Exploration of Drug Science

Open Exploration

Open Access Original Article

Flagellar pocket receptors as entry point for protein-based drugs against kinetoplastid parasites

Lucila Attala[1,2], Cecilia Vernetti[2], Elvio Rodríguez Araya[1,2*]

[1]Laboratorio de Biología y Bioquímica de *Trypanosoma cruzi*, Instituto de Biología Molecular y Celular de Rosario (IBR), Rosario 2000, Argentina

[2]Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Rosario 2000, Argentina

***Correspondence:** Elvio Rodríguez Araya, Laboratorio de Biología y Bioquímica de *Trypanosoma cruzi*, Instituto de Biología Molecular y Celular de Rosario (IBR), Rosario 2000, Argentina. rodriguezaraya@ibr-conicet.gov.ar

Academic Editor: Jinwei Zhang, Chinese Academy of Sciences, China

Received: June 7, 2026 **Accepted:** July 28, 2026 **Published:** September 7, 2026

Cite this article: Attala L, Vernetti C, Rodríguez Araya E. Flagellar pocket receptors as entry point for protein-based drugs against kinetoplastid parasites. Explor Drug Sci. 2026;4:1008179. https://doi.org/10.37349/eds.2026.1008179

Abstract

Aim: Kinetoplastids are flagellated protozoa encompassing multiple parasitic species responsible for severe neglected diseases. Although treatments exist, therapeutic failure and toxic side effects underscore the need for innovative drug development. Recent advances in protein design have accelerated the creation of small protein modules with specific functions, known as miniproteins, with broad pharmacological applications. However, their intrinsic inability to cross biological membranes limits their use against intracellular targets. This work aims to propose and computationally explore a modular delivery strategy that exploits the flagellar pocket (FP) as an entry route to deliver protein-based therapeutics into the parasites.

Methods: Using experimentally determined structures of three FP receptors, we applied a motif-scaffolding pipeline combining RFdiffusion, ProteinMPNN, and AlphaFold2-multimer to design de novo miniprotein modules capable of mimicking the natural cargo recognized by each receptor. Candidate designs were evaluated using a scoring function integrating minimum interaction predicted aligned error (miPAE) and backbone root mean square deviation (RMSD) across five predicted models per design.

Results: The design campaign yielded different outcomes depending on the target. For the transferrin receptor, 67 candidates surpassed the established in silico success thresholds, a pool expected to contain multiple experimentally validated binders. For the invariable surface glycoprotein 65, 17 candidates met the criteria, constituting a tractable experimental panel. Lastly, the haptoglobin-hemoglobin receptor proved a challenging target, with no candidates clearly surpassing both thresholds, likely due to the hydrophilic nature of its binding interfaces and the requirement for direct heme coordination.

Conclusions: This work provides a structural rationale for a novel receptor-mediated intracellular delivery paradigm in kinetoplastid parasites, offering a computational pipeline for generating miniprotein modules ready for experimental validation. We further outline how these delivery modules could be integrated into modular protein-based drugs incorporating protease recognition sequences, cell-penetrating peptides, and

subcellular localization signals, laying the conceptual ground for a new therapeutic approach against these neglected diseases.

Keywords

kinetoplastids, flagellar pocket, miniproteins, protein design, protein-based inhibitors, protein binders, structural bioinformatics

Introduction

Kinetoplastids are a group of flagellated protozoans with various evolutionary stages that cycle between insect vectors and definitive hosts. It includes several parasitic species that affect humans and animals, and the most relevant species from an epidemiological perspective are those known as TriTryps (*Trypanosoma brucei*, *Trypanosoma cruzi*, and *Leishmania* spp.). *T. brucei* is responsible for human sleeping sickness, a debilitating neurological disorder, and Nagana disease in cattle, characterized by weakness and emaciation [1]. *T. cruzi* is the causative agent of Chagas disease, a life-threatening chronic disease that affects the heart and gastrointestinal apparatus [2]. *Leishmania* species are the etiological agents of Leishmaniasis, resulting in cutaneous, mucocutaneous, and visceral infections [3]. In combination, these parasites are estimated to infect several million people worldwide, with current estimates indicating approximately 6–7 million individuals infected with *T. cruzi*, hundreds of thousands of active leishmaniasis cases, and a small but persistent number of human African trypanosomiasis (*T. brucei*) cases due to successful control efforts. Importantly, over 1 billion people live in areas at risk of acquiring kinetoplastid infections, particularly leishmaniasis, with transmission concentrated in low- and middle-income regions where access to diagnosis, treatment, and vector control is limited [4]. Despite the existence of a limited number of therapies, they often lack universal effectiveness and sometimes cause adverse effects that lead to therapy failure or discontinuation [5–7]. Moreover, individuals who do respond to treatment face the challenge of reinfection due to the continuous exposure risk and the immune evasiveness of these parasites, compounded by the absence of effective vaccines [8]. The emergence of drug-resistant strains further complicates the treatment landscape [9, 10]. The imperative need for new therapeutic alternatives is evident, making it necessary to explore novel drug targets, innovative delivery methods, and effective drug design approaches to address the challenges posed by these diseases.

Recent progress in artificial intelligence (AI) methods applied to biology and biotechnology has achieved significant milestones. AlphaFold2 (AF2) significantly increased protein structure prediction accuracy, providing researchers with a groundbreaking opportunity to obtain comprehensive representations of proteins without the lengthy processes of obtaining experimental structures through methods like NMR, X-ray diffraction crystallography or cryogenic electron microscopy (cryo-EM) [11]. Besides AF2, other accurate protein structure prediction models were developed, like RoseTTAFold (RF), RF2, or RoseTTAFoldNA (RFNA) [12, 13]. More recently, AF3 was developed to include all types of biomolecules [14]. In addition to structure prediction, AI methods can also assist in de novo protein design, which aims to create novel proteins that are not present in nature, with desired functions and properties. One such method is RFdiffusion, a guided diffusion model that can generate completely novel protein backbones able to perform various protein design tasks, such as monomer design, oligomer design, binder design, and enzyme active site scaffolding [15]. Another method is ProteinMPNN, a protein sequence design method that can map functional sequences onto given protein backbones with high accuracy, handling single or multiple chains, and coupling the amino acid sequence at different positions [16]. In combination, RFdiffusion and ProteinMPNN have shown outstanding performance in both in silico and experimental tests and have redefined the boundaries of what can be achieved in biotechnology, drug development, vaccine design, and biomaterial engineering [17]. This impact was formally recognized by the Nobel Prize in Chemistry, highlighting the paradigm shift brought by AI-driven protein structure prediction and design [18].

Based on these breakthroughs, a new method for drug design has emerged, where specific protein surfaces are used as targets to design de novo protein binders with the aim of altering the intrinsic function of the target. In particular, de novo-designed miniproteins (MPs) have gained prominence for this purpose, which are small (1–10 kDa) and hyperstable proteins that can bind targets with high affinity and specificity. For example, while the canonical approach of small molecule design as agonists/antagonists against human G protein-coupled receptors (GPCRs) is challenging and can take several years, many MPs have been designed and tested as MRGPRX1 agonists and glucose-dependent insulinotropic polypeptide receptor (GIPR), glucagon-like peptide-1 receptor (GLP1R), glucagon receptor (GCGR), CXCR4 and CGRPR antagonists, resulting in high selectivity, affinity and potency [19]. MPs were also developed against viruses and bacteria. For example, the MP inhibitor TRI2-2, designed to bind the spike protein of SARS-CoV-2, was shown to exhibit pan-neutralization against the variants that evolved during the COVID-19 pandemic, protecting against several strains when administered intranasally [20]. Or the MPs designed to bind the bacterial adhesins FimH and Abp1D/Abp2D from the uropathogenic *Escherichia coli* and *Acinetobacter baumannii*, respectively, which act as antagonists with high specificity and stability, disrupting their ability to bind host receptors [21]. Moreover, MPs can also be designed to disrupt protein-protein interactions (PPIs) involved in the formation of complexes from the same host, like those that interrupt the PPI that assembles the immunomodulator complex PD-1/PD-L1 by targeting PD-L1, with potential applications in cancer therapy [22]. These advances underscore how de novo designed MPs are rapidly transitioning from conceptual tools to viable therapeutic candidates, with some already demonstrating preclinical proof of principle as orally delivered antagonists of the Th17 response, designed to inhibit interleukins IL-23R and IL-17 [23].

The progress in the field and continuously growing capabilities of de novo protein design position MPs as promising candidates for targeting essential kinetoplastid components. However, all the above-mentioned MPs have one thing in common: their targets are all extracellular or cell-surface exposed. This direct accessibility is an advantage that partly explains their appeal as therapeutic targets, as the protein nature of MPs in general makes them impermeable to cell membranes. This poses a significant barrier to efficiently delivering MPs to the target, for example, essential cytosolic or nuclear components of these organisms. Moreover, the thick, impermeable surface coat of kinetoplastids, densely packed with continuously replenished variable surface glycoproteins [24], poses two major challenges: it hinders the intracellular delivery of MPs and limits the effectiveness of targeting surface membrane components because of their rapid turnover. These challenges underscore the need for delivery platforms that enable protein-based therapeutics to reach intracellular targets in kinetoplastids.

Nonetheless, TriTryps mediate nutrient uptake/acquisition and endocytosis exclusively through two specialized surfaces on the cell membrane: the flagellar pocket (FP) and the cytostome-cytopharynx (CC) [25]. The FP is a membrane invagination surrounding the proximal section of the flagellum and is present in all these organisms. It is connected to endomembrane systems that mediate molecular trafficking into and out of the cell. The cytostome, another specialized membrane invagination located near the FP, is found only in the replicative forms of *T. cruzi* (epimastigotes and amastigotes). Several proteins involved in molecular traffic have been identified on the surface of these organelles, and many are actively endocytosed either continuously or in response to external stimuli, like the binding of specific molecules [26]. Meanwhile, the flexibility of protein design allows the integration of modules with diverse functions into one single design, where each module contributes to the function of the whole synthetic protein [27].

In this article, we describe a possible way of using conserved FP receptor proteins as anchors for MP modules that can trigger the internalization of engineered protein-based drugs (Figure 1). By analyzing the experimental and predicted structures of FP proteins known to be actively endocytosed and exploring diverse protein design strategies that take advantage of the natural interactions mediated by these macromolecules, we outline a path toward programmable modular designs capable of hijacking parasite endocytic routes. As an example, we used three FP receptors from *T. brucei*, and we discuss how to adapt the resulting modules to obtain functional designs that allow for proper release and delivery of the intended cargo, laying the ground for a new therapeutic paradigm in the fight against kinetoplastid

diseases. We focused on *T. brucei* because its FP-associated endocytic pathways are comparatively well characterized; extending this strategy to *T. cruzi* and *Leishmania* will require overcoming not only additional physical barriers but, more fundamentally, the current scarcity of knowledge regarding their internalization routes, a gap that we hope this work encourages the field to address.

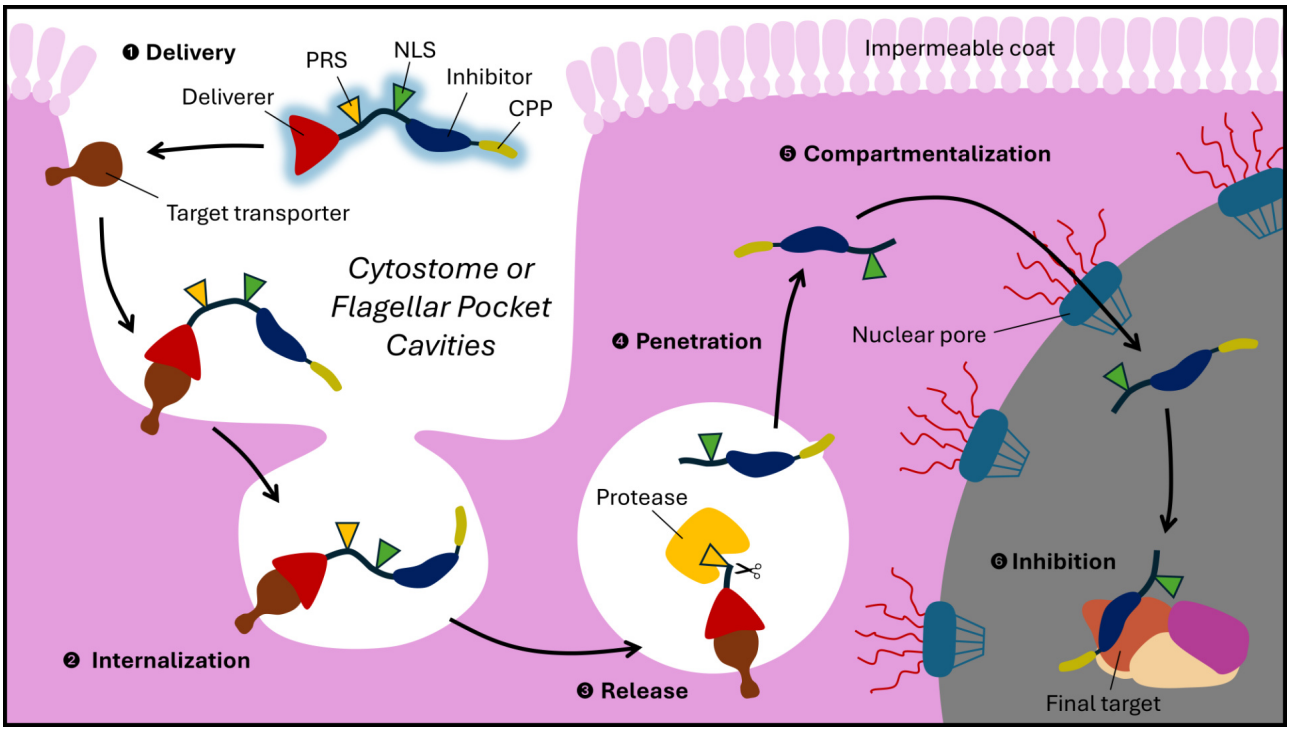

Figure 1. Proposed modular delivery system for protein-based drugs against kinetoplastids. The modular protein-based design is highlighted with a blue glow, composed of a deliverer module (red), a protease recognition sequence (PRS), a nuclear localization signal (NLS), an inhibitor module, and a cell-penetrating peptide (CPP). Successive steps in the delivery process are indicated with numbers. (**1**) Delivery: the deliverer module binds the outer segment of a target transporter located in the CC or the FP. (**2**) Internalization: the binding of the deliverer module triggers the internalization of the transporter bound to the protein-based drug. Steps 1 and 2 are the focus of this work. (**3**) Release: once the internalized vesicle matures, an endo-lysosomal protease recognizes the PRS and cuts the design at a specific point, releasing the inhibitor-containing half. (**4**) Penetration: the CPP destabilizes the membrane and causes the translocation of the inhibitor into the cytosolic compartment [28]. (**5**) Compartmentalization: in this example, an NLS further compartmentalizes the inhibitor by directing it to the nucleus [29]. (**6**) Inhibition: once in the target compartment, in this case the nucleus, the inhibitor binds to the final target, affecting the development of the kinetoplastid parasite. CC: cytostome-cytopharynx; FP: flagellar pocket.

Materials and methods

Experimental and predicted structures of FP receptors

For the transferrin receptor of *T. brucei*, we explored the structures of PDBs 6SOY and 6SOZ solved by X-ray diffraction, both bound to human transferrin [30]. For the protein design process, we used PDB 6SOY, as it is better resolved (2.75 Å, compared to 3.42 Å). For the invariant surface glycoprotein 65 (ISG65) of *T. brucei*, we explored the PDB 7PI6 solved by X-ray diffraction [31] and the AF database model for Q587F5 (Uniprot ID) [32]. As design input, we used PDB 7PI6. For the haptoglobin-hemoglobin receptor (HpHbR) of *T. brucei*, we explored the structures of PDBs 5HU6 and 4WJG solved by X-ray diffraction [33, 34]. As input for the design, we used 5HU6, as it is better resolved (2.90 Å versus 3.10 Å).

Residue contacts determination between cargo and receptor

To identify atoms and residues involved in protein interactions between the receptors and their cargo proteins, we used ChimeraX v1.11.1. For this, the chain of the desired cargo protein was selected, and contacts with the receptor chain were computed using a Van der Waals (VDW) overlap ≥ –0.40. For H-bond detection, we used a distance tolerance of 0.4 Å and an angle tolerance of 20°.

MP design

Binders were designed using RFdiffusion to generate novel backbones using the motif scaffolding method [15], ProteinMPNN to design their sequences [16], and AF2-multimer to model the resulting sequences in complex with their corresponding targets [35].

For *Trypanosoma brucei* transferrin receptor (TbTfR), the RFdiffusion design was run using the following configurations:

contigmap.contigs=[A20-342/0 5-30/C349-370/5-30 B18-337/0]

contigmap.inpaint_seq=[C350,C351,C353,C354,C357,C358,C361,C362,C363,C365,C366,C368,C369,C370]

ppi.hotspot_res=[A221,A222,A228,A248,A266,B140,B141,B142,B150,B151]

For ISG65:

contigmap.contigs=[A27-86/A94-154/A196-229/A251-316/0 5-30/B1164-1174/15-25/B1109-1114/5-30]

contigmap.inpaint_seq=[B1111,B1112,B1165,B1166,B1167,B1168,B1169,B1172,B1173]

ppi.hotspot_res=[A70,A73,A74,A77,A282,A283,A286,A289,A290,A293]

For HpHbR1:

contigmap.contigs=[D38-297/0 10-20/B92-92/3-3/B96-96/13-30/B41-45/10-20]

contigmap.inpaint_seq=[B43]

ppi.hotspot_res=[D157,D160,D161,D164,D200,D201,D203,D56,D57,D59,D60,D63,D66,D67,D70,D64,D61,D204,D153,D68,D71]

For HpHbR2:

contigmap.contigs=[D38-297/0 10-35/C300-305/5-30/C275-280/10-35]

contigmap.inpaint_seq=[C276,C277,C279,C301,C302,C303,C304]

ppi.hotspot_res=[D70,D71,D72,D73,D74,D75,D78,D79,D81,D82,D85,D86,D251,D252,D255,D256,D258,D259,D262]

In all cases, we generated 1000 backbones of length 50–80 and we set guiding_potentials = ["type:binder_ROG,weight:5,min_dist:5", "type:binder_ncontacts,weight:10", "type:interface_ncontacts,weight:5"], guide_scale = 0.5 and guide_decay = "linear" as potentials, except for targets HpHbR1 and HpHbR2, for which we tried different potentials without success. The designs shown in the figures change "type:binder_ncontacts,weight:20", "type:interface_ncontacts,weight:5" for the former and "type:binder_ncontacts,weight:5", "type:interface_ncontacts,weight:10" for the latter. We also tested different sizes for both, leaving 70–90 for HpHbR2. These potentials are soft constraints that we optimized iteratively, based on trial and error, to maximize the number of successful designs. They control the final radius of gyration of the binder (binder_ROG: more weight implying more compact designs), the number of intra-molecular contacts of the binder (binder_ncontacts: more weight implying more intramolecular contacts), and the number of intermolecular contacts between the binder and the target (interface_ncontacts: more weight implying more inter-molecular contacts). See the "Availability of data and materials" section for the source code used in this work.

In all cases, we generated a single sequence per backbone with ProteinMPNN by fixing the target chain and allowing the design of all the positions that weren't a Gly amino acid, omitting Cys residues and using a temperature of 0.0001. Each designed sequence was modelled in complex with the corresponding target with AF2-multimer v3 using LocalColabFold v1.6.1 and the following parameters: --num-models 5, --num-recycle 3, --rank ipTM, --recycle-early-stop-tolerance 0.5, --msa-mode single_sequence, --templates and --

custom-template-path ./pdbs [36]. As input, the paired + unpaired single sequences [37] of both target and designed sequences were given, along with the input PDB as template, stored in the ./pdbs directory.

Metrics extraction, root mean square deviation (RMSD) computation, re-ranking and visualization

For each of the five AF2 predicted structures, the following metrics were extracted from the per-model PDB and JSON output files. Per-residue predicted local distance difference test (pLDDT) was averaged over all residues assigned to the binder chain to yield a mean binder pLDDT value, with higher values indicating greater confidence in the local structure of the designed chain. Predicted aligned error (PAE) matrices were parsed from the AF2 JSON output. The minimum interaction PAE (miPAE) was defined as the minimum value within the binder-to-target submatrix of the full PAE matrix, representing the single most confident predicted inter-chain distance. Interface predicted template modeling score (ipTM) and whole-complex pTM were extracted directly from the JSON score files.

RMSD to the RFdiffusion backbone was computed by superimposing the Cα atoms of the AF2-predicted binder onto the corresponding chain of the RFdiffusion backbone structure, using a least-squares superposition via BioPython's Superimposer class. Lower RMSD values indicate that the predicted structure recapitulates the designed backbone geometry.

Within each design, the five AF2 models were re-ranked by their miPAE (ascending), independently of the AF2 internal ranking. This produced a PAE rank (1 to 5 per design), where PAE rank 1 corresponds to the model with the lowest predicted error in the binder-to-target interaction.

All plots were generated using Plotly and exported as interactive HTML files. 3D scatter plot displays each resulting AF2 model as a point in a three-dimensional space with the following axes: min interaction PAE (x), RMSD to RFdiffusion backbone (y), and PAE rank (z). Models belonging to the same design are connected by a line, visualizing how a single design's predicted properties vary across its five models and PAE ranks. Points were colored by mean binder pLDDT. Hovering over any point highlights all five models from that design, facilitating rapid identification of designs with consistently favorable or inconsistent metrics.

Scoring function

Each design was assigned a scalar score summarizing its structural quality across all five models. For each PAE rank $r \in \{1, 2, 3, 4, 5\}$, the Euclidean distance from the ideal point (0, 0) was computed in the two-dimensional space of miPAE and RMSD to RFdiffusion backbone:

$$d_r = \sqrt{miPAE_r^2 + RMSD_r^2} \quad (1)$$

A decreasing weight (w) was applied to each rank ($w_1 = 0.10$, $w_2 = 0.08$, $w_3 = 0.06$, $w_4 = 0.04$, $w_5 = 0.02$). The rationale behind this weighting scheme was to incorporate structural reproducibility across the five AF2 models while assigning greater importance to higher-confidence predictions. Equal weighting would place the same emphasis on all predicted models, whereas progressively decreasing weights prioritize the highest-ranked prediction while still rewarding designs that remain structurally consistent across all five models. A simple linear decay was adopted as a straightforward compromise between these two objectives. The absolute magnitude of the weights was chosen so that successful designs would typically yield scores close to or greater than one, providing an intuitive threshold for candidate prioritization.

The weighted sum of distances was inverted to yield the score S:

$$S = \frac{1}{\sum_{r=1}^{5} w_r \cdot d_r} \quad (2)$$

Higher raw scores therefore indicate designs whose predicted structures consistently exhibit both low miPAE and close agreement with the RFdiffusion backbone across model ranks.

Results

Transferrin mimetic modules

In the human body, iron is distributed through the bloodstream by transferrin (Tf), a globular protein synthesized by the liver that has two iron-binding sites. Human cells have specialized transmembrane glycoproteins, primarily dimeric in structure, that facilitate cellular iron uptake via receptor-mediated endocytosis of iron-loaded transferrin, known as human transferrin receptors (TfRs: TfR1 and TfR2). For example, when TfR1 binds diferric transferrin at the cell surface, it is internalized into endosomes, releasing the iron due to the low pH, and recycling back to the surface to release apo-transferrin. As a counterpart, kinetoplastid parasites have evolved similar systems to uptake iron from their hosts [38], hence, the involved receptors can be used as targets to design MP modules.

In *T. brucei*, its transferrin receptor (TbTfR) is a heterodimer located at the FP, composed of the proteins ESAG6, anchored to glycosylphosphatidylinositol (GPI), and ESAG7 (Figure 2A) [30]. It does not have sequence similarity to either TfR1 or TfR2 but binds transferrin and is internalized via clathrin-coated pits. The cargo then traffics through Rab5-positive endosomes, and the receptor recycles via Rab11 to the pocket [39].

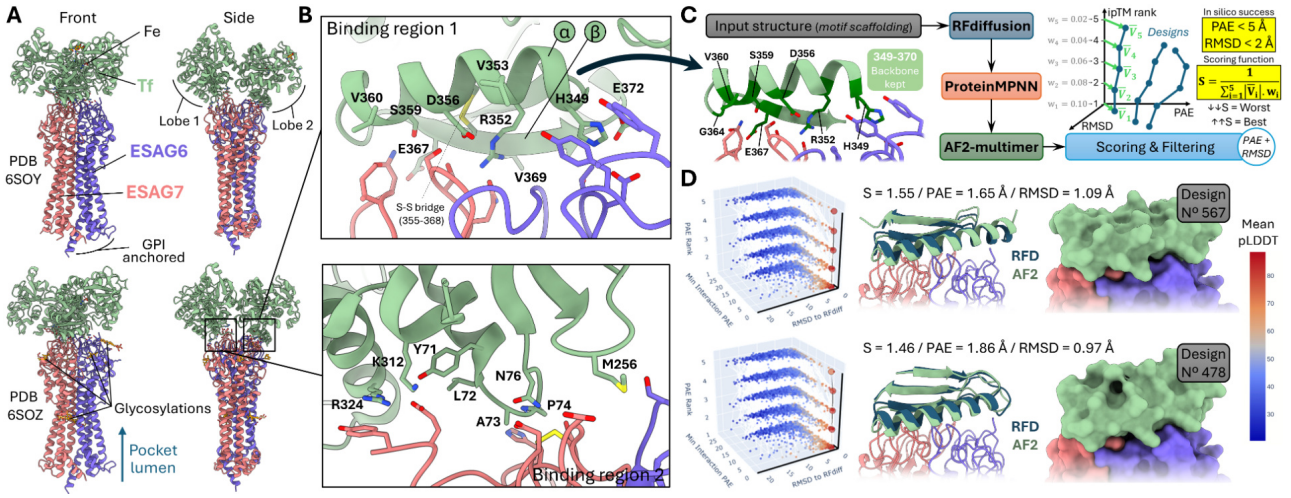

Figure 2. Transferrin receptor of T. brucei and transferrin mimetic designs. (**A**) Front and side (90° rotation) views of ESAG6/ESAG7 crystal structures bound to human Tf (Upper panel: PDB 6SOY; Lower panel: PDB 6SOZ). The GPI anchor binds covalently to ESAG6, attaching the complex to the FP membrane and orienting the Tf binding site to the FP lumen. (**B**) Close up of the binding interfaces from PDB 6SOZ. Sidechains involved in contacts and H-bonds are shown. The corresponding residue numbers and code for Tf are indicated. (**C**) Input structure for the motif scaffolding design process. Dark green residues were kept completely, conserving only the backbone of light green residues. Backbone generation was done with RFdiffusion, sequence generation was done with ProteinMPNN, and AF2-multimer was used to model the resulting sequences in complex with the target receptor. A scoring step was done to select the best designs. The score was derived from the PAE of AF2 and the RMSD between the AF2 model and the RFdiffusion backbone. (**D**) Top two scoring designs. The 3D plot (left panel) shows the distribution of PAEs and RMSDs, highlighting the values of the top designs. Each marker represents an output model from AF2, colored by pLDDT. The superposition of the RFdiffusion backbone is shown in blue and the one from the rank 1 model of AF2 in green (center panel). Solvent exposed surface of the AF2 model (right panel). AF2: AlphaFold2; FP: flagellar pocket; GPI: glycosylphosphatidylinositol; PAE: predicted aligned error; pLDDT: predicted local distance difference test; RFD: RFdiffusion; RMSD: root mean square deviation; Tf: transferrin.

In *Leishmania* species, early biochemical studies detected an integral membrane glycoprotein of ~70 kDa on promastigotes as a single polypeptide that can be purified on transferrin-Sepharose and is recognized by anti-human TfR antibodies. However, its role in receptor-mediated endocytosis and iron acquisition through the incorporation of Tf remains unproven. Instead, *Leishmania* primarily acquires iron through reduction of ferric to ferrous iron (e.g., ferric reductases such as LFR1) followed by transport via dedicated ferrous iron transporters (e.g., LIT1). Additionally, intracellular amastigotes can access host Tf iron indirectly by recruiting TfR-positive endosomal membranes to the parasitophorous vacuole and import already released iron via iron transporters located at the FP [40].

Similarly, *T. cruzi* amastigotes show specific Tf binding and internalization, and an approximately 200 kDa surface protein that cross-reacts with anti-human TfR antibodies was reported in early studies [41]. Evidence indicates that Tf uptake occurs into the FP/CC system and is stored in reservosomes. For example, transferrin-gold nanoparticles were shown to localize in reservosomes and to lysosomal-related organelles after internalization [42]. Several functional studies also indicate that Tf endocytosis in *T. cruzi* is sensitive to disruption of the CC-associated cytoskeleton and membrane cholesterol but is largely independent of classical clathrin-mediated endocytosis, implying an alternative, clathrin-independent receptor-mediated route [43].

Unfortunately, to our knowledge, the coding sequence of putative TfRs from *T. cruzi* and *Leishmania* were never cloned or assigned to gene identifiers of any available genome. These proteins remain poorly characterized, with only limited biochemical evidence available, which prevents access to their protein sequences for structural prediction and analysis. Further evidence is needed; hence, we focused only on TbTfR.

Two experimental structures are available in the Protein Data Bank for TbTfR in complex with human Tf (Figure 2A). One of these structures contains N-linked glycans in exposed positions of the receptor, which are thought to be important for immune evasion [30]. Importantly, none of these glycosylations mediate the recognition of Tf. The binding mode involves two distinct interfaces, one for each Tf lobe. The first binding region (Figure 2B, upper panel) comprises residues from a tightly packed α-helix and β-strand of Tf that interact with residues located in loops from both ESAG6 and ESAG7. Molecular interactions include a network of H-bonds mediated by Tf sidechain atoms from residues S359, D356, R352 and H349 and backbone atoms from residues C368 and S370 (Figure S1). The only hydrophobic interaction is established by residue V360. The Tf residue H349, besides being involved in H-bond formation, engages a strong aromatic interaction (π-π stacking) with a tyrosine of ESAG6 (Y248), indicating that it is particularly important for the interaction (Figure S1). Meanwhile, the second binding region involves residues from four different segments of human Tf (Figure 2B, lower panel). H-bond forming residues are Y71 and R324, which are positioned in two different α-helixes from Tf (Figure S2). Other residues involved in electrostatic and hydrophobic interactions are located in two different loops and another α-helix. These include L72, A73, P74 and N76 in the first loop, M256 in the second loop, and K312 in the α-helix (Figure S2).

For all the protein designs in this article, the design strategy we followed was the scaffolding of binding motifs extracted from the receptor-cargo interfaces onto de novo protein backbones [15]. In this case, from the two binding regions described above, we selected the first as the motif input for the design process. This choice was motivated by the structural compactness of the involved Tf segment, given that residues 349–370 form a contiguous stretch of sequence that folds into a well-defined α-helix and β-strand, making it amenable to motif scaffolding as a single, geometrically coherent unit. In contrast, the second binding region involves residues distributed across four separate backbone segments, making it considerably more challenging to extract and scaffold as an isolated motif without losing the spatial relationships between the interacting residues. For the selected motif, the complete backbone of residues 349–370 was preserved while keeping only the sidechains of residues H349, R352, D356, S359, V360 and E367, which engage directly in contacts with TbTfR (Figure 2C, left panel). RFdiffusion was then used to generate 1000 novel protein backbones that scaffold this motif, producing compact MP structures that position the key interacting residues in the same spatial arrangement as in the native Tf-TbTfR complex. Subsequently, ProteinMPNN was used to design the most probable amino acid sequence compatible with each generated backbone, without redesigning motif residues. The resulting sequences were modeled in complex with TbTfR using AF2-multimer, yielding five predicted models per design. Designs were then scored using a scoring function (Equations 1 and 2) that rewards designs for which the greatest number of models have both a confident interaction with the receptor (low miPAE) and a faithful recapitulation of the RFdiffusion backbone (low AF2 versus RFdiffusion backbone RMSD) prioritizing designs where the sequence assigned by ProteinMPNN repeatedly and consistently folds back into the intended structure and engages TbTfR with low predicted error (Figure 2C, right panel; see Materials and methods for details).

The distribution of the resulting design metrics (miPAE, RMSD against RFdiffusion backbone, and pLDDT) is shown in the scatter plots from Figure 2D. The values for the five AF2 models for each of the two top-scoring designs are highlighted in each plot (markers connected with lines), which gave values of 1.55 and 1.46, with miPAE values of 1.65 Å and 1.86 Å and RMSD values of 1.09 Å and 0.97 Å for the rank 1 AF2 model, respectively (Figure 2D). In both cases, AF2 models closely recapitulate the RFdiffusion backbone geometry, and the designs adopt compact folds with excellent binder-receptor surface complementarity. In terms of tertiary structure, the best designs have backbones that continue the natural β-strand of the motif to form a β-sheet and extend the α-helix with a few extra residues. Notably, the design process proved highly successful, with a considerable number of the generated designs (67 candidates) surpassing the miPAE < 5 Å and RMSD < 2 Å thresholds across the ensemble, a level of success rate that, based on experimental validation studies of RFdiffusion-based binder design campaigns, is strongly predictive of retrieving true binding activity in vitro [15].

ISGs

Many eukaryotic pathogens have evolved to establish and maintain proliferative populations within hosts, and some of them can even maintain a persistent infection that can last for years. These long-term infections only occur if the pathogen population can survive both the innate and adaptive immune responses [31].

Among kinetoplast parasites, African trypanosomes provide a particularly instructive example of how surface architecture supports survival within the host. As extracellular pathogens, they are continuously exposed to the adaptive and innate immune systems. To achieve persistence, they have evolved a remarkable cell surface dominated by many copies of a single variant surface glycoprotein (VSG), which is periodically switched [31]. This VSG coat both protects the plasma membrane from immunoglobulins and underpins a system of antigenic variation, allowing the population to evade the adaptive immune system [44].

However, within the VSG coat are found invariant proteins, including nutrient receptors, and a family of ISGs. Although these invariant proteins are 100- to 200-fold less abundant than the VSG, they are essential for parasite survival [45]. While VSG handles antibody evasion by switching the expressed *VSG* gene or through *VSG* recombination, the parasite relies on less abundant, constant proteins hidden within the coat to mitigate innate attacks. Recent work has shown that ISG65 functions as a receptor for the complement protein C3/C3b, an innate immune component, during the bloodstream stage (BS) of infection. Rather than merely blocking complement deposition, ISG65 operates as an active surface-clearing mechanism: it captures deposited C3b and mediates its internalization and trafficking to the lysosome. Together with the observation that ISG65 reduces C3b surface binding and uptake, this suggests that ISG65 participates in complement handling through receptor-mediated endocytosis [31, 44]. This interpretation is consistent with earlier work showing that ISG65 is internalized and sorted through clathrin-dependent, ubiquitylation-dependent endocytic pathways, a general mechanism that trypanosomes use to remove surface proteins from the plasma membrane [46]. Crucially, while the C3b cargo is destroyed in the lysosome, ISG65 itself exhibits a high recycling rate comparable to VSG, allowing it to return rapidly to the surface and sustain its protective role.

The identification of ISG65 as a critical receptor for human complement C3b has redefined the understanding of innate immune evasion in *T. brucei* [31, 44]. Because this protein is essential for parasite persistence and remains invariant across different strains, it represents a highly attractive target for the development of novel diagnostics or therapeutics [45].

To our knowledge, no putative ISG65 homologs from *T. cruzi* and *Leishmania* have been identified. These proteins appear to be a specialized adaptation of *T. brucei*, specifically restricted to the BS, which prevents structural comparisons across other major trypanosomatids [45]. However, other kinetoplastids deploy functionally analogous invariant surface proteins for immune evasion, such as gp58/68 in *T. cruzi*, which inhibits alternative pathway activation [47], and GP63 in *Leishmania*, which cleaves complement C3 to prevent lysis [48]. These examples demonstrate that utilizing constant surface components to

manipulate host innate immunity is a widely conserved survival strategy among these diverse human pathogens, and may serve as analogous targets like the ISG65.

Structural analysis has demonstrated that ISG65 specifically binds to the thioester domain (TED) of C3b [31]. Furthermore, experimental evidence indicates that ISG65 is involved in reducing trypanosome susceptibility to C3-mediated clearance, effectively acting as a specialized mechanism for innate immune evasion that complements the adaptive shield of the VSG coat [44]. The structural basis for our target is derived from the crystal structure of ISG65 in complex with the human C3d domain (PDB 7PI6), which reveals the characteristic three-helix bundle architecture of the receptor's ectodomain (Figure 3A, upper panel). While this experimental model provides high-resolution data on the binding interface, it lacks the C-terminal regions that were truncated for crystallization [31]. A complete picture of the receptor architecture can be seen in the AF2 model of the Uniprot entry Q587F5 (Figure 3A, lower panel), which includes a small, disordered N-terminal tail and a flexible linker that connects to a C-terminal transmembrane domain that anchors the receptor to the parasite's membrane [44]. The binding mode involves a broad concave interface located within the three-helix bundle, a geometry shaped by a ~20° curvature along the bundle's longest axis. According to the crystal structure, the natural interaction is stabilized by several conserved residues on the C3d side, including L1109, E1110, K1113, P1114, S1164, T1170, K1171, and D1174 (Figure 3B and Figure S3), which mediate thirteen hydrogen bonds, three salt bridges, and 24 residue-residue contacts [31].

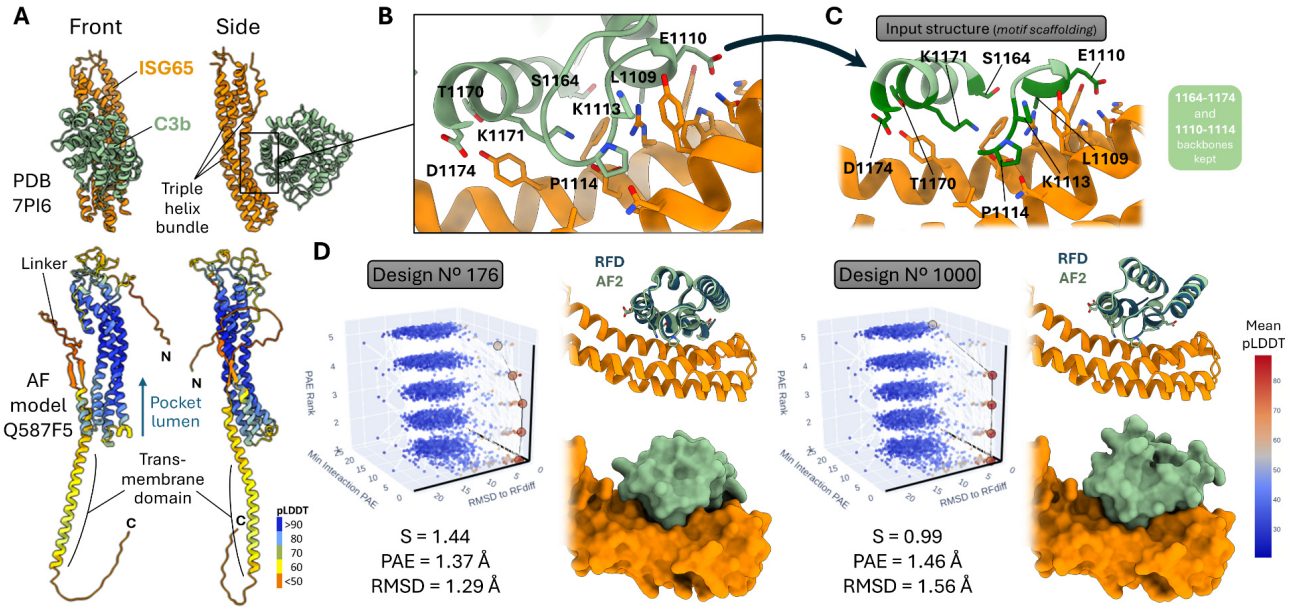

Figure 3. ISG65 receptor of T. brucei and C3d mimetics designs. (**A**) Upper panel: front and side (90° rotation) views of the ISG65 crystal structure bound to human complement C3d (PDB: 7PI6). Lower panel: AF2 model of the full length ISG65 (AF-Q587F5), colored by predicted confidence (pLDDT). (**B**) Close up of the binding interface from PDB 7PI6. Sidechains involved in contacts and H-bonds are shown. The corresponding residue numbers and codes for C3d residues are indicated. (**C**) Input structure for motif scaffolding. Dark green residues were kept completely, conserving only the backbone of light green residues. (**D**) Top two scoring designs. The 3D plot shows the distribution of PAEs and RMSDs, highlighting the values of the top designs. Each marker represents an output model from AF2, colored by pLDDT. The superposition of the RFdiffusion backbone is shown in blue and the one from the rank 1 model of AF2 in green (Upper panel). Solvent exposed surface of the AF2 model (Lower panel). AF2: AlphaFold2; ISG65: invariant surface glycoprotein 65; PAE: predicted aligned error; pLDDT: predicted local distance difference test; RFD: RFdiffusion; RMSD: root mean square deviation.

As input for the motif scaffolding design, we used two stretches of backbones from C3d, 1,110–1,114 and 1,164–1,174, where key interface residues are located (Figure 3C). The idea for this design was to connect them with a backbone of variable size and extend the N- and C-terminal extremes with variable segments. We tested several strategies, and the best results were obtained by inverting the order of the motifs (1,164–1,174 first, instead of 1,110–1,114), maintaining the sidechains of residues L1109, E1110, K1113, P1114, S1164, T1170, K1171 and D1174. As before, ProteinMPNN was used to design sequences

and AF2-multimer for modelling them in complex with the receptor. The results can be seen in Figure 3D, highlighting the top-two scoring designs, which gave *S* values of 1.44 and 0.99, with miPAE values of 1.37 Å and 1.46 Å and RMSD values of 1.29 Å and 1.56 Å for the rank 1 AF2 model. In both cases, AF2 models closely recapitulate the RFdiffusion backbone geometry, and the designs adopt compact folds with good binder-receptor surface complementarity. In this case, the total number of successful designs was 17 (miPAE < 5 Å and RMSD < 2 Å), which was lower than for TfR, but enough to suggest true binding activity in vitro among the candidates.

HpHbR

Heme plays a fundamental role in many cellular processes for the vast majority of living organisms. Structurally, it consists of an iron ion coordinated within a tetrapyrrole ring known as protoporphyrin IX. It is an essential cofactor for proteins involved in oxygen transport and storage, mitochondrial electron transport, drug and steroid metabolism, signal transduction, and antioxidant defense [49]. However, when free, it becomes highly toxic due to its capacity to generate free radicals and cause oxidative damage. In the human blood, heme is normally contained within hemoglobin (Hb) in red blood cells. During intravascular hemolysis, released Hb dissociates into reactive αβ dimers that are neutralized by haptoglobin (Hp), an acute-phase glycoprotein that binds Hb dimers in a virtually irreversible manner, forming the HpHb complex, subsequently cleared by macrophages through CD163-mediated endocytosis, enabling heme degradation and iron recycling [50, 51].

Unlike humans and most eukaryotic organisms, which synthesize heme de novo, kinetoplastid parasites present defects in their biosynthetic pathway. Some species have lost the complete pathway, while others retain only the last three biosynthetic steps [52]. As heme auxotrophs, kinetoplastids have evolved to acquire this essential nutrient from their hosts. For example, it has been established that transmembrane proteins belonging to the heme responsive gene (HRG) family are involved in the transport of free heme from the environment in kinetoplastids [53]. In *T. brucei*, heme acquisition is highly adapted to each stage of its life cycle. During the procyclic stage (PS) in the tsetse fly midgut, heme is obtained from degraded hemoproteins via the TbHRG transporter, localized to the flagellar membrane and the FP [49, 54]. On the other hand, in the BS, the primary route of heme acquisition is bound to Hb, though the parasite's HpHbR, located in the FP, captures HpHb complexes directly from the bloodstream, similarly to the CD163-mediated endocytosis [55]. Once in endo-lysosomes, HpHb is degraded and heme released, where TbHRG would be responsible for transporting it to the cytosol, making it available for essential cellular processes [56]. While heme acquisition mechanisms vary between different kinetoplastid parasites, here we focused on designing binders for the HpHbR from *T. brucei*.

Two experimental structures are available in the Protein Data Bank for the *T. brucei* HpHbR in complex with human HpHb (PDB 5HU6 and PDB 4WJG) (Figure 4A) [33, 34]. This receptor is an elongated protein composed of a three-helix bundle that attaches to the plasma membrane through a GPI anchor covalently bound at its C-terminus. At its membrane-distal end, the receptor widens to form a compact head structure that includes the N-terminus and a 42-residue loop containing two additional helices (Figure 4A, upper panel). Although concentrated within the FP, HpHbR molecules retain lateral mobility within the membrane. This mobility is thought to compensate for the relatively small interaction surface presented by the receptor's narrow and elongated architecture. Indeed, high-affinity ligand recognition is achieved through a bivalent binding mechanism in which two receptor molecules adopt the appropriate orientation to engage a single dimeric HpHb complex, thereby increasing overall binding avidity (Figure 4A, lower panel) [33, 57].

The interaction between HpHbR and its ligand is mediated by two distinct binding regions (Figure 4B). The first binding interface corresponds to the βHb-binding site (Figure 4B, upper panel) and is formed by residues from helices I and II of HpHbR, together with additional contacts contributed by the loop connecting helices III and IV. Notably, the interaction surface involves direct molecular interactions with the heme of βHb (Figure S4), where its carboxylate moieties participate in receptor recognition through several H-bonds. This observation is consistent with previous biophysical studies demonstrating that the

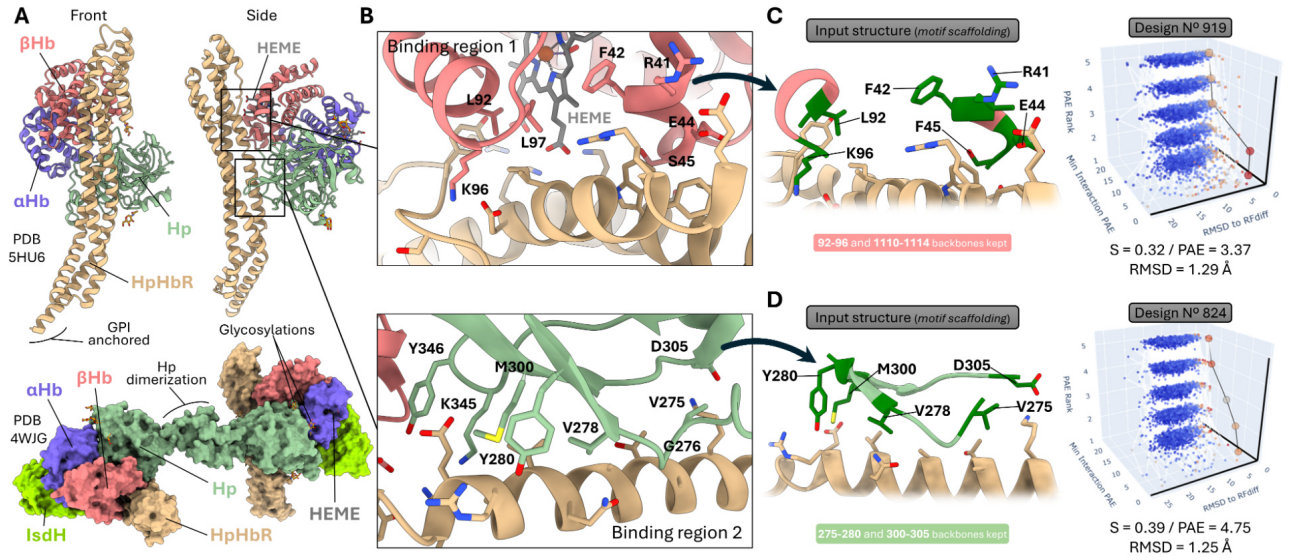

Figure 4. HpHbR of T. brucei and Hp-hemoglobin mimetic designs. (**A**) Upper panel: front and side (90° rotation) views of HpHbR crystal structures bound to human HpHb (PDB 5HU6). The GPI anchor attaches the complex to the FP membrane. Lower panel: Solvent exposed surface of HpHbR structure bound to human HpHb (PDB 4WJG). The dimerization of human Hp, interacting with Hb and HpHbR, is shown in green. The hemoglobin-binding Staphylococcus aureus IsdH NEAT1 domain (shown in bright green) was used as a stabilizer in the crystallization process. (**B**) Close up of the binding interfaces from PDB 5HU6. Sidechains involved in contacts and H-bonds are shown. The corresponding residue numbers and code for Hp and βHb are indicated (Upper panel: βHb-binding region. Lower panel: Hp-binding region). (**C**) Left panel: input structure for design using the first binding region. Dark green residues were kept completely. Only the backbone was kept for pink residues. Right panel: Scatter plot with resulting metrics of the design process, highlighting the top-scoring design and its metrics. (**D**) Left panel: input structure for the second binding region. Dark green residues were kept completely. Only the backbone of light green residues was kept. Right panel: Results highlighting the top-scoring design. Markers were colored by pLDDT (red: high pLDDT; blue: low pLDDT). FP: flagellar pocket; GPI: glycosylphosphatidylinositol; Hp: haptoglobin; HpHb: haptoglobin-hemoglobin; HpHbR: haptoglobin-hemoglobin receptor; PAE: predicted aligned error; pLDDT: predicted local distance difference test; RMSD: root mean square deviation.

presence of heme is required for efficient binding of the HpHb complex by HpHbR, suggesting that direct heme recognition is necessary to ensure the intake of this compound. Besides the interactions mediated by heme, the amino acid residues R41, F42, E44 and S45 from an α-helix of βHb and L92, K96 and L97 from another are involved in the interaction with the receptor. The second binding interface corresponds to the Hp-binding region (Figure 4B, lower panel), which is mainly mediated by residues located in helices I and VI of HpHbR. In this case, no heme is involved, only amino acid residues from three different segments of Hp. Residues V275, G276, V278, Y280 are located in a loop and β-strand, residues M300 and D305 in another loop and β-strand that forms a β-sheet with the first one, and residues K345 and Y346, located in a turn (Figure S5).

For the design process, we tested several strategies without success. For example, we tried to design over the motif from Figure 4C (left panel), conserving the side chains of residues from βHb involved in the first binding region, filling the gap left by heme with a de novo backbone. We also tried using as input the motif from Hp involved in the interaction with the receptor (Figure 4D, left panel). In both cases, we obtained very low scores for the top-scoring designs (0.32 and 0.39, respectively), with very low reproducibility of the metrics across the AF2 models (Figure 4C and 4D, right panel). We even tried to perform several rounds of completely de novo designs (without motif scaffolding), obtaining similar results (data not shown), indicating that this is a particularly difficult target.

For comparison, Figure 5 shows the distribution of AF2 error (as miPAE) and structural discrepancy between AF2 and RFdiffusion (as RMSD) of the rank 1 AF2-multimer models. We can see how the Tf mimetic designs accumulate the most successful models in the lower-left quadrant (67 candidates), followed by ISG65 binders (17 candidates). Meanwhile, binders designed for HpHbR retrieved only one successful candidate for each of the binding regions, but close to the threshold limits.

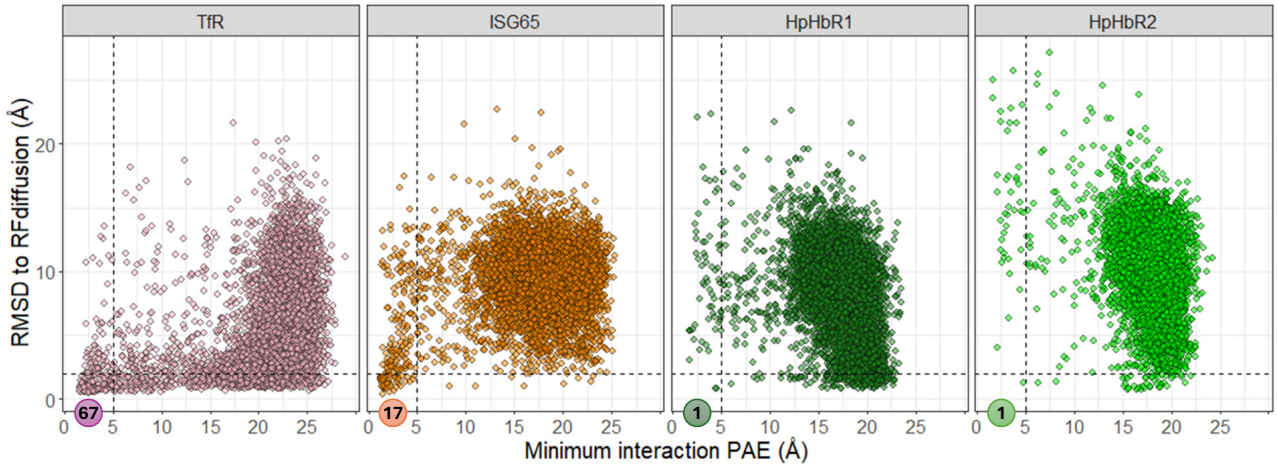

Figure 5. Comparison between results for different targets. Resulting RMSD between AF2 and RFdiffusion backbones and miPAE of rank 1 AF2 models by target. Dashed lines indicate the thresholds of 5 Å for miPAE and 2 Å for RMSD corresponding to in silico success. Circled numbers indicate the number of models surpassing both thresholds for each target. AF2: AlphaFold2; HpHbR1: haptoglobin-hemoglobin receptor 1; miPAE: minimum interaction predicted aligned error; PAE: predicted aligned error; RMSD: root mean square deviation; TfR: human transferrin receptor.

Discussion

In this work, we proposed and computationally explored a strategy to exploit the FP of *T. brucei* as a point of entry for protein-based therapeutic agents. Using experimentally determined structures of three FP receptors (TbTfR, ISG65, and HpHbR), we applied a motif-scaffolding pipeline combining RFdiffusion, ProteinMPNN, and AF2-multimer to design de novo MP modules capable of mimicking the natural cargo recognized by each receptor. Building on those deliverer modules, we further outlined how a modular protein architecture could, in principle, carry additional functional elements [protease recognition sites, cell-penetrating peptides (CPPs), subcellular localization signals, and inhibitory payloads] to ultimately reach and inhibit an intracellular target within the parasite. This work provides a mechanistic and structural rationale for a delivery concept that, to the best of our knowledge, has not been previously applied to kinetoplastid parasites, and offers a concrete pipeline for generating candidate molecules ready for experimental follow-up.

The three receptors explored here yielded markedly different outcomes, and the differences are instructive. For TbTfR, the design campaign produced 67 candidates that simultaneously satisfied the in silico success thresholds of miPAE < 5 Å and RMSD < 2 Å. This pool is sizable enough to constitute a realistic experimental panel: based on the success rate for RFdiffusion-based binder campaigns reported in the original methodology, a set of this magnitude would be expected to contain multiple true binders, and testing fewer than 100 well-filtered designs per target has been shown sufficient to identify micromolar or even nanomolar binders [15]. The outcome for ISG65 was intermediate: 17 candidates passed the same thresholds. While this is a smaller pool, it remains experimentally tractable, and a straightforward path forward would be to increase sampling (running 3,000 to 4,000 RFdiffusion trajectories instead of 1,000), which, given the observed success rate per trajectory, would be expected to yield a pool comparable in size to that obtained for TbTfR.

The situation for HpHbR is qualitatively different. Regardless of the binding design strategy used, the design process consistently returned datasets of poor quality. Crucially, the difficulty does not appear to be confined to the design step itself, as when we attempted to model the natural HpHb-HpHbR complex using AF2-multimer without using templates, and the predictions also failed to recapitulate the known interaction. This systematic failure of the structure prediction model to recognize a well-characterized interface strongly suggests that the obstacle is, at least in part, intrinsic to the binding surface rather than to the particular design strategy employed. A well-known limitation of RFdiffusion documented by its developers is that binding interfaces dominated by polar and charged residues, with fewer than approximately three exposed hydrophobic anchor residues, are substantially harder to target [15]. The

HpHbR binding sites, which are characterized by a prevalence of charged and polar contacts with Hb and Hp, are consistent with this description. Beyond the intrinsic polarity of the interface, the requirement for direct coordination with the heme moiety of Hb at the βHb-binding region introduces an additional complexity, as the cofactor occupies a geometrically critical position in the interface that cannot be straightforwardly replaced by a de novo designed peptide backbone alone. For both binding regions, we think that these features combine to make HpHbR a challenging target for the current pipeline. An alternative approach worth exploring is BindCraft [58], a recently developed pipeline that, rather than using RFdiffusion for backbone generation, employs AF2-guided backpropagation to iteratively optimize binder sequences in co-folding with the target. This co-folding framework may be more appropriate for interfaces where the backbone geometry of the binder and the precise placement of polar groups need to be mutually adjusted. BindCraft reported substantially improved success rates, making it a natural next step for the HpHbR design problem.

More broadly, the field of generative protein design is advancing at an extraordinary pace, and it is likely that emerging frameworks will further improve the success rates achieved by the pipeline presented here. Beyond the tools used in this work, recent methods such as RFdiffusion2 [59], and its successor, RFdiffusion3 [60] (still under peer review), which incorporate native all-atom architectures and finer atomic-level control of molecular interactions; or BoltzGen [61], an open-source all-atom generative model (also still under peer review); represent promising alternatives. As these methods continue to mature and become broadly accessible for therapeutic development, they can be readily incorporated into the modular workflow proposed here, further improving binder design and candidate prioritization.

Regardless of target, the in silico metrics used here (miPAE, RMSD against the RFdiffusion backbone, and pLDDT) are predictors of binding activity, not guarantees of it. Accordingly, experimental validation of the proposed delivery platform should proceed in a stepwise manner, from recombinant production and receptor-binding measurements to uptake studies and, ultimately, evaluation of intracellular cargo delivery and target inhibition. The logical experimental progression for the top-ranked candidates begins with direct binding validation. Recombinant expression of the designed MPs, followed by biophysical assessment using surface plasmon resonance (SPR) or biolayer interferometry (BLI), would provide quantitative affinity constants and confirm that the predicted interaction is realized in solution [58]. For candidates confirmed to bind their target receptor in vitro, the next critical question is whether that binding is sufficient to trigger active internalization. A well-established approach for probing receptor-mediated endocytosis in *T. brucei* involves labelling candidate ligands with fluorescent dyes (e.g., Alexa Fluor conjugates) and incubating live bloodstream-form parasites at permissive temperatures to allow endocytosis to proceed, followed by flow cytometry or confocal fluorescence microscopy to quantify and localize the internalized signal [62]. Co-localization with endo-lysosomal markers can further establish whether the internalized cargo follows the expected trafficking route to the endo-lysosomal compartment.

Although the receptors explored in this work are constitutively internalized and recycled as part of their physiological function, receptor binding does not necessarily guarantee productive endocytosis, as different binders may differentially modulate receptor activity. Consequently, similar to other protein binders targeting receptors [19], some designed mini-proteins could behave as agonists, whereas others might act as antagonists or otherwise interfere with receptor trafficking. While complete inhibition of receptor cycling is therefore considered unlikely, this possibility should be experimentally assessed. Importantly, the availability of multiple high-confidence candidates from each design campaign provides an opportunity to select binders that potentially preserve productive receptor-mediated uptake. If necessary, additional optimization rounds, for example using partial diffusion approaches [63], could also be employed to generate alternative variants with improved functional properties.

Demonstrating productive internalization of a deliverer module is necessary but not sufficient for the full therapeutic concept proposed here. The complete therapeutic concept must ultimately demonstrate that the modular construct undergoes protease-mediated cargo release, escapes the endo-lysosomal compartment, reaches the intended intracellular destination, and produces the expected functional inhibition of its molecular target. Experimental evaluation of these later stages could combine fluorescence-

based localization assays with target-specific functional readouts, such as measurements of enzymatic activity, pathway perturbation, or parasite growth and viability, depending on the inhibitory payload employed. Notably, the central challenge of any receptor-mediated intracellular delivery strategy is escaping the endo-lysosomal compartment before the cargo is degraded, and reaching the intended intracellular target in a functional state. In the modular design we propose, this problem is addressed at three successive levels (Figure 1).

First, a protease recognition site embedded in a flexible linker between the deliverer and the inhibitor modules would be cleaved by endo-lysosomal proteases, releasing the inhibitor-containing half of the construct. Only two protease families are known to be active in the endo-lysosomal compartment of trypanosomatids: serine carboxypeptidases (specifically CBP1A–C), which preferentially cleave C-terminal hydrophobic dipeptide motifs including Phe-Phe [64], and cysteine cathepsins of the rhodesain and Trypanosoma brucei cathepsin B (TbCatB) type, the latter of which has a documented preference for substrates with basic residues at P1 and large hydrophobic residues at P2 [65, 66]. The optimal protease recognition sequence (PRS) for a given construct could therefore be selected from these known specificity profiles. However, it must be acknowledged that the inventory of endo-lysosomal proteases expressed at different life-cycle stages of each parasite species, and the precise substrate preferences of those not yet biochemically characterized, remain incompletely defined. This is one of several fundamental knowledge gaps that currently limit the rational optimization of the release module. Nonetheless, from these two known families, the second proteolytic system of interest would be more accurate to choose, as they were shown to act as endopeptidases. They involve cysteine proteases of the cathepsin family, belonging to clan CA (papain family), which are synthesized as inactive precursors (zymogens) bearing N-terminal propeptides that block the active site until the protein reaches the acidic endo-lysosomal environment. Their primary role is degradation of host-derived proteins captured by endocytosis, including, for example, Tf, a process essential for iron acquisition by the parasite. The two most prominent members are Rhodesian (cathepsin L-like), the most abundantly expressed, and TbCatB (cathepsin B-like), which, despite its lower abundance, is essential for parasite survival [67]. TbCatB is distinguished by the presence of an occluding loop, a structural element that confers dual enzymatic functionality: it can act either as an endopeptidase, cleaving internal peptide bonds, or as an exopeptidase, removing dipeptide units from the C-terminus of the substrate. TbCatB displays a marked preference for basic residues at the P1 position and large hydrophobic residues at P2, with the motif P4(Arg/Lys)-P3(Arg/Lys)-P2(X)-P1(Arg/Lys) identified as the optimal cleavage sequence [65]. These motifs should be omitted in the rest of the modules and incorporated into a linker between the delivery module and the inhibitor module (Figure 1).

Second, escape from the endo-lysosomal lumen into the cytoplasm requires active membrane destabilization, which is the role assigned to the CPP in the modular design. Table 1 shows different CPP types and their characteristics. In particular, histidine-switching CPPs (hsCPPs) are particularly well-suited for this step because their membrane-disrupting activity is pH-conditional. They remain largely neutral at physiological pH, minimizing non-specific uptake during circulation, and become cationic and membrane-active only upon protonation in the acidic endosomal environment [28]. This is a critical property for a system that depends on the parasite's own endocytic machinery to deliver the cargo.

Table 1. CPPs types.

Type	Relevant characteristics	Motif	Ref.
Cationic	High positive charge: electrostatic interactions with membrane	5–30 AA: rich in R/K	[68]
Amphiphilic	Hydrophilic and hydrophobic domains	Amphipathic α-helices or β-sheets	[68]
Hydrophobic	Predominantly nonpolar residues	Hydrophobic residue-rich motifs	[68]
hsCPPs	Modified class of CPPs pH-dependent activation	K/R substitution for H	[28]

Table 1. CPPs types. (*continued*)

Type	Relevant characteristics	Motif	Ref.
Designed with NNJA	Platform to design novel CPPs specific to any cell type using phage display and directed evolution	Variable	[69]

Most important characteristics of each type of CPP signal along with its predicted motif. AA: amino acid; CPPs: cell-penetrating peptides; hsCPPs: histidine-switching cell-penetrating peptides. NNJA: novel peptides for intracellular delivery by hijacking two cell systems.

Third, once in the cytoplasm, subcellular localization signals fused to the inhibitor module would direct it to the appropriate compartment (nucleus, glycosome, mitochondrion, flagellum, etc.) depending on the subcellular localization of the target for inhibition. A non-comprehensive list of possible localization signals is presented in Table 2. For example, for directing heterologous cargo to the nucleus of *T. cruzi*, we could fuse the widely used nuclear localization signal (NLS) from the simian virus 40 (SV40), which we validated experimentally in prior work from our group [70]. One thing to notice is that some signals need to be placed at specific positions of the design for them to work properly, like mitochondrial targeting peptides (mTPs), which need to be located at the N-terminal [71].

Table 2. Subcellular localization signals.

Type	Relevant characteristics	Motif	Ref.
NLS	Mono- (4 or 7 basic AA block) or bipartite (2 basic segments linked by 10–12 AA)	4 AA motif: (B4), (B(B2P)) or K(R/K)X(R/K) /KRRR 7 AA motif: P[(R/K)3X3]	[29, 72]
PTS1	C-terminal tripeptide for glycosomal imports. Recognized by PEX5	(S/C/A)-(K/R/H)-(L/M) Uncharged-positively charged-large hydrophobic residue	[72–74]
PTS2	N-terminal nonapeptide for glycosomal imports. Recognized by PEX7	(R/K)-(L/V/I/Q)-XX-(L/V/I/H)-(L/S/G/A) -X-(H/Q)-(L/A)	[72–74]
mTP	Transient N-terminal signal. Cleaved by mitochondrial processing peptidase	(M/L)RR within first 7–10 AA + hydrophobic region (A, L, V) enriched. Variable length: 5–118 AA	[71]
FLS	Directs protein to flagellum via IFT complexes. Highly variable, no unified consensus	55 AA N-terminal motif (adenylate kinases) or Bipartite C-terminal motif: ≤ 56 AA + 7 AA region/C-terminal 45-residue domain	[72]
mPTS	Glycosomal membrane-bound proteins (different from PTS1/PTS2)	~11 hydrophobic or basic AA adjacent to transmembrane domain	[72, 74]

It describes the most important characteristics of each type of signal along with its predicted motif, where X represents any amino acids. AA: amino acid; NLS: nuclear localization signal; PTS1: peroxisome targeting sequence 1; mTP: mitochondrial targeting peptide; FLS: flagellar localization signal; mPTS: membrane peroxisome targeting sequence.

Beyond the design and delivery challenges, protein-based drugs offer a fundamental advantage over small molecules. Unlike small molecules, which must be chemically synthesized and administered as such, protein-based therapeutics can be delivered in multiple formats: as purified recombinant proteins, as mRNA encoding the therapeutic sequence, or as DNA constructs for stable expression [75]. This versatility opens a therapeutic spectrum that ranges from acute intervention, where recombinant protein administration would clear an active infection, to long-term prophylactic strategies. For example, sustained hepatic expression of a functional protein design, like the one recently achieved using the apoA1-based gene therapy system [76], could maintain plasma levels sufficient to make the host refractory to infection, functioning as a molecular “vaccine”. This is something that small molecules, regardless of their potency, can’t access. Furthermore, the programmable and modular nature of protein-based designs allows fast iterative optimization at the sequence level, accelerating the generation and testing of variants in ways that are not straightforward with traditional medicinal chemistry approaches.

It is also worth noting an important distinction regarding which life-cycle stage and which parasite species are best suited to this delivery concept. In *T. brucei*, the bloodstream form is an extracellular pathogen exposed to the host circulation, making the FP directly accessible to systemically administered protein-based drugs. However, in *T. cruzi*, the replicative forms, epimastigotes and amastigotes, also have a CC complex, which is much more active compared to the FP, but these stages reside within the insect vector

or inside host cells, respectively. On one hand, the intracellular environment of amastigotes presents a fundamentally different pharmacological access problem compared to an extracellular bloodstream parasite like *T. brucei*, while the CC would be more practical as a delivery portal for vector-targeting strategies, rather than for treating the human host. In humans, the FP of bloodstream *T. cruzi* trypomastigotes before entering host cells represents a more tractable target, but it is much less active compared to the FP of *T. brucei*. In *Leishmania* spp., the FP is the only endocytic structure. Nevertheless, amastigotes reside within macrophage phagolysosomes, introducing an additional barrier to drug delivery, like the situation encountered with *T. cruzi* amastigotes. Therefore, targeting may be more straightforward against transient extracellular amastigotes or inside damaged macrophages with permeable membranes, although the duration and accessibility of these stages in vivo remain limiting factors [77]. These considerations highlight the importance of stage-specific biology when translating this delivery paradigm across different kinetoplastid pathogens.

A broader theme that emerges from this work is how much of the basic cell biology required to optimize this delivery concept remains to be characterized. The proteolytic complement of the lysosome, beyond the well-characterized rhodesain and TbCatB, has not been fully inventoried, and the substrate specificities of many additional trypanosomatid hydrolases await biochemical definition. Similarly, the identities of the receptor proteins that mediate endocytosis in the FP of *T. cruzi* and *Leishmania* remain mostly unknown. As we noted, no gene identifiers have been assigned to the putative TfRs of either organism, while the endocytic machinery of the FP in these species is far less characterized than that of *T. brucei* [25, 26]. The CC complex of *T. cruzi*, though highly active in endocytosis, contains virtually no molecularly defined receptor that could serve as a delivery target [25]. These gaps are not peripheral details; they are the information that would allow the rational selection of deliverer modules, PRS sequences, and CPP variants optimized for each parasite and life-cycle stage. Closing them will require sustained basic research investment in kinetoplastid cell biology, and this work is intended, in part, to highlight what that investment would enable. The convergence of increasingly powerful AI-assisted protein design tools with a growing structural understanding of FP and CC receptors creates a genuine opportunity for a new class of protein-based therapeutics against these neglected diseases, but realizing that opportunity will depend on filling the many mechanistic gaps that currently separate computational design from experimental proof of concept.

Abbreviations

AF2: AlphaFold2

AI: artificial intelligence

BLI: biolayer interferometry

BS: bloodstream stage

CC: cytostome-cytopharynx

CPPs: cell-penetrating peptides

cryo-EM: cryogenic electron microscopy

FLS: flagellar localization signal

FP: flagellar pocket

GCGR: glucagon receptor

GIPR: glucose-dependent insulinotropic polypeptide receptor

GLP1R: glucagon-like peptide-1 receptor

GPCRs: G protein-coupled receptors

GPI: glycosylphosphatidylinositol

Hb: hemoglobin

Hp: haptoglobin

HpHb: haptoglobin-hemoglobin

HpHbR: haptoglobin-hemoglobin receptor

HRG: heme responsive gene

hsCPPs: histidine-switching cell-penetrating peptides

ipTM: interface predicted template modeling score

ISG65: invariant surface glycoprotein 65

miPAE: minimum interaction predicted aligned error

MPs: miniproteins

mPTS: membrane peroxisome targeting sequence

mTPs: mitochondrial targeting peptides

NLS: nuclear localization signal

NNJA: novel peptides for intracellular delivery by hijacking two cell systems

PAE: predicted aligned error

pLDDT: predicted local distance difference test

PPIs: protein-protein interactions

PRS: protease recognition sequence

PS: procyclic stage

PTS1: peroxisome targeting sequence 1

RF: RoseTTAFold

RFNA: RoseTTAFoldNA

RMSD: root mean square deviation

SPR: surface plasmon resonance

SV40: simian virus 40

TbCatB: Trypanosoma brucei cathepsin B

TbTfR: *Trypanosoma brucei* transferrin receptor

TED: thioester domain

Tf: transferrin

TfRs: human transferrin receptors

TriTryps: *Trypanosoma brucei*, *Trypanosoma cruzi*, and *Leishmania* spp.

VSG: variant surface glycoprotein

Supplementary materials

The supplementary figures for this article are available at: https://www.explorationpub.com/uploads/Article/file/1008179_sup_1.pdf.

Declarations

Author contributions

LA: Data curation, Formal analysis, Investigation, Writing—original draft, Writing—review & editing. CV: Data curation, Formal analysis, Investigation, Writing—review & editing. ERA: Conceptualization, Formal

analysis, Investigation, Methodology, Project administration, Software, Visualization, Writing—original draft, Writing—review & editing, Supervision. All authors read and approved the submitted version.

Conflicts of interest

The authors declare that they have no conflicts of interest.

Ethical approval

Not applicable.

Consent to participate

Not applicable.

Consent to publication

Not applicable.

Availability of data and materials

All the code to reproduce these results, along the resulting analysis of the designs (like the interactive 3D scatter plots) is available at https://github.com/elviorodriguez/binders_for_trypanosomatid_flagellar_pocket_receptors.

Funding

Not applicable.

Publisher's note

Open Exploration maintains a neutral stance on jurisdictional claims in published institutional affiliations and maps. All opinions expressed in this article are the personal views of the author(s) and do not represent the stance of the editorial team or the publisher.

References

1. Rijo-Ferreira F, Takahashi JS. Sleeping Sickness: A Tale of Two Clocks. Front Cell Infect Microbiol. 2020;10:525097. [DOI] [PubMed] [PMC]
2. Pérez-Molina JA, Molina I. Chagas disease. Lancet. 2018;391:82–94. [DOI] [PubMed]
3. Burza S, Croft SL, Boelaert M. Leishmaniasis – Authors' reply. Lancet. 2019;393:872–3. [DOI] [PubMed]
4. Emergency guidance on use of licensed Ebola vaccine during current outbreak [Internet]. WHO; c2026 [cited 2026 Jun 7]. Available from: https://www.who.int/
5. Lutje V, Probyn K, Seixas J, Bergman H, Villanueva G. Chemotherapy for second-stage human African trypanosomiasis: drugs in use. Cochrane Database Syst Rev. 2021;12:CD015374. [DOI] [PubMed] [PMC]
6. Pérez-Molina JA, Crespillo-Andújar C, Bosch-Nicolau P, Molina I. Trypanocidal treatment of Chagas disease. Enferm infecc microbiol clin (Engl ed,). 2021;39:458–70. [DOI] [PubMed]
7. Taslimi Y, Zahedifard F, Rafati S. Leishmaniasis and various immunotherapeutic approaches. Parasitology. 2016;145:497–507. [DOI] [PubMed]
8. Geiger A, Bossard G, Sereno D, Pissarra J, Lemesre JL, Vincendeau P, et al. Escaping Deleterious Immune Response in Their Hosts: Lessons from Trypanosomatids. Front Immunol. 2016;7:212. [DOI] [PubMed] [PMC]

9. Matovu E, Seebeck T, Enyaru JC, Kaminsky R. Drug resistance in spp., the causative agents of sleeping sickness in man and nagana in cattle. Microbes Infect. 2001;3:763–70. [DOI] [PubMed]
10. Campos MC, Leon LL, Taylor MC, Kelly JM. Benznidazole-resistance in Trypanosoma cruzi: Evidence that distinct mechanisms can act in concert. Mol Biochem Parasitol. 2014;193:17–9. [DOI] [PubMed] [PMC]
11. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596:583–9. [DOI] [PubMed] [PMC]
12. Baek M, McHugh R, Anishchenko I, Jiang H, Baker D, DiMaio F. Accurate prediction of protein–nucleic acid complexes using RoseTTAFoldNA. Nat Methods. 2023;21:117–21. [DOI] [PubMed] [PMC]
13. Baek M, DiMaio F, Anishchenko I, Dauparas J, Ovchinnikov S, Lee GR, et al. Accurate prediction of protein structures and interactions using a three-track neural network. Science. 2021;373:871–6. [DOI] [PubMed] [PMC]
14. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature. 2024;630:493–500. [DOI] [PubMed] [PMC]
15. Watson JL, Juergens D, Bennett NR, Trippe BL, Yim J, Eisenach HE, et al. De novo design of protein structure and function with RFdiffusion. Nature. 2023;620:1089–100. [DOI] [PubMed] [PMC]
16. Dauparas J, Anishchenko I, Bennett N, Bai H, Ragotte RJ, Milles LF, et al. Robust deep learning–based protein sequence design using ProteinMPNN. Science. 2022;378:49–56. [DOI] [PubMed] [PMC]
17. Notin P, Rollins N, Gal Y, Sander C, Marks D. Machine learning for functional protein design. Nat Biotechnol. 2024;42:216–28. [DOI] [PubMed] [PMC]
18. Graham F. Daily briefing: AlphaFold developers share Nobel Prize in Chemistry. Nature. 2024;[Epub ahead of print]. [DOI] [PubMed]
19. Muratspahić E, Feldman D, Kim DE, Qu X, Bratovianu AM, Rivera-Sánchez P, et al. *De novo* design of miniprotein agonists and antagonists targeting G protein-coupled receptors. bioRxiv [Preprint]. 2025 [cited 2026 Jun 7]. Available from: https://www.biorxiv.org/content/10.1101/2025.03.23.644666v3
20. Lee J, Case JB, Park YJ, Ravichandran R, Asarnow D, Tortorici MA, et al. The computationally designed TRI2-2 miniprotein inhibitor protects against multiple SARS-CoV-2 Omicron variants. Commun Biol. 2026;9:224. [DOI] [PubMed] [PMC]
21. Chazin-Gray AM, Thompson TR, Lopatto EDB, Magala P, Erickson PW, Hunt AC, et al. De Novo Design of Miniprotein Inhibitors of Bacterial Adhesins. bioRxiv [Preprint]. 2025 [cited 2026 Jun 7]. Available from: https://www.biorxiv.org/content/10.1101/2025.08.18.670751v1
22. Ciesiołkiewicz A, Lizandra Perez J, Skalniak L, Noceń P, Berlicki Ł. Miniprotein engineering for inhibition of PD-1/PD-L1 interaction. Protein Sci. 2024;33:e5106. [DOI] [PubMed] [PMC]
23. Berger S, Seeger F, Yu TY, Aydin M, Yang H, Rosenblum D, et al. Preclinical proof of principle for orally delivered Th17 antagonist miniproteins. Cell. 2024;187:4305–17.e18. [DOI] [PubMed] [PMC]
24. Borges AR, Link F, Engstler M, Jones NG. The Glycosylphosphatidylinositol Anchor: A Linchpin for Cell Surface Versatility of Trypanosomatids. Front Cell Dev Biol. 2021;9:720536. [DOI] [PubMed] [PMC]
25. Cunha-E-Silva NLD, Alcantara CL, Pereira MG, De Souza W. Three Hungry Tryps: the efficient endocytic pathway of pathogenic trypanosomatids. Trends Parasitol. 2025;41:560–71. [DOI] [PubMed]
26. Field MC, Carrington M. The trypanosome flagellar pocket. Nat Rev Microbiol. 2009;7:775–86. [DOI] [PubMed]
27. Kortemme T. De novo protein design—From new structures to programmable functions. Cell. 2024; 187:526–44. [DOI] [PubMed] [PMC]
28. Zhao Y, Jiang H, Yu J, Wang L, Du J. Engineered Histidine-Rich Peptides Enhance Endosomal Escape for Antibody-Targeted Intracellular Delivery of Functional Proteins. Angew Chem Int Ed. 2023;62: e202304692. [DOI] [PubMed]
29. Canela-Pérez I, López-Villaseñor I, Mendoza L, Cevallos AM, Hernández R. Nuclear localization signals in trypanosomal proteins. Mol Biochem Parasitol. 2019;229:15–23. [DOI] [PubMed]

30. Trevor CE, Gonzalez-Munoz AL, Macleod OJS, Woodcock PG, Rust S, Vaughan TJ, et al. Structure of the trypanosome transferrin receptor reveals mechanisms of ligand recognition and immune evasion. Nat Microbiol. 2019;4:2074–81. [DOI] [PubMed] [PMC]
31. Macleod OJS, Cook AD, Webb H, Crow M, Burns R, Redpath M, et al. Invariant surface glycoprotein 65 of Trypanosoma brucei is a complement C3 receptor. Nat Commun. 2022;13:5085. [DOI] [PubMed] [PMC]
32. Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res. 2021;50:D439–44. [DOI] [PubMed] [PMC]
33. Lane-Serff H, MacGregor P, Lowe ED, Carrington M, Higgins MK. Structural basis for ligand and innate immunity factor uptake by the trypanosome haptoglobin-haemoglobin receptor. eLife. 2014;3: e05553. [DOI] [PubMed] [PMC]
34. Stødkilde K, Torvund-Jensen M, Moestrup SK, Andersen CB. Structural basis for trypanosomal haem acquisition and susceptibility to the host innate immune system. Nat Commun. 2014;5:5487. [DOI] [PubMed]
35. Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, Green T, et al. Protein complex prediction with AlphaFold-Multimer. bioRxiv [Preprint]. 2022 [cited 2026 Jun 7]. Available from: https://www.biorxiv.org/content/10.1101/2021.10.04.463034v2.abstract
36. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all. Nat Methods. 2022;19:679–82. [DOI] [PubMed] [PMC]
37. Bryant P, Pozzati G, Elofsson A. Improved prediction of protein-protein interactions using AlphaFold2. Nat Commun. 2022;13:1265. [DOI] [PubMed] [PMC]
38. Reyes-López M, Piña-Vázquez C, Serrano-Luna J. Transferrin: Endocytosis and Cell Signaling in Parasitic Protozoa. BioMed Res Int. 2015;2015:641392. [DOI] [PubMed] [PMC]
39. Kariuki CK, Stijlemans B, Magez S. The Trypanosomal Transferrin Receptor of Trypanosoma Brucei—A Review. Trop Med Infect Dis. 2019;4:126. [DOI] [PubMed] [PMC]
40. Ansari I, Basak R, Mukhopadhyay A. Hemoglobin Endocytosis and Intracellular Trafficking: A Novel Way of Heme Acquisition by Leishmania. Pathogens. 2022;11:585. [DOI] [PubMed] [PMC]
41. Lima MF, Villalta F. Trypanosoma cruzi receptors for human transferrin and their role. Mol Biochem Parasitol. 1990;38:245–52. [DOI] [PubMed]
42. Eger I, Soares MJ. Endocytosis in Trypanosoma cruzi (Euglenozoa: Kinetoplastea) epimastigotes: Visualization of ingested transferrin–gold nanoparticle complexes by confocal laser microscopy. J Microbiol Methods. 2012;91:101–5. [DOI] [PubMed]
43. Corrêa JR, Atella GC, Batista MM, Soares MJ. Transferrin uptake in Trypanosoma cruzi is impaired by interference on cytostome-associated cytoskeleton elements and stability of membrane cholesterol, but not by obstruction of clathrin-dependent endocytosis. Exp Parasitol. 2008;119:58–66. [DOI] [PubMed]
44. Sülzen H, Began J, Dhillon A, Kereïche S, Pompach P, Votrubova J, et al. Cryo-EM structures of Trypanosoma brucei gambiense ISG65 with human complement C3 and C3b and their roles in alternative pathway restriction. Nat Commun. 2023;14:2403. [DOI] [PubMed] [PMC]
45. Ziegelbauer K, Overath P. Identification of invariant surface glycoproteins in the bloodstream stage of Trypanosoma brucei. J Biol Chem. 1992;267:10791–6. [DOI]
46. Chung WL, Leung KF, Carrington M, Field MC. Ubiquitylation is Required for Degradation of Transmembrane Surface Proteins in Trypanosomes. Traffic. 2008;9:1681–97. [DOI] [PubMed]
47. Lidani KCF, Bavia L, Ambrosio AR, de Messias-Reason IJ. The Complement System: A Prey of Trypanosoma cruzi. Front Microbiol. 2017;8:607. [DOI] [PubMed] [PMC]
48. Chan A, Ayala JM, Alvarez F, Piccirillo C, Dong G, Langlais D, et al. The role of Leishmania GP63 in the modulation of innate inflammatory response to Leishmania major infection. PLOS ONE. 2021;16: e0262158. [DOI] [PubMed] [PMC]

49. Tripodi KE, Menendez Bravo SM, Cricco JA. Role of Heme and Heme-Proteins in Trypanosomatid Essential Metabolic Pathways. Enzyme Res. 2011;2011:873230. [DOI] [PubMed] [PMC]
50. Andersen CB, Torvund-Jensen M, Nielsen MJ, de Oliveira CL, Hersleth HP, Andersen NH, et al. Structure of the haptoglobin–haemoglobin complex. Nature. 2012;489:456–9. [DOI] [PubMed]
51. Nantasenamat C, Prachayasittikul V, Bulow L. Molecular Modeling of the Human Hemoglobin-Haptoglobin Complex Sheds Light on the Protective Mechanisms of Haptoglobin. PLoS ONE. 2013;8: e62996. [DOI] [PubMed] [PMC]
52. Korený L, Lukes J, Oborník M. Evolution of the haem synthetic pathway in kinetoplastid flagellates: An essential pathway that is not essential after all? Int J Parasitol. 2010;40:149–56. [DOI] [PubMed]
53. Tevere E, Di Capua CB, Chasen NM, Etheridge RD, Cricco JA. Trypanosoma cruzi heme responsive gene (TcHRG) plays a central role in orchestrating heme uptake in epimastigotes. FEBS J. 2023;291: 1186–98. [DOI] [PubMed] [PMC]
54. Horáková E, Changmai P, Vancová M, Sobotka R, Van Den Abbeele J, Vanhollebeke B, et al. The Trypanosoma brucei TbHrg protein is a heme transporter involved in the regulation of stage-specific morphological transitions. J Biol Chem. 2017;292:6998–7010. [DOI] [PubMed] [PMC]
55. Horáková E, Lecordier L, Cunha P, Sobotka R, Changmai P, Langedijk CJM, et al. Heme-deficient metabolism and impaired cellular differentiation as an evolutionary trade-off for human infectivity in Trypanosoma brucei gambiense. Nat Commun. 2022;13:7075. [DOI] [PubMed] [PMC]
56. Cabello-Donayre M, Malagarie-Cazenave S, Campos-Salinas J, Gálvez FJ, Rodríguez-Martínez A, Pineda-Molina E, et al. Trypanosomatid parasites rescue heme from endocytosed hemoglobin through lysosomal HRG transporters. Mol Microbiol. 2016;101:895–908. [DOI] [PubMed]
57. Higgins MK, Lane-Serff H, MacGregor P, Carrington M. A Receptor's Tale: An Eon in the Life of a Trypanosome Receptor. PLOS Pathog. 2017;13:e1006055. [DOI] [PubMed] [PMC]
58. Pacesa M, Nickel L, Schellhaas C, Schmidt J, Pyatova E, Kissling L, et al. One-shot design of functional protein binders with BindCraft. Nature. 2025;646:483–92. [DOI] [PubMed] [PMC]
59. Ahern W, Yim J, Tischer D, Salike S, Woodbury SM, Kim D, et al. Atom-level enzyme active site scaffolding using RFdiffusion2. Nat Methods. 2025;23:96–105. [DOI] [PubMed] [PMC]
60. Butcher J, Krishna R, Mitra R, Brent RI, Li Y, Corley N, et al. *De novo* Design of All-atom Biomolecular Interactions with RFdiffusion3. bioRxiv [Preprint]. 2025 [cited 2026 Jun 7]. Available from: https://www.biorxiv.org/content/10.1101/2025.09.18.676967v2
61. Stark H, Faltings F, Choi M, Xie Y, Hur E, O'Donnell T, et al. BoltzGen: Toward Universal Binder Design. bioRxiv [Preprint]. 2026 [cited 2026 Jun 7]. Available from: https://www.biorxiv.org/content/10.1101/2025.11.20.689494v2
62. Umaer K, Bush PJ, Bangs JD. Rab11 mediates selective recycling and endocytic trafficking in *Trypanosoma brucei*. Traffic. 2018;19:406–20. [DOI] [PubMed] [PMC]
63. Vázquez Torres S, Leung PJY, Venkatesh P, Lutz ID, Hink F, Huynh HH, et al. De novo design of high-affinity binders of bioactive helical peptides. Nature. 2023;626:435–42. [DOI] [PubMed] [PMC]
64. Parussini F, García M, Mucci J, Agüero F, Sánchez D, Hellman U, et al. Characterization of a lysosomal serine carboxypeptidase from Trypanosoma cruzi. Mol Biochem Parasitol. 2003;131:11–23. [DOI] [PubMed]
65. O'Brien TC, Mackey ZB, Fetter RD, Choe Y, O'Donoghue AJ, Zhou M, et al. A Parasite Cysteine Protease Is Key to Host Protein Degradation and Iron Acquisition. J Biol Chem. 2008;283:28934–43. [DOI] [PubMed] [PMC]
66. Redecke L, Nass K, DePonte DP, White TA, Rehders D, Barty A, et al. Natively Inhibited *Trypanosoma brucei* Cathepsin B Structure Determined by Using an X-ray Laser. Science. 2013;339:227–30. [DOI] [PubMed] [PMC]
67. Mackey ZB, O'Brien TC, Greenbaum DC, Blank RB, McKerrow JH. A Cathepsin B-like Protease Is Required for Host Protein Degradation in Trypanosoma brucei. J Biol Chem. 2004;279:48426–33. [DOI] [PubMed]

68. Datta N. A Review on the Cell-Penetrating Peptides. Cell Ther Eng Connect. 2025;1:1. [DOI]
69. Liu J, Heddleston J, Perkins DR, Chen JJH, Ghanbarpour A, Smith BW, et al. Discovery of a new class of cell-penetrating peptides by novel phage display platform. Sci Rep. 2024;14:13437. [DOI] [PubMed] [PMC]
70. Rodríguez Araya E, Martínez Peralta G, Attala L, Boselli V, de Hernández A, da Cunha JPC, et al. Exploring Protein Interactomes Using TurboID-Directed Proximity Labeling and Mass Spectrometry. Methods Mol Biol. 2026;3013:315–42. [DOI] [PubMed]
71. Krnáčová K, Vesteg M, Hampl V, Vlček Č, Horváth A. Euglena gracilis and Trypanosomatids Possess Common Patterns in Predicted Mitochondrial Targeting Presequences. J Mol Evol. 2012;75:119–29. [DOI] [PubMed]
72. Pendlebury M, Lukeš J, Hammond MJ. You can go your own way: The targeting signals of trypanosomatid parasites. PLOS Pathog. 2025;21:e1013326. [DOI] [PubMed] [PMC]
73. Galland N, de Walque S, Voncken FG, Verlinde CL, Michels PA. An internal sequence targets Trypanosoma brucei triosephosphate isomerase to glycosomes. Mol Biochem Parasitol. 2010;171: 45–9. [DOI] [PubMed]
74. Moyersoen J, Choe J, Fan E, Hol WG, Michels PA. Biogenesis of peroxisomes and glycosomes: trypanosomatid glycosome assembly is a promising new drug target. FEMS Microbiol Rev. 2004;28: 603–43. [DOI] [PubMed]
75. Vavilis T, Stamoula E, Ainatzoglou A, Sachinidis A, Lamprinou M, Dardalas I, et al. mRNA in the Context of Protein Replacement Therapy. Pharmaceutics. 2023;15:166. [DOI] [PubMed] [PMC]
76. De Giorgi M, Li A, Hurley A, Barzi M, Doerfler AM, Cherayil NA, et al. Targeting the Apoa1 locus for liver-directed gene therapy. Mol Ther - Methods Clin Dev. 2021;21:656–69. [DOI] [PubMed] [PMC]
77. Real F, Florentino PT, Reis LC, Ramos-Sanchez EM, Veras PS, Goto H, et al. Cell-to-cell transfer of Leishmania amazonensis amastigotes is mediated by immunomodulatory LAMP-rich parasitophorous extrusions. Cell Microbiol. 2014;16:1549–64. [DOI] [PubMed] [PMC]