Design of a basigin-mimicking inhibitor targeting the malaria invasion protein RH5

Shira Warszawski[1], Elya Dekel[1], Ivan Campeotto[2], Jennifer M. Marshall[3], Katherine E. Wright[4], Oliver Lyth[4], Orli Knop[1], Neta Regev-Rudzki[1], Matthew K Higgins[2], Simon J Draper[3], Jake Baum[4], and Sarel J Fleishman[1*]

[1]Department of Biomolecular Sciences, Weizmann Institute of Science, Rehovot 7610001, Israel
[2]Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK
[3]Jenner Institute, University of Oxford, Old Road Campus Research Building, Oxford, OX3 7DQ, UK
[4]Department of Life Sciences, Imperial College London, Sir Alexander Fleming Building, Exhibition Road, South Kensington, London, UK
Correspondence: sarel@weizmann.ac.il
Keywords: High-affinity design; Rosetta; deep sequencing; host-pathogen interactions; *Plasmodium falciparum*

Abstract

Many human pathogens use host cell-surface receptors to attach and invade cells. Often, the host-pathogen interaction affinity is low, presenting opportunities to block invasion using a soluble, high-affinity mimic of the host protein. The Plasmodium falciparum reticulocyte-binding protein homolog 5 (RH5) provides an exciting candidate for mimicry: it is highly conserved and its moderate affinity binding to the human receptor basigin (K_D≥1 μM) is an essential step in erythrocyte invasion by this malaria parasite. We used deep mutational scanning of a soluble fragment of human basigin to systematically characterize point mutations that enhance basigin affinity for RH5 and then used Rosetta to design a variant within the sequence space of affinity-enhancing mutations. The resulting seven-mutation design exhibited 1,900-fold higher affinity (K_D approximately 1nM) for RH5 with a very slow binding off rate (0.23 h^{-1}) and reduced the effective Plasmodium growth-inhibitory concentration by at least tenfold compared to human basigin. The design provides a favorable starting point for engineering on-rate improvements that are likely to be essential to reach therapeutically effective growth inhibition.

Introduction

Malaria is the deadliest parasitic disease with over 200 million clinical cases annually[1]. The unicellular parasite, *Plasmodium falciparum*, causes the most fatal and severe form of the disease, and the emergence of *P. falciparum* mutants that are resistant to frontline therapeutic interventions is a major cause of concern[2]. New modes of blocking the parasite are therefore urgently needed. Invasion of host erythrocytes is essential to the life cycle of *Plasmodium* parasites providing opportunities for inhibiting pathogenesis. The interaction between the *P. falciparum* reticulocyte-binding protein homolog 5 (RH5) and the human erythrocyte-surface receptor basigin was shown to be absolutely essential to erythrocyte invasion and *P. falciparum* pathogenesis[3–6]. Indeed, RH5 is highly conserved among all *P. falciparum* field isolates[6–10]; antibodies that target RH5 inhibit pathogenesis; inoculation with RH5 in model animals provides protective immunity[4,6–8,10–12]; and clinical trials are underway to test the efficacy of RH5 subunit vaccines in humans[13].

The molecular structures of RH5 in complex with the extracellular fragment of basigin and with neutralizing antibodies are known[10,14], enabling the computational design of improved vaccine immunogens based on RH5[15]. RH5 folds into a novel kite-like structure, and basigin interacts at a tip of RH5 through two immunoglobulin-like domains that are connected by a flexible linker[10]. Although the interaction surface area between RH5 and basigin (~1,900 Å^2) is above average for protein-protein interactions[16], the molecular contacts between RH5 and basigin are tenuous, comprising mainly hydrogen bonds. Furthermore, the interaction surface on basigin comprises the flexible hinge region that connects the two extracellular immunoglobulin-like domains, potentially reducing affinity due to the loss in conformational freedom upon binding. Accordingly, the interaction affinity between RH5 and basigin is low ($K_D \geq 1$ μM)[10,17], and the application of the extracellular domain blocks invasion in growth assays only at high and therapeutically unrealistic concentrations ($IC_{50} > 10$ μM)[8]. Animal immunizations using RH5 led to the isolation of several high-affinity anti-RH5 antibodies that exhibited much improved growth-inhibitory concentration (IC_{50}=0.1 μM) compared to basigin[12], demonstrating that

binding to RH5 could provide a therapeutically relevant route. Molecular structures, however, revealed that these antibodies interacted with RH5 through surfaces that only partially overlapped with the basigin-binding surface[10], making them susceptible to the emergence of RH5 escape mutants.

The current study seeks to design a variant of basigin that targets the same site on RH5 as basigin but interacts with substantially higher affinity. Such a basigin mimic may have several advantages over antibodies: (1) by interacting with the essential surface of RH5, it would limit the potential emergence of RH5 escape mutants; (2) it would be minimally mutated relative to human basigin, exhibiting a minor risk of immunogenicity; and (3) basigin interacts with RH5 through a 25 kDa fragment, compared to a typical size of 150 kDa for IgGs. Thus, although such an inhibitor would not exhibit the avidity or Fc-mediated immunological response seen in dimeric IgGs, it would require the administration of lower protein mass.

RH5-binding affinity, however, is not the only consideration in developing an effective inhibitor, and binding kinetics are likely to play a major role in determining the neutralization efficacy. Indeed, RH5 is not localized to the parasite cell surface and is only briefly accessible when the parasite and red blood cell are attached[12]. It is therefore thought that the window of opportunity for binding RH5 and blocking invasion may be as short as 10-30 seconds and therefore, among RH5-targeting antibodies, growth inhibition correlates with binding on rates[12].

In recent years, deep mutational scanning has emerged as a powerful strategy to identify affinity-enhancing mutations[18,19]. In this approach, positions on the binder are individually mutated to all possible amino acid identities, combined into a cell-display library and subjected to high-throughput screening for binding to the target molecule, followed by deep sequencing to map the relative contribution of each mutation to affinity. Typically, following mapping, another genetic library is encoded with a set of the affinity-enhancing mutations, and selections isolate multipoint mutants with an improved affinity[19–21]. Here, we used deep mutational scanning to identify mutations on the surface of basigin that would individually increase affinity for RH5. Next, we used Rosetta to design a combination of these affinity-enhancing mutations that would

maximally increase affinity for RH5. Testing only a single design encoding seven mutations, we surprisingly found that this design resulted in a remarkable 1,900-fold higher affinity for RH5. This design provides a proof-of-principle for the rapid generation of high-affinity soluble blockers of essential host-pathogen protein-protein interactions.

Materials & methods

Basigin genetic library construction

Forward and reverse primers with the degenerate codon NNS were generated for the 62 positions on the soluble basigin fragment, essentially as described[22]. Primers were ordered from Sigma (Sigma-Aldrich, Rehovot, Israel) and used to introduce all possible amino acids per position by QuickChange mutagenesis[23]. Next, the PCR product of each position was transformed into yeast (EBY100 cells) and plated on SD-Trp$^-$ as described[24]. As described[25], plates with more than 400 colonies were scraped with 1 ml SDCAA, 50 µl was added to 5 ml SDCAA tube and cells were then grown at 30℃ overnight. The point mutants were divided into four libraries, corresponding to positions that were at most 130 bp apart from one another to enable deep mutational scanning using 150 bp reads.

Yeast surface display selection

Yeast-display experiments were conducted essentially as described [25]. Briefly, yeast cells were grown in selective medium SDCAA overnight at 30℃. The cells were then resuspended in 10 ml induction medium and incubated at 20℃ for 20 h. 10^7 cells were then used for yeast-cell surface display experiments: cells were subjected to primary antibody (mouse monoclonal IgG1 anti-c-Myc (9E10) sc-40, Santa Cruz Biotechnology) for expression monitoring and biotinylated ligand at 725 nM *Pf*RH5ΔNL[10] in PBS-F for 30 min at room temperature. The cells then underwent a second staining with fluorescently labeled secondary antibody (AlexaFluor488 - goat-anti-mouse IgG1 Life Technologies]) for scFv labeling, streptavidin-APC (SouthernBiotech) for ligand labeling for 10 min at 4℃. Next, cell fluorescence was measured, and cells were collected under sorting conditions for expression and top 15% binders. The selection gates were calibrated using the wild-type soluble BSG, and these gates were subsequently applied to the mutant library. Following fluorescence-activated cell sorting (FACS), cells were grown in SDCAA for 1-2 days and plasmids were extracted using Zymoprep Yeast Plasmid Miniprep II kit (Zymo Research).

DNA sequences of tested constructs

>BSG

GCGGGTACCGTTTTCACCACCGTAGAAGACCTGGGTTCTAAAATCCTGCTGACCTGCTCTCTGAACGACTCTGCGACCG
AAGTTACCGGTCACCGTTGGCTGAAAGGTGGTGTTGTTCTGAAAGAAGACGCGCTGCCGGGTCAGAAAACCGAATTCAA
AGTTGATTCTGATGACCAGTGGGGTGAATACTCTTGCGTTTTCCTGCCAGAACCGATGGGCACCGCGAACATCCAACTG
CACGGTCCGCCTCGTGTTAAAGCGGTTAAATCTTCTGAACACATCAACGAAGGTGAAACCGCGATGCTGGTATGCAAAT
CTGAATCTGTTCCGCCGGTTACCGACTGGGCGTGGTACAAAATCACCGACAGCGAAGACAAAGCGCTGATGAACGGTTC
CGAATCTCGTTTCTTTGTTTCTTCTTCTCAGGGTCGTTCTGAGCTGCACATCGAAAACCTGAACATGGAAGCGGACCCT
GGCCAATATCGTTGCAATGGCACCTCTTCTAAAGGTTCTGATCAAGCGATCATCACCCTGCGTGTTCGTTAATAG

>BSG^{des7}

GCGGGCACCATCTTCACTTCTGCGGAAGACCTGGGTTCTAAAATCCTGCTCACCTGTTCTCTGAACGACTCCGCGACTG
AAGTTACCGGTCATCGTTGGCTCAAGGGTGGTGTTGTTCTGAAAGAAGACGCGCTGCCAGGTCAGAAAACCGAATTCAA
AGTTGATTCTGACGACCAGTGGGGTGAATACTCTTGTGTTTTCCTCCCAGAGCCAATGGGTACCGCGAACATCATCCTG
TACGGTCCGCCACGTGTGAAAGCGGTTAAATCTTCTGAACACATCAACGAAGGTGAAACCGCGATGCTGGTTTGTAAAT
CTGAATCTGTTCCGCCGGTTACCGACTGGGCGTGGTATAAGATCACCGACTCTGAAGACAAAGCGCTCATGAACGGTTC
TGAGTCCCGCTTTTTCGTTTCTTCTTCCTTCGGTCGTTCCGAACTGCACATTGAAAACCTGAACATGGAAGCGGACCCA
GGCCAATATCGCTGTAACGGCACTTCTCAGAAGGGTTCCGACCAAGCGATCATCACCCTGCGTGTTCGTCTCGAGGGCG
GCGGATCCGAACAAAAGCTTATTTCTGAAGAGGACTTGTAATAG

DNA preparation for deep sequencing

To connect the DNA adaptors for deep sequencing, the plasmids extracted from the libraries were amplified using Phusion High-Fidelity DNA Polymerase (ThermoFisher) in a two-step PCR protocol.

PCR 1:

(barcode: CTCTTTCCCTACACGACGCTCTTCCGATCT)

>forward (seg1):**<barcode>**AGGGTCGGCTAGCCATATG

>forward (seg2):**<barcode>**GCCGGGTCAGAAAACCGAA

>forward (seg3):**<barcode>**ATCCAACTGCACGGTCCG

>forward (seg4):**<barcode>**GGTTCCGAATCTCGTTTCTTTG

>reverse: **CTGGAGTTCAGACGTGTGCTCTTCCGATCT**CATCTACACTGTTGTTATCAGATCT

The PCR product for each population (expressed and top 15% of binders for each of the four libraries) was cleaned using Agencourt AMPure XP (Beckman Coulter, Inc.) and 1 µl from a 1:10 dilution was taken to the next PCR step for index labeling using KAPA Hifi DNA-polymerase (Kapa Biosystems, London, England):

>forward: AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGC

>reverse: CAAGCAGAAGACGGCATACGAGAT<index>GTGACTGGAGTTCAGACGTGTGC

Top 15% - index: CAATAGTC

Expressed - index: TTGAGCCT

All the primers were ordered as PAGE-purified oligos. The concentration of the PCR product was measured using Qu-bit assay (Life Technologies, Grand Island, New York).

Deep-sequencing runs

DNA samples were run on an Illumina MiSeq using 150-bp paired-end kits. The FASTQ sequence files were obtained for each run, and customized scripts were used to generate the selection heat maps from the data as previously described[22]. Briefly, the script starts by translating the DNA sequence to amino acid sequence; eliminates sequences that harbor more than one amino acid mutation relative to wild type and also sequences that failed the QC test; counts each variant in each population; and eliminates variants with fewer than 100 counts in the reference population (to reduce statistical uncertainty).

Sequencing analysis

To derive the mutational landscapes, we computed the frequency $p^{i,j}$ of each mutant relative to wild-type in the selected and reference pools, where i is the position and j is the substitution, relative to wild-type:

$$p^{i,j} = \frac{count^{i,j}}{count_{wild-type}} \quad \text{(Eq. 1)}$$

where *count* is the number of reads for each mutant. The selection coefficients were then computed as the ratio:

$$S^{i,j} = \frac{(P^{i,j})_{selected}}{(P^{i,j})_{reference}} \quad \text{(Eq. 2)}$$

where *selected* refers to the top 15% binding population and reference refers to the reference population (Expression). The resulting $S^{i,j}$ values were then transformed to fold-enrichment values by:

$$Enrichment\ fold = -ln(S^{i,j}) \quad \text{(Eq. 3)}$$

Computational methods

All Rosetta design simulations used git version fb77c732b4f08b6c30572a2ef7760ad3bb4535ca of the Rosetta biomolecular modeling software, which is freely available to academics at http://www.rosettacommons.org. RosettaScripts[26] and command lines are available in Supplemental Data.

We refined the RH5-bound basigin structure (PDB entry: 4U0Q) by four iterations of side-chain packing and side-chain and backbone minimization[27].

The experimental mutation tolerance map was used to define a sequence space with all beneficial mutations, followed by a combinatorial sequence optimization. Starting from the refined structure, we imposed coordinate restraints relative to the coordinates in the PDB entry and implemented four iterations of sequence design, sidechain, and backbone minimization, while alternating soft and hard repulsive potentials[27]. The energy difference between the refined structure and the optimized high-affinity mutant was calculated using the talaris2014 energy function[28].

BSG and BSG^{des7} production and purification

BSG and BSG^{des7} were transformed into pet21 H-Tev plasmid by RF cloning[29]. *E. coli* strain Origami B(DE3) cells were induced with 1 mM (0.2mM for BSG^{des7}) IPTG at OD600 = 0.6, transferred to 20 ℃ (16 ℃ for BSG^{des7}) and harvested after 20 h. The cells were sonicated at 20 mM HEPES pH 7.5 and 150 mM NaCl, following centrifugation at 20,000g for 1 hour. Next, the supernatant was loaded on nickel gravity column, washed with 20 mM HEPES pH 7.5, 150 mM NaCl and 20mM imidazole and eluted with 20 mM HEPES pH 7.5, 150mM NaCl and 300mM imidazole. The eluted sample was dialyzed overnight at 4℃ in 20 mM HEPES pH 7.5, 150mM NaCl. To remove the His-tag, Tev-His was added at 4℃ overnight. On the following day, the protein with Tev-His was loaded on nickel gravity column and eluted with 20 mM HEPES pH 7.5, 150 mM NaCl and 20mM imidazole and purified by gel filtration over a HiLoad 26/260 superdex 75 pg column.

Apparent T_m and aggregation onset measurements

All experiments were performed on a ViiA 7 real-time PCR machine (Applied Biosystems). BSG and BSG^{des7} were diluted to 10 µM in a 1:5,000 dilution of SYPRO Orange in HBS and 20 µL of the solution was measured. The temperature was ramped from 25 °C to 100 °C at 0.05 °C/s.

Surface-plasmon resonance

Surface plasmon resonance experiments on BSG and BSG^{des7} were carried out on a Biacore T200 instrument (GE Healthcare) at 25 °C with HBS-N EP+ (10 mM Hepes, 150 mM NaCl, 3 mM EDTA, 0.005% vol/vol surfactant P20 [pH 7.4]). For binding analysis, 1,000 response units (RU) of *Pf*RH5ΔNL[10] were captured on a CM5 sensor chip. Samples of different protein concentrations were injected over the surface at a flow rate of 30 µL/min for 500 s, and the chip was washed with buffer for 2,000 s. If necessary, surface regeneration was performed with 30 s injection of 4M Urea (BSG^{des7}) at a flow rate of 30 µL/min. One flow cell contained no ligand and was used as a reference. The acquired data were

analyzed using the device's software, and due to the very slow dissociation rates of BSG^{des7}, only kinetic analyses were performed.

Parasite culture

Parasite line NF54 was used for the growth assay (generously provided by Malaria Research Reference Reagent Resource Center [MR4]), while parasite line D10-PHG was used for the merozoite invasion assay[30]. Parasites were grown in pooled (A+/A+) donor RBCs provided by the Israeli blood bank (Magen David Adom blood donations in Israel) at 4% hematocrit and incubated at 37 ℃ in a gas mixture of 1% O_2, 5% CO_2 in N_2. Parasites were maintained in RPMI medium pH 7.4, 25 mg/ml HEPES, 50 µg/ml hypoxanthine, 2 mg/ml sodium bicarbonate, 20 µg/ml gentamicin and 0.5% AlbumaxII[31]. *P. falciparum* cultures were tested for mycoplasma once a month using a commercially available kit (Mycoplasma Detection Kit—QuickTest [Tivanbiotech]).

Growth assays

NF54 parasites were plated at 1% trophs, 4% hematocrit in 6-well plates with 2 mL total medium. Following a 3 hr recovery, treatments were introduced and parasitemia was monitored using FACS. Parasites were stained with nuclear dye Hoechst 33342 (Invitrogen) titrated to 2-4 µM and RNA dye Thiazole Orange (Sigma) was diluted 1:100000 from 1 mg/ml stock. Cells were incubated for 30 min at 37 ℃. Parasitemia was monitored over two growth cycles, and qualitative evaluation was done with Giemsa smears.

Merozoite invasion assays

Parasites were grown in O+ donor RBCs. Merozoite assays were performed as previously described[32]. In brief, late-stage parasites were magnetically purified using the Magnetic Cell Separation (MACS) system (Miltenyi Biotec), then incubated in culture medium supplemented with 10 µM E64 for <4 hours. Parasites were centrifuged for 5 minutes at 1,900xg and resuspended in a small volume of culture medium, before rupturing using a 1.2-µm syringe filter (Sartorius Stedim Biotech). Merozoites were rapidly added to erythrocytes that were pre-incubated with additives: buffer (20 mM HEPES pH 7.5 and 150 mM NaCl); heparin (a known invasion inhibitor); or BSG or BSG^{des7} protein at two concentrations (189 µg/ml and 38 µg/ml). After 12 hours, erythrocytes were washed in PBS, stained for 20 minutes in the dark with SYBRGreen (1:2,000 in PBS; Sigma-Aldrich), then washed again three times before analysis on a BD Fortessa flow cytometer. 100,000 cells per sample were acquired using the FITC channel.

Growth-inhibitory assays

GIAs were performed using 3D7 clone *P. falciparum* parasites, including monoclonal antibody controls[12]), as previously described[15].

Results

Mutational tolerance mapping of basigin

Following visual inspection of the RH5-basigin molecular structure (Protein Data Bank [PDB] entry: 4u0q), we selected 64 positions on the two N-terminal immunoglobulin domains of basigin for deep mutational scanning. The positions comprised the entire RH5-binding surface as well as second-shell positions (Fig. 1a).

The basigin fragment and the single-point mutants were cloned into a yeast-display vector[25], expressed on the yeast cell-surface, and incubated with biotinylated RH5 (at 725 nM concentration, approximately ½ of the wild-type dissociation constant)[20]. We then used fluorescence-activated cell sorting (FACS) to select the top 15% of the binding population[25] (Fig. 1b). To measure the baseline propensities of each mutant in the starting library, we also screened the expressed population without selecting for binding[20]. Plasmids from both populations were purified, PCR-amplified, and subjected to single-read deep sequencing, resulting in 8.4 million high-quality reads. We then determined the enrichment ratio for each mutant relative to human basigin in both the affinity-selected and baseline populations (Fig. 1c). Four mutations that resulted in substantial improvements in affinity occurred at the core of the RH5-binding surface (Val26Ile, Gln100Ile, His102Tyr, and Ser190Gln), whereas three others were in the periphery (Thr29Ser, Val30Ala, and Gln164Phe). Two of these positions (100 and 102) form close atomic contacts with one another, suggesting that multipoint mutations at least in these positions may exhibit complex epistatic relationships[33].

Past applications of deep mutational scanning to improve binding affinity used an iterative experimental selection of high-affinity variants from libraries comprising combinations of affinity-enhancing mutations[19,21,34]. Here, instead, we tested whether Rosetta design could circumvent such iterative experimental screening by finding a single optimal combination of affinity-enhancing mutations. We started by identifying nine positions on the RH5-binding surface in which mutations exhibited at least

fourfold enrichment relative to the wild-type identity according to deep mutational scanning. We then used Rosetta to find an optimal combination of mutations within this set. The theoretical sequence space of all combinations of mutations at nine positions is very large (10^{12} different sequences), whereas the space of affinity-enhancing variants as inferred from deep mutational scanning is much smaller (14,400 sequences; Table 1), thus focusing design calculations on a small set of sequences that are likely to be enriched for improved binding affinity. In preliminary calculations, we noticed that the design procedure converged on variants that were at most three mutations relative to human basigin, reflecting the tendency of design algorithms to favor the crystallographically determined structure[33,35,36]. To force additional mutations, we visually inspected the basigin-RH5 structure and eliminated the wild-type identities from the design choices at six of the nine positions that appeared permissive to mutation and used Rosetta to find an optimal combination of the remaining mutations. This resulted in a single design, comprising seven mutations that we called BSG^{des7} (Table 1).

We noted that most of the mutations in BSG^{des7} relative to wild type (4 out of 7) were not ranked as the most enriched in the deep mutational scanning experiment (Table 1); indeed, at position 98, the least favored identity (out of five) was selected by Rosetta. We used Rosetta design calculations to force the design of the most-favored mutation according to the experiment at each of the seven positions, finding that this design exhibited energy that was higher (worse) by a substantial 7 Rosetta energy units compared to BSG^{des7}. Thus, although the deep mutational scanning data are essential for focusing design calculations on a productive subset of the theoretical sequence space, the ranking of mutants according to the experiment may not be accurate enough for predicting favorable multipoint mutant combinations.

BSG^{des7} exhibits almost 1,900-fold higher affinity for RH5

We solubly expressed BSG^{des7} and measured its affinity for RH5 using surface-plasmon resonance (SPR). Strikingly, BSG^{des7} exhibited 1,900-fold higher affinity for RH5 compared to human basigin (K_D=1.15 and 2,200 nM, respectively). The kinetic analysis suggested that the main difference in affinity came from a vast decrease in binding off-rate (0.23 h^{-1} compared to 0.13 s^{-1}) (Fig. 2a). SPR single-cycle kinetic analysis confirmed the very large affinity improvement, though with a twofold higher K_D = 3.3 nM (Fig. S1). A comparison of the structure of BSG and the model of BSG^{des7} revealed that the design increased the hydrophobicity at the core of the interface through the introduction of two radical polar-to-hydrophobic mutations (Gln100Ile and His102Tyr) providing a likely explanation for the vast improvement in affinity (Fig 2b). Examination of the experimental mutation-tolerance map suggested that the human identities at

these positions, particularly position 100, are likely to be suboptimal for binding RH5[37]; for instance, nearly all aliphatic point mutations at position 100 relative to human BSG exhibited large gains in affinity. Other mutations in BSG^{des7} relative to BSG, for instance, Ser190Gln, improved polar interactions across the binding interface. Thus, we concluded that owing to the suboptimal molecular contacts across the BSG-RH5 interface, large gains in affinity could be readily designed.

The introduction of seven simultaneous mutations could severely destabilize the protein[38]. We found, however, that BSG^{des7} exhibited a high apparent melting temperature (50 °C compared to 55 °C for BSG), demonstrating that the design procedure maintained protein stability (Fig. 2c). Indeed, the small reduction in thermal stability is to be expected given the radical nature of some of the mutations (Table 1).

BSG^{des7} improves growth-inhibitory concentrations

Encouraged by the large improvement in affinity, we next tested the ability of BSG^{des7} to inhibit the invasion of *P. falciparum* parasites (NF54 strain) to human red blood cells. We started by measuring parasitemia levels for parasites incubated with human BSG or BSG^{des7} at three concentrations using fluorescence-activated cell sorting (FACS) following two invasion cycles. A clear decrease in parasitemia was evident for cultures treated with 1 μM concentration of the design, as was measured by both FACS analysis[39] and Giemsa smears (Fig 3a). Importantly, five-day treatment with BSG^{des7} eliminated live invading parasites (no evidence for ring-stage parasites).

To confirm that invasion is the step where the parasite life cycle is inhibited by BSG and BSG^{des7}, we tested the direct effect on invasion by adding merozoites to red blood cells that were pretreated with BSG^{des7} or BSG. Both proteins inhibited invasion, with BSG^{des7} observed to be slightly more potent (Fig. 3b).

We next used assays of growth-inhibition activity (GIA) to assess the inhibitory concentration of BSG^{des7}, finding that it exhibited an inhibitory concentration (IC_{50}) of 5 μM, compared to only 50 μM for human basigin (Fig. 3c). Although the tenfold improvement in inhibitory concentration was encouraging, it was substantially less than expected based solely on the three orders of magnitude improvement in affinity. This result is consistent with suggestions made previously that growth inhibition depends crucially on binding on-rate, and indeed, antibody 2AC7, which exhibits 24-fold higher binding on-rate[12] compared to BSG^{des7} also inhibits parasitemia with nearly 200-fold improved inhibitory constant. Therefore, we

concluded that substantial improvement in growth-inhibitory concentrations would demand improvements in binding on rate or potentially the fusion of multiple BSG^{des7} monomers to generate an avid inhibitor[6].

Discussion

Many host-pathogen interactions are transient and therefore exhibit only low affinity. Furthermore, host proteins are highly unlikely to have been evolutionarily selected for high-affinity interactions with the pathogen; it is far more likely that selection pressures operated if at all, to weaken interactions with pathogens. Host target proteins, therefore, provide potential starting points for engineering and design of potent inhibitors of cell invasion. In this study, we demonstrated a straightforward path for substantial improvement of affinities through systematic mutation analysis followed by one-shot design calculations. Our structural analysis demonstrated that indeed much of the binding surface on basigin, including at the very core of its interaction with RH5, is suboptimal, and that mutations can readily improve binding affinity. In an *in vivo* setting, it would be essential to verify that such design does not lead to undesired side effects due to interactions with other host proteins or to an immune reaction.

The large increase in binding affinity translated, however, to only a tenfold improvement in inhibitory concentration. Recent analyses of anti-RH5 antibody binding kinetics and inhibitory concentrations suggested that binding on-rate makes an essential contribution to inhibition[12]. Red-blood-cell invasion by parasites is a very fast process (10-30 seconds)[40], meaning that the exposure of RH5 to serum is short and that inhibitors must engage RH5 quickly in order to block invasion. Our results also do not exclude the possibility that the design binds off-target proteins in serum or on the parasite, leading to a lower effective concentration. BSG^{des7} may be used as a starting point to increase RH5-binding on-rate in order to produce an effective parasitemia blocker.

Acknowledgment

SJF and NRR are supported by a research grant from the Benoziyo Endowment Fund for the Advancement of Science. SJF is additionally supported by a European Research Council Consolidator Grant (713255) and by a gift from Sam Switzer and family. O.L. is supported by a Ph.D. scholarship from the UK Medical Research Council (MR/K501281/1). K.E.W. is supported through a Sir Henry Wellcome Postdoctoral Fellowship (107366/Z/15/Z). Research in the Baum lab is supported by an Investigator

Award to JB from Wellcome (100993/Z/13/Z). M.K.H. is a Wellcome Investigator. S.J.D. is a Jenner Investigator; a Lister Institute Research Prize Fellow and a Wellcome Trust Senior Fellow (106917/Z/15/Z).

References

1. Organization WH. *World Malaria Report 2015*. World Health Organization; 2016.

2. Goldberg DE, Siliciano RF, Jacobs WR Jr. Outwitting evolution: fighting drug-resistant TB, malaria, and HIV. *Cell*. 2012;148(6):1271-1283.

3. Draper SJ, Angov E, Horii T, et al. Recent advances in recombinant protein-based malaria vaccines. *Vaccine*. 2015;33(52):7433-7443.

4. Douglas AD, Williams AR, Illingworth JJ, et al. The blood-stage malaria antigen PfRH5 is susceptible to vaccine-inducible cross-strain neutralizing antibody. *Nat Commun*. 2011;2:601.

5. Baum J, Chen L, Healer J, et al. Reticulocyte-binding protein homologue 5--an essential adhesin involved in invasion of human erythrocytes by Plasmodium falciparum. *Int J Parasitol*. 2009;39(3):371-380.

6. Crosnier C, Bustamante LY, Bartholdson SJ, et al. Basigin is a receptor essential for erythrocyte invasion by Plasmodium falciparum. *Nature*. 2011;480(7378):534-537.

7. Bustamante LY, Bartholdson SJ, Crosnier C, et al. A full-length recombinant Plasmodium falciparum PfRH5 protein induces inhibitory antibodies that are effective across common PfRH5 genetic variants. *Vaccine*. 2013;31(2):373-379.

8. Williams AR, Douglas AD, Miura K, et al. Enhancing blockade of Plasmodium falciparum erythrocyte invasion: assessing combinations of antibodies against PfRH5 and other merozoite antigens. *PLoS Pathog*. 2012;8(11):e1002991.

9. Manske M, Miotto O, Campino S, et al. Analysis of Plasmodium falciparum diversity in natural infections by deep sequencing. *Nature*. 2012;487(7407):375-379.

10. Wright KE, Hjerrild KA, Bartlett J, et al. Structure of malaria invasion protein RH5 with erythrocyte basigin and blocking antibodies. *Nature*. 2014;515(7527):427-430.

11. Douglas AD, Baldeviano GC, Lucas CM, et al. A PfRH5-based vaccine is efficacious against heterologous strain blood-stage Plasmodium falciparum infection in aotus monkeys. *Cell Host Microbe*. 2015;17(1):130-139.

12. Douglas AD, Williams AR, Knuepfer E, et al. Neutralization of Plasmodium falciparum merozoites by antibodies against PfRH5. *J Immunol*. 2014;192(1):245-258.

13. Payne RO, Silk SE, Elias SC, et al. Human vaccination against RH5 induces neutralizing antimalarial antibodies that inhibit RH5 invasion complex interactions. *JCI Insight*. 2017;2(21). doi:10.1172/jci.insight.96381

14. Chen L, Xu Y, Healer J, et al. Crystal structure of PfRh5, an essential P. falciparum ligand for invasion of human erythrocytes. *Elife*. 2014;3. doi:10.7554/eLife.04187

15. Campeotto I, Goldenzweig A, Davey J, et al. One-step design of a stable variant of the malaria invasion protein RH5 for use as a vaccine immunogen. *Proc Natl Acad Sci U S A*. 2017;114(5):998-1002.

16. Conte LL, Chothia C, Janin J, Lo Conte L, Chothia C, Janin J. The atomic structure of protein-protein recognition sites. *J Mol Biol*. 1999;285(5):2177-2198.

17. Wong W, Huang R, Menant S, et al. Structure of Plasmodium falciparum Rh5-CyRPA-Ripr invasion complex. *Nature*. 2019;565(7737):118-121.

18. Fowler DM, Araya CL, Fleishman SJ, et al. High-resolution mapping of protein sequence-function relationships. *Nat Methods*. 2010;7(9):741-746.

19. Whitehead TA, Chevalier A, Song Y, et al. Optimization of affinity, specificity and function of designed influenza inhibitors using deep sequencing. *Nat Biotechnol*. 2012;30(6):543-548.

20. Whitehead TA, Baker D, Fleishman SJ. Computational design of novel protein binders and experimental affinity maturation. *Methods Enzymol*. 2013;523:1-19.

21. Schreiber G, Fleishman SJ. Computational design of protein-protein interactions. *Curr Opin Struct Biol*. 2013;23(6):903-910.

22. Elazar A, Weinstein J, Biran I, Fridman Y, Bibi E, Fleishman SJ. Mutational scanning reveals the determinants of protein insertion and association energetics in the plasma membrane. *Elife*. 2016;5. doi:10.7554/eLife.12125

23. Liu H, Naismith JH. An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol*. 2008;8:91.

24. Gietz RD, Schiestl RH. Frozen competent yeast cells that can be transformed with high efficiency using the LiAc/SS carrier DNA/PEG method. *Nat Protoc*. 2007;2(1):1-4.

25. Chao G, Lau WL, Hackel BJ, Sazinsky SL, Lippow SM, Wittrup KD. Isolating and engineering human antibodies using yeast surface display. *Nat Protoc*. 2006;1(2):755-768.

26. Fleishman SJ, Leaver-Fay A, Corn JE, et al. RosettaScripts: a scripting language interface to the Rosetta macromolecular modeling suite. *PLoS One*. 2011;6(6):e20161.

27. Goldenzweig A, Goldsmith M, Hill SE, et al. Automated Structure- and Sequence-Based Design of Proteins for High Bacterial Expression and Stability. *Mol Cell*. 2018;70(2):380.

28. O'Meara MJ, Leaver-Fay A, Tyka MD, et al. Combined covalent-electrostatic model of hydrogen bonding improves structure prediction with Rosetta. *J Chem Theory Comput*. 2015;11(2):609-622.

29. Unger T, Jacobovitch Y, Dantes A, Bernheim R, Peleg Y. Applications of the Restriction Free (RF) cloning procedure for molecular manipulations and protein expression. *J Struct Biol*. 2010;172(1):34-44.

30. Wilson DW, Crabb BS, Beeson JG. Development of fluorescent Plasmodium falciparum for in vitro growth inhibition assays. *Malar J*. 2010;9:152.

31. Trager W, Jensen JB. Human malaria parasites in continuous culture. *Science*. 1976;193(4254):673-675.

32. Boyle MJ, Wilson DW, Richards JS, et al. Isolation of viable Plasmodium falciparum merozoites to define erythrocyte invasion events and advance vaccine and drug development. *Proc Natl Acad Sci U S A*. 2010;107(32):14378-14383.

33. Khersonsky O, Lipsh R, Avizemer Z, et al. Automated Design of Efficient and Functionally Diverse Enzyme Repertoires. *Mol Cell*. 2018;72(1):178-186.e5.

34. Leung I, Dekel A, Shifman JM, Sidhu SS. Saturation scanning of ubiquitin variants reveals a common hot spot for binding to USP2 and USP21. *Proc Natl Acad Sci U S A*. 2016;113(31):8705-8710.

35. Netzer R, Listov D, Lipsh R, et al. Ultrahigh specificity in a network of computationally designed protein-interaction pairs. *Nat Commun*. 2018;9(1):5286.

36. Fleishman SJ, Whitehead TA, Ekiert DC, et al. Computational design of proteins targeting the conserved stem region of influenza hemagglutinin. *Science*. 2011;332(6031):816-821.

37. Ferreiro DU, Komives EA, Wolynes PG. Frustration in Biomolecules. *Q Rev Biophys*. 2013;47(4):1-97.

38. Goldenzweig A, Fleishman SJ. Principles of Protein Stability and Their Application in Computational Design. *Annu Rev Biochem*. 2018;87:105-129.

39. Dekel E, Rivkin A, Heidenreich M, et al. Identification and classification of the malaria parasite blood developmental stages, using imaging flow cytometry. *Methods*. 2017;112:157-166.

40. Weiss GE, Crabb BS, Gilson PR. Overlaying Molecular and Temporal Aspects of Malaria Parasite Invasion. *Trends Parasitol*. 2016;32(4):284-295.

Figure legends

Figure 1. **Mutational scanning of the soluble fragment of human basigin.** (**a**) Basigin (cartoon) and RH5 (surface) structure (PDB entry: 4u0q). Positions that were subjected to deep mutational scanning are colored in orange. (**b**) Schematic representation of the experimental method. 64 positions were selected for single point saturation mutagenesis. The PCR products were transformed into yeast and the top 15% and expressed population was selected for deep mutational scanning. The fold enrichment values were calculated for each point mutation. (**c**) The deep mutational scanning map shows that most positions do

not increase binding affinity and mutations in only 9 of 64 positions improved affinity. Mutations in blue improved binding affinity; mutations in red were deleterious and ones in gray contained insufficient data (NA) for statistical analysis. One-letter codes indicate wild type amino acid identities at each position. Positions that increased the binding affinity are marked by arrows.

Figure 2. **Molecular modeling and functional analysis of BSG^{des7}.** (**a**) SPR kinetic analysis of RH5 binding with twofold dilutions of RH5 from a maximal concentration of 8 μM and threefold dilutions of RH5 from a maximal concentration of 333 nM for BSG^{des7} (kinetic fits shown in gray; k_a=6.3*10^4 M^{-1}s^{-1}, k_d = 0.14 s^{-1}, K_D=2.2 μM, in agreement with ref. [10,14]) and k_a=5.6*10^4 M^{-1}s^{-1}, k_d = 6.4*10^{-5} s^{-1} K_D=1.15 nM for BSG and BSG^{des7}, respectively. (**b**) BSG^{des7} is predicted to exhibit improved core packing and reduced electrostatic strain at the core of the interface. (**c**) Thermal denaturation of BSG and BSG^{des7}.

Table 1: Sequence space of mutations that improve binding relative to BSG and the BSG^{des7} sequence, rank-ordered from left-to-right according to the enrichment values observed in deep mutational scanning.

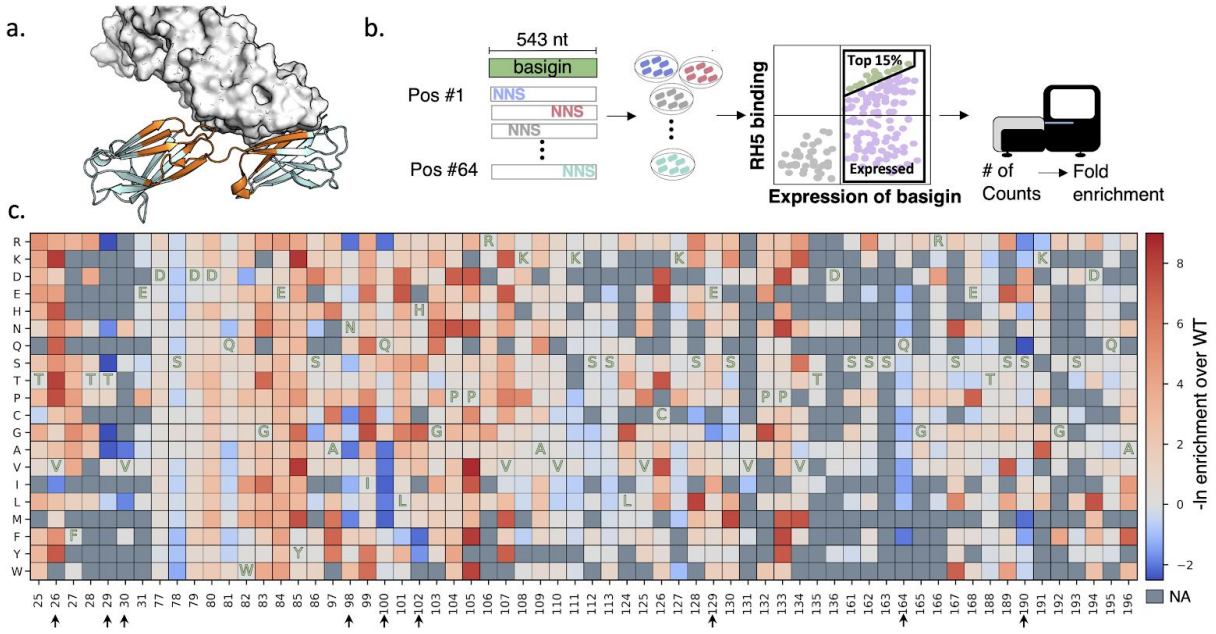
Figure 3. *P. falciparum* growth inhibition assays with BSG and BSG^{des7}. (**a**) *P. falciparum*-infected cells (RBC) were monitored for parasitemia levels during two invasion cycles. Incubation with 5 μM BSG^{des7} reduces parasitemia levels drastically relative to BSG. Blood smears show that in the non-treated cells, there are live parasites, while there are no ring-stage parasites in 5 μM BSG^{des7}. (**b**) A direct merozoite invasion assay shows that BSG^{des7} inhibits invasion compared to the positive control treated cells. *P. falciparum* merozoites were isolated and added to RBC in the presence of normal basigin or BSG^{des7}. The experiment was performed in triplicate (mean and standard deviation are shown). Both inhibit invasion, with BSG^{des7} showing slightly higher potency. (**c**) GIA assays against 3D7 clone *P. falciparum* parasites show an improvement of an order of magnitude in the growth-inhibitory concentration of BSG^{des7} relative to BSG. Mouse antibodies QA1, 9AD4, and 2AC7 exhibit approximately 200-fold lower growth-inhibitory concentrations than BSG^{des7}. The experiment was performed in triplicate (mean and standard deviation are shown) and nonlinear regression was used to determine IC50.

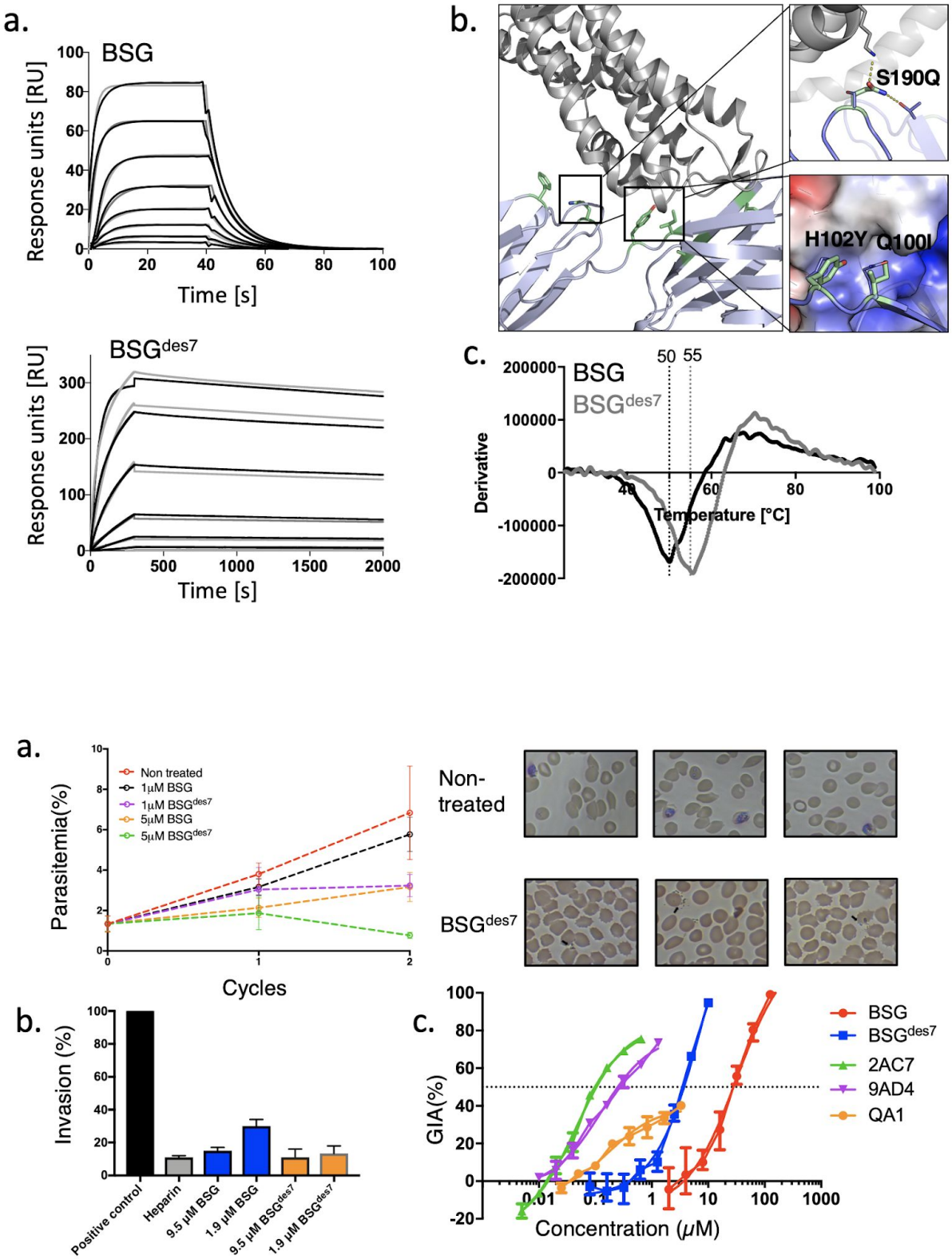
Figure S1. Binding data from single-cycle kinetic measurements of the interaction between BSG^{des7} and RH5. SPR single-cycle kinetic analysis with fivefold dilutions of RH5 from a maximal concentration of 625 nM. Kinetic fits shown in black. k_a=4.14*10^4 M^{-1}s^{-1}, k_d = 1.25*10^{-4} s^{-1} and K_D=3.3 nM.

Tables

Position	**26**	**29**	**30**	**98**	**100**	**102**	**129**	**164**	**190**
BSG	V	T	V	N	Q	H	E	Q	S
Sequence space	I	GSRAN	AL V	RAMCN	IMRAVL	FY	GE	FI	QMRF
BSG^{des7}	I	S	A	N	I	Y	E	F	Q

Figures


a.
b.
543 nt
basigin
Pos #1
NNS
NNS
NNS
Pos #64
NNS
RH5 binding
Top 15%
Expressed
Expression of basigin
# of Counts
Fold enrichment
c.
-ln enrichment over WT
NA


a.
BSG
Response units [RU]
Time [s]
BSGdes7
b.
S190Q
H102Y
Q100I
c.
BSG
BSGdes7
Derivative
Temperature [°C]
50
55
a.
Non treated
1μM BSG
1μM BSGdes7
5μM BSG
5μM BSGdes7
Parasitemia(%)
Cycles
Non-treated
BSGdes7
b.
Invasion (%)
Positive control
Heparin
9.5 μM BSG
1.9 μM BSG
9.5 μM BSGdes7
1.9 μM BSGdes7
c.
GIA(%)
Concentration (μM)
BSG
BSGdes7
2AC7
9AD4
QA1

Supplementary Figure 1

