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

Shira Warszawski1, Elya Dekel1, Ivan Campeotto2, Jennifer M. Marshall1, Katherine E. Wright4, Oliver Lyth4, Orli Knop1, Neta Regev-Rudzki1, Matthew K Higgins2, Simon J Draper3, Jake Baum4 and Sarel J Fleishman1*

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

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 ($K_D \geq 1 \mu M$) 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 2,500-fold higher affinity ($K_D < 1 \text{ nM}$) for RH5 with a very slow binding off rate ($0.23 \text{ h}^{-1}$) 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. Designed mimics may provide therapeutic advantages over antibodies, since the mimics bind to essential surfaces on the target pathogen proteins, reducing the likelihood for the emergence of escape mutants.
Introduction

Malaria is the deadliest parasitic disease with over 200 million clinical cases annually\(^{(1)}\)\((\text{Organization, 2016})\). 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\(^{(2)}\)\((\text{Goldberg, Siliciano, & Jacobs, 2012})\). New modes of blocking the parasite are therefore urgently needed. Invasion of host erythrocytes is essential to the life cycle of *Plasmodium* parasites and therefore provides 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\(^{(3)}\)\((\text{Baum et al., 2009; Crosnier et al., 2011; Douglas et al., 2011; Draper et al., 2015})\). Indeed, RH5 is highly conserved among all *P. falciparum* field isolates\(^{(4)}\)\((\text{Bustamante et al., 2013; Crosnier et al., 2011; Manske et al., 2012; Williams et al., 2012; Wright et al., 2014})\); antibodies that target RH5 inhibit pathogenesis; inoculation with RH5 in model animals provides protective immunity\(^{(5)}\)\((\text{Bustamante et al., 2013; Crosnier et al., 2011; Douglas et al., 2015, 2011, 2014; Williams et al., 2012; Wright et al., 2014})\); and clinical trials are underway to test the efficacy of RH5 subunit vaccines in humans\(^{(6)}\)\((\text{Payne et al., 2017})\).

The molecular structures of RH5 in complex with the extracellular fragment of basigin and with neutralizing antibodies are known\(^{(7)}\)\((\text{Chen et al., 2014; Wright et al., 2014})\), enabling computational design of improved vaccine immunogens based on RH5\(^{(8)}\)\((\text{Campeotto et al., 2017})\). 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\(^{(9)}\)\((\text{Wright et al., 2014})\). Although the interaction surface area between RH5 and basigin (~1,900 Å\(^2\)) is above average for protein-protein interactions\(^{(10)}\)\((\text{Conte et al., 1999})\), 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 loss in conformational
freedom upon binding. Accordingly, the interaction affinity between RH5 and basigin is low ($K_D \geq 1 \mu M$)(Wong et al., 2019; Wright et al., 2014), and the application of the extracellular domain blocks invasion in growth assays only at high and therapeutically unrealistic concentrations ($IC_{50} > 10\mu M$). Animal immunizations using RH5 led to isolation of several high-affinity anti-RH5 antibodies that exhibited much improved growth-inhibitory concentration ($IC_{50}=0.1 \mu M$) compared to basigin((Douglas et al., 2014), 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 partly overlapped with the basigin-binding surface((Wright et al., 2014), making them susceptible to the emergence of RH5 escape mutants.

The current study therefore 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.

In recent years, deep mutational scanning has emerged as a powerful strategy to identify affinity-enhancing mutations(Fowler et al., 2010; Whitehead et al., 2012). 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 improved affinity(Schreiber & Fleishman, 2013; Whitehead, Baker, & Fleishman, 2013; Whitehead et al., 2012). 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.
Surprisingly, a design comprising seven affinity-enhancing mutations resulted in a remarkable 2,500-fold higher affinity for RH5. This design provides a proof-of-principle for 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 (Elazar et al., 2016). Primers were ordered from Sigma (Sigma-Aldrich, Rehovot, Israel) and used to introduce all possible amino acids per position by QuickChange mutagenesis (Liu & Naismith, 2008). Next, the PCR product of each position was transformed into yeast (EBY100 cells) and plated on SD-Trp as described (Gietz & Schiestl, 2007). As described (Chao et al., 2006), plates with more than 400 colonies were scraped with 1 ml SDCAA, 50 μl were added to 5 ml SDCAA tube and cells were then grown at 30°C 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 (Chao et al., 2006). Briefly, yeast cells were grown in selective medium SDCAA overnight at 30°C. The cells were then resuspended in 10 ml induction medium and incubated at 20°C for 20 h. 10⁷ cells were then used for yeast-cell surface display experiments: cells were subjected to primary antibody (mouse monoclonal IgG1 anti-e-Myc (9E10) sc-40, Santa Cruz Biotechnology) for expression monitoring and biotinylated ligand at 725 nM PfRH5ΔNL (Wright et al., 2014) 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°C. 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

GCGGTACCGTTTTTCACCCAGTAGACCTGGTTCTAATAATCTCTGCTGACCTGCTCTCTGAACGACTCTGCGACCG
AAGTTACCGGTCACCGTTGGCTGAAAGGTGGTGTTGTTCTGAAAGAAGACGCGCTGCCGGGTCAGAAAACCGAATTCA
AGTTGATTCTGATAGACCAGTGGGGTGAATACTCTTGCGTTTTCCTGCCAGAACCGATGGGCACCGCGAACATCCAACTG
CGAACTCTGGTTCTTCTCTTCTTCGAGGCTGTTCTGCTGACTCACTACACATGAAACGTGAACCTGGCTTCTCTGCAGCT
GGAGATCCGAACAAAAGCTTATTTCTGAAGAGGACTTGTATAAGTCTCTTCCTACACGACGCTCTTCTCTTCTCTGCG

>BSG<des7

GCGGGCAACATTTCTACCTTCGCAGGAAAGCTCTGGTTCTAATAATCTCTGCTGACCTGCTCTCTGAACGACTCTCGAGCT
AAGTTACCGGTCATCGTTGGCTCAAGGGTGGTGTTGTTCTGAAAGAAGACGCGCTGCCAGGTCAGAAAACCGAATTCA
AGTTGATTCTGATAGACCAGTGGGGTGAATACTCTTGCGTTTTCCTGCCAGAACCGATGGGCACCGCGAACATCCAACTG
CGAACTCTGGTTCTTCTCTTCTTCGAGGCTGTTCTGCTGACTCACTACACATGAAACGTGAACCTGGCTTCTCTGCAGCT
GGAGATCCGAACAAAAGCTTATTTCTGAAGAGGACTTGTATAAGTCTCTTCCTACACGACGCTCTTCTCTTCTCTGCG

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>GCCGGGTCAAGAAAACCGA
>forward (seg3):<barcode>ATCCAACTGACGCTTGCC
>forward (seg4):<barcode>GTTTCGAATCTCGTTTCTTG
>reverse: CTGGAGTTCAGACGTGTCTCTTCCGATCTCATCTACACTGTGTTATCAGATCT

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: AATGATACGGCGACCCGTATACACTGT
>reverse: CAAGCAGAAGACGGCATACGT
}
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 (Elazar et al., 2016). 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:

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

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

$$S_{i,j} = \frac{(p_{i,j})_{\text{selected}}}{(p_{i,j})_{\text{reference}}} \quad \text{(Eq. 2)}$$

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:

$$\text{Enrichment fold} = \log(S_{i,j}) \quad \text{(Eq. 3)}$$

**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 (Fleishman, Leaver-Fay, et al., 2011) 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 (Goldenzweig et al., 2018).
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 (Goldenzweig et al., 2018). The energy difference between the refined structure and the optimized high-affinity mutant was calculated using the talaris2014 energy function (O’Meara et al., 2015).

**BSG and BSG\textsuperscript{des7} production and purification**

BSG and BSG\textsuperscript{des7} were transformed into pet21 H-Tev plasmid by RF cloning (Unger, Jacobovitch, Dantes, Bernheim, & Peleg, 2010). *E. coli* strain Origami B(DE3) cells were induced with 1 mM (0.2 mM for BSG\textsuperscript{des7}) IPTG at OD\textsubscript{600} = 0.6, transferred to 20 °C (16 °C for BSG\textsuperscript{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°C in 20 mM HEPES pH 7.5, 150mM NaCl. To remove the His-tag, Tev-His was added at 4°C 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\textsuperscript{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\textsuperscript{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 \( P/RH5\Delta NL \) (Wright et al., 2014) 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

---

John Wiley & Sons, Inc.
4M Urea (BSG\textsuperscript{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 kinetic fits 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 (Wilson, Crabb, & Beeson, 2010). 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 °C in gas mixture of 1% O\textsubscript{2}, 5% CO\textsubscript{2} in N\textsubscript{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 gentamycin and 0.5% AlbumaxII (Trager & Jensen, 1976). \textit{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 °C. 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 (Boyle et al., 2010). 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\textsuperscript{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 (Douglas et al., 2014), as previously described (Campeotto et al., 2017).

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 fragments 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 each of the single-point mutants were cloned into a yeast-display vector (Chao et al., 2006), expressed on the yeast cell-surface, and incubated with biotinylated RH5 (at 725 nM concentration, approximately ½ of the wild-type dissociation constant) (Whitehead et al., 2013). We then used fluorescence-activated cell sorting (FACS) to select the top 15% of the binding population (Chao et al., 2006) (Fig. 1b). To measure the baseline propensities of each mutant in the starting library, we also screened the expressed population without selecting for binding (Whitehead et al., 2013). 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).

Past applications of deep mutational scanning to improve binding affinity used iterative experimental selection of high-affinity variants from libraries comprising combinations of affinity-enhancing mutations (Leung, Dekel, Shifman, & Sidhu, 2016; Schreiber & Fleishman, 2013; Whitehead et al., 2012). 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 ($<10^5$), thus focusing design calculations on a small set of sequences that is 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 (Fleishman, Whitehead, et al., 2011; Khersonsky et al., 2018; Netzer et al., 2018). To force additional mutations, we visually inspected the basign-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).

BSG$^{des7}$ exhibits 2,500-fold higher affinity for RH5

We solubly expressed BSG$^{des7}$ and tested its affinity for RH5 using surface-plasmon resonance (SPR) (Fig. 2a). Strikingly, BSG$^{des7}$ exhibited 2,500-fold higher affinity for RH5 compared to human basigin ($K_D$=0.8 and 2,200 nM, respectively). 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$^{-1}$). A comparison of the structure of BSG and the model of BSG$^{des7}$ revealed that the design substantially increased 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 these positions, particularly position 100, are likely to be highly frustrated (Ferreiro, Komives, & Wolynes, 2013); 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. We therefore 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 (Goldenzweig & Fleishman, 2018). 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).

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\textsuperscript{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 (Dekel et al., 2017) and Giemsa smears (Fig 3a). Importantly, five-day treatment with BSG\textsuperscript{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\textsuperscript{des7}, we tested the direct effect on invasion by adding merozoites to red blood cells that were pretreated with BSG\textsuperscript{des7} or BSG. Both proteins inhibited invasion, with the designed BSG observed to be slightly more potent (Fig. 3b).

We next used assays of growth-inhibition activity (GIA) to assess the inhibitory concentration of BSG\textsuperscript{des7}, finding that it exhibited an inhibitory concentration (IC\textsubscript{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 (Douglas et al., 2014) than BSG\textsuperscript{des7} also inhibits parasitemia with nearly 200-fold improved inhibitory constant. We therefore concluded that substantial improvement in growth-inhibitory concentrations would demand improvements in binding on rate or potentially the fusion of multiple BSG\textsuperscript{des7} monomers to generate an avid inhibitor (Crosnier et al., 2011).

**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 frustrated, 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 (Douglas et al., 2014). Red-blood cell invasion by parasites is a very fast process (10-30 seconds) (Weiss, Crabb, & Gilson, 2016), meaning that the exposure of RH5 to serum is short, and that inhibitors must engage RH5 quickly in order to block invasion. BSG<sup>des7</sup> may be used as a starting point to increase RH5-binding on rate in order to produce an effective parasitemia blocker.

**Acknowledgement**

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). O.L. is supported by a PhD 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**


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 to yeast and the top 15% and expressed population were selected for deep mutational scanning. The fold enrichment values were calculated for each point mutation. (c) Deep mutational scanning map shows that most positions do not increase binding affinity and only 9 of 64 positions improved binding. Mutations in blue improved binding affinity; mutations in red were deleterious and ones in gray contained insufficient data 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 modelling and functional analysis of BSG\textsuperscript{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 111 nM for BSG\textsuperscript{des7} (kinetic fits shown in gray). BSG exhibited $k_a=6.3 \times 10^4$ M$^{-1}$s$^{-1}$, $k_d=0.14$ s$^{-1}$, and $K_D=2.2$ µM. BSG\textsuperscript{des7} exhibited $k_a=7.5 \times 10^4$ M$^{-1}$s$^{-1}$, $k_d=6.5 \times 10^{-5}$ s$^{-1}$, and $K_D=0.86$ nM. (b) BSG\textsuperscript{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\textsuperscript{des7}.

Table 1: Sequence space of mutations that improve binding relative to BSG and the BSG\textsuperscript{des7} sequence.

<table>
<thead>
<tr>
<th>Position</th>
<th>26</th>
<th>29</th>
<th>30</th>
<th>98</th>
<th>100</th>
<th>102</th>
<th>129</th>
<th>164</th>
<th>190</th>
</tr>
</thead>
</table>

Figure 3. P. falciparum growth inhibition assays with BSG and BSG\textsuperscript{des7}. (a) P. falciparum-infected red-blood cells (RBC) were monitored for parasitemia levels during two invasion cycles. Incubation with 5 µM BSG\textsuperscript{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\textsuperscript{des7}. (b) A direct merozoite invasion assay shows that BSG\textsuperscript{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\textsuperscript{des7}. The experiment was performed in triplicate (mean and standard deviation are shown). Both inhibit invasion, with BSG\textsuperscript{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\textsuperscript{des7} relative to BSG. Mouse antibodies QA1, 9AD4 and 2AC7 exhibit approximately 200-fold lower growth-inhibitory concentrations than BSG\textsuperscript{des7}. The experiment was performed in triplicate (mean and standard deviation are shown) and nonlinear regression was used to determine IC50.
<table>
<thead>
<tr>
<th>BSG</th>
<th>V</th>
<th>T</th>
<th>V</th>
<th>N</th>
<th>Q</th>
<th>H</th>
<th>E</th>
<th>Q</th>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequence space</td>
<td>I</td>
<td>RNSGA</td>
<td>LA</td>
<td>V</td>
<td>MACRN</td>
<td>AVILMR</td>
<td>FY</td>
<td>GE</td>
<td>IF</td>
</tr>
<tr>
<td>BSG&lt;sub&gt;Δns7&lt;/sub&gt;</td>
<td>I</td>
<td>S</td>
<td>A</td>
<td>N</td>
<td>I</td>
<td>Y</td>
<td>E</td>
<td>F</td>
<td>Q</td>
</tr>
</tbody>
</table>

**Figures**

a. [Image of a figure]

b. [Image of another figure]

c. [Image of a third figure]