Computational backbone design enables soluble engineering of transferrin receptor apical domain

Dick J. Sjöström | Sarah A. Berger | Gustav Oberdorfer | Sinisa Bjelic

Abstract
Supply of iron into human cells is achieved by iron carrier protein transferrin and its receptor that upon complex formation get internalized by endocytosis. Similarly, the iron needs to be delivered into the brain, and necessitates the transport across the blood-brain barrier. While there are still unanswered questions about these mechanisms, extensive efforts have been made to use the system for delivery of therapeutics into biological compartments. The dimeric form of the receptor, where each subunit consists of three domains, further complicates the detailed investigation of molecular determinants responsible for guiding the receptor interactions with other proteins. Especially the apical domain’s biological function has been elusive. To further the study of transferrin receptor, we have computationally decoupled the apical domain for soluble expression, and validated the design strategy by structure determination. Besides presenting a methodology for solubilizing domains, the results will allow for study of apical domain’s function.

KEYWORDS
iron transport, protein design, Rosetta, transferrin, transferrin receptor

1 INTRODUCTION

Transferrin receptor 1 (TfR) together with its ligand, transferrin (Tf), supply cells with iron and are responsible for maintaining the physiological levels of iron in humans.1,2 The receptor is a homodimer with each subunit consisting of helical, protease-like, and apical domains.3 The dimeric interface is formed by the helical domain in one subunit, and the protease-like domain in the second subunit, with the functional interface responsible for binding of Tf situated at the helical domains. Upon complex formation, the Tf/TfR system gets internalized into cells via receptor mediated endocytosis. The interaction between the receptor and Tf is fine-tuned by another protein, known as hemochromatosis protein (HFE), thus keeping the intracellular concentration of iron at the physiologically relevant levels.4-6 Both the helical and protease-like domains consequently have biological relevance for either iron delivery or receptor structural cohesion. For the apical domain that does not seem to be the case, although it has been implied in binding and internalization of the iron carrier ferritin.7 The recently solved complex structure between the human ferritin and TfR8 will support the further investigation of the detailed molecular mechanism.

The experimentally determined structures of TfR have been reported both in apo form and complexed with different ligands.9,10,11 The central role of the TfR and Tf/HFE in iron regulation has led to the evolutionary pressure to conserve residues essential for the interface formation across the mammalian species. On the other hand, despite the lack of clear functional relevance of the apical domain, its surface has been exploited for binding by

© 2020 The Authors. Proteins: Structure, Function, and Bioinformatics published by Wiley Periodicals LLC.
exogenous opportunistic pathogens. The domain is utilized by viruses and parasites to enter into cells. Machupo virus, for example, interacts with glycoprotein 1 (MG1) to form a complex with the apical domain. Similarly, malarial parasite, Plasmodium vivax, forms the majority of its interactions with the apical domain in addition to interacting with the protease domain of the receptor. The evolutionary pressure for residue conservation in the apical domain is much lower; therefore, exogenous agents often exhibit a clear species preference.

The receptor's ability to internalize large protein molecules has found many biotechnological applications, where efforts have been made to transport small drugs, proteins, modified virus capsids, and nanocarriers across cell membranes. The small molecules or proteins are often coupled to different TfR binders and delivered into cells during endocytosis. The development of such carriers would largely benefit from an easy supply to Tf receptor and experimental assays. It would thus be highly beneficial if faster turn-around strategies existed to develop new protein binders to the receptor. In an effort to advance such experimental development and characterization, we have reengineered the apical domain of the TfR for soluble expression. Besides decoupling the domain from the receptor, we have applied computational methods to optimize the length of unstructured loops and to explore the amino acid sequence space to minimize nonspecific oligomerization properties of the rebuilt protein. We show that a computationally designed apical domain expresses recombinantly in Escherichia coli as soluble protein. In addition, the experimentally determined structure had well conserved structural properties at its interface forming region. The solubilized TfR domain is hence a valuable resource for further elucidation of biological function, and in development of TfR-based biotechnological applications.

2 | MATERIALS AND METHODS

2.1 | Computational rebuilding of Tf receptor apical domain

The apical domain of the Tf receptor starts at Gin$^{197}$ and finishes at Ser$^{378}$ as determined from the visual assessment of the experimental structure, PDB ID: 3kas. Residues Tyr$^{219}$ and Tyr$^{222}$, the large aromatic amino acids that constitute an important contribution to a hydrophobic patch interacting with the protease-like domain, were allowed to design to polar amino acids in an effort to prevent aggregation. The domain was additionally trimmed in-between residues Phe$^{297}$ and Ile$^{323}$ which form an unstructured, exposed loop interacting with the helical domain in the opposite unit of the homodimer. Additionally, Phe$^{297}$ was fixed to its native identity during the design process, and Phe$^{298}$ was allowed to design to polar residues. The latter one forms an extension of the hydrophobic surface together with the two abovementioned tyrosine residues Tyr$^{219}$ and Tyr$^{222}$. Positions at Ser$^{227}$ and Leu$^{229}$ were fixed to lysine amino acids each as they had been found to be solvent exposed after initial, loop-modeling trials. The total length of the rebuilt loop was 7 amino acids and was modeled with Rosetta software application RosettaRemodel protocol with talaris2013 energy function. After limited screening for loop conditions, final modeling consisted of 100 trajectories. The limited number of trials was sufficient due to the fixed amino acid and the short length of the loop. The resulting models were sorted by total energy, and the lowest energy models were visually inspected for optimal backbone conformations. Finally, the sequence from the next best model was chosen, which resulted in AP01 design.

Further modeling of the apical domain was carried out by rebuilding larger portions of the loop in-between the residues Met$^{283}$ and Ile$^{299}$, and designing only the last 3 amino acids before the C-terminal isoleucine residue. The loop was furthermore trimmed upstream of the three designed residues, one position at the time. The trajectories were analyzed by RMSD in comparison to the lowest energy model, where we sought after a converged folding funnel toward the lowest energy state. All modeling was carried out with a RosettaRemodel protocol using Ref2015 energy functions, but with 24 000 trajectories for each loop length. The increased number of trajectories was necessary to get sufficient sampling of energy landscapes. The best loop length with respect to the conformational energy landscape of the remodeled loop resulted in the design AP02. Both designs of the rebuilt apical domain, AP01 and AP02, were characterized experimentally.

2.2 | Expression and purification of AP01 and AP02

AP01 and AP02 were cloned in pET29b(+) at Ndel and XhoI restriction sites. NovaBlue competent cells were used for plasmid amplification; heat-shock protocol was used transformation with AP01/pET29b(+) plasmid. The gene insert was confirmed by colony PCR and DNA sequencing (Mix2Seq kit, Eurofins, Luxemburg). Tuner (DE3) competent cells (Invitrogen) were used for subsequent protein expression. Overnight culture of the confirmed clone was diluted in LB to OD$_{600}$ = 0.1, and grown at 37°C shaking at 180 rpm. When the optical density OD$_{600}$ reached 0.6, the temperature was lowered to 20°C, and the protein was induced by 0.2 mM IPTG for 18 hours. Cells were collected at 8000g for 10 minutes. Each gram cell pellet was dissolved in 5 mL 100 mM HEPES, 500 mM NaCl, pH 7.4, and sonicated four times at 40% amplitude for 20 seconds, with 20 seconds pause in between pulses using a Libra Cell 100 sonicator. Lysed cells were centrifuged at 15000g for 30 minutes. Supernatant was filtered through a 0.22 μm filter and loaded on a HisPur Cobalt resin column (Thermo Scientific), preequilibrated with 100 mM HEPES, 500 mM NaCl, pH 7.4. The column was washed with 100 mM HEPES, 500 mM NaCl, 10 mM imidazole, pH 7.4, and eluted with 100 mM HEPES, 500 mM NaCl, 500 mM imidazole, pH 7.4, according to the protocol. Elutes containing protein of the correct size, were confirmed with SDS-PAGE, concentrated, and further purified by size exclusion chromatography (SEC) over HiLoad 16/60 Sephacryl S100HR column. Eluted protein was stored in 20 mM HEPES, 150 mM NaCl, and 10% glycerol. AP01 was analyzed over an HPLC C3 column (Agilent ZORBAX 300 SB-C3, 4.6 × 50 mm, 5 μm) by MS using ESI accordingly to a protocol described previously.
2.3 | Structure determination

Additional purification of AP01 was carried out for protein preparations for structure determination. AP01 expressed as described above was purified by Ni-NTA affinity chromatography, and collected fractions were concentrated and further cleaned up by SEC (Superdex 10/300 column, GE Healthcare) using a buffer containing 20 mM Tris-Cl pH 8.0, 150 mM NaCl, 1 mM TCEP. The cleanest fractions, as determined by SDS–PAGE, were pooled and concentrated to 13 mg/mL (OD_{280}) for crystallization setups.

Crystallization drops were setup with commercial crystallization screens using the vapor diffusion method employing an Oryx 7 crystallization robot (Douglas Instruments Ltd.) and incubated at 289 K. The drop volume was 1 μL, with different ratios of protein and precipitant solution (1:1, 1:2, and 2:1). These setups resulted in diffraction quality crystals grown in 0.2 M ammonium fluoride and 20% v/v PEG 3350 (Molecular Dimensions SG-1 screen, condition 78). The obtained crystals were harvested from mother liquor with CryoLoops (Hampton Research), and flash frozen in liquid nitrogen without the use of any cryo-protecting agent. Diffraction data was collected at 100 K on beamline P11 at DESY, Hamburg (Germany). A full dataset (180°) was collected to 2.0 Å resolution from a monoclinic crystal (space group \( P_{12_1} \)). The collected data were processed using XDS\textsuperscript{23,24} with the provided input file from the beamline. Structure determination was performed by molecular replacement using PHASER\textsuperscript{25} with the design model of AP01 as a search template. The best solution was refined in reciprocal space with PHENIX\textsuperscript{26} with 10% of the data used for \( R_{\text{free}} \) and by real-space fitting steps against \( \sigma_{\text{A}} \)-weighted 2Fo–Fc and Fo–Fc electron density maps using COOT.\textsuperscript{27} Water molecules were placed automatically into difference electron maps, and accepted or rejected according to geometry criteria and their B-factors, defined in the refinement protocol. Noncrystallographic symmetry (NCS) restraints were applied during all refinement steps. The final model could be refined to \( R = 23.66\% \) and \( R_{\text{free}} = 27.01\% \).

No electron density was observed for the last 13 residues in both chains. Moreover, density of residues 25-27 and 97-99 in chain A as well as residues 102-107 in chain B was very poor. However, all residues were modeled to their NCS counterparts, which were visible. Three sodium ions were included into the model. Details of data collection, processing, and refinement are summarized in Table 1.

3 | RESULTS AND DISCUSSION

In this report, we present the design of solubilized apical domain derived from human Tf receptor (Figure 1A). Computational protein engineering was carried out to trim the domain, and to improve the protein solubility by redesigning unstructured loop regions. Furthermore, the hydrophobic patches on the surface, that resulted from decoupling from the receptor protease domain, have been optimized to minimize aggregation by designing them into polar amino acids. In total, two different designs have been built, AP01 and AP02, with the major difference being in the redesigned loop region stretching in-between residues Met\textsuperscript{283} and Ile\textsuperscript{329}. This loop region, besides contacting the protease domain, is naturally extending toward the opposite subunit of the TfR dimer. Both AP01 and AP02 were bacterially expressed, with AP01 being the more soluble design. To validate our design strategy and confirm the receptor remodeling, we determined the structure of AP01 by X-ray crystallography.

### Table 1 Data collection and refinement statistics for AP01 (PDB ID: 6y76)

| Wavelength (Å) | Resolution range | 30.7–1.985 (2.056–1.985) |
|----------------|------------------|--------------------------|
| Space group    | \( P_{12_1} \)   |                           |
| Unit cell      | 33.12119.558 33.167 | 90 112.249 90             |
| Total reflections | 53 762 (4426) |                           |
| Unique reflections | 15 946 (1375) |                           |
| Multiplicity   | 3.4 (3.2)       |                           |
| Completeness (%) | 96.61 (82.20) |                           |
| \(<1/\sigma_i\)> | 7.61 (2.38) |                           |
| Wilson B-factor | 21.33           |                           |
| \( R_{\text{merge}} \) | 0.1188 (0.395) |                           |
| \( R_{\text{max}} \) | 0.1414 (0.4714) |                           |
| \( R_{\text{p.i.m.}} \) | 0.07602 (0.2547) |                           |
| CC\textsubscript{1/2} | 0.991 (0.875) |                           |
| CC*            | 0.998 (0.966)   |                           |
| Reflections used in refinement | 15 936 (1372) |                           |
| Reflections used for \( R_{\text{free}} \) | 1595 (138) |                           |
| \( R_{\text{work}} \) | 0.2344 (0.2772) |                           |
| \( R_{\text{free}} \) | 0.2745 (0.3291) |                           |
| CC (work)      | 0.915 (0.864)   |                           |
| CC (free)      | 0.889 (0.700)   |                           |
| Number of non-hydrogen atoms | 2566 |                           |
| Macromolecules | 2359            |                           |
| Ligands        | 3               |                           |
| Solvent        | 204             |                           |
| Protein residues | 312             |                           |
| RMS (bonds)    | 0.004           |                           |
| RMS (angles)   | 0.82            |                           |
| Ramachandran favored (%) | 96.10 |                           |
| Ramachandran allowed (%) | 3.90 |                           |
| Ramachandran outliers (%) | 0.00 |                           |
| Rotamer outliers (%) | 1.14 |                           |
| Clashscore     | 5.87            |                           |
| Average B-factor | 24.73           |                           |
| Macromolecules | 24.35           |                           |
| Ligands        | 31.13           |                           |
| Solvent        | 29.07           |                           |

Note: Statistics for the highest-resolution shell are shown in parentheses.
3.1 | Computational design of AP01 and bacterial expression

The redesigned apical domain of the receptor was built with the RosettaRemodel application, which is part of the Rosetta protein engineering suite. The application allows for easy access to algorithms to vary protein backbone lengths, in addition to protein sequence optimization (Figure 1B). To decouple the apical domain, we had to carry out loop remodeling in-between residues Phe297 and Ile323 thus removing a large unstructured segment extending toward the helical domain of the receptor. Besides loop remodeling, an aromatic patch consisting of two tyrosine residues and a phenylalanine were optimized to polar amino acids serine, threonine, and asparagine. In the native protein, these aromatics are interlaced into the opposite domain on the receptor, resembling hot-spot residues of protein–protein interfaces. The final, designed protein AP01 thus contained both changes in backbone length as well as design of specific sequence positions. To validate the computational design, AP01 was bacterially expressed and purified by metal affinity and SEC (Figure 2). The major peak eluted at 69 mL according to the size exclusion chromatogram, which corresponds to ~16 000 Da according to the calibration, indicating soluble protein in a monomeric state in the solution. The molecular weight was approximated by SDS–PAGE to be around 18 000 Da which was confirmed by mass spectrometry (determined molecular weight of 18 026 corresponds to the theoretically calculated molecular weight of 18 027 Da).

3.2 | Assessment of the AP01 experimental structure

To evaluate our design strategy and confirm the structural cohesion of the AP01, we experimentally determined its structure. The overall structure of the AP01 recapitulated the core of the apical domain well with C\(_\alpha\) RMSD of ~1.2 Å (calculated over 134 residues not part of the restructured loop; Figure 3A). The AP01 design did contain two major differences in comparison to the computational model. First, there is a rearrangement of the N-terminal segment that is in direct contact with virus glycoprotein known to bind to the TfR apical domain (Figure 3B). Upon binding, the glycoprotein induces a conformational change in the segment allowing it to form hydrogen bonding interactions. This type of conformational bias upon binding has been seen in the apical domain, and was explained in detail previously.\(^\text{12}\) Second, the change in the rebuilt backbone is more relevant from the protein design perspective, as it directly affects our computational strategy (Figure 3C). The modeled backbone of AP01 contained a stretch of 7 amino acids, which simply connected the backbone to eliminate the extended loop interacting with the opposite unit of the receptor homodimer. Despite the limited length of the rebuilt loop, we find in the structure a large rearrangement of the backbone that spans from Met283 to Ile329 (Figure 3C). This backbone segment forms unstructured loop that directly contacts the protease domain. The lack of the adjacent surface most likely destabilizes the loop sufficiently, so it adopts a new conformation seen in the AP01 structure. We cannot, however, exclude that the crystal packing may be partially responsible for the stabilization of the loop in the observed conformation.

3.3 | Redesigning AP01 to optimize the energy landscape of the rearranged loop

It has been shown for other systems that proteins may have alternate conformational states as identified by Rosetta methodology,\(^\text{28,29}\) similarly to what has been employed here. Since the AP01 showed conformational flexibility in the loop adjacent to the rebuilt backbone...
FIGURE 2  Remodeled apical domain of the transferrin receptor, AP01, expresses solubly in bacteria. A. Size exclusion chromatography (SEC) identified the apical domain monomer, AP01, at the correct elution volume ~70 mL (blue line); AP02 eluted at a much lower peak intensity at approximately the same elution volume (orange line). B, SDS-PAGE protein gel of the highest peak fractions from AP01 SEC, shows protein of about 18 kDa. C, Mass spectrometry determined AP01 protein envelope. D, Mass spectrometry derived AP01 mass at 18026 Da, which agrees with the theoretical mass of 18 027 Da [Color figure can be viewed at wileyonlinelibrary.com]

FIGURE 3  Experimentally determined structure of AP01 shows preserved core in comparison with the design. A, Overlay of the AP01 experimental structure and the apical domain of PDB ID: 3kas. The large loop protrudes out of the apical domain and is known to interact with the opposite subunit of the receptor homodimer. B, Rearrangement of the N-terminal segment known to be involved in binding of the viral glycoproteins. C, A large loop rearrangement occurs in the experimental structure in comparison to the designed conformation. Experimental structure in gray; designed model in yellow [Color figure can be viewed at wileyonlinelibrary.com]
similarly to alternate states, it was rebuilt by including rearranged positions spanning 21 residues in-between Met^{283} to Ile^{329} (PDB ID: 3kas numbering, Figure 4). The modeled loop was sampled for different backbone conformations from the loop fragment library. The rebuilt 21 residues were furthermore trimmed down one position at the time until converged solutions were found, as assessed by well-

**FIGURE 4** Computational optimization of the loop indicated a folding funnel at 12 residues length. The reconnected loop in AP01 showed different conformations in the experimental structure and the design. To optimize the loop length, it was shortened one position at the time and rebuilt from backbone fragments. The loop energetics was evaluated by folding funnels as described by model energy as a function of RMSD when compared to the lowest energy conformation. The initial loop showed large flexibility that finally converged to an energy landscape resembling a sharp folding funnel at loop length of 12 residues. Additional trimming resulted in less defined energy landscape with fewer solutions (loop length of 11). The 12 residues loop was also shown to be sensitive to amino acid changes which allowed for design; loop length = 12 (1) and 12 (2) describe two different loop design strategies. Loop sequence is defined by the box insert, where X denotes a position allowed to design to any amino acid [Color figure can be viewed at wileyonlinelibrary.com]
defined folding funnels. In this way, the loop flexibility will eventually decrease due to the restrictions in available degrees of freedom. The lower limit that was found to be able to connect the loop backbones was as low as 11 amino acids during the rebuilding process. In addition, to allow for limited design process, three residue positions C-terminally in the loop (introduced in AP01) were allowed to optimize its amino acid sequence, while the rest of the loop was held to the native identities. The conformational sampling indicates that the energetic landscape of the loop, when sampled around the native length, is rather flat; trimming the loop down restricts the possible backbone conformations resulting in fewer solutions. Interestingly, the conformational landscape focuses to a folding funnel at the loop length equal to 12 positions. Further trimming of the loop leads to fewer solutions as well as to a less defined folding energy landscape, for example, loop length equal to 11 in Figure 4.

Having identified the loop length of 12 residues as the most favorable solution, additional exploration of the energy landscape was carried out where we fixed all residues to their designed sequence, but for 1 or 2 positions (Figure 4). In one case of the designed loop, the first position (DQTKFPIVNPNTN), which is an aspartic acid, was allowed to change (loop denoted as 12 (1) in Figure 4), while in the other case both first and fourth positions were allowed to redesign, that is, lysine in addition to the aspartic residue (loop denoted as 12 (2) in Figure 4). Backbone remodeling with this strategy gave only satisfactory solutions in the first case, where a mutation to valine showed the best energetics. This designed sequence, called AP02, was chosen for experimental characterization. Compared to the AP01 design, soluble protein for AP02 was obtained only in lower quantities (Figure 1A).

3.4 | Effect of the decoupled apical domain on interactions with binding partners

Apical domain of TfR is known to interact with viral glycoproteins and malaria parasite reticulocyte binding protein, for which there are available complexed experimental structures. To investigate the impact of backbone rearrangement on binding, we overlaid the AP01 on the apical domain in PDB ID: 3kas12 and PDB ID: 6d0413 structures, which contain the receptor bound to Machupo virus glycoprotein (MGP1) and reticulate binding protein (RBP2b), respectively. Overall, the interface structure is not significantly affected by the conformational change of the loop observed in AP01. The noticeable effect of the rearrangement is the loss of a salt bridge between a glutamate residue in the receptor and a lysine in the binding partner; the salt-bridge is actually formed in both complexes by the same glutamate, Glu294, with the lysine residues MGP1–Lys169 and RBP2b–Lys600 (Figure 5). Furthermore, both lysines occupy approximately the same space with

![Figure 5](image_url)

**Figure 5** The loop rearrangement in AP01 leads to a loss of salt-bridge interactions in interacting protein partners of TfR apical domain. A, TfR Glu294 interacting with MGP1 Lys169 (PDB ID: 3kas). B, Computational redesign leads to a loss of the native salt-bridge interaction were the loop rearrangement leads to about 27 Å difference between the sidechains of AP01 corresponding residue Glu99 and MGP1 Lys169. C, TfR Glu294 interacting with RBP2b Lys600 (PDB ID: 6d04). D, Salt-bridge interactions is similarly lost between the AP01 Glu99 and RBP2b Lys600. For panels C and D, the AP01 structure was overlaid on the apical domain of the receptor in PDB ID: 3kas and PDB ID: 6d04, respectively, before the distance measurements were carried out. Lysine residues occupy approximately the same space with backbone Cα atoms being only 1 Å apart after superimposing the receptor in complexed structures PDB ID: 3kas and PDB ID: 6d04. AP01 or TfR in gray; MGP1 or RBP2b in magenta.
backbone Cα atoms being only 1 Å apart, indicating that there is an evolutionary pressure on TR to conserve Glu294 site, although there is no clear evidence for it from known endogenous binding partners. Studies have reported that the iron carrier complex ferritin may bind to the receptor apical domain, but detailed experimental analysis is lacking.

Additionally, the bacterially expressed apical domain is missing posttranslational glycosylations. The known site, Asn271, is on the opposite side of the interacting surfaces, and therefore not involved in the binding of neither MGP1 nor RBPb2. Glycosylation on Asn217 extends, together with unstructured native loop, toward the other TR subunit in the dimer; it seems more important thus for TR structural stability. The absence of the glycans on the AP01 should most likely not impact its functions or further optimization.

4 | CONCLUSIONS

We have solubilized TR apical domain for bacterial expression by decoupling it from the receptor by two different computational strategies. Initially, minimal design was carried out to remove hot spot residues to the protease domain, and to reconnect the unstructured loop that makes contacts with the opposite subunit of TR homodimer. For the computationally designed protein, AP01, an X-ray crystallographic structure was determined. The structure is in great agreement with the design model for all parts except for a conformational change in the loop that is adjacent to the protease domain. Further loop length optimization resulted in the second design, AP02, which was trimmed by 9 positions. It was also soluble, but seemed more prone to aggregation. When compared to the known interaction partners of TR apical domain, AP01 retained most of the structural determinants responsible for the complex formation, except for the salt bridge between Glu294 and the lysine residue in the binding partner. The designed proteins, AP01 and AP02, are thus relevant starting points for further binding optimization, for example, toward pathogenic epitopes in an effort to devise future protein therapeutics.

ACCESSION NUMBER

Coordinates and structure factors have been deposited in the Protein Data Bank with accession number 6y76.

ACKNOWLEDGMENTS

Sinisa Bjelic is grateful for financial support from FHL and Vinnova (2015-04912), for the work Anneli Lundgren carried out during initial testing of AP01 expression and purification, and Dr. Panagiota Georgoula for performing AP01 LC–MS analysis. GO is supported by an ERC-StG (Acronym: HelixMold, Grant-Agreement number: 802217) and the Austrian Science Fund (FWF) project (Project number P30826). The computations were enabled by resources provided by SNIC at Lunarc partially funded by the Swedish Research Council through grant agreement no. SNIC 2019/3-320.

CONFLICT OF INTEREST

The authors declared no potential conflicts of interest.

AUTHOR CONTRIBUTIONS

Dick J. Sjöström: designed the study, characterized AP01 and AP02, and wrote the manuscript. Sarah A. Berger: determined the experimental structure. Gustav Oberdorfer: designed the study and wrote the manuscript. Sinisa Bjelic: designed the study and wrote the manuscript.

ORCID

Sinisa Bjelic https://orcid.org/0000-0002-9300-614X

REFERENCES

1. Andrews NC, Schmidt PJ. Iron homeostasis. Annu Rev Physiol. 2007;69:69-85.
2. Eckeroth BE, Steere AN, Chasteen ND, Everse SJ, Mason AB. How the binding of human transferrin primes the transferrin receptor potentiating iron release at endosomal pH. Proc Natl Acad Sci USA. 2011;108(32):13089-13094.
3. Lawrence CM, Ray S, Babaynovshev M, Gallowe R, Borhani DW, Harrison SC. Crystal structure of the ectodomain of human transferrin receptor. Science. 1999;286(5440):779-782.
4. Lebron JA, Bennett MJ, Vaughn DE, et al. Crystal structure of the hemochromatosis protein HFE and characterization of its interaction with transferrin receptor. Cell. 1998;93(1):111-123.
5. West AP Jr, Giannetti AM, Herr AB, et al. Mutational analysis of the transferrin receptor reveals overlapping HFE and transferrin binding sites. J Mol Biol. 2001;313(2):385-397.
6. Giannetti AM, Bjorkman PJ. HFE and transferrin directly compete for transferrin receptor in solution and at the cell surface. J Biol Chem. 2004;279(24):25866-25875.
7. Li L, Fang CJ, Ryan JC, et al. Binding and uptake of H-ferritin are mediated by human transferrin receptor-1. Proc Natl Acad Sci USA. 2010;107(8):3505-3510.
8. Montemiglio LC, Testi C, Ceci P, et al. Cryo-EM structure of the human ferritin-transferrin receptor 1 complex. Nat Commun. 2019;10(1):1121.
9. Testi C, Boffi A, Montemiglio LC. Structural analysis of the transferrin receptor multifaceted ligand(s) interface. Biophys Chem. 2019;254:106242.
10. Bennett MJ, Lebron JA, Bjorkman PJ. Crystal structure of the hereditary haemochromatosis protein HFE complexed with transferrin receptor. Nature. 2000;403(6765):46-53.
11. Cheng Y, Zak O, Aisen P, Harrison SC, Walz T. Structure of the human transferrin receptor-transferrin complex. Cell. 2004;116(4):565-576.
12. Abraham J, Corbett KD, Farzan M, Choe H, Harrison SC. Structural basis for receptor recognition by New World hemorrhagic fever arenaviruses. Nat Struct Mol Biol. 2010;17(4):438-444.
13. Gruszczyn J, Huang RK, Chan LJ, et al. Cryo-EM structure of an essential Plasmodium vivax invasion complex. Nature. 2018;559(7712):135-139.
14. Johnsen KB, Bak M, Kempen PJ, et al. Antibody affinity and valency impact brain uptake of transferrin receptor-targeted gold nanoparticles. Theranostics. 2018;8(12):3416-3436.
15. Lee JH, Engler JA, Collawn JF, Moore BA. Receptor mediated uptake of peptides that bind the human transferrin receptor. Eur J Biochem. 2001;268(7):2004-2012.
16. Patel MM, Patel BM. Crossing the blood-brain barrier: recent advances in drug delivery to the brain. CNS Drugs. 2017;31(2):109-133.
17. Wiley DT, Webster P, Gale A, Davis ME. Transcytosis and brain uptake of transferrin-containing nanoparticles by tuning avidity to transferrin receptor. Proc Natl Acad Sci USA. 2013;110(21):8662-8667.

18. Yu YJ, Atwal JK, Zhang Y, et al. Therapeutic bispecific antibodies cross the blood-brain barrier in nonhuman primates. Sci Transl Med. 2014;6(261):261ra154.

19. Leaver-Fay A, Tyka M, Lewis SM, et al. ROSETTA3: an object-oriented software suite for the simulation and design of macromolecules. Methods Enzymol. 2011;487:545-574.

20. Huang PS, Ban YE, Richter F, et al. RosettaRemodel: a generalized framework for flexible backbone protein design. PLoS One. 2011;6(8):e24109.

21. Alford RF, Leaver-Fay A, Jeliazkov JR, et al. The Rosetta all-atom energy function for macromolecular modeling and design. J Chem Theory Comput. 2017;13(6):3031-3048.

22. Bjelic S, Nivon LG, Celebi-Olcum N, et al. Computational design of enone-binding proteins with catalytic activity for the Morita-Baylis-Hillman reaction. ACS Chem Biol. 2013;8(4):749-757.

23. Diederichs K, Karplus PA. Improved R-factors for diffraction data analysis in macromolecular crystallography. Nat Struct Biol. 1997;4(4):269-275.

24. Kabsch W. Xds. Acta Crystallogr D Biol Crystallogr. 2010;66(Pt 2):125-132.

25. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. J Appl Cryst. 2007;40(Pt 4):658-674.

26. Adams PD, Afonine PV, Bunkoczi G, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr. 2010;66(Pt 2):213-221.

27. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr. 2010;66(Pt 4):486-501.

28. Andre I, Bjelic S. Computational assessment of folding energy landscapes in heterodimeric coiled coils. Proteins. 2018;86(7):790-801.

29. Tyka MD, Keedy DA, Andre I, et al. Alternate states of proteins revealed by detailed energy landscape mapping. J Mol Biol. 2011;405(2):607-618.

How to cite this article: Sjöström DJ, Berger SA, Oberdorfer G, Bjelic S. Computational backbone design enables soluble engineering of transferrin receptor apical domain. Proteins. 2020;88:1569–1577. https://doi.org/10.1002/prot.25974