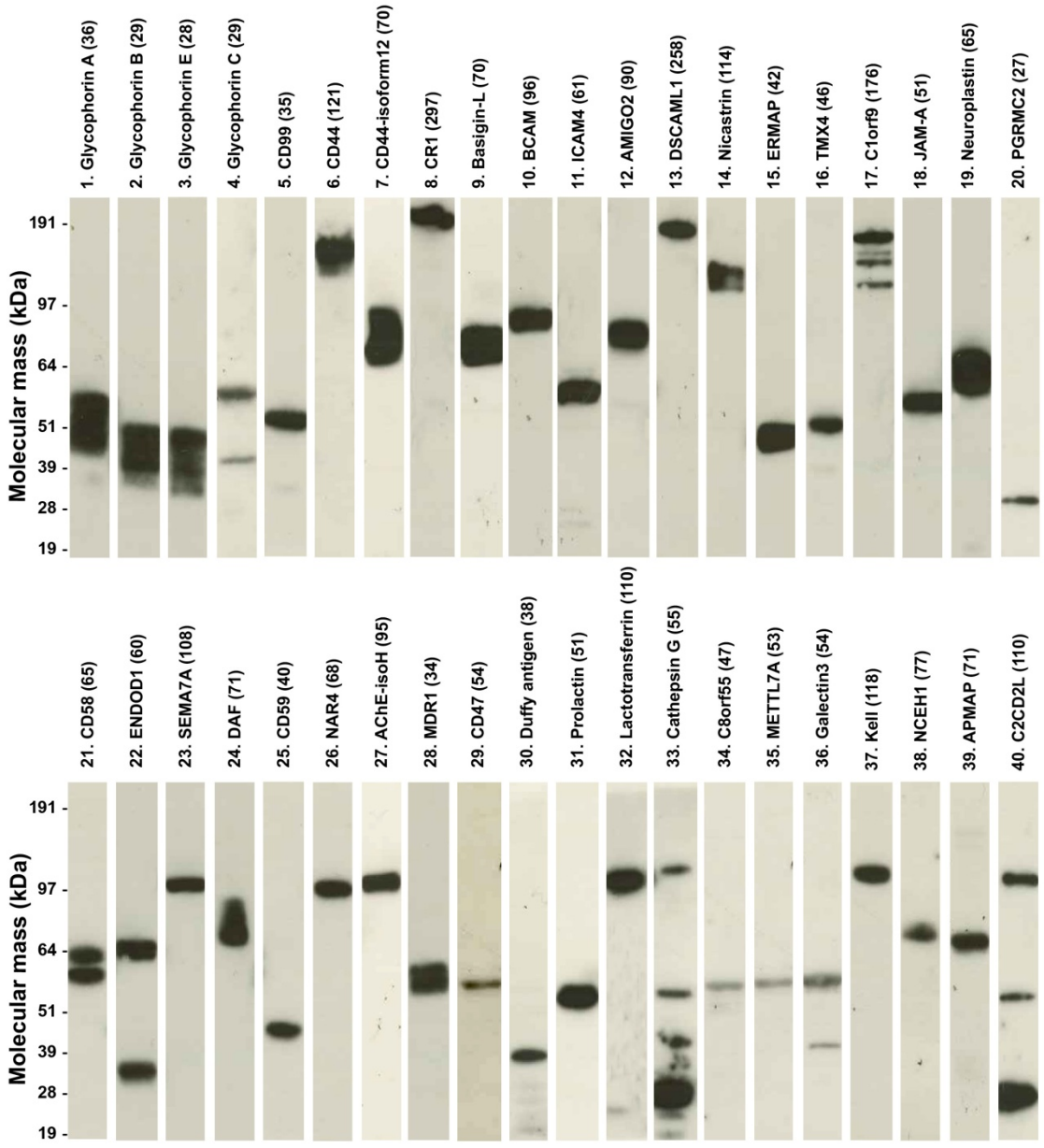
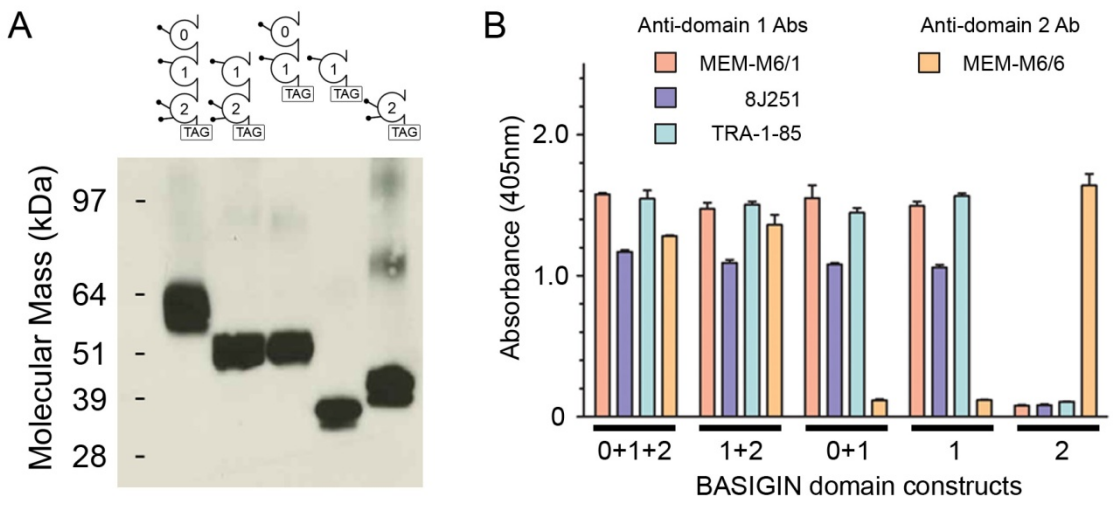
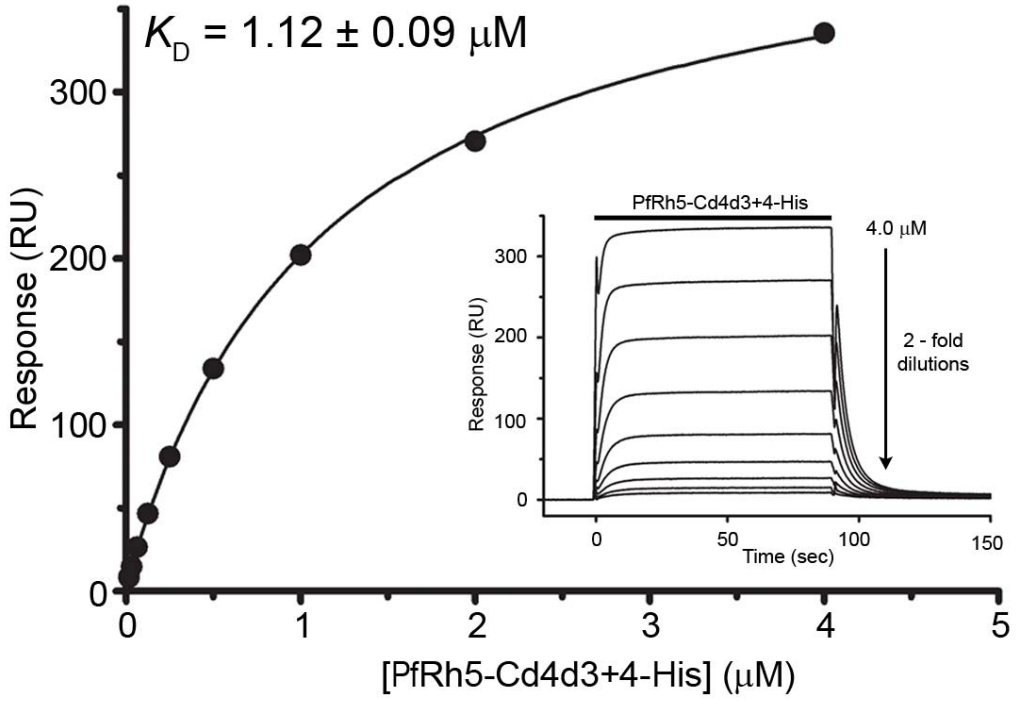
doi:10.1038/nature10606



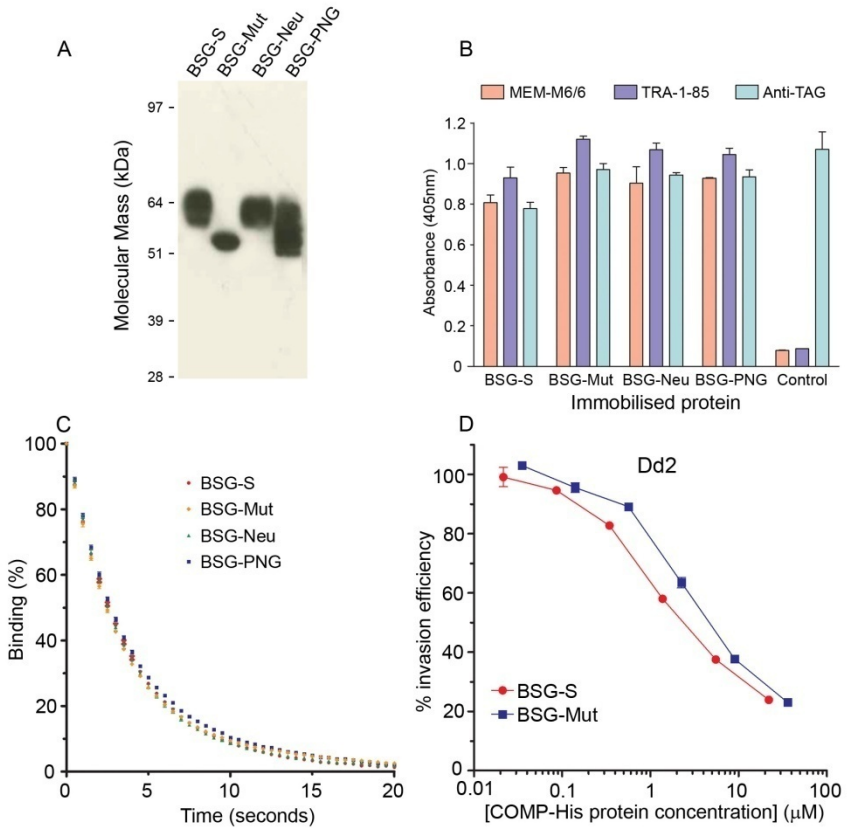
Supplementary Figure 1. An ectodomain protein library of the human erythrocyte. The soluble recombinant monomeric biotinylated baits were expressed by transient transfection of HEK293E cells. Supernatants containing the proteins were harvested, expression levels equalised, resolved under reducing conditions by SDS-PAGE, blotted, and probed using streptavidin-HRP. The expected molecular weight for the recombinant protein is indicated in brackets above each lane; this includes the Cd4d3+4-bio tag (25kDa) and an average of 3kDa for each predicted N-linked glycosylation site. Some proteins are larger than the expected size (such as the Glycophorins) possibly due to the presence of additional glycosylation or stable oligomeric forms; other proteins show evidence of proteolytic processing such as Cathepsin G and C2CD2L.



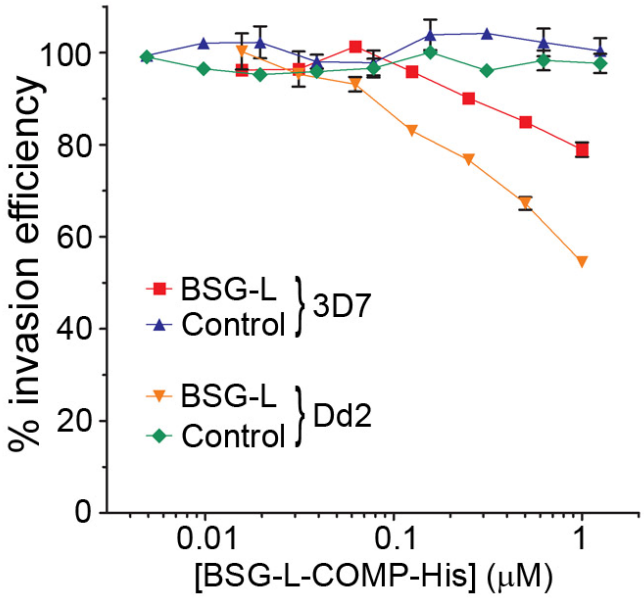
Supplementary Figure 2. Biochemical characterization of the BSG domain truncations. (A) Schematic diagrams of the soluble recombinant domain deletion constructs produced; "TAG" represents the recombinant protein tag used on the bait proteins (rat Cd4d3+4-bio which has a mass of ~25 kDa). A western blot showing expression of the biotinylated soluble recombinant deletion constructs detected using streptavidin-HRP. (B) Reactivity determined by ELISA of a panel of anti-BSG monoclonal antibodies to the domain deletion constructs captured on a streptavidin-coated plate suggesting that the fragments are correctly folded. Means are shown, error bars represent SEM; $n = 3$.



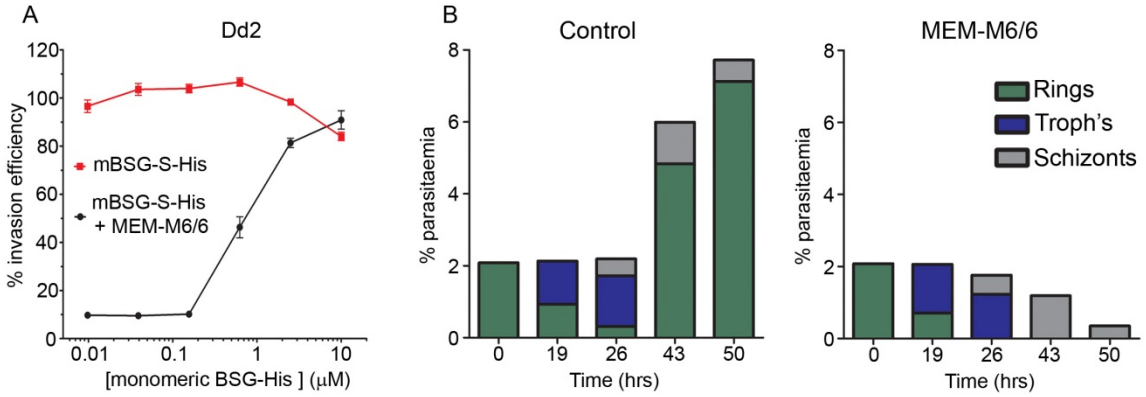
Supplementary Figure 3. Equilibrium binding analysis of the PfRh5-BSG-S interaction using surface plasmon resonance. Serial dilutions of purified PfRh5-Cd4d3+4-His were injected (solid bar, inset) through flow cells containing 150RU Cd4d3+4-bio (used as a reference) and 325 RU (approximate molar equivalent) of BSG-S-Cd4d3+4-bio captured on streptavidin-coated sensor chips. Reference-subtracted binding data were plotted as a binding curve and an equilibrium dissociation constant calculated using non-linear regression fitting of a simple Langmuir binding isotherm to the data (solid line). A K_D of 1.12 ± 0.09 μM (mean ± SEM) was calculated from three experiments with independent protein preparations (Supplementary Table 2). One representative experiment of three is shown.



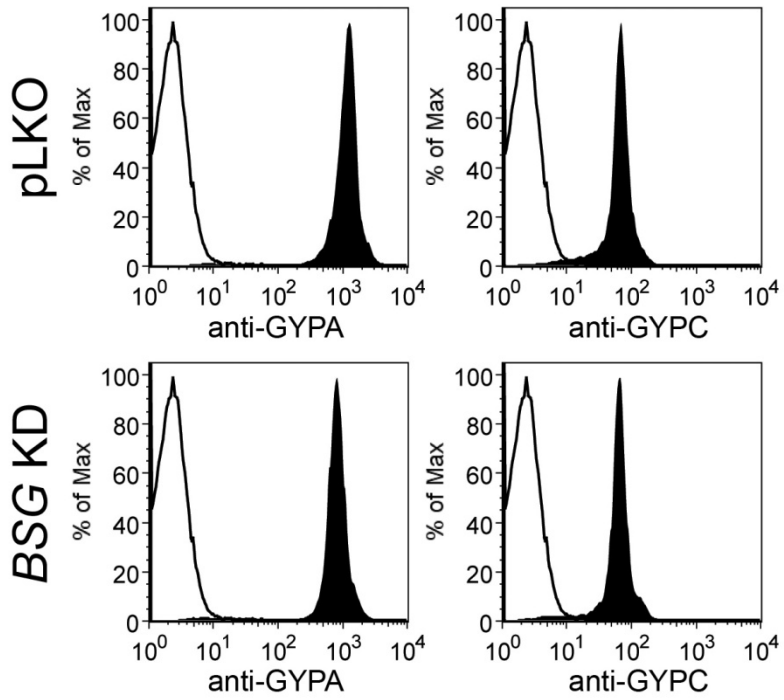
Supplementary Figure 4. Glycans on BSG do not contribute to the binding of PfRh5 or anti-BSG monoclonal antibodies, nor to invasion-blocking activity. (A) Western blot of recombinant soluble monomeric BSG-S-Cd4d3+4-bio (BSG-S) and a mutant form of BSG (BSG-Mut) where all three asparagine residues within the predicted N-linked glycosylation sites were mutated to aspartic acid. The smaller mass of BSG-Mut is consistent with the lack of added glycans. BSG-S-Cd4d3+4-bio was also treated with neuraminidase (BSG-Neu) or PNGaseF (BSG-PNG) which removes sialic acid residues and N-linked glycans, respectively. PNGaseF treatment did not fully remove all glycans from the sample as shown by the presence of intermediate masses between the native (BSG-S) and glycan-free (BSG-Mut) forms. (B) The epitopes for the invasion-blocking anti-BSG monoclonal antibodies, MEM-M6/6 and TRA-1-85, were still present on the glycan-free BSG-Mut or after treatment of BSG-S with neuraminidase or PNGase. (C) Normalised dissociation curves of PfRh5 binding to the different forms of BSG using surface plasmon resonance demonstrated that the modification or absence of glycans on BSG do not quantitatively affect PfRh5 binding. Each data point is the mean ± SEM of three injected PfRh5 concentrations (1.2, 0.6 and 0.3 μM). Only every fifth data point is shown for clarity. (D) Erythrocyte invasion of the *P. falciparum* Dd2 strain was inhibited by both fully glycosylated (BSG-S) and glycan-free (BSG-Mut) pentamers. Each data point shows mean ± SEM, $n = 3$.



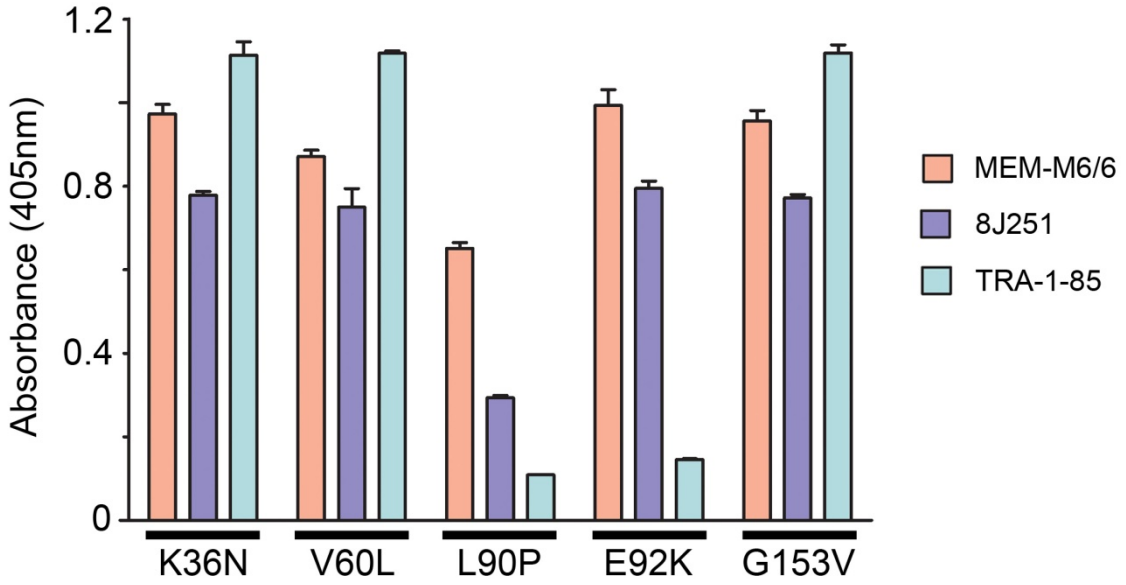
Supplementary Figure 5. Soluble pentamerised BSG-L ectodomains inhibit 3D7 and Dd2 *P. falciparum* erythrocyte invasion. Purified BSG-L-Cd4d3+4-COMP-His (BSG-L-COMP-His) recombinant protein inhibited the invasion of 3D7 and Dd2 *P. falciparum* strains relative to a Cd4d3+4 control in a dose-dependent manner. Each data point shows mean ± SEM, $n = 3$.



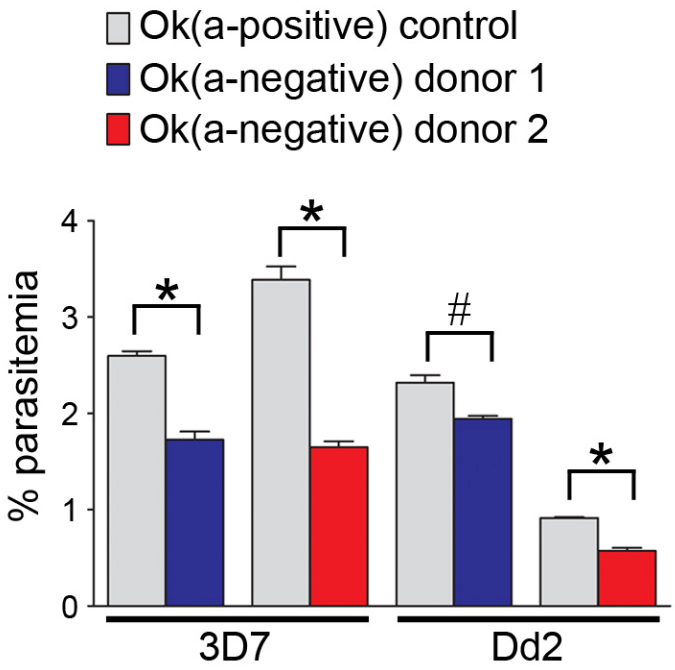
Supplementary Figure 6. The inhibitory effect of anti-BSG monoclonal antibody MEM-M6/6 can be relieved by preabsorption with soluble BSG and specifically blocks schizont to ring stage progression. (A) Serial dilutions of purified monomeric (m)BSG-S-Cd4d3+4-His were added directly or preincubated with 2.5 μg/ml MEM-M6/6 prior to addition to invasion cultures. Increasing concentrations of mBSG-S-Cd4d3+4-His relieved the inhibitory effect of the MEM-M6/6 antibody demonstrating specificity of the antibody to BSG (black circles). Monomeric BSG-S-Cd4d3+4-His which binds with low affinity to PfRh5, in contrast to the more avid pentameric BSG-S-Cd4d3+4-COMP-His (see Fig. 2a, b), did not inhibit invasion except by a small amount at high concentrations (red squares). Each data point shows mean ± SEM, n = 3. (B) Synchronised 7G8 *P. falciparum* blood ring-stage cultures were left untreated (control, left panel) or treated with 10 μg/ml MEM-M6/6 monoclonal antibody (right panel). Cultures were analysed by direct microscopic observation of Giemsa-stained thin smears over the course of 50 hours and the stages of parasite development quantified (green = rings, blue = trophozoites, grey = schizonts). The culture in the presence of the MEM-M6/6 antibody was able to progress from rings to schizonts, but, in contrast to the control, no further ring-stage parasites could be observed after 26 hours suggesting that the antibody prevented schizont to ring-stage progression by blocking invasion.



Supplementary Figure 7. BSG-knockdown erythrocytes are fully mature. Erythrocytes differentiated from hematopoietic stem cells transduced with lentiviruses containing shRNA targeting *BSG* (lower panels) or sequence-scrambled control (pLKO, top panels) expressed equivalent levels of the terminal differentiation erythrocyte markers, GYPA and GYPC.



Supplementary Figure 8. Reactivity of the BSG-S variants to anti-BSG monoclonal antibodies. Each variant of BSG-S was expressed as a soluble recombinant biotinylated bait, captured on a streptavidin-coated plate before being probed with the anti-BSG monoclonal antibodies shown. K36N, V60L and G153V bound all antibodies suggesting that they were correctly folded. TRA-1-85, the antibody used to distinguish the Ok^{a-} blood group did not bind the polymorphism responsible for this blood group, E92K, as expected. Variant L90P showed reduced binding of domain 1-specific antibodies 8J251 and TRA-1-85 suggesting misfolding of this domain although reactivity to the anti-domain 2-specific antibody, MEM-M6/6 was only slightly reduced. Data points are means; error bars represent SEM; $n = 3$.



Supplementary Figure 9. *P. falciparum* invasion is reduced in Ok^{a-} compared to Ok^{a+} erythrocytes: repeat assay. A repeat of the erythrocyte invasion assays for both Ok^{a-} donors and their controls. The reduced levels of parasitemia relative to the first assay (Figure 3d) most likely reflect a reduction in the quality of the sample. Invasion frequencies (% parasitemia) of Ok^{a-} blood cells by *P. falciparum* strains 3D7 and Dd2 were reduced relative to the Ok^{a+} control in two unrelated donors. Mean ± SEM, $n = 3$; *, $P \leq 0.0004$; #, $P = 0.0052$, unpaired one-tailed *t* test.

No.	Receptor name	Trunc.	Type	Accession number	Blood group
1	GLYCOPHORIN A	P90	I	P02724	MNS
2	GLYCOPHORIN B	P60	I	P06028	MNS
3	GLYCOPHORIN E	N51	I	P15421	
4	GLYCOPHORIN C	D58	III	P04921	Gerbich
5	MIC2/CD99	P124	I	P14209	Xg
6	CD44	P648	I	P16070	Indian
7	CD44_isoform 12	P267	I	P16070-12	Indian
8	COMPLEMENT RECEPTOR 1_S	L2423	I	NP_000642	Knops
9	BASIGIN_L	L322	I	P35613_1	Ok
10	BCAM	G548	I	P50895	Lutheran
11	ICAM4/CD242	T239	I	Q14773	Lands.-Wiener
12	AMIGO2	T398	I	Q86SJ2	
13	DSCAML1_1	E1585	I	Q8TD84_1	
14	NICASTRIN	L670	I	Q92542	
15	ERMAP	S154	I	Q96PL5	Scianna/Radin
16	THIOREDOXIN-RELATED TM	P184	I	Q9H1E5	
17	C1ORF9	R1010	I	Q9UBS9	
18	JAM-A	V238	I	Q9Y624	
19	NEUROPLASTIN	P223	I	Q9Y639	
20	PROGESTERON RECEPTOR COMPONENT 2	A46	I	O15173	
21	LFA3/CD58	L218	I	P19256	
22	ENDOD1	P338	I	O94919	
23	SEMAPHORIN7A	A644	GPI	O75326	JMH
24	DAF	S353	I/GPI	P08174	Cromer
25	MACIF/CD59	N102	GPI	P13987	
26	ECTO-ADP-RIBOSYLTRANSFERASE 4/NAR4	A285	GPI	Q93070	Dombrock
27	ACETYLCHOLINESTERASE_ISOFORM H	T581	GPI	P22303-2	Yt
28	MULTIDRUG RESISTANCE PROTEIN 1/MRP1	T36	17TM	P33527-2	
29	CD47_ISOFORM OA3-293	P139	5TM	Q08722	
30	DUFFY_ISOFORM1	P63	7TM	Q16570-2	
31	PROLACTIN	n.a.	sec.	P01236	
32	LACTOTRANSFERRIN	n.a.	sec.	P02788	
33	CATHEPSIN G	n.a.	sec.	P08311	
34	C8ORF55	n.a.	sec.	Q8WUY1	
35	METHYLTRANSFERASE-LIKE PROTEIN 7A	n.a.	sec.	Q9H8H3	
36	GALECTIN-3/LGALS3	n.a.	sec.	P17931	
37*	KELL BLOOD GROUP GLYCOPROTEIN	F69-W732	II	P23276	Kell
38*	NEUTRAL CHOLESTEROL ESTER HYDROLASE	L23-L408	II	Q6PIU2	
39*	ADIPOCYTE PLASMA MEMBRANE-ASSOC	P64-V416	II	Q9HDC9	
40*	C2 DOMAIN-CONTAINING PROTEIN2-LIKE	Q32-L706	II	O14523	

Supplementary Table 1. The human erythrocyte ectodomain protein library. Listed for each protein within the library is: the ectodomain truncation residue (Trunc.); the protein type (I= type I, II = type II, III = type III, GPI = GPI-linked, #TM = multipass transmembrane protein with number of predicted transmembrane spanning regions, sec = secreted); accession number, and associated blood group. * type II proteins were produced as baits only; n.a. = not applicable.

			Equilibrium analysis				Kinetic analysis								
							Association rate constant				Dissociation rate constant				
	Ligand	Exp. #	K_D (μM)	fit error	Mean (μM)	SEM	k_a ($M^{-1}s^{-1}$)	fit error	Mean ($M^{-1}s^{-1}$)	SEM	k_d (s^{-1})	fit error	Mean (s^{-1})	SEM	K_{Dcalc} (μM)
BSG-S	Reference	1	1.28	± 0.03			197000	± 1000			0.251	± 0.002			
		2	1.10	± 0.01	**1.12**	0.09	212600	± 900	**212000**	8000	0.2637	± 0.0006	**0.251**	0.007	1.18
		3	0.984	± 0.002			225000	± 1600			0.2391	± 0.0009			
	K36N	1	1.25	± 0.03			193000	± 8000			0.2417	± 0.0005			
		2	1.10	± 0.02	**1.13**	0.06	220400	± 900	**213000**	10000	0.2632	± 0.0006	**0.259**	0.009	1.22
		3	1.05	± 0.01			226600	± 600			0.2723	± 0.0004			
	V60L	1	0.921	± 0.005			246300	± 200			0.2577	± 0.0002			
		2	1.410	± 0.001	**1.1**	0.1	192000	± 900	**220000**	20000	0.251	± 0.001	**0.261**	0.007	1.19
		3	1.040	± 0.001			216600	± 400			0.2733	± 0.0003			
	L90P	1	no binding				no binding				no binding				
		2	no binding				no binding				no binding				
	E92K	1	2.8	± 0.1			139800	± 400			0.3633	± 0.0006			
		2	2.48	± 0.05	**2.5**	0.2	159400	± 400	**156000**	9000	0.3907	± 0.0006	**0.377**	0.008	2.42
		3	2.17	± 0.04			169500	± 300			0.3758	± 0.0002			
	G153V	1	1.02	± 0.03			205000	± 1000			0.247	± 0.001			
		2	1.33	± 0.02	**1.17**	0.09	192000	± 900	**202000**	5000	0.251	± 0.001	**0.254**	0.005	1.26
		3	1.152	± 0.008			208500	± 500			0.2636	± 0.0005			
BSG-L	Reference	1	0.71	± 0.02			208200	± 800			0.1436	± 0.0003			
	K36N	1	0.60	± 0.02			236600	± 60			0.13760	± 0.00004			
	V60L	1	0.72	± 0.03			238100	± 700			0.1489	± 0.0002			
	L90P	1	no binding				no binding				no binding				
	E92K	1	0.91	± 0.02			167800	± 500			0.2241	± 0.0004			
	G153V	1	0.490	± 0.001			208200	± 800			0.131	± 0.003			

Supplementary Table 2. Summary of the biophysical binding data for PfRh5 binding BSG-S and BSG-L and their naturally-occurring sequence variants. Both the equilibrium and kinetic measurements were calculated from surface plasmon resonance studies using serial dilutions of PfRh5 as an analyte and BSG as the immobilised ligand. Experimental details are described in the Methods. For BSG-S and its variants (except L90P), the experiment was performed three times (Exp. #) using independently produced protein samples. The parameters from each experiment are derived by globally fitting a simple 1:1 binding model to a family of binding curves produced from a dilution series of the PfRh5 protein. For BSG-L and its variants, the experiment was performed once. The k_d:k_a ratio (K_{Dcalc}) calculated from the kinetic data provides an independent measurement of the equilibrium dissociation constant and agrees well with the values obtained from the equilibrium binding.

	PfRh5 Amino acid position									
Parasite strain	**48**	**197**	**203**	**204**	**347**	**358**	**362**	**407**	**410**	**429**
3D7	E	S	C	I	N	Y	E	I	I	K
W2mef	E	S	C	I	N	Y	E	I	I	K
Dd2	E	S	C	I	N	Y	E	I	M	K
Tm284	E	S	C	I	N	Y	E	I	M	K
7G8	E	S	Y	I	N	Y	E	I	I	K
HB3	E	S	Y	I	N	Y	E	I	I	K
Haiti	E	S	Y	I	N	Y	E	I	I	K
JAV	E	S	Y	I	N	Y	E	I	I	K
Ecu110	E	S	Y	I	N	Y	E	I	I	K
ECP	E	S	Y	I	N	Y	E	I	I	K
Liberia II	E	S	Y	I	N	Y	E	I	I	K
D10	E	S	Y	I	N	Y	E	I	I	K
Mont S1	E	Y	Y	I	N	Y	E	I	I	K
PF120	E	Y	Y	I	N	Y	E	I	I	K
BT3	E	Y	Y	I	N	Y	E	I	I	N
FCR3	E	Y	Y	I	N	F	E	I	I	K
Geneve	E	Y	Y	I	N	F	E	I	I	K
FCR-1/FVO	E	Y	Y	I	N	F	D	I	I	K
Malayan Camp	K	S	Y	I	Y	Y	E	I	I	N
Palo Alto	K	S	Y	R	N	Y	E	I	I	N
GB4	E	S	Y	K	N	Y	E	V	I	K
Santa Lucia	E	S	Y	I	D	Y	E	I	I	K

Supplementary Table 3. Position and identity of PfRh5 protein sequence variants. The table is adapted from Table 1 Hayton *et al.* Cell Host Microbe 65 p231 2007 and includes data from Baum *et al.* Int. J. Parasitol. 39 p371 2009. The PfRh5 sequence from FCR3 was obtained from this study. Strain names highlighted in bold were used in the invasion blocking experiment shown in Figure 2e. Shaded residues differ from the 3D7 strain reference sequence.

SNP number	nucleotide variant	Amino acid polymorphism	Location of polymorphism
rs11551906	A>G	T16A	Signal peptide
rs14704	G>T	K36N	Domain 1
rs2229662	G>T	V60L	Domain 1
rs55911144	T>C	L90P	Domain 1
rs104894669	G>A	E92K	Domain 1
rs1803203	G>T	G153V	Domain 2
rs55805128	C>T	R203C	Membrane proximal

Supplementary Table 4. Non-synonymous single nucleotide polymorphisms (SNPs) located within the predicted extracellular regions of BSG. Seven non-synonymous polymorphisms have been identified in human BSG, five of which are located within the immunoglobulin superfamily (IgSF) domains (see Fig. 3a). The two variants T16A and R203C were not included in the analysis since they are located outside of the IgSF domains and are therefore unlikely to affect PfRh5 binding. Population frequency data exist for only two of these SNPs: rs104894669 – the Ok^{a-} blood group identified in eight Japanese families, and rs2229662. rs2229662 has only been reported as a heterozygote with minor allele (T) frequency between 0.01 and 0.06; elevated frequencies are seen in CEU (Utah residents with Northern and Western European ancestry) = 0.05 and MEX (Los Angeles residents with Mexican ancestry) = 0.06. All amino acid positional numbering is based on BSG-S: domain 1: residues 25-103 and domain 2: residues 105-199.