1 **Triggering closure of a sialic acid TRAP transporter substrate binding** 2 **protein through binding of natural or artificial substrates**

3

4 Martin F. Peter¹, Christian Gebhardt², Janin Glaenger¹, Niels Schneberger¹, Marijn de
5 Boer³, Gavin H. Thomas⁴, Thorben Cordes^{2,3}, Gregor Hagelueken^{1,*}

6

7 ¹Institute of Structural Biology, University of Bonn, 53127 Bonn, Germany

8

9 ²Physical and Synthetic Biology, Faculty of Biology, Ludwig-Maximilians-Universität
10 München, 82152 Planegg-Martinsried, Germany

11

12 ³Molecular Microscopy Research Group, Zernike Institute for Advanced Materials,
13 University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands

14

15 ⁴Department of Biology (Area 10), University of York, York YO10 5YW, UK

16

17

18 *Corresponding author

19 E-mail: hagelueken@uni-bonn.de

20

Abstract

The pathogens *Vibrio cholerae* and *Haemophilus influenzae* use tripartite ATP-independent periplasmic transporters (TRAPs) to scavenge sialic acid from host tissues. They use it as a nutrient or to evade the innate immune system by sialylating surface lipopolysaccharides. An essential component of TRAP transporters is a periplasmic substrate binding protein (SBP). Without substrate, the SBP has been proposed to rest in an open-state, which is not recognised by the transporter. Substrate binding induces a conformational change of the SBP and it is thought that this closed state is recognised by the transporter, triggering substrate translocation. Here we use real time single molecule FRET experiments and crystallography to investigate the open- to closed-state transition of VcSiaP, the SBP of the sialic acid TRAP transporter from *V. cholerae*. We show that the conformational switching of VcSiaP is strictly substrate induced, confirming an important aspect of the proposed transport mechanism. Two new crystal structures of VcSiaP provide insights into the closing mechanism. While the first structure contains the natural ligand, sialic acid, the second structure contains an artificial peptide in the sialic acid binding site. Together, the two structures suggest that the ligand itself stabilises the closed state and that SBP closure is triggered by physically bridging the gap between the two lobes of the SBP. Finally, we demonstrate that the affinity for the artificial peptide substrate can be substantially increased by varying its amino acid sequence and by this, serve as a starting point for the development of peptide-based inhibitors of TRAP transporters.

Keywords

Membrane transporter, TRAP, smFRET, crystal structure, integrative structural biology

Abbreviations

SBP: Substrate binding protein

TRAP: tripartite ATP-independent periplasmic

Introduction

The tripartite-ATP independent-periplasmic (TRAP) transporters are a large class of membrane transporters that are widespread in bacteria and archaea. They have unique properties in being a structural and functional mix of ATP-binding cassette (ABC) transporters and secondary-active transporters, both of which are well studied. TRAPs translocate a large variety of small substrates, most of which contain a carboxylic acid group^{1–3}. They rely on a periplasmic high-affinity substrate binding protein (SBP, in TRAP transporters also termed P-domain) that is assumed to scavenge and deliver substrate molecules to the actual transporter domains in the inner membrane^{4, 5}. For TRAP transporters, the latter are known as Q- and M-domains and consist of 4 and 12 transmembrane (TM) helices, respectively. In some representatives, the Q- and M-domains are fused into one peptide chain by an additional TM-helix⁵.

The best studied TRAP transporters are currently the *N*-acetylneuraminic acid (commonly known as sialic acid or Neu5Ac) transporters HiSiaPQM and VcSiaPQM from the two human pathogens *Haemophilus influenzae* and *Vibrio cholerae*^{4, 5}. For these two pathogens, the TRAP transporter is the only uptake route for the exogenously synthesized sialic acid, which is present in host tissues. After uptake, the small sugar molecule is used to sialylate the bacterial lipopolysaccharides as a strategy to evade detection by the hosts innate immune system^{6–8}. To date, the structures of the transmembrane domains are unknown, but, high-resolution crystal structures of the open form of the P-domains of HiSiaP and VcSiaP and of the closed form of HiSiaP have been determined^{9–11}.

Like other SBPs¹², the TRAP transporter P-domains have a characteristic bilobal fold with a large substrate binding cleft. The two lobes are connected by a long α -helix, which forms the backbone of the structure. Substrate binding induces large conformational changes of the whole protein by more than 15 Å and the formation of a characteristic kink in the backbone helix⁹. The closing motion and the binding mechanism of the SBP have often been compared to a Venus flytrap^{12, 13}. A PELDOR spectroscopy study (pulsed electron double resonance, also known as DEER) on spin labelled VcSiaP indicated that in frozen solution and without substrate, the P-domain “lurks” in the open state, similar to some ABC transporter SBPs^{14–20}. The substrate binding cleft has two highly conserved arginine residues (Arg125 and Arg145 in VcSiaP), which coordinate the carboxylic acid group of the substrate^{10, 11, 21}. In addition, an intricate water network around the bound substrate is an important determinant of substrate affinity²².

1 Still, numerous questions concerning the role of the P-domains for TRAP transporters
2 and concerning SBPs in general remain unanswered. Most importantly, it is still unclear
3 how the sialic acid triggers the transition to the closed state. Considering that TRAP SBPs
4 are potential drug targets, it is of high interest to find the requirements and specifics of a
5 substrate that can initiate the transition.

6 Here, we use a combination of single-molecule FRET spectroscopy and X-ray
7 crystallography to study the conformational changes of single VcSiaP molecules in real
8 time and to investigate the molecular details of the closing mechanism.

9

Results

Open-closed dynamics of VcSiaP at room temperature

Binding of the substrate to the periplasmic SBP is the first step of the proposed mechanism for TRAP transporters and involves large scale conformational changes of the protein⁵. PELDOR/DEER experiments on frozen VcSiaP solutions indicated that the protein switches its conformation in a substrate dependent manner¹⁸. Here, we used single-molecule Förster-resonance energy transfer (smFRET) spectroscopy to investigate this important step of the mechanism by real-time measurements of VcSiaP dynamics under more physiological conditions. Such experiments should allow comparisons to SBPs from the ABC transporter family^{14, 17}. For this purpose, we employed the Q54C/L173C double mutant of VcSiaP that was used for the PELDOR/DEER experiments¹⁸. The mutant protein was labelled with maleimide-derivatives of AlexaFluor 555 (donor) and AlexaFluor 647 (acceptor). The distance between both label positions was expected to decrease upon ligand binding, i.e., the transition from the open- to closed state should lead to a change from intermediate to high FRET-efficiency. On average we obtained a labelling efficiency > 90% with ~1.83 fluorophores per protein (0.68 donors and 1.15 acceptors) for Alexa555/647 and a resulting yield of donor-acceptor containing VcSiaP of ~49% (with 10% donor only and ~41% acceptor only molecules).

With this at hand, we studied the substrate dependence of the conformational changes in VcSiaP (Q54C/L173C) using solution-based FRET. This allowed us to simultaneously check whether the addition of the fluorescence label had any impact on the affinity to Neu5Ac and which conformational states of VcSiaP are seen in solution at room temperature. The relative populations of the open- and closed-conformation were determined in a titration experiment for different substrate concentrations, allowing us to determine a K_D value of 300 ± 100 nM (Fig 1A-C). This value was in very good agreement with a previously published ITC experiment¹¹ and a control ITC experiment for our particular wild-type construct (Figure 1E). Thus, the presence of the fluorescence labels does not influence the Neu5Ac affinity and the conformational changes in VcSiaP are directly coupled to the substrate binding to the protein (Figure 1DE).

To derive the binding mechanism of VcSiaP, we performed smFRET experiments with surface-immobilized protein via scanning confocal microscopy. For this purpose, we attached VcSiaP to a glass-surface using its N-terminal His₆-tag (Figure 2A). We analysed the conformational dynamics between open- and closed-state in real time using

1 apo conditions (no ligand), 500 nM Neu5Ac ($\sim K_D$) and 1 mM Neu5Ac (saturated ligand
2 concentration). The FRET-states observed of VcSiaP (Q54C/L173C) at apo and saturated
3 ligand conditions are fully compatible with diffusion-based experiments (Figure 3) and
4 show the population of the open (absence of substrate) and closed (1 mM Neu5Ac) state
5 (Figure 2BC). For concentrations close to the K_D -value we observed discrete and
6 stochastic transitions between low- and high-FRET efficiency states (Figure 2B, middle
7 panel). The low and high FRET states have an average apparent FRET efficiency of 0.791
8 ± 0.012 (95% confidence interval) and 0.895 ± 0.010 , respectively. They are similar to
9 the average efficiencies in the absence (0.807 ± 0.003) and presence of saturating (0.898
10 ± 0.003) concentrations of Neu5Ac (Figure 2BC). Therefore, the transitions can be
11 interpreted directly as transitions between the open (low E^*) and closed (high E^*)
12 conformation.

13 To characterize the kinetics of the conformational transitions, we modelled the
14 underlying state-trajectory of the fluorescence trajectories by a Hidden Markov Model
15 (HMM ²³) with two states, open and closed (Figure 2B; middle panel). From the state
16 assignment of the HMM the individual lifetimes of the closed state were extracted. The
17 distribution of these lifetimes was fitted with an exponential distribution and has an
18 average lifetime of 122 ± 16 ms (95% confidence interval) (Figure 2D). Thus, once
19 VcSiaP is in the closed conformation, it takes on average around 120 ms to re-open and
20 release the ligand. Within the detection limit of smFRET and a total duration of almost
21 20 min recording time (850 fluorescence trajectories, i.e., 700 apo and 150 holo state),
22 we observed that VcSiaP is exclusively in the open state when the ligand is absent and
23 only closes when the ligand is bound (Figure 2B). Such a behavior would be compatible
24 with an induced-fit (Venus flytrap) mechanism of ligand binding assuming that
25 fluctuations between both conformations are not faster than the 5 ms time resolution we
26 could achieve in the presented experiments. If conformational sampling was indeed
27 (much) faster than this, other binding mechanism are still possible, where open-to-closed
28 state transitions are not only driven by ligand.

29

30 **Crystal structure of substrate-bound VcSiaP**

31 To create an atomic model of the closed state of VcSiaP that was observed in the smFRET
32 experiments, we co-crystallised the protein with Neu5Ac. Single crystals were observed
33 after three months of incubation at 20 °C. The crystal structure was solved at 1.6 Å
34 resolution by molecular replacement ²⁴, using the closed state of HiSiaP as the search

1 model (PDB-ID: 3B50¹⁰). The electron density map clearly revealed a Neu5Ac molecule
 2 in the binding pocket, tightly bound by residues R125, R145 and N185 (Figure 3BC). The
 3 structure contains residues 1-299 of the mature protein (i.e. after cleavage of the signal
 4 peptide, Figure 3A) and was refined to R/R_{free} values of 0.221/0.257 (Table 1). While the
 5 main interactions between protein and substrate are similar to other TRAP SBPs, subtle
 6 differences in the water network around the substrate were detected (SI Figure 1). These
 7 differences might explain the comparably low affinity of VcSiaP towards Neu5Ac¹¹.
 8 Overall, the structure is similar to the HiSiaP/Neu5Ac complex with an r. m. s. d. value
 9 of 0.7 Å for 279 Ca atoms. Still, the fact that the HiSiaQM membrane transporter
 10 discriminates between VcSiaP and HiSiaP *in vitro*²⁵ implies that important structural
 11 differences exist. We hence used the ENDscript server²⁶ to identify conserved areas on
 12 the surface of the closed state structure. The result is shown in Figure 3D (see SI for the
 13 underlying alignment). Clearly, residues surrounding the closed substrate binding cleft
 14 are strongly conserved. Considering the notion that it is the closed state of the SBP which
 15 is most likely recognised by the QM domains⁵, this conserved and continuous surface
 16 region is likely involved in the P-QM interaction, as proposed for the YiaMNO TRAP
 17 transporter²⁷.

18 Within this region, there was no obvious candidate that immediately explains the
 19 described discrimination between VcSiaP and HiSiaP, but, we noticed a surface patch
 20 with a number of reversed charges (D25, E28, K21, K36 in VcSiaP corresponding to K26,
 21 K29, E22, E37 in HiSiaP). This feature represents an interesting candidate for further
 22 experiments addressing this question (Figure 3D). The magenta lines in Figure 3E
 23 illustrate the spatial relation between conserved residue pairs of SBPs mapped onto the
 24 closed state structure of VcSiaP²⁸. Strikingly, although the two lobes clearly touch each
 25 other in the closed state (Figure 3D), there is only one weak evolutionary coupling of
 26 residues across the gap between the two lobes. Hence, it appears that the substrate is the
 27 “glue” that holds the two lobes of the SBP together, fitting nicely to the smFRET results
 28 above.

29 **A 9-mer peptide can artificially trigger SBP closing**

30 From a structural and mechanistic perspective, it remains poorly understood how the large
 31 conformational change from the opened- to the closed state conformation is triggered. It
 32 has been proposed that a conserved triad of residues (R125, E184 and H207, VcSiaP
 33 numbering) in the hinge region is involved¹¹. But, our previous experiments have shown
 34

1 that these residues have no influence on the overall conformational change.¹⁸ Mutants of
 2 these residues had a reduced affinity for Neu5Ac but were still able to adopt the closed-
 3 state structure. A serendipitous hint towards an answer to this question was obtained
 4 during our efforts to crystallize various substrate binding site mutants of VcSiaP with
 5 attached EPR nitroxide spin labels (R1). The crystal structure of one of them, a R125A
 6 mutant, was solved at a resolution of 2.1 Å by molecular replacement using the VcSiaP
 7 wild-type as search model (PDB-ID: 4MAG,¹¹) and refined to R/R_{free} values of
 8 0.199/0.232 (Table 1). Inspection of the initial electron density map surprisingly revealed
 9 a peptide in the substrate binding site (SI Figure 2). Also, the R125A mutation was clearly
 10 visible and the two R1 spin labels at positions 54 and 173 (far away from the peptide: SI
 11 Figure 3) were well defined in the electron density (SI Figure 2).

12 During refinement, it became clear that the peptide is connected to a neighboring
 13 molecule in the crystal and is in fact a part of the latter molecules' affinity tag. Thus, two
 14 neighboring SBPs form an interlocked dimer in the hexagonal crystal packing (Figure
 15 4A). The "tag peptide" is involved in many interactions with the SBP domain, especially
 16 at the entrance of the Neu5Ac binding cleft (Figure 4B). The N-terminal His₆-tag itself is
 17 not visible in the electron density and therefore apparently disordered. Interestingly,
 18 position D₋₁₄ of the tag peptide interacts with R145 of the SBP. The latter is slightly
 19 disordered in our structure, but all known substrates of TRAP transporters interact with
 20 this residue²¹. Further, there are several interactions between Y₋₁₅ of the tag peptide and
 21 R69 of the SBP, T₋₁₀ of the tag peptide and S42 of the SBP as well as multiple
 22 hydrophobic interactions (Figure 4B). As mentioned above, the crystal structure is of the
 23 R125A Q54R1 L173R1 mutant of VcSiaP. Modeling the arginine residue back into the
 24 structure shows that its presence would not interfere with the tag peptide binding site and
 25 DLS measurements show a similar behavior in dimerization between the tagged wildtype
 26 and R125A mutant (SI Figure 3). We used difference distance matrices to analyze the
 27 conformational change induced by peptide binding and compared them to the changes
 28 induced by Neu5Ac binding (Figure 4C). This comparison reveals clear signs of closing
 29 of the peptide-bound SBP. Although the scale of the two half-matrices is different, their
 30 overall pattern is strikingly similar. Thus, the open to Neu5Ac-bound movement and the
 31 open to peptide-bound movement are closely related with the same rigid bodies moving
 32 relative to each other, albeit on different length scales. It appears that the peptide, a
 33 completely artificial substrate, has triggered closing and its presence has trapped a
 34 previously undescribed intermediate state of the SBP.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34

Characterization and strengthening of the peptide-SBP interaction

The discovery that a peptide can bind to the SBP opens an interesting road towards the design of a peptide-based inhibitor for TRAP transporter SBPs. As a first step into this direction, we designed a peptide-scan, i.e. a cellulose membrane with a set of 9-mer peptides attached via spot synthesis²⁹. Binding of the SBP to these short 9-mer peptides was detected by a fluorescence labelled SBP, where the label was attached to an artificially introduced cysteine on the “backside” of the protein (Q245C-DyLight800). The results of these experiments are shown in Figure 5A.

Since the tagged SBP is a monomer in gel filtration experiments, we knew that the interaction seen in the crystal lattice is rather weak and the peptide sequence needed to be optimized to see the interaction in solution. Indeed, no fluorescence signal was observed at the position of the tag peptide on the cellulose membrane (Figure 5A, peptide A1). Similarly, negative results were obtained for variations of the tag peptide sequence (not shown). However, inspired by the “scoop-loop” theory for ABC transporters³⁰, we tentatively modeled various periplasmic loops of the QM-domains into the substrate binding cleft by hand and found that the loop connecting TMs 1 and 2 of the Q-domains provided a surprisingly good fit to the binding pocket (SI Figure 4). While this peptide did not bind to the SBP in our experiment (Figure 5A, peptide A2), several variations that we had included in the screen did, resulting in a striking increase in binding strength. For example, simply introducing a positive charge C-terminal of the central double glutamate led to a drastic increase in affinity (peptides A7 vs A8, peptides B2 vs B3 and peptides B6 vs B7 vs B8, magenta 1 in Figure 5AB). In terms of structure, this makes sense, because the introduced lysine residue can potentially interact with E16 deep in the substrate binding pocket (SI Figure 4). The binding could be further strengthened by introducing a second positive charge (peptides B10 and B11 in Figure 6) that can potentially interact with E66 in the SBP. The hydrophobic C-terminal LLF sequence is also important for binding (peptide A9 vs peptide C9, magenta 2 in Figure 5AB). In the crystal structure, this part of the peptide is in contact with a hydrophobic patch on the surface of the SBP. The N-terminal sequence PSLW (peptide A9 vs C10, magenta 3) and the arginine at position 10 of the peptide are also important (peptides B9 vs A9, magenta 4). The latter two features can however not be rationalized by the structural model of the

1 peptide, indicating that the conformation of the peptide in the SBP might change
2 depending on the sequence. To get a rough impression about the strength of the
3 interaction, we performed a competition experiment with 1 mM Neu5Ac using an
4 identical peptide membrane. However, both experiments showed similar or slightly lower
5 fluorescence intensities. Thus, Neu5Ac had an influence on the interaction but due to the
6 very high local peptide concentrations on the membrane, it is difficult to estimate a K_D
7 value from such experiments.

8 Therefore, we further substantiated the peptide binding in solution with dynamic light
9 scattering (DLS) experiments. For this purpose, the sequence connecting the His₆-tag of
10 VcSiaP was replaced with the sequence of the strong binding peptide identified in the
11 peptide-scan (peptide B10, Figure 5AB). We then performed a series of DLS experiments
12 with different concentrations of both types of tagged (HisTEV and peptide B10) and
13 untagged VcSiaP protein (Figure 5C) and in competition with the natural substrate sialic
14 acid. For the untagged protein, a nearly constant hydrodynamic radius was measured for
15 all different concentrations and, as expected, the addition of sialic acid did not
16 significantly change the observed hydrodynamic radius. Overall, the radii for the
17 untagged samples corresponded reasonably well to the monomeric radius calculated from
18 the crystal structure of the open VcSiaP conformation (Figure 5C). In contrast, for both
19 tagged proteins, we observed a clear dependence of the hydrodynamic radii on the protein
20 concentration (Figure 5CD). At low concentrations, the radii were close to that of the
21 untagged protein, while for higher concentrations the radii increased significantly and
22 converged close to the expected value for the dimeric crystal structure. In line with the
23 peptide-scan experiment, the “optimized tag” led to a stronger increase than the original
24 tag. For both constructs the hydrodynamic radii slightly decreased after addition of sialic
25 acid, maybe because other interactions stabilize the dimer, once it is formed (Figure 4A).
26 Fitting the average hydrodynamic radii of the different concentrations with a binding
27 isotherm allowed us to estimate a K_D of ~30 μ M for the B10-peptide interaction, which
28 is two magnitudes lower than the affinity towards the sialic acid. It should be noted that
29 the K_D for the isolated peptide is likely lower, because the dimerization may entail a
30 cooperative effect.

Discussion

The conformational dynamics between the open- and closed-states of SBPs have only recently been studied for ABC-transporter related SBPs^{14, 16, 31} and were completely unknown for TRAP transporter SBPs. Our above described FRET experiments and kinetic analysis reveal that VcSiaP has a ligand dependent substrate-binding rate. Contrary to some SBPs of ABC importers^{15, 17, 32, 33}, no ligand-free closed or ligand-bound open state of VcSiaP were detected in our experiments. Thus, VcSiaP likely uses an induced-fit binding mechanism³⁴. As for ABC importers³⁵, this ligand binding mechanism presumably allows the transmembrane domains of the transporter to discriminate between the ligand-free and ligand-bound SBP and thus to initiate transport, while avoiding futile transport cycles. It should be noted that the time-resolution of ~5 ms used in the present study is not sufficient to fully exclude a mechanism involving very fast intrinsic conformational dynamics < 1 ms or transitions that are so rare that they require a longer observation time to be detected.

The high-resolution crystal structure of the closed-state of VcSiaP presented above completes the basic structural information for the two best studied model systems of TRAP transporter SBPs, VcSiaP and HiSiaP. The tag peptide bound to sialic acid binding site is structurally completely different from the natural substrate. Still, it appears to trigger the same closing mechanism, as evident from the difference distance matrices in Figure 4C. One might therefore speculate, that an important prerequisite for triggering the closing mechanism is a substrate that can physically bridge the gap between the two lobes of the SBP and thereby provides the trigger needed to overcome the open-locked state energetic barrier of SBPs. It is interesting that D₁₄ of the peptide emulates the interaction with natural substrates (Figure 4B). One possible mechanism is therefore that the substrate first binds to a residue that acts as a substrate filter²¹ and if it is large enough to bridge the gap to the second lobe and make sufficient interactions, the P-domain closes. Small molecules would hence not be able to induce the closed conformation, as exemplified by a sulfate molecule tightly bound to arginine 145 in the open structure of VcSiaP (PDB-ID: 4MAG)¹¹. This fits nicely to the very weak evolutionary coupling between the residues lining the inner surfaces of the gap (Figure 3E). Overall, the structure of the peptide bound VcSiaP protein is reminiscent of the complex between the maltose transporter MalFGK₂ and its SBP, where a non-conserved periplasmic loop protrudes into the binding cleft of the SBP³⁰. A similar observation has been made in the complex of the vitamin B12 transporter (BtuCDF) and its SBP³⁶. The feature has been

1 interpreted as a “scoop loop”, which helps to dislocate the bound maltose molecule from
 2 the SBP (i.e. the maltose binding protein), or to prevent reassociation of the released
 3 ligand with the SBP before it can form a productive interaction with the membrane
 4 domains ³⁰. It will be interesting to investigate in further studies, if the ability of VcSiaP
 5 to bind peptides in the binding site regions are an indication of such a “scoop loop”
 6 mechanism for TRAP transporters.

7 As mentioned above, TRAP transporters are present in many important pathogens such
 8 as *V. cholerae* and *H. influenzae*. Since TRAPs are absent in humans, they are potential
 9 targets for the development of new antibiotics. The mere fact that a TRAP transporter
 10 SBP is able to bind a peptide sequence opens an interesting avenue at developing
 11 inhibitors for TRAP transporters, similar to the successful inhibition of vitamin B12
 12 transporter SBP BtuF with a nanobody, bound to the binding site and blocking the
 13 substrate binding ³⁷. Our peptide screens have shown that with only 32 rationally designed
 14 peptides, a strong increase in affinity could be accomplished. With an estimated K_D in the
 15 range of 30 μ M, the tested peptide sequences are still far away from affinities of
 16 compounds required for drug development. But, due to the modular structure of peptides,
 17 even short sequences give rise to an easily accessible and very large pool of possible
 18 structures ³⁸. The X-ray structure of the peptide linked SBP dimer immediately suggests
 19 that screening approaches such as Bac2Hybrid or phage display assays ³⁹ might be
 20 efficient ways to find and characterize such inhibitors in further experiments.

21

22 **Conclusion**

23 We have used smFRET to analyse the closing mechanism of VcSiaP in real time and
 24 found that the closing movement of the TRAP transporter SBP is substrate induced and
 25 closely related to SBPs from ABC transporters. The comparison of two ligand-bound
 26 structures of VcSiaP, one with the natural and one with a peptide ligand suggests that the
 27 closing mechanism of VcSiaP is triggered by physically bridging the gap between the two
 28 lobes of the protein. The peptide-bound structure of VcSiaP allowed us to design stronger
 29 binding peptides that can serve as starting points to find possible peptide-based inhibitors
 30 for TRAP transporters.

Material & Methods

Expression and Purification

The proteins were expressed and purified as described before ⁴⁰. The genes for substrate binding proteins VcSiaP and HiSiaP were cloned into a pBADHisTEV vector with a N-terminal 6xHis-TEV tag (Huanting Liu, University of St Andrews). All mutations were performed according to a protocol by Liu et al ⁴¹. The mutants were expressed in M9-minimal media to prevent the presence of the substrate sialic acid during expression from LB media ⁴⁰. To purify the protein, a Ni²⁺-affinity chromatography was followed by ion-exchange chromatography and size-exclusion chromatography. If necessary, the His-tag was cleaved off by an overnight incubation at 4 °C with 1:50 ratio protein to TEV protease and separated by a second Ni²⁺-affinity chromatography. The purity of protein containing fractions were checked with SDS-PAGE after each step. The purified proteins were concentrated, flash frozen and stored at - 80 °C.

Crystallization and Structure Determination

For crystallization, the purified proteins were concentrated to ~ 25 mg/mL. The crystallization was carried out with vapour diffusion method by using sitting drop 96 well crystallization plates. The substrate bound approaches were previously incubated for 30 minutes on ice with a 2 times excess of sialic acid. In every drop, 0.5 µL protein solution and 0.5 µL screen solution were mixed and the crystal plates were grown for several weeks at room temperature under exclusion of light and vibrations. The drops were regularly checked with a microscope or a protein crystallization imager (Rock Imager 1000, Formulatrix US). For crystallization approaches with substrate Neu5Ac, one crystal was obtained in commercially available screen Morpheus (Molecular Dimensions, UK) well F4 and for the peptide-bound approaches in self-designed optimization screen, based on JCSG-plus (Molecular Dimensions, UK) well A7/D7 (optimized condition: 0.2 M lithium sulfate, 0.1 M Tris (pH 8.5) and 1.26 M ammonium sulfate). The obtained crystals were harvested with a cryo-loop from the initial drop, soak quickly into a mixture of 35 % glycerol reservoir solution for cryo protectant and flash frozen into liquid nitrogen. The diffraction data for VcSiaP dimer were collected at beamline BL14.3 of BESSYII (Berlin, Germany), with a MarMOSAIC 225 CCD detector. The data for VcSiaP with sialic acid were collected at beamline PX1 (Zuerich, Switzerland) with EIGER 16M detector. The diffraction data were integrated with XDS ⁴² and the structures were solved by PHASER ²⁴ with molecular replacement using PDB-ID 4MAG ¹¹ for VcSiaP dimer and PDB-ID

1 3B50 for sialic acid bound VcSiaP¹⁰. The structures were refined with PHENIX⁴³ and
2 COOT⁴⁴ and the quality of the model was checked after each step with MolProbity⁴⁵.

3 4 *ITC measurement*

5 The ITC experiments were performed with a MicroCal PEAQ-ITC machine from
6 Malvern Panalytical (UK). Previously to every measurement, the sample cell was
7 equilibrated three times with 300 μ L standard protein buffer (50 mM Tris, 50 mM NaCl,
8 pH=8). The protein solution was diluted to 100 μ M in a total volume of 350 μ L with
9 standard buffer. After transferring the protein solution into the sample cell without any air,
10 the concentration of the remaining solution was determined via UV-Vis absorption with
11 a NanoDrop 2000 (ThermoFisher Scientific, US). The syringe was automatically loaded
12 with 1.2 mM sialic acid in standard protein buffer. The experiment was carried out
13 isothermally at 25 °C and 19 injections of the ligand with a volume of 2 μ L per injections
14 were titrated into the sample cell. The experimental data were recorded and analysed with
15 commercial software from Malvern Panalytical. The heat pulses were integrated and the
16 isotherm curve for each measurement were fitted with the one set of sites fitting mode to
17 get the experimental and thermodynamic parameters.

18 19 *Protein labelling*

20 Stochastic labelling with the dyes Alexa555 and Alexa647 maleimide (ThermoFisher)
21 was performed as done previously¹⁶. In brief, the purified proteins were first treated
22 with 10 mM Dithiothreitol (DTT) for 30 min to fully reduce oxidized cysteines. After
23 dilution of the protein sample to a DTT concentration of 1 mM the reduced protein was
24 immobilized on a Ni²⁺-Sephacrose resin and washed with ten column volumes of buffer A
25 (50 mM Tris-HCl, pH 7.4, 50 mM KCl). The resin was incubated 1-8 h at 4 °C with a
26 dye-to-protein ratio of ~20 dissolved in buffer A. Subsequently, unbound dyes were
27 removed by washing with twenty column volumes of buffer A. Elution of the proteins
28 was done with buffer A and 400 mM imidazole. The labelled proteins were further
29 purified by size-exclusion chromatography (Superdex 200, GE Healthcare) using
30 buffer A.

31
32
33
34

Solution-based smFRET experiments

Solution based smFRET experiments⁴⁶ were performed on a homebuilt confocal ALEX⁴⁷ microscope as previously described^{16,17}. All sample solutions were measured with 100 µl drop on a coverslip with concentration of around 50 pM in buffer A.

The fluorescent donor molecules were excited by a diode laser OBIS 532-100-LS (Coherent, USA) at 532 nm operated at 60 µW. The fluorescent acceptor molecules are excited by a diode laser OBIS 640-100-LX (Coherent, USA) at 640 nm operated at 25 µW at the sample in alternation mode (50 µs alternation period). The lasers were combined by a polarization maintaining single-mode fiber P3-488PM- FC-2 (Thorlabs, USA). The laser light is guided into the epi-illuminated confocal microscope (Olympus IX71, Hamburg, Germany) by dual-edge beamsplitter ZT532/640rpc (Chroma/AHF) focussed by a water immersion objective (UPlanSApo 60x/1.2w, Olympus Hamburg, Germany). The emitted fluorescence is collected through the objective and spatially filtered using a pinhole with 50 µm diameter and spectrally split into donor and acceptor channel by a single-edge dichroic mirror H643 LPXR (AHF). Fluorescence emission was filtered (donor: BrightLine HC 582/75 (Semrock/AHF), acceptor: Longpass 647 LP Edge Basic (Semrock/AHF) and focused on avalanche photodiodes (SPCM-AQRH-64, Excelitas). The detector outputs were recorded by a NI-Card PCI-6602 (National Instruments, USA).

Data analysis was performed using home written software package as described in¹⁴. Single-molecule events were identified using an All-Photon-Burst-Search algorithm with a threshold of 15, a time window of 500 µs and a minimum total photon number of 50⁴⁸. Photon count data were extracted, background subtracted and plotted in ES-histogram with a minimum of 150 counts.

E-histogram of double-labelled FRET species with Alexa555 and Alexa647 was extracted by selecting 0.25<S<0.75. E-histograms of open state without ligand (apo) and closed state with saturation of Neu5Ac at 200 µM concentration (holo) were fitted with a

Gaussian distribution $A e^{-\frac{(E-\mu)^2}{2\sigma^2}}$ resulting in ($\mu_o = 0.866$; $\sigma_o = 0.059$) and ($\mu_c = 0.932$; $\sigma_c = 0.031$) for open and closed state. The closed fraction in the titration experiment was calculated as $r_c = \frac{A_c \sigma_c}{A_o \sigma_o + A_c \sigma_c}$ from the double Gaussian fit to the data

$A_o e^{-\frac{(E-\mu_o)^2}{2\sigma_o^2}} + A_c e^{-\frac{(E-\mu_c)^2}{2\sigma_c^2}}$, where mean and standard deviations were fixed to the respective stated values.

Surface-based smFRET experiments

Confocal scanning experiments were performed at room temperature and using a home-built confocal scanning microscope as described previously¹⁴. In brief, surface scanning was performed using a XYZ-piezo stage with 100 μ m $\times$ 100 μ m $\times$ 20 μ m range (P-517-3 CD with E-725.3CDA, Physik Instrumente). The detector signal was registered using a HydraHarp 400 picosecond event timer and a module for time-correlated single photon counting (both Picoquant).

Surface immobilization was done as follows. Microscope cover-slides (No. 1.5, Marienfeld) were cleaned by sonication in ethanol, acetone and MQ water for 10 min each, followed by plasma etching (Plasma Etch, PE-25- JW) for 10 min. Functionalization with PEG-Silane (6-9 PE units) and Biotin-PEG-Silane (MW3400, Laysan Inc.) was done in toluene (55 $^{\circ}$ C; o/n). The surfaces were incubated for 10 min with 0.2 mg/ml neutravidin (Invitrogen) in buffer A (50 mM Tris-HCl, pH 7.4, 50 mM KCl). Unbound neutravidin was washed away with buffer A. Next, the surface was incubated with 1 nM biotinylated penta-His antibody (Qiagen) in buffer A for 10 min. Then, labelled proteins were immobilized by incubating with 5-100 pM of labelled protein in buffer A for 5 min. Unbound proteins were washed away with buffer A.

All surface-based smFRET measurements were done as described previously^{16, 17}. In brief, measurements were carried out in buffer A supplemented with 1 mM of 6-Hydroxy-2,5,7,8-tetramethylchromane-2-carboxylic acid (Trolox) and 10 mM Cysteamine. Data were recorded with constant 532 nm excitation at an intensity of 0.5 μ W ($\sim$ 125 W/cm²). Scanning images of 10 $\times$ 10 μ m were recorded with 50 nm step size and 2 ms integration time at each pixel. After each surface scan, the positions of labelled proteins were identified manually. The position information was used to subsequently generate fluorescence trajectories with a time resolution of 5 ms. The donor and acceptor fluorescence counts were used to calculate the apparent FRET efficiency, via $F_{DA}/(F_{DA}+F_{DD})$, where F_{DA} and F_{DD} are the uncorrected acceptor and donor photon counts. The apparent FRET efficiency time-traces were further analysed with hidden Markov modelling as described previously¹⁷. In the absence of ligand and in the presence of 1 mM Neu5Ac the time-traces were fitted with one state, and in the presence of 500 nM Neu5Ac with two states. From the fit, the individual lifetimes could be extracted as well as the average apparent FRET efficiency of each state.

Peptide Screens

To detect the interaction of VcSiaP with several peptides on the screen surface, a construct VcSiaP Q245C was expressed and purified as described before. To avoid dimerization of the introduced cysteine, every buffer was supplemented with 1 mM TCEP (Tris(2-carboxyethyl)phosphine). Before labelling, the TCEP was removed through a size-exclusion chromatography step on a SD 200 10/300 with standard protein buffer without TCEP. Immediately after elution, the protein fractions were combined and supplemented with four times excess of DyLight800 Maleimide Fluorophore Label (ThermoFisher Scientific, US). The solution was incubated overnight under light exclusion at 4 °C. On the next day, the remaining free label was removed with a PD-10 column and standard protein buffer. The protein elution fraction was concentrated up to 1 mg/mL.

The peptide screen was designed based on the interaction in the VcSiaP dimer structure and ordered from peptides&elephants (Germany). The peptides were synthesized in single spots on a cellulose membrane. Before incubation with the target protein, the membrane was soaked with water, washed three times for 15 seconds with methanol and again soaked into water. For reduction of unspecific binding, the peptide screen was incubated for 1 h at room temperature with 1 % BSA in standard protein buffer (50 mM Tris, 50 mM NaCl, pH=8). After three washing steps with buffer A for each 10 minutes the screen was incubated for 1 h at room temperature in 10 mL buffer A with 1 µg VcSiaP Q245-DyLight800. Afterwards the screen was again washed three times for 10 minutes with buffer A. The fluorescence onto the screen was detected with an Odyssey Imaging System at a wavelength of 800 nm.

DLS measurements

For the determination of the hydrodynamic radius R_0 of different VcSiaP constructs, protein solutions of different concentrations were prepared for each construct in the same manner. The purified protein solution was concentrated to a final concentration of 40 mg/ml. Subsequently, to one half of the volume 100 mM sialic acid solution (dissolved in 50 mM NaCl, 50 mM Tris, pH=8) was added to a final concentration of 10 mM. The other half of the concentrated protein solution was diluted with 1/10 of buffer (50 mM NaCl, 50 mM Tris, pH=8). Both of the solutions were then diluted step-wise to result in a concentration series with protein concentrations of 40 mg/ml, 30 mg/ml, 20 mg/ml, 10 mg/ml, and 5 mg/ml. These solutions were centrifugated at 15000 x g for 15 min to remove aggregates and measured in a DLS-device (DynaPro® NanoStar® (Wyatt Technology Corporation)). For each condition, three measurements of 10 µl of VcSiaP

1 solution were done at a sample temperature of $T = 25\text{ }^{\circ}\text{C}$ and three measurement cycles
 2 of each 20 single data acquisitions with acquisition times of $T_{\text{acq}} = 3\text{ s}$. For the calculation
 3 of the hydrodynamic radii based on crystal structures, the online tool ‘HullRad’⁴⁹ was
 4 used ($R(\text{dim})=3.61\text{ nm}$, $R(\text{monomer, apo})=2.59\text{ nm}$, $R(\text{monomer, holo})=2.53\text{ nm}$).

5

6 **Accession numbers**

7 PDB: 7A5Q

8 PDB: 7A5C

9

10 **Glossary**

11 TRAP transporters: Tripartite ATP-independent periplasmic transporters.

12 SBP: Substrate binding protein

13 ABC transporter: ATP binding cassette transporter

14

15

1 **Acknowledgements**

2 This project was financed by the German Research Foundation (Deutsche
3 Forschungsgemeinschaft, DFG) in project no. HA 6805/5-1 (to GH), GRK2062 – project
4 number C03 – (to T.C.), SFB 863 – Project number A13 – 111166240 (to TC) and the
5 European Research Council (ERC-StG “SM-IMPORT”, No. 638536 to TC). MFP
6 acknowledges a PhD fellowship from the Konrad Adenauer Stiftung. CG acknowledges
7 a PhD fellowship from the Studienstiftung des deutschen Volkes. GH thanks Olav
8 Schiemann (University of Bonn) for support. GH thanks Matthias Geyer (University of
9 Bonn) for helpful discussions. JG and GH thank Anton Schmitz (caesar, Bonn) for help
10 with the fluorescence detection of the peptide screens. The authors thank beamline staff
11 at SLS (Zuerich) and BESSY (Berlin) for support during data collection.

12

1 **Author contributions**

2 GH conceived this study. MFP, CG, JG, GHT, TC and GH have designed the experiments
3 and analyzed the data. MFP and NS expressed and purified the protein constructs and
4 determined the protein structures, JG performed the peptide scan experiments. CG and
5 MdB performed single-molecule FRET experiments. MFP and GH wrote the manuscript
6 with input from all authors.

7

8

References

1. Vetting MW, Al-Obaidi N, Zhao S, San Francisco B, Kim J, Wichelecki DJ, Bouvier JT, Solbiati JO, Vu H, Zhang X, Rodionov DA, Love JD, Hillerich BS, Seidel RD, Quinn RJ, Osterman AL, Cronan JE, Jacobson MP, Gerlt JA, Almo SC. Experimental strategies for functional annotation and metabolism discovery: targeted screening of solute binding proteins and unbiased panning of metabolomes. 2015;54:909-931.
2. Forward JA, Behrendt MC, Wyborn NR, Cross R, Kelly DJ. TRAP transporters: a new family of periplasmic solute transport systems encoded by the dctPQM genes of *Rhodobacter capsulatus* and by homologs in diverse gram-negative bacteria. *Journal of Bacteriology* 1997;179:5482-5493.
3. Fischer M, Zhang QY, Hubbard RE, Thomas GH. Caught in a TRAP: substrate-binding proteins in secondary transport. *Trends in microbiology* 2010;18:471-478.
4. Rosa LT, Bianconi ME, Thomas GH, Kelly DJ. Tripartite ATP-Independent Periplasmic (TRAP) Transporters and Tripartite Tricarboxylate Transporters (TTT): From Uptake to Pathogenicity. *Frontiers in cellular and infection microbiology* 2018;8:33.
5. Mulligan C, Fischer M, Thomas GH. Tripartite ATP-independent periplasmic (TRAP) transporters in bacteria and archaea. *FEMS Microbiology Reviews* 2011;35:68-86.
6. Severi E, Hood DW, Thomas GH. Sialic acid utilization by bacterial pathogens. *Microbiology* 2007;153:2817-2822.
7. Mulligan C, Leech AP, Kelly DJ, Thomas GH. The Membrane Proteins SiaQ and SiaM Form an Essential Stoichiometric Complex in the Sialic Acid Tripartite ATP-independent Periplasmic (TRAP) Transporter SiaPQM (VC1777-1779) from *Vibrio cholerae*. 2012;287:3598-3608.
8. Bouchet V, Hood DW, Li J, Brisson J-R, Randle GA, Martin A, Li Z, Goldstein R, Schweda EKH, Pelton SI, Richards JC, Moxon ER. Host-derived sialic acid is incorporated into *Haemophilus influenzae* lipopolysaccharide and is a major virulence factor in experimental otitis media. *Proceedings of the National Academy of Sciences of the United States of America* 2003;100:8898-8903.
9. Müller A, Severi E, Mulligan C, Watts AG, Kelly DJ, Wilson KS, Wilkinson AJ, Thomas GH. Conservation of structure and mechanism in primary and secondary transporters exemplified by SiaP, a sialic acid binding virulence factor from *Haemophilus influenzae*. *Journal of Biological Chemistry* 2006;281:22212-22222.
10. Johnston JW, Coussens NP, Allen S, Houtman JCD, Turner KH, Zaleski A, Ramaswamy S, Gibson BW, Apicella MA. Characterization of the N-acetyl-5-neuraminic acid-binding site of the extracytoplasmic solute receptor (SiaP) of nontypeable *Haemophilus influenzae* strain 2019. *Journal of Biological Chemistry* 2008;283:855-865.
11. Gangi Setty T, Cho C, Govindappa S, Apicella MA, Ramaswamy S. Bacterial periplasmic sialic acid-binding proteins exhibit a conserved binding site. *Acta Crystallographica Section D* 2014;70:1801-1811.
12. Berntsson RPA, Smits SHJ, Schmitt L, Slotboom D-J, Poolman B. A structural classification of substrate-binding proteins. *FEBS Letters* 2010;584:2606-2617.
13. Quijcho FA, Spurlino JC, Rodseth LE. Extensive features of tight oligosaccharide binding revealed in high-resolution structures of the maltodextrin transport/chemosensory receptor. *Structure* 1997;5:997-1015.
14. Gouridis G, Schuurman-Wolters GK, Ploetz E, Husada F, Vietrov R, De Boer M, Cordes T, Poolman B. Conformational dynamics in substrate-binding domains

- 1 influences transport in the ABC importer GlnPQ. *Nature Structural & Molecular*
2 *Biology* 2015;22:57.
- 3 15. Tang C, Schwieters CD, Clore GM. Open-to-closed transition in apo maltose-
4 binding protein observed by paramagnetic NMR. *Nature* 2007;449:1078-1082.
- 5 16. de Boer M, Gouridis G, Vietrov R, Begg SL, Schuurman-Wolters GK, Husada F,
6 Eleftheriadis N, Poolman B, McDevitt CA, Cordes T. Conformational and
7 dynamic plasticity in substrate-binding proteins underlies selective transport in
8 ABC importers. *Elife* 2019;8:e44652.
- 9 17. de Boer M, Gouridis G, Muthahari YA, Cordes T. Single-molecule observation of
10 ligand binding and conformational changes in FeuA. *Biophysical journal*
11 2019;117:1642-1654.
- 12 18. Glaenger J, Peter MF, Thomas GH, Hagelueken G. PELDOR Spectroscopy
13 Reveals Two Defined States of a Sialic Acid TRAP Transporter SBP in Solution.
14 *Biophysj* 2017;112:109-120.
- 15 19. Van Der Velde JHM, Oelerich J, Huang J, Smit JH, Jazi AA, Galiani S, Kolmakov
16 K, Gouridis G, Eggeling C, Herrmann A. A simple and versatile design concept
17 for fluorophore derivatives with intramolecular photostabilization. *Nature*
18 *communications* 2016;7:10144.
- 19 20. Husada F, Gouridis G, Vietrov R, Schuurman-Wolters GK, Ploetz E, de Boer M,
20 Poolman B, Cordes T. Watching conformational dynamics of ABC transporters
21 with single-molecule tools. *Biochemical Society Transactions* 2015;43:1041-1047.
- 22 21. Fischer M, Hopkins AP, Severi E, Hawkhead J, Bawdon D, Watts AG, Hubbard
23 RE, Thomas GH. Tripartite ATP-independent Periplasmic (TRAP) Transporters
24 Use an Arginine-mediated Selectivity Filter for High Affinity Substrate Binding.
25 *2015*;290:27113-27123.
- 26 22. Darby JF, Hopkins AP, Shimizu S, Roberts SM, Brannigan JA, Turkenburg JP,
27 Thomas GH, Hubbard RE, Fischer M. Water networks can determine the affinity
28 of ligand binding to proteins. *Journal of the American Chemical Society*
29 2019;141:15818-15826.
- 30 23. Rabiner LR. A tutorial on hidden Markov models and selected applications in
31 speech recognition. *Proceedings of the IEEE* 1989;77:257-286.
- 32 24. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ.
33 Phasercrystallographic software. *Journal of Applied Crystallography* 2007;40:658-
34 674.
- 35 25. Mulligan C, Geertsma ER, Severi E, Kelly DJ, Poolman B, Thomas GH. The
36 substrate-binding protein imposes directionality on an electrochemical sodium
37 gradient-driven TRAP transporter. *Proceedings of the National Academy of*
38 *Sciences* 2009;106:1778-1783.
- 39 26. Robert X, Gouet P. Deciphering key features in protein structures with the new
40 ENDscript server. *Nucleic acids research* 2014;42:W320-W324.
- 41 27. Ovchinnikov S, Kamisetty H, Baker D. Robust and accurate prediction of residue-
42 residue interactions across protein interfaces using evolutionary information. *eLife*
43 2014;3:1061-1021.
- 44 28. Marks DS, Colwell LJ, Sheridan R, Hopf TA, Pagnani A, Zecchina R, Sander C.
45 Protein 3D structure computed from evolutionary sequence variation. *PloS one*
46 2011;6:e28766.
- 47 29. Frank R. The SPOT-synthesis technique: synthetic peptide arrays on membrane
48 supports—principles and applications. *Journal of immunological methods*
49 2002;267:13-26.
- 50 30. Oldham ML, Khare D, Quirocho FA, Davidson AL, Chen J. Crystal structure of a
51 catalytic intermediate of the maltose transporter. *Nature* 2007;450:515-521.

- 1 31. Kim E, Lee S, Jeon A, Choi JM, Lee H-S, Hohng S, Kim H-S. A single-molecule
2 dissection of ligand binding to a protein with intrinsic dynamics. *Nature chemical*
3 *biology* 2013;9:313-318.
- 4 32. Van Meervelt V, Soskine M, Singh S, Schuurman-Wolters GK, Wijma HJ,
5 Poolman B, Maglia G. Real-time conformational changes and controlled
6 orientation of native proteins inside a protein nanoreactor. *Journal of the American*
7 *Chemical Society* 2017;139:18640-18646.
- 8 33. Duan X, Hall JA, Nikaido H, Quijcho FA. Crystal structures of the
9 maltodextrin/maltose-binding protein complexed with reduced oligosaccharides:
10 flexibility of tertiary structure and ligand binding. *Journal of molecular biology*
11 2001;306:1115-1126.
- 12 34. Mao B, Pear MR, McCammon JA, Quijcho FA. Hinge-bending in L-arabinose-
13 binding protein. The "Venus's-flytrap" model. *Journal of Biological Chemistry*
14 1982;257:1131-1133.
- 15 35. Doeven MK, Van den Bogaart G, Krasnikov V, Poolman B. Probing receptor-
16 translocator interactions in the oligopeptide ABC transporter by fluorescence
17 correlation spectroscopy. *Biophysical journal* 2008;94:3956-3965.
- 18 36. Hvorup RN, Goetz BA, Niederer M, Hollenstein K, Perozo E, Locher KP.
19 Asymmetry in the structure of the ABC transporter-binding protein complex
20 BtuCD-BtuF. *Science* 2007;317:1387-1390.
- 21 37. Mireku SA, Sauer MM, Glockshuber R, Locher KP. Structural basis of nanobody-
22 mediated blocking of BtuF, the cognate substrate-binding protein of the
23 *Escherichia coli* vitamin B12 transporter BtuCD. *Scientific reports* 2017;7:1-12.
- 24 38. Heninot A, Collins JC, Nuss JM. The Current State of Peptide Drug Discovery:
25 Back to the Future? *Journal of Medicinal Chemistry* 2018;61:1382-1414.
- 26 39. Battesti A, Bouveret E. The bacterial two-hybrid system based on adenylate
27 cyclase reconstitution in *Escherichia coli*. *Methods (San Diego, Calif)*
28 2012;58:325-334.
- 29 40. Glaenger J, Peter MF, Hagelueken G. Studying structure and function of
30 membrane proteins with PELDOR/DEER spectroscopy - The
31 crystallographers' perspective. *Methods (San Diego, Calif)* 2018;147:163-
32 175.
- 33 41. Liu H, Naismith JH. An efficient one-step site-directed deletion, insertion, single
34 and multiple-site plasmid mutagenesis protocol. *BMC Biotechnology* 2008;8:91.
- 35 42. Kabsch W. Automatic-Indexing of Rotation Diffraction Patterns. 1988;21:67-71.
- 36 43. Adams P, Grosse-Kunstleve R, Hung L, Ioerger T, McCoy A, Moriarty N, Read
37 R, Sacchettini J, Sauter N, Terwilliger T. PHENIX: building new software for
38 automated crystallographic structure determination. *Acta Crystallographica*
39 *Section D* 2002;58:1948-1954.
- 40 44. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta*
41 *Crystallographica Section D* 2004;60:2126-2132.
- 42 45. Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ,
43 Murray LW, Richardson JS, Richardson DC. MolProbity: all-atom structure
44 validation for macromolecular crystallography. *Acta Crystallographica Section D*
45 2010;66:12-21.
- 46 46. Lerner E, Cordes T, Ingargiola A, Alhadid Y, Chung S, Michalet X, Weiss S.
47 Toward dynamic structural biology: Two decades of single-molecule Förster
48 resonance energy transfer. *Science* 2018;359
- 49 47. Kapanidis AN, Lee NK, Laurence TA, Doose S, Margeat E, Weiss S.
50 Fluorescence-aided molecule sorting: analysis of structure and interactions by

- 1 alternating-laser excitation of single molecules. Proceedings of the National
2 Academy of Sciences 2004;101:8936-8941.
- 3 48. Eggeling C, Fries JR, Brand L, Günther R, Seidel CAM. Monitoring
4 conformational dynamics of a single molecule by selective fluorescence
5 spectroscopy. Proceedings of the National Academy of Sciences 1998;95:1556-
6 1561.
- 7 49. Fleming PJ, Fleming KG. HullRad: fast calculations of folded and disordered
8 protein and nucleic acid hydrodynamic properties. Biophysical journal
9 2018;114:856-869.
- 10 50. Laskowski RA, Swindells MB. LigPlot+: multiple ligand-protein interaction
11 diagrams for drug discovery. Journal of chemical information and modeling
12 2011;51:2778-2786.
- 13 51. Hagelueken G, Abdullin D, Schiemann O. mtsslSuite: Probing Biomolecular
14 Conformation by Spin-Labeling Studies. Methods in Enzymology 2015;563:595-
15 622.
- 16

Figure captions

Figure 1: Steady-state Neu5Ac binding by VcSiaP. **A)** Apparent FRET efficiency histograms of VcSiaP with labels Alexa555/647 for different Neu5Ac substrate concentrations (0 (top) to 200 μ M (bottom)). The fraction of molecules in the closed state r_c was determined by a global fit of 2 Gaussians for the open and closed state, shown in dark gray and green, respectively. **B)** Representative 2D ES-histograms from data in A) at 4 different concentrations. **C)** Fraction of closed VcSiaP r_c depending on ligand concentration extracted from Gaussian fits to data in A). The data points were fitted to the function $r_c = \frac{c_L}{K_D + c_L}$ with a dissociation constant of $K_D = 300 \pm 100$ nM (95% confidence interval). **D+E)** ITC data of Neu5Ac titrated to VcSiaP solution (left) and corresponding binding isotherm (right).

Figure 2. Study of immobilized VcSiaP proteins for analysis of sub-second dynamics. **A)** Schematic of surface immobilization procedure. **B)** Representative fluorescence trajectories of VcSiaP (Q54C/L173C) labelled with AlexaFluor 555 and AlexaFluor 647 under different conditions as indicated in the figure. The top panel shows the apparent FRET efficiency (blue) and donor (green) and acceptor (red) photon counts in the bottom panel. Orange lines are the most probable state-trajectory of the Hidden Markov Model (HMM). **C)** Apparent FRET efficiency histogram of all fluorescence trajectories recorded in the absence and presence of 1 mM Neu5Ac. **D)** Histogram and cumulative distribution function (CDF; inline panel) for the lifetime of the closed conformation (high FRET state) from all fluorescence trajectories recorded in the presence 500 nM Neu5Ac ($N_{\text{traces}} = 45$). The black line is an exponential fit.

Figure 3: Structural classification of sialic acid-bound VcSiaP. **A)** Outline of the primary sequence of VcSiaP with structural motifs and positions of selected amino acids. **B)** Alignment of open (white, PDB-ID: 4MAG¹¹) and closed (green) conformations (this work) of VcSiaP. The substrate from the closed structure is represented as blue-coloured spheres. **C)** Ligplot analysis of the substrate binding interactions from VcSiaP. The water molecules are represented as magenta spheres⁵⁰. **D)** Surface conservation of TRAP transporter SBPs, calculated with the ENDscript server²⁶ and mapped onto the closed VcSiaP structure. The proteins that are used for this analysis are listed in a supporting file. The protein is shown as grey surface model and red regions indicate conserved surface regions. The two views are rotated by 180°. **E)** Evolutionary couplings²⁸ analysis of the closed VcSiaP structure as model. The protein is shown as a green cartoon and the bound substrate as blue-red ball-sticks. Interactions between pairs of residues that are identified as conserved from the analysis are shown as magenta lines.

Figure 4: Crystal structure of an intertwined dimer of VcSiaP. **A)** Two interlocked monomers of VcSiaP are shown in cartoon representation (green and red). The N-terminus of one domain is bound to the substrate binding cleft of the other domain. The N-termini are indicated by spheres. Numbers in circles and the amino acid sequence at the bottom clarify, which part of the N-terminal tag is bound to the neighboring molecule. The magenta sphere indicates residue Q245, which was used to mutate and attach a fluorescence label for binding studies (see below). **B)** Close-up of the peptide-VcSiaP interactions in the dimeric crystal structure. **C)** Difference distance matrices between the substrate-free and peptide-bound VcSiaP (top-left half) and the substrate-free and -sialic

1 acid bound VcSiaP structures (bottom-right half). The coloring of the two matrices
2 (different scales) indicates the extend of the conformational changes between each pair
3 of structures. Rigid bodies are indicating by dashed lines. The matrices were created with
4 mtsslWizard ⁵¹ (<http://www.mtsslsuite.isb.ukbonn.de>).
5

6 **Figure 5: Characterization and strengthening of peptide-VcSiaP interactions. A)**
7 Spots with synthesized peptides on cellulose membranes after incubation with
8 fluorescence labelled VcSiaP and detection of fluorescence at 800 nm. The experiment
9 was performed with two identical peptide screens, one without addition of sialic acid (left)
10 and one after incubation of VcSiaP with sialic acid (right). **B)** Sequences of the peptide
11 screens from A). Single mutations are indicating in red letters, mutations with clearly
12 different binding of VcSiaP are highlighted in magenta and explained in the main text.
13 **C)** DLS measurements for hydrodynamic radius determination of different VcSiaP
14 constructs at certain concentration. For each protein construct, the measurements were
15 carried out with (dark icons) and without (bright icons) substrate. **D)** Histograms of the
16 radius distributions in mass percent for different concentrations from the untagged protein
17 (red) and the protein with strong binding peptide B10 (green). The calculated radii from
18 the crystal structures are shown as black bars.
19
20

1 Tables

2 Table 1: Data collection and refinement statistics.

	Native sialic acid-bound	Artificial peptide-bound
Space group	P2 ₁ 2 ₁ 2 ₁	P6 ₃
Unit cell (Å)	48.4, 106.7, 134.5	118.7, 118.7, 113.8
Wavelength (Å)	1.000009	0.894290
Total reflections	152048 (40060)	493043 (42386)
Unique reflections	79907 (7732)	46188 (3974)
Completeness (%)	99.6 (97.8)	100.0 (100.0)
Multiplicity	7.2 (5.2)	10.7 (10.7)
Resolution (Å)	44.05-1.68 (1.74-1.68)	59.4-2.20 (2.27-2.20)
I/sigma (I)	7.1 (1.3)	16.6 (1.8)
Wilson B-factor (Å²)	23.74	41.4
R_{merge}	0.089 (0.92)	0.085 (1.48)
R_{work}	0.212 (0.383)	0.188 (0.316)
R_{free}	0.248 (0.363)	0.234 (0.333)
Rmsd bonds (Å)	0.013	0.007
Rmsd angles (°)	1.27	0.98
Ramachandran favored/allowed/forbidden	98.66/1.34/0	98.8/1.2/0
Unfavorable rotamers (%)	0.19	0
Molprobrity Score ⁴⁵	1.26	1.33
PDB-ID	7A5Q	7A5C

3

4

Figure 1

Figure 1: Steady-state Neu5Ac binding by VcSiaP. **A)** Apparent FRET efficiency histograms of VcSiaP with labels Alexa555/647 for different Neu5Ac substrate concentrations (0 (top) to 200 μM (bottom)). The fraction of molecules in the closed state r_c was determined by a global fit of 2 Gaussians for the open and closed state, shown in dark gray and green, respectively. **B)** Representative 2D ES-histograms from data in A) at 4 different concentrations. **C)** Fraction of closed VcSiaP r_c depending on ligand concentration extracted from Gaussian fits to data in A). The data points were fitted to the function $r_c = \frac{c_L}{K_D + c_L}$ with a dissociation constant of $K_D = 300 \pm 100$ nM (95% confidence interval). **D+E)** ITC data of Neu5Ac titrated to VcSiaP solution (left) and corresponding binding isotherm (right).

Figure 2

Figure 2. Study of immobilized VcSiaP proteins for analysis of sub-second dynamics.

A) Schematic of surface immobilization procedure. **B)** Representative fluorescence trajectories of VcSiaP (Q54C/L173C) labelled with AlexaFluor 555 and AlexaFluor 647 under different conditions as indicated in the figure. The top panel shows the apparent FRET efficiency (blue) and donor (green) and acceptor (red) photon counts in the bottom panel. Orange lines are the most probable state-trajectory of the Hidden Markov Model (HMM). **C)** Apparent FRET efficiency histogram of all fluorescence trajectories recorded in the absence and presence of 1 mM Neu5Ac. **D)** Histogram and cumulative distribution function (CDF; inline panel) for the lifetime of the closed conformation (high FRET state) from all fluorescence trajectories recorded in the presence 500 nM Neu5Ac ($N_{traces} = 45$). The black line is an exponential fit.

Figure 3

Figure 3: Structural classification of sialic acid-bound VcSiaP. **A)** Outline of the primary sequence of VcSiaP with structural motifs and positions of selected amino acids. **B)** Alignment of open (white, PDB-ID: 4MAG¹¹) and closed (green) conformations (this work) of VcSiaP. The substrate from the closed structure is represented as blue-coloured spheres. **C)** Ligplot analysis of the substrate binding interactions from VcSiaP. The water molecules are represented as magenta spheres⁵⁰. **D)** Surface conservation of TRAP transporter SBPs, calculated with the ENDscript server²⁶ and mapped onto the closed VcSiaP structure. The proteins that are used for this analysis are listed in a supporting file. The protein is shown as grey surface model and red regions indicate conserved surface regions. The two views are rotated by 180°. **E)** Evolutionary couplings²⁸ analysis of the closed VcSiaP structure as model. The protein is shown as a green cartoon and the bound substrate as blue-red ball-sticks. Interactions between pairs of residues that are identified as conserved from the analysis are shown as magenta lines.

Figure 4

Figure 4: Crystal structure of an intertwined dimer of VcSiaP. **A)** Two interlocked monomers of VcSiaP are shown in cartoon representation (green and red). The N-terminus of one domain is bound to the substrate binding cleft of the other domain. The N-termini are indicated by spheres. Numbers in circles and the amino acid sequence at the bottom clarify, which part of the N-terminal tag is bound to the neighboring molecule. The magenta sphere indicates residue Q245, which was used to mutate and attach a fluorescence label for binding studies (see below). **B)** Close-up of the peptide-VcSiaP interactions in the dimeric crystal structure. **C)** Difference distance matrices between the substrate-free and peptide-bound VcSiaP (top-left half) and the substrate-free and -sialic acid bound VcSiaP structures (bottom-right half). The coloring of the two matrices (different scales) indicates the extent of the conformational changes between each pair of structures. Rigid bodies are indicating by dashed lines. The matrices were created with mtsslWizard⁵¹ (<http://www.mtsslsuite.isb.ukbonn.de>).

Figure 5

Figure 5: Characterization and strengthening of peptide-VcSiaP interactions. A) Spots with synthesized peptides on cellulose membranes after incubation with fluorescence labelled VcSiaP and detection of fluorescence at 800 nm. The experiment was performed with two identical peptide screens, one without addition of sialic acid (left) and one after incubation of VcSiaP with sialic acid (right). **B)** Sequences of the peptide screens from A). Single mutations are indicating in red letters, mutations with clearly different binding of VcSiaP are highlighted in magenta and explained in the main text. **C)** DLS measurements for hydrodynamic radius determination of different VcSiaP constructs at certain concentration. For each protein construct, the measurements were carried out with (dark icons) and without (bright icons) substrate. **D)** Histograms of the radius distributions in mass percent for different concentrations from the untagged protein (red) and the protein with strong binding peptide B10 (green). The calculated radii from the crystal structures are shown as black bars.

1

2

3

4

5

6

7

8

9

10

Supporting information

11

12

13

SI Figure 1: Interaction between sialic acid and TRAP SBPs. The substrate interaction interfaces were calculated with *LigPlot*⁵⁰ and the corresponding sialic acid-bound SBPs (VcSiaP from *Vibrio cholerae* (this work), HiSiaP from *Haemophilus influenzae* (PDB-ID: 3B50¹⁰), PmSiaP from *Pasturella multocida* (PDB-ID: 4MMP¹¹) and FnSiaP from *Fusobacterium nucleatum* (PDB-ID: 4MNP¹¹))

SI Figure 2: Structure refinement and comparison of peptide-VcSiaP interaction.

A) Difference electron density (green, $\sigma=5.0$) reveals a clear peptide-chain like discrepancy in the active binding of VcSiaP (cartoon, red). **B)** The N-terminal affinity tag could be successfully built into the difference electron density from A) and connect to a neighboured VcSiaP monomer ($\sigma=1.5$). **C)** The spin label MTSSL attached to a cysteine at position Q54C ($\sigma=1.5$). **D)** The spin label MTSSL attached to a cysteine at position L173C ($\sigma=1.5$). **E)** Mutation R125A from previous studies to validate the overall protein structure for this mutant.

SI Figure 3: Analysis of the dimeric VcSiaP interface. **A)** Surface representation of the dimeric VcSiaP with affinity tag bound to the active binding site of the opposite VcSiaP. One chain is represented in green, the other one in red. The spin label positions at positions Q54C and L173C are represented in blue. **B)** Modelling of Arg125 back into the crystal structure of the dimer showing no obvious interruptions into peptide binding. For a comparison of artificial peptide and natural substrate, the sialic acid (magenta) was modelled into the dimeric protein structure. **C)** Hydrodynamic radii from DLS experiments for VcSiaP wildtype with HisTEV (dark red), VcSiaP wildtype without HisTEV (light red) and VcSiaP R125A Q54R1 L173R1 with HisTEV (blue), which was used for crystallization of VcSiaP dimer.

SI Figure 4: Characterisation of the interaction between peptides and VcSiaP active binding site. **A)** Analysis of the peptide-active binding site interaction with LigPlot ⁵⁰. The peptide is shown in green, the amino acids of the active binding site in pink. Strong interactions are represented as dashed, pink lines. **B)** Representation of the peptide inside the active binding cleft of VcSiaP. The peptide is shown in green sticks, the surface of VcSiaP with electrostatics. **C)** Same as B) but with model of a strong binding peptide (cyan sticks) from Figure 5 A9. **D)** Hypothetical profile of the membrane domains Q (yellow) and M (blue) from the TRAP transporter VcSiaPQM, the C-terminus be located on the cytoplasm side and the N-terminus on the periplasm site ⁷.