-HS1 to have the best production properties, yielding ~50-fold more protein than PfrH5ΔNL-HS2 and approximately fourfold more than PfrH5ΔNL-HS3. PfrH5ΔNL-HS1 was produced with a final yield of ~1.3 mg from each liter of *E. coli* culture. To determine whether these improvements in expression properties are limited to prokaryotic expression systems, we also generated a stable *Drosophila melanogaster* Schneider 2 (S2) cell line that expresses PfrH5ΔNL-HS1, allowing us to assess its expression levels in what is currently the leading system for PfrH5 expression (11, 14). Here too, we observed an increased yield of protein, with PfrH5ΔNL-HS1 expressed at levels threefold to fourfold higher than PfrH5ΔNL (Fig. 1C). Therefore, stability design resulted in the first versions of PfrH5 that express in a stable, soluble, and folded form in *E. coli* and were also shown to have the potential to increase the level of expression of PfrH5 in insect cell systems.

We next assessed the structural integrity and functionality of PfrH5ΔNL-HS1 purified from bacteria. We used surface plasmon resonance to show that this protein bound to basigin with an affinity of 0.41 μ M (Fig. 1D), comparable to the 0.29 μ M observed for PfrH5ΔNL (Fig. S2). In addition, basigin binding with an

affinity of ~0.3 μ M was retained after lyophilization and resuspension of both PfrH5ΔNL and PfrH5ΔNL-HS1, increasing the options for vaccine preparation (Fig. S2). To ensure that PfrH5ΔNL-HS1 contains the epitopes required to elicit an inhibitory immune response, we raised mouse polyclonal IgG and tested their ability to neutralize parasites in an in vitro assay of growth inhibition activity (GIA) (Fig. 1E). IgG raised against PfrH5ΔNL-HS1 showed a strong inhibitory effect, similar to that for IgG raised against PfrH5ΔNL. Indeed, these polyclonal IgG were qualitatively comparable, requiring very similar amounts of PfrH5-specific IgG to neutralize 50% of parasites. Finally, we determined the crystal structure of PfrH5ΔNL-HS1 in complex with the Fab fragment from 9AD4, an inhibitory monoclonal antibody raised previously against PfrH5ΔNL (13), to 2.35-Å resolution (Fig. 1F, Fig. S3, and Table S2). Composite omit maps showed clear electron density for mutated residues, largely in their designed positions (Fig. S3). However, there were no other significant changes in the PfrH5 structure with r.m.s.d. values of 0.7 Å for the backbone C α positions and 1.0 Å for the complex, when comparing PfrH5ΔNL-HS1 and PfrH5ΔNL. The stabilized and bacterially expressed PfrH5ΔNL-HS1 variant therefore retains the structure, ligand binding, and immunogenic properties of native PfrH5.

To assess the thermal stability of PfrH5ΔNL-HS1 (defined as the temperature at which one-half of the protein is inactive), we first conducted circular dichroism with a thermal melt, following ellipticity at 220 nm as a measure of α -helicity (Fig. 2A). Although ellipticity for PfrH5ΔNL was reduced by 50% at 48 °C, the stabilized variant, PfrH5ΔNL-HS1, showed the same loss of ellipticity at 66 °C, showing a ~20 °C improvement in thermal stability. To assess whether this loss of ellipticity is reversible, we heated PfrH5ΔNL-HS1 to 75 °C and cooled to 20 °C before repeating circular dichroism measurements, and observed no recovery of secondary structure, showing that thermal denaturation is irreversible (Fig. S4). As circular dichroism measures secondary-structure content rather than protein functionality, we also conducted an experiment in which we incubated the PfrH5 variants at different temperatures for an hour before returning them to room temperature and testing basigin

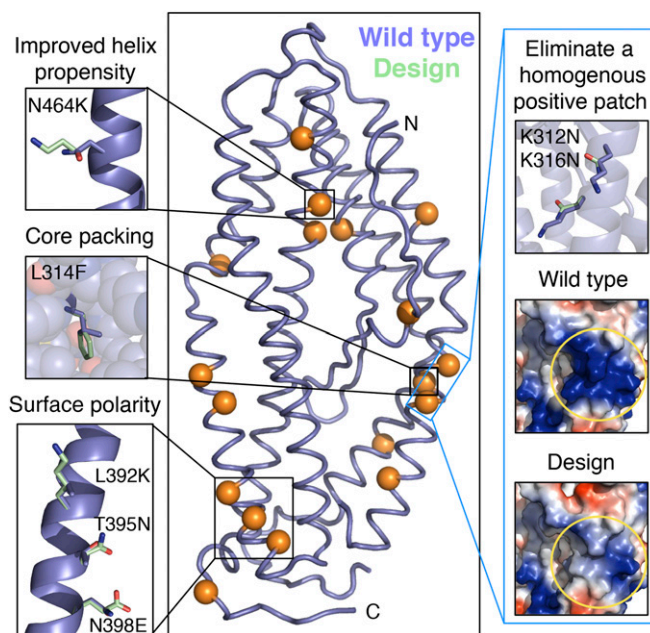


Fig. 3. Structural insights into thermal stability of PfrH5ΔNL-HS1. Structure of wild-type PfrH5 is shown in blue, and the 18 mutated positions, throughout PfrH5ΔNL-HS1, are indicated by orange spheres in the central panel. Surrounding thumbnails highlight stabilizing effects of selected mutations.

binding by surface plasmon resonance (Fig. 2*B* and *C*). P_{FRH5}ΔNL lost 50% of its basigin-binding capacity at ~47 °C, whereas P_{FRH5}ΔNL_{HS1} showed a similar loss at ~57 °C, showing a ~10 °C improvement in thermal stability. We have therefore designed a variant of P_{FRH5} that retains both the structure and immunogenicity of the native protein, while allowing expression in *E. coli* and increased thermal stability.

To assess the reasons for its higher bacterial expression levels and improved thermal stability, we compared the structure of P_{FRH5}ΔNL_{HS1} with that of P_{FRH5}ΔNL. The comparison showed several of the hallmarks expected for stabilizing mutations (Fig. 3 and Table S3). Of the 18 mutations, 15 occurred on the surface, of which 8 improved surface polarity, and an additional 2 (Lys312Asn and Lys316Asn) eliminated a homogenous positively charged patch. Indeed, such patches have been associated with aggregation and poor stability (24, 25). Furthermore, P_{FRH5} comprises a mostly helical backbone (75% of the sequence), and seven mutations increased helix-forming propensity relative to P_{FRH5}. Last, the design contains three mutations that improve hydrophobic core packing (Asp183Glu, Leu314Phe, and Ser467Ala). It is by combining a large set of mutations in a single variant that significant gains in expression levels and thermal stability can be achieved.

Discussion

In conclusion, we have used a computational method to design improved expression properties and thermal stability into a leading malaria vaccine candidate, P_{FRH5}. This generated an improved immunogen that could be scaled economically for production in *E. coli* and is more thermally stable for storage and delivery. This immunogen will be valuable in future generations of malaria vaccines. The approach used here is also highly applicable for the development of vaccine immunogens against other stages of the life cycle of *Plasmodium falciparum* or other pathogens. Indeed, a major challenge in applying PROSS to P_{FRH5} was the lack of sequence variation in this protein across different parasite strains, requiring a modified strategy for selecting distant sequence homologs. Although this will be an issue for some other vaccine candidates, such as highly conserved transmission-blocking targets, the majority of pathogen surface proteins, from viral surface proteins to the diverse *P. falciparum* erythrocyte membrane protein 1 (PfEMP1) family of *Plasmodium* surface proteins (26), are highly variable as they evolve under diversifying selection pressure to avoid detection by the acquired immune system. With judicious selection of natural variants on which to base the design process, the methods outlined here will be applicable to many of these cases, whether highly conserved or highly diverse.

Current methods for vaccine immunogen stabilization rely on laborious cycles of design and experimental testing. By contrast, the design strategy implemented here allows for rapid and significant enhancement of stability and expression levels in one test cycle. We therefore expect this method to contribute to vaccine immunogen production across a wide range of the most deadly diseases affecting humanity.

Materials and Methods

Collection and Filtering of Homologous Sequences. The P_{FRH5} structure was downloaded from the Protein Data Bank (PDB) [ID code 4WAT (23)], and homologous sequences were collected using CSI-BLAST (27, 28) to search the nonredundant (nr) database in May 2015, with *e* value of <10^{−4}, three iterations, a maximum of 500 sequences, and default values on all other parameters. Hits were clustered using cd-hit (29) at 98% threshold and default parameters. Hits from genera other than *Plasmodium* were excluded. Hits were also

excluded if their sequence identity to the query was lower than 15% or if they showed more than 1% gaps in the aligned segment. Of the remaining sequences, two sets of hits were defined: a strict set containing only hits sharing 19% sequence identity to the query or more (P_{FRH5}ΔNL_{HS1}), and a permissive set containing all remaining hits (8 and 14 hits, respectively, including the query sequence) (P_{FRH5}ΔNL_{HS2} and P_{FRH5}ΔNL_{HS3}). The strict alignment contained the following UniProt entries: Q8IFM5 (PDB ID code 4WAT), A0A078K5N4, B4X6H6, K6VIX0, A0A060RXZ9, W7J6M4, A0A024WYW5, and Q7YWE8; the permissive alignment contained the entries in the strict alignment and the following additional entries: I6QQT7, C1IW27, I6RGY9, A5K940, A5K696, and A0A060S1Z4. MUSCLE (30) was used with default parameters to derive a multiple sequence alignment from each set of hits.

PROSS Stability Design. To preserve the P_{FRH5} binding interface with its natural target basigin [PDB ID code 4U0Q (11)] and with two neutralizing antibodies, 9AD4 and QA1 [PDB ID codes 4U0R and 4U1G (11), respectively], 63 residues within 5 Å of all three interfaces were held fixed throughout all Rosetta simulations (Table S1).

Two independent runs of the PROSS algorithm (22) were carried out based on the strict and permissive alignments. As a template structure, we used a version of P_{FRH5} with both the flexible N terminus (residues 1–140) and a disordered loop (residues 248–296) removed (P_{FRH5}ΔNL) (11). Because P_{FRH5} is large (>400 aa), we decided to experimentally test designs with 15 mutations or more. Using the strict alignment, only one design had more than 15 mutations (P_{FRH5}ΔNL_{HS1}), and using the permissive alignment, two designs were selected for experimental testing based on visual inspection (P_{FRH5}ΔNL_{HS2} and P_{FRH5}ΔNL_{HS3}) (Table S1). Datasets S1 and S2 contain the strict and permissive alignments, respectively, as well as the three design variants.

Design Model Analysis. Sequence and structural features of P_{FRH5}ΔNL_{HS1} were compared with the sequence and structure of wild-type P_{FRH5}ΔNL. Mutations were defined as improving helical propensity if they were in helical regions and if Rosetta energy calculations showed ΔΔ*G*_{calc} of less than −0.15 Rosetta energy units for the energy term that accounts for the compatibility of the amino acid identity with the local backbone ϕ and ψ dihedral angles (p_{aa}pp). Positions were defined as buried if they had >21 and >75 neighboring nonhydrogen atoms within 10 and 12 Å, respectively, according to the Rosetta Features Reporter (31, 32).

Expression, Purification, and Characterization of Thermally Stabilized RH5 Variants. A gene for *Plasmodium falciparum* RH5 spanning from K141 to Q526 with both the flexible N terminus (residues 1–140) and a disordered loop (residues 248–296) removed (P_{FRH5}ΔNL) was available from a previous study (11). Synthetic genes were designed to express the equivalent regions from the thermally stabilized variants with codon use optimized for *Drosophila melanogaster* expression. They were expressed in the Rosetta-gami strain of *E. coli* and were purified using immobilized affinity and size exclusion chromatography. P_{FRH5}ΔNL and basigin were expressed as described (11). Determination of the structure of the P_{FRH5}ΔNL_{HS1}:9AD4 complex and its biophysical and immunological characterization were characterized using protocols similar to those described (11) and are detailed in *SI Materials and Methods*. All procedures on mice were performed in accordance with the terms of the UK Animals (Scientific Procedures) Act Project License and were approved by the University of Oxford Animal Welfare and Ethical Review Body.

ACKNOWLEDGMENTS. We acknowledge the beamline staff of I04-1 at the Diamond Light Source synchrotron for their support during data collection. We are grateful for the assistance of Julie Furze, David Pattinson, Jing Jin, and Alexander Douglas (Jenner Institute, University of Oxford); Wian de Jongh (ExpreS²ion Biotechnology); and David Staunton and Ed Lowe (Department of Biochemistry, University of Oxford). M.K.H. and I.C. are supported by a Wellcome Trust Investigator Award. The S.J.F. laboratory is supported by a Starter Grant from the European Research Council, the Israel Science Foundation through an individual grant, its Center of Research Excellence in Structural Cell Biology, and the joint Israel-India program, and a charitable donation from Sam Switzer and family. S.J.D. is a Jenner Investigator, a Lister Institute Research Prize Fellow, and a Wellcome Trust Senior Fellow (106917/Z/15/Z).

1. World Health Organization (2015) *World Malaria Report 2015* (World Health Organization, Geneva).
2. Birkett AJ, Moorthy VS, Loucq C, Chitnis CE, Kaslow DC (2013) Malaria vaccine R&D in the decade of vaccines: Breakthroughs, challenges and opportunities. *Vaccine* 31(Suppl 2):B233–B243.
3. RTS,S Clinical Trials Partnership (2015) Efficacy and safety of RTS,S/AS01 malaria vaccine with or without a booster dose in infants and children in Africa: Final results of a phase 3, individually randomised, controlled trial. *Lancet* 386(9988):31–45.

4. Draper SJ, et al. (2015) Recent advances in recombinant protein-based malaria vaccines. *Vaccine* 33(52):7433–7443.
5. Douglas AD, et al. (2011) The blood-stage malaria antigen P_{FRH5} is susceptible to vaccine-inducible cross-strain neutralizing antibody. *Nat Commun* 2:601.
6. Baum J, et al. (2009) Reticulocyte-binding protein homologue 5—an essential adhesin involved in invasion of human erythrocytes by *Plasmodium falciparum*. *Int J Parasitol* 39(3):371–380.

7. Crosnier C, et al. (2011) Basigin is a receptor essential for erythrocyte invasion by *Plasmodium falciparum*. *Nature* 480(7378):534–537.
8. Bustamante LY, et al. (2013) A full-length recombinant *Plasmodium falciparum* PfrH5 protein induces inhibitory antibodies that are effective across common PfrH5 genetic variants. *Vaccine* 31(2):373–379.
9. Williams AR, et al. (2012) Enhancing blockade of *Plasmodium falciparum* erythrocyte invasion: Assessing combinations of antibodies against PfrH5 and other merozoite antigens. *PLoS Pathog* 8(11):e1002991.
10. Manske M, et al. (2012) Analysis of *Plasmodium falciparum* diversity in natural infections by deep sequencing. *Nature* 487(7407):375–379.
11. Wright KE, et al. (2014) Structure of malaria invasion protein RH5 with erythrocyte basigin and blocking antibodies. *Nature* 515(7527):427–430.
12. Douglas AD, et al. (2015) A PfrH5-based vaccine is efficacious against heterologous strain blood-stage *Plasmodium falciparum* infection in Aotus monkeys. *Cell Host Microbe* 17(1):130–139.
13. Douglas AD, et al. (2014) Neutralization of *Plasmodium falciparum* merozoites by antibodies against PfrH5. *J Immunol* 192(1):245–258.
14. Hjerrild KA, et al. (2016) Production of full-length soluble *Plasmodium falciparum* RH5 protein vaccine using a *Drosophila melanogaster* Schneider 2 stable cell line system. *Sci Rep* 6:30357.
15. Joyce MG, et al. (2016) Iterative structure-based improvement of a fusion-glycoprotein vaccine against RSV. *Nat Struct Mol Biol* 23(9):811–820.
16. Stewart-Jones GB, et al. (2015) A cysteine zipper stabilizes a pre-fusion F glycoprotein vaccine for respiratory syncytial virus. *PLoS One* 10(6):e0128779.
17. Kwon YD, et al. (2015) Crystal structure, conformational fixation and entry-related interactions of mature ligand-free HIV-1 Env. *Nat Struct Mol Biol* 22(7):522–531.
18. Ingale J, et al. (2014) Hyperglycosylated stable core immunogens designed to present the CD4 binding site are preferentially recognized by broadly neutralizing antibodies. *J Virol* 88(24):14002–14016.
19. Steichen JM, et al. (2016) HIV vaccine design to target germline precursors of glycan-dependent broadly neutralizing antibodies. *Immunity* 45(3):483–496.
20. Sanders RW, et al. (2002) Stabilization of the soluble, cleaved, trimeric form of the envelope glycoprotein complex of human immunodeficiency virus type 1. *J Virol* 76(17):8875–8889.
21. Azoitei ML, et al. (2011) Computation-guided backbone grafting of a discontinuous motif onto a protein scaffold. *Science* 334(6054):373–376.
22. Goldenzweig A, et al. (2016) Automated structure- and sequence-based design of proteins for high bacterial expression and stability. *Mol Cell* 63(2):337–346.
23. Chen L, et al. (2014) Crystal structure of PfrH5, an essential *P. falciparum* ligand for invasion of human erythrocytes. *eLife* 3:3.
24. Kramer RM, Shende VR, Motl N, Pace CN, Scholtz JM (2012) Toward a molecular understanding of protein solubility: Increased negative surface charge correlates with increased solubility. *Biophys J* 102(8):1907–1915.
25. Chan P, Curtis RA, Warwicker J (2013) Soluble expression of proteins correlates with a lack of positively-charged surface. *Sci Rep* 3:3333.
26. Lau CK, et al. (2015) Structural conservation despite huge sequence diversity allows EPCR binding by the PfEMP1 family implicated in severe childhood malaria. *Cell Host Microbe* 17(1):118–129.
27. Angermüller C, Biegert A, Söding J (2012) Discriminative modelling of context-specific amino acid substitution probabilities. *Bioinformatics* 28(24):3240–3247.
28. Biegert A, Söding J (2009) Sequence context-specific profiles for homology searching. *Proc Natl Acad Sci USA* 106(10):3770–3775.
29. Li W, Godzik A (2006) Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22(13):1658–1659.
30. Edgar RC (2004) MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5:113.
31. Leaver-Fay A, et al. (2013) Scientific benchmarks for guiding macromolecular energy function improvement. *Methods Enzymol* 523:109–143.
32. O'Meara MJ, et al. (2015) Combined covalent-electrostatic model of hydrogen bonding improves structure prediction with Rosetta. *J Chem Theory Comput* 11(2):609–622.
33. Walter TS, et al. (2006) Lysine methylation as a routine rescue strategy for protein crystallization. *Structure* 14(11):1617–1622.
34. Miura K, et al. (2009) Anti-apical-membrane-antigen-1 antibody is more effective than anti-42-kilodalton-merozoite-surface-protein-1 antibody in inhibiting *plasmodium falciparum* growth, as determined by the in vitro growth inhibition assay. *Clin Vaccine Immunol* 16(7):963–968.
35. Kabsch W (2010) Xds. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):125–132.
36. Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* 62(Pt 1):72–82.
37. Winn MD, et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):235–242.
38. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Cryst* 40(Pt 4):658–674.
39. Vagin AA, et al. (2004) REFMAC5 dictionary: Organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2184–2195.
40. Bricogne G, et al. (2011) *BUSTER, Version 2.10.0* (Global Phasing, Cambridge, UK).
41. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486–501.

Supporting Information

Campeotto et al. 10.1073/pnas.1616903114

SI Materials and Methods

Collection and Filtering of Homologous Sequences. The PfrH5 structure was downloaded from the Protein Data Bank (PDB) [ID code 4WAT (23)], and homologous sequences were collected using CSI-BLAST (27, 28) to search the nonredundant (nr) database in May 2015, with e value $< 10^{-4}$, three iterations, a maximum of 500 sequences, and default values on all other parameters. Hits were clustered using cd-hit (29) at 98% threshold and default parameters. Hits from genera other than *Plasmodium* were excluded. Hits were also excluded if their sequence identity to the query was lower than 15% or if they showed more than 1% gaps in the aligned segment. Of the remaining sequences, two sets of hits were defined: a “strict” set containing only hits sharing 19% sequence identity to the query or more (PfrH5 Δ NL_{HS1}), and a “permissive” set containing all remaining hits (8 and 14 hits, respectively, including the query sequence) (PfrH5 Δ NL_{HS2} and PfrH5 Δ NL_{HS3}). The strict alignment contained the following UniProt entries: Q8IFM5 (PDB ID code 4WAT), A0A078K5N4, B4X6H6, K6VIX0, A0A060RXZ9, W7J6M4, A0A024WYW5, and Q7YWE8; the permissive alignment contained the entries in the strict alignment and the following additional entries: I6QQT7, C1IW27, I6RGY9, A5K940, A5K696, and A0A060S1Z4. MUSCLE (30) was used with default parameters to derive a multiple sequence alignment from each set of hits.

PROSS Stability Design. To preserve the PfrH5 binding interface with its natural target basigin [PDB ID code 4U0Q (11)] and with two neutralizing antibodies, 9AD4 and QA1 [PDB ID codes 4U0R and 4U1G (11), respectively], 63 residues within 5 Å of all three interfaces were held fixed throughout all Rosetta simulations (Table S1).

Two independent runs of the PROSS algorithm (22) were carried out based on the strict and permissive alignments. As a template structure, we used a version of PfrH5 with both the flexible N terminus (residues 1–140) and a disordered loop (residues 248–296) removed (PfrH5 Δ NL) (11). Because PfrH5 is large (>400 aa), we decided to experimentally test designs with 15 mutations or more. Using the strict alignment, only one design had more than 15 mutations (PfrH5 Δ NL_{HS1}), and using the permissive alignment two designs were selected for experimental testing based on visual inspection (PfrH5 Δ NL_{HS2} and PfrH5 Δ NL_{HS3}) (Table S1).

Design Model Analysis. Sequence and structural features of PfrH5 Δ NL_{HS1} were compared with the sequence and structure of wild-type PfrH5 Δ NL. Mutations were defined as improving helical propensity if they were in helical regions and if Rosetta energy calculations showed $\Delta\Delta G_{calc}$ less than -0.15 Rosetta energy units for the energy term that accounts for the compatibility of the amino acid identity with the local backbone ϕ and ψ dihedral angles (p_aa_pp). Positions were defined as buried if they had >21 and >75 neighboring nonhydrogen atoms within 10 and 12 Å, respectively, according to the Rosetta Features Reporter (31, 32).

Expression and Purification of Thermally Stabilized RH5 in *E. coli*. A gene for *Plasmodium falciparum* RH5 spanning from K141 to Q526 with both the flexible N terminus (residues 1–140) and a disordered loop (residues 248–296) removed (PfrH5 Δ NL) was available from a previous study (11). The gene had been codon optimized for expression in *Drosophila melanogaster* as a C-terminal hexahistidine-tagged protein.

Synthetic genes of the PfrH5 variants were designed to match the boundaries of PfrH5 Δ NL and were codon optimized for expression

in *Drosophila melanogaster*, giving the constructs PfrH5 Δ NL_{HS1}, PfrH5 Δ NL_{HS2}, and PfrH5 Δ NL_{HS3}. PCRs were performed using the same primers for all of the variants and for PfrH5 Δ NL, and products were cloned into a modified pET15b vector (Novagen) encoding an N-terminal hexahistidine tag and tobacco etch virus (TEV) protease cleavage site. This cloning strategy generated hexahistidine-tagged constructs pET15b-PfrH5 Δ NL, pET15b-PfrH5 Δ NL_{HS1}, pET15b-PfrH5 Δ NL_{HS2}, and pET15b-PfrH5 Δ NL_{HS3}.

These constructs were introduced into the Rosetta-gami B(DE3) pLysS *E. coli* expression strain (Novagen). Bacterial cultures were grown overnight at 37 °C in 50 μ g/mL ampicillin, 30 μ g/mL kanamycin, 34 μ g/mL chloramphenicol, and 12.5 μ g/mL tetracycline, and inoculated in a 1:50 ratio into 1 L of fresh 2xYT medium supplemented with the same antibiotic concentrations. Cultures were grown at 37 °C until an OD₆₀₀ of 0.5–0.7 and gene expression was induced by addition of 0.5 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) for 4 h at 37 °C. Cultures were pelleted at 4,000 $\times g$ for 10 min and resuspended in 50 mM Tris-HCl, pH 7.5, 200 mM NaCl (Buffer A) supplemented with EDTA-free protease inhibitors (Roche). Bacterial cell lysis was performed using a cell disruptor at a pressure of 30 kpsi, followed by centrifugation at 20,000 $\times g$ for 40 min. The soluble fraction of each construct was loaded into a gravity flow column containing nickel-nitrilotriacetic acid resin (Ni²⁺-NTA) (Qiagen) previously equilibrated with Buffer A. The Ni²⁺-NTA resin was washed with 10 column volumes (CV) of Buffer A containing 30 mM imidazole, while the proteins were eluted in 3 CV of Buffer A containing 0.7 M imidazole. The imidazole was removed by buffer exchange in PBS using a PD-10 desalting column (GE Healthcare), and protein analysis was performed by Western blot using a rabbit antibody anti-hexahistidine (Abcam) and a goat anti-rabbit HRP-conjugated antibody (Abcam). As the variant PfrH5 Δ NL_{HS1} showed the highest level of soluble protein expression, only this construct was scaled up to expression in 6-L scale and purification as above. After Ni²⁺-NTA affinity purification, PfrH5 Δ NL_{HS1} was further purified by size exclusion chromatography using a Superdex 200 16/60 column (GE Healthcare) previously equilibrated with Buffer A. Size exclusion column multiangle light scattering (SEC-MALS) analysis was performed using a S200 10/300 column (GE Healthcare), and data were analyzed using ORYX software (Oryx Systems).

Expression of PfrH5 Δ NL and PfrH5 Δ NL_{HS1} in S2 Cells. The pExpreS²-2.1 vector containing wild-type PfrH5 for expression in *Drosophila* S2 cells (pExpreS²-2.1-PfrH5 Δ NL) was already available from a previous study (11). The PfrH5 Δ NL_{HS1} gene was cloned into the pExpreS²-2.1 vector using the Gibson cloning kit (NEB), giving the construct pExpreS²-2.1-PfrH5 Δ NL_{HS1}. S2 cells were transfected with pExpreS²-2.1-PfrH5 Δ NL_{HS1} as recommended by the supplier (ExpreS²ion Biotechnologies). Briefly, after thawing, S2 cells were grown to a density of 2.0×10^6 cells per mL in Ex-cell medium (SAFC Biosciences) supplemented with 10% (vol/vol) FCS (Life Technologies) and 10 mg/mL penicillin–streptomycin (Sigma) medium. The cells were then spun at 300 $\times g$ for 10 min and resuspended to a density of 2.0 million per milliliter in fresh Ex-cell medium with 10 mg/mL penicillin–streptomycin. Transfection was then performed by adding 12.5 μ g of DNA and 50 μ L of ExpreS² TR 5 \times reagent (ExpreS²ion Biotechnologies) to 2.5 mL of resuspended cells at 2.0 million cells per mL for 3 h at 25 °C. At the end of the incubation, FCS was added to the medium to a final concentration of 10% (vol/vol), and cells were transferred to a Cell-star cell culture flask (Greiner) and grown in a static incubator at

25 °C. At regular intervals of 4 d, one-half of the medium (2.5 mL) was removed and replaced with an equivalent amount of fresh Ex-cell medium containing 10 mg/mL penicillin-streptomycin and 10% (vol/vol) FCS. In the third week, 10 mL of Ex-cell medium supplemented with 2 mg/mL zeocin (Life Technologies) was added to the S2 cells upon transfer from Cellstar cell culture flask to a T25 flask. Cell growth was performed in a shaking incubator at 25 °C, and cells were expanded every 4 d by dilution through the addition of Ex-cell medium to a final density of 8 million cells per mL. When the desired volume of S2 cells was reached, they were pelleted at $300 \times g$ for 10 min and resuspended to a density of 8 million cells per mL in Ex-cell medium containing 10 mg/mL penicillin-streptomycin (Sigma) before growth for 5 d at 25 °C.

Detection of PfrH5 Proteins by Western Blot. At the end of the 5 d of expression, the cells were harvested at $2,000 \times g$ for 30 min, and protein expression was assessed by Western blot analysis using the supernatant after centrifugation, which was normalized between PfrH5 Δ NL and PfrH5 Δ NL_{HS1} on the basis of the cell density at the time of the cell harvesting. Western blot was performed using a rabbit anti-hexahistidine primary antibody (Abcam) and a goat anti-rabbit HRP-conjugated secondary antibody (Abcam). Western blot analysis was performed using a LI-COR system.

Expression and Purification of Basigin. Basigin was produced as described (11). In brief, residues 22–205 were expressed from a modified pET15b in bacterial strain Origami B (DE3) (Novagen) by incubation overnight at 25 °C after induction with 1 mM IPTG. The protein was purified by Ni²⁺-NTA (Qiagen) affinity chromatography, followed by buffer exchange into PBS using a PD-10 desalting column (GE Healthcare), and overnight cleavage with His-tagged TEV protease at 4 °C, before a second Ni²⁺-NTA column. The flow-through was concentrated using an Amicon Ultra centrifugal filter device (molecular mass cutoff, 3,000 Da). Finally, gel filtration was performed with a Superdex 200 16/60 column (GE Healthcare) in 20 mM Hepes (pH 7.5) and 150 mM NaCl.

Expression and Purification of 9AD4 Monoclonal Antibody for Crystallization. The anti-PfrH5 monoclonal antibody, 9AD4, was described previously (13). The hybridoma was grown in DMEM (Sigma) supplemented with 4 mM L-glutamine (Sigma), 0.01 M Hepes (Life Technologies), 100 U of penicillin, 0.1 mg/mL streptomycin (Sigma), and 20% (vol/vol) FCS (Gibco). It was then transferred into CD Hybridoma medium (Life Technologies) with glutamine, penicillin, and streptomycin. The cells were harvested after 7–10 d. The cell culture supernatant was exchanged into 20 mM phosphate, pH 7.0, with a tangential flow filtration device (Pall).

The sample was then loaded onto a HiTrap Protein G HP column (GE Healthcare), eluted in 0.1 M glycine-HCl (pH 3.0), and immediately neutralized with 0.1 M Tris (pH 8.0). The sample was exchanged into 100 mM phosphate (pH 6.4), 300 mM NaCl, 2 mM EDTA, 5 mM L-cysteine (pH 6.4), and 1.5 mM β -mercaptoethanol using PD-10 columns (GE Healthcare).

Antibody Fab fragments were generated by addition of papain agarose (Sigma) and overnight incubation at 37 °C. The papain agarose was removed by centrifugation, and the sample loaded onto a HiTrap Protein A HP column (GE Healthcare). Fab fragments did not bind to the column, and were gel filtered on a Superdex 200 16/60 column (GE Healthcare) in 20 mM Hepes (pH 7.5) and 150 mM NaCl.

Heat Treatment and Lyophilization of PfrH5 Δ NL and PfrH5 Δ NL_{HS1} Variants. Purified protein samples of PfrH5 Δ NL and PfrH5 Δ NL_{HS1} variants were normalized to a final 16 μ M concentration and divided into 100- μ L aliquots. Each aliquot was incubated for 1 h in a PCR machine at the fixed temperature of 40, 45, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, and 75 °C. Each aliquot was flash cooled in liquid N₂ at the end of the heat treatment, and all of the aliquots were

defrosted at the same time at room temperature before analysis by surface plasmon resonance (SPR). An additional aliquot of each protein sample was flash-cooled in liquid N₂ and lyophilized overnight. The samples were resuspended in the same volume in deionized filtered water before SPR experiments.

Surface Plasmon Resonance. SPR experiments were carried out using a BIAcore T200 instrument (GE Healthcare). Experiments were performed at 20 °C using a buffer containing 20 mM Hepes (pH 7.4), 150 mM NaCl, 3 mM EDTA, 0.005% Tween 20, 2 mg/mL dextran, and 1 mg/mL salmon sperm DNA (Sigma-Aldrich). Basigin was immobilized on a CM5 chip (GE Healthcare) by amine coupling (GE Healthcare kit) to a total of 800 RU.

Lysine methylation of gel exclusion purified PfrH5 Δ NL and PfrH5 Δ NL_{HS1} was performed by incubation with 0.02 M dimethylamine-borane complex (ABC) and 0.04 M formaldehyde at 4 °C for 1 h in the dark (33). After 1 h, an additional 0.02 M ABC and 0.04 M formaldehyde were added for a further hour. Finally, 0.01 M ABC was added and the reaction was incubated at 4 °C overnight in the dark. The reaction mixture was quenched by addition of Tris-HCl (pH 7.5) to a final concentration of 50 mM, and the proteins were injected into a S200 16/60 column previously equilibrated with 20 mM Hepes, pH 7.5, and 150 mM NaCl. Gel-filtered protein was concentrated to 16 μ M using an Amicon 10-kDa cutoff concentrator (Millipore). A concentration series of PfrH5 Δ NL or PfrH5 Δ NL_{HS1} (8, 4, 2, 1, 0.5, 0.25, 0.125, and 0.0625 μ M) was injected over the basigin-coated chip for 120 s at 30 μ L/min, followed by a 300-s dissociation time. The chip surface was then regenerated with 30 s of 2 M NaCl. Specific binding was obtained by subtracting the response from a blank surface from that of the basigin-coated surface. The kinetic sensorgrams were fitted to a global 1:1 interaction model, allowing determination of the dissociation constant, K_D , using BIAevaluation software 1.0 (GE Healthcare). Data were plotted using the software GraphPad 7.0 (Prism).

Circular Dichroism Analysis. PfrH5 Δ NL and PfrH5 Δ NL_{HS1} proteins were buffer exchanged into 20 mM phosphate, pH 7.5, 150 mM NaF, using a PD-10 desalting column (GE Healthcare) and concentrated to 50 μ g/mL. Circular dichroism (CD) spectra were recorded between 200 and 250 nm with a temperature ramp increasing by increments of 2 °C from 20 to 90 °C. Data analysis was performed using the software GraphPad 7.0 (Prism).

Mouse Immunizations and Antibodies. Six-week-old BALB/c mice (Harlan) were immunized intramuscularly with two doses of 20 μ g of protein in Addavax (Invivogen) 2 wk apart. Blood samples were harvested by exsanguination 2 wk after the second immunization. Polyclonal IgG was purified from pooled serum samples on Pierce Protein G agarose (Thermo Fisher Scientific) and buffer exchanged using Amicon Ultra-15 Centrifugal Filter Units (Merck Millipore) into incomplete *P. falciparum* culture media. IgG concentration was determined using a Nanodrop 2000 (Thermo Scientific). All procedures on mice were performed in accordance with the terms of the UK Animals (Scientific Procedures) Act Project License and were approved by the University of Oxford Animal Welfare and Ethical Review Body.

In Vitro Assay of Growth Inhibitory Activity. In vitro-cultured *P. falciparum* (clone 3D7) was grown in O Rh⁺ erythrocytes and 10% (vol/vol) human serum. The previously described standardized growth inhibitory activity (GIA) assay from the NIH GIA Reference Center was used to assess the ability of antibodies to inhibit *P. falciparum* growth in vitro (13, 34). Briefly, each test IgG was incubated with synchronized *P. falciparum* parasites for a single growth cycle, and relative parasitemia levels were quantified by biochemical determination of parasite lactate dehydrogenase. The purified polyclonal IgG was tested in triplicate in a twofold dilution series from 2 mg/mL down to 8 μ g/mL. The data points are

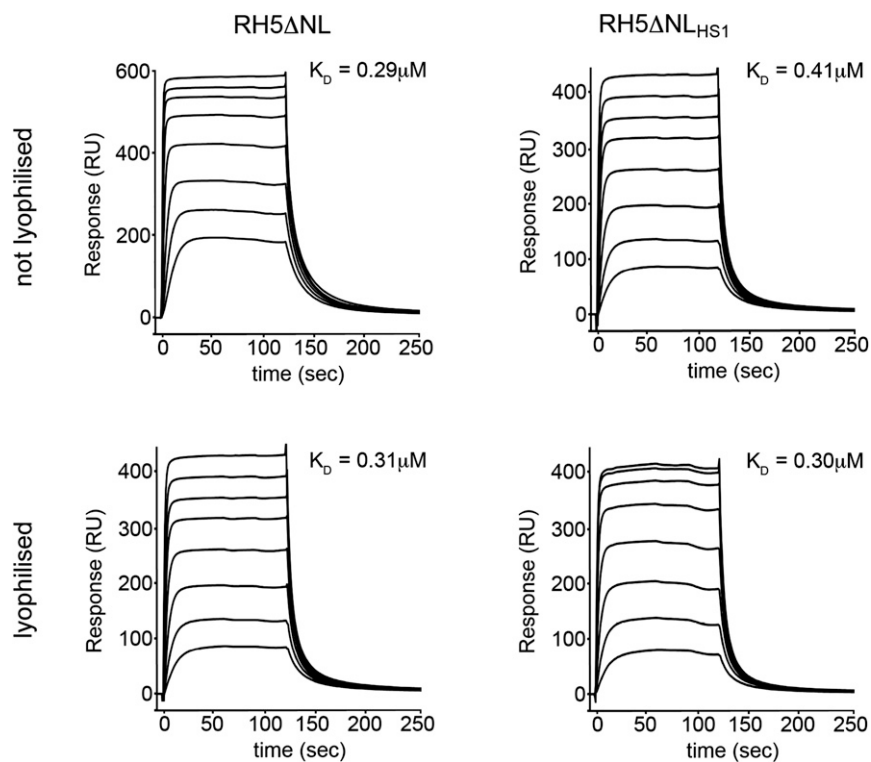


Fig. S2. SPR characterization of PfrH5ΔNL and PfrH5ΔNL_{HS1}. SPR analysis of the binding of twofold dilutions of PfrH5ΔNL and PfrH5ΔNL_{HS1} to basigin, from a maximal concentration of 8 μM. The same analysis was performed after lyophilizing both PfrH5ΔNL and PfrH5ΔNL_{HS1} before measurement.

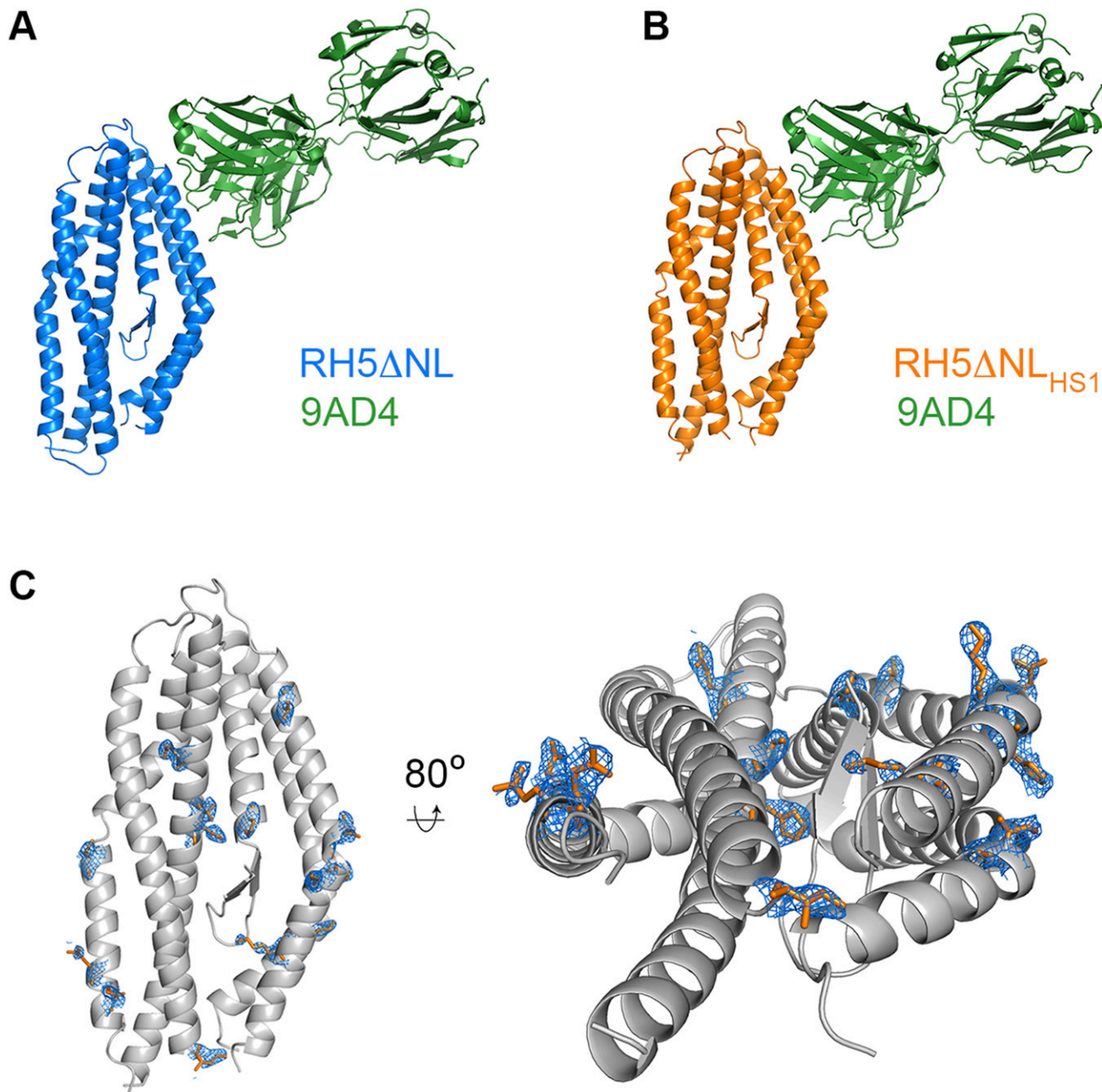


Fig. S3. Structural characterization of PfrH5ΔNL and PfrH5ΔNL_{H51}. The structures of (A) PfrH5ΔNL and (B) PfrH5ΔNL_{H51}, bound to the Fab fragment of the monoclonal antibody 9AD4. (C) The structure of PfrH5ΔNL_{H51} in gray, with residues that are different to those in PfrH5ΔNL shown as orange sticks. A composite omit map, contoured at 1.0 sigma level, shows the electron density for the residues that differ between PfrH5ΔNL and PfrH5ΔNL_{H51}.

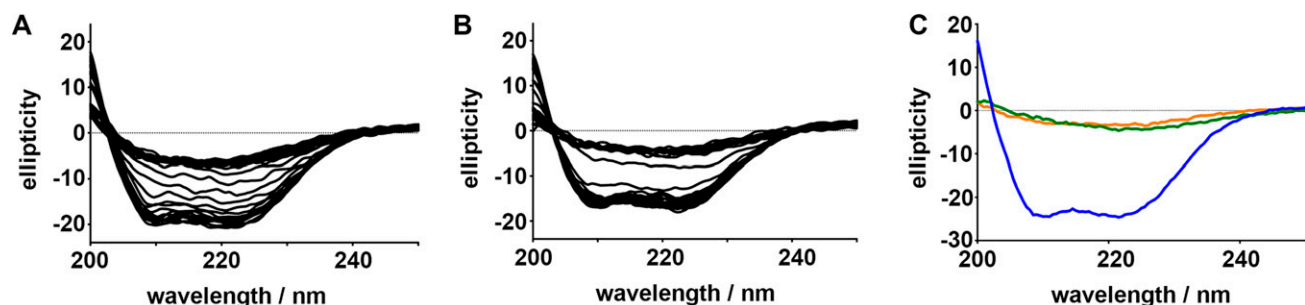


Fig. S4. CD analysis of PfrH5ΔNL and PfrH5ΔNL_{H51}. Determination of the effect of temperature on the molecular ellipticity of (A) PfrH5ΔNL and (B) PfrH5ΔNL_{H51}. (C) Thermal denaturation of PfrH5ΔNL_{H51} is irreversible. CD spectra of PfrH5ΔNL_{H51} at 20 °C (blue), at 75 °C (green), and after heating to 75 °C and returning to 20 °C (orange).

Table S1. Computational parameters used to generate designed variants and lists of mutations

Designed variant	Positions held fixed ^{*,†}	$\Delta\Delta G_{calc}$ cutoffs, R.e.u. [‡]	Mutations [*]
PfrH5ΔNL _{H51} (strict MSA)	147, 149, 193, 196, 202, 205, 206, 209, 212, 213, 216, 327, 328, 331, 334, 335, 337, 338, 339, 340, 341, 342, 344, 452, 456, 194, 197, 200, 201, 202, 203, 204, 207, 222, 225, 226, 242, 243, 244, 245, 246, 247, 248, 249,	−0.45 (18)	I157L [§] , D183E, A233K, M304F, K312N, L314F, K316N, M330N, S370A, S381N, T384K, L392K, T395N, N398E, R458K, N464K, S467A, F505L
PfrH5ΔNL _{H52} (permissive MSA)	250, 346, 347, 349, 350, 352, 357, 358, 362, 447, 448, 449, 496, 198, 343, 345, 348, 353, 451	−0.75 (25)	D183E, N191I, S192A, A233N, K236H, N308K, L314F, L369N, S370N, S381N, T384K, L392D, T395R, N398K, H414L, L444E, N445D, S463A, N464K, S467A, I470R, H474D, H495N, F505L, K511P
PfrH5ΔNL _{H53} (permissive MSA)		−1.25 (15)	D183E, N191I, S192A, A233N, L369N, S381N, L392D, N398K, N445D, S463A, N464K, S467A, I470R, H474D, K511P

*Numbering according to PDB ID code 4WAT.

[†]Residues within 5 Å of either basigin, QA1, or 9AD4 were held fixed through all design simulations.

[‡]PROSS energy threshold of designs that were selected for experimental validation. The cutoffs defined the subsets of individually stabilizing mutations. Each defined subset resulted in a single variant. R.e.u. refers to Rosetta energy units. Values in parentheses show the number of mutations relative to PfrH5ΔNL.

[§]Two sequences contribute information at position 157, and both have ILE at this position. LEU was allowed and selected because the PSSM accounts also for BLOSUM62 statistics.

Table S2. Data collection and refinement statistics

9AD4–PfRH5ΔNL _{HS1} complex	
Data collection	
Space group	<i>P</i> 2 ₁
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> , Å	39.66, 85.55, 132.87
α , β , γ , °	90.00, 90.91, 90.00
Resolution, Å	42.77 (2.43–2.35)*
<i>R</i> _{sym} or <i>R</i> _{merge}	0.096 (0.554)
<i>I</i> / σ (<i>I</i>)	13.9 (3.1)
Completeness, %	98.6 (97.9)
Redundancy	6.5 (6.9)
Refinement	
No. reflections	69,953
<i>R</i> _{work}	0.176
<i>R</i> _{free}	0.222
No. atoms	6,079
Protein	5,747
Ligand/ion	2
Water	330
R.m.s. deviations	
Bond lengths, Å	0.010
Bond angles, °	1.11
Ramachandran plot	
Allowed region	93.6%
Additional allowed	6.0%
Generously allowed	0.3%
Disallowed	0%
PDB ID code	5MI0

*Values in parentheses are for highest-resolution shell.

Table S3. PfRH5ΔNL_{HS1} structure features compared with PfRH5ΔNL

Structural feature	Contributing mutations*
Mutations at buried positions	D183E, L314F, S286A
Mutations improving helical propensity	I157L, D183E, S370A, T384K, N398E, N464K, S467A
Surface mutations increasing polarity	A233K, M330N, S381A, T384K, L392K, T395N, N398E, N464K
Mutations in homogenous positive patch	K312N, K316N

*Numbering according to PDB ID code 4WAT.

Other Supporting Information Files

[Dataset S1 \(TXT\)](#)

[Dataset S2 \(TXT\)](#)