Structural basis for ligand and innate immunity factor uptake by the trypanosome haptoglobin-haemoglobin receptor

Harriet Lane-Serff[a], Paula MacGregor[b], Edward D. Lowe[a], Mark Carrington[b,1] and Matthew K. Higgins[a,1]

[a] Department of Biochemistry, South Parks Road, OX1 3QU, Oxford

[b] Department of Biochemistry, Tennis Court Road, CB2 1QW, Cambridge

[1]To whom correspondence should be addressed. E-mails: mc115@cam.ac.uk or matthew.higgins@bioch.ox.ac.uk

The authors declare that no competing interests exist

Abstract

The haptoglobin-haemoglobin receptor (HpHbR) of African trypanosomes allows acquisition of haem and provides an uptake route for trypanolytic factor-1, a mediator of innate immunity against trypanosome infection. Here we report the structure of *Trypanosoma brucei* HpHbR in complex with human haptoglobin-haemoglobin (HpHb), revealing an elongated ligand-binding site that extends along its membrane distal half. This contacts haptoglobin and the β-subunit of haemoglobin, showing how the receptor selectively binds HpHb over individual components. Lateral mobility of the glycosylphophatidylinositol-anchored HpHbR, and a ~50° kink in the receptor, allows two receptors to simultaneously bind one HpHb dimer. Indeed, trypanosomes take up dimeric HpHb at significantly lower concentrations than monomeric HpHb, due to increased ligand avidity that comes from bivalent binding. The structure therefore reveals the molecular basis for ligand and innate immunity factor uptake by trypanosomes, and identifies adaptations that allow efficient ligand uptake in the context of the complex trypanosome cell surface.

Introduction

African Animal Trypanosomiasis is one of the major constraints on the productivity of pastoralists in sub-Saharan Africa and can be caused by infection by a range of trypanosome species (Shaw, 2004), while infections of humans are caused by only two subspecies of *Trypanosoma brucei* (Laveran, 1902; Pays E and Vanhollebeke, 2009). The disease is persistent as the host immune system is usually unable to clear the infection. This is due to the trypanosome having evolved a population survival strategy based on autoregulation of parasitaemia and antigenic variation (MacGregor et al., 2011; Horn, 2014). The trypanosomes also internalize and degrade surface bound immunoglobulin (Engstler et al., 2007; Pal et al., 2003), increasing the survival of an individual cell and thereby increasing the likelihood of transmission. Both of these strategies require a densely packed cell surface coat of variant surface glycoprotein (VSG) that acts as a barrier, preventing access of host immunoglobulins to the plasma membrane (Schwede et al., 2010). This coat also undergoes antigenic variation through expression of a single VSG gene from a genomic repertoire of hundreds (Horn, 2014).

Although the VSG coat restricts immunoglobulin access, it must be permissive for receptor-mediated binding and uptake of macromolecular ligands. *T. brucei*, and the closely related *T. congolense*, have receptors for both transferrin (TfR) for iron (Schell et al., 2011; Steverding et al., 1994; Jackson et al., 2013) and haptoglobin-haemoglobin (HpHbR) for haem (Vanhollebeke et al., 2008; Higgins et al., 2013). These are held on the external face of the plasma membrane by

covalent attachment of the C-terminal carboxyl group to a glycosylphosphatidyl-inositol to form a GPI-anchor. All have free movement in the lateral plane of the membrane although the receptors are concentrated in the flagellar pocket, an invagination of the plasma membrane at the base of the flagellum and the site of all endocytosis (Mussmann et al., 2004; Vanhollebeke et al., 2008).

Humans, together with a few other primates, display innate immunity to most trypanosome species (Laveran, 1902) through the action of trypanolytic factors-1 and -2 (TLF1 and TLF2) (Raper et al., 1996; Raper et al., 1999; Hager et al., 1994). Although containing different scaffold components, these factors both include apolipoprotein L1 (ApoL1) together with complexes of haemoglobin bound to haptoglobin related protein (HprHb) (Vanhamme et al., 2003; Pérez-Morga et al., 2005). TLF1 enters trypanosomes via receptor-mediated endocytosis, through binding of the HprHb component to HpHbR (Widener et al., 2007; Drain et al., 2001; Vanhollebeke et al., 2008). This delivers ApoL1 to the endosome where it causes lysosomal swelling and cell death (Pérez-Morga et al., 2005). In contrast, the uptake route for TLF2 is unclear as, unlike TLF1, it is able to kill HpHbR null mutants (Capewell et al., 2013; Uzureau et al., 2013).

Just two subspecies of *T. brucei* (*T. b. rhodesiense* and *T. b. gambiense*) have evolved counter measures to the Trypanolytic Factors, allowing them to cause Human African Trypanosomiasis (Pays et al., 2014). In the case of human-infective group 1 *T. b. gambiense,* a unique point polymorphism is found in HpHbR (Symula et al., 2012) that reduces the monovalent affinity for ligand by 20-fold (Higgins et al., 2013). This contributes to resistance to TLF1, illustrating

98 the importance of HpHbR.

99

100 Haptoglobin-haemoglobin is an elongated 'dumbell-shaped' complex consisting
101 of a dimer of haptoglobin molecules, each joined to an αβ haemoglobin dimer
102 (Andersen et al., 2012). Trypanosomes take up this HpHb complex, but not the
103 individual components (Vanhollebeke et al., 2008). The structure of the *T.*
104 *congolense* HpHbR is an elongated three helical bundle with a small membrane
105 distal head (Higgins et al., 2013). Residues involved in HpHb binding are part of a
106 small conserved patch ~25 Å below the tip of the receptor, but details of ligand
107 binding and uptake were not characterized.

108

109 Here, we present the structure of *T. brucei* HpHbR. We show that the receptor
110 adopts a similar architecture to its *T. congolense* homologue, but with a ~50°
111 kink a third of the way along from the membrane proximal end. We also present
112 the structure of TbHpHbR in complex with HpHb, revealing the molecular basis
113 for ligand binding and selectivity. Finally we show that the kink allows two
114 independent membrane attached receptors to interact with a single dimeric
115 HpHb molecule, and confirm using cell uptake experiments that this causes
116 dimeric ligand to be taken up with greater efficiency than monomeric ligand.
117 This reveals the molecular basis for the uptake of HpHb and trypanolytic factor-
118 1, and identifies adaptations in the trypanosome receptor that allow efficient
119 ligand uptake in the context of the tightly packed VSG coat.

120

121

122

Results

TbHpHbR binds to the HpSP domain:Hb head structure

To provide detailed molecular knowledge of the mechanism of uptake of haptoglobin-haemoglobin and trypanolytic factor-1 (TLF1), we aimed to determine the structure of *T. brucei* HpHbR (TbHpHbR), alone and bound to a human haptoglobin-haemoglobin complex.

TbHpHbR is longer that its homologue from *T. congolense* due to the presence of an additional C-terminal membrane-proximal domain. We therefore used the previously determined structure of *T. congolense* HpHbR (Higgins et al., 2013) to design a construct containing the corresponding region of TbHpHbR (residues 36-299). This region of the protein is identical in the human infective *T. b. rhodesiense*.

Haptoglobin-haemoglobin consists of a dimer of haptoglobin chains, each interacting with an αβ dimer of haemoglobin, and adopts a dimeric 'dumbell-shaped' architecture (Andersen et al., 2012). At each end, a serine protease (HpSP) domain of haptoglobin forms a stable complex with a haemoglobin dimer. Dimerisation occurs through an interface formed by the CCP domains of haptoglobin, linking together these HpSPHb 'heads'.

Previous studies have shown that TbHpHbR interacts with the HpHb complex, but not with either haptoglobin or haemoglobin alone (Vanhollebeke et al.,

2008), suggesting that that the receptor most likely binds to the heads of HpHb, where its two constituent components come together. We therefore designed a human haptoglobin construct containing just the SP domain (residues 148-406). This was expressed in baculovirus-infected insect cells and was combined with haemoglobin extracted from human blood to assemble HpSPHb complexes. We used surface plasmon resonance to determine the affinity of these HpSPHb complexes for TbHpHbR, and showed binding with an affinity of 0.7 μM (Figure 1 – figure supplement 1), similar to the 1μM affinity observed for intact human HpHb (Higgins et al, 2013).

Proteolytic cleavage of haptoglobin normally occurs in the endoplasmic reticulum after residue R102 but this cleavage event did not occur in the insect cell expressed HpSP domain. However this did not affect the affinity for TbHpHbR. The shortened TbHpHbR construct and the HpSPHb complex therefore interact together with the same affinity as the full-length components, providing reagents for structural determination. These findings also confirm that TbHpHbR binds to the 'head' structure of dimeric HpHb, raising the possibility of two receptors simultaneously interacting with one HpHb complex.

Determination of the structure of TbHpHbR alone and in complex with HpSPHb

To investigate the molecular basis for HpHb binding by TbHpHbR, crystallisation plates were set up for HpSPHb, TbHpHbR and a complex containing TbHpHbR

bound to HpSPHb. Crystals of HpSPHb diffracted to 2.05Å and were of space group $P3_121$ with one complex in the asymmetric unit. Crystals of TbHpHbR diffracted to 1.85 Å resolution and were of space group $P2_1$ with two molecules in the asymmetric unit. Crystals of the TbHpHbR:HpSPHb complex were of space group C2 and diffracted to 3.1 Å resolution with a single complex in the asymmetric unit (Table 1).

The structure of human HpSPHb was determined using molecular replacement with the equivalent region of porcine HpHb (pdb: 4F4O) as a search model. The structure of the TbHpHbR:HpSPHb complex was then determined through molecular replacement using HpSPHb as a search model, allowing a poly-alanine model of TbHpHbR to be built. This model was then used as a molecular replacement search model to determine the structure of TbHpHbR using higher-resolution data obtained from crystals of the receptor alone. Both structures were then completed using iterative cycles of model building and refinement (Table 2).

The structure of the T. brucei *haptoglobin-haemoglobin receptor*

Like *T. congolense* HpHbR, the *T. brucei* receptor is elongated, consisting primarily of a three-helical bundle (Figure 1): helix I (red; residues 42-110), helix II (orange; residues 116-182), and helix V (dark blue; residues 224-296) with a total length of 118Å. At the membrane distal end, the receptor widens to form a compact head structure that includes the N-terminus and a 42-residue

198 loop containing two further helices, helix III (yellow: residues 186-196) and helix
199 IV (green: residues 207-218). The upper part of the structure is extremely
200 similar to that from *T. congolense*, with the membrane distal halves of the two
201 receptors aligning with a root mean square deviation of 1.1 Å (Figure 1 – figure
202 supplement 2).
203
204 The most dramatic difference between the *T. brucei* and *T. congolense* receptors
205 is a ~50° kink in TbHpHbR, located approximately one third of the way along the
206 receptor from the membrane proximal end. Each of the three helices are affected,
207 with the backbone carbonyl groups of Asp88, Ala89, Glu123, Asn124, Asp270
208 and Ala271 no longer forming hydrogen bonds. This kink is not caused by
209 flexibility, but is a rigid feature of the receptor, as it adopts the same
210 confirmation in crystals of receptor alone, and in crystals of its complex with
211 HpSPHb (Figure 2A), and is also observed in molecular envelopes derived from
212 small angle x-ray scattering (Figure 1, Table 3). Instead it is caused by changes in
213 the pattern of hydrophobic and hydrophilic residues around the kink site in each
214 of the three helices. The three long helices of the *T. congolense* receptor are
215 characterised by an alternating pattern of hydrophobic and hydrophilic residues,
216 leading to continuous hydrophobic strips along the length of each helix, that pack
217 in the core of the helical bundle, stabilising its fold. In the *T. brucei* receptor, this
218 pattern is disturbed at each kink site, breaking the organisation of the helix and
219 leading to an alteration in the surface that displays the hydrophobic patch,
220 resulting in the kink (Figure 1C). This stabilises the kink and makes it a rigid
221 feature of the receptor structure.
222

The structure of TbHpHbR in complex with haptoglobin-haemoglobin

The structure of the TbHpHbR:HpSPHb complex reveals an unexpected binding mode in which the ligand-binding surface extends along more than half of the length of the receptor (Figure 2, Figure 2 – Figure Supplement 1). Residues previously identified as playing a role in HpHb binding in TcHpHbR, such as S59 (Higgins et al., 2013), lie ~35 Å from the membrane distal tip of the receptor and directly contact haemoglobin. However, this is the upper part of the binding site, with residues from haptoglobin interacting as far as 70 Å from the membrane distal tip. This arrangement is confirmed by small angle x-ray scattering, with complexes of HpSPHb bound to either *T. brucei* or *T. congolense* receptors showing a similar architecture to that observed in the crystal (Figure 2 – Figure Supplement 2, Table 3).

The haptoglobin-haemoglobin complex covers a total area of ~1250 Å^2 of the receptor and can be divided into two distinct regions (Figure 2B). The membrane distal part, (~745 Å^2) contacts the β-subunit of the haemoglobin dimer with no contacts between the receptor and the haemoglobin α-subunit. The membrane proximal region (~505 Å^2) forms a binding surface for haptoglobin. The involvement of both haemoglobin and haptoglobin in binding explains why the receptor binds HpHb but not haptoglobin alone. Modelling suggests that the lack of haemoglobin binding is due to steric clashes of the receptor with the second αβ dimer of haemoglobin when the β-subunit of a haemoglobin tetramer is

docked onto the receptor with the binding mode observed in the TbHpHbR:HpSPHb complex (Figure 2 – Figure Supplement 3). Therefore the conformation of the receptor and the presence of two distinct binding sites allows the receptor to specifically select HpHb over its two constitutive components.

The haemoglobin β-subunit makes a number of direct interactions, mostly hydrogen bonds, with the receptor (Figure 2C, Table 4). Side chains from helix I of the receptor make the majority of these contacts, with additional interactions from helix II and the loop that links helices III and IV. These features lie along a groove on haemoglobin that is formed by helices C and F of the β-subunit. The haem group also makes direct contacts with the receptor, with the propionate chains contacting residues K56, S59, K164, R199 and Y200 of the receptor. These interactions, mediated by haem, form ~140 Å^2 of the ~745 Å^2 total contact area of Hb.

The haptoglobin subunit also interacts with helix I of the receptor, through a predominantly hydrophobic contact, mediated by three loops that emerge from the C-terminal β-sheet of haptoglobin (Figure 2B, Table 4). The structure of human haptoglobin from this complex aligns with that from porcine Hp with a root mean square deviation of just 0.5 Å and reveals no significant structural change on receptor binding (Figure 2 – Figure Supplement 4). The alignment also confirms that the natural cleavage of Hp does not affect TbHpHbR binding, as residues in the loop that contains the cleavage site are not close to the receptor.

Rather than haptoglobin, trypanolytic factor-1 (TLF1) contains haptoglobin-related protein (Hpr) and binding of HprHb complex to TbHpHbR results in TLF1 uptake. The HprSP domain contains a total of sixteen amino acid substitutions when compared with the HpSP domain. Mapping these onto the structure shows that none of these differences lie in residues that contact the receptor (Figure 3A). Indeed HprSPHb complexes, prepared using the same protocols as HpSPHb complexes, bound to the receptor with an affinity of 1.7 μM, as determined by surface plasmon resonance (Figure 3B), comparable to the 0.7 μM affinity of the receptor for HpSPHb. This suggests that HprHb, and as a result, TLF1, will have a shared binding mode with HpHb.

A model for haptoglobin-haemoglobin uptake in the context of the VSG layer

The haptoglobin-haemoglobin receptor operates in the context of the VSG layer, a dense coat of surface protein that covers the trypansosome surface. It is therefore initially surprising that the location of the binding site for bulky HpHb complexes extends some 70 Å from the membrane distal tip of the receptor, and below the surface of the VSG layer. However, one consequence of the kink in the *T. brucei* receptor is to increase its effective diameter, pushing apart VSG molecules. In addition, the orientation of the kink is precisely arranged to increase exposure of the HpHb binding site to the surface, making it more accessible for ligand binding.

Docking of TbHpHbR:HpSPHb structures onto the structure of dimeric porcine HpHb reveals another consequence of the kink. This modelling suggests that two receptors can bind simultaneously to a single HpHb dimer, resulting in a C-shaped complex with a parallel arrangement of the membrane proximal parts of the two receptors (Figure 4A). Indeed, this arrangement was confirmed in solution by small angle x-ray scattering. Native (dimeric) human HpHb was mixed with TbHpHbR, and gel filtration was performed, with SAXS data collected from samples as they emerged from the column. The resultant scattering curves confirmed the assembly of a complex containing two receptors and one HpHb *in vitro*. This data was used to generate a molecular envelope for the complex, which confirmed the C-shaped architecture (Figure 4B, Figure 4 – Figure Supplement 1, Table 3). Additional support for the formation of this complex in solution came from multi-angle light scattering (SEC-MALLS), which revealed masses of 30kDa for the receptor, 150kDa for HpHb and 210kDa for the complex, showing that two receptors bind to each HpHb in solution (Figure 4 – Figure Supplement 2). This arrangement, in which two independent GPI-anchored receptors can bind simultaneously to an HpHb dimer, would increase avidity for the ligand and decrease the ligand concentration required for efficient uptake.

To test this hypothesis, we performed uptake experiments using *T. brucei*. HpHb (dimeric), HpSPHb (monomeric) and bovine serum albumin (BSA, as a control to assess fluid phase uptake), each fluorescently labelled. In addition we prepared a null cell line, TbHpHbR$^{-/-}$, in which both copies of the receptor were disrupted (Figure 4 – Figure supplement 3). These reagents allowed us to investigate the concentration dependence of ligand uptake. In wild type *T. brucei* cells, uptake of

HpHb reached saturation below a concentration of 4 nM (Figure 4C). In contrast, the uptake of HpSPHb was negligible at a concentration of 4 nM and continued to increase at 62.5 nM (Figure 4D). Uptake of HpSPHb and HpHb observed in wild type cells was due to the TbHpHbR, as expected (Vanhollebeke et al., 2008), as uptake of both ligands into TbHpHbR$^{-/-}$ cells was comparable to that of BSA in the range of ligand concentrations assayed (0-62.5 nM).

Therefore uptake of dimeric HpHb into trypanosomes occurs efficiently at a significantly lower concentration than that of monomeric HpSPHb. As the monovalent affinities of the receptor for HpHb and HpSPHb are indistinguishable, as measured by surface plasmon resonance, this suggests that more efficient uptake of HpHb is caused by the dimeric ligand simultaneously binding to two receptors. Indeed, measurements of the binding of HpHb to immobilised receptor, using a very high surface density of HpHbR to measure bivalent binding, gave an affinity of 4.5 nM (Higgins et al., 2013), which is in the same range as the concentration at which HpHb uptake becomes saturated in live parasites. It therefore appears as though TbHpHbR has evolved a kink to increase accessibility of its ligand-binding site and to allow simultaneous binding of two receptors to one HpHb ligand, increasing ligand avidity and uptake efficiency.

Discussion

The external surface of an African trypanosome is covered with a tightly packed layer of variant surface glycoprotein that shields epitopes that lie close to the plasma membrane from antibody binding (Schwede et al., 2011). Receptors such as those required for the uptake of transferrin and haptoglobin-haemoglobin complexes must operate within the context of this coat, with their structures organised such that ligand binding sites are not masked by the VSG layer.

Here, the first structure of a trypanosome receptor in complex with its ligand is presented: that of the *T. brucei* haptoglobin-haemoglobin receptor bound to haptoglobin-haemoglobin. Remarkably, the ligand-binding site extends more that half the way along the receptor, forming distinct binding surfaces for the β-subunit of haemoglobin and for haptoglobin. The simultaneous binding of both components explains the specificity of the receptor for haptoglobin-haemoglobin complexes over each individual component. However, the extent of the binding surface places it below the top of the VSG layer, apparently increasing the likelihood that it will be masked by VSG.

However, a ~50° rigid kink occurs as an adaptation in the three helical bundle of the *T. brucei* receptor and we propose that it has two main functional consequences. Firstly, the direction of the kink is precisely arranged to bend the receptor such that the ligand-binding site becomes more exposed at the membrane surface. The kink will also increase the effective diameter of the receptor perpendicular to the membrane. This combination is likely to increase

the separation of VSG molecules in the region of the receptor and to increase the accessibility of the binding site for bulky macromolecular ligands such as HpHb and trypanolytic factors. Increased separation of VSG molecules by a trypanosome receptor is not a novel phenomena, with bulky glycan chains attached to the transferrin receptor proposed to have a similar effect (Mehlert et al., 2012), suggesting that different receptors increase the accessibility of binding sites for bulky ligands by different means.

The precise nature of the integration of TbHpHbR into the VSG layer remains unresolved as the effect of the C-terminal domains of both the receptor and VSG on the vertical disposition of each molecule remains unknown. An attractive model is that the ligand, whether HpHb or TLF1, is bound above the top of the VSG. However, the dimensions of the structures of VSG and the TbHpHbR:HpHb complex suggests that this may not be the case and that the HpHb ligand is held at least partially within the VSG layer (Figure 5). The TLF1 ligand is ~four times the size of HpHb and this must, at least partially, protrude above the top of the VSG layer.

A second consequence of the kink is to allow two receptors to simultaneously bind to one dimeric HpHb complex when their membrane-proximal C-termini are membrane attached. The affinity of a single receptor for HpHb is modest, at ~1 μM, with a rapid off-rate. By enabling two receptors to bind simultaneously, the kink will increase the ligand avidity, changing the effective affinity to something in the low nanomolar range, as previously measured for bivalent binding by TbHpHbR (Higgins et al., 2013) and to live cells (Drain et al., 2001).

What is the likelihood of two receptors interacting with a single ligand on the cell surface? The receptor copy number is 200-400 and it is concentrated in the flagellar pocket (Vanhollebeke et al, 2008). The surface area of the flagellar pocket in live cells is 4.3 μm^2 (Grünfelder et al, 2002). If there are 200 receptor molecules in the membrane of the flagellar pocket then the density is one receptor per 0.022 μm^2. The diffusion constant for HpHbR is unknown, but the diffusion constant for another GPI-anchored protein, VSG, has been measured by fluorescence recovery after photobleaching to be 0.01 $\mu m^2/s$ (Bulow et al, 1988); this means that the receptor will contact a receptor molecule approximately every 2 secondS. The $t_{1/2}$ for release of monovalently bound HpHb is 70-100 s (Higgins et al, 2013) (Fig 1 – Figure Supplement 1) so if it is assumed that the receptor has a similar diffusion coefficient to the VSG then it is very likely that a monovalently bound ligand will become bivalently bound. Indeed our analysis of uptake of dimeric HpHb and monomeric HpSPHb into *T. brucei* confirmed that this increased avidity does occur *in vivo*, with HpHb uptake saturating at a concentration below 4 nM, while HpSPHb uptake is far from saturation at 62.5 nM.

Whether other trypanosome receptors use a similar avidity-increase mechanism to improve the efficiency of ligand uptake remains to be seen, and whether it is required will depend upon the receptor affinity for monomer and the sera concentration of the nutrient. However, it is clear from the example of TbHpHbR that this mode of ligand binding is potentially applicable to other GPI-anchored cell surface proteins.

While the evolution of the kink allows increased accessibility of the binding site for HpHb, it would also increase accessibility for the large TLF1 complex. One mechanism used by human infective *T. b. gambiense* to avoid TLF1-mediated innate immunity is a point polymorphism in HpHbR that reduces the monovalent affinity for HpHb by 20-fold (Higgins et al., 2013) and subsequently reduces TLF1 uptake (Kieft et al., 2010). It remains to be seen whether TLF1 contains one or multiple HprHb complexes, and whether these are in a conformation to allow bivalent binding. However, even if bivalent binding does occur in TLF1, the difference in affinity due to the *T. b. gambiense* polymorphism will be amplified under conditions where two receptors bind to a single ligand.

As the kink in the receptor appears to have a number of functional consequences that facilitate ligand uptake, it is surprising that the *T. congolense* receptor lacks such a kink. One reason for this difference might be the lack of a C-terminal domain in *T. congolense* surface proteins (Higgins et al., 2013). Perhaps the direct attachment of the ligand-binding domain to a GPI-anchor provides enough flexibility to allow the receptors to adopt an angle that allows simultaneous uptake, or perhaps the VSG coat of *T. congolense* parasites is less densely packed. These questions will need further study.

In conclusion, we present the first structure of a trypanosome receptor in complex with its ligand and reveal a number of adaptations that are tailored to facilitate efficient ligand binding in the context of the VSG coat. These will decrease the packing of VSG molecules in the immediate vicinity of the receptor

and increase accessibility of the ligand-binding site. They also allow two receptors to bind to a single ligand, thereby increasing avidity and dramatically decreasing the ligand concentration at which efficient uptake occurs. While different adaptations might facilitate each of these goals in different receptors, we would expect them to be general principles, frequently used by the parasite to aid nutrient uptake and survival.

Materials and Methods

T. brucei *HpHbR cloning, expression and purification*

Full-length *T. b. brucei* HpHbR, without the N-terminal signal sequence and C-terminal GPI-anchor addition sequence had been previously cloned for expression in a modified pET-15b to generate a polypeptide with an N-terminal hexahistidine tag and a cleavage site for TEV protease (Higgins et al., 2013). To produce a truncated construct for expression of the N-terminal ligand-binding domain, a stop codon was inserted after residue R299 using a polymerase chain reaction based mutagenesis protocol, using oligonucleotide `GAGATGAAGCGCTAGGGGAACCCGATC` and its reverse-complement. Mutagenesis was carried out as described for the Quikchange mutagenesis method (Stratagene) and the plasmid sequence verified.

The protein was expressed in *E. coli* Origami B, induced with 1 mM IPTG and incubated overnight at 18 °C. The protein was purified by Ni^{2+}-NTA affinity chromatography and cleaved overnight with his-tagged TEV protease at 4 °C in PBS with 3 mM oxidized glutathione, 0.3 mM reduced glutathione, followed by reverse Ni^{2+}-NTA affinity chromatography. The protein was concentrated by Amicon Ultra centrigugal filter device (10,000 MWCO), and gel filtered using a Superdex 75 16/60 column (Life Technologies) into 20 mM HEPES pH 7.5, 150 mM NaCl.

HpSP and HprSP cloning, expression and purification

Synthetic genes encoding the SP domains of human Hp (148-406) and Hpr (90-348) were cloned into a modified pAcGP67A vector to generate a polypeptide with an N-terminal hexahistidine tag and a cleavage site for TEV protease. These were transfected into Sf9 insect cells using the BaculoGold Baculovirus DNA transfection protocol (BD Biosciences). Following selection of virus using plaque assays, the third amplification of recombinant virus was used to infect Sf9 insect cells. After 3 days, the cells were centrifuged for 15 min at 6000 *g*. After filtering, the supernatant was buffer exchanged into 20 mM Tris pH 8, 300 mM NaCl using a tangential flow apparatus (Pall Corporation), followed by Ni^{2+}-NTA affinity chromatography. The protein was concentrated using an Amicon Ultra centrifugal filter device (10,000 MWCO). When used for crystallisation, HpSP was deglycosylated by incubation with endoglycosidase Hf (Sigma) and

endoglycosidase F3 at enzyme:protein ratios of 1:25 in 1 mM $CaCl_2$, 1 mM $MgCl_2$, 100 mM HEPES pH 7.5 at 37 °C for 3 hours.

Purification of Hb and formation of the HpHb complex

Hb was isolated from human blood by sonication, followed by anion exchange chromatography using a Mono Q column (Life Technologies). HpHb was made by mixing full length Hp 1-1 (Sigma) with purified Hb, and isolating the complex by gel filtration using a Superdex 200 16/60 column in 20 mM HEPES pH 7.5, 150 mM NaCl.

HpSPHb and HprSP complex formation

HpSP or HprSP at a three-fold molar excess was mixed with Hb, and diluted five-fold into 20 mM Tris pH 8, 500 mM NaCl, 15 mM imidazole. The complex was purified by Ni^{2+}-NTA affinity chromatography, washed using the dilution buffer, and eluted into PBS containing 200 mM imidazole. The complexes were then concentrated using an Amicon Ultra centrifugal filter device (10,000 MWCO), and purified by gel filteration using a Superdex 200 16/60 column (Life Technologies) in 20 mM HEPES pH 7.5, 150 mM NaCl.

Crystallisation, data collection and structure determination of the HpSPHb complex

HpSPHb was concentrated to 15 mg ml^{-1} for crystallization. Crystals were obtained after eight hours in sitting drop with a well solution containing 0.2 M NaCl, 0.1 M sodium cacodylate pH 6.5 and 2 M ammonium sulphate. These were cryoprotected by transfer into well solution with the addition of 30% v/v glycerol before cryo-cooling using liquid nitrogen. Data were collected on beamline I04-1 at the Diamond light source and were integrated and scaled using iMosflm (Battye et al., 2011) and scala (Evan, 1993) from the CCP4 suite (Winn et al., 2004), giving a final resolution of 2.05 Å. The structure was determined by molecular replacement using Phaser (McCoy et al., 2007) with the equivalent region of the porcine HpHb (pdb: 4F4O) as a search model. The structure was rebuilt and refined using Coot (Emsley et al., 2010) and REFMAC (Murshudov et al., 2011).

Crystallisation, data collection and structure determination of the TbHpHbR:HpSPHb complex

TbHpHbR was mixed with HpSPHb and purified by gel filtration using a Superdex 200 16/600 column into a buffer containing 150 mM NaCl and 20 mM HEPES pH 7.5. It was concentrated to a final concentration of 15 mg ml^{-1} and

crystallised at 18 °C using sitting drops with a well solution containing 12.5% v/v MPD, 0.03 M NaBr, 0.03M NaI, 0.03M NaF, 0.1 M MES/imidazole pH 6.5, 12.5% w/v PEG 1000, 12.5% w/v PEG 3350 from the Morpheus screen (Molecular Dimensions). Crystals formed after 10 days. Seed beads (Hampton Research) were used to create seeds from these crystals. These were used to seed a plate containing 100 nl of protein, 50 nl of the Morpheus well solution, and 50 nl of the Silver Bullet additive screen (Hampton Research). Crystals grew after 8 days in the well containing additives 0.2% w/v 2,2'-thiodiglcolic acid, 0.2% w/v apidic acid, 0.2% w/v benzoic acid, 0.2% w/v oxalic acid anhydrous, 0.2% w/v terephthalic acid. These were cryo-cooled in liquid nitrogen in the Morpheus well solution.

Data were collected on beamline I03 at the Diamond light source. Data reduction was performed using XDS (Kabsch, 2010) and the structure was solved by molecular replacement with the HpSPHb structure as a search model using Phaser (McCoy et al, 2007). Automatic model building in Buccaneer (Cowtan, 2006) was used to identify the positions of the receptor helices, leading to a cycle of model building and refinement in Coot (Emsley et al, 2010) and Buster (Bricogne et al., 2011). The coordinates from the higher resolution structures of both HpHbSP and TbHpHbR, also determined during this study, were used to provide restraints during refinement, leading to improved stereochemistry of the resultant model.

565

566

567 ***Crystallisation, data collection and structure determination of TbHpHbR***

568

569 The receptor was concentrated to 12.5 mg ml^{-1} for crystallization. Crystals were
570 obtained at 18 °C after 7 days using sitting drops with a well solution of 0.15 M
571 KBr, 30 % w/v PEG 2000 MME from the JCSG+ screen (Molecular Dimensions).
572 These were partially dehydrated and cryoprotected by transfer into the well
573 condition with addition of 30% v/v glycerol before cryo-cooling in liquid
574 nitrogen.

575

576 Data were collected on beamline I03 at the Diamond light source. Data reduction
577 was performed using iMosflm (Battye et al., 2011) and scala (Evans, 1993) from
578 the CCP4 data processing suite (Winn et al., 2004). Molecular replacement was
579 performed using Phaser (McCoy et al., 2007), with a the structure of TbHpHbR
580 taken from the TbHpHbR:HpSPHb complex as a search model. A cycle of
581 refinement and model building was carried out using REFMAC (Murshudov et al.,
582 2011) and Coot (Emsley et al., 2010).

583

584

585

586 ***Small-angle X-ray scattering (SAXS) data collection and processing***

587

588 SAXS data for TbHpHbR alone and in complex with HpSPHb, and for the complex
589 of TcHpHbR with HpSPHb, were collected at the PetraIII P12 beamline at

590 Deutsches Elektronen-Synchrotron using a wavelength of 1.24 Å. SAXS data for
591 the complex of the receptor with dimeric HpHb were collected at beamline BM29
592 at the European Synchrotron Radiation Facility using a wavelength of 0.9 Å. SAXS
593 data for HpHb alone was collected at beamline B21 at the Diamond Light Source
594 with a wavelength of 1.0 Å. In all cases, scattering was detected using a Pilatus
595 image reader at 20 °C.
596
597 The receptors alone and in complex with HpSPHb, as well as HpHb alone, were
598 prepared at concentrations of 2.0, 1.0, 0.5, 0.25 and 0.125 mg ml^{-1} in 20 mM
599 HEPES pH 7.5, 150 mM NaCl. Twenty consecutive frames of 10 s each were
600 recorded for each protein sample with a buffer sample measured between each,
601 except HpHb, for which 180 consecutive 1 s frames were taken. . Any images
602 where the data had been affected by protein radiation damage were excluded
603 from further processing.
604
605 The complex with dimeric HpHb was prepared and analysed by size-exclusion
606 column (SEC)-SAXS, using a Superdex 200 10/300 column (Life Technologies) in
607 20 mM HEPES pH 7.5, 150 mM NaCl with a running speed of 0.4 ml min^{-1}. One
608 frame was collected every 2 s. Frames corresponding to the peak seen on the UV
609 trace were selected, and a curve representing the scattering due to buffer was
610 produced by averaging ten frames from the beginning of the run.
611
612 For each dataset, PRIMUS (Konarev et al., 2003; Petoukhov et al., 2007) was used
613 to normalize data to the intensity of the incident beam, for averaging of
614 equivalent images, and to subtract scattering due to buffer. Where Guinier plots

615 revealed aggregation due to high concentration, data were removed (Guinier et
616 al., 1955). Composite curves were generated by scaling and merging the data
617 sets.
618
619 AutoRg calculated the distance distribution function (P(r)) using an indirect
620 Fourier transform, allowing estimation of the radius of gyration (R_g), the
621 maximum particle dimension (D_{max}) and the Porod volume (Porod,1982) by
622 GNOM (Petoukhov et al., 2007). Initial models of the shape were generated using
623 DAMMIF (Franke et al., 2009), and averaged using the DAMAVER programme
624 suite (Konarev et al., 2006). DAMMIN then produced a final model by minimising
625 differences between experimental data and scattering of the model. The
626 envelope model was produced using Situs (Birmanns et al., 2011), and feature-
627 based docking of the crystal structures was completed using Sculptor (Birmanns
628 et al., 2011).
629
630
631 ***Surface plasmon resonance***
632
633 Measurements were performed on a Biacore T200 instrument with a constant
634 flow rate of 30 μg ml^{-1}. A CM5 chip was prepare by flowing over a 1:1 mixture of
635 ethyl-dimethylaminopropyl-carbodiimide and N-hydroxysuccinimide. TbHpHbR
636 was diluted into 10 mM sodium acetate pH5 to a final concentration of 0.1 μM.
637
638 Ligands were diluted into HBS (20mM HEPES pH 7.5, 150mM NaCl, 0.005% v/v
639 Tween 20). Both channels were equilibrated with HBS before injection of

640 binding partner, and the level of specific binding obtained from a subtraction of
641 the response from channel 2 from that of channel 1. Values for K_D were obtained
642 by equilibrium binding analysis using the BIAevaluation software.
643
644 ***Size exclusion chromatography-multiangle laser light scattering***
645 Purified samples were loaded onto a Superose 6 10/300 column (GE
646 Healthcare), then analysed using a laser light scattering detected at 662 nm
647 wavelength at 8 scattering angles between 20.6° and 149.1° using a Heleos 8
648 instrument (Wyatt Technology). ASTRA 6.1 (Wyatt Technology) was used to
649 calculate molecular weights using the Zimm equation. The samples were loaded
650 at concentrations of 10 μM for HpHb and 40 μM for TbHpHbR.
651
652
653 ***Trypanosome cell culture and construction of* T. brucei *HpHbR$^{-/-}$ cell line***
654
655 *T. brucei* blood stream form cells were grown in HMI-9 at 37 °C with 5% CO_2
656 (Hirumi and Hirumi, 1989). The linear DNAs used to replace HpHbR genes by
657 homologous recombination were produced by PCR. First, one allele was replaced
658 in procyclic form Lister 427 cells using a PCR product that contained 80 bp from
659 upstream of the HpHbR gene followed by the blasticin resistance cassette
660 followed by 80 bp from downstream of the HpHbR gene. The same approach was
661 used with a G418 resistance cassette. Genomic DNA was prepared from blasticin
662 or G418 resistant cell lines and used as a template for a PCR using
663 oligonucleotides 500 bp upstream and downstream of the HpHbR gene. These

second PCR products were used to serially transfect Lister 427 bloodstream form cells expressing VSG118.

Uptake assays for monitoring uptake of fluorescently labelled ligands into live cells

When used for uptake experiments, Hp, HpSP and BSA were labelled with Alexa Fluor 488 using the protein labelling kit (Life Technologies). The manufacturer's protocol was adapted, extending the reaction time to overnight at 4 °C to increase labelling efficiency. Hp and HpSP were then subsequently mixed with Hb to form complex as above. For each assay, $5x10^6$ wild type Lister 427 or *HpHbR*$^{-/-}$ cells were resuspended in 100 µl of serum free HMI-9 with 1% BSA and incubated with 2 µM protease inhibitor FMK-024 for 10 minutes at 37 °C. Cells were then incubated with 1-62.5nM fluorescently labelled protein for 2 hours at 37 °C before being washed twice in serum free HMI-9 with 1% BSA. Cells were fixed in 4% paraformaldehyde for 10 minutes at room temperature and resupended in PBS. Uptake was assayed by flow cytometry using a BD FACScan™ and quantified on FlowJo software. Mode increase in fluorescence was measured relative to a no ligand negative control and all assays were carried out in triplicate.

Acknowledgements

This work was supported by Medical Research Council Project Grant MR/L008246. HL is funded by the Wellcome Trust PhD program in Structural Biology. MKH is a Wellcome Investigator. We thank David Staunton for help with biophysical methods and Paul McKean (Lancaster University) for the gene deletion method. The structures are deposited in the protein data bank as 4X0I, 4X0J and 4X0L.

References

Andersen CB, Torvund-Jensen M, Nielsen MJ, de Oliveira CL, Hersleth HP, Andersen NH, Pedersen JS, Andersen GR, Moestrup SK. 2012. Structure of the haptoglobin-haemoglobin complex. *Nature* **489**:456-9.

Battye TG, Kontogiannis L, Johnson O, Powell HR, Leslie AG. 2011. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr*. **D67**:271–281.

Birmanns S, Rusu M, Wriggers W. 2011. Using Sculptor and Situs for simultaneous assembly of atomic components into low-resolution shapes. *J. Struct. Biol.* **173**:428–435.

Bricogne G, Blanc E, Brandl M, Flensburg C, Keller P, Paciorek W, Roversi P, Sharff A, Smart OS, Vonrhein C, Womack TO. 2011. *BUSTER version 2.10.0*. Global Phasing Ltd., Cambridge, UK.

712 Bulow R, Overath P, Davoust J. 1988. Rapid lateral diffusion of the variant surface
713 glycoprotein in the coat of *Trypanosoma brucei*. *Biochemistry* **27**:2384-2388.

714 Capewell P, Clucas C, DeJesus E, Kieft R, Hajduk S, Veitch N, Steketee PC, Cooper
715 A, Weir W, MacLeod A. 2013. The TgsGP gene is essential for resistance to human
716 serum in *Trypanosoma brucei gambiense*. *PLoS Pathog.* **9**:e1003686.

717 Cowtan K. 2006. The Buccaneer software for automated model building. *Acta*
718 *Crystallogr.* **D62**:1002–1011.

719 Drain J, Bishop JR, Hajduk SL. 2001. Haptoglobin-related protein mediates
720 trypanosome lytic factor binding to trypanosomes. *J. Biol. Chem*. **276**:30254-60.

721 Emsley P, Lohkamp B, Scott WG, CowtanK. 2010. Features and development of
722 Coot. *Acta Crystallogr*. **D66**:486–501.

723 Engstler M, Pfohl T, Herminghaus S, Boshart M, Wiegertjes G, Heddergott N,
724 Overath P. 2007. Hydrodynamic flow-mediated protein sorting on the cell
725 surface of trypanosomes. *Cell* **131**: 505-5.

726 Evans PR. Data reduction. 1993. In: Sawyer L, Isaacs N, Bailey S, editors.
727 *Proceedings of the CCP4 Study Weekend*. Daresbury Laboratory, Warrington, UK.
728 p114–122.

729 Franke D, Svergun DI. 2009. DAMMIF, a program for rapid ab-initio shape
730 determination in small-angle scattering. *J. Appl. Crystallogr*. **42**:342–346.

731 Grünfelder CG, Engstler M, Weise F, Schwarz H, Stierhof YD, Boshart M, Overath
732 P. 2002. Accumulation of a GPI-anchored protein at the cell surface requires

733 sorting at multiple intracellular levels. *Traffic*. **3**:547-559

734 Guinier A, Fournet G. 1955. *Small-Angle Scattering of X-rays*. Wiley.

735 Hager KM, Pierce MA, Moore DR, Tytler EM, Esko JD, Hajduk SL. 1994.
736 Endocytosis of a cytotoxic human high density lipoprotein results in disruption
737 of acidic intracellular vesicles and subsequent killing of African trypanosomes. *J.*
738 *Cell Biol*. **126**:155-67.

739 Higgins MK, Tkachenko O, Brown A, Reed J, Raper J, Carrington M. 2013.
740 Structure of the trypanosome haptoglobin-hemoglobin receptor and implications
741 for nutrient uptake and innate immunity. *Proc Natl Acad Sci U S A* **110**:1905-10.

742 Hirumi H, Hirumi K. 1989. Continuous cultivation of *Trypanosoma brucei* blood
743 stream forms in a medium containing a low concentration of serum protein
744 without feeder cell layers. *J Parasitol*. **75**:985-9.

745 Horn D. 2014. Antigenic variation in African trypanosomes. *Mol Biochem*
746 *Parasitol*. **195:**123-9.

747 Jackson AP, Allison HC, Barry JD, Field MC, Hertz-Fowler C, Berriman M. 2013. A
748 cell-surface phylome for African trypanosomes. *PLoS Negl Trop Dis.* **7**:e2121.

749 Kabsch W. 2010. XDS. *Acta Cryst.* **D66**:125-132.

750 Kieft R, Capewell P, Turner CM, Veitch NJ, MacLeod A, Hajduk S. 2010.
751 Mechanism of *Trypanosoma brucei gambiense* (group 1) resistance to human
752 trypanosome lytic factor. *Proc Natl Acad Sci U S A.* **107**:16137-41.

753 Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI. 2003. PRIMUS: a

754 Windows PC-based system for small-angle scattering data analysis. *J. Appl.*
755 *Crystallogr*. **36**:1277–1282.

756 Konarev PV, Petoukhov MV, Volkov VV, Svergun DI. 2006. ATSAS 2.1, a program
757 package for small-angle scattering data analysis. *J. Appl. Crystallogr*. **39**:277–286.

758 Laveran A. 1902. De l'action du sérum humain sur le trypanosome de Nagana (*Tr.*
759 *brucei*). *C. R. Acad. Sci*. **134**: 735-739.

760 Leslie AG. 2006. The integration of macromolecular diffraction data. Acta
761 Crystallogr. **D62**:48–57.

762 MacGregor P, Savill NK, Hall D, Matthews KR. 2011. Transmission stages
763 dominate trypanosome within-host dynamics during chronic infections. *Cell Host*
764 *Microbe* **9**:310-8.

765 McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. 2007.
766 Phaser crystallographic software. *J. Appl. Crystallogr.* **40**:658–674.

767 Mehlert A, Wormald MR, Ferguson MA. 2012. Modeling of the N-glycosylated
768 transferrin receptor suggests how transferrin binding can occur within the
769 surface coat of *Trypanosoma brucei*. *PLoS Pathog*. **8**:e1002618.

770 Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls A, Winn
771 MD, Long F, Vagin AA. 2011. REFMAC5 for the refinement of macromolecular
772 crystal structures. *Acta Crystallogr.* **D67**:355–367.

773 Mussmann R, Engstler M, Gerrits H, Kieft R, Toaldo CB, Onderwater J, Koerten H,
774 van Luenen HG, Borst P. 2004. Factors affecting the level and localization of the
775 transferrin receptor in *Trypanosoma brucei*. *J Biol Chem* **279**:40690-8.

776 Pal A, Hall BS, Jeffries TR, Field MC. 2003. Rab5 and Rab11 mediate transferrin
777 and anti-variant surface glycoprotein antibody recycling in *Trypanosoma brucei.*
778 *Biochem J* **374**:443-51.
779
780 Pays E, Vanhollebeke, B 2009. Human innate immunity against African
781 trypanosomes. *Curr. Opin. Immunol*. **21**:493-498.
782
783 Pays E, Vanhollebeke B, Uzureau P, Lecordier L, Pérez-Morga D. 2014. The
784 molecular arms race between African trypanosomes and humans. *Nat Rev*
785 *Microbiol.* **12**:575-84.
786
787 Pérez-Morga D, Vanhollebeke B, Paturiaux-Hanocq F, Nolan DP, Lins L, Homblé F,
788 Vanhamme L, Tebabi P, Pays A, Poelvoorde P, Jacquet A, Brasseur R, Pays E.
789 2005. Apolipoprotein L-I promotes trypanosome lysis by forming pores in
790 lysosomal membranes. *Science* **309**:469-72.

791 Petoukhov MV, Konarev PV, Kikhney AG, Svergun DI. 2007. ATSAS 2.1 - towards
792 automated and web-supported small-angle scattering data analysis. *J. Appl.*
793 *Crystallogr*. **40**:s223–s228.

794 Porod G. 1982. *General Theory Small Angle X-ray Scattering*. Academic Press.

795 Raper J, Nussenzweig V, Tomlinson S. 1996. The main lytic factor of

796 *Trypanosoma. b. brucei* in normal human serum is not high density lipoprotein. *J.*
797 *Exp. Med.* **183**:1023-1030.

798 Raper J, Fung R, Ghiso G, Nussenzweig V, Tomlinson S. 1999. The
799 characterisation of a novel trypanosome lytic factor in human serum. *Infection*
800 *and Immunity.* **67**:1910-1916.

801 Schell D, Evers R, Preis D, Ziegelbauer K, Kiefer H, Lottspeich F, Cornelissen AW,
802 Overath P. 1991. A transferrin-binding protein of *Trypanosoma brucei* is encoded
803 by one of the genes in the variant surface glycoprotein gene expression site.
804 *EMBO J.* **10**:1061-6.

805 Schwede A, Carrington M. 2010. Bloodstream form Trypanosome plasma
806 membrane proteins: antigenic variation and invariant antigens. *Parasitology*
807 **137**:2029-39.

808 Schwede A, Jones N, Engstler M, Carrington M. 2011. The VSG C-terminal domain
809 is inaccessible to antibodies on live trypanosomes. *Mol Biochem Parasitol*
810 **175**:201-4.

811 Shaw A. 2004. The economics of African Trypanosomiasis. In: Maudlin I, Holmes
812 P, Miles M, editors. *The Trypanosomiases*. CABI. p369-402.

813 Steverding D, Stierhof YD, Chaudhri M, Ligtenberg M, Schell D, Beck-Sickinger AG,
814 Overath P. 1994. ESAG 6 and 7 products of *Trypanosoma brucei* form a
815 transferrin binding protein complex. *Eur J Cell Biol.* **64**:78-87.

816 Symula RE, Beadell JS, Sistrom M, Agbebakun K, Balmer O, Gibson W, Aksoy S,
817 Caccone A. 2012. *Trypanosoma brucei gambiense* Group 1 is distinguished by a

818 unique amino acid substitution in the HpHb receptor implicated in human serum
819 resistance. *PLoS Negl Trop Dis*. **6**: e1728.

820 Uzureau P, Uzureau S, Lecordier L, Fontaine F, Tebabi P, Homblé F, Grélard A,
821 Zhendre V, Nolan DP, Lins L, Crowet JM, Pays A, Felu C, Poelvoorde P,
822 Vanhollebeke B, Moestrup SK, Lyngsø J, Pedersen JS, Mottram JC, Dufourc EJ,
823 Pérez-Morga D, Pays E. 2013. Mechanism of *Trypanosoma brucei gambiense*
824 resistance to human serum. *Nature* **501**:430-4.

825 Vanhamme L, Paturiaux-Hanocq F, Poelvoorde P, Nolan DP, Lins L, Van Den
826 Abbeele J, Pays A, Tebabi P, Van Xong H, Jacquet A, Moguilevsky N, Dieu M, Kane
827 JP, De Baetselier P, Brasseur R, Pays E. 2003. Apolipoprotein L-I is the
828 trypanosome lytic factor of human serum. *Nature* **422**:83-7.

829 Vanhollebeke B, De Muylder G, Nielsen MJ, Pays A, Tebabi P, Dieu M, Raes M,
830 Moestrup SK, Pays E. 2008. A haptoglobin-hemoglobin receptor conveys innate
831 immunity to *Trypanosoma brucei* in humans. *Science* **320**:677-81.

832 Widener J, Nielsen MJ, Shiflett A, Moestrup SK, Hajduk SL. 2007. Hemoglobin is a
833 co-factor of human trypanosome lytic factor. *PLoS Pathog*. **3**:1250-61.

834 Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM,
835 Krissinel EB, Leslie AGW, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS,
836 Potterton EA, Powell HR, Read RJ, Vagin A, Wilsonc KS. 2011. Overview of the
837 CCP4 suite and current developments. *Acta Crystallogr*. **D67**:235–242.

838

839

Figure Legends

Figure 1 – The structure of the **T. brucei** *haptoglobin-haemoglobin receptor*

A. The structure of the *T. brucei* haptoglobin-haemoglobin receptor, with helix I (red), helix II (orange) and helix V (blue). These three helices form an elongated bundle with a ~50° kink towards the membrane proximal C-terminal end. The inset shows a molecular envelope derived from small angle x-ray scattering. **B.** The structure of the *T. congolense* haptoglobin-haemoglobin receptor (Higgins et al., 2013) for comparison. **C.** A change in the pattern of hydrophobic residues results in a rigid kink in the three helical bundle of the TbHpHbR. Corresponding regions of the structures of TbHpHbR and TcHpHbR are shown with side chains of the hydrophobic residues that pack in the core of the bundle coloured red and residues at the kink sites in TbHpHbR coloured green. Also shown are sequence alignments of TbHpHbR and TcHpHbR for these regions of each helix, coloured in the same way.

Figure 1 – figure supplement 1 – Surface plasmon resonance analysis of the binding of HpSPHb to TbHpHbR

Surface plasmon resonance signals for two fold dilutions of HpSPHb complex from a maximum concentration of 16 μM, binding to a surface coated with the truncated version of *T. brucei* HpHbR.

Figure 1 – figure supplement 2 – Alignment of the TbHpHbR and TcHpHbR structures

Structural alignment of *T. brucei* HpHbR (blue) with *T. congolense* HpHbR (red). The membrane distal (upper) halves of the receptors align with a root mean square deviation of 1.1 Å while the membrane proximal (lower) halves differ due to the presence of a ~50° kink in TbHpHbR.

Figure 2 – The structural basis for haptoglobin-haemoglobin binding by TbHpHbR

A The structure of the complex between *T. brucei* HpHbR (blue) bound to its ligand, HpSPHb (haptoglobin is yellow, the β-subunit of haemoglobin is red and the α-subunit of haemoglobin is orange). **B.** The complex viewed from the membrane proximal end, showing the contacts made by haptoglobin and the β-subunit of haemoglobin. **C.** A view of the haemoglobin-binding site showing direct contacts between the haem and the receptor. Residues from the receptor that directly contact the haemoglobin subunit are shown as sticks and are numbered.

Figure 2 – figure supplement 1 – stereoview of the TbHpHbR in complex with HpHb

Figure 2 – figure supplement 2 – Small angle x-ray scattering of complexes of TcHpHbR and TbHpHbR with HpSPHb

A. The structure of the TbHpHbR:HpSPHb complex docked into an *ab initio* molecular envelopes calculated from scattering data. **B.** The theoretical scattering calculated from *ab initio* reconstructions (blue for HpSPHb, red for TbHpHbR and purple for TbHpHbR:HpSPHb), superimposed into experimental scattering data. Guinier plots are shown as an insert. **C.** Distance distribution functions of HpSPHb (blue), TbHpHbR (red) and TbHpHbR:HpSPHb (purple) derived from small angle x-ray scattering. **D.** A model of the TcHpHbR:HpSPHb complex docked into an *ab initio* molecular envelope calculated from scattering data. **E.** The theoretical scattering calculated from an *ab initio* reconstruction of the TcHpHbR:HpSPHb complex. **F.** Distance distribution function of TcHpHbR:HpSPHb derived from small angle x-ray scattering.

Figure 2 – figure supplement 3 – Clashes between TbHpHbR and a haemoglobin tetramer explain why the receptor does not bind to haemoglobin

A model for a complex of TbHpHbR bound to haemoglobin. This was derived by docking a haemoglobin tetramer onto the receptor with the β-subunit binding to the receptor as in the TbHpHbR:HpSPHb complex. TbHpHbR is shown in blue, the α-subunits of haemoglobin are orange and the β-subunits are red. A close up of the model is shown in the right hand panel with side chains involved in clashes shown as sticks.

Figure 2 – figure supplement 4 – The region affected by haptoglobin cleavage is not involved in interaction with TbHpHbR.

A. The structures of the HpSPHb region of porcine HpHb (red) aligned to the equivalent region of human HpSPHb from the structure of the TbHpHbR:HpSPHb complex (yellow). The structures align with a root mean square deviation of ~0.5 Å. The major difference is circled and lies around the site at which haptoglobin is cleaved during a processing event in the endoplasmic reticulum, which is disordered in the TbHpHbR:HpSPHb complex. **B.** A structural alignment of the porcine HpSPHb structure onto the TbHpHbR:HpSPHb structure. The region that is structurally altered by cleavage is circled and is not involved in contacts with the receptor. This is confirmed by surface plasmon resonance data (Figure 1 – figure supplement 1) which shows that TbHpHbR binds with similar affinity to HpSPHb as to previously measured native, cleaved HpHb.

Figure 3 – Differences between haptoglobin and haptoglobin-related protein do not alter affinity for TbHpHbR

A. The structure of the TbHpHbR:HpSPHb complex is shown with the receptor in blue and haptoglobin in yellow. Side chains in haptoglobin that are different in haptoglobin-related protein are highlighted in pink and are not involved in making interactions with the receptor. **B.** Surface plasmon resonance signals for two fold dilutions of HprSPHb complex from a maximum concentration of 8 μM, binding to a surface coated with *T. brucei* HpHbR. The measured affinity of 1.7 μM can be compared with the affinity of 0.7 μM for HpSPHb.

Figure 4 – Simultaneous binding of two receptors to each HpHb dimer leads to more efficient uptake into trypanosomes

A. A model for a complex of one HpHb dimer bound to two receptors, generated by docking the structure of the TbHpHbR:HpSPHb complex onto that of porcine HpHb (Andersen et al., 2012). The receptors are organized such that two receptors, both associated with the membrane through attachment at their C-termini, can simultaneously bind to one HpHb dimer. **B.** An *ab initio* molecular envelope derived from small angle x-ray scattering analysis of the TbHpHbR:HpHb complex supports the formation of a complex containing one HpHb dimer bound to two receptors. **C.** Uptake of fluorescently labelled dimeric HpHb into live cells was monitored via flow cytometry across a range of 1-62.5 nM. Uptake saturated by 4 nM in wild type cells whereas no uptake was observed in the HpHbR null cell line. No fluid phase uptake of labelled BSA was observed at these concentrations. **D.** Uptake of fluorescently labelled monomeric HpSPHb was not readily detected until 62.5 nM, at which point uptake had not saturated. HpSPHb uptake at 62.5 nM was lost in the HpHbR null cell line. Each uptake assay was carried out in triplicate. Error bars represent standard error of the mean, n=3.

Figure 4 – figure supplement 1 – Small angle x-ray scattering of HpHb, alone and in complex with TbHpHbR

A. An *ab initio* molecular envelopes calculated from scattering data from the HpHb complex. **B.** The theoretical scattering calculated from *ab initio* reconstructions for HpHb superimposed into experimental scattering data. Guinier plots are shown as an insert. **C.** Distance distribution functions of HpHb derived from small angle x-ray scattering. **D.** An *ab initio* molecular envelopes calculated from scattering data from the TbHpHbR:HpHb complex. **E.** The theoretical scattering calculated from *ab initio* reconstructions for TbHpHbR:HpHb superimposed into experimental scattering data. Guinier plots are shown as an insert. **F.** Distance distribution functions of TbHpHbR:HpHb derived from small angle x-ray scattering.

Figure 4 – figure supplement 2 – SEC MALLS data to assess the stoichiometry of the TbHpHbR:HpHb complex

Multi-angle light scattering (MALLS) measurements of TbHpHbR (red), HpHb (blue) and the TbHpHbR:HpHb complex (green). The molecular weights determined from scattering data (~30kDa for TbHpHbR, ~150kDa for HpHb and ~210kDa for the TbHpHbR:HpHb complex) show the formation of a complex containing two receptors bound to a single HpHb.

***Figure 4 – figure supplement 3 – Establishment and characterization of a HpHb$^{-/-}$ cell line of* T. brucei**

TbHpHbR null cell lines were generated in *T. b. brucei* Lister 427 bloodstream form (BSF) cells. **A.** The TbHbHbR gene was knocked out in Lister 427 BSF cells by replacement of one allele with a blasticidin resistance gene and the other allele with a neomycin resistance gene in four independent clones, as confirmed by Southern blot (P = Parental cell line, 1-4 = TbHpHbR null clones 1-4). The schematic depicts the original (top) and replacement (middle and lower) TbHpHbR loci and the positions of Southern blot probes and restriction enzyme sites used. Expected fragment sizes are annotated. **B.** Growth of the TbHpHbR null clones was monitored *in vitro* over 192 hours. Parental L427 cells grew with a mean doubling time of 8.4 hours whereas the TbHpHbR null clones had an increased mean doubling time of 11.5-12.0 hours. Error bars represent standard deviation, n= 8.

Figure 5 – Comparison of the structure of the TbHpHbR:HpHb complex with that of a variant surface glycoprotein

A comparison of the dimensions of the TbHpHbR:HpHb complex with that of the N-terminal domains of the variant surface glycoproteins (shown in grey). This suggests that HpHb will lie at least partially within the VSG layer when bound to two receptors.

Table 1 - Crystallographic data collection statistics

	HpSPHb	Tbb HpHbR	TbbHpHbR:HpSPHb
Beamline	Diamond I04-1	Diamond I03	Diamond I03
Space Group	$p3_121$	$p2_1$	c2
Cell dimensions (Å)	a=b=96.6, c=132.77	a=27.90, b=47.79, c=203.38, β=92.79	a=223.4, b=56.59, c=65.29, β=92.99
Resolution (Å)	2.05	1.85	3.1
Wavelength (Å)	0.916	0.9763	0.9750
R_{PIM} (%)	8.1 (37.4)	4.5 (42.9)	6.3 (72.6)
I/ σ(I)	8.7 (2.3)	10.2 (2.0)	9.8 (1.6)
Completeness (%)	99.8 (100)	97.4 (96.5)	96.9 (97.1)
Multiplicity	9.6 (10.2)	3.1 (3.1)	3.2 (3.3)

Table 2 - X-ray refinement statistics

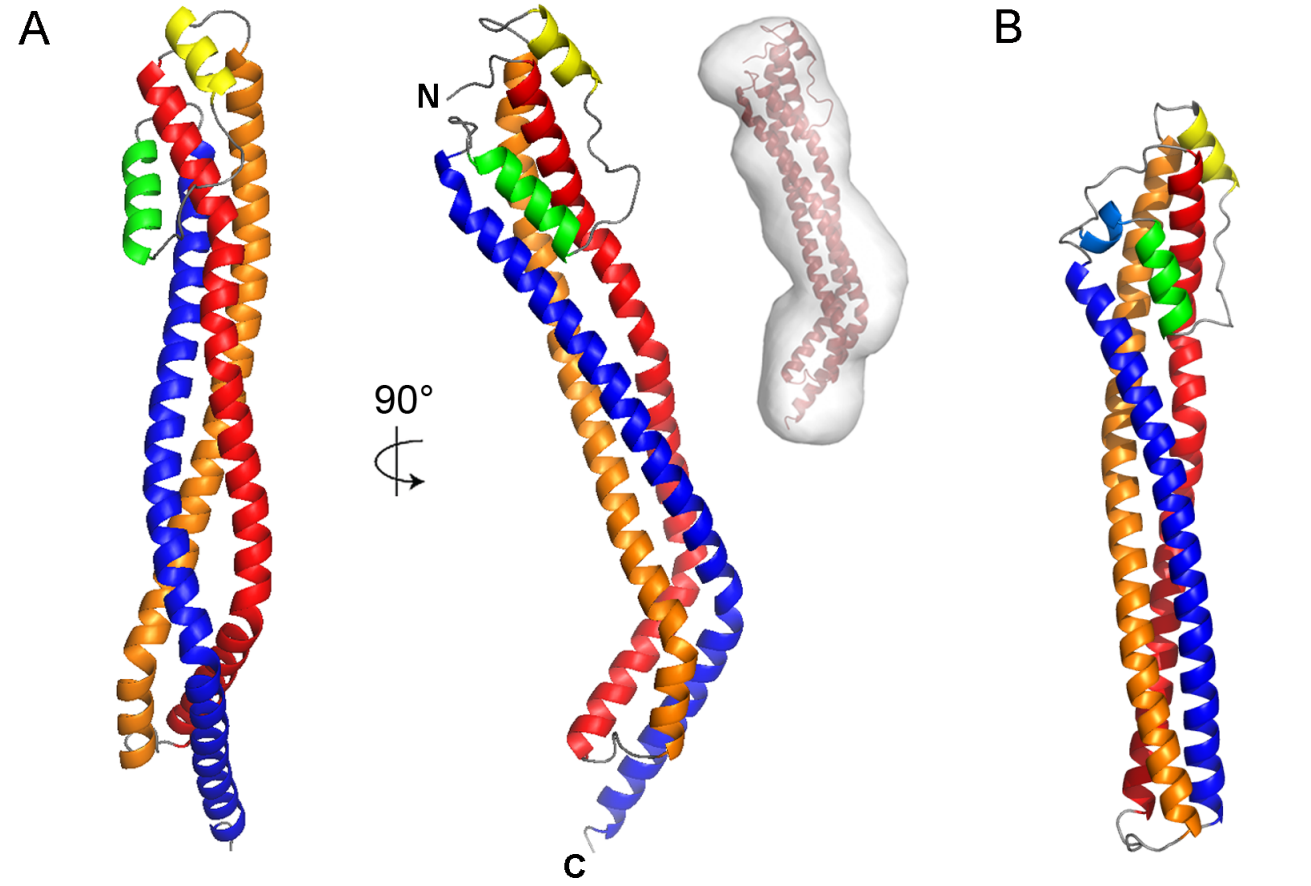
Complex	HpSPHb	Tbb HpHbR	TbbHpHbR:HpSPHb
Resolution (Å)	2.05	1.85	3.1
No. reflections	43170	44685	17302
R_{work} / R_{free} (%)	18.0 / 22.4	19.84 / 23.95	19.5 / 21.7
No. of protein residues in model	544	523	782
rmsd bond lengths (Å)	0.020	0.017	0.012
rmsd bond angles (°)	2.0	1.6	1.5
Ramachandran plot			
Allowed region	89.0%	98.8%	92.5%
Additional allowed region	11%	1.2%	7.5%
Generously allowed region	0%	0%	0%
Disallowed region	0%	0%	0%

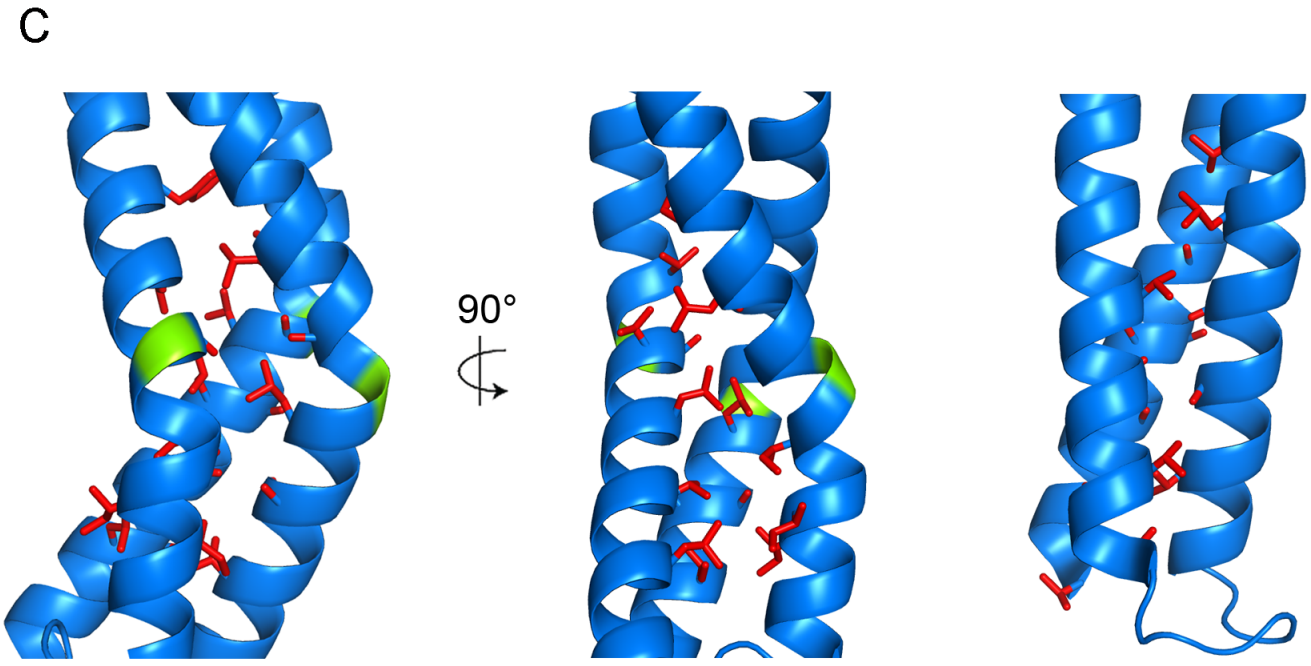
Table 3 - Small angle x-ray scattering statistics

	M_W (kDa)	R_G (nm)	D_{max} (nm)	Volume (nm^3)	Mw_{app} (kDa)
HpSPHb	59.7	2.6	7.5	75	36
TbHpHbR	32.2	3.5	11.5	44	22
TbHbHbR:HpSPHb	91.8	3.2	10.8	110	55
TbHpHbR:HpSPHb	89.6	3.8	12.0	140	70
HpHb	152	5.6	18.2	214	107
TbHpHbR:HpHb	217	6.3	16.5	370	185

Table 4: Interactions between TbHpHbR and HpSPHb

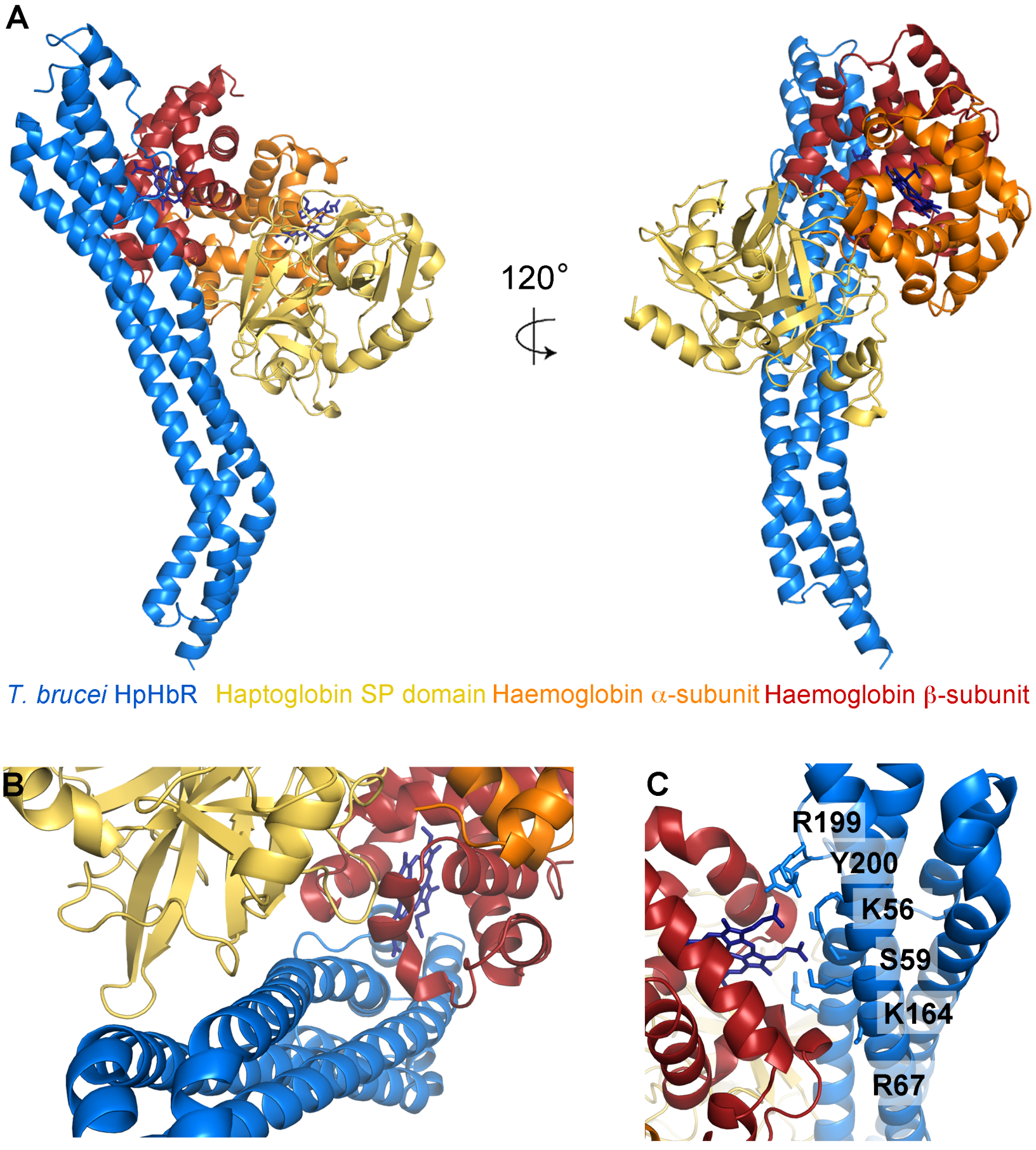
Receptor		HpSPHb			
Residue	**Group**	**Chain**	**Residue**	**Group**	**Interaction**
		Hbβ			
K56	side chain	B	Heme144	O1D	Hydrogen bond
E57	side chain	B	K96	Side chain	Salt bridge
S59	side chain	B	Heme144	O1D/O2D	Hydrogen bond
I60	side chain	B	Patch		Hydrophobic
R67	side chain NH1	B	R41	Backbone CO	Hydrogen bond
E70	side chain OE1/OE2	B	R41	Side chain NE/NH2	Salt bridge
S161	side chain	B	K60	Side chain	Hydrogen bond
S161	side chain	B	S45	Backbone CO	Hydrogen bond
K164	side chain	B	Heme144	O2D	Hydrogen bond
R199	side chain NE	B	Heme144	O2A	Hydrogen bond
Y200	side chain OH	B	Heme144	O2A	Hydrogen bond
S203	backbone CO	B	K96	Side chain	Hydrogen bond
		HpSP			
S73	side chain	C	K345	Side chain	Hydrogen bond
V74	hydrophobic	C	Patch		Hydrophobic
Q75	OE1	C	G276	Backbone CO	Hydrogen bond
A78	side chain	C	Patch		Hydrophobic
A82	side chain	C	Patch		Hydrophobic
K85	side chain	C	D305	Side chain O2D	Salt bridge

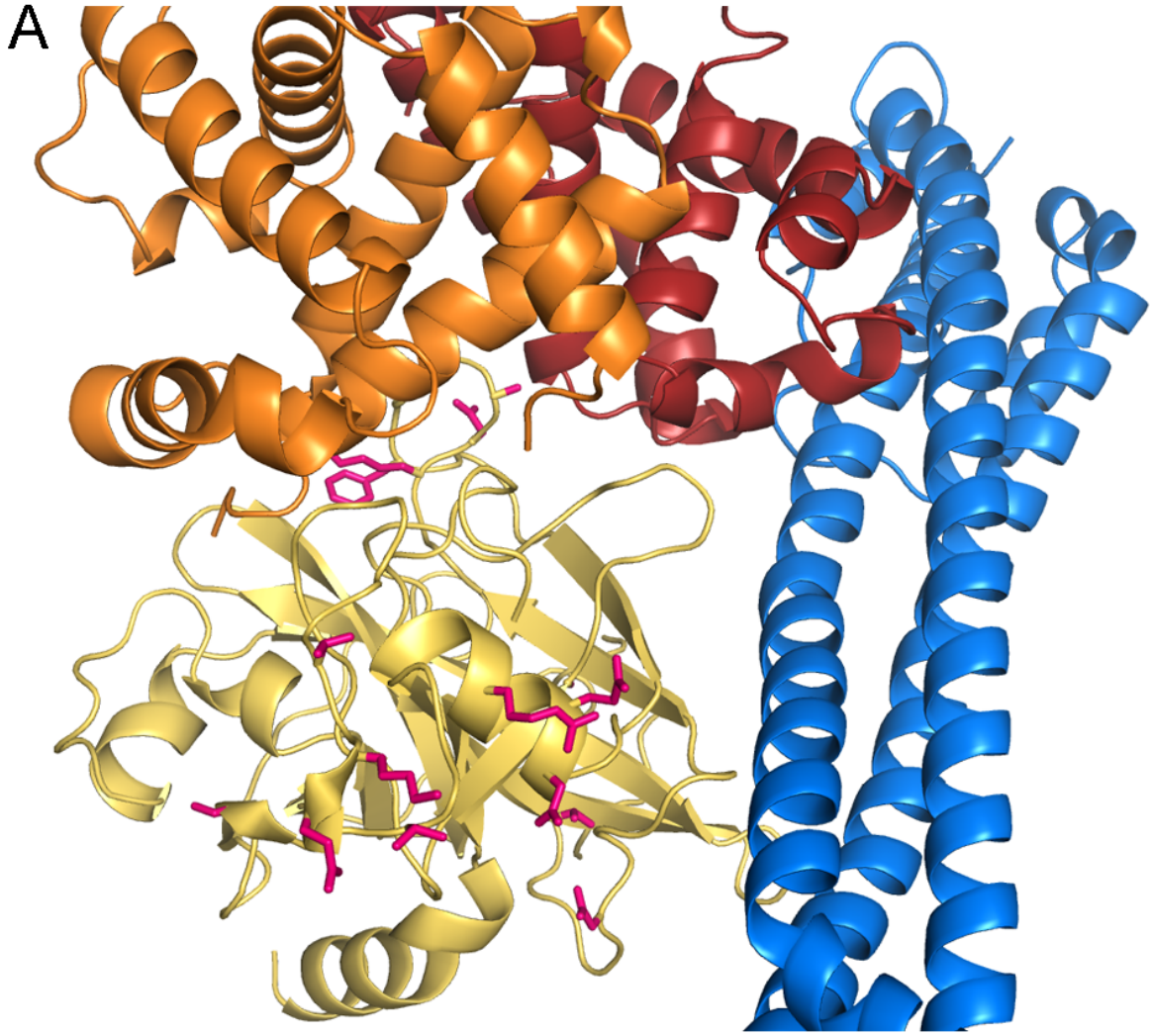


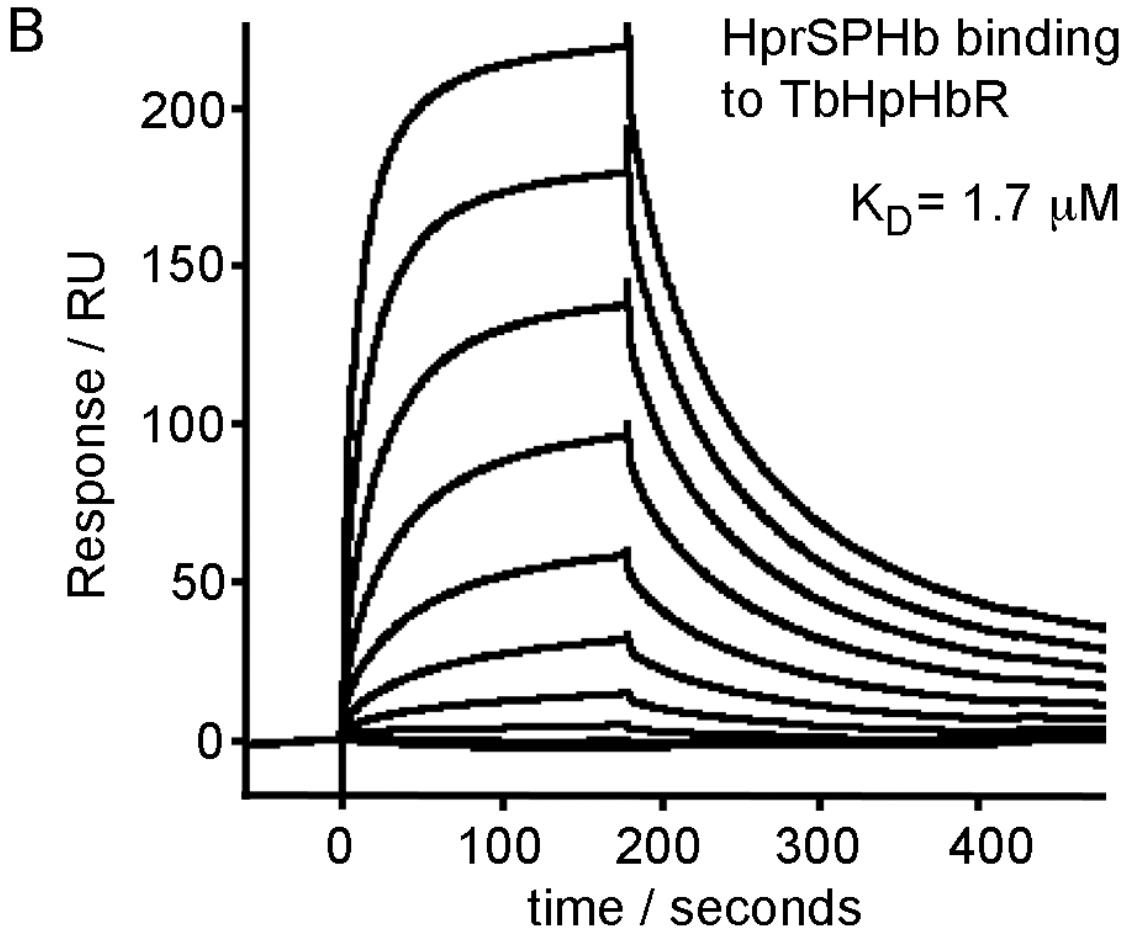


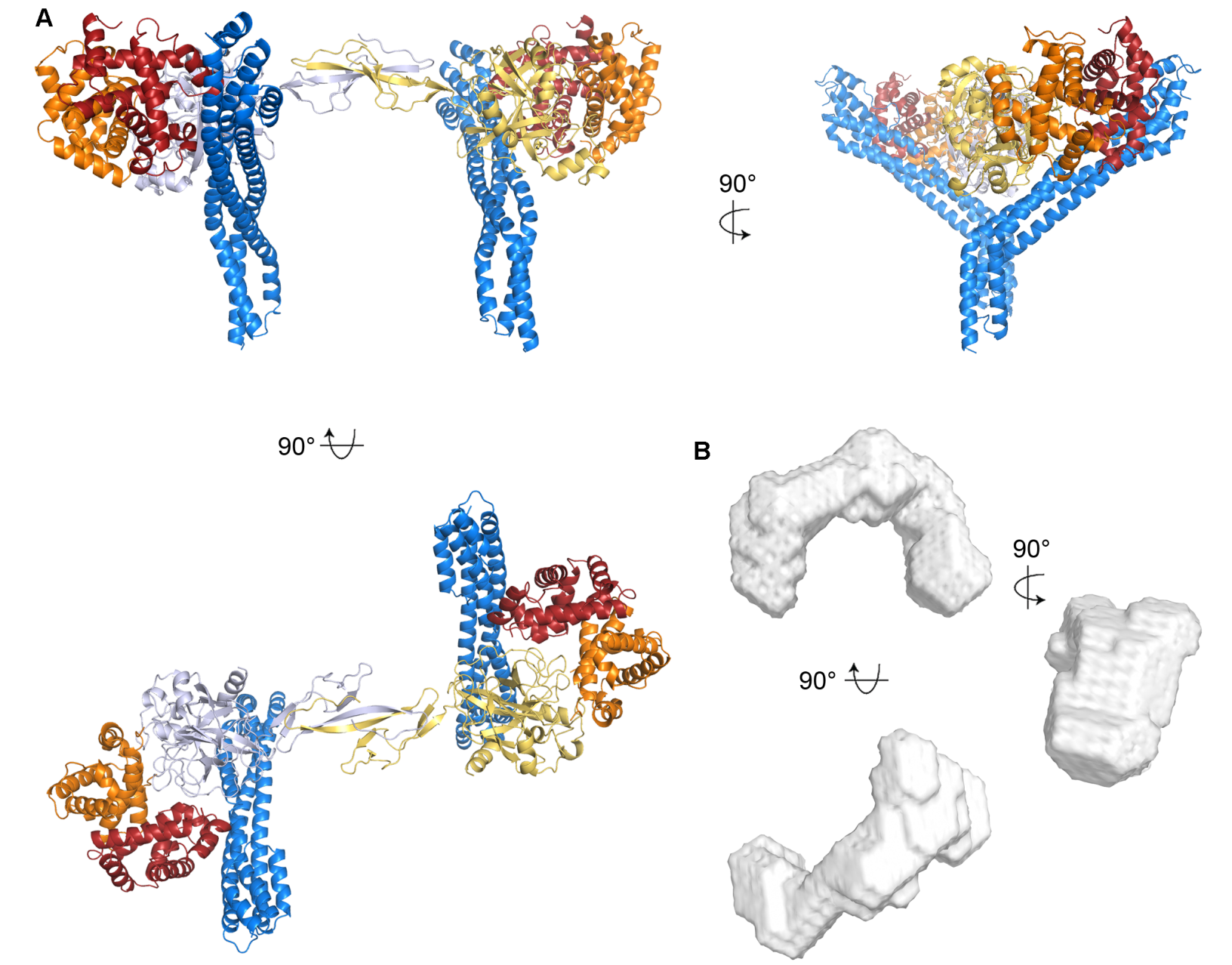
TbHpHbR **TcHpHbR**

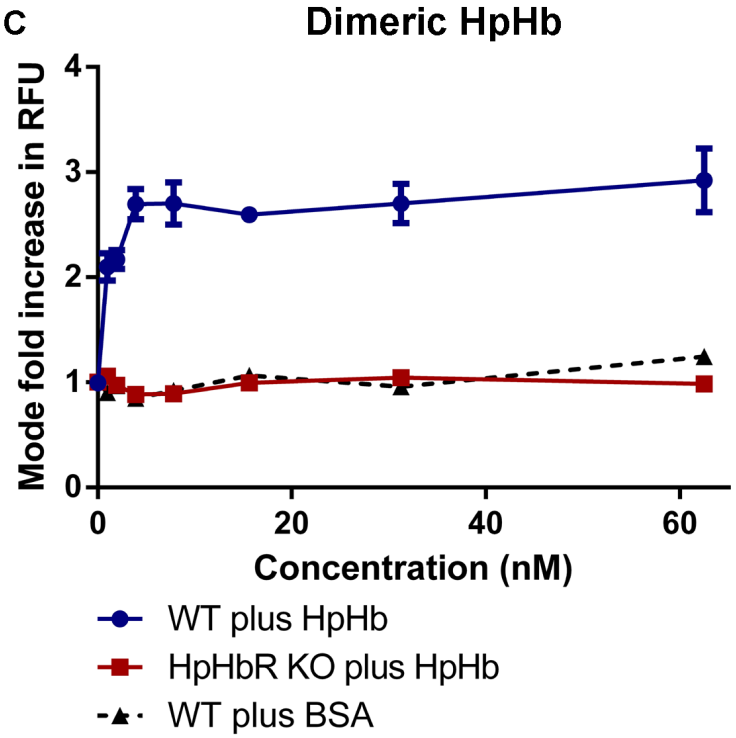
```
Helix I:   TcHpHbR   60-GGAAVERASAAVDRV
           TbHpHbR   90-VSKAVEALTRVDVAL

Helix II:  TcHpHbR   86-KVRKAAAVAHHALEHV
           TbHpHbR  120-IIENITSARENLALFN

Helix V:   TcHpHbR  233-AEKAAEEAETESSKV
           TbHpHbR  268-LKDAEVSLAAAAAEI
```

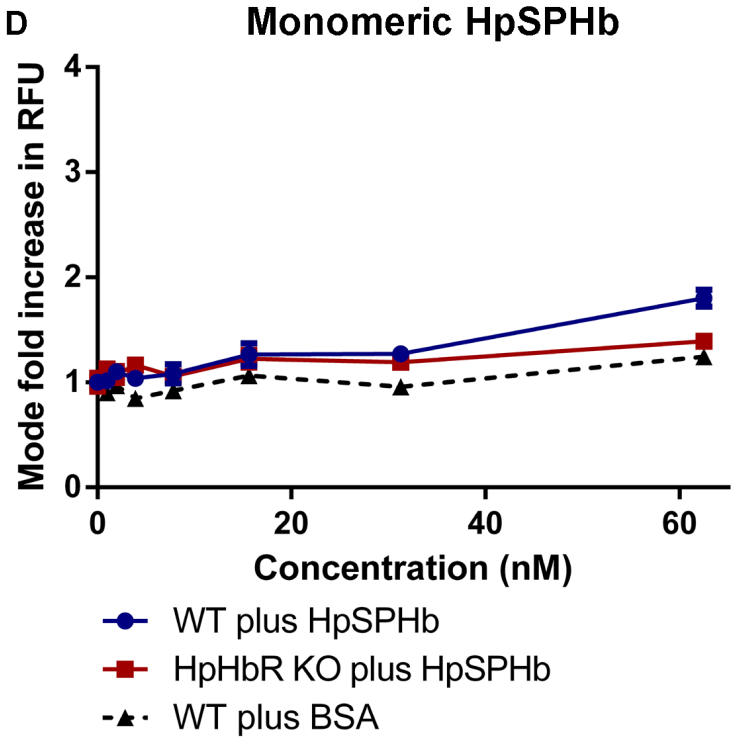

A
120°
T. brucei HpHbR
Haptoglobin SP domain
Haemoglobin α-subunit
Haemoglobin β-subunit
B
C
R199
Y200
K56
S59
K164
R67


A


B
HprSPHb binding
to TbHpHbR
K_D= 1.7 μM
Response / RU
200
150
100
50
0
0
100
200
300
400
time / seconds


A
90°
90°
B
90°
90°


C
Dimeric HpHb
Mode fold increase in RFU
0
1
2
3
4
0
20
40
60
Concentration (nM)
WT plus HpHb
HpHbR KO plus HpHb
WT plus BSA


D
Monomeric HpSPHb
Mode fold increase in RFU
0
1
2
3
4
0
20
40
60
Concentration (nM)
WT plus HpSPHb
HpHbR KO plus HpSPHb
WT plus BSA