Structure-function relationships in substrate binding protein dependent secondary transporters

Dissertation

zur Erlangung des Doktorgrades

der Naturwissenschaften

vorgelegt beim Fachbereich 14

Biochemie, Chemie und Pharmazie

der Johann Wolfgang Goethe-Universität

in Frankfurt am Main

von

Anja Roden

aus Wittlich

Frankfurt am Main 2023

(D 30)

Vom Fachbereich Biochemie, Chemie und Pharmazie (FB14)
der Johann Wolfgang Goethe-Universität als Dissertation angenommen.

Dekan: Prof. Dr. Clemens Glaubitz

Gutachter:

Dr. Eric Geertsma

Prof. Dr. Klaas Martinus Pos

Datum der Disputation: 28.07.2023

Table of contents

Summary

Tripartite transport systems exhibit increased substrate affinity and transport rates as they depend on substrate binding proteins (SBPs). In contrast to primary transporters, SBP dependent secondary transport systems are under-researched although already discovered in 1996 with TRAPRc-DctPQM from *Rhodobacter capsulatus* (Madden 2002; Jacobs et al. 1996). Within the family of the TRAP transport systems three subfamilies evolved, the TRAP-TRAP, TRAP-TAXI and TRAP-TPAT transporter (L. T. Rosa, Bianconi, et al. 2018). To reduce the complexity of the naming, we will hereafter use the terms TRAP, TAXI and TPAT. Following the nomenclature of TRAPRc-DctPQM the subunits of TRAP transporter are most often referred to as P for the SBP and Q and M for the two membrane domains (Jacobs et al. 1996). TRAP transport systems are found in bacteria and those of the TAXI family additionally in archaea because of which TAXI transport systems are believed to represent an ancient form of TRAP transporter (Mulligan, Fischer, and Thomas 2011; L. T. Rosa, Bianconi, et al. 2018). Detailed *in vitro* characterization is limited to three N-acetylneuraminic acid TRAP transport systems, derived from *Haemophilus influenzae, Vibrio cholerae* and *Photobacterium profundum* (Mulligan et al. 2012, 2009; Davies, Currie, and North 2023). Within the SBP dependent secondary transport systems, the TAXI family is the least studied group. Structurally only one TAXI SBP was solved, named TtGluBP from *Thermus thermophilus* (Takahashi et al. 2004) and functional characterization of a TAXI transport system is absent from the literature (Mulligan, Fischer, and Thomas 2011). We selected homologs of this group for functional characterization. Using BLAST, we selected the homologs based on their similarity to TRAPRc-DctPQM from *Rhodobacter capsulatus.* After the identification of homologs, we amplified TRAP-QM domains to overexpress and detergent-solubilize them in analytical tests. This was successful for transport systems deriving from *Desulfotomaculum carboxydivorans*, *Shimwellia blattae*, *Natrialba asiatica*, *Proteus mirabilis* and *Marinobacter hydrocarbonoclasticus*. We then expressed and purified successfully all five corresponding SBPs. To identify potential substrates of the tripartite transport systems, we exposed the SBPs to members of a compound library created from known and similar ligands of TRAP SBPs and studied their thermal melting by differential scanning fluorimetry (DSF). A wide range of mostly C_4- and C_5-dicarboxylates, amino acids (aa), and sugars were tested. After studying the gene neighborhood of TAXIPm-PQM from *Proteus mirabilis*, α-ketoglutarate and its direct precursor α-hydroxyglutarate were included into the compound library, motivated by the presence of the *IhgO* gene, coding for a putative L-2-hydroxyglutarate oxidase, located downstream of the TAXIPm-QM and TAXIPm-P genes in the same operon.

For TAXIPm-P from *Proteus mirabilis* a profound increase in melting temperature was observed in DSF in the presence of the C_5-dicarboxylates α-hydroxyglutarate and α-ketoglutarate. The binding appears to be very specific as compounds that are structurally closely related, such as glutarate or glutamate, did not change the thermostability of the protein. When different α-ketoglutarate concentrations were supplied, a destabilization of the protein was observed above 50 µM.

For TAXIMh-P from *Marinobacter hydrocarbonoclasticus* a distinct increase in melting temperature was observed in DSF in the presence of the C_4-dicarboxylates fumarate, succinate, and L-malate.

Similarly, to the deorphanization of TAXIPm-P, genes located in close proximity to the transporter provided information about the ligand for TAXIMh-P as well. *DcuB*, located in the operon next to TAXIMh-P and TAXIMh-QM, encodes an antiporter which imports the C_4-dicarboxylates fumarate, malate, aspartate, and D-tartrate in exchange for succinate. After the substrates are imported, fumarase, and fumarate reductase convert L-malate to fumarate and fumarate to succinate. L-malate, fumarate, and succinate are all structurally and metabolically connected (Kim 2006) simultaneously to α-hydoxyglutarate and α-ketoglutarate (Brunengraber 2007).

Two further proteins selected in the expression screenings derived from *Desulfotomaculum carboxydivorans* (Dc) and *Shimwellia blattae* (Sb) were tested in DSF supplying different substrates from the compound library. As none of the tested substrates bound to the respective SBPs, the following work focused on the two transporters TAXIMh-PQM from *Marinobacter hydrocarbonoclasticus* and TAXIPm-PQM from *Proteus mirabilis*. To confirm the ligand binding and determine the affinities of the SBPs for the ligands, we performed isothermal titration calorimetry (ITC). For TAXIMh-P, the binding of L-malate and fumarate was confirmed, but binding of succinate could not be confirmed. This could originate from destabilization of the protein under the tested conditions, or from little to no heat effect with the binding of succinate to the SBP (Boudker and Oh 2015). K_D-values of 10.9 µM for L-malate and 31.2 µM for fumarate were determined. Additionally, the transport of fumarate was confirmed for TAXIMh-PQM as we observed robust uptake in glucose-energized cells using an inward-directed Na^+-gradient. By determining the initial transport rates at different fumarate concentrations, we obtained an apparent K_m of TAXIMh-PQM for fumarate of 14.6 µM. While whole cell uptake assays demonstrated fumarate uptake by TAXIMh-PQM, we did not succeed to demonstrate this *in vitro*. This was although the purification of TAXIMh-PQM was successful, and a high reconstitution efficiency was obtained (75 %). Also, the transport assay was based on a protocol previously established for the TRAP transporter TRAPHi-SiaPQM (Mulligan et al. 2009). Despite the theoretically optimal conditions no significant transport above the background signal was observed for reconstituted TAXIMh-PQM. Therefore, in a further investigation of this study, we focused on TAXIPm-PQM from *Proteus mirabilis*.

For the SBP TAXIPm-P from the TAXIPm-PQM system, the binding of α-ketoglutarate was confirmed by ITC, showing a K_D-value of 8.4 µM. To additionally determine which substrate species preferably binds, we performed ITC for TAXIPm-P at pH 5.5, 6.0 and 7.5. At these pH-values α-ketoglutarate occurs in different amounts in the single anionic and di-anionic form due to its pK_a-values of 1.9 and 4.4 (Tokonami et al. 2013). With the change of the pH-values, the titrations resulted in similar K_D-values between 16.6 µM and 8.4 µM which does not reflect the 70-fold change in concentration of the single anionic form. Therefore, we suggest that the most dominant, the di-anionic form, which was abundant in all scenarios, is most likely the substrate. Furthermore, the transport of α-ketoglutarate was verified *in vivo* and additionally *in vitro* for TAXIPm-PQM. In our *in vivo* transport assays, we observed robust TAXIPm-PQM dependent uptake of α-ketoglutarate in glucose-energized cells, indicating that α-ketoglutarate is a substrate for the membrane-embedded transport domain as well. As the initial transport assays showed successful α-ketoglutarate uptake by TAXIPm-PQM in both the presence and absence of Na^+-ions, we assumed a potential proton coupling.

Therefore α-ketoglutarate transport by TAXIPm-PQM was tested in the presence and absence of the protonophore CCCP. The presence of CCCP diminished the proton gradient and the uptake, implying that the transport is coupled to the proton motive force. We determined the initial transport rate at different α-ketoglutarate concentrations and calculated an apparent affinity of 31 µM at pH 7.5 and of 84 µM when the pH was shifted to 6.0. To further characterize the transport of TAXIPm-PQM *in vitro*, the SBP TAXIPm-P and the transport unit TAXIPm-QM were both purified and reconstituted into liposomes. For this, we followed the protocol previously established for the TRAP transporter TRAPHi-SiaPQM (Mulligan et al. 2009) and the knowledge gained from the *in vivo* experiments, that TAXIPm-PQM is most likely coupled to the proton motive force. Despite apparent good conditions no significant transport of α-ketoglutarate above the background signal was observed for reconstituted TAXIPm-QM in the initial screenings. So, we evaluated the differences to the well-transporting TAXIPm-PQM *in vivo* in more detail. The TAXIPm-QM membrane domain is in both cases located in the membrane, *in vivo* and *in vitro*. There might be a difference in the SBP, however, as it is located *in vivo* in the periplasmic space and *in vitro* outside the proteoliposomes. The periplasmic space *in vivo* is a confined environment which results in comparably high SBP concentrations. For primary transport systems it is known, that some SBPs appear anchored to the membrane through N-terminal lipid moieties, by incorporation of transmembrane helices in the membrane or as a direct fusion at the membrane domain (van der Heide and Poolman 2002). It is believed that the anchoring of the SBP increases its local concentration at the membrane domain to an even higher extent than soluble SBPs can achieve as the anchor allows two-dimensional movements only (Doeven et al. 2004).

To determine whether the SBP TAXIPm-P requires anchoring to the membrane, we performed a bioinformatic analysis using SignalP (Almagro Armenteros et al. 2019; Teufel et al. 2022). In total, we analyzed 215 TAXI and 636 TRAP SBP sequences from the TRAP database (Mulligan, Kelly, and Thomas 2007) for which transmembrane-anchors were identified in 19 % of the TAXI transport systems and in 7 % of the TRAP transport systems. For TAXIPm-P an N-terminal transmembrane-anchor was predicted. Furthermore, we performed protein sequence alignments of the *E. coli* signal peptidase and *Proteus mirabilis* signal peptidases which showed that they have identities of 59 % whereby the catalytic residues of the *E. coli* signal peptidase (Mark Paetzel 2013) also appear in the *Proteus mirabilis* signal peptidases. Thereby this analysis implied that an *E. coli* signal peptidase can recognize and cleave a potential cleavage site in a signal peptide when it originates from *Proteus mirabilis*. So, we assume that if the membrane-tethered TAXIPm-P is present after expression in *E. coli*, it very likely occurs membrane-tethered in *Proteus mirabilis* as well. To experimentally investigate the possibility of an anchored SBP in the TAXIPm-PQM system, we expressed the TAXIPm-P SBP protein in two versions, one with and one without the endogenous signal peptide from *Proteus mirabilis*. We thereby observed that the construct expressed with endogenous signal sequence showed a different electrophoretic mobility in the gel which supported the idea of a transmembrane-anchored TAXIPm-P. The different size of the constructs and the consequent absence of a signal peptide cleavage was additionally confirmed by analysis of the constructs in mass spectrometry.

To further investigate the TAXIPm-PQM tripartite system *in vitro*, we reconstituted the TAXIPm-P including its membrane-anchor to TAXIPm-QM proteoliposomes. While detergent is required to solubilize membrane proteins, high concentrations of detergent will also destabilize lipid membranes. In addition, destabilization of the liposomes may enable P to be embedded in an inside-out orientation, which may allow substrate export by QM with a similar inside-out orientation. Since we aimed to mimic the *in vivo* occurrence, the SBP should only attach to the TAXIPm-QM proteoliposomes from the outside, therefore further destabilization should be prevented. For this purpose, we purified the SBP in the absence of detergent, assuming that the membrane-anchors would form micelle-like structures and thereby prevent each other from denaturation. Subsequent analysis showed that the aggregation was indeed reversible.

Next, we re-examined the *in vitro* uptake, this time with reconstituted SBP TAXIPm-P next to reconstituted TAXIPm-QM. This *in vitro* assay resulted in successful α-ketoglutarate uptake when we applied an inward directed pH gradient. To investigate the influence of the number of P molecules on the transport behavior, we reconstituted TAXIPm-P in different ratios to TAXIPm-QM. An increasing amount of most likely unphysiological high TAXIPm-P resulted in an earlier transport reduction, so we assume that besides the liganded TAXIPm-P, the unliganded TAXIPm-P also binds to TAXIPM-QM which reduces uptake or enables efflux as observed for TRAPHi-SiaPQM (Mulligan et al. 2009).

To investigate whether TAXIPm-PQM is a unidirectional importer similar to TRAPHi-SiaPQM (Mulligan et al. 2009), we tested to what extent an excess of α-ketoglutarate influences its transport. While the concentration of the SBP TAXIPm-P had a direct effect on α-ketoglutarate transport, the abundant concentration of substrate appeared to have no effect on transport, confirming the unidirectional import.

After the general functional characterization of TAXIPm-PQM, structural characterization of the entire complex and in different conformations would help to further investigate the substrate and proton coordination and also the interaction between the SBP and membrane domain. To initialize the work on this, nanobodies against TAXIPm-QM were generated in alpacas in collaboration with the Nanobody Service facility from the University of Zurich (www.nsf.uzh.ch), enriched at the MPI CBG, Dresden (Eric Geertsma Group, www.mpi-cbg.de) and selected within this study by ELISA. The nanobodies were then tested for potential inhibition of the α-ketoglutarate uptake by TAXIPm-PQM in *in vivo* transport assays. Several candidates were selected from this step, purified and sent to the MPI CBG, Dresden (Eric Geertsma Group, www.mpi-cbg.de) to aid structural characterization in subsequent steps.

As no TRAP structures were determined in complex so far, one can only speculate about the interface of the SBP and membrane domain and the docking site for the SBP. The periplasmic loops within the Q membrane domain were suggested to be important for the interaction and successful transport (Peter et al. 2022; Davies, Currie, and North 2023). This study confirmed this hypothesis. To investigate the interaction between the SBP and membrane domain, we constructed variants of TAXIMh-PQM from *Marinobacter hydrocarbonoclasticus* and TRAPHi-SiaPQM from *Haemophilus influenzae*. The exchange of five amino acid residues in the first periplasmic loop of the Q membrane

domain of TAXIMh-QM abolished the transport of fumarate *in vivo* completely. The finding suggests that we inserted mutations in the docking site for the SBP TAXIMh-P which can therefore not transfer the substrate to the membrane domain. In contrast, we deleted the complete membrane domain TRAPHi-QM in TRAPHi-SiaPQM and still detected transport of N-acetylneuraminic acid. From this, we concluded that TRAPHi-SiaP delivers the substrate to another *E. coli* transport system. It is conceivable that a transport protein that otherwise interacts with a different SBP can also interact with TRAPHi-SiaP if present. A similar behavior was previously suggested for TTT transport systems. As SBPs were often found overrepresented, it was concluded that individual membrane domains interact with different SBPs (L. T. Rosa, Springthorpe, et al. 2018). Therefore, we constructed a knock-out for the only endogenous *E. coli* TRAP transporter, YiaMNO, but still obtained N-acetylneuraminic acid uptake when the SBP TRAPHi-SiaP was present. With that, we propose that TRAPHi-SiaP interacts in our transport assays with a membrane domain that normally transports a substrate other than N-acetylneuraminic acid and presumably belongs to the primary transport systems. The membrane domain is potentially part of an ABC transport system, as this protein family represents the largest SBP dependent protein group in *E. coli* with 50 importers out of a total of 65 functional ABC transport systems (Moussatova et al. 2008).

Overall, this work provides new insights into the relevance of SBP dependent secondary transport systems, especially in the thus far under-researched subgroup of TAXI transporters. Importantly, we identified and characterized the TAXI transport system TAXIPm-PQM from *Proteus mirabilis.* We demonstrated that, in contrast to previously characterized SBP dependent secondary transport systems, TAXIPm-PQM is a proton coupled system and transports the C_5-dicarboxylate α-ketoglutarate. Since initially the transport of α-ketoglutarate could only be demonstrated *in vivo* but not *in vitro* using established protocols (Mulligan et al. 2009), we investigated in detail the differences between the *in vivo* and *in vitro* assay. This resulted in a bioinformatic analysis of TRAP and TAXI signal peptides, which strongly implied that TAXIPm-P requires a transmembrane anchor to allow for transport. We then provided TAXIPm-P surface tethered to the membrane in *in vitro* transport assays and confirmed the prediction of our bioinformatic analysis that TAXIPm-PQM deploys a membrane-anchored instead of a soluble SBP. Furthermore, the TAXI transport system TAXIMh-PQM from *Marinobacter hydrocarbonoclasticus* transports fumarate only if both membrane domains Q and M are present. For further characterization, Michaelis-Menten kinetics and affinities were determined for both TAXI transport systems TAXIPm-PQM from *Proteus mirabilis* and TAXIMh-PQM from *Marinobacter hydrocarbonoclasticus*. In addition, nanobodies were selected for the membrane domain TAXIPm-QM from *Proteus mirabilis* to stabilize different conformations which can serve in subsequent structural elucidation studies. Furthermore, the TRAP SBP TRAPHi-SiaP from *Haemophilus influenzae* was shown to interact not only with its corresponding membrane domain TRAPHi-SiaQM but with at least one additional transporter. It was thereby excluded that TRAPHi-SiaP transfers N-acetylneuraminic acid to the only native *E. coli* TRAP transporter TRAPEc-YiaMNO and suggested to rather interact with a SBP dependent ABC transport system as this protein family represents the largest SBP dependent protein group in *E. coli* (Moussatova et al. 2008).

Zusammenfassung

Dreiteilige Transportsysteme weisen durch ihre Abhängigkeit von Substratbindeproteinen (SBP) eine erhöhte Substrataffinität und Transportrate auf. Im Gegensatz zu primären Transportern sind SBP-abhängige sekundäre Transportsysteme noch wenig erforscht, obwohl sie bereits 1996 mit TRAPRc-DctPQM aus *Rhodobacter capsulatus* entdeckt wurden (Madden 2002; Jacobs et al. 1996). Innerhalb der Familie der TRAP-Transportsysteme haben sich drei Unterfamilien entwickelt, die TRAP-TRAP, TRAP-TAXI und TRAP-TPAT Transporter (L. T. Rosa, Bianconi, et al. 2018). Um die Komplexität der Namensgebung zu reduzieren, werden wir im Folgenden die Begriffe TRAP, TAXI und TPAT verwenden. In Anlehnung an die Nomenklatur von TRAPRc-DctPQM werden die Untereinheiten des TRAP-Transporters meist als P für das SBP und Q und M für die beiden Membrandomänen bezeichnet (Jacobs et al. 1996). TRAP-Transportsysteme finden sich in Bakterien und die der TAXI-Familie zusätzlich in Archaeen, weshalb man annimmt, dass TAXI-Transportsysteme eine Vorläufer-Form von TRAP-Transportern darstellen (Mulligan, Fischer, and Thomas 2011; L. T. Rosa, Bianconi, et al. 2018). Die detaillierte *in vitro* Charakterisierung beschränkt sich auf drei N-Acetylneuraminsäure-TRAP-Transportsysteme, die aus *Haemophilus influenzae*, *Vibrio cholerae* und *Photobacterium profundum* stammen (Mulligan et al. 2012, 2009; Davies, Currie, and North 2023). Innerhalb der SBP-abhängigen sekundären Transportsysteme ist die TAXI-Familie die am wenigsten untersuchte Gruppe. Strukturell wurde nur ein TAXI-SBP gelöst, das TtGluBP aus *Thermus thermophilus* (Takahashi et al. 2004), und eine funktionelle Charakterisierung eines TAXI-Transportsystems ist in der Literatur nicht zu finden (Mulligan, Fischer, and Thomas 2011). Wir haben daher Homologe der TAXI Gruppe für die funktionelle Charakterisierung ausgewählt. Mithilfe von BLAST wählten wir sie auf der Grundlage ihrer Ähnlichkeit mit TRAPRc-DctPQM aus *Rhodobacter capsulatus* aus und amplifizierten die TRAP-QM-Domänen, um sie in analytischen Tests zu überexprimieren und mit Detergenz zu solubilisieren. Dies war erfolgreich für Transportsysteme aus *Desulfotomaculum carboxydivorans*, *Shimwellia blattae*, *Natrialba asiatica*, *Proteus mirabilis* und *Marinobacter hydrocarbonoclasticus*. Die anschließende Expression und Reinigung der zugehörigen SBP war für alle fünf Systeme erfolgreich. Um potenzielle Substrate der dreiteiligen Transportsysteme zu identifizieren, setzten wir die SBP C_4- und C_5-Dicarboxylaten, Aminosäuren und Zuckern aus und untersuchten ihr thermisches Schmelzen mittels differentieller Scanning Fluorimetrie (DSF). Nach der Untersuchung der Gen-Nachbarschaft von TAXIPm-PQM aus *Proteus mirabilis* wurden α-Ketoglutarat und sein direkter Vorläufer α-Hydroxyglutarat in die Ligandenbibliothek aufgenommen.

Bei TAXIPm-P aus *Proteus mirabilis* wurde im DSF in Gegenwart der C_5-Dicarboxylate α-Hydroxyglutarat und α-Ketoglutarat ein starker Anstieg der Schmelztemperatur beobachtet. Die Bindung scheint sehr spezifisch zu sein, da strukturell verwandte Liganden, wie Glutarat oder Glutamat, die Thermostabilität des Proteins nicht veränderten. Wurden verschiedene α-Ketoglutarat-Konzentrationen zugeführt, so wurde oberhalb von 50 µM eine Destabilisierung des Proteins beobachtet. Für TAXIMh-P aus *Marinobacter hydrocarbonoclasticus* wurde ein deutlicher Anstieg der Schmelztemperatur im DSF in Gegenwart der C_4-Dicarboxylate Fumarat, Succinat und L-Malat

beobachtet. Ähnlich wie bei der Deorphanisierung von TAXIPm-P lieferte ein Gen, das sich in unmittelbarer Nähe des Transporters befindet, Hinweise auf diese Liganden.

Zwei weitere Proteine, die in den Expressionsscreenings aus *Desulfotomaculum carboxydivorans* (Dc) und *Shimwellia blattae* (Sb) ausgewählt wurden, wurden im DSF untersucht. Da keines der getesteten Liganden an die jeweiligen SBP band, konzentrierten sich die folgenden Arbeiten auf die beiden Transporter TAXIMh-PQM von *Marinobacter hydrocarbonoclasticus* und TAXIPm-PQM von *Proteus mirabilis*. Um die Affinitäten der SBP für die Liganden zu bestimmen, führten wir isothermale Titrationskalorimetrie (ITC) durch. Für TAXIMh-P wurde die Bindung von L-Malat und Fumarat bestätigt, die Bindung von Succinat konnte jedoch nicht bestätigt werden. Dies könnte auf eine Destabilisierung des Proteins unter den getesteten Bedingungen oder auf fehlende Wärmeumsetzung bei der Bindung von Succinat an das SBP zurückzuführen sein (Boudker and Oh 2015). Es wurden K_D-Werte von 10,9 µM für L-Malat und 31,2 µM für Fumarat ermittelt. Zusätzlich wurde der Transport von Fumarat für TAXIMh-PQM bestätigt, da wir eine robuste Aufnahme in Glukose-energetisierten Zellen unter Verwendung eines nach innen gerichteten Na^+-Gradienten beobachteten. Durch Bestimmung der initialen Transportraten bei verschiedenen Fumaratkonzentrationen erhielten wir einen scheinbaren K_m-Wert von TAXIMh-PQM für Fumarat von 14,6 µM. Während die Aufnahme von Fumarat durch TAXIMh-PQM in ganzen Zellen nachgewiesen werden konnte, ist es uns nicht gelungen, dies *in vitro* zu zeigen. Dies war der Fall, obwohl die Reinigung von TAXIMh-PQM erfolgreich war und eine hohe Rekonstitutionseffizienz erzielt wurde (75 %). Außerdem basierte der Transporttest auf einem Protokoll, das zuvor für den TRAP-Transporter TRAPHi-SiaPQM erstellt wurde (Mulligan et al. 2009). Trotz der theoretisch optimalen Bedingungen wurde für rekonstituiertes TAXIMh-PQM kein signifikanter Transport oberhalb des Hintergrundsignals beobachtet. Daher konzentrierten wir uns in der weiteren Untersuchung dieser Studie auf TAXIPm-PQM aus *Proteus mirabilis*.

Für das SBP TAXIPm-P aus dem TAXIPm-PQM-System wurde die Bindung von α-Ketoglutarat mittels ITC bestätigt, wobei ein K_D-Wert von 8,4 µM ermittelt wurde. Um zusätzlich zu bestimmen, welche Substratspezies bevorzugt bindet, führten wir ITC für TAXIPm-P bei pH 5,5, 6,0 und 7,5 durch und titrierten dabei α-Ketoglutarat, das bei den pH-Werten zu unterschiedlichen Anteilen einfach anionisch und di-anionisch vorliegt (pK_a-Werte 1,9 und 4,4) (Tokonami et al. 2013). Die Titrationen ergaben ähnliche K_D-Werte zwischen 16,6 µM und 8,4 µM, was nicht die 70-fache Änderung der Konzentration der einzelnen anionischen Form beim Wechsel von pH 7.5 zu pH 5.5 widerspiegelt. Daher vermuten wir, dass die dominanteste, die di-anionische Form, die in allen Szenarien reichlich vorhanden war, höchstwahrscheinlich das Substrat ist. Außerdem wurde der Transport von α-Ketoglutarat *in vivo* und zusätzlich *in vitro* für TAXIPm-PQM überprüft. In unseren *in vivo* Transportversuchen beobachteten wir eine robuste TAXIPm-PQM-abhängige Aufnahme von α-Ketoglutarat in Glukose-energetisierten Zellen, was darauf hindeutet, dass α-Ketoglutarat auch ein Substrat für die in die Membran eingebettete Transportdomäne ist. Die Anwesenheit von CCCP verringerte die Aufnahme, was darauf hindeutet, dass der Transport an die protonenmotorische Kraft gekoppelt ist. Wir bestimmten die anfängliche Transportrate bei verschiedenen α-Ketoglutarat-Konzentrationen und berechneten eine scheinbare Affinität von 31 µM bei pH 7,5 und von 84 µM,

wenn der pH auf 6,0 verschoben wurde. Um den Transport von TAXIPm-PQM *in vitro* weiter zu charakterisieren, wurden sowohl das SBP TAXIPm-P als auch die Transporteinheit TAXIPm-QM gereinigt und in Liposomen rekonstituiert. Dabei folgten wir dem bereits für den TRAP-Transporter TRAPHi-SiaPQM etablierten Protokoll (Mulligan et al. 2009) und der aus den *in vivo* Experimenten gewonnenen Erkenntnis, dass TAXIPm-PQM höchstwahrscheinlich an die protonenmotorische Kraft gekoppelt ist. Trotz scheinbar guter Bedingungen wurde bei den ersten Screenings für rekonstituiertes TAXIPm-QM kein signifikanter Transport von α-Ketoglutarat über das Hintergrundsignal hinaus beobachtet. Daher haben wir die Unterschiede zum gut funktionierenden *in vivo* Transport genauer untersucht. Die TAXIPm-QM-Membrandomäne befindet sich in beiden Fällen in der Membran, *in vivo* und *in vitro*. Einen Unterschied könnte es jedoch bei den SBP geben, da sie *in vivo* im begrenzten periplasmatischen Raum, aber *in vitro* außerhalb der Proteoliposomen lokalisiert sind. Bei primären Transportsystemen kennt man SBP, die durch N-terminale Lipideinheiten oder durch den Einbau von Transmembranhelices in die Membran verankert auftreten (van der Heide and Poolman 2002). Es wird vermutet, dass die Verankerung des SBP die lokale Konzentration an der Membrandomäne in einem noch höheren Maße erhöht, als es lösliche SBP erreichen können, da der Anker nur zweidimensionale Bewegungen zulässt (Doeven et al. 2004).
Um festzustellen, ob das SBP TAXIPm-P eine Verankerung an der Membran benötigt, haben wir eine bioinformatische Analyse mit SignalP durchgeführt (Almagro Armenteros et al. 2019; Teufel et al. 2022) Insgesamt analysierten wir 215 TAXI- und 636 TRAP-SBP-Sequenzen aus der TRAP-Datenbank (Mulligan, Kelly, and Thomas 2007), wobei Transmembran-Anker in 19 % der TAXI-Transportsysteme und in 7 % der TRAP-Transportsystemen identifiziert wurden; dabei wurde für TAXIPm-P ein N-terminales Transmembran Segment vorhergesagt. Weiterhin haben wir in Proteinsequenzvergleichen gezeigt, dass die katalytischen Reste der *E. coli* Signalpeptidase (Mark Paetzel 2013) auch in den *Proteus mirabilis* Signalpeptidasen vorkommen. Daher sollte eine *E. coli* Signalpeptidase eine potenzielle Schnittstelle in einem Signalpeptid erkennen und spalten, auch wenn sie von *Proteus mirabilis* stammt. Wir gehen also davon aus, dass, wenn das membrangebundene TAXIPm-P nach der Expression in *E. coli* vorhanden ist, es sehr wahrscheinlich auch in *Proteus mirabilis* membrangebunden vorkommt. Um die Möglichkeit eines verankerten SBP im TAXIPm-PQM-System experimentell zu untersuchen, exprimierten wir das TAXIPm-P SBP in zwei Versionen, eine mit und eine ohne das endogene Signalpeptid aus *Proteus mirabilis*. Dabei stellten wir fest, dass das mit endogener Signalsequenz exprimierte Konstrukt eine andere elektrophoretische Mobilität im Gel aufwies, was die Idee eines transmembranverankerten TAXIPm-P unterstützte. Die unterschiedliche Größe der Konstrukte und das damit verbundene Fehlen einer Signalpeptidabspaltung wurde zusätzlich durch eine massenspektrometrische Analyse der Konstrukte bestätigt.
Um das dreiteilige TAXIPm-PQM-System *in vitro* weiter zu untersuchen, haben wir das TAXIPm-P einschließlich seines Membranankers in TAXIPm-QM-Proteoliposomen rekonstituiert. Während Detergenzien zur Solubilisierung von Membranproteinen erforderlich sind, destabilisieren hohe Konzentrationen von Detergenzien auch Lipidmembranen. Darüber hinaus kann die Destabilisierung der Liposomen dazu führen, dass P in die Liposomen eingebettet wird, was den Substratexport durch

QM ermöglicht. Da wir das *in vivo* Vorkommen nachahmen wollten, sollte sich das SBP nur von außen an die TAXIPm-QM-Proteoliposomen anlagern, weshalb eine weitere Destabilisierung verhindert werden sollte. Zu diesem Zweck reinigten wir das SBP in Abwesenheit von Detergenz, in der Annahme, dass die Membrananker mizellenartige Strukturen bilden und sich dadurch gegenseitig an der Denaturierung hindern. Die anschließende Analyse zeigte, dass die Aggregation reversibel war.

Als nächstes untersuchten wir erneut die *in vitro* Aufnahme, diesmal mit rekonstituiertem SBP TAXIPm-P neben rekonstituiertem TAXIPm-QM. Dieser *in vitro* Test führte zu einer erfolgreichen α-Ketoglutarat-Aufnahme, wenn wir einen nach innen gerichteten pH-Gradienten anlegten. Um den Einfluss der Anzahl der P-Moleküle auf das Transportverhalten zu untersuchen, rekonstituierten wir TAXIPm-P in verschiedenen Verhältnissen zu TAXIPm-QM. Eine zunehmende Menge an wahrscheinlich unphysiologisch hohem TAXIPm-P führte zu einer früheren Transportreduktion, so dass wir davon ausgehen, dass neben dem Liganden gebundenen TAXIPm-P auch das Liganden ungebundene TAXIPm-P an TAXIPm-QM bindet, was die Aufnahme reduziert bzw. den Efflux ermöglicht, wie es für TRAPHi-SiaPQM beobachtet wurde (Mulligan et al. 2009). Um zu untersuchen, ob TAXIPm-PQM ähnlich wie TRAPHi-SiaPQM (Mulligan et al. 2009) ein unidirektionaler Importer ist, haben wir getestet, inwieweit ein Überschuss an α-Ketoglutarat seinen Transport beeinflusst. Während die Konzentration des SBP TAXIPm-P eine direkte Auswirkung auf den α-Ketoglutarat-Transport hatte, schien die sehr hohe Konzentration des Substrats keine Auswirkung auf den Transport zu haben, was den unidirektionalen Import bestätigte.

Nach der funktionellen Charakterisierung von TAXIPm-PQM würde eine strukturelle Charakterisierung des gesamten Komplexes und in verschiedenen Konformationen helfen, die Substrat- und Protonenkoordination sowie die Interaktion zwischen dem SBP und der Membrandomäne zu untersuchen. Daher wurden in Zusammenarbeit mit der Nanobody Service Facility der Universität Zürich (www.nsf.uzh.ch) Nanobodies gegen TAXIPm-QM in Alpakas generiert, am MPI CBG Dresden (Eric Geertsma Group, www.mpi-cbg.de) angereichert und im Rahmen dieser Studie mittels ELISA selektiert. Die Nanobodies wurden dann auf eine mögliche Inhibition der α-Ketoglutarat-Aufnahme durch TAXIPm-PQM in *in vivo* Transportversuchen getestet. Aus diesem Schritt wurden mehrere Kandidaten ausgewählt, gereinigt und an das MPI CBG Dresden (Eric Geertsma Group, www.mpi-cbg.de) geschickt, um die strukturelle Charakterisierung in den nachfolgenden Schritten zu unterstützen.

Da bisher keine TRAP-Strukturen im Komplex bestimmt wurden, kann man über das Interface von SBP und Membrandomäne und die Andockstelle für das SBP nur spekulieren. Um die Interaktion zwischen SBP und Membrandomäne zu untersuchen, konstruierten wir Varianten von TAXIMh-PQM aus *Marinobacter hydrocarbonoclasticus* und TRAPHi-SiaPQM aus *Haemophilus influenzae*. Der Austausch von fünf Aminosäureresten in der ersten periplasmatischen Schleife der Q-Membrandomäne von TAXIMh-QM hob den Transport von Fumarat *in vivo* vollständig auf. Diese Analyse deutet darauf hin, dass wir Mutationen in die Andockstelle für das SBP TAXIMh-P eingefügt haben, die daher das Substrat nicht an die Membrandomäne übertragen kann. Im Gegensatz dazu beobachteten wir, dass wir in TRAPHi-SiaPQM die Membrandomänen QM komplett deletieren

konnten und trotzdem N-Acetylneuraminsäure Transport stattfand. Daraus schlossen wir, dass TRAPHi-SiaP das Substrat an ein anderes Transportsystem von *E. coli* weiterleitet. Daher konstruierten wir ein Knock-out für den einzigen endogenen *E. coli* TRAP-Transporter, YiaMNO. Auch ohne YiaMNO erfolgte noch N-Acetylneuraminsäure Aufnahme, solange das SBP TRAPHi-SiaP vorhanden war. Es ist denkbar, dass ein Transportprotein, das ansonsten mit einem anderen SBP interagiert, auch mit TRAPHi-SiaP interagieren kann, wenn es vorhanden ist. Ein ähnliches Verhalten wurde auch für TTT-Transportsysteme vorgeschlagen, in denen SBP überrepräsentierte sind, woraus geschlossen wurde, dass einzelne Membrandomänen mit verschiedenen SBP interagieren (L. T. Rosa, Springthorpe, et al. 2018). Wir schlagen vor, dass TRAPHi-SiaP in unseren Transportassays mit einer Membrandomäne interagiert, die normalerweise ein anderes Substrat als N-Acetylneuraminsäure transportiert. Diese Membrandomäne ist möglicherweise Teil eines ABC-Transportsystems, da diese Proteinfamilie mit 50 Importern von insgesamt 65 funktionalen ABC-Transportsystemen die größte SBP-abhängige Proteingruppe in *E. coli* darstellt (Moussatova et al. 2008).

Insgesamt bietet diese Arbeit neue Einblicke in die Bedeutung von SBP-abhängigen sekundären Transportsystemen, insbesondere in der bisher wenig erforschten Untergruppe der TAXI-Transporter. Wichtig ist, dass wir das TAXI-Transportsystem TAXIPm-PQM aus *Proteus mirabilis* identifiziert und charakterisiert haben. Wir konnten zeigen, dass TAXIPm-PQM im Gegensatz zu bisher charakterisierten SBP-abhängigen sekundären Transportsystemen ein protonengekoppeltes System ist, das das C_5-Dicarboxylat α-Ketoglutarat transportiert. Da der Transport von α-Ketoglutarat zunächst nur *in vivo*, nicht aber *in vitro* mit etablierten Protokollen nachgewiesen werden konnte (Mulligan et al. 2009), haben wir die Unterschiede zwischen dem *in vivo* und dem *in vitro* Assay im Detail untersucht. Dies führte zu einer bioinformatischen Analyse der TRAP- und TAXI-Signalpeptide, die stark darauf hindeutete, dass TAXIPm-P einen Transmembrananker benötigt, um den Transport zu ermöglichen. Wir haben dann TAXIPm-P in *in vitro* Transportversuchen an die Membran gebunden und die Vorhersage unserer bioinformatischen Analyse bestätigt, dass TAXIPm-PQM ein membranverankertes statt ein lösliches SBP einsetzt. Außerdem transportiert das TAXI-Transportsystem TAXIMh-PQM aus *Marinobacter hydrocarbonoclasticus* Fumarat nur, wenn beide Membrandomänen Q und M vorhanden sind. Zur weiteren Charakterisierung wurden Michaelis-Menten-Kinetik und Affinitäten für die beiden TAXI-Transportsysteme TAXIPm-PQM aus *Proteus mirabilis* und TAXIMh-PQM aus *Marinobacter hydrocarbonoclasticus* bestimmt. Darüber hinaus wurden für die Membrandomäne TAXIPm-QM aus *Proteus mirabilis* Nanobodies ausgewählt, um möglicherweise verschiedene Konformationen zu stabilisieren, die in späteren Strukturaufklärungsstudien verwendet werden können. Darüber hinaus konnte gezeigt werden, dass das TRAP-SBP TRAPHi-SiaP aus *Haemophilus influenzae* nicht nur mit der zugehörigen Membrandomäne TRAPHi-SiaQM, sondern mit mindestens einem weiteren Transporter interagiert. Damit konnte ausgeschlossen werden, dass TRAPHi-SiaP N-Acetylneuraminsäure auf den einzigen nativen *E. coli* TRAP-Transporter TRAPEc-YiaMNO überträgt, und es wurde vermutet, dass TRAPHi-SiaP eher mit einem SBP-abhängigen ABC-Transportsystem interagiert, da diese Proteinfamilie die größte SBP-abhängige Proteingruppe in *E. coli* darstellt (Moussatova et al. 2008).

1 Introduction

1.1 Membrane transport

Biological membranes form a semi-permeable barrier around and compartmentalization within cells that are composed of amphiphilic phospholipids, proteins, and carbohydrates. Due to hydrophobic effects, the phospholipids form a bilayer in an aqueous environment which prevents free diffusion of substances (Lipowsky 1991). Membranes are involved in many cellular processes, including cell signaling and energy production, and they regulate the movement of ions and metabolites in and out of cells (Deamer 2016). To facilitate transport across membranes, a variety of proteins evolved. More than 20000 proteins are assigned in the Transporter Classification Database to more than 1500 transporter families (Saier, Tran, and Barabote 2006; Saier et al. 2021). More than 60 % of membrane transporters are targets of FDA-approved small-molecule drugs (Santos et al. 2017). According to their transport mechanism, membrane transport systems are categorized into channels, primary and secondary transport systems. Channels allow movement of molecules and ions in the direction of concentration gradients. Primary transport systems perform active transport of their substrate across the membrane. The largest group within primary transport systems are ATP-binding cassette transporters which transport their substrates independent of concentration gradients, but dependent on ATP (Guidotti 1996; Moussatova et al. 2008). The electrochemical gradient that is built up over the membrane by primary transport systems provides an indirect energy source for secondary active transport systems which couple substrate movement to ion flux and thereby allow import or export of substrates (Shi 2013). Within the primary and secondary transport systems mostly importing systems exist which increase their affinity and specificity for the substrate through SBPs (Reuss and Altenberg 2008). Such SBP dependent secondary transport systems are investigated in this study.

1.2 SBPs

Solute uptake systems in prokaryotes and archaea commonly involve a SBP to increase their affinity and specific uptake rate (Bosdriesz et al. 2015). SBPs are best known in the context of primary transport systems, for example ATP-binding cassette importers (Davidson and Maloney 2007; Lewinson and Livnat-Levanon 2017), but they are also found as part of secondary transporters (Forward et al. 1997; Jacobs et al. 1996) whose transport mechanism is less well understood. Known SBP dependent secondary transport systems are symporters that use Na^+-gradients to drive the import of amino acids (aa), N-acetylneuramic acids, (aromatic) dicarboxylates, α-keto acids, and even hydrophobic compounds (Brautigam et al. 2012; Deka et al. 2012; Gautom et al. 2021; Vetting et al. 2015; Mulligan et al. 2012, 2009).

Architecturally, SBPs consist of two globular domains composed of both β-sheets and α-helices with a substrate-binding cleft in the center (L. T. Rosa et al. 2017). They occur in two states, the open and closed conformations, whereby the conformational change into the closed state resembles the mechanism of ‘Venus fly traps’; it is induced by substrate binding (Peter et al. 2023; Scheepers, Lycklama a Nijeholt, and Poolman 2016). Having SBPs that enhance the specificity and affinity for a certain substrate is especially important when substrates are in low abundance. Thus, many

transporting systems take the advantage of SBPs, thereby primarily allowing unidirectional transport (Mulligan et al. 2009). SBPs occur in membrane-bound and soluble form in Gram-negative and Gram-positive bacteria as well as in archaea (Berntsson et al. 2010). SBPs can be classified into several general groups (Figure 1). Cluster A members are ATP binding cassette (ABC) Type II transporter SBPs, the special characteristic is one α-helix serving as hinge. Cluster B is divided in several subgroups, with SBPs belonging to ABC type I transporting systems, Natriuretic peptide receptors, and G-protein-coupled receptors. In general, proteins belonging to this class, show three interconnecting strands between the two domains. Cluster C contains SBPs with high sequence similarity to that of the ABC type I transporter family. Proteins belonging to cluster C entail an additional domain when compared to proteins from other clusters. Cluster D is divided into several subgroups containing SBPs all belonging to ABC type I transporting systems. Cluster D proteins are recognized by two short hinges. Cluster E contains SBPs belonging to Tripartite ATP independent periplasmic transport systems (TRAP) and Tripartite Tricarboxylate transport systems (TTT), most likely SBPs of TRAP associated extracytoplasmic immunogenic (TAXI) and Tetratricopeptide repeat protein associated transporting systems (TPAT) are also to be assigned to this group. Within this cluster, E-I type proteins contain a characteristic α-helix and two β-sheets that span both domains. E-II type proteins reveal a shortened form of this characteristic α-helix. Cluster F contains proteins belonging to Type I ABC transporting systems. They contain two relatively long hinges that provide flexibility. Cluster G proteins cannot be distinctly assigned to one of the other SBP clusters; they possess a relatively large size, similar to cluster C proteins, and long hinges, but additionally a metal-binding site (Figure 1) (Berntsson et al. 2010; Scheepers, Lycklama a Nijeholt, and Poolman 2016).

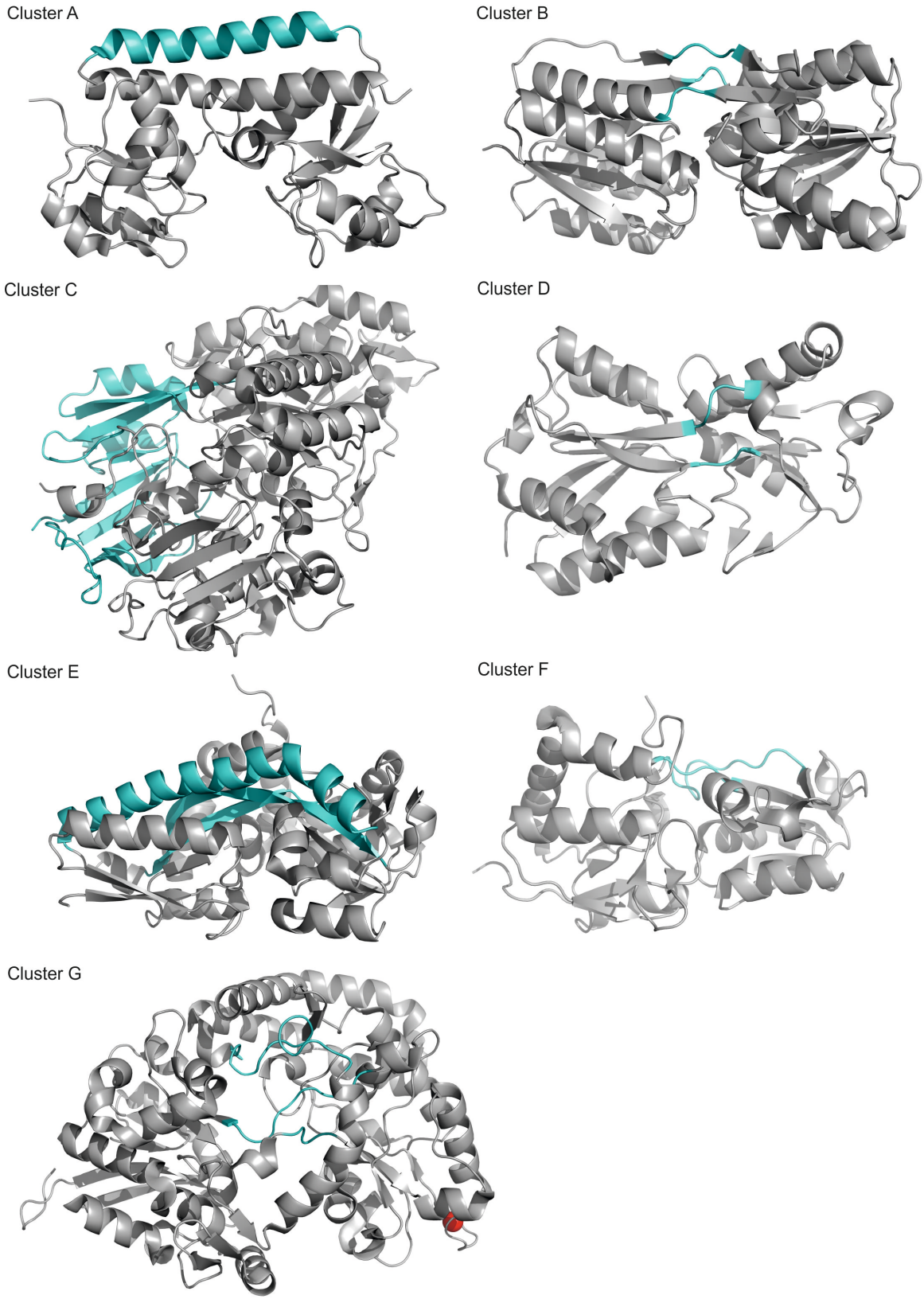


Figure 1 **Classification of SBPs according to Berntsson and Scheepers.** Cluster A representative vitamin B12 binding SBP BtuF from *E. coli* (PDB code: 1N2Z). Cluster B ribose binding SBP RBP from *E. coli* (PDB code: 1DRJ). Cluster C peptide binding SBP OppA from *Lactococcus lactis* (PDB code: 3DRF). Cluster D molybdate binding SBP ModA from *Pseudominas aeruginosa* (PDB code: 7T5A). Cluster E N-acetylneuraminic acid binding SBP SiaP from *Haemophilus influenzae* (PDB code: 2WXV). Cluster F histidine binding SBP HisJ from *E. coli* (PDB code: 1HSL). Cluster G unknown ligand SBP from *Bifidobacterium longum* with bound magnesium ion, displayed in red (PDB code: 3OMB). All characteristic features of the respective cluster are highlighted in teal. Figure adapted from Berntsson. (Berntsson et al. 2010; Scheepers, Lycklama a Nijeholt, and Poolman 2016). The structures were created using Pymol 1.8.4.1 (Schrödinger, LLC).

1.2.1 SBP dependent primary transport systems

SBP dependent primary transport systems are well characterized in the context of ATP-binding cassettes (Davidson and Maloney 2007; Lewinson and Livnat-Levanon 2017). They hydrolyze ATP for transport and use the SBP to obtain high affinity and specificity for their substrate. Their overall structure comprises two architecturally identical transmembrane domains with two nucleotide binding domains attached on the cytoplasmic site (Figure 2). The position of the SBP on the periplasmic site ensures unidirectional substrate translocation to import only (Figure 2).

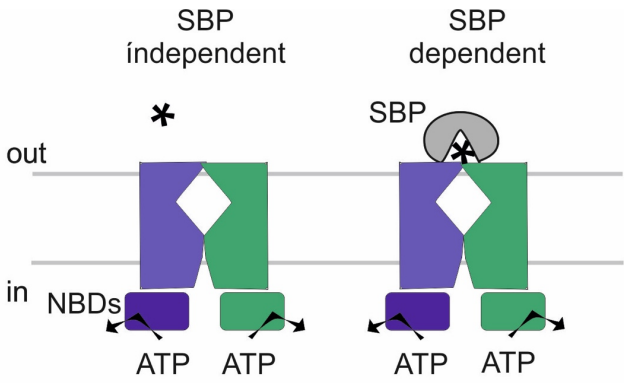


Figure 2 **SBP independent and SBP dependent primary transport systems.** Primary transport systems contain two architecturally identical transmembrane domains and two nucleotide binding domains (NBDs). These systems hydrolyze ATP for substrate translocation in which there are SBP independent and dependent systems. The transport system is embedded in a membrane, which is illustrated with two lines. The SBP is in the periplasm (out) and the NBDs in the cytoplasm (in). The substrate is depicted as a star.

Whereas ABC exporting systems exist in all kingdoms of life, ABC importers are found in bacteria and archaea only, thereby being SBP dependent or independent (Saurin, Hofnung, and Dassa 1999). The SBP on the periplasmic side ensures the import of the substrates by applying high specificity and affinity to the transport process. The SBPs evolved to different configurations, mostly found soluble in the periplasmic space of Gram-negative organisms and membrane-anchored in Gram-positive organisms and archaea which lack an outer membrane (van der Heide and Poolman 2002). The anchoring is realized either through N-terminal lipid moieties, by incorporation of transmembrane helices in the membrane, or the SBP appear as a direct fusion at the membrane domains, either individually or multiple times per transporter (Figure 3) (van der Heide and Poolman 2002). The oligopeptide transporter from *Lactococcus lactis* for example contains an N-terminal lipid moiety. OppA binds the peptides and delivers them to the membrane domains OppB and OppC to allow transport, in the case of bradykinin with an affinity of 0.1 µM. But not only the affinity of the SBP to the substrate, but also that of the SBP to the membrane domain is pivotal for efficient transport (Doeven et al. 2004). In ABC transport systems type I, for example, this affinity is usually low with K_m > 50 µM (Gul and Poolman 2012). It is believed that the anchoring of the SBP increases its local concentration at the membrane domain to a higher extent than soluble SBPs can achieve as the anchor allows two-dimensional movements only. Consequently, lower substrate concentrations are needed to achieve saturation of the transporter. For this transporter complex, the SBP alone determines the specificity as every peptide that it can bind is transported (Doeven et al. 2004).

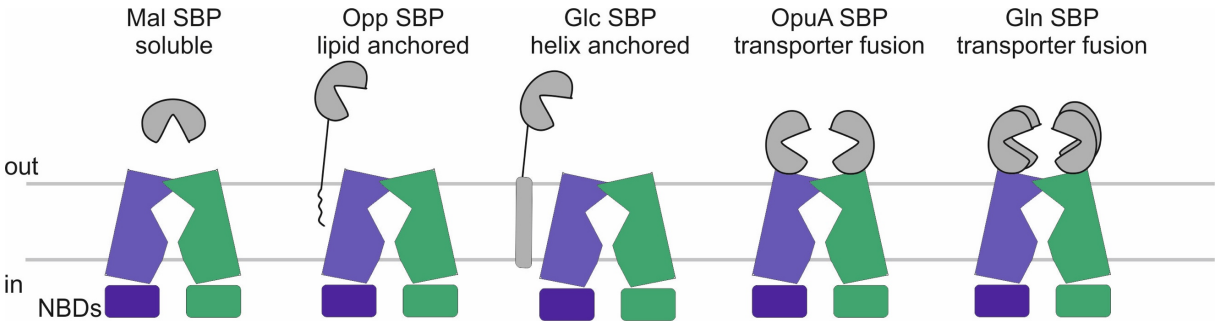


Figure 3 **Different types of SBPs.** ABC transporter uptake systems contain different types of SBPs in the periplasmic space to determine the direction of transport to import only. The SBPs exist in a soluble form in the maltose/maltodextrin transporter from *E. coli* (Mal), anchored through a lipid moiety in the oligopeptide transporter from *Lactococcus lactis* (Opp) or transmembrane helix within the membrane in the glucose transporter from *Sulfolobus solfataricus* (Glc) or directly fused to the TMDs in the glutamine/glutamic acid transporter from *Lactococcus lactis* (Gln). The transport systems are embedded in a membrane, which is illustrated with two lines. The SBPs are in the periplasm (out) and the nucleotide binding sites (NBDs) in the cytoplasm (in). The substrate is depicted as a star. Figure adapted van der Heide and Hofmann (van der Heide and Poolman 2002; Hofmann et al. 2019).

1.2.2 SBP dependent secondary transport systems

In general, secondary-active transport systems couple substrate movement to downhill ion flux, thereby allowing import or export of substrates. These ions are typically Na^+ or H^+ which allow transport against a concentration gradient, often in combination with the membrane potential. Next to conventional secondary transport systems, TRAP and TTT transport systems exist that involve a SBP in the transport process (Figure 4) (Winnen, Hvorup, and Saier 2003). With the representative TRAPHi-SiaPQM from *Haemophilus influenzae*, TRAP transport systems were shown to facilitate transport using the elevator type transport mechanism (section 1.3.6) (Peter et al. 2022). To enable the elevator like transport mechanism, the large membrane domain SiaM contains two domains, the scaffold and the transport domain whereby the scaffold domain is fixed in the membrane and the transport domain moves up and down to translocate the substrate. SiaQ extends the scaffold domain of SiaM (section 4.8.1) (Peter et al. 2022). While conventional secondary transport systems are found in all kingdoms of life, TRAP and TTT transport systems are widespread among archaea and bacteria only. Architecturally, TRAP and TTT transport systems contain three protein domains which are referred to as P for the SBP, Q for the small membrane domain and M for the large membrane domain following the nomenclature of the first reported TRAP transporter DctPQM from *Rhodobacter capsulatus* (Figure 4) (Jacobs et al. 1996). In contrast to primary transport systems, in which different types of SBPs were characterized (Figure 3), for secondary transport systems only transporters with soluble SBPs have thus far been characterized in detail (Mulligan et al. 2009; L. T. Rosa, Bianconi, et al. 2018).

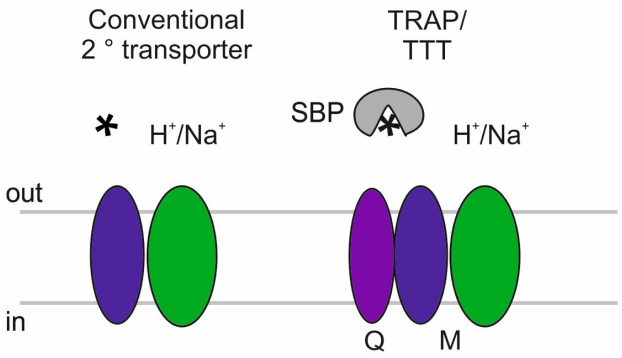


Figure 4 **Conventional and TRAP/TTT secondary transport systems.** Conventional secondary transport systems contain a TMD and commonly use Na^+- or H^+-gradients for substrate translocation. TRAP and TTT transport systems contain two membrane domains, named Q and M and depend on a SBP. The scaffold domains are displayed in magenta and purple and the transport domains in green. The membrane is illustrated with two lines and is located between periplasm (out) and cytoplasm (in). The substrate is depicted as a star.

Within the family of the TRAP transport systems three subfamilies evolved, the TRAP-TRAP, TRAP-TAXI and TRAP-TPAT transporter. To reduce the complexity of the naming, we will hereafter use the terms TRAP, TAXI and TPAT. The membrane domains of the TTT transport systems evolved independently from the TRAP transport systems thereby sharing no sequence similarities (Winnen, Hvorup, and Saier 2003). In general, TRAP and TTT SBPs are both found in bacteria and those of the TRAP subfamily TAXI additionally in archaea (Mulligan, Fischer, and Thomas 2011). TTT SBPs are often overrepresented in bacteria which implies that the same membrane domains are most likely associated with multiple SBPs (L. T. Rosa, Bianconi, et al. 2018). In contrast to this, thus far, no TTT SBPs have been found in archaea which could correspond to the TTT membrane domains present in archaea, such as TctA. This lack of common TTT SBPs implies that the tripartite systems evolved after the divergence of archaea and eukaryotes from bacteria (Winnen, Hvorup, and Saier 2003).
While all tripartite SBP dependent transporter systems generally consist of the three domains P, Q and M, the gene order and associated components differ (Figure 5). A fusion of the SBP with the Q membrane domain is rare but found in the *Rhizobium* sp. strain (Kelly and Thomas 2001). Also, the downstream fusion of Q to M is less frequently observed but present for example in *Roseovarius* sp. 217 (Mulligan, Kelly, and Thomas 2007). In the TRAP subfamily TAXI fusions of the membrane domains Q and M are common (L. T. Rosa, Bianconi, et al. 2018). The TPAT transport systems involve the additional T component, which most likely delivers the substrate to the SBP P which would otherwise not be able to pass the periplasm (Brautigam et al. 2012). Proteins belonging to the TTT family have an additional gene nearby, named D and located opposite to the TTT-PQM operon. The deletion of the additional gene leads to a loss of TTT facilitated transport as it encodes a transcriptional activator relevant for the expression of the TTT transporter (Widenhorn, Somers, and Kay 1989).

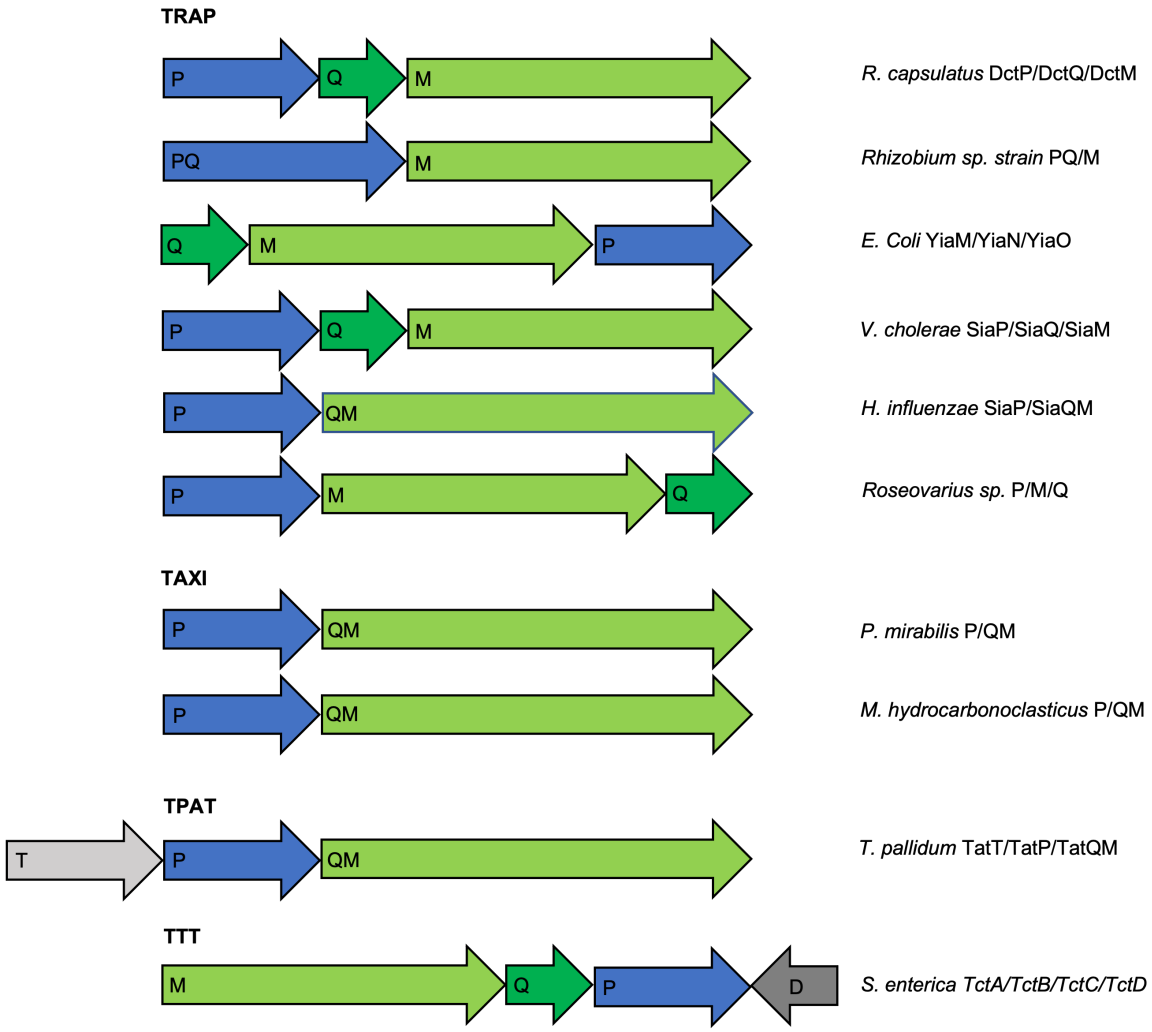


Figure 5 **Different configurations of the gene order in tripartite ATP independent transport systems.** The genes are organized differently in the tripartite transport systems. In TRAP transport systems, arrangements are found in which the SBP P and the membrane domains Q and M appear in all orders, with QM sometimes occurring fused, while P and Q rarely fuse. Especially in the TRAP subfamily TAXI, Q and M appear fused. The TRAP subfamily TPAT has an additional component, the T protein. The TTT transport systems also have an additional component, named D, which is encoded in the opposite direction to the P, Q and M genes. Figure adapted from Kelly and Mulligan (Kelly and Thomas 2001; Mulligan, Kelly, and Thomas 2007).

The overall architecture of the membrane domains Q and M appears highly conserved within the TRAP family, whereas different architectures exist for the SBPs (Figure 6). The SBPs are absolutely required for transport and the sole determinant for substrate specificity (Mulligan et al. 2009). One explanation for the variety of SBPs might be that they had to evolve to be able to bind a broad spectrum of substrates (Brautigam et al. 2012). Variations have evolved either due to convergent evolution (Fischer et al. 2010; L. T. Rosa, Bianconi, et al. 2018) or orthology and divergence (Winnen, Hvorup, and Saier 2003; L. T. Rosa, Bianconi, et al. 2018). The two subfamilies TAXI and TPAT were named according to differences in the SBPs of the TRAP transport systems. The conventional TRAP SBP such as SiaP have a distinctive long α-helix that spans both domains. These SBPs form the group E-I according to the classification of Scheepers *et al.* (Figure 6, A) (Scheepers, Lycklama a Nijeholt, and Poolman 2016). The SBPs of the TTT transporting systems appear structurally different with a shorter α-helical hinge, thus belonging to E-II (Figure 6, B) (Scheepers, Lycklama a

Nijeholt, and Poolman 2016). In general, TTT transporter differ from TRAP transport systems in that they share no sequence similarities within the TMDs and in that they have an additional gene in the operon, which encodes for a transcriptional regulator of TTT expression (Widenhorn, Somers, and Kay 1989). However, the third group describes the TRAP subgroup TAXI whose SBPs have not yet been officially assigned, but Clustal Ω alignments suggest similarities to the SBP of the TTT family (Figure 7) (Sievers et al. 2011; Lemoine et al. 2019). The SBPs of the TAXI transport systems have a long hinge which connects domains I and II instead of the long α-helix in TRAP transport systems, thereby allowing more flexibility between the open and closed conformation (Figure 6 C). TPAT transporting systems have also not yet been officially assigned to any group, but their SBPs seem similar to the TRAP SBPs (Figure 6, D). From an evolutionary point of view, this group probably developed to be able to shuttle hydrophobic substrates through the aqueous environment of the periplasm (Brautigam et al. 2012).

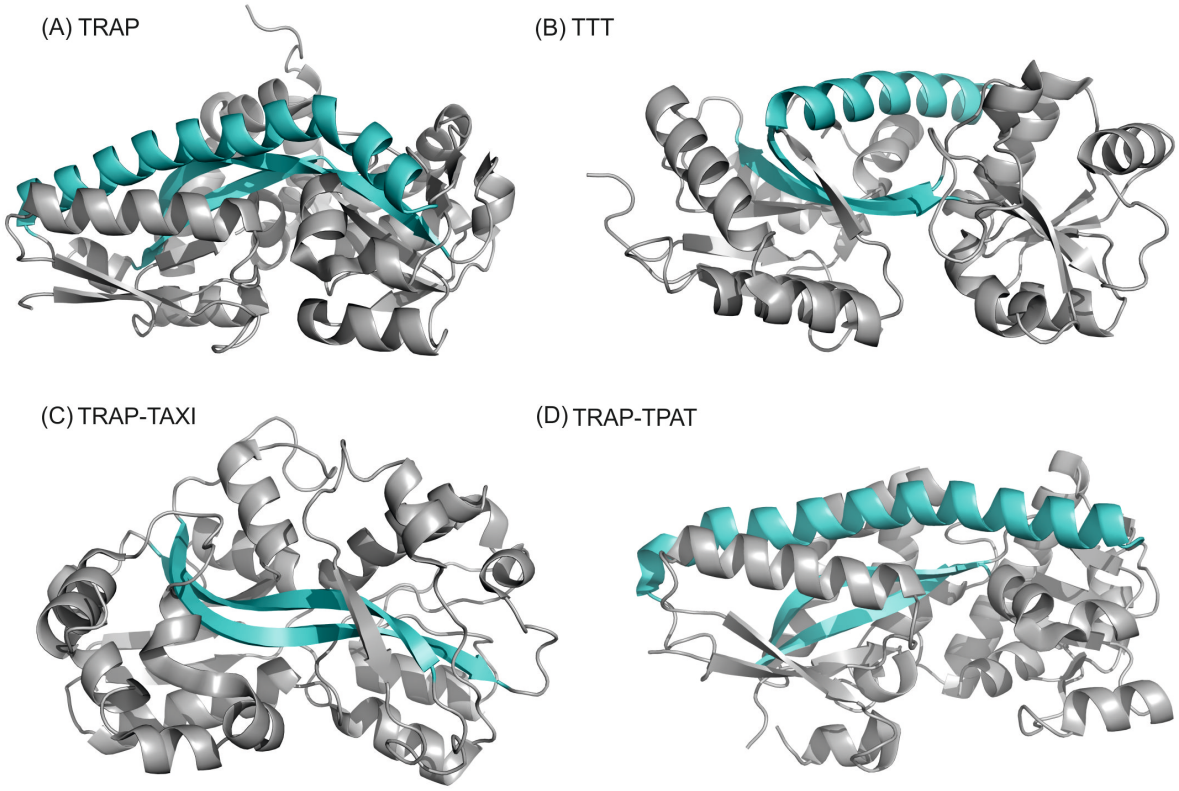


Figure 6 **Structures of the SBPs within the TRAP family.** (A) Structure of the N-acetylneuraminic acid binding TRAP-SBP SiaP from *Haemophilus influenzae* (PDB code: 2wxv) with the characteristic long α-helix and two β-sheets spanning both domains of the SBP. (B) Structure of the citrate binding TTT-SBP TctC from *Polaromonas sp.* Js666 (PDB code: 4X9T) containing a shorter α-helix than TRAP SBP. (C) Structure of the glutamate/glutamine binding TAXI-SBP from *Thermus thermophilus* (PDB code: 1us4) containing a long hinge which connects domain I and II instead of the long α-helix in TRAP transport systems. (D) Structure of the TPAT-SBP TatP from *Treponema pallidum* which is similar in structure to a TRAP SBP but functions together with a T protein to shuffle the substrate through the periplasm. The ligand is unknown. (PDB code: 4di4). The characteristic long α-helices and connecting β-sheets spanning both domains are highlighted in teal. The structures were created using Pymol 1.8.4.1 (Schrödinger, LLC).

Within the tripartite transport systems, the protein sequences of TRAP and TPAT cluster together, the same as they do in TTT and TAXI, whereby the protein sequences of the TRAP SiaP from *Haemophilus influenzae* and the TAXI GluBP from *Thermus thermophilus* show very low sequence similarity (Figure 7).

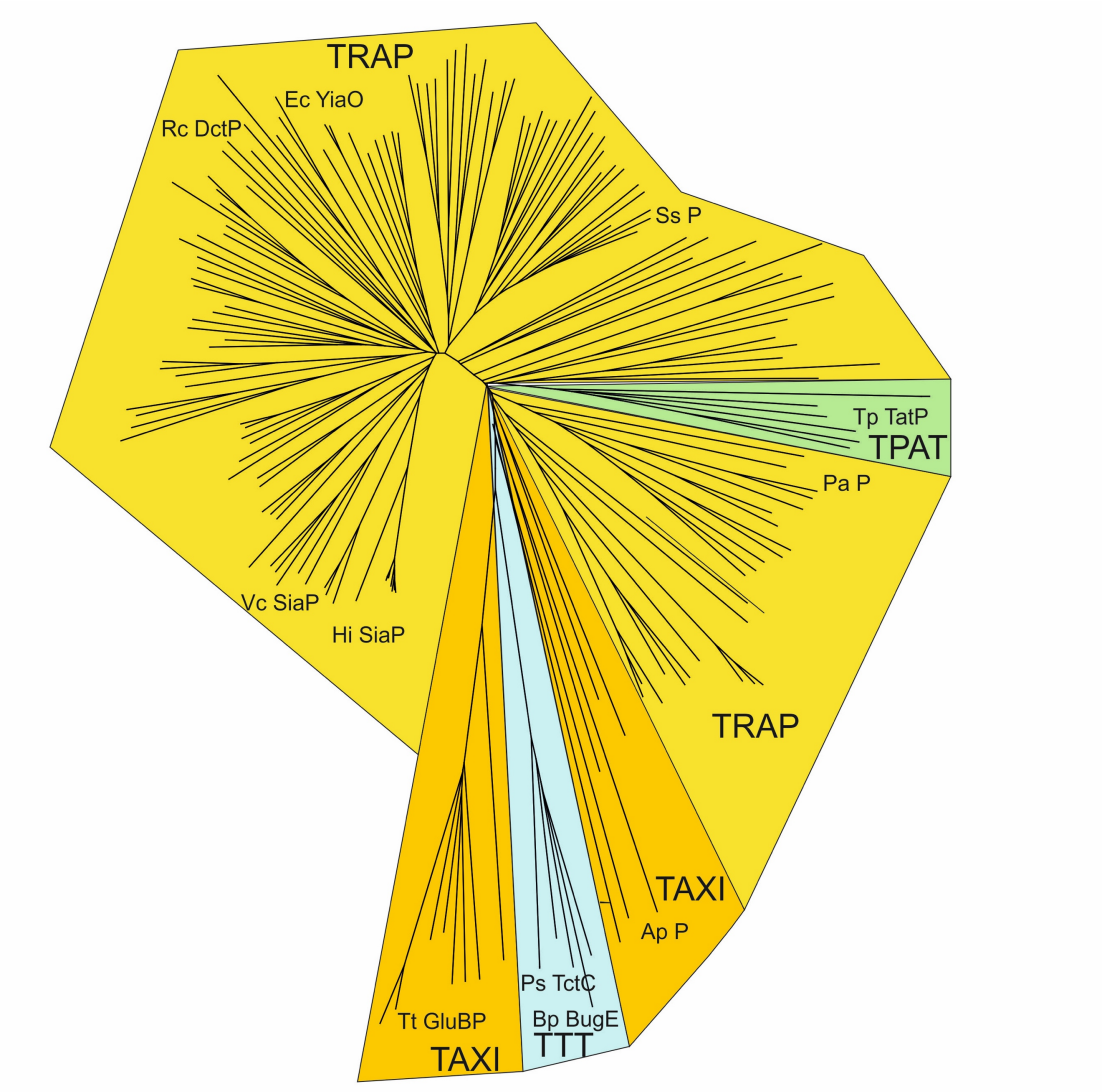


Figure 7 **Phylogenetic tree of SBPs.** The phylogenetic tree was created on the basis of SBP sequence alignments using Clustal Ω. Two clusters form, one containing the TRAP and TPAT SBPs and the other containing the TTT and TAXI SBPs. Abbreviations are as follows: *Haemophilus influenzae* (Hi), *Vibrio cholerae* (Vc), *Rhodobacter capsulatus* (Rc), *Escherichia coli* (Ec), *Sulfitobacter sp.* (Ss), *Treponema pallidum* (Tp), *Pseudomonas aeruginosa* (Pa), *Aeropyrum pernix* (Ap), *Bordetella pertussis* (Bp), *Polaromonas sp.* (Ps), *Thermus thermophilus* (Tt). The sequence alignment was performed using Clustal Ω (Sievers et al. 2011) and the phylogenetic tree was created using MEGA (Hall 2013).

Some SBPs appear to be very specific while others allow the binding of similar substrates (L. Rosa et al. 2019; L. T. Rosa et al. 2017). MatC from a TRAP transporter in *Rhodopseudomonas palustris* shows a high specificity for malate as substitutions can only be made to the second C atom (L. Rosa et al. 2019). In contrast, other SBPs are less specific, such as Bug27 from a TTT transporter in *Bordetella pertussis*, which binds nicotinamide, benzoate and citrate in addition to nicotinate, or AdpC of a TTT transporter in *Rhodopseudomonas palustris* which also binds a broader range, e.g. dicarboxylates with six to nine carbons, for example adipate (L. T. Rosa et al. 2017).

1.3 TRAP, TRAP-TAXI, TRAP-TPAT and TTT transport systems

TRAP transport systems and both subclasses, the TRAP-TAXI and TRAP-TPAT proteins, have not yet been fully investigated. Detailed *in vitro* characterization is limited to the N-acetylneuraminic acid transporter TRAP-SiaPQM from *Haemophilus influenzae*, *Photobacterium profundum* and *Vibrio cholerae* (Davies, Currie, and North 2023; Mulligan et al. 2009, 2012). Bioinformatic analysis shows that both TRAP and TTT systems comprise a similar overall architecture of two membrane domains and one SBP. However, they form two different families as they do not show significant identities in the membrane domains. Recently, the structures of the first TRAP-QM membrane domains, TRAPHi-SiaQM from *Haemophilus influenzae* and TRAPPp-SiaQM from *Photobacterium profundum* were determined (Peter et al. 2022; Davies, Currie, and North 2023). The structure of the SBP TRAPHi-SiaP from *Haemophilus influenzae* was resolved over a decade ago (Johnston et al. 2008), and structures of multiple TRAP SBPs are known (Vetting et al. 2015). Thus far, no structure of any TRAP SBP and membrane domain in complex has been resolved, and no structure of a TTT membrane domain has been determined. In general, similar architectures can be seen in the topologies of TRAP-TRAP, TRAP-TAXI, TRAP-TPAT and TTT systems. The number of predicted TM-segments of the M-domains varies, with 17 for TRAPHi-SiaQM, 18 for TTTPs-TctAB, 18 for TAXIPm-QM and 20 for TAXIMh-QM (Figure 8). However, the – at the time – predicted number of TM-segments with 17 was recently shown to resemble 16 with solving the structure of TRAPHi-SiaQM (Peter et al. 2022). TRAP-TPATs have not been sufficiently annotated so far, so that no complete system has thus far been identified. The four examples of TRAPHi-SiaQM, TTTPs-TctAB, TAXIPm-QM and TAXIMh-QM contain fused Q- and M-domains which all have at least one additional helix between the cytoplasmic C-terminus of the Q-domain and the periplasmic N-terminus of the M-domain. The difference in the number of TM segments results in the N-terminal end of the M-domains located either periplasmic or cytoplasmic (Figure 8), which may be of functional importance. The large membrane domain TctA from the TTT system TctCBA is reported to contain a varying number of TM-segments throughout different organisms, ranging from nine to 12 in bacteria, and seven to 11 in archaea (Winnen, Hvorup, and Saier 2003). Since repetitive conserved motifs were found, it was assumed that the TctA containing 12 TM segments had an early precursor containing six TM segments that was duplicated (Winnen, Hvorup, and Saier 2003). As so far, no TRAP-TAXI systems have been characterized in detail, in this study, the screening for the identification of biochemical superior homologs was focused on TRAP-TAXI transporter.

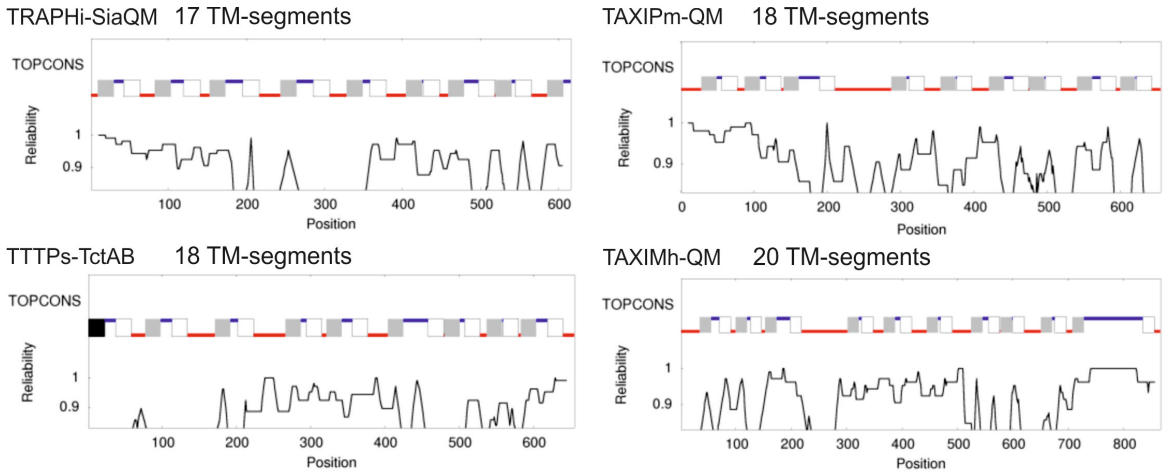


Figure 8 **Topcon analysis of TRAP, TTT and TAXI membrane domains.** To evaluate the difference in topology in TRAP, TTT and TRAP-TAXI transport systems, a Topcon analysis was performed (Tsirigos et al. 2015). The amino acid sequences of the QM-domains were used from *Haemophilus influenzae* (TRAPHi-QM), Polaromonas sp. JS666 (TTTPs-TctAB), Proteus mirabilis (TAXIPm-QM) and Marinobacter hydrocarbonoclasticus (TAXIMh-QM), respectively. The x-axis shows the amino acid positions and the y-axis the reliability of the TOPCON prediction. The reliability is illustrated in a black line. Black, grey and white boxes resemble transmembrane segments. Protein sequences predicted to be in the cytoplasm are depicted in red lines and protein sequences predicted to be in the periplasm are depicted as blue lines.

1.3.1 Tripartite ATP independent periplasmic (TRAP) transporter family

The first discovered SBP dependent tripartite ATP independent periplasmic system was the C_4-dicarboxylate TRAP transporter DctPQM from *Rhodobacter capsulatus* (Jacobs et al. 1996). By using the SBP, it increases its affinity to L-malate, in this case to 50 nM (L. T. Rosa et al. 2017). Besides C_4- and C_5-dicarboxylates, mostly amino acids or sugars are transported by TRAP transport systems (Vetting et al. 2015). The only TRAP transporter found in *E. coli*, YiaMNO, was found to transport 2,3-diketogulonic acid (Mulligan, Fischer, and Thomas 2011). SiaPQM from *Haemophilus influenzae*, is the model protein for the TRAP family in which the SBP SiaP binds N-acetylneuraminic acid with nanomolar affinity and delivers it to the membrane domain SiaQM. With the transport of N-acetylneuraminic acid, SiaPQM ensures that the pathogen is no longer recognized by the immune system as *Haemophilus influenzae* uses it to sialylate the lipopolysaccharide on the cell surface. SiaPQM which is a contributing factor in severe diseases such as respiratory tract infections and meningitis is therefore clinically relevant (Skoczyńska et al. 2005; Severi et al. 2005). As SiaPQM is a secondary transporter, it was expected to catalyze bidirectional substrate transport like other secondary transporting systems. However, as it was demonstrated to facilitate substrate import only, it operates unidirectionally (Mulligan et al. 2009). Only when unphysiologically high concentrations of the SBP are present can the export of the substrate be induced, as SiaP can form a complex with SiaQM, regardless of whether with N-actylneuraminic acid is in the binding pocket or not. The SBP therefore determines whether import or export occurs (Mulligan et al. 2009).

SiaP is a monomeric protein that is composed of two globular domains that consist of both ß-sheets and α-helices whereby one α-helix is elongated thereby spanning both domains (Figure 9, A). A cleft is formed between the two domains, in which the substrate N-acetylneuraminic acid binds (Figure 9, A). The salt bridge between the positive amino acid Arg147 and the carboxylate group of the substrate is essential for substrate coordination and therefore strongly conserved in TRAP SBPs (Figure 9, B) (Mulligan, Fischer, and Thomas 2011).

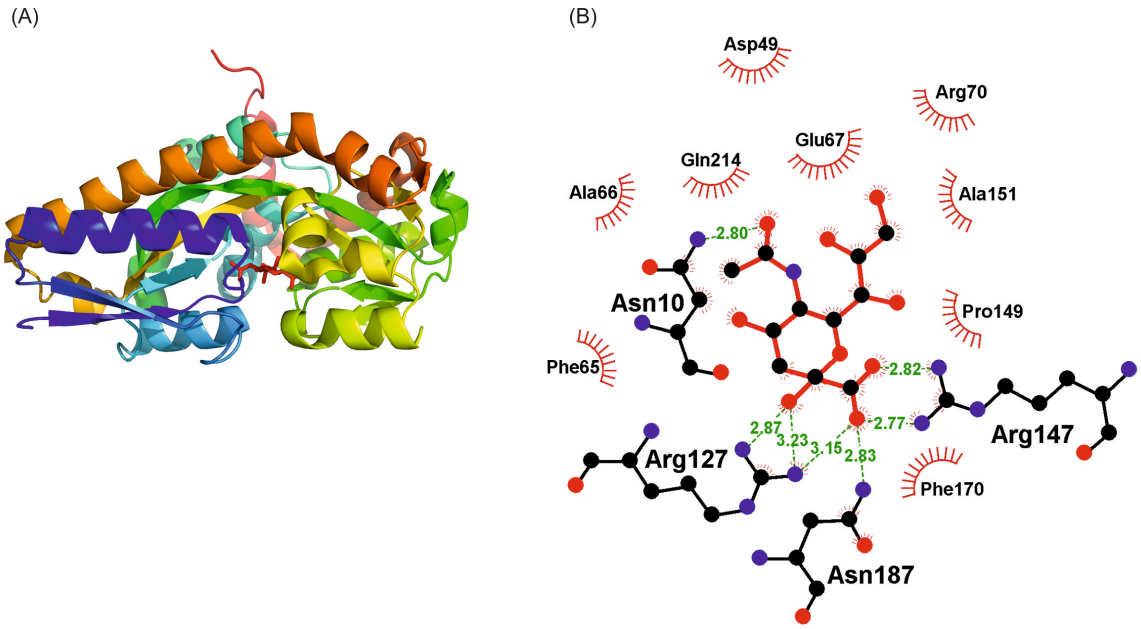


Figure 9 **Structure of the N-acetylneuraminic acid binding protein TRAPHi-SiaP**. (A) The structure of the N-acetylneuraminic acid binding protein SiaP (PDB code: 2WXV) is shown in rainbow. (B) The amino acid residues directly assumed to be involved in the binding are shown as molecules and depicted in black letters. Hydrogen bonds are depicted as a green line, their distances to the docked ligand are indicated in Å. Neighboring hydrophobic amino acid residues are depicted in black letters next to red semicircles. Carbon atoms are shown in black, oxygen atoms in red and nitrogen atoms in blue. The structure was created using Pymol 1.8.4.1 (Schrödinger, LLC) and the substrate binding analysis using LigPlot+ (Laskowski and Swindells 2011).

Very recent structural data of the membrane domain SiaQM revealed that TRAP transporters belong to elevator-type proteins (section 1.3.6) (Peter et al. 2022). An elevator type protein needs a firm anchoring in the membrane facilitated by the scaffold domain to ensure that the transport domain can translocate substrates across the membrane in an elevator-like manner (Drew and Boudker 2016). The assumption that SiaQM uses the elevator type transport mechanism is based on the high structural similarity of the transport domain SiaM with the elevator-type sodium-coupled dicarboxylate transporter VcINDY (Peter et al. 2022). Prior to the structural analysis of SiaQM, it was believed that all elevator proteins oligomerize as di- or trimers like the dimeric VcINDY (Peter et al. 2022; Holzhueter and Geertsma 2020). The oligomerization is in some cases also relevant for biogenesis as, for example, for the UapA purine transporter from *Aspergillus nidulans* it is essential for correct membrane trafficking (Holzhueter and Geertsma 2020). Additionally the oligomerization most likely ensures a stable anchoring of the protein in the membrane to allow the elevator-like up and down movement of the transport domain (Peter et al. 2022). However, SiaQM is the first reported monomeric elevator type protein. Instead of oligomerization, SiaQM uses the Q domain as structural extension of the scaffold domain in M to provide stability for the whole protein complex (Davies, Currie, and North 2023). There is no clear explanation why the SiaQM membrane domain is composed of a heterodimer of the two units Q and M, rather than a homodimer of SiaM only. The existence of a heterodimer poses the question of whether Q is also functionally needed for its structural task. SiaQM from *Haemophilus influenzae* consists of 16 transmembrane helices of which four are assigned to the smaller membrane domain Q, 11 to the large membrane domain M and one to the connecting unit between Q and M (Figure 10). Within the transport domains, two helical

hairpins are found which serve as molecular gates thereby coordinating substrate and Na^+-ions (Figure 10) (Peter et al. 2022).

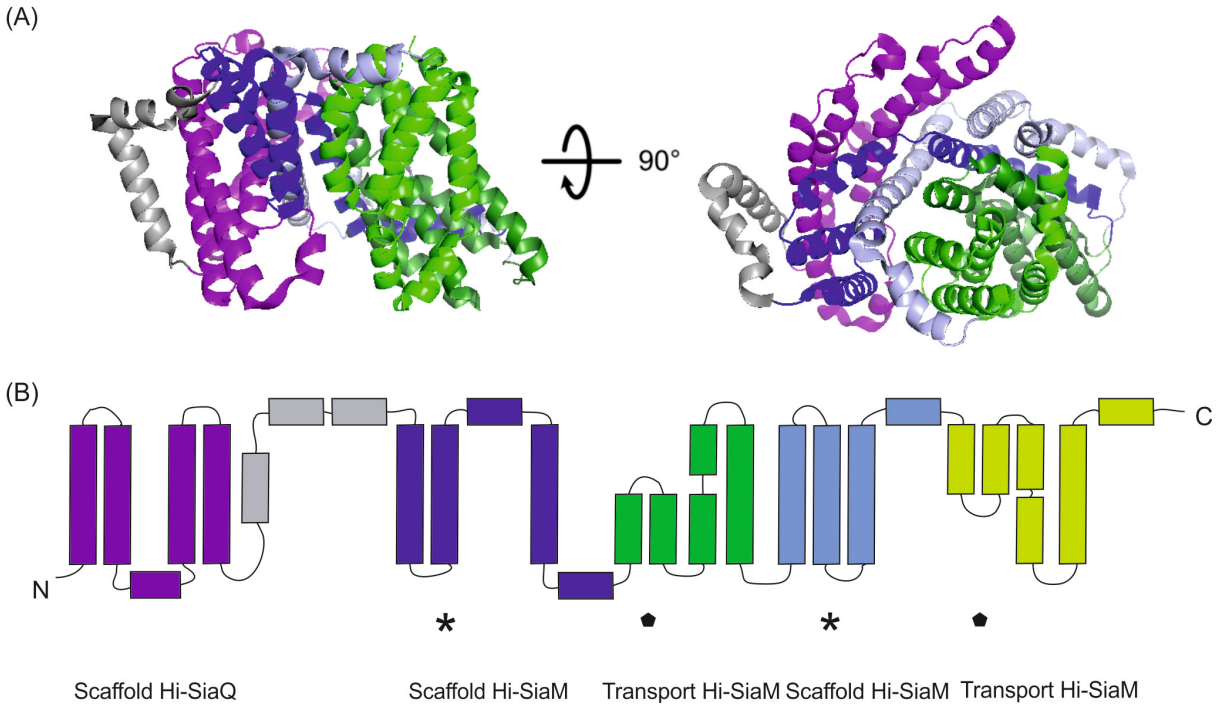


Figure 10 **Structure and topology plot of TRAPHi-SiaQM.** (A) Structure of TRAPHi-SiaQM showing a view on SiaQM in the plane of the membrane (left panel) and the top view (right panel; PDB code: 7QE5). Q is depicted in magenta, the connector unit between Q and M in grey, the scaffold domains 1 and 2 in dark purple and light blue and the transport domain 1 and 2 in dark and light green, respectively. (B) Topology plot of TRAPHi-SiaQM based on structure, adapted from (Peter et al. 2022). The same color scheme for the different domains is used as in A. The stars mark the two barrier helices, which are a recurring feature in the scaffold domains of elevator proteins. The hairpins in the transport domains are marked with octagons (Peter et al. 2022). The structures were created using Pymol 1.8.4.1 (Schrödinger, LLC).

1.3.2 TRAP-associated extracytoplasmic immunogenic (TAXI) proteins

TAXI transport systems are a subfamily of the TRAP family, named after an immunogenic protein from *Brucella* (Mayfield et al. 1988; L. T. Rosa, Bianconi, et al. 2018). The TRAP-TAXI proteins have a distinct class of SBPs (Kelly and Thomas 2001) and their membrane domains Q and M are predominantly fused (L. T. Rosa, Bianconi, et al. 2018). GluBP from Thermus is a representative of this group for which the structure is known (Takahashi et al. 2004), but this protein has not been functionally characterized. Bioinformatic analysis revealed that TAXI transport systems appear more frequently in archaea than bacteria (Mulligan, Fischer, and Thomas 2011; Kelly and Thomas 2001). TRAP-TAXI SBPs appear also particularly often membrane-tethered which allows for high local concentrations at the membrane (Doeven et al. 2004). Additionally it prevents diffusion out of the cell in archaea and Gram positive bacteria (Mulligan, Fischer, and Thomas 2011).

1.3.3 Tetratricopeptide repeat protein associated transport (TPAT) systems

When compared to TRAP transporters, the tetratricopeptide repeat protein associated transport systems (TPATs) have an additional T-component, which is believed to channel substrates from the outer membrane to the SBP, enabling them to move hydrophobic substrates through the aqueous

periplasmic space and ultimately allowing their transport. This subgroup might have evolved to enable transport of a wide range of substrates (Brautigam et al. 2012). Their membrane components Q and M appear fused or separated. The first structure of a TPAT transporter was determined by crystallization of the T and P complex from the tatTPQ/M of *T. pallidum* in which three monomeric TatPs can bind to a trimeric TatT. In studying the structure of the whole complex, the cleft of TatP aligns very well with the pore of TatT, so it was assumed that they interact with one another to transfer the substrate (Brautigam et al. 2012). Both proteins contain lipid moieties, so it was concluded that both components are lipid-anchored, most likely TatT to the outer and TatP to the inner membrane (Figure 11). The significantly high proportion of hydrophobic amino acids located inside the cleft of TatP suggests that the transported substrates are hydrophobic (Brautigam et al. 2012). No TPAT system has been functionally characterized.

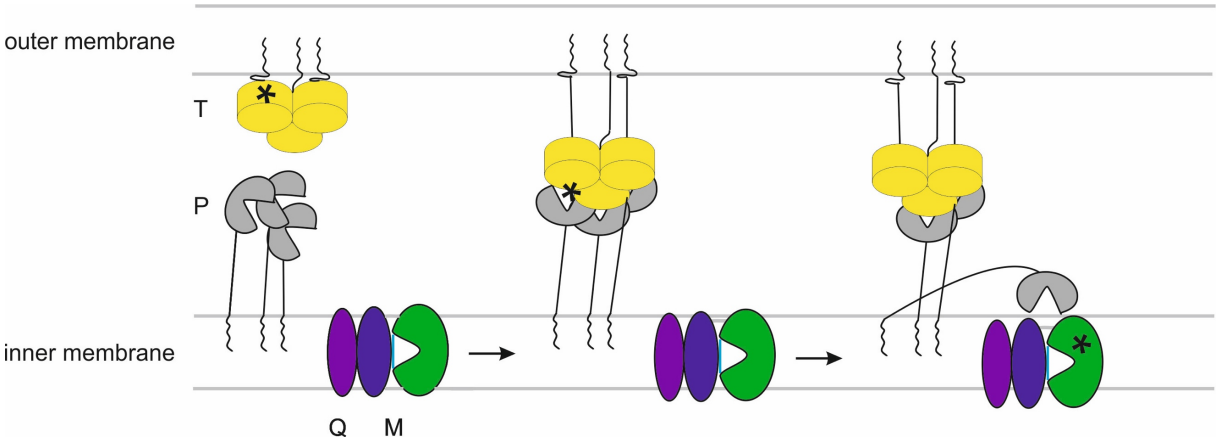


Figure 11 **Schematic illustration of a TPAT transporter.** The Tetratricopeptide repeat protein associated transporter consists of a T-component, the SBP P and the membrane domain QM. The T and P components are believed to be anchored to the outer and inner membrane, respectively. They are believed to facilitate movement of hydrophobic substrates through the periplasmic space to the membrane components Q and M which enable transport over the inner membrane. The figure was adapted from Brautigam et al. (Brautigam et al. 2012).

1.3.4 Tripartite tricarboxylate transport (TTT) systems

TTT systems form their own family having a similar assembly as TRAP transport systems and their subfamilies but significant sequence differences in the membrane domains. So far, TTT systems have commonly been linked to the transport of carboxylates. The model protein for the tripartite tricarboxylate transporting systems group is the citrate transporter TctCBA from *Salmonella typhimurium* (Widenhorn, Somers, and Kay 1989). When compared to TRAP transport systems, TTT systems have an additional D-component in the gene neighborhood, located opposite to the TTT-PQM operon. The deletion of the additional gene leads to a loss of TTT facilitated transport as it encodes a transcriptional activator relevant for the expression of the TTT system (Widenhorn, Somers, and Kay 1989). The SBP TctC appears in a dimeric form and showed dependence on Na^+ for citrate binding which resulted in a Na^+ dependence for the transport of fluorocitrate, citrate and isocitrate by TctCBA next to a dependence on membrane potential with K_m-values of 3, 50 and 170 µM, respectively (Winnen, Hvorup, and Saier 2003). Homologs of the large membrane domain of TTT proteins were found in both bacteria and archaea, whereas the whole tripartite systems were found in bacteria only. The sole existence of the large membrane domain in ancestors explains

higher similarities in the membrane domains compared to the SBPs which evolved later independently (Winnen, Hvorup, and Saier 2003; Batista-García et al. 2014). A noticeable number of SBPs were found in *Bordetella pertussis* (Antoine et al. 2003), many of which could not be clearly assigned to a transporter, suggesting that several SBPs may deliver to the same transmembrane domains (L. T. Rosa et al. 2017). Structures of a number of TTT SBPs have been solved, but functional details are outstanding (L. T. Rosa, Bianconi, et al. 2018).

1.3.5 Transport mechanism in primary transport systems

With more than 2000 members, ATP-binding cassette (ABC) transport systems represent the largest group within primary active transport systems for substrate translocation (Oldham, Chenb, and Chen 2013). They work unidirectionally as either ABC importers or exporters, whereby the importing systems commonly involve SBPs to determine the transport direction (Hvorup et al. 2007). The transport of SBP dependent ABC importers starts in the periplasm with the high affinity binding of the substrate to the SBP which thereby changes its conformation from an open to a closed conformation. The liganded SBP then binds to the membrane domain which was, in the absence of a ligand-bound SBP, in the inward-facing state with two separated nucleotide-binding domains (NBDs). After the binding of the ligand-bound SBP the NBDs can cooperatively bind ATP which leads to a reorientation of the membrane domain to the outward-facing state with subsequent substrate transfer. The reorientation reduces the space between the NBDs and enables ATP hydrolysis, followed by dissociation of the NBD dimer which induces conformational changes from outward to inward facing thereby releasing the substrate into the cytoplasm (Figure 12) (Oldham, Chenb, and Chen 2013; Lewinson et al. 2010).

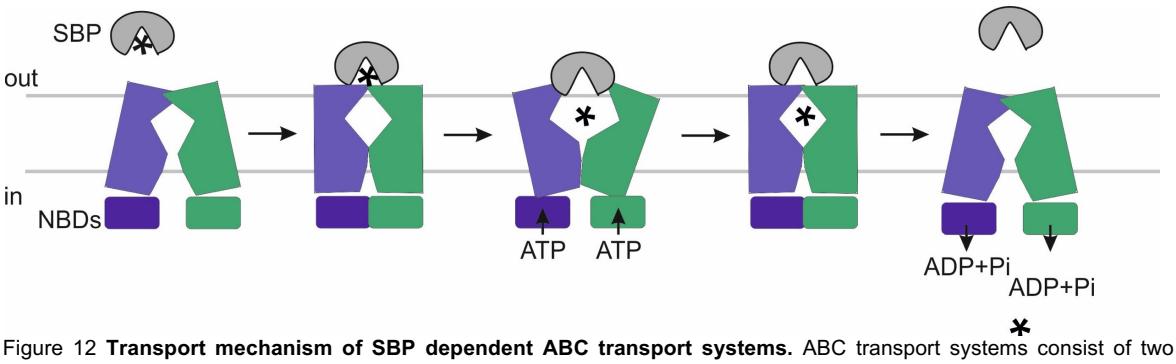


Figure 12 **Transport mechanism of SBP dependent ABC transport systems.** ABC transport systems consist of two architecturally identical transmembrane domains with two nucleotide binding domains (NBDs) attached on the cytoplasmic site. For substrate translocation, some ABC transport systems are SBP dependent. In all cases ATP is hydrolyzed into ADP+P_i. The transport systems are embedded in a membrane, which is illustrated with two lines. The SBPs are in the periplasm (out) and the nucleotide binding sites (NBDs) in the cytoplasm (in). The substrate is depicted as a star. Figure adapted from Oldham and Hofmann (Oldham, Chenb, and Chen 2013; Hofmann et al. 2019).

Among ABC transporters, SBP dependent importers are most common, as for example the MBP-MalFGK2 which shows specific unidirectional import of maltose through MalFGK2 (Figure 13). Both the SBP and the membrane domain determine the specificity, but the SBP ensures high affinity for the substrate (Oldham, Chenb, and Chen 2013). While ions and monosaccharides freely pass porins in the outer membrane, malto-oligosaccharides are transported by dedicated maltoporin. It is then the SBP that binds maltose like a cap to the cavity of the membrane domain to ensure unidirectional

transport (Oldham et al. 2007). The complex is then stabilized in an energetically higher conformation which initiates ATP hydrolysis (Shilton 2008). BtuCD-BtuF is also a unidirectional SBP dependent importer (Figure 13); it transports vitamin B12. Within the transport cycle, ATP binding facilitates the conformational change of the membrane domain to the outward-facing state which allows the transfer of the ligand from BtuF to the binding cleft of BtuCD which is subsequently closed (Yang et al. 2018; Hvorup et al. 2007).

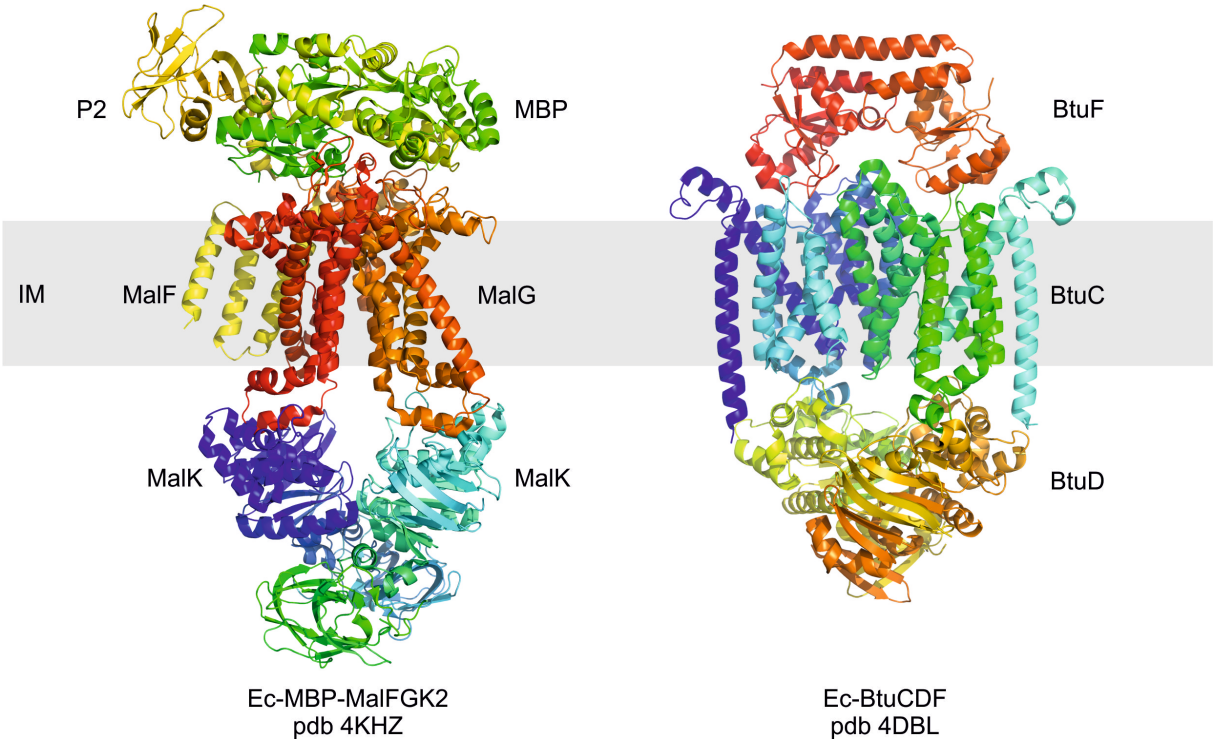


Figure 13 **Structure of MBP-MalFGK2 and BtuCDF.** Structure of the maltose binding protein MBP attached to the maltose transporter complex MalFGK2 (left) and the vitamin B12 binding protein BtuF attached to the transporter BtuCDF (right) in rainbow (PDB codes: 4KHZ, 4DBL, respectively). The transport systems are embedded in the inner membrane (IM), which was approximated and shaded in grey. The structures were created using Pymol 1.8.4.1 (Schrödinger, LLC).

Generally ABC transporters operate unidirectionally, but some allow for bidirectional transport when an excess of substrate is present (Hosie et al. 2001). Aap from *Rhizobium leguminosarum* and the His transporter from *S. typhimurium* are both SBP dependent ABC transporters that facilitate uptake and efflux of amino acids. Kinetic assays showed high affinity uptake of solutes compared to low affinity efflux as the latter was only possible after the addition of a substrate excess (Hosie et al. 2001). Interestingly, TAP which is a unidirectional SBP independent ABC peptide importer was converted into a bidirectional transporter. Efflux was not observed when the outward substrate concentration gradient was generated by preloading proteoliposomes with substrate, but a single mutation that prevented ATP-binding enabled both import and export. This result suggests that ATP binding is needed to set the direction of transport, but does not facilitate the transport in general (Grossmann et al. 2014). Next to ABC importing systems, NaAtm1 from *Novosphingobium aromaticivorans* represents a unidirectional ABC exporter. The transporter undergoes

conformational changes during the transition from the inward-facing to the outward-facing state and vice-versa that affect the substrate binding site which allows binding of glutathione derivatives to the inward-facing conformation only (Fan, Kaiser, and Rees 2020).

1.3.6 Transport mechanism in secondary transport systems

For secondary transport there are three proposed transport mechanisms, the rocker switch, rocking bundle and elevator mechanism (Figure 14) (Drew and Boudker 2016). Rocker switch proteins have two architecturally very similar domains. The domains open and close against each other and thus move the substrate from one side to the other. Rocking-bundle proteins have two different domains, one of which is more strongly fixed in the membrane and the other which can open and close against it. Elevator proteins can be thought of as one domain sitting very firmly in the membrane and the second domain moving up and down like an elevator (Figure 14). Elevator-type proteins consist of a scaffold and a transport domain. The transport domain moves vertically along the scaffold domain, carrying the substrate from one side of the membrane to the other. The task of the scaffold domain is to provide stability and support for the transport domain. As the complete transporter needs a firm anchoring in the membrane, it usually exists as a dimer or trimer. In proteins that use the elevator-type mechanism, an inverted repeat scheme usually appears in their topology (Forrest 2015; Mulligan et al. 2017). These repeats are duplications with inherent functionality which can hence be reversed (Forrest 2015).

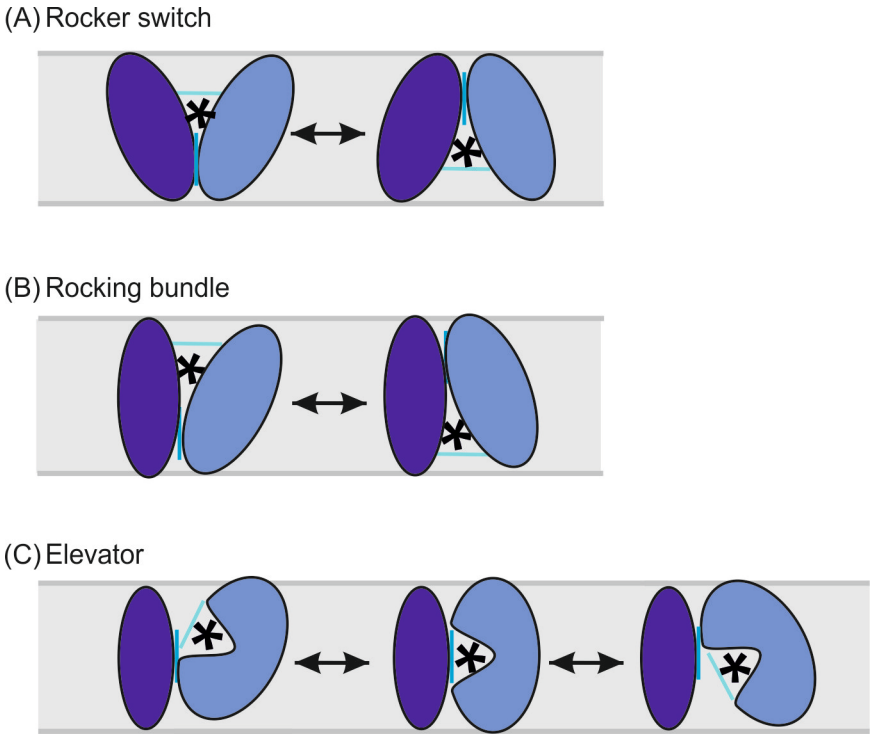


Figure 14 **Transport mechanisms and architecture of secondary transporting systems.** (A) Proteins that use the rocker switch mechanism consist of two architecturally identical domains rocking in opposition to each other. (B) In rocking-bundle proteins the domains are structurally distinct with one fixed in the membrane and the other rocking against it. (C) Elevator-type proteins, like rocking bundle proteins, have structurally distinct domains. The scaffold domain is fixed in the membrane and the transport domain shifts up and down like an elevator. Inside-out and right-side-out conformations are shown here. The substrate is depicted as asterisk. Figure adapted from Drew, Boudker (Drew and Boudker 2016).

One example of the elevator-type proteins is the secondary C_4- and C_5-dicarboxylate transporter VcINDY from *Vibrio cholerae* which is both pH dependent and sodium dependent. In proteoliposome transport assays the additional application of a negative membrane potential led to highest transport rates, indicating that the transport by VcINDY is electrogenic (Mulligan et al. 2014). At least one net positive charge is transported per transport cycle, suggesting that at least two Na^+-ions are co-transported with the single anionic species of the model substrate succinate and at least three Na^+-ions with the di-anionic species (Mulligan et al. 2014). As the other elevator type proteins, VcINDY contains two scaffold and two transport domains. These are separated from each other in the amino acid sequence but positioned spatially proximate to each other. Within the transport domains two helical hairpins are found in some elevator-type proteins that serve as molecular gates, thereby coordinating substrate and Na^+-ions (Figure 15) (Peter et al. 2022).

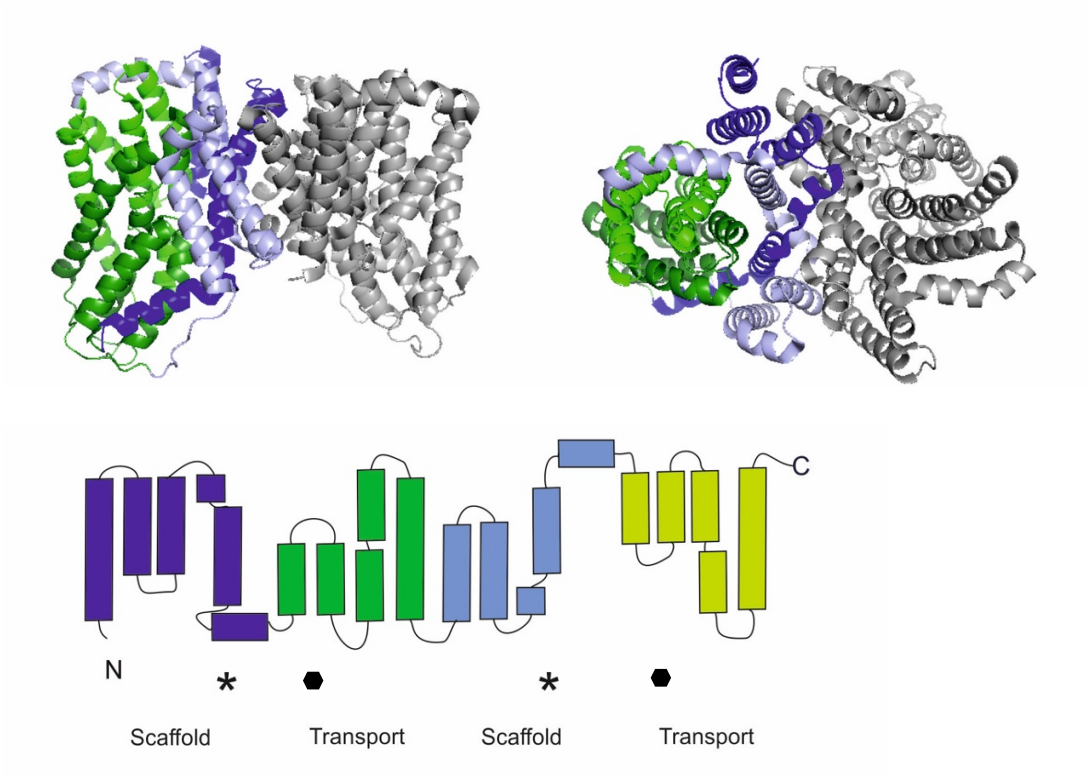


Figure 15 **Architectural design in elevator proteins is exemplarily shown with VcIndy**. The scaffold domains of VCINDY from *Vibrio cholerae* are displayed in dark purple and light blue and the transport domains in dark and light green in both, the structure, and the topology plot (PDB code: 7T9F). The barrier helices, characteristic of elevator type proteins, are marked with asterisks. The hairpins in the transport domains are marked with octagons. The structures were created using Pymol 1.8.4.1 (Schrödinger, LLC).

Most recently, TRAP transporters were assigned to the family of elevator proteins as well, based on the structure of SiaQM from *Haemophilus influenzae* (Peter et al. 2022). In the structural analysis of VcINDY and SiaQM, high similarities were found (Peter et al. 2022). However, in contrast VcINDY and most elevator-type proteins, TRAP transport systems appear monomeric and not dimeric in structure. Another elevator-type protein possibly occurring as a monomer has so far only been reported with CcdA from *Thermus thermophilus*, as the structure was solved as a monomer here (Garaeva and Slotboom 2020). Biochemical data, however, indicate a dimeric structure (Bocian-Ostrzycka et al. 2017). This is most likely required as the oligomerization is essential to form a rigid

scaffold as it was shown for the SLC26 family, for example (Chang et al. 2019). In the case of TRAP transporting systems, the membrane domain contains an additional membrane domain of four transmembrane helices, which ensure a firm anchoring in the membrane, thereby enabling transport like an elevator (Davies, Currie, and North 2023). All other elevator type proteins form dimers or even trimers as in the case of SLC1 such as GltPh and SLC28 such as hCNT3 proteins (Grewer, Gameiro, and Rauen 2014; Stecula et al. 2017). An exception seems to be NhaA which is active as a monomer under most tested conditions. However the functional relevance of the monomer remains unclear, as it was shown that the dimeric NhaA enables higher transport rates when exposed to high salt and pH (Rimon, Tzubery, and Padan 2007). So, structurally, elevator type proteins are highly comparable, but functionally, different transport mechanisms are described. In the SLC26 family, for example, there are transport systems that function as exchanges or uniporters. In addition, some function as chloride channels and one as a motor protein (Rapp, Bai, and Reithmeier 2017; Walter, Sawicka, and Dutzler 2019; Dallos and Fakler 2002). In addition to their monomeric form, the TRAP transporting systems, assigned to the elevator-type proteins, exhibit the additional feature that they depend on an SBP which delivers the substrate to the membrane domain with high affinity and specificity (Figure 16). This mechanism is well characterized for primary ABC transporting systems which couple the use of a SBP and the hydrolysis of ATP for transport. However, TRAP transporter facilitate transport without the hydrolysis of ATP. Instead, they make use of a sodium- or proton-gradient, often combined with membrane potential so that the energy comes from the electrochemical gradient created by pumping ions out of the cell. The SBP P binds the substrate and delivers it to the membrane domain QM which then changes to the outside open conformation followed by the occluded and inside open conformations, resulting in the substrate entering the cytoplasm (Figure 16). While conventional secondary transporting systems work bidirectionally, TRAP transporting systems are believed to import only due to the SBP-dependency, thereby being unidirectional. Bidirectional transport was observed for the TRAP transporter SiaPQM from *Haemophilus influenzae* only under unphysiological conditions when a large excess of SBP was present (Mulligan et al. 2009).

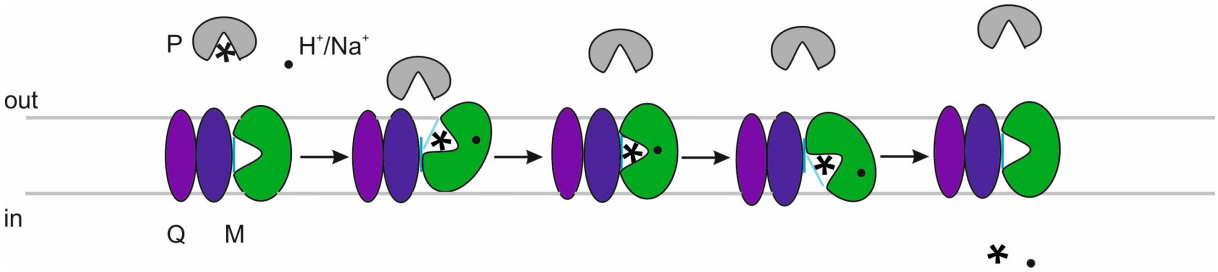


Figure 16 **Transport mechanism of TRAP transporter.** TRAP transport systems use the elevator mechanism to transport their substrate over the inner membrane by symport of at least one sodium ion or proton. The SBP P binds the substrate and delivers it to the membrane domain QM which then changes to the outside open conformation followed by the occluded and inside open conformations, resulting in the substrate entering the cytoplasm. The substrate is depicted as asterisk. Figure adapted from (Mulligan et al. 2009).

So far, no structure of a complete complex of TRAP SBP and membrane domains has been solved, but studies using artificial modelling and comparison of SiaPQM from *Haemophilus influenzae* and VcINDY from *Vibrio cholerae* allow assumptions about the potential interface and residues involved

in the binding of N-acetylneuraminic acid and Na^+ (Figure 17) (Peter et al. 2022). The interface is predicted to span a surface area of approximately 1980 $Å^2$, whereby the N-terminus of SiaP presumably binds the Q- and the C-terminus of SiaP the M-domain (Figure 17) (Peter et al. 2022). The interaction with Q probably predominates, especially with the periplasmic loop of Q between TM3 and TM4, since these contain complementary charges to SiaP (Davies, Currie, and North 2023). The Q and M domains form a tight connection most likely realized through multiple hydrophobic side chains which is not only the case for the fused membrane domain in SiaQM of *Haemophilus influenzae* but also for the unfused version in *Vibrio cholerae* (Peter et al. 2022). The substrate N-acetylneuramic acid has access to the positive residues within the SBP, but the translocation through the membrane is only possible after binding of the SBP to the membrane domain which thereby undergoes a conformational change and presents the residues responsible for substrate and Na^+-binding (Davies, Currie, and North 2023).

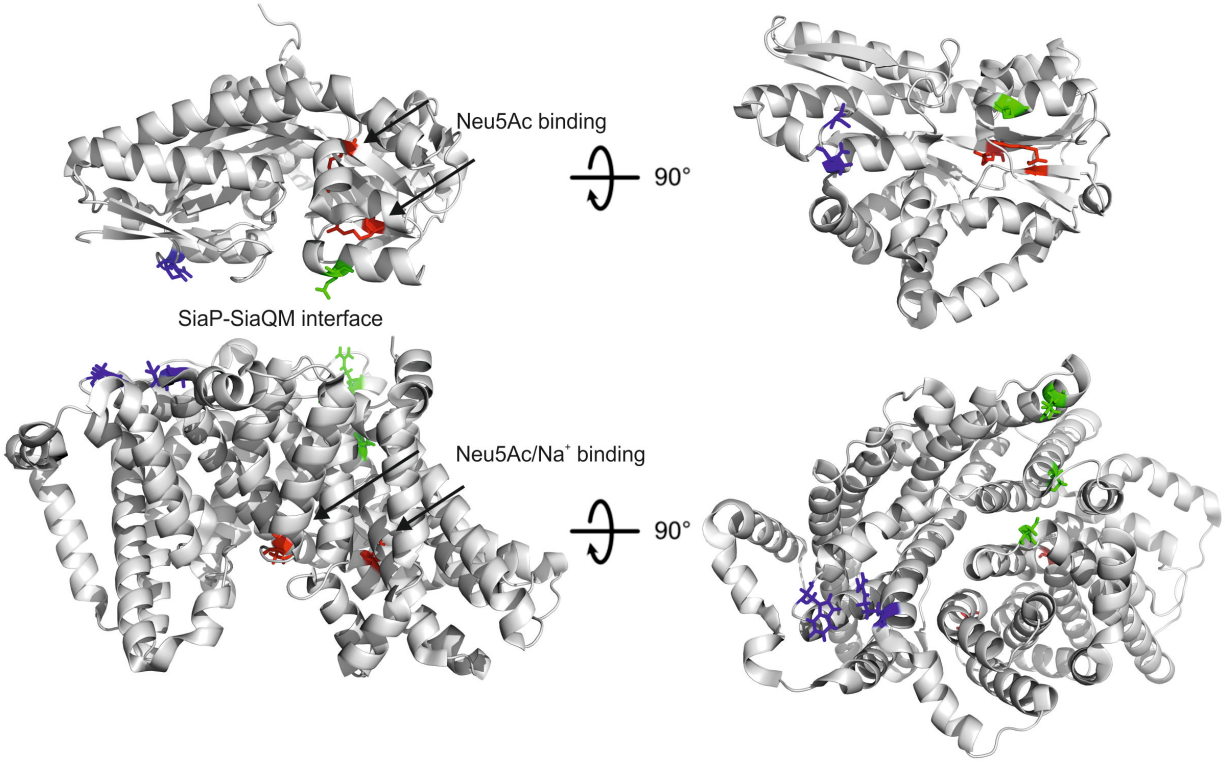


Figure 17 **Structure of SiaP and SiaQM from *Haemophilus influenzae*.** Side and top view of TRAPHi-SiaQM and TRAPHi-SiaP are shown. Residues presumably involved in the TRAPHi-SiaP and TRAPHi-SiaQM interface are highlighted in blue and green and residues involved in binding of N-acetylneuraminic acid (Neu5Ac) and Na^+ are highlighted in red (PDB code: SiaP 2XWV, SiaQM 7QE5) (Peter et al. 2022). The structures were created using Pymol 1.8.4.1 (Schrödinger, LLC).

1.4 Nanobodies

Conventional antibodies are characterized by a Y-shape structure, which is formed by two heavy and two light chains. The antigen binding sites, named paratopes, are located at both ends. From this basic structure, further types of antibodies can be derived. The heavy chain only antibodies, as the name implies, consist only of the two heavy chains and are found in camelids such as llamas, alpacas, and camels and a similar form is found in sharks (Steeland, Vandenbroucke, and Libert 2016). From the heavy chain only antibodies, so-called nanobodies are generated by amplification

of the variable antigen binding domains (Figure 18). They are divided in constant regions next to regions that are responsible for the specific recognition of the antigens (Salvador, Vilaplana, and Marco 2019; Pardon et al. 2014). The complementary determining region 3 is particularly important as it is longer and highly variable compared to the one in conventional antibodies which leads to nanobodies binding even more specifically to unique epitopes (Bannas, Hambach, and Koch-Nolte 2017). Therefore, nanobodies can bind conformational epitopes, whereas antibodies are mostly limited to linear epitopes (Forsström et al. 2015).

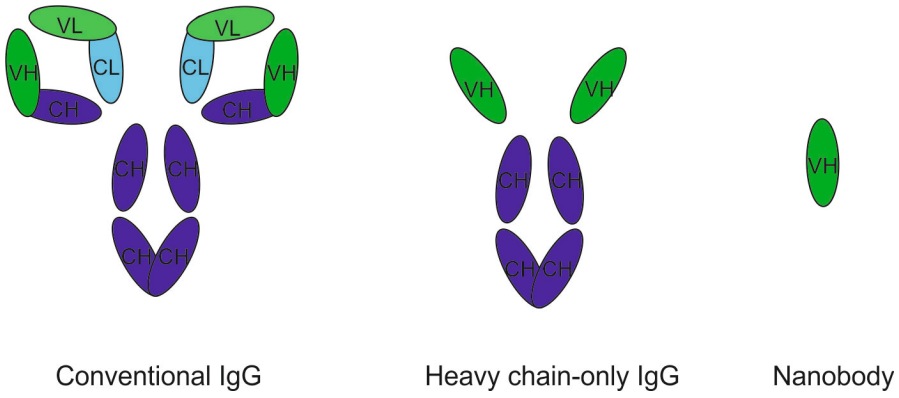


Figure 18 **Schematic illustration of antibody formats.** The conventional antibody IgG, heavy chain-only antibody IgG and nanobody are shown, adapted from Steeland (Steeland, Vandenbroucke, and Libert 2016). Constant heavy chains are labelled as CH, variable heavy chains as VH, constant light chains as CL and variable light chains as VL.

Nanobodies have developed the status of a very popular tool in recent years by being used in different kinds of biochemical assays. They can be selectively generated from the blood of alpacas after immunization with the target protein. A library is constructed and the nanobodies are enriched in two rounds of phage display. Single clones are expressed and nanobodies are selected by ELISA with a subsequent sequence analysis of the CDR regions. The most promising nanobodies are used for larger expression scales in *E. coli* and purified in subsequent steps for use in further assays (Figure 19) (Zimmermann et al. 2018; Salvador, Vilaplana, and Marco 2019)**.** The generation from immune responses of animals is advantageous compared to *in vitro* generation because it is extremely broad and diverse leading to large libraries with high binding affinities and stability; *in vitro* generation on the other hand, often leads to a selection bias. Also, a large-scale expression in *E. coli* uses the standard biochemical repertoire and is low in terms of production costs (Salvador, Vilaplana, and Marco 2019).

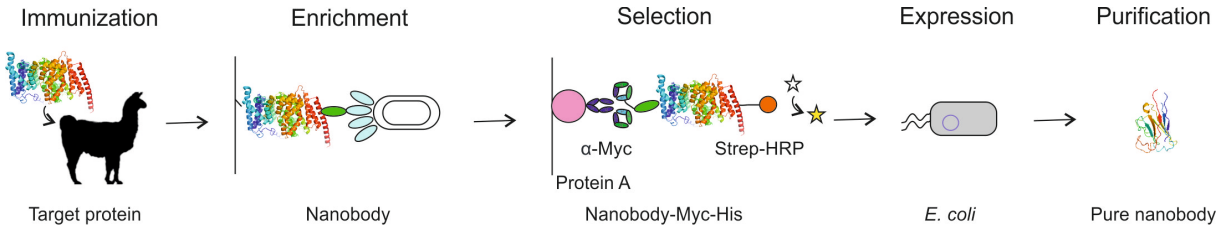


Figure 19 **Schematic illustration of nanobody production.** Alpacas are immunized with the target protein followed by isolation of the lymphocytes from the blood. A library is constructed and the nanobodies are enriched in two rounds of phage display, followed by transformation of *E. coli* cells. Single clones are expressed and nanobodies are selected by ELISA with a subsequent sequence analysis of the CDR regions. The most promising nanobodies are used for larger expression scales in *E. coli* purified and used in further assays. The figure was adapted from Zimmermann and Salvador (Zimmermann et al. 2018; Salvador, Vilaplana, and Marco 2019).

Like antibodies, nanobodies show high affinities and specificities for the target protein, in addition to other advantages. Due to their small size of about 14 kDa, they interfere less sterically when compared to conventional antibodies and can reach epitopes that are very narrow. An IgG antibody has a size of 150 kDa and a heavy chain only antibody of 75 kDa (Bannas, Hambach, and Koch-Nolte 2017). The small size of nanobodies can also be used to improve, for example, the resolution in high resolution microscopy in which the nanobodies are fluorescently labelled to detect target structures, as fluorophore displacement is avoided (Traenkle and Rothbauer 2017). Further advantages of individual nanobodies might include increased solubility, high chemical, and thermal stabilities, high yield purification, deep tissue penetration, fast blood clearance and low immunogenicity, which make them interesting for both diagnostics and therapeutics (Salvador, Vilaplana, and Marco 2019). Thus, nanobodies are used in many biochemical assays, which address both functional and structural questions. Binding to the target protein can lead to transport inhibition, which allows interesting aspects of transport mechanisms to be elucidated (Salvador, Vilaplana, and Marco 2019). As they take on the role of stabilizing certain conformations of the target protein thereby bringing homogeneity related to the conformations into the sample, they often facilitate protein structure determination (Steyaert and Kobilka 2011). Especially for membrane proteins, nanobodies can also be regarded as crystallization chaperones since their hydrophilic character usually leads to more crystal contacts. Since a variety of different nanobodies are isolated from an alpaca immunization which can bind to different epitopes of the target protein, different conformations are specifically stabilized. The stabilization of different conformations can especially aid to elucidate structures that do not correspond to the preferred conformation of the protein (Shohei Koide 2009). Nanobodies are also successfully used in cryo-EM, which is increasingly used for structural elucidation. The nanobodies facilitate the determination of structures by cryo-EM as they increase the size of the rigid core of the protein; it is often challenging to solve structures of particle sizes below 80 - 100 kDa in cryo-EM. Additionally, the size of the nanobodies is often increased to megabodies by adding protein scaffolds, while retaining the nanobody's specificity, affinity, and ability to bind conformational epitopes (Uchański et al. 2021). Additionally, within cryo-EM studies of nanobody-protein complexes Fab modules can be used to further increase the size and serve as fiducial markers by binding the conserved scaffold of nanobodies (Bloch et al. 2021).

1.5 Aim of thesis

This PhD thesis provides new insights into the transport mechanism of SBP dependent secondary transporting systems. In particular, the functions of previously unexplored homolog proteins of the TRAP transporter DctPQM of *Rhodobacter capsulatus* were investigated. So far, only a few proteins belonging to this group have been functionally characterized in detail, none of which belong to the TRAP subgroup TAXI. For this purpose, the screening for the identification of biochemical superior homologs was focused on the TAXI transporter. We identified their ligands by thermal shifts assays and verified their transport by whole cell uptake assays. In order to investigate the transport of substrates under defined conditions, a proteoliposome-based transport assay was established for a TAXI transporter of *Proteus mirabilis*, in which not only the α-ketoglutarate transport could be demonstrated, but also the strict proton coupling and need for a membrane-anchored SBP. Furthermore, it was shown that a TAXI transporter from *Marinobacter hydrocarbonoclasticus* transports the identified substrate fumarate *in vivo* only if both membrane domains Q and M are present. Even the replacement of one of the two periplasmic loops by a glycine serine linker leads to complete loss of transport with expression remaining present. The interaction of SBPs with several transporting systems was shown for the SBP SiaP of the TRAP SiaPQM from *Haemophilus influenzae*. It could be excluded that it functions with the only native *E. coli* TRAP transporter YiaMNO, so that the transport presumably occurred through an ABC transporter. Furthermore, Michaelis-Menten kinetics and affinities were considered for two TAXI transporting systems. In addition, nanobodies were selected for one identified TAXI target protein, which stabilize different conformations and can thus serve as a basis for structural elucidation. We therefore expand on previously published data of the TRAP family and provide new insights into the relevance of SBPs and, in general, in the thus far less investigated subgroup of TAXI transporter.

2 Materials

2.1 Chemicals

Chemicals and suppliers used in this study are listed in Table 1.

Table 1: **Chemicals.**

Chemical	**Supplier**
Acetic acid	Carl Roth
Aceton	Carl Roth
Acrylamide/bisacrylamide, 30 %/0.8 % (w/v)	Carl Roth
Adenosine-5'-triphosphate (ATP) disodium salt	Carl Roth
Agarose	Sigma-Aldrich
Agar powder	AppliChem
Ammonium peroxodisulfate (APS)	Carl Roth
Ampicillin sodium salt	Carl Roth
L(+)-Arabinose	Carl Roth
Bromophenol blue	Carl Roth
Calcium chloride dihydrate ($CaCl_2 \cdot 2H_2O$)	Carl Roth
Carbonylcyanid-3-chlorphenylhydrazone (CCCP)	Sigma-Aldrich
Charcoal activated	Carl Roth
Chloramphenicol	Carl Roth
Citric acid	Carl Roth
Coomassie Brilliant Blue R-250 (CBB)	AppliChem
7-Diethylamino-3-(4′-maleimidylphenyl)-4-methylcoumarin) (CPM)	Sigma-Aldrich
Dimethyl sulfoxide (DMSO)	AppliChem
DL-Dithiothreitol (DTT)	Carl Roth
Ethanol ≥99.8 %, p. a.	Carl Roth
Ethylendiamine tetraacetic acid (EDTA)	Carl Roth
Fluorescein-5-maleimide	Fisher Scientific
D(+)-Glucose	Carl Roth
Glycerol 86 %	Carl Roth
Glycerol ≥99.7%, p.a.	Carl Roth
Glycine	Carl Roth
HEPES	Carl Roth
HisPur NiNTA Resin	Thermo Scientific
Hydrochloric acid solution 6N	Carl Roth
Imidazole	Carl Roth
InstantBlue® Coomassie Protein Stain (ISB1L)	abcam
Isopropanol	Carl Roth
Isopropyl β-D-1-thiogalactopyranoside (IPTG)	Carl Roth

Kanamycin sulfate	Carl Roth
α-Ketoglutaric acid	Sigma-Aldrich
Magnesium chloride ($MgCl_2$)	Carl Roth
Magnesium sulfate ($MgSO_4$)	Carl Roth
ß-Mercaptoethanol	Carl Roth
MES	Carl Roth
Methanol	Carl Roth
Methoxypolyethylene glycol maleimide 5,000 (PEG-maleimide)	Sigma-Aldrich
Milk Powder Rapilait	Mikros
N-ethylmaleimide (NEM)	Sigma-Aldrich
PEG-maleimide 5000 (PEG-5K)	Fluka
Phenylmethylsulfonyl fluoride (PMSF)	Carl Roth
L-α-phosphatidylcholine from soybean, Type II-S, 14-23 % choline basis (soy PC)	Sigma-Aldrich
Potassium chloride (KCl)	Carl Roth
Potassium dihydrogen phosphate (KH_2PO_4)	Carl Roth
di-Potassium hydrogen phosphate (K_2HPO_4)	Carl Roth
Pyrimidine	Sigma-Aldrich
D(+)-Saccharose	Carl Roth
Serva DNA Stain G	SERVA
Sodium chloride (NaCl)	Carl Roth
Sodium dodecyl sulfate (SDS)	Carl Roth
di-Sodium hydrogen phosphate (Na_2HPO_4)	Carl Roth
Sypro Orange Protein stain	Sigma-Aldrich
Tetramethylethylenediamine (TEMED)	Carl Roth
3,3′,5,5′-Tetramethylbenzidine (TMB)	provided by Dr. Eric Geertsma
TRIS	Carl Roth
Tryptone BioChemica	AppliChem
Tween-20	Merck Millipore
Valinomycin	Sigma-Aldrich
Yeast extract BioChemica	AppliChem

2.2 Radioactive substrates

Radioactive substrates used for *in vivo* and *in vitro* transport assays in this study are listed in Table 2.

Table 2: **Radioactive substrates.**

Substrate	Supplier
[1,4-^{14}C] - fumaric acid	Hartmann Analytik
[1,4-^{14}C] - fumaric acid	Movarek
[^{14}C(U)] - α-ketoglutaric acid sodium salt	American Radiolabeled Chemicals, Inc.
[^{14}C(U)] - α-ketoglutaric acid sodium salt	BIOTREND
[^{14}C(U)] - α-ketoglutaric acid sodium salt	Hartmann Analytik
[6-^{3}H] - sialic acid	American Radiolabeled Chemicals, Inc.

2.3 Detergents and lipids

Detergents and lipids used for solubilization, purification and reconstitution of membrane proteins are listed in Table 3.

Table 3: **Detergents and lipids.**

Detergent/lipid	Supplier
n-Decyl-β-maltoside (DM)	GLYCON Biochemicals
n-Dodecyl-β-maltoside (DDM)	GLYCON Biochemicals
E. coli Extract polar	Avanti
L-α-phosphatidylcholine from soybean, Type II-S, 14-23 % choline basis (soy PC)	Sigma-Aldrich
Triton X-100	Carl Roth

2.4 Enzymes and Antibodies

Enzymes and antibodies used in this study are listed in Table 4.

Table 4: **Enzymes and antibodies.**

Enzyme/Antibody	Supplier
BirA	provided by Dr. Eric Geertsma
Anti-6xHis peroxidase-conjugate	F. Hoffmann-La Roche AG
DNase I	AppliChem
Goat anti-mouse IgG alkaline phosphatase conjugate	Sigma-Aldrich
HRV 3C protease	provided by Dr. Eric Geertsma
Lysozyme BioChemica	AppliChem
Monoclonal anti-c-Myc antibody	provided by Dr. Eric Geertsma
Phusion HF polymerase	New England Biolabs
SapI	New England Biolabs
Streptavidin-peroxidase polymer	provided by Dr. Eric Geertsma
Trypsin	Carl Roth
T4 DNA ligase	Thermo Scientific

2.5 Commercially available kits

Commercially available kits used in this study are listed in Table 5.

Table 5: **Commercially available kits.**

Kit	Supplier
DNA Clean&Concentrator kit	ZymoResearch
GenElute™ Bacterial Genomic DNA	Sigma-Aldrich
Immobilon Western Chemiluminescent HRP Substrate	Merck KGaA
Zymoclean™ Gel DNA Recovery kit	ZymoResearch
Zippy™ Plasmid Miniprep kit	ZymoResearch

2.6 Bacterial Growth Media, Buffers and Solutions

All bacterial growth media and buffers used in this study are listed in Table 6 and Table 7, respectively.

Table 6: **Bacterial growth media and their composition.**

Bacterial growth medium	Composition
Antifoam	10 % Polypropylene glycol P 2,000, autoclaved
Lysogeny Broth (LB) Medium	10.0 g/L tryptone, 5.0 g/L yeast extract, 5.0 g/L NaCl, autoclaved

LB Agar Medium	10.0 g/L tryptone, 5.0 g/L yeast extract, 5.0 g/L NaCl, 15 g/L agarose, autoclaved
Terrific Broth (TB) Medium	24 g/L tryptone, 12 g/L yeast extract, 5.0 g/L glycerol, 12.5 g/L K_2HPO_4, 2.3 g/L KH_2PO_4, 1 mM $MgSO_4$, phosphates and magnesium autoclaved separately

Table 7: **Buffers and solutions.**

Buffer	**Composition**
3C Protease cleavage buffer	20 mM HEPES, pH 7.5, 150 mM NaCl, 10 % (w/v) glycerol, 0.2 % (w/v) DM, 0.2 mg/mL 3C protease
Coomassie blue stain solution	2 g/L Coomassie Brilliant Blue R-250, 45 % (v/v) methanol, 10 % (v/v) glacial acetic acid
Coomassie blue destain solution	0.5 % (v/v) methanol, 0.7 % (v/v) glacial acetic acid
DNA loading dye (10X)	50 % (w/v) glycerol, 0.05 % (w/v) bromophenol blue, 100 mM EDTA
ELISA developing buffer	51.5 mM Na_2HPO_4, 24.3 mM citric acid, 0.006 % (v/v) H_2O_2, 1 mg/mL TMB
Fast Prep buffer	50 mM KPi, pH 7.5, 10 % (w/v) glycerol, 1mM $MgSO_4$, 1mM PMSF, 20 µg/mL DNase I
Imidazole elution buffer	20 mM HEPES, pH 7.5, 150 mM NaCl, 10 % (w/v) glycerol, 0.2 % (w/v) DM, 300 mM imidazole
Imidazole wash buffer	20 mM HEPES, pH 7.5, 500 mM NaCl, 10 % (w/v) glycerol, 0.2 % (w/v) DM, 50 mM Imidazole
Milk-PBST buffer	2.5 % (w/v) milk powder in PBST
PBS buffer (10X)	80 g/L NaCl, 2 g/L KCl, 14.4 g/L Na_2HPO_4, 2.4 g/L KH_2PO_4
PBST buffer	0.1 % (v/v) Tween-20 in PBS (1x)
Periplasmic extraction buffer	20 % (w/v) saccharose, 50 mM TRIS/HCl pH 8.0, 0.5 mM EDTA, 0.5 µg/mL lysozyme
Purification buffer membrane protein	20 mM HEPES, pH 7.5, 150 mM NaCl, 10 % (w/v) glycerol, 0.2 % (w/v) DM/0.05 % DDM/0.05 % Triton-100
Purification buffer soluble protein	20 mM HEPES, pH 7.5, 150 mM NaCl, 10 % (w/v) glycerol
Resuspension buffer	60 mM $CaCl_2$, 15 % (w/v) glycerol
SEC running buffer membrane protein	20 mM HEPES, pH 7.5, 150 mM NaCl, 0.2 % (w/v) DM/0.05 % DDM/0.05 % (w/v) Triton-100, filtered and degassed
SEC running buffer soluble protein	20 mM HEPES, pH 7.5, 150 mM NaCl, filtered and degassed
SDS PAGE Running buffer (10X)	144 g/L glycine, 30.2 g/L TRIS, 5 g/L SDS
SDS PAGE sample loading buffer, reducing (5X)	10 % (w/v) SDS 10 mM dithiothreitol or β-mercaptoethanol, 20 % (v/v)

	glycerol, 0.2 M TRIS-HCl, pH 6.8, 0.05 % (w/v) bromophenolblue
SDS PAGE sample loading buffer, non-reducing (5X)	10 % (w/v) SDS, 20 % (v/v) glycerol, 0.2 M TRIS-HCl, pH 6.8, 0.05 % (w/v) bromophenolblue
Solubilisation buffer	20 mM HEPES, pH 7.5, 150 mM NaCl, 10 % (w/v) glycerol
TAE buffer (50x)	2 M Tris, 1 M acetic acid, 50 mM EDTA pH 8.0
TBS (10x)	44.2 g/L TRIS, 87.7 g/L NaCl, adjusted pH to 7.4
TBS-BSA	0.5 % (w/v) BSA in TBS (1x)
TBS-BSA-D	0.5 % (w/v) BSA 0.1 % (w/v) DM in TBS (1x)
TBS-D	0.261 % (w/v) DM in TBS (1x)
Uptake buffer proteoliposomes	50 mM KPi or NaPi, 2 mM $MgSO_4$, pH 6.0/7.5
Uptake buffer whole cells	50 mM KPi or NaPi, 2 mM $MgSO_4$, 0.1% (w/v) glucose, pH 6.0/7.5
Western Blot transfer buffer	48 mM TRIS, 39 mM glycine, 20 % (v/v) methanol

Bacterial strains

Bacterial strains used in this study are listed in Table 8.

Table 8: **Bacterial strains.**

Strain	Purpose	Genotype	Supplier
E. coli DB3.1	Vector construction	F⁻ *gyrA462 endA1 ∆(sr1-recA) mcrB mrr hsdS20 glnV44 (=supE44) ara14 galK2 lacY1 proA2 rpsL20 xyl5 leuB6 mtl1*	Provided by Dr. Eric Geertsma (Yanisch-Perron *et al.*, 1985)
E. coli MC1061	Expression	F⁻ λ⁻ Δ*(ara-leu)7697* [*araD139*]B/r Δ*(codB-lacI)3 galK16 galE15* e14⁻ *mcrA0 relA1 rpsL150*(Str^R) *spoT1 mcrB1 hsdR2*(r^-m^+)	Provided by Dr. Eric Geertsma (Casadaban and Cohen, 1980)
E. coli IMW424	Transport assay TAXIMh-PQM	*λ- F- fnr-, P1-sensitiv, ΔdctA::spcR, dcuA::spcR, dcuB::kan^R*	Provided by Dr. Eric Geertsma (Ok Bin Kim, 2006)
E. coli JW2571-1	Transport assay TAXIPm-PQM	F⁻ λ⁻ Δ*(araD-araB)567* Δ*lacZ4787*(::rrnB-3) Δ*kgtP754::kan rph-1* Δ*(rhaD-rhaB)568 hsdR514*	Keio collection
E. coli JW3193-1	Transport assay TRAPHi-SiaPQM	F⁻ λ⁻ *Δ(araD-araB)567 ΔlacZ4787*(::rrnB-3) *ΔnanT752::kan rph-1 Δ(rhaD-rhaB)568 hsdR514*	Keio collection
E. coli JW5651-1	Transport assay TRAPHi-SiaPQM	*F- λ- Δ(araD-araB)567 ΔlacZ4787(::rrnB-3) ΔyiaN761::kan rph-1 Δ(rhaD-rhaB)568 hsdR514*	Keio collection
E. coli JW5651-1 ΔnanT	Transport assay TRAPHi-SiaPQM	*F- λ- Δ(araD-araB)567 ΔlacZ4787(::rrnB-3) ΔyiaN761::kan rph-1 Δ(rhaD-rhaB)568 hsdR514* ΔnanT	This study

2.7 Plasmids

Plasmids used in this study for cloning and overexpression are listed in Table 9.

Table 9: **Plasmids.**

Plasmid name	Purpose	Characteristics	Reference
pINIT_cat	Mutagenesis of TRAPs and TAXIs, sequencing vector	pBR322 origin, cam resistance	Geertsma et al. (2011)
pBXC3H	Expression screening of TRAPRc-DctQM and TRAPRc-DctP homologs, expression of TRAPRc-DctP homologs	pBR322 origin, ampicillin (amp) resistance, arabinose-inducible promoter, C-terminal fusion with HRV-3C protease cleavage site and 10x His-tag	Geertsma et al. (2011)
pBXNH3	Expression screening of TRAPRc-DctQM and TRAPRc-DctP homologs, expression of TRAPRc-DctP homologs	pBR322 origin, amp resistance, arabinose-inducible promoter, N-terminal fusion with HRV-3C protease cleavage site and 10x His-tag	Geertsma et al. (2011)
p7XC3GH	Expression screening of TAXINa-P	pBR322 origin, kan resistance, IPTG-inducible T7 promoter, C-terminal fusion with HRV-3C protease cleavage site, GFP fusion and 10x His-tag	Geertsma et al. (2011)
pBXC3GH	Expression of TAXI and TRAP membrane domains	pBR322 origin, amp resistance, arabinose-inducible promoter, C-terminal fusion with HRV-3C protease cleavage site, GFP fusion and 10x His-tag	Geertsma et al. (2011)
pES7	Mutagenesis of TRAPHi-SiaPQM	pSC101 origin, amp resistance, IPTG-inducible lac promoter, 6x His-tag	Mulligan et al. (2009)
pEXC3sfGH	Transport assays of TRAPs und TAXIs	pSC101 origin, amp resistance, IPTG-inducible lac promoter, C-terminal fusion with HRV-3C protease cleavage site, superfolder GFP fusion and 10x His-tag	Roden, unpublished
pBXCA3GH	Expression of TAXIPm-QM for nanobody screening	pBR322 origin, amp resistance, arabinose-inducible promoter, C-terminal fusion with Avi-tag,	Geertsma et al. (2011)

		HRV-3C protease cleavage site, GFP fusion and 10x His-tag	
pBXNA3GH	Expression of TAXIPm-QM for nanobody screening	pBR322 origin, amp resistance, arabinose-inducible promoter, N-terminal fusion with Avi-tag, C-terminal fusion with HRV-3C protease cleavage site, GFP fusion and 10x His-tag	Geertsma et al. (2011)
pSB_init	Nanobody sequencing	pBR322 origin, cam resistance, arabinose-inducible promoter, PelB leader sequence, C-terminal Myc-tag and 6x His-tag	Zimmermann et al. (2018)
pDX_init	Expression of nanobodies	pBR322 origin, amp resistance, IPTG-inducible lac promoter, PelB leader sequence, C-terminal 6x His-tag and EPEA-tag	Zimmermann et al. (2018)

2.8 Primers

Primers used in this study for general sequencing are listed in Table 10. The primers were generated, and the sequencing was performed at Microsynth AG, Göttingen.

Table 10: **Sequencing primers used in this study.**

Primer name	**Primer sequence (5' to 3')**
pINIT_FW	GAG TAG GAC AAA TCC GC
pINIT_rev2	TGG CAG TTT ATG GCG GGG CGT
intGFPrevSEQ	CAT TAA CAT CAC CAT CTA ATT CAA CAA GAA
145_3C_rev	GGA CCT TGA AAC AAA ACT TCT A
pBAD_forward	ATG CCA TAG CAT TTT TAT CC
pBAD24_rev	GCT GAA AAT CTT CTC TCA TCC G
pDX_seq_for	GAG CGG ATA ACA ATT TCA CA

Primers used in this study for amplification of homologues proteins to the membrane domain of the TRAP transporter DctQM from *Rhodobacter capsulatus* are listed in Table 11. The primers were generated at Microsynth AG, Göttingen.

Table 11: **Primers used to amplify TRAPRc-DctQM homologs.**

Construct	Primer name	Primer sequence (5' to 3')
TRAPHi-SiaQM	62_SiaQM_for	ATA TAT GCT CTT CTA GTC TAA TCC CTC TTT AGT GTG CCG TAA TG
TRAPHi-SiaQM	63_SiaQM_rev	TATATAGCTCTTCATGCTGGTATCAATAGATTTGGCACAA ATGT
TRAPEc-YiaMN	186_YiaM-FX_f	ATA TAT GCT CTT CTA GTA AAA AAA TAC TCG AAG CAA TAC TGG CG
TRAPEc-YiaMN	189_YiaN-FX_r	TAT ATA GCT CTT CAT GCA TTA ATC CAT TTC AAA GGG AGG ATG AT
TAXIDc-QM	FX forward	ATA TAT GCT CTT CTA GTT TAG TTA GCA CTG ACA GAT GG TTG TTT
TAXIDc-QM	FX reverse	TAT ATA GCT CTT CCT GCC TGA CTA AGC ATA ATT TGA CAA CAC CG
TAXISb-QM	FX forward	ATA TAT GCT CTT CTA GTT CAG GCA TCC ACA TGA CCG GTA ACT GG
TAXISb-QM	FX reverse	TAT ATA GCT CTT CAT GCA TGT GCA GTA ATA ATG TTG AGC TGA AT
TAXINa-QM	FX forward	ATA TAT GCT CTT CTA GTA AAA TTG CGT TCG CGA TCC AAA CCG AC
TAXINa-QM	FX reverse	TAT ATA GCT CTT CAT GCG TCA CCG GTC GGC GTG TCG ATG GC
TAXIPm-QM	FX forward	ATA TAT GCT CTT CTA GTC TAG TTT TCT CCA GAT ACT TGC CCT AC
TAXIPm-QM	FX reverse	TAT ATA GCT CTT CAT GCA TGA CTA ATA ATG ACC ACC TAA CTC CG
TAXIMh-QM	68_MarHy29_QM_for	ATA TAT GCT CTT CTA GTA CCA AAG AAC AAT CCC ATG CCA GCG GT
TAXIMh-QM	69_MarHy29_QM_rev	ACT GAC GCT CTT CAT GCA GCA GCA GCT AGT GCC GGC GCA CGC TT

Primers used in this study for amplification of the SBPs belonging to the homologues proteins of the TRAP transporter DctQM from *Rhodobacter capsulatus* are listed in Table 12. The primers were generated at Microsynth AG, Göttingen.

Table 12: **Primers used to amplify SBP belonging to TRAPRc-DctQM homologs.**

Construct	Primer name	Primer sequence (5' to 3')
TRAPHi-SiaP No SS	58_SiaP_for	ATATATGCTCTTCTAGTGCTGATTATGACTTGAAATTCGG TATG
TRAPHi-SiaP	59_SiaP_rev	TAT ATA GCT CTT CAT GCT GGG TTG ATT GCT TCA ATT TGT TTT AA
TRAPEc-YiaO No SS	190_YiaO-FX_f	ATA TAT GCT CTT CTA GTA AAT TAC GCT CTG TAA CCT ACG CAT TA
TRAPEc-YiaO	191_YiaO-FX_r	TAT ATA GCT CTT CAT GCT TGC ACC TCA TCC ACG GCT TTC AGC AG
TAXIDc-P SS	744DCAsigseq_ for	ATA TAT GCT CTT CCA GTT TGA AAT TAA CTA AAC GTG GAC TTG CT
TAXIDc-P No SS	749DCAnosig_f or	ATA TAT GCT CTT CTA GTG GTT CTT CTA CCG GTG AGC AAA AGG CT
TAXIDc-P	754DCA_rev	TAT ATA GCT CTT CAT GCC TTC TCG TTA AAG TAT TTC TCT GCA CC
TAXISb-P SS	745SBLsigseq-f_for	ACT GAC GCT CTT CTA GTG CAT CCC GGA AAA CAA AAG TAT TGC TT
TAXISb-P No SS	750SBLnosig_fo r	ATA TAT GCT CTT CTA GTG GTA AAA GAG AAT CAG AAC AAG ACC
TAXISb-P	755SBL_rev	TAT ATA GCT CTT CCT GCC TGC TCA TGA TGT GAT CCT CCC TCT AC
TAXINa-P SS	741NASsigseq_ for	ATA TAT GCT CTT CTA GTG AGG ATC GAC CGA TTT CCA CGA ACA GG
TAXINa-P No SS	746NASnosig_f or	ATA TAT GCT CTT CTA GTG GGT GTC TCG GCG GTA ATG AGG GGG AT
TAXINa-P	751NAS_rev	TAT ATA GCT CTT CAT GCA GGC TCG TCA GCA CGT TCG AAC TCG TC
TAXIPm-P SS	742PMIsigseq_f or	ATA TAT GCT CTT CTA GTA GAA TAA GAA AAA CAA AGC TTG CCA TC
TAXIPm-P No SS	747PMInosigse q_for	ATA TAT GCT CTT CTA GTA TCG GTA AAA AAG AGA ATG CCA ATA AG
TAXIPm-P	752PMI_rev	TAT ATA GCT CTT CAT GCT TTA TTA TTA GTT TGC GTT TGA GCT TC
TAXIMh-P SS	64_MarHy29Pwi thsignalsequenc e_for	ATA TAT GCT CTT CTA GTT TTT TAC TAA CTG GCC TGG CCC TTG CT

TAXIMh-P SS	No	748MHYnosig_for	ATA TAT GCT CTT CTA GTC AAG ATC GTG CCG GCT GGC CTG AGA GC
TAXIMh-P		65_MarHy29_P_rev	TAT ATA GCT CTT CAT GCT TGC TTG ATG GCA GCG TCT TCG ATG GG

Primers used in this study to generate TAXIMh-QM mutants are listed in Table 13. The primers were generated at Microsynth AG, Göttingen.

Table 13: **Mutagenesis primers used to generate TAXIMh-QM mutants.**

Primer name	Primer sequence (5' to 3')
64_MarHy29Pwithsignalsequence_for	ATA TAT GCT CTT CTA GTA CCA AAG AAC AAT CCC ATG CCA GCG GT
69_MarHy29_QM	ACT GAC GCT CTT CAT GCA GCA GCA GCT AGT GCC GGC GCA CGC TT
80_M29PQMdelQMfor	GGA CTT ATC ATG ACC AAA GAA CAA TCC GTG CCG GAT CTG GTT CTG ATC GT
81_M29PQMdelQPrev	GCA CGG ATT GTT CTT TGG TCA TGA TAA GTC
82_M29PQMparQMfor	GGA CTT ATC ATG ACC AAA GAA CAA TCC GGT GCT GCG AGG CAT CGC
83_M29PQMparQPrev	CAC CGG ATT GTT CTT TGG TCA TGA TAA GT
132_M29-1-1xGS_r	GCT TCC ACC CCC GCC CAC GTT CAG GGT GTA GAG GT
110_M29-3a-1xGS_f	GGC GGG GGT GGA AGC TGG TCG TTC CGT ATT GTG CA

Primers used in this study to generate TRAPHi-SiaQM mutants are listed in Table 14. The primers were generated at Microsynth AG, Göttingen.

Table 14: **Mutagenesis primers used to generate TRAPHi-SiaQM mutants.**

Primer name	Primer sequence (5' to 3')
58_SiaP_for	ATATATGCTCTTCTAGTGCTGATTATGACTTGAAATTCGG TATG
63_SiaQM_rev	TATATAGCTCTTCATGCTGGTATCAATAGATTTGGCACAA ATGT
54_SiaPQM-mutant_delQ_M_for	GGA GTG AGT ATG AAA TAT ATT AAT AAG CAA ACC CTA AAC TTT AAG ACT GG
55_SiaPQM_mutant_delQ_P_rev	TTT GCT TAT TAA TAT ATT TCA TAC TCA CTC CGT TT
56_SiaPQM_mutant_parQ_Mfor	GGA GTG AGT ATG AAA TAT ATT AAT AAG GAT TTT TTA ACT AAT CTA ATG CCC GAA
57_SiaPQM_mutant_parQ_P_rev	AAT CCT TAT TAA TAT ATT TCA TAC TCA CTC CGT TT

126_Sia1-2xGS_rev	CCA CCA CCA GAA CCA CCA CCA CCG CGT GAA AGA ATT TGA GCG AT
127_Sia1-2xGS_for	GGT TCT GGT GGT GGT GGT TCT GAA GAA CTC GCC AAG CTC TTA
128_Sia2-2xGS_rev	CCG CCG CCA CTG CCG CCG CCG CCA CCG TTA AAA GTA CGA ATA CC
129_Sia2-2xGS_for	GGC AGT GGC GGC GGC GGC TCA ATT TCT GAA AAA TGG ATT TTC
130_Sia2-3xGS_rev	GAG CCG CCG CCG CCA CTG CCA CCG CCT CCA ATA CCG AAA TGA ATA AAT AAG A
131_Sia2-3xGS_for	GCG GCG GCG GCT CCG GAG GCG GCG GGA GCA TTT CTG AAA AAT GGA TTT TC
163_SiaP_R147A_r	CTG CAT TTG GCA CTG CAA GTT TTA AGC C
164_SiaP_R147A_f	GAA AGG CTT AAA ACT TGC AGT GCC AAA TGC AGC AAC

Primers used in this study to generate an *E. coli* strain deficient of the nanT and YiaMNO transporter are listed in Table 15. The primers were generated at Microsynth AG, Göttingen.

Table 15: **Mutagenesis and sequencing primers used to generate knockout strain.**

Primer name	Primer sequence (5’ to 3’)
70_E.coli-genome-bef-nanT-for	GCT GAT GGT GGT ATC GGC AG
71_E.coli-genome-behind-nanT-rev	TAT AGG CCG TGA TGC GTA CC
195_YiaM_KO_for	GAA GCA ATA CTG GCG ATT AAT CTC GCC GTA CTT TCC TGG TGT AGG CTG GAG CTG CTT C
196_YiaO_KO_rev1	GTC GCT AAA GGC TTT GCG GTC GAG ATC GGT GAT GAC CAT ATG AAT ATC CTC CTT AG

Primers used in this study to generate the pEXC3sfGH vector from the pES7-SiaPQM and pBXC3sfGH vector are listed in Table 16. The primers were generated at Microsynth AG, Göttingen.

Table 16: **Mutagenesis primers used to generate pEXC3sfGH vector**.

Primer name	Primer sequence (5’ to 3’)
88_pES7-SapIdel_f	TCA GTG AGC GAG GAA GCG GAA CAG CGC CCA ATA CGC AAA C
89_pES7-SapIdel_r	GTT TGC GTA TTG GGC GCT GTT CCG CTT CCT CGC TCA CTG A
217_pEXwoSSic_r	ATG TAT ATC TCC TTC TTT TAC GGT ACC
218_pEXwoSSic_f	GCT CTT CTG CAT TAG AAG TTT TGT TTC

219_pEXcccdBic_f	GAA GGA GAT ATA CAT AGT TGA AGA GCG ACC TGC AG
220_pEXccdBic_r	CTA ATG CAG AAG AGC TGA ACT AGT GGA TCC CCA AAA AAG AAG

Primers used in this study to generate N-terminal and C-terminal Avi-tagged TAXIPM-QM for nanobody selection are listed in Table 17. The primers were generated at Microsynth AG, Göttingen.

Table 17: **Mutagenesis primers used to generate N-terminal and C-terminal Avi-tagged TAXIPM-QM for selection.**

Primer name	Primer sequence (5' to 3')
198_26T-avi-FX_f (C-term)	ATA TAT GCT CTT CTA GTA GTA CTA ATA ATG ACC ACC TAA CTC CG
199_26T-avi-FX_r (C-term)	TAT ATA GCT CTT CAT GCT TCA TGC CAT TCG ATT TTC TGA GCT TC
200_avi-26T-FX_f (N-term)	ATA TAT GCT CTT CTA GTG GTC TGA ACG ATA TCT TCG AAG CTC AG
201_avi-26T-FX_r (N-term)	TAT ATA GCT CTT CAT GCG TTT TCT CCA GAT ACT TGC CCT ACA TT

2.9 Proteins

Proteins that were overexpressed in large-scale in this study are listed in Table 18. The corresponding DNA sequences can be found with the listed NCBI reference. The DNA sequences were obtained in this study by PCR amplification of open reading frames from the genomic DNA of respective organisms.

Table 18: **NCBI sequence reference of TRAPRc-DctPQM homologs that were selected within this study.**

Protein	Source organism	DSM	NCBI reference
TRAPHi-SiaP	*Haemophilus influenzae*, gram -	4690	WP_005694430.1
TRAPHi-SiaQM	*Haemophilus influenzae*, gram -	4690	WP_005694432.1
TRAPEc-YiaO	*Escherichia coli*, gram -	18039	WP_000776887.1
TRAPEc-YiaM	*Escherichia coli*, gram -	18039	WP_000721664.1
TRAPEc-YiaN	*Escherichia coli*, gram -	18039	WP_049904699.1
TAXIDc-P	*Desulfotomaculum carboxydivorans*, gram +	14880	WP_003542660.1
TAXIDc-QM	*Desulfotomaculum carboxydivorans*, gram +	14880	WP_013810889.1
TAXISb-P	*Shimwellia blattae*, gram-	4481	WP_002440971.1
TAXISb-QM	*Shimwellia blattae,* gram -	4481	WP_002440973.1
TAXINa-P	*Natrialba asiatica*, gram -	12278	WP_049904699.1
TAXINa-QM	*Natrialba asiatica*, gram -	12278	ELY99376.1
TAXIPm-P	*Proteus mirabilis*, gram -	4479	WP_231310397.1

TAXIPm-QM	*Proteus mirabilis*, gram -	4479	WP_004247951.1
TAXIMh-P	*Marinobacter hydrocarbonoclasticus, gram -*	8798	WP_011785813.1
TAXIMh-QM	*Marinobacter hydrocarbonoclasticus, gram -*	8798	WP_014420714.1

The properties of proteins used for more detailed characterization in this study are listed in Table 19. The values were obtained from Expasy Prot Param after providing the sequences originating from NCBI. The extinction coefficient is given in two values, the first applies under oxidizing conditions and the second under reducing conditions. Under oxidizing conditions we assume that all cysteines form cystines.

Table 19: **Properties of proteins that were selected in assays within this study.**

Protein	**Length (aa)**	**Molecular weight (kDa)**	**Extinction coefficient (M^{-1} cm^{-1})**
TRAPHi-SiaP	307	34	23840
TRAPHi-SiaQM	616	68	66600/66350
TRAPEc-YiaO	328	*36*	25440
TRAPEc-YiaM	157	16	29575/29450
TRAPEc-YiaN	425	425	65555/65430
TAXIDc-P	305	32	19370
TAXIDc-QM	652	69	62925/62800
TAXISb-P	330	36	45380
TAXISb-QM	650	69	84020/83770
TAXINa-P	316	35	53985/53860
TAXINa-QM	697	73	68885/68760
TAXIPm-P	322	35	37360
TAXIPm-QM	655	*70*	71405/71280
TAXIMh-P	304	33	45045/44920
TAXIMh-QM	731	78	92375/92250

2.10 Equipment and Devices

All equipment and devices used are listed in Table 20 with their respective suppliers.

Table 20: **Equipment, devices and their suppliers.**

Equipment		Supplier
Centrifuges	Heraeus Biofuge 13	Thermo Scientific Inc.
	Heraeus Megafuge 1.0R	Thermo Scientific Inc.
	Heraeus Multifuge 1S-R	Thermo Scientific Inc.
	Microcentrifuge Micro Star 17	VWR International
	Optima TLX-120 Ultracentrifuge	Beckman Coulter
	Sorvall WX 80 Ultracentrifuge	Thermo Scientific Inc.
Electro blotter	Trans-Blot Turbo Transfer System	Bio-Rad Laboratories, Inc.
Electrophoresis power supply	Power Supply MP-300V	Major Science
Electrophoresis system/ Electrophoresis supplies	Mini-PROTEAN Tetra Cell System	Bio-Rad Laboratories, Inc.
Extruder	LiposoFast	AVESTIN, Inc.
Fast-Prep homogenizer	FastPrep 24 Classic	MP Biomedicals, LLC
Fermenter	NLF 30L	Bioengineering AG
FPLC system	Azura SEC System	KNAUER Wissenschaftliche Geräte GmbH
Glass beads	Cell Disruption Media 0.1 mm Glass Beads	Scientific Industries
Imaging systems	Fusion FX7 Imager	Vilber
	ImageQuant LAS 4000 Imager	GE Healthcare Life Sciences
Incubator	Heraeus B6120 Incubator	Thermo Scientific Inc.
Microplate reader	Infinite M200 PRO Multimode Microplate Reader	Tecan Group AG
Microwave	Sharp Inverter	Sharp
OD spectrophotometer	Ultrospec 10 Cell Density Meter	Biochrom
PCR cycler	Mastercycler ep gradient S	Eppendorf AG
Pressure cell homogenizer	FPG12804	Stanstedt fluid power LTD
	Gaulin	APV
qPCR	Rotor Gene Q PCR Cycler	QIAGEN N. V.
Rocking platform	Rocking platform	VWR International LLC
Scintillation counter	Hidex 300SL	Hidex Deutschland Vertrieb GmbH
SEC column	Superdex 75 increase 10/300 GE column	GE Healthcare

SEC column	Superdex 200 increase 10/300 GE column	GE Healthcare
Shaker	New Brunswick Scientific C24 Benchtop Incubator Shaker	Marshall Scientific
Sonicator	Sonifier 250	Branson
Thermomixers	Thermomixer 5436	Eppendorf
	Thermomixer pro	CellMedia
UV/Vis spectrometer	NanoDrop spectrophotometer ND-100	Thermo Scientific Inc.
Vacuum pump	KNF Laboport	KNF lab
Weighing scale	Scale Quintix	Sartorius AG
Western Blot	TransBlot Turbo Transfer System	Bio-Rad Laboratories, Inc.

Consumables used in this study are listed with the suppliers in Table 21.

Table 21: **Consumables.**

2.11 Consumables

Item	**Supplier**
Bio-Beads SM2	Bio-Rad Laboratories, Inc.
Cell disruption media 0.1 mm glass beads	Scientific Industries, Inc.
Cellulose nitrate filter 450 nm pore diameter	Sartorius AG
Concentrator Amicon Ultra-4/15 Centrifugal Filters, 3/10/30/50 kDa MWCO	Merck Millipore
CryoPure tubes	Sarstedt
Desalting Column Econo-pac 10DG	Bio-Rad Laboratories, Inc.
Desalting Column Microspin 6	Bio-Rad Laboratories, Inc.
Dialysis membrane, MWCO 8 kD	Carl Roth GmbH + Co. KG
Fast prep tubes 2 mL	Bio-Rad Laboratories, Inc.
Glass Econo columns	Bio-Rad Laboratories, Inc.
Graduated Tips	Starlab
HisPurTM Nickel nitrilotriacetic acid resin	Thermo Fisher Scientific, Inc.
Liquid nitrogen	Linde
Micro tube 1.5 mL/2 mL	Sarstedt
Multiply-µ Strip 0.2 mL chain, PCR tubes	Sarstedt
Nunc-ImmunoTM MicroWellTM 96-well solid plates	Sigma-Aldrich Chemie GmbH
PeqGOLD 1 kb DNA ladder	PEQLAB Biotechnologie GmbH
Petridish 92*16 mm	Sarstedt
Pipettes	Eppendorf AG
Polycarbonate filter 400 nm pore diameter	AVESTIN, Inc.
Poly-prep Chromatography columns	Bio-Rad Laboratories, Inc.

Potter-Elvehjem PTFE pestle	VWR International LLC
Potter-Elvehjem glass tube	VWR International LLC
Roti-Fluoro PVDF membrane 0.2 µm	Carl Roth GmbH + Co. KG
Rotiszint Eco Plus scintillation liquid	Carl Roth GmbH + Co. KG
SERVA Triple Color Protein Standard II	AppliChem GmbH
Sterile needles	B. Braun Melsungen AG
Sterile syringes 1/10/25 mL	B. Braun Melsungen AG
Sterile pipettes 5/10/25/50 mL	Sarstedt
Streptavidin PVT	PerkinElmer, Inc.
Syringe sterile filter 0.2 nm	VWR International LLC
Tubes 15 mL/50 mL	Sarstedt
Ultracentrifuge tubes 7*20 mm	Beckman Coulter
Ultracentrifuge tubes 13*56 mm	Beckman Coulter
Ultracentrifuge tubes for Ti45 rotor	Beckman Coulter
Whatmann paper	Thermo Fisher Scientific Inc.

3 Methods

3.1 Biology

3.1.1 Alpaca immunization

In collaboration with the Nanobody Service facility from the University of Zurich two alpacas, named Adonis and Blitz, were immunized (www.nsf.uzh.ch). For this purpose, detergent-solubilized TAXIPm-QM protein was purified by immobilized metal affinity chromatography (IMAC) and size exclusion chromatography (SEC) after which protein containing fractions were mixed. Over a period of six weeks 100 µg of the protein was injected once a week combined with Gerbu Adjuvant Fama^{TM} in a 1:1 volumetric ratio to boost the immune response. To ensure a good protein quality after storage of the protein at -80 °C, a test sample was thawed and analyzed in the size exclusion chromatography. Blood samples were analyzed four days after each injection in an enzyme linked immunosorbent assay (ELISA) with secondary antibodies specific to IgG1a, IgG1b, IgG2, IgG3a and IgG3b. The titers of the complete antibodies, abbreviated as IgG1A and IgG1b, were only examined for comparison; they were not further processed. The blood samples were stored at 4 °C until proceeding with the generation of the immune libraries (section 4.11.1).

3.2 Microbiology

3.2.1 Large-scale preparation of chemically competent *E. coli* cells

5 mL LB medium were inoculated with *E. coli* MC1061 (Casadaban et al., 1980) or *E. coli* DB3.1 cells (Hartley et al., 2000) from glycerol stocks and incubated overnight at 37° C, shaking. The next day, 500 mL 2-YT medium (20 g/L tryptone, 5 g/L yeast extract, 0.1 M NaCl, 10 mM $MgSO_4$) were inoculated with 1 % (v/v) of the pre-culture in baffled 1 L flasks and incubated at 37°C, shaking until an OD_{600} of 0.6 – 1.0 was reached. The cells were transferred into centrifuge tubes and treated at 4 °C in the following steps or stored on ice. All buffers were prepared and pre-cooled on ice. The cells were pelleted by centrifugation at 5000 g for 20 min at 4° C, resuspended in 750 mL ice-cold TfB I buffer (30 mM KAc, 50 mM $MnCl_2$, 0.1 M KCl, 10 mM $CaCl_2$, 15 % (w/v) glycerol). The centrifugation step was repeated, and the pellet was afterwards resuspended in 75 mL TfB II buffer (10 mM Na-MOPS pH 7.0, 75 mM $CaCl_2$, 10 mM KCl, 15 % (w/v) glycerol). 100 µL aliquots were flash frozen in liquid nitrogen and stored at -80 °C until transformation (section 3.2.3).

3.2.2 Small-scale preparation of chemically competent *E. coli* cells

5 mL LB medium cells were inoculated with *E. coli* cells from glycerol stocks and incubated overnight at 37° C, shaking. The next day, 1 % (v/v) of the pre-cultures was used to inoculate 10 mL of LB medium in 50 mL tubes with gas-permeable lid and incubated at 37°C, shaking until an OD_{600} of 0.5 was reached. A 0.1 M $CaCl_2$ solution was prepared freshly and stored on ice. 1 mL aliquots were transferred to sterile 2 mL tubes and pelleted by centrifugation for approximately 1 min at 9000 g at 4 °C. The cells were then resuspended in 500 µL of the 0.1 M $CaCl_2$ solution, incubated on ice for 30 min and again centrifuged for approximately 1 min at 9000 g at 4 °C. After this second

centrifugation step, the cells were resuspended in 100 µL of the 0.1 M $CaCl_2$ solution and used on the same day for transformation (section 3.2.3).

3.2.3 Transformation of chemical competent *E. coli* cells

50 to 100 ng plasmid DNA was added to chemical competent *E. coli* cells followed by gentle mixing and incubation for 30 min on ice. Then a heat shock transformation was performed for 30 sec at 42 °C. 400 µL of pre-warmed LB medium was added and the cells were incubated for 1 h at 37 °C, shaking. Afterwards the cells were centrifuged for 1 min at 9000 g at 4 °C, resuspended in 150 µL LB medium and transferred to LB agar plates containing the respective antibiotics. The plates were incubated overnight at 37°C.

3.2.4 Preparation of electrocompetent *E. coli* cells

5 mL LB medium cells were inoculated with *E. coli* cells from glycerol stocks and incubated overnight at 37° C, shaking. The next day, 1 % (v/v) of the pre-cultures was used to inoculate 10 mL of LB medium and incubated at 37° C or at 30 °C for temperature sensitive plasmids, shaking, until an OD_{600} of approximately 0.5 was reached. A 1 mM Hepes solution and 10 % (w/v) glycerol were prepared and stored on ice. 1 mL cell aliquots were transferred to sterile 2 mL tubes, incubated for 15 min on ice and pelleted by centrifugation for approximately 7 min at 2200 g at 4 °C. The cells were resuspended in 1 mL of 1 mM Hepes, pH 7.5, pelleted again, resuspended in 500 µL 10 % (w/v) glycerol and pelleted again with a final resuspension in 100 µL 10 % (w/v) glycerol. The cells were used on the same day for transformation (section 3.2.5).

3.2.5 Transformation of electrocompetent *E. coli* cells

100 ng plasmid DNA was added to electrochemical *E. coli* cells in pre-cooled electroporation cuvettes (gap width 0.1 cm) followed by gentle mixing and incubation for 10 min on ice. The subsequent transformation was performed with a 2000 V pulse after which 1 mL of pre-warmed LB medium was immediately added to the cells. The cells were incubated at 37 °C or at 30 °C for temperature sensitive plasmids, shaking. Afterwards the cells were centrifuged for 1 min at 9000 g at 4 °C, resuspended in 150 µL LB medium and transferred to LB agar plates containing the respective antibiotics. The plates were incubated overnight at 37 °C or at 30 °C if the transformation involved temperature sensitive plasmids.

3.2.6 Selection of *E. coli* colonies from LB-agar plates

E. coli single colonies were transferred from LB-agar plates in 5 mL LB medium supplied with the respective antibiotics and incubated overnight at 37 °C, shaking. The next day, either plasmids were isolated and sequenced (section 3.3.1) or the cells were stored (3.2.7).

3.2.7 Storage of *E. coli* cells

For long term storage of *E. coli* cells, cultures were transferred to cryo tubes, supplemented with 20 % (w/v) glycerol, flash frozen in liquid nitrogen and stored at -80 °C. They were used to inoculate

pre-cultures with the aim of expression screenings (section 3.4.1) or protein expression (section 3.4.2 and 3.4.4).

3.2.8 Generation of a double knockout *E. coli* strain

For the generation of an *E. coli* double knockout strain with nanT and YiaMNO knockout, the strain from the Keio collection JW3193-1 that already contains one knockout was cured from its kanamycin resistance. This allowed a subsequent knockout of the second gene and insertion of a new kanamycin resistance using an established protocol (Datsenko and Wanner 2000).

3.2.8.1 Amplification of FRT flanked resistance gene

A PCR reaction was performed to amplify the kanamycin resistance gene from the pKD4 plasmid, flanked with two FRT (FLP recombinase recognition target) sites and homology extensions of the target gene that should later be exchanged for the kanamycin resistance cassette (Table 22 and Table 23). Therefore, the primers were generated in such a way that they contained homologous parts of YiaMNO (Table 15, section 2.8).

Table 22: **Composition of PCR reaction for amplification of FRT flanked resistance gene.**

Item	**Amount**
Phusion HF buffer (5x)	10 µL
dNTPs (10mM) (thaw on ice)	1 µL
,,Forward"-Primer (5 µM)	5 µL
,,Backward"-Primer (5 µM)	5 µL
Dimethylsulfoxid (DMSO)	2 µL
Template plasmid pKD4	1 µL (50 ng)
Phusion Polymerase (add at the end)	0.5 µL
Milli-Q H_2O	to 50 µL

Table 23: **Protocol of PCR reaction for amplification of FRT flanked resistance gene.**

Step	**Temperature**	**Time**
Initial Denaturation	98 °C	3 min
Denaturation/Annealing Touchdown/Extension (14 Cycles)	98 °C/68 °C (-0.5 °C per cycle)/72 °C	10 sec/30 sec/120 sec
Denaturation/Annealing/Extension (14 Cycles)	98 °C/58 °C/72 °C	10 sec/30 sec/120 sec
Final Extension	72 °C	10 minutes
Hold	4 °C	forever

The resulting linear pKD4-PCR product was purified by agarose gel electrophoresis using the Zippy gel extraction kit according to the manufacturer's instructions, followed by gel extraction and storage at -20 °C.

3.2.8.2 Curation of single knockout strain from kanamycin resistance

To cure the *E. coli* strain JW3193-1 that already contains one knockout from its kanamycin resistance, 5 mL LB medium supplemented with 25 µg/mL kanamycin were inoculated with the *E. coli* strain from the Keio collection and grown overnight at 37 °C. The next day, the overnight culture was used to prepare electrocompetent cells, followed by electroporation to introduce the temperature sensitive plasmid pCP20 which contains ampicillin and chloramphenicol resistance. The cells were plated on LB agar plates containing 100 µg/mL ampicillin and incubated for 20 h at 30 °C. The next day, single colonies were picked and used to inoculate 5 mL LB medium which were incubated overnight at 44 °C, shaking, to induce the expression of the FLP recombinase and simultaneously remove the pCP20 plasmid. The next day, the *E. coli* cultures were diluted 100x in LB medium of which 50 µL were plated on LB agar plates and incubated for 20 h at 30 °C. The next day, single colonies were transferred to a 0.85 % (w/v) NaCl solution of which 10 µL were plated on four different plates: LB, LB supplemented with 100 µg/mL ampicillin, LB + 34 µg/mL chloramphenicol, LB + 25 µg/mL kanamycin. Cultures which showed growth on LB agar plates without antibiotics only were cured from all plasmids and the kanamycin selection gene.

3.2.8.3 Introduction of Red recombinase system in single knock out strain

Electrocompetent cells were generated from the *E. coli* strain JW3193-1 that was cured from the kanamycin resistance cassette (section 3.2.4 and 3.2.8.2). The cells were transformed with the pKD46 plasmid by electroporation (section 3.2.5), plated on LB agar plates containing ampicillin and grown for 20 h at 30 °C. The next day, single colonies were picked and used to inoculate 5 mL LB medium. The liquid culture was incubated for 20 h at 30 °C, shaking and the next day stored in 20 % (w/v) glycerol at -80 °C (section 3.2.7).

3.2.8.4 Transformation of single knockout strain expressing Red recombinase

Electrocompetent cells were prepared from the *E. coli* strain JW3193-1 containing the pKD46 plasmid (section 3.2.4). The expression of the Red recombinase was induced by the addition of 1 mM L-arabinose. The cells were transformed by electroporation with the linear pKD4-PCR product directed against the YiaMNO gene (section 3.2.5 and 3.2.8.1). 1 mL pre-warmed LB medium containing 0.1 % (w/v) glucose was immediately added, followed by a 1 h incubation at 37 °C, shaking. The cells were plated on LB agar plates containing 0.1 % (w/v) glucose and 25 µg/mL kanamycin and incubated overnight at 37 °C. The next day, single colonies were transferred to 5 mL LB medium, supplemented with 0.1 % (w/v) glucose and 50 µg/mL kanamycin. The cultures were grown overnight at 37 °C, shaking and the next day stored in 20 % (w/v) glycerol at -80 °C (section 3.2.7).

3.2.8.5 Validation of YiaMNO gene knockout

5 mL LB medium containing 50 µg/mL kanamycin were inoculated with *E. coli* cells containing both the nanT and presumably YiaMNO knockout. The next day, the genomic DNA was extracted using the GenEluteTM Bacterial Genomic DNA Kit according to the manufacturer's instructions and a PCR reaction was performed (Table 24 and Table 25) for which the primers were designed in such a way that they would either amplify the kanamycin resistance cassette or the YiaMNO gene (Table 15, section 2.8).

Table 24: **Composition of PCR reaction for validation of YiaMNO gene knockout.**

Item	**End conc**
Thermo Pol buffer (10x)	5 µL
dNTPs (10mM) (thaw on ice)	1 µL
,,Forward"-Primer (5 µM)	5 µL
,,Backward"-Primer (5 µM)	5 µL
Dimethylsulfoxid (DMSO)	2 µL
Chromosomal DNA	1 µL
Taq Polymerase (add at the end)	0.5 µL
Milli-Q H_2O	to 50 µL

Table 25: **Protocol of PCR reaction for validation of YiaMNO gene knockout.**

Step	**Temperature**	**Time**
Initial Denaturation	98 °C	3 min
Denaturation/Annealing Touchdown/Extension (14 Cycles)	98 °C/68 °C (-0.5 °C per cycle)/72 °C	10 sec/15 sec/60 sec
Denaturation/Annealing/Extension (14 Cycles)	98 °C/56 °C/72 °C	10 sec/15 sec/60 sec
Final Extension	72 °C	10 minutes
Hold	4 °C	forever

The PCR product was separated by agarose gel electrophoresis, followed by gel extraction using the Zippy gel extraction kit according to the manufacturer's instructions and sequencing.

3.3 Molecular biology

3.3.1 Plasmid isolation

5 mL LB medium containing the *E. coli* cells with the respective plasmid were centrifuged at 3000 g for 7 min. The supernatant was discarded, and the pellet resuspended in 500 µL Milli-Q H_2O from which the plasmid was isolated using the Zippy plasmid miniprep kit according to the manufacturer's

instructions. In the final step, the plasmid DNA was eluted from the column using 30 µL Milli-Q H_2O and, if necessary, increased, depending on the concentrations subsequently determined spectrophotometrically, since approximately 1 µL should be pipetted in the following cloning procedure. Re-constructed plasmids or inserts cloned into the plasmid with the pInitial backbone were sequenced at Microsynth AG, Göttingen.

3.3.2 Genomic DNA isolation

To analyze the modified genomic DNA of *E. coli* knockout mutants, 5 mL LB medium containing the *E. coli* cells were centrifuged at 3000 g for 7 min. The supernatant was discarded, and the genomic DNA was isolated using the GenElute Bacterial Genomic DNA kit according to the manufacturer's instructions. In the final step, the genomic DNA was eluted from the column using Milli-Q H_2O and, if necessary, increased, depending on the concentrations subsequently determined spectrophotometrically, since approximately 1 µL should be pipetted in the following PCR reactions.

3.3.3 PCR amplification of open reading frames

Genomic DNA of bacterial and archaeal organisms was purchased from the Keio collection of the German Collection of Microorganisms and Cell Cultures GmbH (DSMZ). To obtain the DNA encoding DctP and DctQM homologous proteins from *Rhodobacter capsulatus*, the open reading frames deriving from the Keio collection were amplified by PCR from a variety of genomes. The primers were constructed using a web-based tool (FXcloning.org) and ordered at the Microsynth AG, Göttingen (Table 11 and Table 12 , section 2.8). The web-based tool automatically introduces SapI sites into the primers which are therefore suitable for FX cloning. The open reading frames were amplified using a touchdown PCR protocol (Table 27).

Table 26: **Composition of PCR reaction for amplification of open reading frames.**

Item	**Amount**
Phusion HF buffer (5x)	10 µL
dNTPs (10mM) (thawed on ice)	1 µL
,,Forward"-Primer (10 µM)	5 µL
,,Backward"-Primer (10 µM)	5 µL
Dimethylsulfoxid (DMSO)	2 µL
Ethylene glycol	3 µL
Template genomic DNA	50-100 ng
Phusion Polymerase (add at the end)	0.5 µL
Milli-Q H_2O	to 50 µL

Table 27: **Protocol of PCR reaction for amplification of open reading frames.**

Step	Temperature	Time
Initial Denaturation	98 °C	3 min
Denaturation/Annealing Touchdown/Extension (14 Cycles)	98 °C/68 °C (-0.5 °C per cycle)/72 °C	10 sec/15 sec/60 sec
Denaturation/Annealing/Extension (14 Cycles)	98 °C/56 °C/72 °C	10 sec/15 sec/60 sec
Final Extension	72 °C	10 minutes
Hold	4 ° C	forever

The DNA obtained from the PCR reactions was purified via agarose gel electrophoresis, followed by gel extraction using the Zippy gel extraction kit according to the manufacturer's instructions and stored on ice or at -20 °C until used for cloning.

3.3.4 Construction of FX Cloning expression vector pEXC3sfGH

pES7SiaPQM was obtained from Dr. Eric Geertsma and changed into an FX cloning suitable expression vector. For this purpose, two SapI sites in the pES7 vector backbone were eliminated by site directed mutagenesis (section 3.3.7). Additionally, the gene encoding for the protein SiaPQM from *Haemophilus influenzae* was exchanged with a DNA sequence deriving from the plasmid pBXC3sfGH by infusion cloning (section 3.3.6 and Table 16, section 2.8). The resulting plasmid pEXC3sfGH is a low copy plasmid as it contains a pSC101 ori, a lac promoter and ampicillin resistance. Furthermore, it is FX cloning suitable as it contains two SapI sites which flank the DNA sequence that should be exchanged by different DNA sequences in the following cloning procedures. Additionally, we introduced DNA sequences encoding for a HRV 3C protease cleavage site, a decaHis tag and a sfGFP fusion protein on the vector backbone.

3.3.5 FX cloning

For FX cloning, the DNA sequences encoding the target proteins were first cloned into the pInitial vector and sequenced at Microsynth AG, Göttingen. In further steps subcloning to the expression vectors pBXC3H, pBXNH3, pBXC3sfGH and pEXC3sfGH was performed (Geertsma and Dutzler 2011). The pInitial vector contains two SapI sites, which allows it to be linearized with simultaneous removal of the ccdB-cassette. Therefore, the primers for amplifying the DNA sequence of the target insert were generated with the help of the web tool fxcloning.org (section 2.8) which attach corresponding SapI sites in a Touchdown PCR reaction (Table 28 and Table 29).

Table 28: **Composition of PCR reaction for FX cloning.**

Item	Amount
Phusion HF buffer (5x)	10 µL
dNTPs (10mM) (thawed on ice)	1 µL
,,Forward"-Primer (10 µM)	5 µL
,,Backward"-Primer (10 µM)	5 µL
Dimethylsulfoxid (DMSO)	2 µL
Ethylene glycol	3 µL
Template DNA	50-100 ng
Phusion Polymerase (add at the end)	0.5 µL
Milli-Q H_2O	to 50 µL

Table 29: **Protocol of PCR reaction for FX cloning.**

Step	Temperature	Time
Initial Denaturation	98 °C	3 min
Denaturation/Annealing Touchdown/Extension (14 Cycles)	98 °C/65 °C (-0.5 °C per cycle)/72 °C	10 sec/15 sec/60 sec
Denaturation/Annealing/Extension (14 Cycles)	98 °C/56 °C/72 °C	10 sec/15 sec/60 sec
Final Extension	72 °C	10 minutes
Hold	4 ° C	forever

50-100 ng of the pInitial vector were mixed with 5x molar excess of the amplified insert, supplemented with 5U SapI and the CutSmart buffer and filled up to 10 µL. The mix was digested for 1 h at 37 °C, followed by 20 min heat inactivation at 65 °C. Then, 1 mM ATP and 2.5U T4 DNA ligase were added and the mix was incubated for 1 h at 25 °C, followed by 20 min heat inactivation at 65 °C. 5 µL of the ligated DNA mix were used for heat shock transformation of chemically competent *E. coli* MC1061 cells (section 3.2.3). After selection on LB agar plates containing 34 µg/mL chloramphenicol, the correctness of the DNA sequences was verified by sequencing. When these were confirmed, they were subcloned into the expression vectors. For this purpose, 250 ng of the pInitial vector containing the target insert were mixed with 50 ng of the respective expression vector pBXC3H, pBXNH3, pBXC3sfGH or pEXC3sfGH, digested with 10U SapI and ligated with T4 DNA ligase as previously performed for the pInitial constructs. 5 µL of the ligated DNA mix was again used for heat shock transformation of chemically competent *E. coli* MC1061 cells (section 3.2.3). The selection of the final constructs was performed on LB agar plates containing 100 µg/mL ampicillin. All resulting FX cloning constructs contain a tunable arabinose promoter and allow after expression

of the target protein from the expression vector elution for purification with HRV 3C protease which simultaneously cuts off both the deca His tag and sfGFP fusion.

3.3.6 Infusion cloning

For infusion cloning, two PCR reactions were performed, one to linearize the vector and one to attach overhangs to the template DNA (Table 30). The primers to attach the overhangs to the template were designed having at the 5' end 15 bases homologues to 15 bases of the linearized vector and at the 3' end bases homologous to the template DNA (section 2.8).

Table 30: **Composition of PCR reactions for Infusion cloning.**

Item	**Amount PCR vector**	**Amount PCR insert**
Phusion HF buffer (5x)	10 µL	10 µL
dNTPs (10mM) (thawed on ice)	1 µL	1 µL
,,Forward"-Primer (5 µM)	5 µL	5 µL
,,Backward"-Primer (5 µM)	5 µL	5 µL
Dimethylsulfoxid (DMSO)	2 µL	2 µL
Template DNA or vector	1 µL (50 ng)	1 µL (50 ng)
Phusion Polymerase (add at the end)	0.5 µL	0.5 µL
Milli-Q H_2O	to 50 µL	to 50 µL

Step	**Temperature**	**Time PCR vector**	**Time PCR insert**
Initial Denaturation	98 °C	3 min	3 min
Denaturation/Annealing Touchdown/Extension (14 Cycles)	98 °C/68 °C (-0.5 °C per cycle)/72 °C	10 sec/15 sec/120 sec	10 sec/15 sec/60 sec
Denaturation/Annealing/Extension (14 Cycles)	98 °C/56 °C/72 °C	10 sec/15 sec/120 sec	10 sec/15 sec/60 sec
Final Extension	72 °C	10 minutes	10 minutes
Hold	4 ° C	forever	forever

The amplified DNA and linearized vector were both purified by agarose gel electrophoresis and extracted from the gel using the Zippy gel extraction kit according to the manufacturer's instructions. All items for the infusion cloning reaction were mixed (Table 31).

Table 31: **Items for infusion cloning reaction.**

Item	Amount
5x Infusion Enzyme mix	2 µL
Linearized vector	10-200 ng
Purified PCR fragment	50-200 ng
Milli-Q H_2O	to 10 µL

The infusion cloning mix was incubated for 15 min at 50 °C and stored on ice or at -20 °C until transformation of *E. coli* DB3.1 cells.

3.3.7 Site directed mutagenesis

Site directed mutagenesis was performed in a megaprimer approach to create mutations in inserts (Forloni, Liu, and Wajapeyee 2019). These inserts were beforehand cloned and sequenced in the FX cloning vector pInitial (section 3.3.5). For this purpose, mutagenesis primer were designed using the web tool PrimerX (https://www.bioinformatics.org/primerx/cgi-bin/protein_1.cgi). Two touchdown PCR reactions were performed to generate megaprimers, one using the pINIT_FW primer in combination with the reverse QuickChange primer and the second using the forward QuickChange primer in combination with the pINIT_rev2 primer. These megaprimers were combined and amplified in a third PCR reaction using the pINIT_FW and pINIT_rev2 primer. The touchdown protocol decreases the annealing temperature stepwise to reduce unspecific binding (Don et al. 1991).

Table 32: **Composition of PCR reaction for site directed mutagenesis.**

Item	Amount PCR 1 + 2	Amount PCR 3
Phusion HF buffer (5x)	10 µL	10 µL
dNTPs (10mM) (thawed on ice)	1 µL	1 µL
,,Forward"-Primer (10 µM) 150 or 152	5 µL	5 µL
,,Backward"-Primer (10 µM) 151 or 153	5 µL	5 µL
Dimethylsulfoxid (DMSO)	2 µL	2 µL
Ethylene glycol	3 µL	3 µL
Template DNA	1 ng	1 µL + 1 µL (PCR 1 + 2)
Phusion Polymerase (add at the end)	0.5 µL	0.5 µL
Milli-Q H_2O	to 50 µL	to 50 µL

Table 33: **Protocol of PCR reaction for site directed mutagenesis.**

Step	Temperature	Time PCR 1 + 2	Time PCR 3
Initial Denaturation	98 °C	3 min	3 min
Denaturation/Annealing Touchdown/Extension (14 Cycles)	98 °C/65 °C (-0.5 °C per cycle)/72 °C	10 sec/30 sec/60 sec	10 sec/30 sec/120 sec
Denaturation/Annealing/Extension (14 Cycles)	98 °C/56 °C/72 °C	10 sec/30 sec/60 sec	10 sec/30 sec/120 sec
Final Extension	72 °C	10 minutes	10 minutes
Hold	4 ° C	forever	forever

The amplified DNA was purified by agarose gel electrophoresis and extracted from the gel using the Zippy gel extraction kit according to the manufacturer's instructions. The purified DNA was stored on ice or at -20 °C until FX cloning into the plasmid pINIT_cat (section 3.3.5).

3.3.8 DNA purification

To purify the target DNA from PCR reactions, the PCR products were supplemented with 10x DNA loading dye and separated by agarose gel electrophoresis which contained 1.5 % (w/v) agarose, supplemented with Serva DNA stain G and was run for 30 to 60 min at 10 V/cm. The comparison with a marker protein revealed the fragment of the correct size which was cut from the gel and extracted using the ZymoResearch Gel DNA Recovery kit according to the manufacturer's instructions. In the final step, the DNA was eluted with 12 µL Milli-Q H_2O and the concentration was spectrophotometrically determined.

3.3.9 Generation of immune libraries

For the generation of TAXIPm-QM specific nanobodies, two immune libraries from the blood samples were prepared, performed by Dr. Eric Geertsma at the MPI CBG in Dresden (Eric Geertsma Group, www.mpi-cbg.de). For this, the RNA of the heavy chain only antibodies was purified, followed by synthesis of the corresponding cDNA. From this the nanobody open reading frames were obtained by PCR and cloned into the pDX phage display vectors for enrichment (Hofmann et al. 2019). For phage display based nanobody enrichment and selection, phages presenting nanobodies were applied on a plate which contained the target antigen TAXIPm-QM. Those phages that recognized the antigen were selected for transformation of *E. coli* TG1 cells and subsequent amplification. The bacteria were transformed in a 10-fold dilution series and at the same time the phages selected in round 1 were used for a second round of selection, which should lead to an enrichment of phages carrying the nanobody gene sequence specific for the TAXIPm-QM protein (Pardon et al. 2014). The resulting libraries were named A1, A2, B1 and B2 and stored at -80 °C until shipped to the University of Frankfurt.

3.3.10 Preparation of nanobody DNA sequences for selection

The preparation of nanobody DNA sequences for selection was performed by Christiane Ruse, University of Frankfurt under my supervision. Four immune libraries of nanobodies directed against TAXIPm-QM were obtained from the MPI CBG Dresden (Eric Geertsma Group, www.mpi-cbg.de) with the genes encoding the nanobodies in the phage display vector pDX_init. To introduce a C-terminal Myc-tag to the nanobody DNA sequence which is used for binding in the selection via ELISA (section 3.5.3), the genes were subcloned from pDX_init into pSB_init using the FX cloning technique (section 3.3.5), followed by transformation in *E. coli* MC1061 cells (section 3.2.3). From this, 95 clones were picked from each phage display library and used to inoculate 1.2 mL TB medium supplemented with 25 µg/mL chloramphenicol in 96 deep-well plates. Additionally, *E. coli* MC1061 cells carrying a plasmid with the gene encoding a UraA specific sybody, pSB_init-*sy45*, was used, kindly obtained from Dr. Benedikt Kuhn, MPI CBG Dresden (Eric Geertsma Group, www.mpi-cbg.de). These pre-cultures were incubated for 4 h at 37 °C, shaking. Afterwards, 1 mL TB medium, supplemented with 25 µg/mL chloramphenicol were inoculated with 50 µL for each sample and incubated for 2 h at 37 °C, shaking, followed by a reduction of the temperature to 22 °C and further incubation for 1.5 h. The expression of the nanobodies was then induced by adding 0.01% (w/v) arabinose. The expression was performed overnight, shaking. The next day, the cells were harvested by centrifugation for 20 min at 3,200 *g* at 4 °C. The supernatant was discarded, and the cell pellets were stored at -20 °C until the nanobodies specific for TAXIPm-QM were selected.

3.4 Biochemistry

3.4.1 Small-scale expression screening

To screen for best expression conditions or in general for proteins that can be well expressed in *E. coli*, a small-scale expression approach was performed. For this, 5 mL of pre-warmed TB medium, supplemented with 100 µg/mL ampicillin were inoculated with MC1061 *E. coli* cells containing the respective plasmids and incubated overnight at 37 °C, shaking. The next day, 15 mL of the same medium were inoculated with 1 % (v/v) of the pre-culture, incubated at 37 °C until an OD_{600} of 0.6 to 0.8 was reached, induced with either 1 mM IPTG or 0.01 % (w/v) arabinose and then continued to grow for either 3 h at 30 °C or overnight at 25 °C, shaking. In the further procedure, the *E. coli* cells were pelleted for 10 min at 2500 g, 4° C, after which the supernatant was discarded. The cell pellet was resuspended in 500 µL of 50 mM KPi, pH 7.5 and then adjusted to achieve the same OD_{600} for all samples. Depending on the expression levels, approximately 2.5 mg of total protein for each sample was pelleted and the supernatant discarded. We thereby assume that the amount from OD_{600} = 1 corresponds to 0.3 mg/mL protein (Geertsma, Groeneveld, et al. 2008). The pellet was supplied with 400 µL of ice-cold Fast Prep buffer containing 50 mM KPi pH 7.5, 10 % glycerol, supplemented with 1 mM $MgSO_4$, 1 mM PMSF and trace amounts of DNAse I. Additionally, 300 mg glass beads were added, and the cells were lysed using the Fast Prep homogenizer. Two cycles of cell disruption were performed at room temperature with five minutes in between to incubate the *E. coli* cells on ice. Afterwards the supernatant was used for analysis by *anti*-His Western Blot (3.4.17).

3.4.2 Soluble protein expression

Large-scale expression of soluble proteins was performed in either 5 L flasks or 9 L fermenter with ampicillin supplemented TB medium using the FX cloning expression vector pBXNH3, pBXC3H or pSB_init. The medium was inoculated with 1 % (v/v) overnight pre-culture. The cells were grown at 37 °C until an OD_{600} of 1 to 1.5 was reached, then the temperature was reduced to 25 °C. When the lower temperature was reached, the cells were induced for gene expression using 0.01 % (w/v) arabinose overnight or 1 mM IPTG and continued to cultivate overnight (Rosano and Ceccarelli 2014; Guzman et al. 1995). The next day, cells were harvested for 20 min at 6000 g and frozen at -20 °C or stored on ice until protein purification (section 3.4.3).

3.4.3 Soluble protein purification

For purification of soluble proteins, cell pellets were resuspended in 20 mM Hepes pH 7.5, 150 mM NaCl, 10 % (w/v) glycerol and supplemented with 1 mg/mL lysozyme, trace amounts of DNAse, 2 mM $MgSO_4$ and 15 mM imidazole. All purification steps were performed either on ice or at 4 °C. The sample was homogenized and incubated for 1 h, stirring. Afterwards the cells were lysed using a pressure cell homogenizer, 1 mM PMSF was added, and non-broken cells were removed by ultracentrifugation. The proteins were then purified via IMAC. For this, Ni-NTA beads were washed with 20 column volume Milli-Q H_2O and pre-equilibrated with 20 column volume of 20 mM Hepes pH 7.5, 150 mM NaCl, 2 mM MgSO4 and 10% (w/v) glycerol. The supernatant was incubated with 1 mL Nickel-NTA resin per 20 mL lysate for 1 h batch binding and then loaded on a BioRad gravity flow column. After the flow through was separated from the Ni-NTA beads, they were subsequently washed with 10 column volumes of the membrane protein purification buffer. Afterwards, the purification buffer was supplemented with 300 mM imidazole and used to elute the proteins. To cleave His-tags from the target proteins, HRC 3C protease was added and the sample was dialyzed against imidazole free buffer overnight. The next day, a Re-IMAC was performed using 20 mM Hepes, pH 7.5, 150 mM NaCl, 10 % (w/v) glycerol to remove His-tags and the HRC 3C protease from the solution. Then, the protein concentration was increased in the respective Amicon Ultra centrifugation tube thereby reducing the overall volume. In the final purification step, a size exclusion chromatography was performed on a Superdex 200 30/100 GL Increase column or Superdex 75 30/100 GL Increase column in a buffer containing 20 mM Hepes pH 7.5 and 150 mM NaCl. The fractions that contained pure target protein were pooled and supplemented with 10 % (w/v) glycerol. If higher, the protein concentration was diluted to a maximum of 10 mg/mL, flash frozen in liquid nitrogen and stored at -80 °C until further usage.

3.4.4 Membrane and membrane-anchored protein expression

Large-scale expression and membrane preparation of membrane and membrane-anchored proteins were performed following the same protocol in either 5 L flasks or 9 L fermenter. The proteins were expressed in with ampicillin supplemented TB medium using the FX cloning expression vectors pBXC3sfGH, pBXC3H, pBXCA3GH or pBXNA3GH. The medium was inoculated with 1 % (v/v)

overnight pre-culture. The cells were grown at 37 °C until an OD_{600} of 1 to 1.5 was reached, then the temperature was reduced to 25 °C. When the lower temperature was reached, the cells were induced for gene expression using 0.01 % (w/v) arabinose and continued to cultivate overnight (Guzman et al. 1995). The next day, cells were harvested for 20 min at 6,000 g and frozen at -20 °C or stored on ice until membrane preparation.

3.4.5 Membrane preparation

To prepare the membrane vesicles for membrane protein purification, the *E. coli* cells were resuspended after expression in a buffer containing 50 mM KPi pH 7.5, 150 mM NaCl, supplemented with 1 mg/mL lysozyme, trace amounts of DNAse, 2 mM $MgSO_4$, homogenized and incubated for 1 h, stirring. All steps were performed on ice or at 4 °C. Afterwards the cells were lysed using a pressure cell homogenizer, 1 mM PMSF was added, and non-broken cells were removed by centrifugation for 30 min at 15,000 g. Afterwards the membrane vesicles were collected by ultracentrifugation for 1 h at 140,000 g and resuspended in 50 mM KPi pH 7.5, 150 mM NaCl with a final concentration of 10 % (w/v) glycerol using a Potter-Elvehjem homogenizer. To store the membrane vesicles, they were diluted to 500 mg/mL, flash frozen in 5 to 10 mL aliquots in liquid nitrogen and kept at -80 °C until further usage.

3.4.6 Membrane protein purification

For purification of membrane proteins 20 g vesicles were thawed in a water bath, stirring and resuspended to 200 mL final volume in 20 mM Hepes, pH 7.5, 150 mM NaCl, 10 % (w/v) glycerol. 3.2 g DM was supplemented to reach a final concentration of 1.6 % (w/v) DM. All steps were performed on ice or at 4 °C. The solubilized membrane proteins were purified via IMAC. For this, Ni-NTA beads were washed with 20 column volume Milli-Q H_2O and pre-equilibrated with 20 column volume of 20 mM Hepes pH 7.5, 150 mM NaCl, 2 mM $MgSO_4$ which was adjusted to 10% (w/v) glycerol and 0.2 % (w/v) DM. The sample was incubated for 1 h batch binding after which it was loaded on a BioRad gravity flow column. After the flow through was separated from the Ni-NTA beads, they were subsequently washed with 10 column volumes of the membrane protein purification buffer. To elute the proteins from the Ni-NTA beads, the sample was incubated for 2 h in the same buffer, supplemented with HRV 3C protease. To separate the beads including His-tags and GFP fusions from target proteins, the samples were again loaded on a BioRad gravity flow column and washed in an appropriate amount of purification buffer. All eluting fractions were collected. To later analyze how much protein was still attached on the Ni-NTA beads, the purification buffer was supplemented with 300 mM imidazole and an additional protein fraction was collected. The elution fractions containing the target protein were concentrated in the respective Amicon Ultra centrifugation tubes thereby reducing the overall volume. In the final purification step, a size exclusion chromatography was performed on a Superdex 200 30/100 GL Increase column in a buffer containing 20 mM Hepes pH 7.5 and 150 mM NaCl, supplemented to a final concentration of 0.2 % (w/v) DM. The fractions that contained pure target protein were pooled and adjusted to 10 % (w/v)

glycerol. If higher, the protein concentration was diluted to a maximum of 10 mg/mL, flash frozen in liquid nitrogen and stored at -80 °C until further usage (section 3.2.7).

3.4.7 Membrane-anchored protein purification

For the purification of membrane-anchored proteins, essentially the same protocol was used as for the purification of membrane proteins, but different detergents were used. 10 g membrane vesicles of membrane-anchored proteins were resuspended in 100 mL of a buffer containing 20 mM Hepes, pH 7.5, 150 mM NaCl, adjusted to 10 % (w/v) glycerol and supplemented with 1.6 g DDM to reach a final concentration of 1.6 % (w/v) DDM. The purification buffer during the IMAC was either supplemented to a final concentration of 0.05 % (w/v) TritonX-100 or 0.05 % (w/v) DDM. In the following size exclusion chromatography, the buffer was initially supplemented with the same detergent as well. In the optimization of the protocol, it was not added to prevent the destabilization of lipids in the following reconstitution. The fractions that contained pure target protein were pooled and adjusted to 10 % (w/v) glycerol. If higher, the protein concentration was diluted to a maximum of 10 mg/mL, flash frozen in liquid nitrogen and stored at -80 °C until further usage (section 3.2.7).

3.4.8 Preparation of liposomes

Liposomes were prepared from soybean L-a-phosphatidylcholine (soy PC) by resuspension of 1 g lipids in 5 mL chloroform, followed by formation of a thin biofilm using a rotary evaporator. The lipid film was resuspended in 50 mM KPi pH 7.5, supplemented with 2 mM $MgSO_4$, such that the final concentration was 20 mg/mL using a sonicator to detach the lipids from the round bottom flask and bring them into solution. Afterwards a canula was used to homogenize the lipid solution which was then 5x flash frozen to form large multilamellar liposomes and stored in liquid nitrogen. On the day of reconstitution, the liposomes were thawed and extruded 11x through 400 nm pore size polycarbonate filters.

3.4.9 Reconstitution of membrane protein into liposomes

Membrane proteins that were purified via IMAC and SEC were reconstituted for transport assays to soy PC 1:40 (w/w) protein-to-lipid ratio. For this, preformed liposomes were destabilized by successively adding approximately 10 µL of 10 % (w/v) Triton X-100 per 5 mL lipid solution until R_{Sat} was obtained, thereby pivoting. The protein was then added, and the detergent was subsequently removed by gradual addition of BioBeads (Geertsma, Nik Mahmood, et al. 2008). The resulting proteoliposomes were collected by ultracentrifugation at 200,000 g for 30 min at 15 °C and resuspended in 50 mM KPi or NaPi pH 7.5, supplemented with 2 mM $MgSO_4$ to a lipid concentration of 20 mg/mL. To randomize the orientation of the reconstituted protein between inside-out and right-side-out, the proteoliposomes were sonicated in six cycles (10 sec on/50 sec off). In the final step, the size of the proteoliposomes was increased by five freeze/thaw cycles in liquid nitrogen. They were then analyzed in a reconstitution efficiency test (section 3.4.12) and stored in liquid nitrogen until used for the additional reconstitution of a membrane-anchored protein (section 3.4.11) or for transport assays (section 3.5.5).

3.4.10 Exchange of proteoliposome inside buffer

All proteoliposomes generated contained 50 mM KPi pH 7.5, supplemented with 2 mM $MgSO_4$ as inside buffer. To mimic a positive membrane potential, the inside buffer was exchanged to 50 mM NaPi pH 7.5, supplemented with 2 mM $MgSO_4$ and the outside buffer containing 50 mM KPi pH 6.0, was supplemented with 2 mM $MgSO_4$ and 500 nM valinomycin. To exchange the inside buffer, the proteoliposomes containing TAXIPM-QM were thawed and collected by ultracentrifugation for 30 min at 80,000 g at 15 °C. They were then resuspended in the desired inside buffer to a concentration of 20 mg/mL and five cycles of freezing in liquid nitrogen with subsequent thawing were performed. The proteoliposomes were then analyzed in a reconstitution efficiency test (section 3.4.12) and stored in liquid nitrogen until used for the additional reconstitution of a membrane-anchored protein (section 3.4.11).

3.4.11 Reconstitution of membrane-anchored protein to proteoliposomes

On the same day of the transport assay (section 3.5.5), the proteoliposomes containing the reconstituted membrane protein were thawed and extruded in in 50 mM KPi or NaPi pH 7.5, supplemented with 2 mM $MgSO_4$. Afterwards the proteoliposomes were mixed with SBP that was purified from the membrane of *E. coli* cells and size exclusion chromatography pure. Thereby ratios of reconstituted membrane protein to membrane-anchored protein of 1:0.1, 1:1 or 1:10 were set and the samples were incubated for 15 min at room temperature, rocking. The proteoliposomes were then collected by ultracentrifugation at 200,000 g for 30 min at 15 °C, resuspended in 50 mM KPi pH 7.5, supplemented with 2 mM $MgSO_4$ to a lipid concentration of 100 mg/mL and stored on ice until used for transport assays. In the optimization of the protocol, no additional freeze/thaw cycles were performed to ensure that the SBPs were surface tethered from the periplasmic side only.

3.4.12 Reconstitution efficiency test

Reconstitution efficiency was tested by solubilization and ultracentrifugation of the proteoliposomes. For this purpose, 50 µL of the prepared proteoliposomes at 20 mg/mL were thawed and adjusted to a final concentration of 1 % (w/v) DDM. To solubilize the proteins, the samples were incubated for 1 h on ice while stirring every 15 minutes. Afterwards the samples were ultracentrifuged at 15 ° C for 10 min at ~200000 g. The well-folded protein remained in the supernatant which was then transferred to a new tube and adjusted to a final concentration of 10 % (w/v) SDS and 5x SDS buffer, followed by analysis in a 12 % SDS gel. The efficiency was evaluated by quantifying the loss of the protein. For this purpose, band intensities were measured using the program ImageJ (Schneider, Rasband, and Eliceiri 2012) and the percentage of how much protein was lost during solubilization was calculated by comparing the samples from before and after ultracentrifugation.

3.4.13 *In vitro* biotinylation of avi-tag

For biotinylation of the TAXIPm-QM protein with C- and N-terminal avi-tag, a Bir-A based *in vitro* approach was followed (B. T. Kuhn et al. 2020). For this, 6 µM IMAC-pure protein was mixed with 9 µM D-biotin (ratio of 1:1.5) and supplemented with 10 mM MgOAc, 5 mM ATP and 300 nM BirA enzyme (molar ratio of target protein:BirA of 20:1). The biotinylation mix was incubated overnight at 4 °C, rotating. The next day, a Re-IMAC was performed to separate BirA from the biotinylated protein using 1 mL of in purification buffer equilibrated Ni-NTA, followed by SEC.

3.4.14 Protein separation

To separate proteins according to their size, protein samples were mixed with 5x reducing or non-reducing sample buffer and SDS-Polyacrylamide gel electrophoresis (SDS PAGE) was performed. Prior to this, SDS polyacrylamide gels were prepared with a resolving gel containing either 10 % or 12 % (v/v) acrylamide, 375 mM Tris pH 8.8, 0.1% SDS (w/v), 0.1% APS (w/v) and 6.6 mM TEMED and a stacking gel containing 5 % (v/v) acrylamide,125 mM Tris pH 6.5, 0.1 % SDS (w/v), 0.1% APS (w/v) and 6.6 mM TEMED. The gels were casted in a Mini-PROTEAN Tetra Handcast System (BioRad) and transferred to a Mini-PROTEAN Tetra Vertical electrophoresis cell (BioRad). The protein samples as well as a protein standard were applied, and the electrophoresis was performed for 60 to 90 minutes at 120 V. After complete separation the gels were washed in Milli-Q H_20, and proteins were visualized by *in gel* fluorescence (section 3.4.15) or Coomassie staining (section 3.4.16) or further used for *anti*-His Western blotting (section 3.4.17).

3.4.15 Protein visualization by *in gel* GFP fluorescence

Proteins with GFP or sfGFP fusion were visualized after separation in a SDS PAGE electrophoresis using the *in gel* GFP fluorescence. The GFP fluorescence signals were detected at the ImageQuant LAS 4000 using the excitation wavelength of 460 nm with the filter 510DF10 and the marker signals were detected using the excitation wavelength of 630 nm with the filter R670BP. To determine the correct protein size the images were combined using Power Point (Microsoft). To quantify and compare expression levels, signal intensities were measured using ImageJ (Schneider, Rasband, and Eliceiri 2012).

3.4.16 Protein visualization by Coomassie staining

To visualize proteins unspecific after separation in a SDS PAGE electrophoresis, the gels were incubated either for 2 h in a self-prepared Coomassie Brilliant blue solution, followed by a 3 h de-staining step or for 10 min in InstantBlue® Coomassie Protein Stain (ISB1L), in both cases rocking. The de-staining solution was exchanged regularly and replaced by Milli-Q H_2O in a final washing step. The gels which were treated with the InstantBlue® Coomassie Protein Stain were washed in Milli-Q H_2O only. Afterwards, the gels were digitalized using the trans illumination channel of the ImageQuant LAS 4000.

3.4.17 Protein visualization by *anti*-His Western blotting

To visualize His-tagged proteins specifically after separation in a SDS PAGE electrophoresis, *anti*-His Western blotting was performed. For this, PDVF membranes were activated for 15 sec in 100 % methanol, washed for 2 min in Milli-Q H_2O and pre-equilibrated for 15 min the Western Blot transfer buffer. Additionally, two filter paper were pre-equilibrated, also for 15 min the Western Blot transfer buffer. The gel and membrane were assembled in between the two filter papers in a Bio-Rad Trans-Blot Turbo machine. The protein transfer from the gel to the membrane was performed for 30 min at 25 V/0.1 A. Afterwards the membrane, gel and filter papers were disassembled, and the membrane was blocked using 2.5 % (w/v) milk powder in PBS buffer supplemented with 0.1 % (v/v) Tween-20 overnight at 4° C or for 1 h at room temperature. In the next step, the membrane was washed with PBST, transferred into a small plastic bag, supplied with 0.05U conjugated *anti*-His horseradish peroxidase antibody and incubated for 1 h at room temperature, rocking. The small plastic bag ensured contact between the proteins on the membrane and the antibody. The membrane was washed three times for five min with PBST and then transferred to the imaging system ImageQuant LAS 4000. 500 µL luminol and 500 µL hydrogen peroxide solution from the Immobilon Western HRP Substrate kit were evenly distributed on the membrane and the resulting chemiluminescence of His-tagged proteins was recorded in a time series from 1 to 8 min.

3.5 Functional assays

3.5.1 Differential scanning fluorimetry (DSF)

To identify potentially transported substrates of TRAP transporter, melting temperatures were determined using DSF (Vetting et al. 2015). As usually a stabilization of proteins occurs when ligands bind, the melting temperatures of the TRAP SBPs alone and supplemented with a variety of potential ligands were compared (Vetting et al. 2015). For this purpose, SBPs belonging to different TRAP transporter were purified by IMAC, followed by Re-IMAC and SEC (section 3.4.3). Potential ligands were dissolved in the SEC elution buffer so that ligand and protein were in the same buffer. 40 µM ligand were mixed with 10 µM of the respective SBPs in PCR tubes if not stated otherwise. To study the stabilization behavior of the SBP from TAXIPm-PQM in different substrate concentrations, the concentration of α-ketoglutarate was varied from 6.25 µM to 100 µM. 20x of the commercially available 5000x Sypro Orange concentrate was added to a final volume of 25 µL, adjusted to 10 % glycerol and the tubes were placed in a Rotor-Gene Q PCR cycler. In the corresponding software, the excitation was adjusted to 470 nm and the emission to 555 nm for recording the Sypro Orange fluorescence. The detector gain setting was adjusted to -3 to ensure that the temperature curves did not reach saturation and therefore allowed a turning point determination. Starting the method, the samples were incubated for 1 min at each 1 °C step with a temperature gradient from 25 up to 95 °C. After the measurement, the first derivative was used to analyze the melting temperatures in the software OriginPro 2019b (OriginLab Corporation).

3.5.2 Isothermal titration calorimetry (ITC)

To determine the affinity and binding parameters of ligands identified in DSF to the SBPs of the TAXIPm-P and TAXIMh-P protein ITC was performed (Velazquez-Campoy, Leavitt, and Freire 2010). For this purpose, the SBPs were purified by IMAC and Re-IMAC and adjusted to 10 % glycerol with subsequent storage at -80 °C (section 3.4.3). On the day of the experiment, the proteins were thawed and further purified by SEC in 20 mM Hepes, 150 mM NaCl at the different pH-values 5.5, 6.0 or 7.5. They were then concentrated to 30 µM, degassed, pre-warmed to 18 °C and filled into the reaction cell of the ITC machine. The ligand was solved in the same buffer as the respective protein preparation, adjusted to 600 µM and filled in the syringe of the ITC machine. When the reaction cell achieved 20 °C, the method started with titrating the ligand in 10 µL steps to the protein. Like this, the ligand concentration in the cell ranged from 4.3 µM to 111.4 µM. To record the raw heat exchange data for background correction, the ligands were titrated into buffer instead of protein solution. The binding isotherm was obtained from the integration of the heat pulses, normalized per mole of injection and fitted by a one-site binding model. Individual K_D-values were indicated. The software packages NITPIC, SEDPHAT and GUSSI were used to analyze and plot the data. The evaluation was performed using the software NITPIC for automatic baseline correction and subtraction of dilution heat, the software Sedphat to set the binding parameters to heteroassociation as we assume 1:1 binding and do a global fit and the software GUSSI for plotting the data (Brautigam et al. 2016).

3.5.3 Enzyme linked immunosorbent assay (ELISA)

For the selection of TAXIPm-QM specific nanobodies, the enzyme linked immunosorbent assay was performed by Christiane Ruse, University of Frankfurt under my supervision. For this, two Nunc Maxisorp Immunoplates for both libraries were coated with 100 µL of 5 µg/mL Protein A by incubation overnight at 4 °C. The next day, the solution was discarded, the wells were washed with 250 µL TBS, 300 µL TBS-BSA were added and incubated for 30 min at room temperature to block the remaining surface area of the plates. Afterwards, the wells were washed three times with 250 µL TBS, incubated for 20 min with 100 µL of a 1:2,000 anti-c-Myc antibody dilution in TBS-BSA-D (0.5 % (w/v) BSA, 0.1 % (w/v) DM) and then washed three times, using 250 µL TBS-D (0.261 % DM) per well. Then, the cell pellets of the expression cultures were resuspended in 100 µL periplasmic extraction buffer and incubated for 60 min on ice. Afterwards, 900 µL TBS, supplemented with 1 mM $MgCl_2$ were added and the plates were centrifuged for 20 min at 3,200 *g* at 4 °C. 20 µL of the supernatant containing the periplasmic extracts with the nanobodies or the sybody Sy45 were mixed with 80 µL of TBS-BSA-D and applied to the ELISA plates. Each sample was transferred to two wells so that the binding of the nanobodies could be tested against both the target protein TAXIPm-QM and a negative control. Then the plates were incubated for 20 min and washed with 250 µL TBS-D three times. After this, 100 µL of a 50 nM solution of previously isolated and biotinylated target TAXIPm-QM protein or the negative control UraA in TBS-BSA-D were added. So UraA was used as negative control testing unspecific binding exemplarily with several TAXIPm-QM specific nanobodies and additionally as positive control where the UraA specific sybody Sy45 was applied, both UraA and Sy45 kindly provided by Dr. Benedikt Kuhn, MPI CBG Dresden (Eric Geertsma Group, www.mpi-

cbg.de) (B. T. Kuhn et al. 2018; B. T. Kuhn 2020). Additionally, several wells were treated without any target protein to later set a background threshold. The samples were incubated for 20 min and washed three times with 250 µL TBS-D per well. Then, 100 µL of a 1:5,000 dilution of a streptavidin-horseradish peroxidase polymer in TBS-BSA-D were applied, followed by incubation for 20 min and three times washing with 250 µL TBS-D. Afterwards, 100 µL of ELISA developing buffer were added to the samples. The plates were monitored until some samples changed to blue color of which the absorbance was measured using a Tecan Microplate reader at 650 nm. Thereby, the samples with signals with 1.3x higher absorbance than background were classified as significant. From these, the plasmids were isolated from pre-culture plates using the Zippy Plasmid Miniprep Kit and sequenced using the pBAD_forward sequencing primer at the Microsynth AG, Göttingen. From the results, the nanobody DNA sequences were evaluated. DNA sequences corresponding in size to nanobodies and showing variations in the three CDR regions were identified using a Python program, kindly provided by Dr. Eric Geertsma, MPI CBG Dresden (Eric Geertsma Group, www.mpi-cbg.de) and assigned as most promising candidates for following experiments.

3.5.4 Radioactive transport assay into whole cells

In vivo radioactive transport assays of α-ketoglutarate by TAXIPm-PQM, fumarate by TAXIMh-QM and N-acetylneuraminic acid by SiaPQM were performed into *E. coli* JW2571-1 (*ΔkgtP*), *E. coli* IMW424 (*ΔdctA, ΔdctB*) and *E. coli* JW3193-1 (*ΔnanT*), respectively. Pre-cultures with the respective cells carrying the genes of TAXIPm-PQM, TAXIMh-PQM or SiaPQM on the pBXC3sfGH or pEXC3sfGH plasmids were grown overnight in LB medium, supplemented with 100 µg/mL ampicillin. The next day, 10 mL of the same medium were inoculated with 1 % of the pre-culture. The cells were grown in tubes with gas-permeable lid at 37 °C until an OD_{600} of 0.6 to 0.8 was reached, followed by induction with 0.01 % arabinose for the pBXC3sfGH or 1 mM IPTG for the pEXC3sfGH plasmid. The proteins were overexpressed for 2 h at 37 °C, shaking. Afterwards all steps were performed on ice or at 4 °C. The cells were harvested for 10 min at 2,500 g, washed twice using 50 mM KPi, pH 7.5, supplemented with 2 mM $MgSO_4$ and adjusted to an OD_{600} of 100. Approximately 1 mg total protein was transferred to a new tube estimating the amount from OD_{600} = 1 corresponds to 0.3 mg/mL protein (Geertsma, Groeneveld, et al. 2008). These cells were stored at -20 °C until the analysis of protein expression levels (section 3.4.15). To perform the transport assay, the water bath was pre-warmed to 30 °C in which the reaction tube was placed containing freshly prepared 50 mM KPi buffer pH 7.5 or pH 6.0, supplemented with 2 mM $MgSO_4$, 0.1 % (w/v) glucose and 2 µM [^{14}C]-α-ketoglutarate, [^{14}C]-fumarate or [^{3}H]-N-acetylneuraminic acid. The mix was pre-incubated for 2 min. To dissipate the proton gradient, 10 µM Carbonyl cyanide m-chlorophenylhydrazone (CCCP) was added to individual tubes, followed by an additional incubation for 2 min. Then, the cells carrying overexpressed protein were added to a final OD of 3 to start the transport while stirring. After a time series of 0.5, 1, 2 and in some cases 4, 6 and 8 min, the transport reaction was stopped by transferring 100 µL into 2 mL ice-cold 50 mM KPi buffer, supplemented with 2 mM MgSO4 and immediate filtration over a 0.45 µm pore size nitrocellulose acetate filter. Additionally, 100 µL containing 50 mM KPi buffer pH 7.5 or pH 6.0, supplemented with 2 mM $MgSO_4$, 0.1 % (w/v) glucose

and 2 µM of the respective substrate were prepared and the cells were added to the side of the tube to be immediately flushed with the ice-cold stop buffer to resemble the 0 min data point. The filters were then washed with another 2 mL ice-cold buffer and dissolved in 4 mL scintillation liquid overnight. The next day, the radioactivity was measured using a scintillation counter. An additional total count measurement of the sample without filtration served as reference value. For kinetic characterization and determination of K_m/v_{max}-values, differing substrate concentrations were used. To achieve concentrations of up to 102.4 µM, radioactive substrate, referred to as hot substrate, was mixed with non-radioactive substrate, referred to as cold substrate, and the values measured in the scintillation counter were extrapolated to determine the transport activity. For the analysis and quantification of TAXIPM-PQM and TAXIMh-PQM expression levels from the *E. coli* cells used for radioactive transport assays, cells were thawed and lysed in a buffer containing 50 mM KPi pH 7.5, 150 mM NaCl, 10 % (w/v) glycerol, 1 mM $MgSO_4$, 1 mM PMSF and trace amounts of DNAse. ~300 mg glass beads were added prior to cell disruption and the Fast Prep cell disrupter was used twice for 0.5 min with 5 min incubation on ice in between. Afterwards 100 µL of the supernatant were incubated for 1 h with 1 % DDM and ultracentrifuged for 10 min at 100,000 g to pellet aggregated protein. To analyze the protein quality, the samples from before and after the ultracentrifugation were separated in a 12 % SDS gel from which the expression levels of the proteins were visualized by recording the *in gel* GFP fluorescence (section 3.4.15), followed by quantification using the software ImageJ (Schneider, Rasband, and Eliceiri 2012).

3.5.5 Radioactive transport assay into proteoliposomes

In vitro radioactive transport assays were performed for SiaPQM, TAXIPm-PQM and TAXIMh-PQM. In all experiments for SiaPQM and TAXIMh-PQM the transmembrane domains QM were reconstituted into liposomes (section 3.4.9) and P was supplied as a soluble protein (section 3.4.3). For TAXIPm-PQM, the same protocol was performed initially and then optimized to reconstituted TAXIPm-QM with membrane-anchored TAXI-Pm-P (section 3.4.11). On the day of the *in vitro* transport experiment, the QM proteoliposomes were extruded 11x using 400 nm pore size polycarbonate filters, followed for individual TAXIPm samples by reconstitution of the SBP P. In both cases, the proteoliposomes were then collected by ultracentrifugation for 30 min at 80,000 g at 15 °C. Afterwards they were resuspended to 100 mg/mL in 50 mM KPi or NaPi at pH 6.0 or 7.5, supplemented with 2 mM $MgSO_4$ to stabilize the lipids. To perform the transport assay, the water bath was pre-warmed to 20 °C in which the reaction tube was placed containing 50 mM KPi or 50 mM NaPi buffer at pH 6.0 or 7.5, supplemented with 2 mM MgSO4 and 5 µM [^{3}H]-N-acetylneuraminic acid, 30 µM [^{14}C]-fumarate or 20 µM [^{14}C]-α-ketoglutarate in the *in vitro* transport assays for SiaPQM, TAXIMh-PQM or TAXIPm-PQM, respectively. To mimic a negative membrane potential, 500 nM valinomycin was added to individual samples containing proteoliposomes with 50 mM KPi pH 7.5 inside buffer, supplemented with 2 mM $MgSO_4$ and 50 mM NaPi pH 6.0 outside buffer, supplemented with 2 mM $MgSO_4$. To achieve a positive membrane potential, the inside buffer of the QM containing proteoliposomes was exchanged to 50 mM NaPi pH 7.5, supplemented with 2 mM $MgSO_4$ prior to the reconstitution of the SBP P (section 3.4.10) and the transport assay was performed with 50 mM

KPi pH 6.0 outside buffer, supplemented with 2 mM $MgSO_4$ and 500 nM valinomycin. In all samples, the buffer containing the radioactive substrate was pre-incubated for 1 min, then the proteoliposomes containing 4.2 µg QM protein per data point were added to start the transport reaction while stirring. In most cases, a time series with data points at 0.5, 1, 2, 4 and 6 min was recorded which required a total reaction volume of 600 µL containing 25 µL proteoliposomes. The transport was then stopped by transferring 100 µL aliquots into 2 mL ice-cold 50 mM KPi or NaPi buffer at 7.5, which was supplemented with 2 mM $MgSO_4$ and immediate filtration over a 0.45 µm pore size nitrocellulose acetate filter. Additionally, 100 µL containing 50 mM KPi buffer pH 7.5 or pH 6.0, supplemented with 2 mM $MgSO_4$ and the respective substrate were prepared and the proteoliposomes were added to the side of the tube to be immediately flushed with the ice-cold stop buffer to resemble the 0 min data point. The filters were again washed with 2 mL ice-cold buffer and dissolved in 4 mL scintillation liquid overnight. The next day, the radioactivity was measured using a scintillation counter. An additional total count measurement of the sample without filtration served as reference value. To verify the presence of the proteins during the transport assay, a defined volume of each sample was flash frozen and stored in liquid nitrogen until being solubilized and evaluated from a gel as described for the reconstitution efficiency (section 3.4.12).

4 Results

4.1 Homolog screening of TRAP transport systems

To identify unknown proteins with superior biochemical properties, a homolog screening of TRAP transport systems was performed. It is advantageous to have a large selection of proteins at the beginning of the screening, as membrane proteins are generally difficult to produce in a folded state and often have poor stability in detergent. Homologs were selected based on their similarity with TRAPRc-DctPQM from *Rhodobacter capsulatus* which was the first discovered TRAP transport system (Jacobs et al. 1996). Based on BLAST, we selected homologs from proteins with amino acid similarities to TRAPRc-DctQM ranging from 11 to 89 % (Figure 20) (Madden 2002). This resulted in an initial set of 72 TRAP homologs which contains TRAP-TRAP, TRAP-TAXI and TRAP-TPAT proteins (Figure 20). The aim in this study was to select homologs part of the TAXI family that can be characterized, as this is the least studied group within these tripartite systems. We specifically focused on those systems that have fused QM domains to prevent loss of one of the subunits during the purification procedure. The number of candidates was additionally reduced by selecting TRAP-QM domains in our homolog search for which we could unambiguously identify the associated SBP in the same operon (Figure 109, section 6).

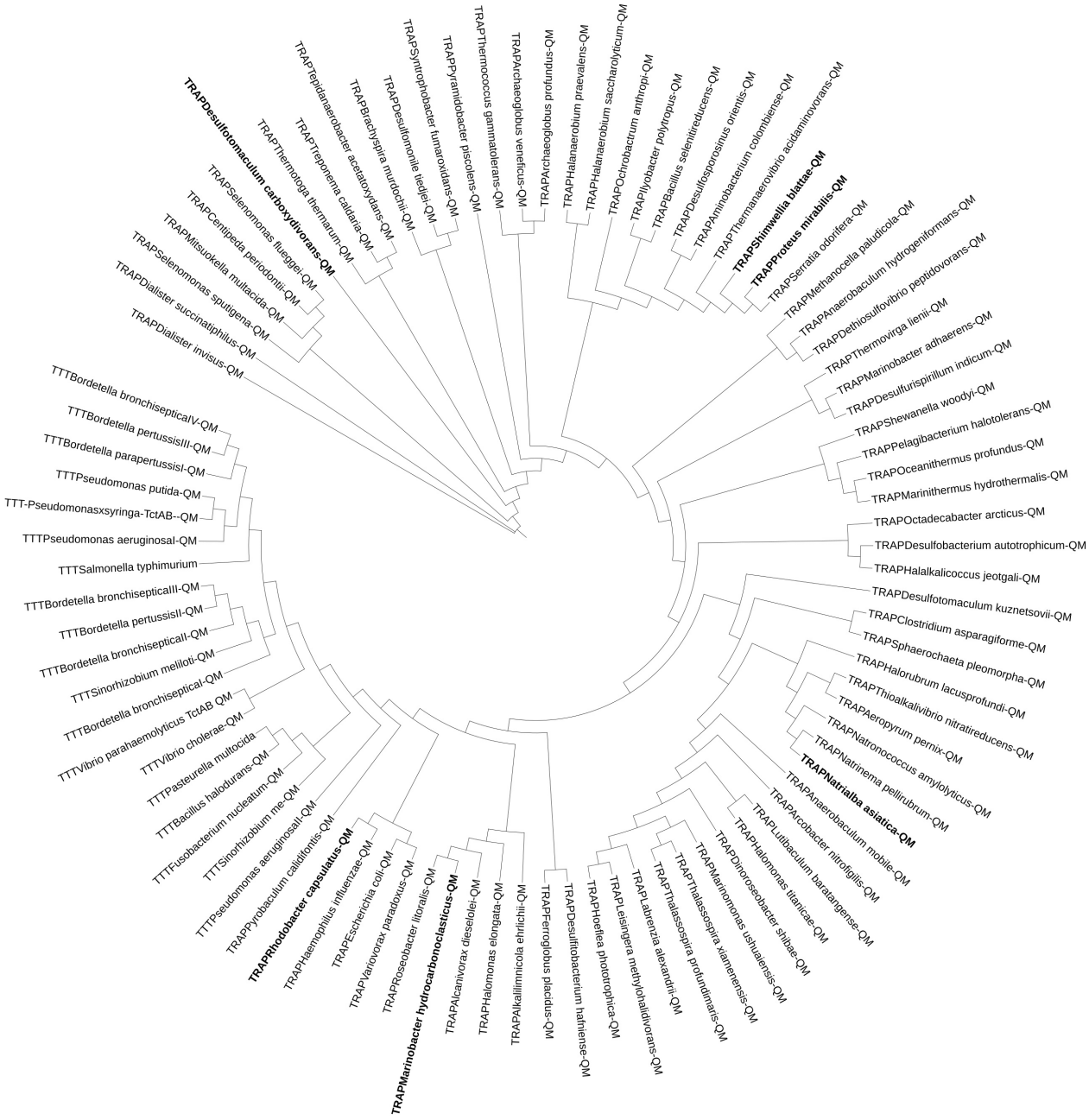


Figure 20 **Phylogenetic tree of proteins homologues to TRAPRc-DctQM from *Rhodobacter capsulatus*.** Sequence alignments of TRAPRc-DctQM homologs were carried out using Clustal Ω (Sievers et al. 2011). TTT proteins were included to illustrate their sequence dissimilarity with TRAP proteins. The corresponding phylogenetic tree was created using NGPhylogeny (Lemoine et al. 2019). *Rhodobacter capsulatus* and five homologs that were later selected in expression screenings are highlighted in bold: *Desulfotomaculum carboxydivorans, Shimwellia blattae, Proteus mirabilis, Natrialba asiatica* and *Marinobacter hydrocarbonoclasticus.*

For the selected homologs, we subsequently amplified the QM membrane-domains to overexpress and detergent-solubilize them in the further course in analytical expression tests. To aid simultaneous protein detection and an estimation of the folding state, we determined whether expression of the target proteins with a C-terminal GFP was feasible (Geertsma, Groeneveld, et al. 2008). The C-terminal GFP fusion would not only allow detection of expressed target protein but also allow the determination of the folding state of the target protein. Using the approach requires that the protein has a cytoplasmic C-terminus, as GFP folds well and becomes fluorescent in the cytoplasm only

(Drew et al. 2003). If the C-terminus of the target protein is located in the periplasm, however, than an adapted GFP version, the so-called superfolder GFP (sfGFP) can be used which still allows for detection of the target protein but does not serve as folding indicator (Dinh and Bernhardt 2011). In the absence of any high-resolution TRAP protein structures, we predicted the location of several proteins of the C-terminus using Topcons (Figure 21) (Tsirigos et al. 2015). From the total analysis more than 80 % were predicted to contain a cytoplasmic C-terminus.

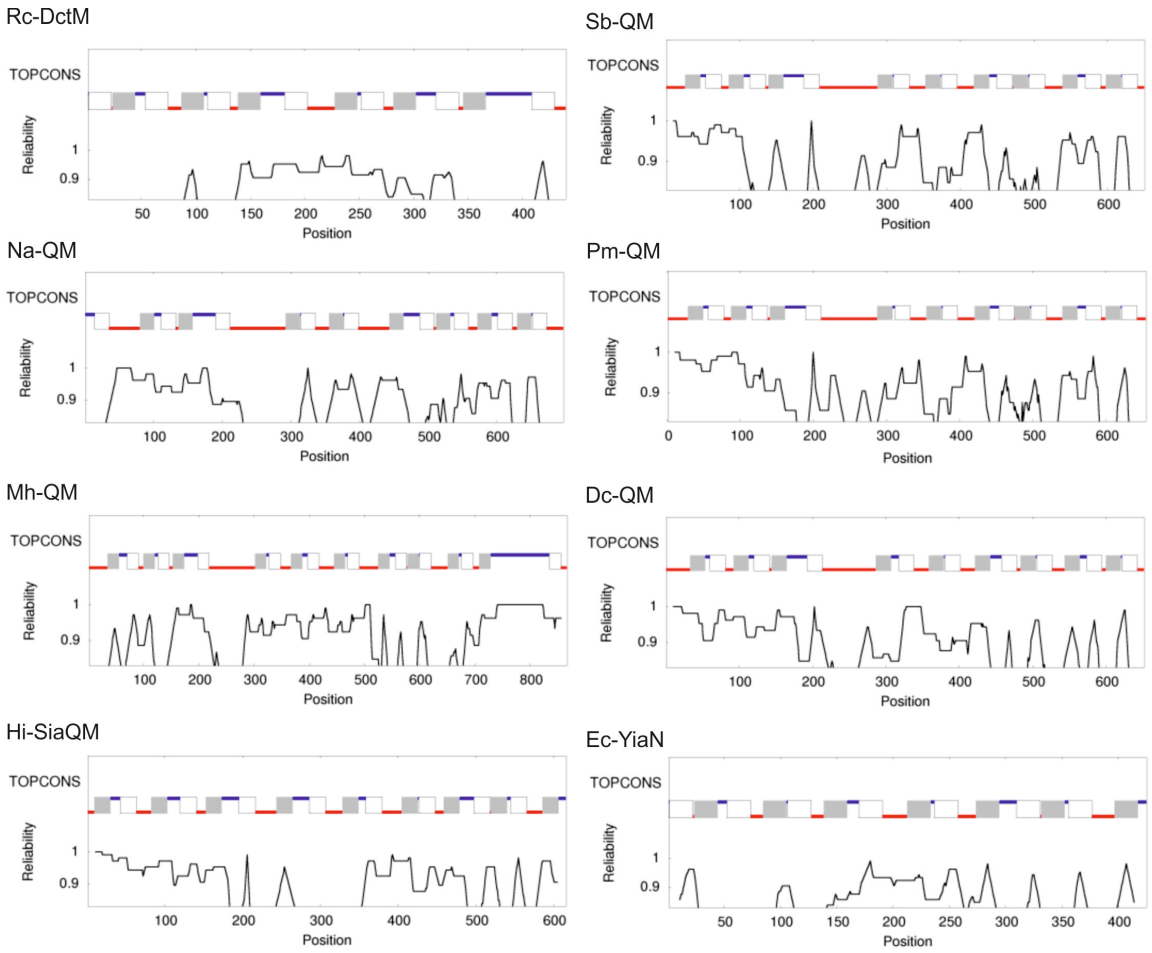


Figure 21 **Topcon analysis of TRAP membrane domains.** To evaluate how likely the C-terminus of TRAP transporters is periplasmatic or cytoplasmatic, a Topcon analysis was performed. The amino acid sequences of the large membrane domains were used from *Rhodobacter capsulatus* (Rc-DctM), *Shimwellia blattae* (Sb-QM), *Natrialba asiatica* (Na-QM), *Proteus mirabilis* (Pm-QM), *Marinobacter hydrocarbonoclasticus* (Mh-QM), *Desulfotomaculum carboxydivorans* (Dc-QM), *Haemophilus influenzae* (Hi-QM) and *Escherichia coli* (Ec-YiaN), respectively. The x-axis shows the amino acid positions and the y-axis the reliability of the Topcon prediction. The predicted topology is depicted above the graph. The grey and white boxes resemble transmembrane segments. The amino acid stretches that are predicted to be in the cytoplasm or in the periplasm are depicted as red or blue lines, respectively (Tsirigos et al. 2015).

As the bioinformatic analysis suggested that both cytoplasmic and periplasmic C-termini occurred within this homolog set, we avoided the expression of the proteins with a C-terminal GFP fusion and chose to later use the sfGFP fusion protein to aid protein expression and purification (Geertsma, Groeneveld, et al. 2008; Drew et al. 2003). To initially screen broadly for TRAP membrane domains that show highest expression in the most native form, we used expression vectors fused either to an N- or a C-terminal HRV 3C protease-cleavable decaHis-tag only (pBXC3H and pBXNH3) (Geertsma and Dutzler 2011). Due to the general poor expression levels of TRAP transmembrane domains and

strongly reduced cell growth upon induction, we performed these tests in 15 mL culture volumes to obtain sufficient biomass for starting our analysis with 2.5 mg total protein. In comparison, many membrane proteins can be detected when starting with an initial amount of 1 mg total protein and this amount can be easily obtained with 700 µl culture volumes (Marino et al. 2017). Following disruption of the cells by bead-beating, the proteins were immediately solubilized in 1% (w/v) DDM and subsequently submitted to ultra-centrifugation to remove aggregated material prior to detection by Western blotting. This procedure allowed us to detect well-folded protein only. In general, poor protein production was detected for the homologs upon expression with an N-terminal His-tag (Figure 22). The target proteins have calculated molecular weights between 68 and 80 kDa. Next to the TRAP-QM domain from *Haemophilus influenzae* (Hi), a signal was obtained for the TRAP-QM domains derived from *Desulfotomaculum carboxydivorans* (Dc), *Shimwella blattae* (Sb) and *Proteus mirabilis* (Pm).

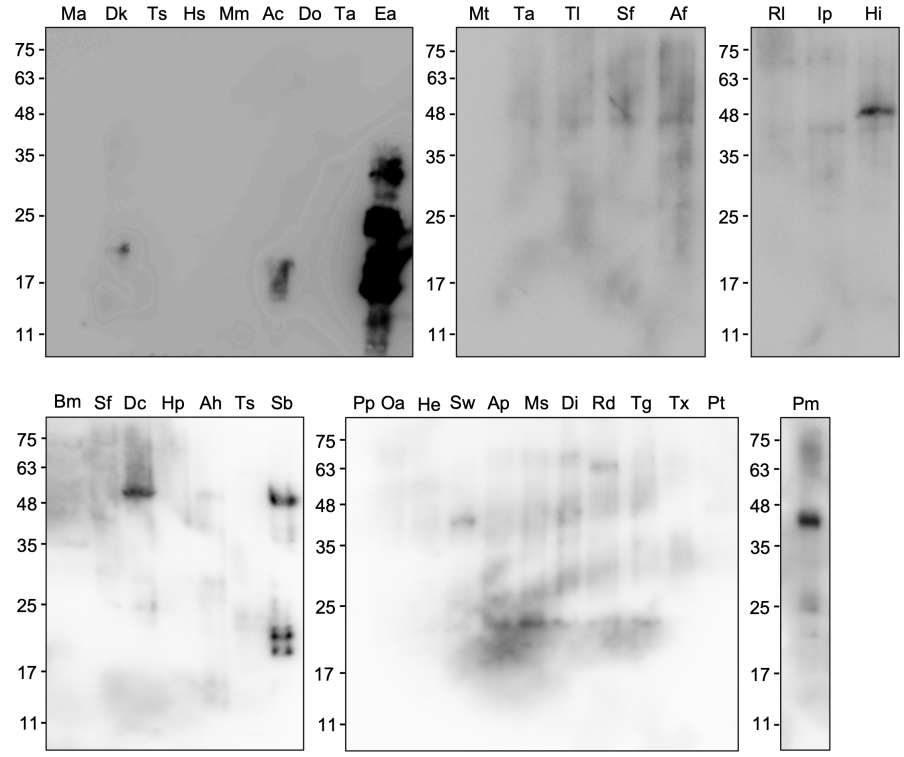


Figure 22 ***Anti*-His Western blot of TRAP-QM domains expressed from the pBXNH3 plasmid.** To select for TRAP-QM domains that can be expressed and solubilized, a protein expression screening was performed. The TRAP-QM domains were expressed in *E. coli* whole cells from the pBXNH3 plasmid at 25 °C overnight following induction with 0.01 % arabinose. Detergent-solubilized whole cell lysates were submitted to ultracentrifugation and the supernatants were analyzed by *anti*-His Western blots. The membrane domains derived from genomic DNA of the organisms from *Marinobacter adhaerens* (Ma), *Desulfotomaculum kuznetsovii* (Dk), *Thermococcus sibiricus* (Ts), *Hoeflea sp.* (Hs), *Mitsuokella multacida* (Mm), *Aeropyrum camini* (Ac), *Desulfosporosinus orientis* (Do), *Thermanaerovibrio acidaminovorans* (Ta), *Enterocloster asparagiformis* (Ea), *Marinithermus hydrothermalis* (Mt), *Thermanaerovibrio acidaminovorans* (Ta), *Thermovirga lienii* (Tl), *Selenomonas flueggei* (Sf), *Archaeoglobus fulgidus* (Af), *Roseobacter litoralis* (Rl), *Ilyobacter polytropus* (Ip), *Haemophilus influenzae* (Hi), *Brachyspira murdochii* (Bm), *Syntrophobacter fumaroxidans* (Sf), *Desulfotomaculum carboxydivorans* (Dc), *Halanaerobium praevalens* (Hp), *Acetomicrobium hydrogeniformans* (Ah), *Thermococcus sibiricus* (Ts), *Shimwellia blattae* (Sb), *Pyramidobacter piscolens* (Pp), *Octadecabacter arcticus* (Oa), *Halomonas elongata* (He), *Shewanella woodyi* (Sw), *Archaeoglobus profundus* (Ap), *Marinomonas sp.* (Ms), *Dialister invisus* (Di), *Roseobacter denitrificans* (Rd), *Thermococcus gammatolerans* (Tg), *Thalassospira xiamenensis* (Tx), *Pseudothermotoga thermarum* (Pt) and *Proteus mirabilis* (Pm).

However, using the pBXC3H plasmid as expression vector, we achieved in general higher expression levels and identified several TRAP-QM domains amenable to overexpression (Figure 23). In initial *anti*-His Western blots significant signals were obtained for the TRAP-QM domains deriving from *Desulfotomaculum carboxydivorans* (Dc), *Dialister invisus* (Di), *Thermococcus sibiricus* (Ts), *Thermovirga lienii* (Tl), *Shimwellia blattae* (Sb), *Natrialba asiatica* (Na), *Proteus mirabilis* (Pm) and *Marinobacter hydrocarbonoclasticus* (Mh).

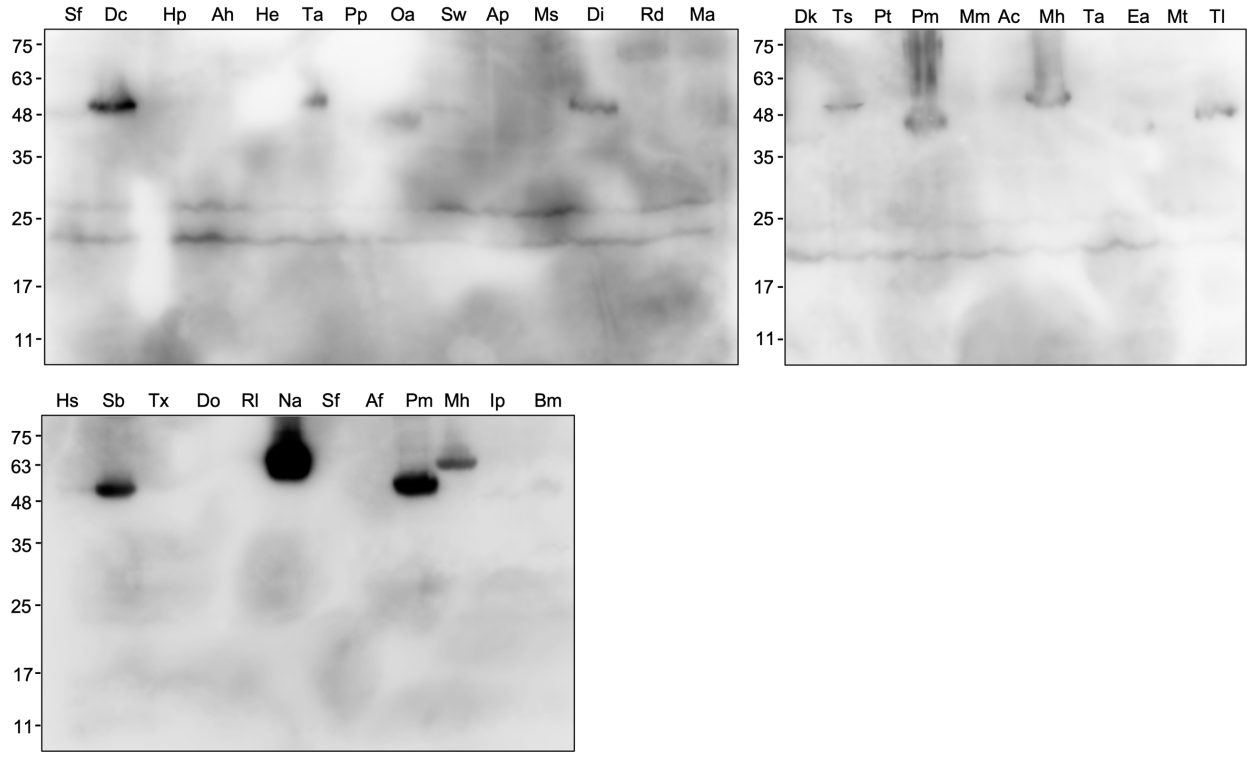
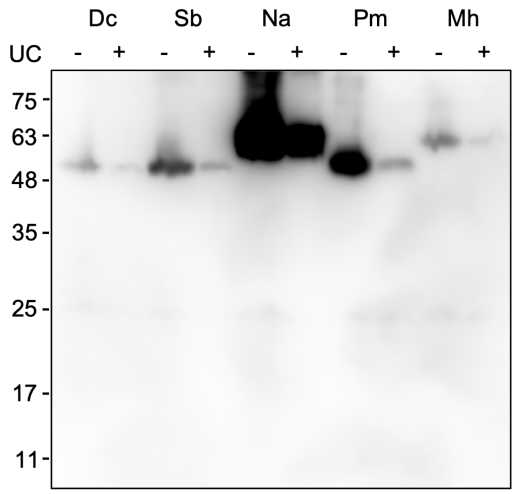


Figure 23 ***Anti*-His Western blot of TRAP-QM domains expressed from the pBXC3H plasmid.** To select for TRAP-QM domains that can be expressed and solubilized, a screening was performed. The TRAP-QM domains were expressed in *E. coli* whole cells from the pBXC3H plasmid at 25 °C overnight following induction with 0.01 % arabinose. Detergent-solubilized whole cell lysates were submitted to ultracentrifugation and the supernatants were analyzed by *anti*-His Western blots. The TRAP-QM domains derive from the following organisms in Western blot 1: *Selenomonas flueggei* (Sf), *Desulfotomaculum carboxydivorans* (Dc), *Halanaerobium praevalens* (Hp), *Acetomicrobium hydrogeniformans* (Ah), *Halomonas elongata* (He), *Thermanaerovibrio acidaminovorans* (Ta), *Pyramidobacter piscolens* (Pp), *Octadecabacter arcticus* (Oa), *Shewanella woodyi* (Sw), *Archaeoglobus profundus* (Ap), *Marinomonas sp.* (Ms), *Dialister invisus* (Di), *Roseobacter denitrificans* (Rd), *Marinobacter adhaerens* (Ma). In Western blot 2: *Desulfotomaculum kuznetsovii* (Dk), *Thermococcus sibiricus* (Ts), *Pseudothermotoga thermarum* (Pt), *Proteus mirabilis* (Pm), *Mitsuokella multacida* (Mm), *Aeropyrum camini* (Ac), *Marinobacter hydrocarbonoclasticus* (Mh), *Thermanaerovibrio acidaminovorans* (Ta), *Enterocloster asparagiformis* (Ea), *Marinithermus hydrothermalis* (Mt), *Thermovirga lienii* (Tl). In Western blot 3: *Halobacterium salinarum* (Hs), *Shimwellia blattae* (Sb), *Thalassospira xiamenensis* (Tx), *Desulfosporosinus orientis* (Do), *Ruegeria lacuscaerulensis* (Rl), *Natrialba asiatica* (Na), *Selenomonas flueggei* (Sf), *Archaeoglobus fulgidus* (Af), *Proteus mirabilis* (Pm), *Marinobacter hydrocarbonoclasticus* (Mh), *Ilyobacter polytropus* (Ip), *Brachyspira murdochii* (Bm).

Expression tests of TRAP-QM domains that were positive in the initial analysis were repeated and expression levels were directly compared. Protein overexpression could be reproduced, and reasonable amounts were obtained for five homologues proteins, named Dc from *Desulfotomaculum carboxydivorans*, Sb from *Shimwellia blattae*, Na from *Natrialba asiatica*, Pm from *Proteus mirabilis* and Mh from *Marinobacter hydrocarbonoclasticus* (Figure 24). The band intensity decreased for all five proteins after ultracentrifugation, showing that a large fraction of the protein was not functionally expressed. However, a significant amount of well-folded protein could be recovered (Figure 24).



Organism	Molecular weight of QM
Desulfotomaculum carboxydivorans	72 kDa
Shimwellia blattae	72 kDa
Natrialba asiatica	77 kDa
Proteus mirabilis	72 kDa
Marinobacter hydrocarbonoclasticus	80 kDa

Figure 24 ***Anti*-His Western blot of five expressed and solubilized TRAP-QM domains.** *E. coli* whole cells were used to express QM domains of different TRAP transporters. The expression of the proteins was performed at 25 °C for overnight from the pBXC3H plasmid following induction with 0.01 % arabinose. Detergent-solubilized whole cell lysates were submitted to ultracentrifugation. The protein samples from before (-) and after (+) ultracentrifugation (UC) were applied on a 12 % SDS PAGE followed by detection of the target proteins by *anti*-His Western blotting. The QM domains were derived from the following organisms: *Desulfotomaculum carboxydivorans* (Dc), *Shimwellia blattae* (Sb), *Natrialba asiatica* (Na), *Proteus mirabilis* (Pm) and *Marinobacter hydrocarbonoclasticus* (Mh). The table shows the molecular weight calculated from the amino acid sequences of the TRAP-QM domains from organisms corresponding to the protein samples on the *anti*-His Western blot.

4.2 TRAP-QM large scale purifications

To characterize proteins not only *in vivo*, but also *in vitro*, it is essential to purify them while maintaining the folded state. For this purpose, individual TRAP-QM domains were expressed in 1 to 5 L cultures and purified by IMAC, followed by SEC. The initial purifications of Dc-QM, Pm-QM and Mh-QM, from the organisms *Desulfotomaculum carboxydivorans*, *Proteus mirabilis* and *Marinobacter hydrocarbonoclasticus* were promising, as they showed the main peak in the SEC chromatogram at the expected elution volume between 12 and 14 mL. This equals a molecular weight range in the protein standard solution of approximately 50 to 100 kDa (Figure 110, section 6). The Na-QM domain from *Natrialba asiatica*, that was expressed and solubilized to high amounts in initial screenings (Figure 24), eluted to high amounts in the void peak (9 mL). However, Na-QM was still a promising candidate for characterization as it also eluted to high amounts at the expected volume between 12 and 14 mL (Figure 25). The initial test purifications were carried out by Melanie Engelin.

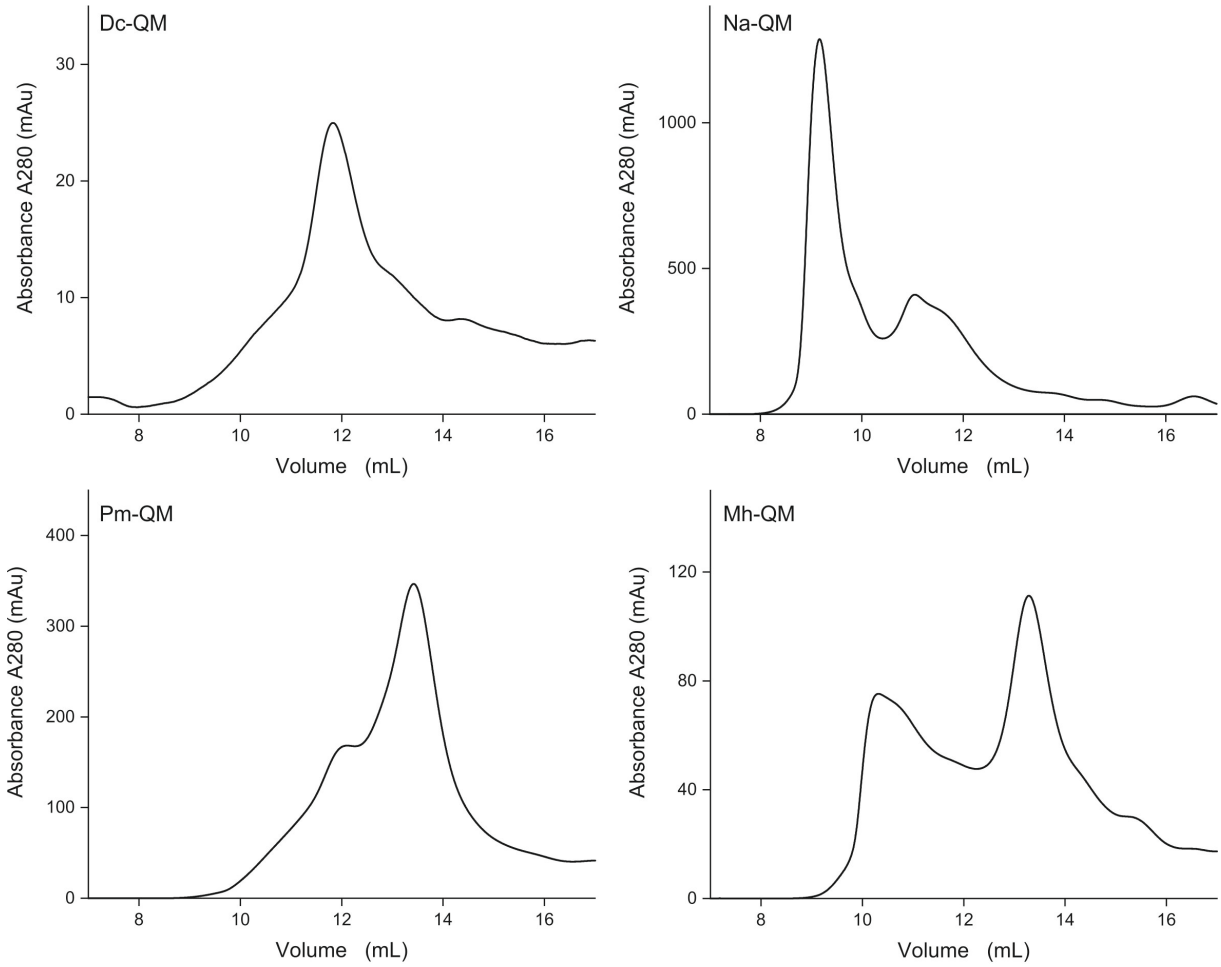


Figure 25 **Purification of the QM domains of selected transporter.** The QM domains of selected proteins from *Desulfotomaculum carboxydivorans* (Dc-QM), *Natrialba asiatica* (Na-QM), *Proteus mirabilis* (Pm-QM) and *Marinobacter hydrocarbonoclasticus* (Mh-QM) were expressed in *E. coli* MC1061, solubilized in 1.6 % DM and purified by IMAC followed by SEC as a preliminary test. The SEC were carried out on a Superdex 200 30/100 GL Increase column with a running buffer containing 20 mM Hepes pH 7.5, 150 mM NaCl, 0.2 % DM. The initial test purifications were carried out by Melanie Engelin, Goethe University Frankfurt. The plots were created using OriginPro 2019b (OriginLab Corporation).

4.3 Bioinformatic analysis of SBPs

Within the TRAP transporter family, three subgroups have been identified, named (TRAP-)TRAP, (TRAP-)TAXI and (TRAP-)TPAT, which are believed to differ structurally in the SBPs (Kelly and Thomas 2001). Functional *in vitro* characterization has so far only been reported for conventional TRAP transporters (Mulligan et al. 2009; Davies, Currie, and North 2023), so it was of particular interest to us to investigate proteins of these subgroups in detail. To classify the proteins that we selected in the initial homolog screening (section 4.1) a Clustal Ω Alignment was performed in which a large number of SBPs of conventional TRAP transporter and subgroups were compared (Figure 26) (Sievers et al. 2011). For comparison, SBPs of TTT transporters, whose SBPs are structurally similar to TRAPs, were additionally included. For the analysis, several amino acid sequences of annotated SBPs were aligned that were derived the NCBI database (Benson et al. 2013). In general, two clusters formed, one including the conventional TRAP SBPs combined with the TRAP-TPAT SBPs and the second containing the TTT and TRAP-TAXI SBPs. The first cluster included the SBP from the TRAP SBPs *Haemophilus influenzae* (Hi-SiaP; WP_005689956.1), *Vibrio cholerae* (Vc-SiaP; WP_000849284.1), *Rhodobacter capsulatus* (Rc-DctP; WP_013068721.1), *E. coli* (YiaO;

WP_162827541.1), *Xanthobacter autotrophicus* (Xa-P; ABS68597.1) and *Ruegeria pomeroyi* (Rp-P; WP_044028089.1) and the TRAP-TPAT SBP from *Treponema pallidum* (Tp-TatP; WP_010882401.1). The second cluster contained the SBPs from the TRAP-TAXI SBPs *Aeropyrum pernix* (Ap-P; WP_010867075.1) and *Thermus thermophilus* (Tt-GluBP; WP_011228475.1) and the TTT SBP from *Bordetella pertussis* (Bp-BugE; WP_003814324.1) and *Polaromonas sp. JS666* (Ps-TctC; WP_011484418.1). The five SBPs selected in our study from *Desulfotomaculum carboxydivorans* (Dc-P), *Shimwellia blattae* (Sb-P), *Natrialba asiatica* (Na-P), *Proteus mirabilis* (Pm-P) and *Marinobacter hydrocarbonoclasticus* (Mh-P) positioned among the TRAP-TAXIs and TRAP-TTTs (Figure 26). We could exclude, that they belong to the TTT family since the protein sequences of QM domains from TTT transporters differ significantly from those of the TRAPs and our previous Clustal Ω alignments of the QM domains showed that these five transporters belong to the TRAP and not to the TTT family (Figure 20, section 4.1) (Sievers et al. 2011). Therefore, we assigned Dc-P, Sb-P, Na-P, Pm-P and Mh-P to the TRAP-TAXIs (Figure 26).

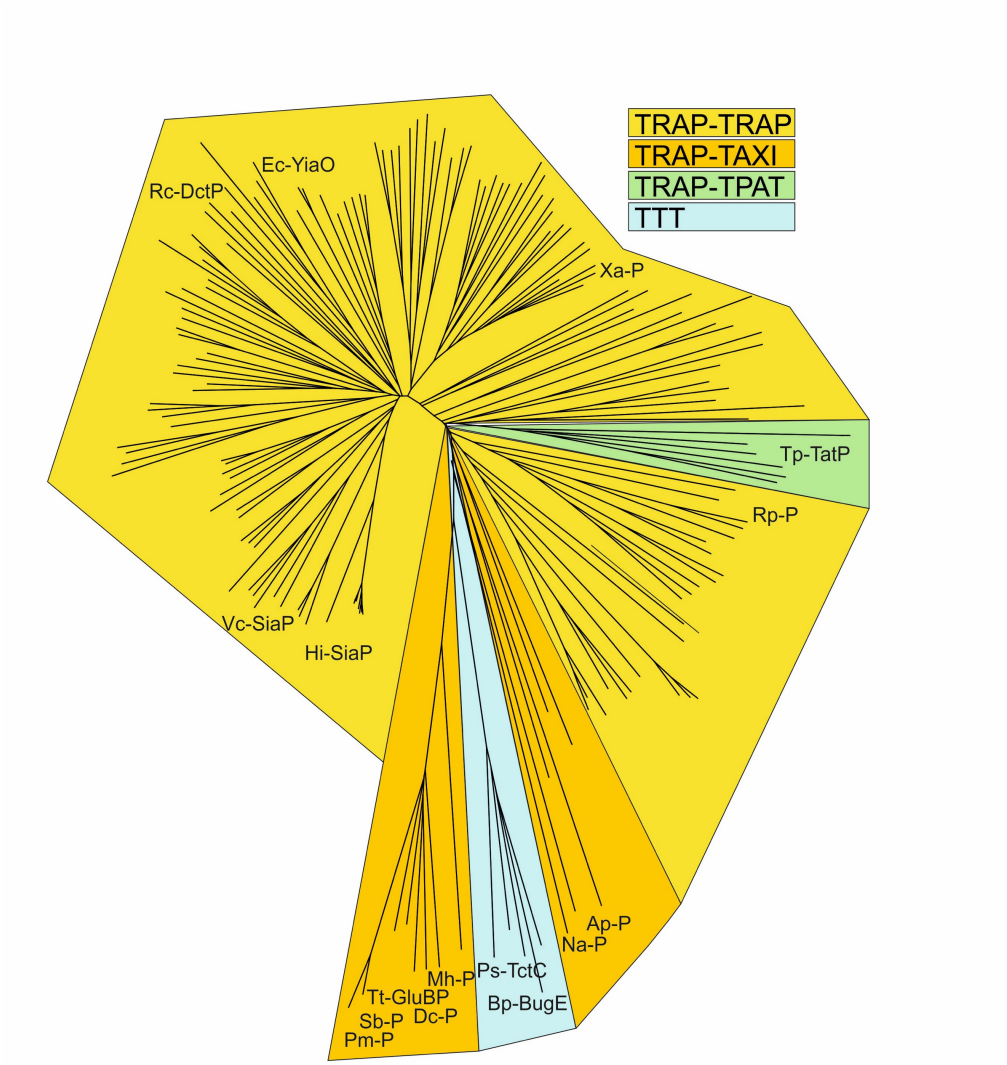


Figure 26 **Phylogenetic radiation tree of TRAP, TRAP-TAXI, TRAP-TPAT and TTT SBPs.** To assign the SBP from *Desulfotomaculum carboxydivorans* (Dc-P), *Shimwellia blattae* (Sb-P), *Natrialba asiatica* (Na-P), *Proteus mirabilis* (Pm-P) and *Marinobacter hydrocarbonoclasticus* (Mh-P) to a SBP class, amino acid sequence alignments were performed which included previously annotated SBPs from *Haemophilus influenzae* (Hi-SiaP; WP_005689956.1), *Vibrio cholerae* (Vc-SiaP; WP_000849284.1), *Rhodobacter capsulatus* (Rc-DctP; WP_013068721.1), *E. coli* (YiaO; WP_162827541.1), *Xanthobacter autotrophicus* (Xa-P; ABS68597.1), *Treponema pallidum* (Tp-TatP; WP_010882401.1), *Ruegeria pomeroyi* (Rp-P; WP_044028089.1), *Aeropyrum pernix* (Ap-P; WP_010867075.1), *Bordetella pertussis* (Bp-BugE; WP_003814324.1), *Polaromonas sp. JS666* (Ps-TctC; WP_011484418.1) and *Thermus thermophilus* (Tt-GluBP; WP_011228475.1) (Vetting et al. 2015; Benson et al. 2013). Clustal Ω and MEGA were used to carry out the alignments and to create the phylogenetic radiation tree, respectively (Sievers et al. 2011; Hall 2013).

Additionally, we superimposed the representative SBP models from TRAP, TRAP-TAXI, TRAP-TPAT and TTT SBPs on TAXIPm-P *Proteus mirabilis*. The high protein sequence similarity of TAXITt-GluBP and TAXIPm-P, representated by the low RMSD value of 3.2 Å, assigns TAXIPm-P to the TAXI SBPs (Figure 27).

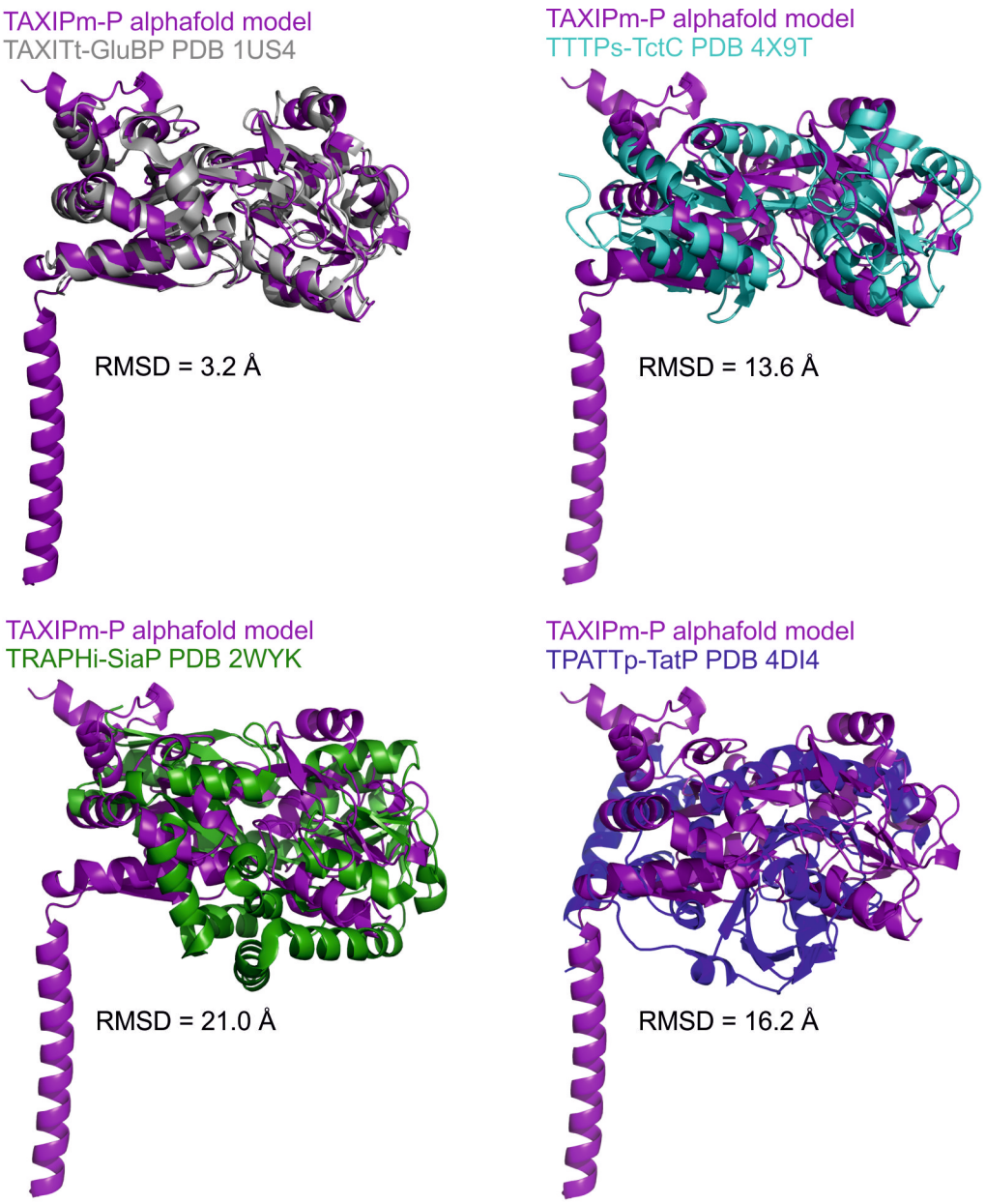


Figure 27 **Superimposition of representative SBP models on TAXIPm-P**. To assign the SBP from TAXIPm-P to a group, representative SBP models from TRAP, TRAP-TAXI, TRAP-TPAT and TTT were superimposed on TAXIPm-P. The TAXI SBP TAXITt-GluBP from *Thermus thermophilus* (PDB code: 1US4) is shown in grey, the TTT SBP TTTPs-Tct from *Polaromonas sp. JS666* (PDB code: 4X9T) is shown in teal, the TPAT SBP TPATTp-TatP from *Treponema pallidum* (PDB code: 4DI4) is shown in blue and the TRAP SBP TRAPHi-SiaP from *Haemophilus influenzae* (PDB code: 2WYK) is shown in green. The superimposition was performed and the RMSD values calculated using Pymol 1.8.4.1 (Schrödinger, LLC).

We further analyzed the N-terminal signal sequences of the SBPs as these determine whether the protein occurs soluble in the periplasm or attached to the membrane by an anchor. This might bring functional next to structural insights. In Gram-negative organisms, signal peptides predominantly direct the proteins to the Sec- or Tat-translocation pathway where they are cleaved by signal peptidase I after passing the membrane (Auclair, Bhanu, and Kendall 2012). Mostly found in Gram positive organisms and archaea signal peptides are cleaved by signal peptidase II in the presence of a lipid anchor (Auclair, Bhanu, and Kendall 2012). In the absence of recognition sites for signal

peptidases, the peptides are not cleaved which often results in the whole protein remaining attached to the membrane by a transmembrane segment.

For the analysis of the N-terminal signal sequences, we performed a SignalP analysis of 34 SBPs, thereby differentiating (TRAP-)TRAP, (TRAP-)TAXI, (TRAP-)TPAT and TTT SBPs. The analysis resulted in signal peptidase I cleavage site predictions for all TRAP and most TPAT SBPs. For all other types of SBPs the predictions varied with either no clear result or predictions for both signal peptidase I or II cleavage sites leading to the usage of the Sec- or Tat-translocation pathway or surface tethering mediated by a lipid anchor (Table 34).

Table 34: **SignalP predictions for cleavage of signal sequences in TRAP, TAXI, TTT und TPAT SBPs.** The table shows the results of an analysis of 34 signal sequences performed using SignalP. Predictions for signal peptidase I cleavage sites using the Sec- or Tat- translocation pathway are abbreviated with Sec/SPI or Tat/SPI, respectively. The prediction for a signal peptidase II cleavage site leading to a lipid anchor is abbreviated with LIPO (Sec/SPII) and all unclear results with 'other'.

	TRAP	TAXI	TTT	TPAT
Sec/SPI	10/10	3/12	-	6/7
Tat/SPI	-	1/12	-	-
LIPO (Sec/SPII)	-	1/12	-	-
Other	-	7/12	6/6	1/7

For the representative proteins of (TRAP-)TRAP, (TRAP-)TAXI, (TRAP-)TPAT and TTT SBPs a clear signal peptidase I clevage site was obtained for TRAPHi-SiaP from *Haemophilus influenzae*, TAXITt-GluBP from *Thermus thermophilus* and TTTPs-TctC *Polaromonas sp. JS666.* The signal sequence of TPATTp-TatP from *Treponema pallidum* could not be assigned to a known signal peptide and therefore labeled with 'other' which suggests an anchoring of the SBP to the membrane (Figure 28) (Almagro Armenteros et al. 2019).

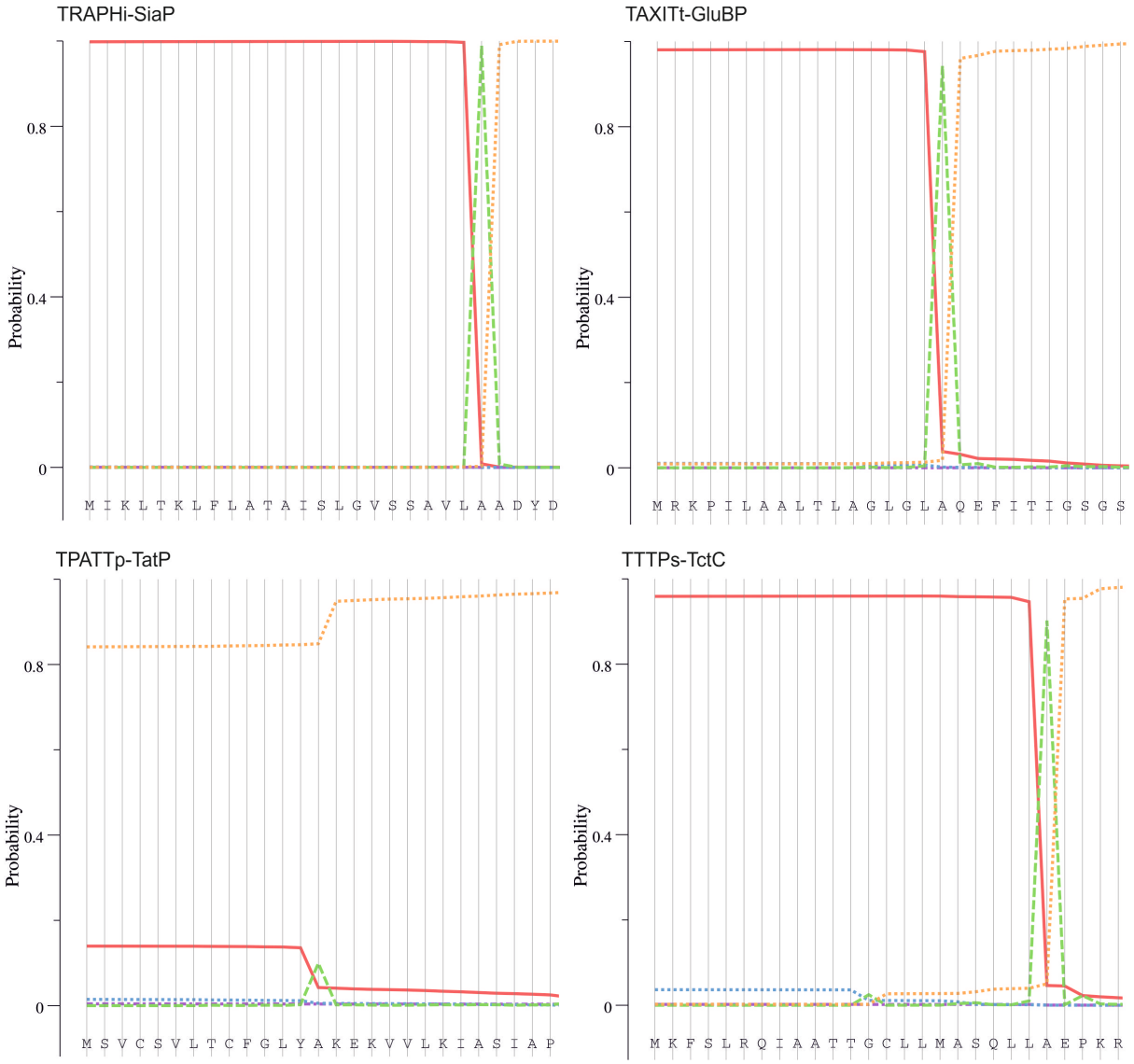


Figure 28 **Signal sequence analysis in different types of SBPs**. The N-terminal amino acid sequences of SBPs from *Haemophilus influenzae* (TRAPHi-SiaP), *Thermus thermophilus* (TAXITt-GluBP), *Treponema pallidum* (TPATTp-TatP) and *Polaromonas sp. JS666* (TTTPs-TctC) are shown on the x-axis. The y-axis shows the probability of the different types of signal sequences thereby differentiating the Sec- and Tat-translocation pathway including signal peptidase I cleavage sites, shown in red and purple lines, respectively. The prediction for a lipid anchor including a signal peptidase II cleavage site is shown as a blue line or other unclear results, shown as orange line. The positions of the cleavage sites are additionally displayed as a green line (Almagro Armenteros et al. 2019).

Selected TAXI sequences from *Desulfotomaculum carboxydivorans* (TAXIDc-P), *Shimwella blattae* (TAXISb-P), *Natrialba asiatica* (TAXINa-P), *Proteus mirabilis* (TAXIPm-P) and *Marinobacter hydrocarbonoclasticus* (TAXIMh-P) were additionally analyzed using SignalP. For all of them the prediction was ‘unclear’ as the sequences could not be assigned to known signal peptides (Figure 29). This hints that all the analyzed SBPs contain N-terminal transmembrane segments through which they remain surface-tethered to the membrane.

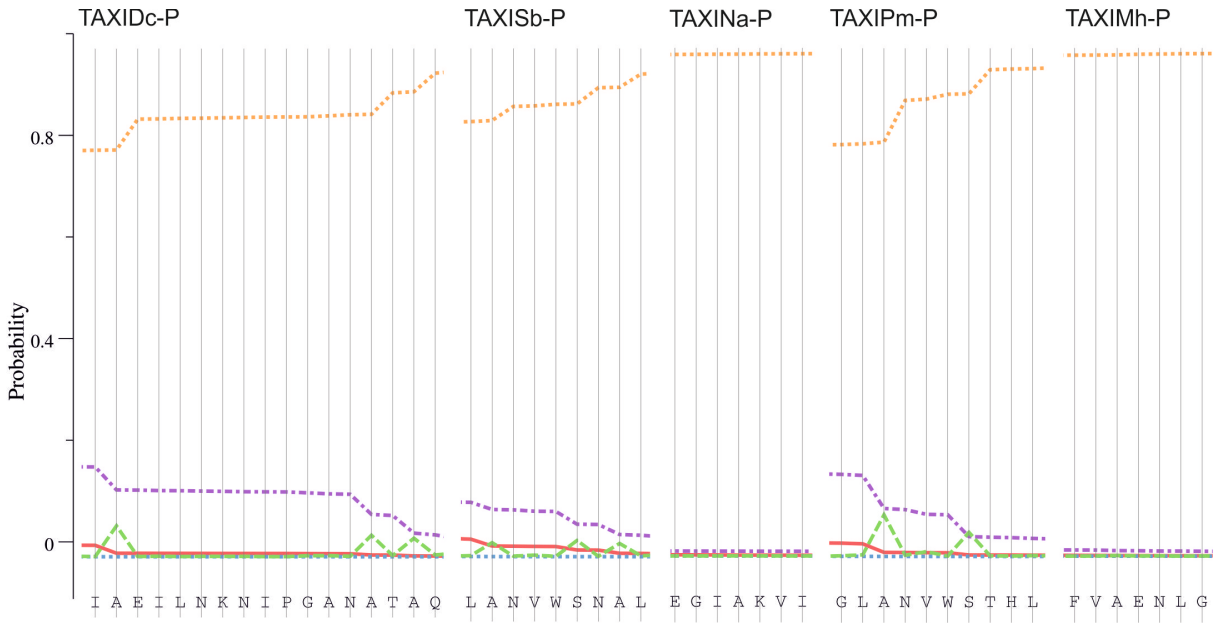
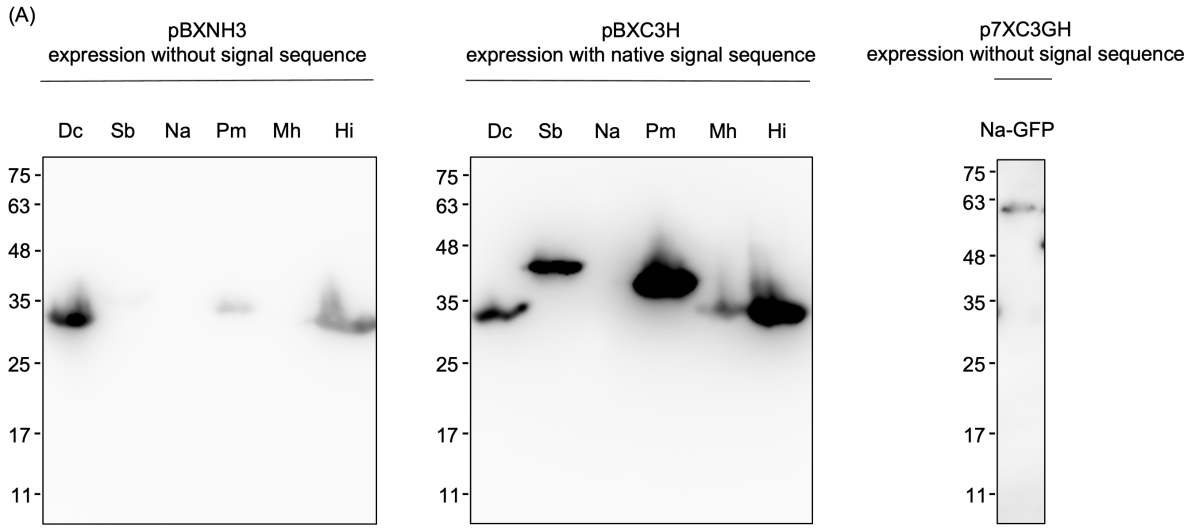


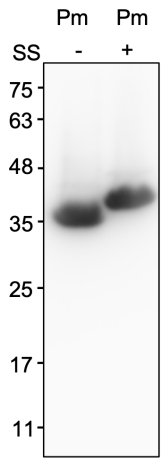
Figure 29 **Signal sequence analysis of selected TAXI SBPs**. The relevant sections of the protein sequences for the evaluation of the signal peptide are shown from the SignalP output. The protein sequences derive from the SBPs from *Desulfotomaculum carboxydivorans* (TAXIDc-P), *Shimwella blattae* (TAXISb-P), *Natrialba asiatica* (TAXINa-P), *Proteus mirabilis* (TAXIPm-P) and *Marinobacter hydrocarbonoclasticus* (TAXIMh-P). The protein sequences are shown on the x-axis, respectively. The y-axis shows the probability of the different types of signal sequences thereby differentiating the Sec- and Tat-translocation pathway including signal peptidase I cleavage sites, shown in red and purple lines, respectively. The prediction for a lipid anchor including a signal peptidase II cleavage site is shown as a blue line or other unclear results shown as orange line. The cleavage sites in general are additionally displayed as a green line. Small sections of the protein sequences are shown as the predictions are unambiguous (Almagro Armenteros et al. 2019).

4.4 SBP expression analysis

For TRAP transporters, the complete functional unit is composed of the Q and M membrane domains plus the corresponding SBP P. Thus, after expressing the QM-domains, we determined whether the corresponding SBPs could be produced, either in the cytoplasm or in the periplasm following secretion by their native signal sequence. Low to high expression levels were detected for the SBPs from *Desulfotomaculum carboxydivorans* (Dc), *Shimwellia blattae* (Sb), *Natrialba asiatica* (Na), *Proteus mirabilis* (Pm) and *Marinobacter hydrocarbonoclasticus* (Mh) and protein production was most successful using the native signal sequence (Figure 30, A). As the SBP of *Natrialba asiatica* could not be detected under any condition, we additionally screened intracellular expression from the T7 promoter and as a C-terminal fusion to GFP (Figure 30, A) (Geertsma and Dutzler 2011). The constructs expressed with signal sequence should have the same size in the gel as the constructs without signal sequence, assuming that the *E. coli* Signal Peptidase I is able to recognize and cleave these. However, the SBPs of *Shimwella blattae* (Sb) and *Proteus mirabilis* (Pm) appear to migrate slower in the gel after expression with native signal peptide when compared to the corresponding constructs without signal peptide, suggesting that the signal peptides are not cleaved by signal peptidase I. Also, from the calculated molecular sizes of 34.8 kDa and 35.4 kDa for the cleaved SBPs of *Shimwella blattae* (Sb) and *Proteus mirabilis* (Pm) the proteins should run higher in the gel (Figure 30, C). This suggests that their signal peptides were not removed, leaving the SBPs anchored to the membrane. In contrast, the SBPs of *Desulfotomaculum carboxydivorans* (Dc) and *Marinobacter hydrocarbonoclasticus* (Mh) seem to migrate corresponding to their calculated molecular sizes of

33.6 and 33.4 kDa, suggesting that their signal peptides are cleaved (Figure 30, A, C). It should be noted that the previous sequence analysis by SignalP however could not assign their signal peptides to a signal peptidase I cleavage site (Figure 29, section 4.3). TRAPHi-SiaP from *Haemophilus influenzae* serves as reference as it has already been shown experimentally that here the signal sequence is removed by signal peptidase I in *E. coli* (Mulligan et al. 2009). The SBP of *Natrialba asiatica* (Na) could only be expressed without a signal sequence in a plasmid called p7XC3GH (Figure 30, A), so no information about the possible cleavage of the signal sequence can be inferred here. In a detailed analysis, we could show for the SBP of *Proteus mirabilis* (Pm), that the signal peptide was indeed not cleaved by the *E. coli* signal peptidase I (Figure 30, B). For this, an *anti*-His Western Blot was performed on the basis of a 10 % gel instead of the 12 % gel on the example, as the 10 % gel yields better resolution in the here important region between 30 and 40 kDa (Figure 30, C). This time, the construct was expressed with and without signal sequence both from the plasmid pBXC3H to facilitate direct comparison. A clear difference was detected in the size between the two constructs which supports the hypothesis that the SBP remains anchored in the membrane (Figure 30, B).





(C)

Organism	Molecular weight of TAXI-P w/o SS	Molecular weight of TAXI-P SS	Molecular weight of TAXI-P w/o SS GFP fusion
Desulfotomaculum carboxydivorans	33.6 kDa	36.1 kDa	-
Shimwellia blattae	36.3 kDa	38.8 kDa	-
Natrialba asiatica	34.8 kDa	38.0 kDa	61.7 kDa
Proteus mirabilis	35.4 kDa	37.8 kDa	-
Marinobacter hydrocarbonoclasticus	33.4 kDa	39.5 kDa	-
Haemophilus influenzae	33.8 kDa	33.8 kDa	-

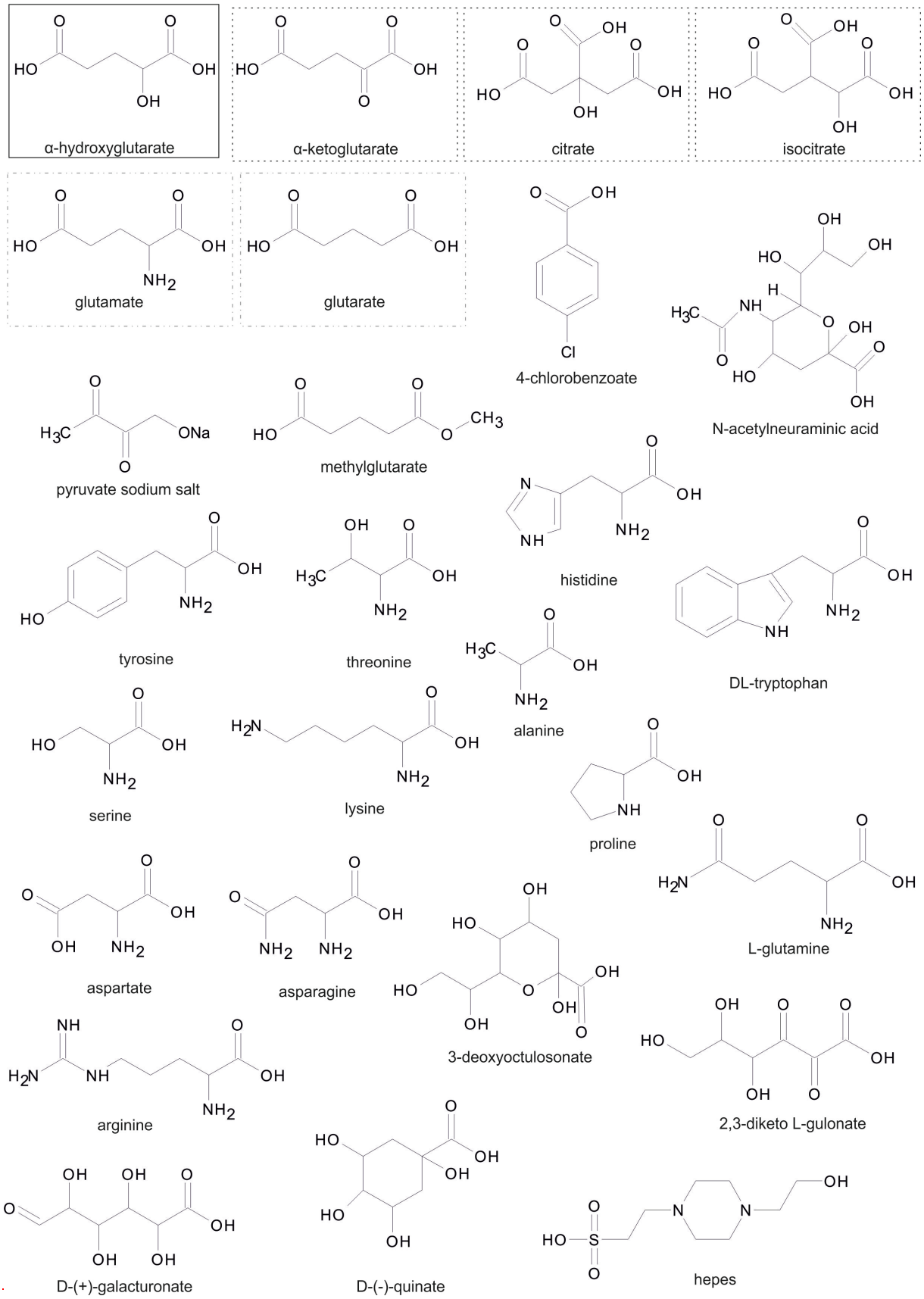
Figure 30 ***Anti*-His Western blot of selected TAXI SBPs and the TRAP SBP TRAPHi-SiaP.** (A) Initial expression screening for the SBPs that correspond to the TAXI TMD shown in figure Figure 24. *E. coli* cells were used to express the SBP at 25 °C overnight. The plasmid pBXC3H was used for the expression of the constructs with their native signal sequence and the plasmids pBXNH3 and p7XC3GH for those without signal sequence. Whole cell lysates were submitted to ultracentrifugation and the supernatants were analyzed by *anti*-His Western blots. The proteins were derived from the following organisms: *Desulfotomaculum carboxydivorans* (Dc), *Shimwellia blattae* (Sb), *Natrialba asiatica* (Na), *Proteus mirabilis* (Pm) and *Marinobacter hydrocarbonoclasticus* (Mh). SiaP from the TRAP transporter TRAPHi-SiaPQM from *Haemophilus influenzae* (Hi) served as positive control. The proteins were applied on a 12 % gel and detected in an *anti*-His Western blot. (B) The SBP from *Proteus mirabilis* TAXIPm-P (Pm) was expressed in *E. coli* cells at 25 °C overnight for both constructs without (-) and with (+) signal sequence (SS) from the plasmid pBXC3H at 25 °C overnight. The proteins were applied on a 10 % gel and detected in an *anti*-His Western blot. (C) The molecular weights calculated from the amino acid sequences of the SBPs are shown in the table for constructs without signal sequence (w/o SS) and with signal sequence (SS). The molecular weight of the TAXI-P from *Natrialba asiatica* equals the protein without signal sequence, but with GFP fusion due to expression from the p7XC3GH plasmid.

4.5 Deorphanization of TRAP transporters by DSF

While homolog screening is an efficient method to single out biochemical superior protein variants, it comes with the challenge that often very little if any functional information on the candidates is available. A robust deorphanization strategy is thus required. In this respect, studying the thermal melting of the target protein using DSF is the method of choice due to its comparably high-throughput and the fact that it does not require labeled ligands. Instead, the fluorescent dye Sypro Orange is

excited at 500 nm thereby providing a high signal to noise ratio as very few molecules interfere at this wavelength. The method itself is particularly attractive in the characterization of transporters that function with an SBP, such as TRAP transporters, as only the soluble SBP and not the membrane domain is required. So, we used DSF to identify binding of substrates to the SBPs of TRAP transporters to pre-screen for potential transport candidates. During the experiment, the temperature rise leads to the gradual unfolding of the protein and consequent exposure of hydrophobic residues. This facilitates binding of Sypro Orange which results in a sharp increase of the fluorescence (Scott, Spry, and Abell 2016). As with high temperatures the protein aggregates and the dye dissociates, the fluorescence decreases again (Vetting et al. 2015). When substrate is added, the binding to the protein is expected to alter the thermostability which leads to unfolding of the protein at lower or higher temperatures thereby shifting the fluorescence curve. Therefore, changes in the melting temperature of a certain protein can provide information about the range of substrates it binds (Niesen, Berglund, and Vedadi 2007; Kranz and Schalk-Hihi 2011). In substrate binding dependent transport systems, the binding of ligands to the SBPs may provide information about what is transported by their respective transport domains (Scott, Spry, and Abell 2016).

We created a compound library of potential ligands for TRAP-SBPs based on known substrates for TRAP systems (Figure 31) (Vetting et al. 2015; Mulligan, Fischer, and Thomas 2011) which we tested in DSF experiments for individual TAXI-SBPs. We added compounds from the analysis of the gene neighborhoods as genes that encode a certain transporter and the enzyme that metabolizes the transported molecule are often part of the same operon to assure their co-localization and -regulation (Vetting et al. 2015). The oligosaccharide/H^+ symporter lactose permease LacY as part of the LacIOZYA operon is an example of a protein located in the same operon near proteins proceeding with the conversion of a certain substrate or transport of its successor (Guan and Kaback 2006). For TAXIPm-P we additionally examined α-hydroxyglutarate as the gene encoding the hydroxyglutarate oxidase is located in the direct gene neighborhood of TAXIPm-PQM (Figure 95, section 5.1). Also, α-ketoglutarate, citric acid and isocitric acid were tested as they are metabolically closely related to α-hydroxyglutarate as well as glutamate and glutarate as they are structurally related (Figure 31). For TAXIMh-P a gene encoding an antiporter which exchanges L-malate and fumarate for succinate was found in the gene neighborhood (Figure 98, section 5.2). As these ligands were already part of the compound library from known TRAP SBP ligands, they were also tested (Figure 31).


α-hydroxyglutarate
α-ketoglutarate
citrate
isocitrate
glutamate
glutarate
4-chlorobenzoate
N-acetylneuraminic acid
pyruvate sodium salt
methylglutarate
histidine
tyrosine
threonine
DL-tryptophan
alanine
serine
lysine
proline
aspartate
asparagine
L-glutamine
3-deoxyoctulosonate
arginine
2,3-diketo L-gulonate
D-(+)-galacturonate
D-(-)-quinate
hepes

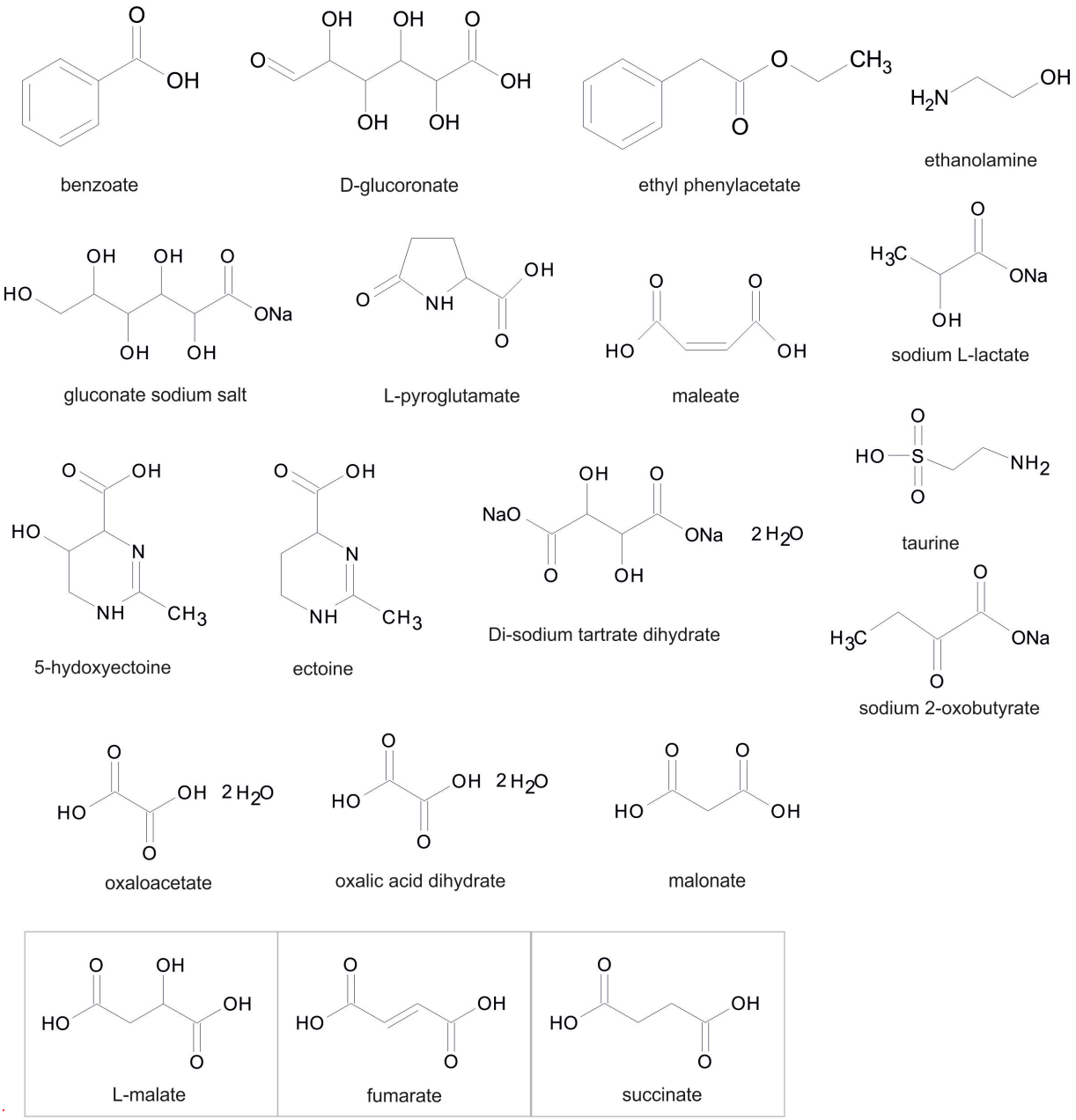


Figure 31 **Chemical structures of potential ligands for TRAP-SBPs.** The compound library of potential ligands for TRAP-SBPs was created from common TRAP substrates and randomly from C_4-and C_5-dicarboxylates and amino acids to test their interaction in DSF experiments. The boxed substrates were selected from an analysis of the genomic DNAs from *Proteus mirabilis* and *Marinobacter hydrocarbonoclasticus*. α-hydroxyglutarate, framed with a solid black line, was added to the compound library after the analysis of the gene neighborhood of TAXIPm-PQM from *Proteus mirabilis*. The metabolically closely related substrates α-ketoglutarate, citrate and isocitrate are framed with a black dashed line and the structurally related substrates glutamate and glutarate with a grey dashed line. L-malate, fumarate and succinate, framed with solid grey lines, were potential substrates for TAXIMh-PQM from *Marinobacter hydrocarbonoclasticus* after its analysis of the gene neighborhood. The chemical structures were created using ACD/ChemSketch (Advanced Chemistry Development, Inc. 2021.2.1).

4.5.1 Validation of the deorphanization strategy

To validate the envisioned deorphanization strategy, we studied the thermal melting of the SBP TRAPHi-SiaP from the TRAP transporter TRAPHi-SiaPQM from *Haemophilus influenzae.* It is known that TRAPHi-SiaP specifically binds N-acetylneuraminic acid and delivers the substrate to the transport domain TRAPHi-SiaQM (Mulligan, Fischer, and Thomas 2011; Mulligan et al. 2009). TRAPHi-SiaP was expressed in the cytoplasm of *E. coli* in TB medium and purified to high homogeneity (Figure 32).

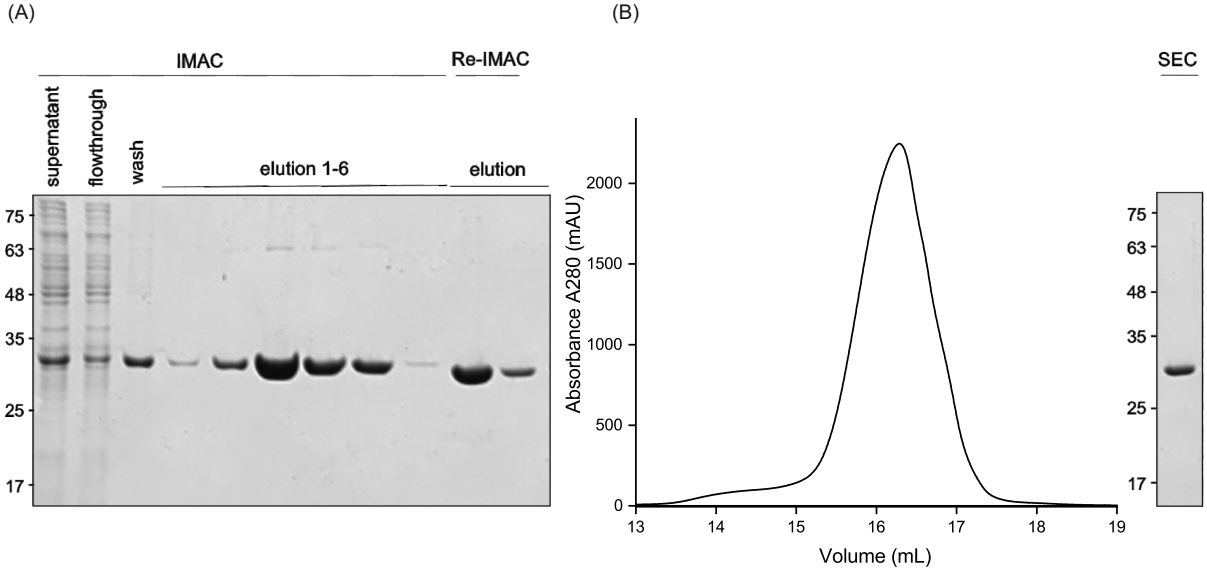


Figure 32 **Purification of the SBP TRAPHi-SiaP from *Haemophilus influenzae.*** (A) The SBP TRAPHi-SiaP was expressed in *E. coli* MC1061 and purified by IMAC followed by 3C-cleavage and Re-IMAC. (B) The size exclusion chromatography profile shows a monodisperse peak. It was carried out on a Superdex 200 30/100 GL Increase column with a running buffer containing 20 mM Hepes pH 7.5, 150 mM NaCl. The purity of TRAPHi-SiaP was visualized in a Coomassie protein stain of a 12 % SDS PAGE. The plot was created using OriginPro 2019b (OriginLab Corporation).

The melting point of unliganded TRAPHi-SiaP was 54 °C (Figure 33). In the presence of the substrate N-acetylneuraminic acid at 100 µM, the protein was stabilized by 6 °C. Comparable shifts in melting temperatures were obtained in previous DSF screens for the stabilization of certain proteins for which the binding was additionally confirmed using crystallization (Vetting et al. 2015). Other tested compounds, such as tartaric acid or taurine, did not alter the thermal stability of TRAPHi-SiaP in DSF (Figure 33).

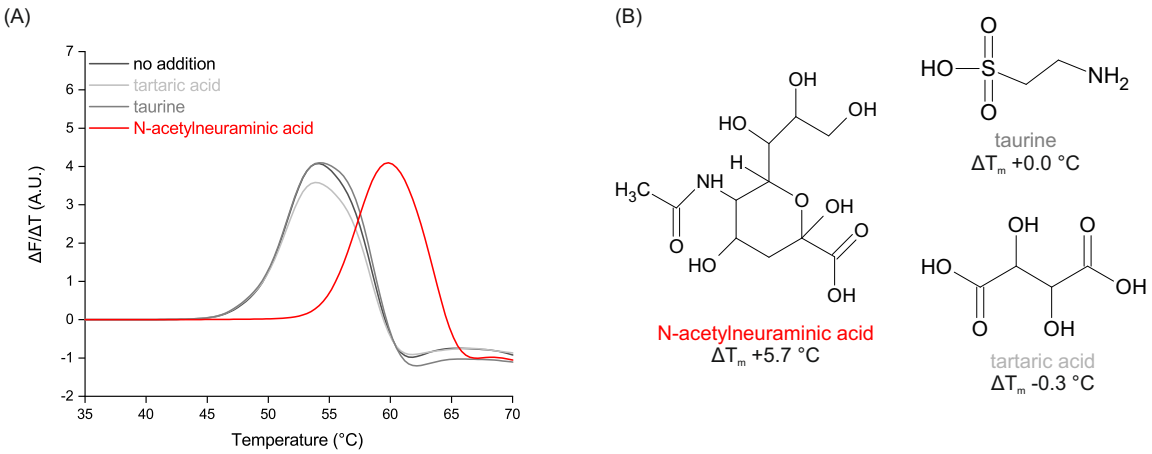


Figure 33 **Purification and DSF of the SBP TRAPHi-SiaP from *Haemophilus influenzae*.** The SBP TRAPHi-SiaP was expressed in *E. coli* MC1061 and purified by IMAC followed by Re-IMAC and SEC. (A) The size exclusion chromatography profile shows a monodisperse peak. It was carried out on a Superdex 200 30/100 GL Increase column with a running buffer containing 20 mM Hepes pH 7.5, 150 mM NaCl. (B) The purity of TRAPHi-SiaP was visualized in a Coomassie protein stain of a 12 % SDS PAGE. (C) The thermal stabilization of TRAPHi-SiaP after addition of its substrate N-acetylneuraminic acid is increased showing a 6 °C higher melting point using 10 µM protein and 40 µM substrate. Since tartaric acid and taurine do not bind TRAPHi-SiaP, they do not change the thermal stabilization. This demonstrates that DSF can be used as a method to deorphanize SBP from TRAP transporters. (D) Chemical structure of N-acetylneuraminic acid, taurine and tartaric acid. The plot was created using OriginPro 2019b (OriginLab Corporation). The chemical structures were created using ACD/ChemSketch (Advanced Chemistry Development, Inc. 2021.2.1).

The thermal shift was reproduced very well and only N-acetylneuraminic acid, the known substrate of TRAPHi-SiaPQM, led to a significant shift in the thermal stability of TRAPHi-SiaP. This highlights that DSF is extremely efficient in identifying potential substrates in TRAP systems.

4.5.2 Deorphanization of TAXIMh-P

As anticipated based on the results of the expression screening (Figure 30), the SBP TAXIMh-P from the TAXI transporter from *Marinobacter hydrocarbonoclasticus* was produced well and to high biochemical purity (Figure 34).

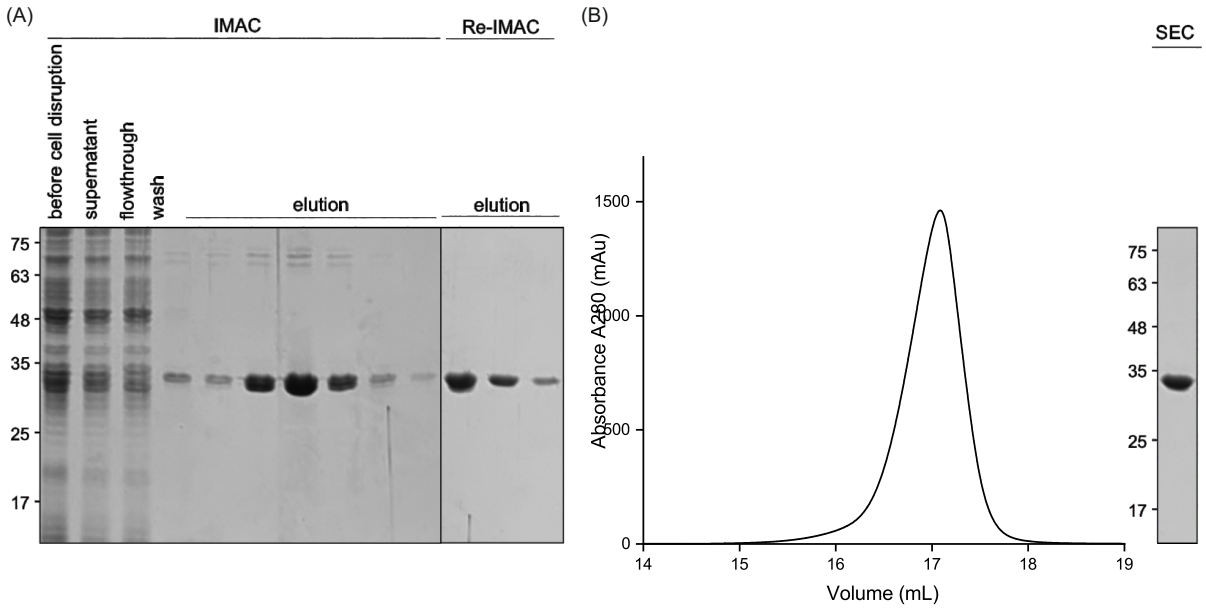


Figure 34 **Purification of the TAXI-SBP TAXIMh-P from *Marinobacter hydrocarbonoclasticus*.** The SBP TAXIMh-P was expressed in *E. coli* MC1061 and purified by IMAC followed by Re-IMAC and SEC. (A) The size exclusion chromatography profile was carried out on a Superdex 200 30/100 GL Increase column with a running buffer containing 20 mM Hepes pH 7.5, 150 mM NaCl. (B) The purity of TAXIMh-P was visualized in a Coomassie protein stain of a 12 % SDS PAGE. The plot was created using OriginPro 2019b (OriginLab Corporation).

We exposed TAXIMh-P to members of a substrate compound library created from known ligands of TRAP SBPs (Figure 31) and studied its thermal melting (Figure 35, A). Most known ligands of TRAP transporters are C_4-or C_5-dicarboxylates, but in rare cases sugars and amino acids are also recognized (Vetting et al. 2015; Mulligan, Fischer, and Thomas 2011). In addition to these we tested substrates that were considered relevant after inspecting the genomic neighborhood of the TRAP transporter. Purified, unliganded TAXIMh-P had a melting point of 47.5 °C and with a small shoulder at around 52 °C which could originate from traces of contaminations or endogenous ligands being present which are not visible in the SDS gel (Figure 34). However, the biphasic thermal melting does not influence the identification of potential substrates, so the slight shoulder in the melting curve of the apo protein is neglected. When being exposed to a range of ligands, an increase in melting temperature was observed in the presence of fumarate, L- malate and succinate (Figure 35, A, C). The addition of fumarate increased the thermal stability of the SBP TAXIMh-P by 7.7 °C, succinate led to an increase of 10.3 °C and L-malate to an increase of 12.3 °C (Figure 35, A). This binding of the three ligands to TAXIMh-P implies that TAXIMh-QM can transport all three structurally similar substrates (Figure 35, B), but it is also conceivable that one SBP serves different transporters (Antoine et al. 2003; L. T. Rosa, Springthorpe, et al. 2018).

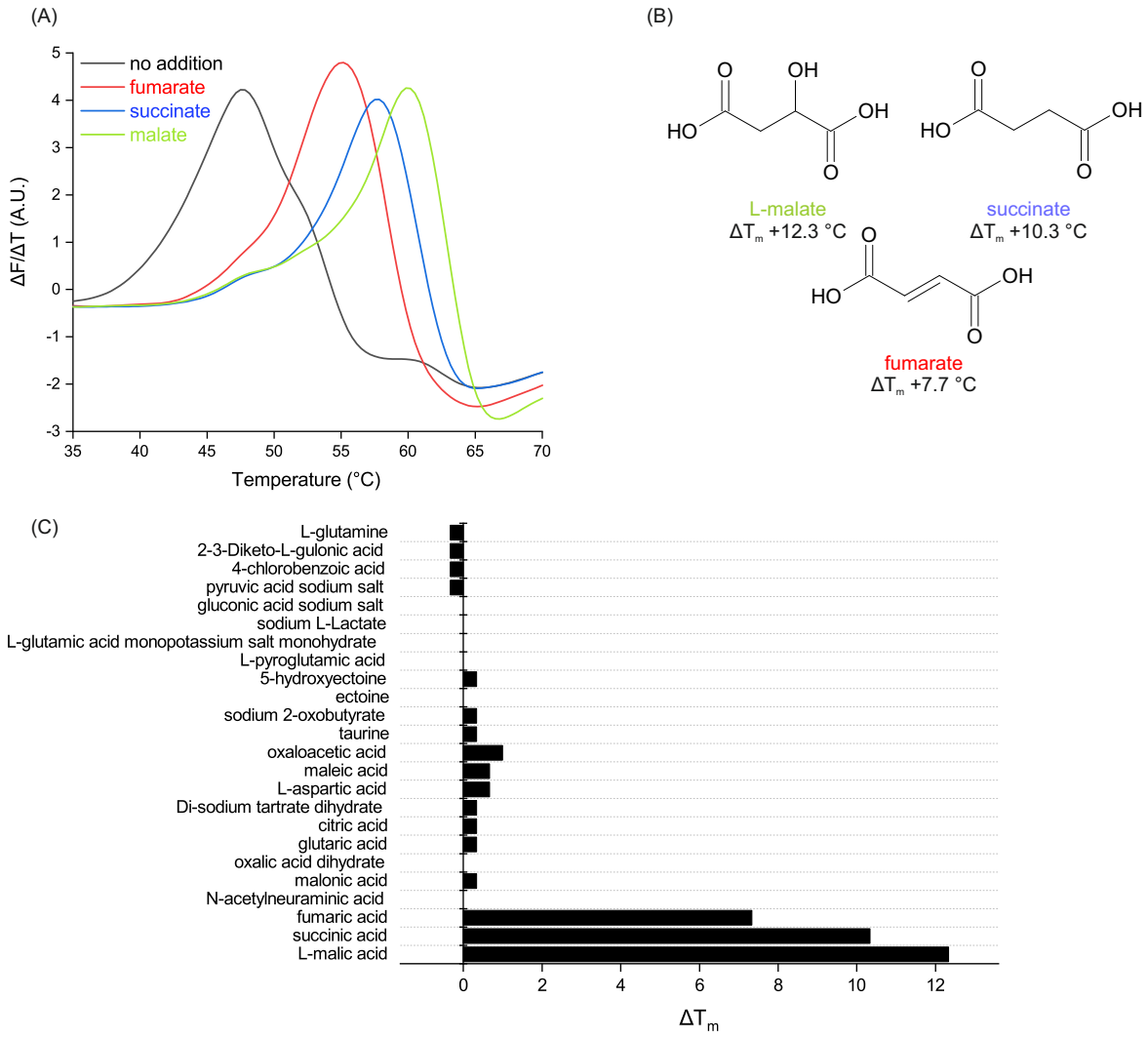


Figure 35 **DSF of the TAXI-SBP TAXIMh-P from *Marinobacter hydrocarbonoclasticus*.** (A) DSF was used to identify potential substrates of the whole transporter TAXIMh-PQM from *Marinobacter hydrocarbonoclasticus* using 10 µM protein and 40 µM potential substrate. (B) The chemical structures of L-malate, succinate and fumarate are shown. (C) The bar chart shows all tested potential substrates and the respective thermal stabilization of the protein. The plots were created using OriginPro 2019b (OriginLab Corporation). The chemical structures were created using ACD/ChemSketch (Advanced Chemistry Development, Inc. 2021.2.1).

To gain insight into potential substrate binding sites, the binding of fumarate, succinate and L-malate to the SBP TAXIMh-P was simulated using Autodockvina and visualized using Ligplot+ (Figure 36) (Laskowski and Swindells 2011). For this, smiles string codes of the respective substrates were created in chemsketch and imported into Chimera, where they were provided either single or di-anionic. Additionally, the structural model of TAXIMh-P was generated using Alphafold (Jumper et al. 2021) and imported into Chimera as well. The docking was then simulated by using the Autodockvina built-in tool in Chimera (Trott and Olson 2010). Thereby the complete SBP was specified as potential receptor into which the software docked the ligand at the matching binding site. In the simulated docking, all three ligands, fumarate, succinate, and L-malate were similarly oriented and depicted in the expected substrate binding cleft (Figure 36), where for other TRAP substrates binding was structurally shown (Mulligan, Fischer, and Thomas 2011). Of the involved residues, it is often the backbone carbonyl groups that are shown in binding distance to the substrates (Figure 36),

which implies that hydrogen bonds are more relevant in the interaction of ligand and SBP than charged side chains for example.

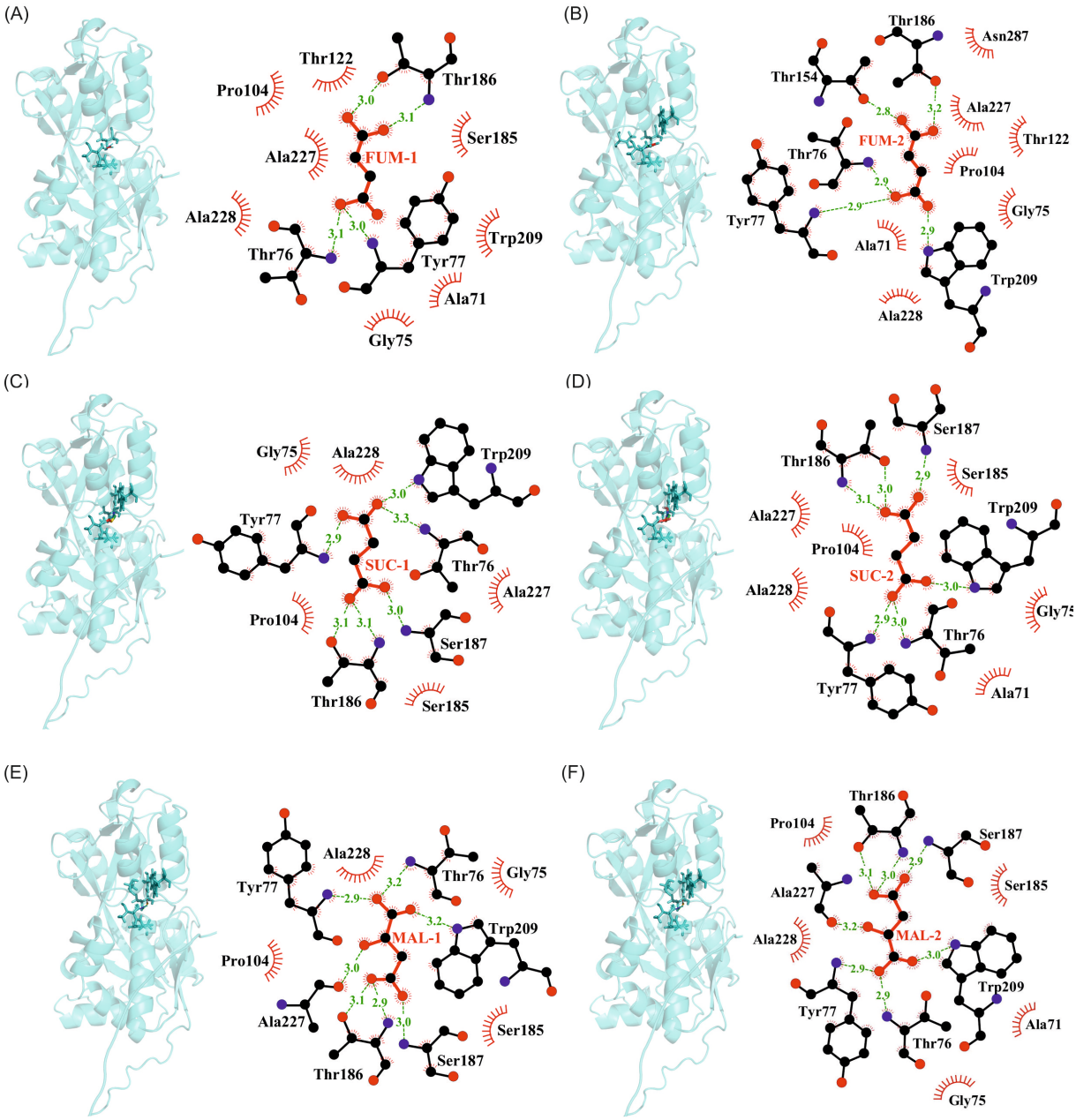


Figure 36 **Autodocking of TAXIMh-P with fumarate, succinate and L-malate.** To gain insight into potential substrate binding sites the binding of fumarate (FUM-1; FUM-2), succinate (SUC-1; SUC-2) and L-malate (MAL-1; MAL-2) to the SBP TAXIMh-P was simulated using the Chimera implemented tool Autodockvina (Trott and Olson 2010) and visualized using Ligplot^{+} (Laskowski and Swindells 2011). A, C and E show the docked single anionic ligands and B, D and F show the docked di-anionic ligands. The amino acids directly assumed to be involved in the binding are shown as molecules and depicted in green letters including their position in the SBP. Their distance to the docked ligand is indicated in Å. Neighboring hydrophobic amino acids are depicted in black letters next to red semicircles.

4.5.3 Deorphanization of TAXIPm-P

As anticipated based on the results of the expression screening (Figure 30) and successful studies with the SBP of of *Marinobacter hydrocarbonoclasticus* TAXIMh-P (section 4.5.2), the SBP TAXIPm-

P from the transporter from *Proteus mirabilis* was also produced well and to high biochemical purity (Figure 37).

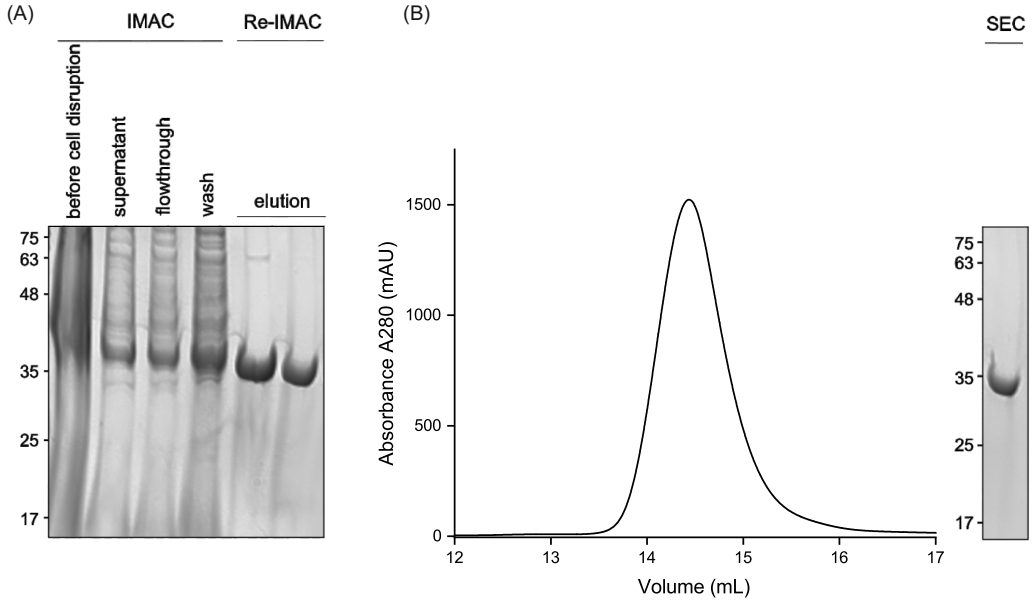


Figure 37 **Purification TAXI-SBP TAXIPm-P from *Proteus mirabilis*.** The SBP TAXIPm-P was expressed in *E. coli* MC1061 and purified by IMAC followed by Re-IMAC and SEC. (A) The size exclusion chromatography profile was carried out on a Superdex 200 30/100 GL Increase column with a running buffer containing 20 mM Hepes pH 7.5 and 150 mM NaCl. (B) The purity of TAXIPm-P is visualized in a Coomassie protein stain of a 12 % SDS PAGE. The plot was created using OriginPro 2019b (OriginLab Corporation).

We exposed TAXIPm-P to members of a compound library created from known ligands of TRAP SBPs (Figure 31) and studied its thermal melting by DSF (Figure 38). A wide range of mostly C_4- and C_5-dicarboxylates and also amino acids and sugars were tested as potential substrates. Purified, unliganded TAXIPm-P shows a slightly biphasic melting curve with the main melting point of 56 °C and a minor shoulder near 47 °C. When being exposed to a range of substrates, a significant increase in melting temperature was observed in the presence of the C_5-dicarboxylates α-hydroxyglutarate and α-ketoglutarate (Figure 38, A). The two ligands are very similar in structure and part of the same metabolic pathway with α-hydroxyglutarate being a direct precursor of α-ketoglutarate. The binding of α-hydroxyglutarate leads to a temperature shift of 5.3 °C and binding of α-ketoglutarate to a temperature shift of 3.9 °C (Figure 38, A). The binding appears to be very specific as compounds that are structurally closely related such as glutarate or glutamate, did not change the thermostability of the protein (Figure 38, B, C). The binding of α-hydroxyglutarate and α-ketoglutarate to TAXIPm-P imply that they represent genuine substrates of the transporter TAXIPm-QM as well. Ligand screening was performed together with Melanie Engelin, Goethe University Frankfurt.

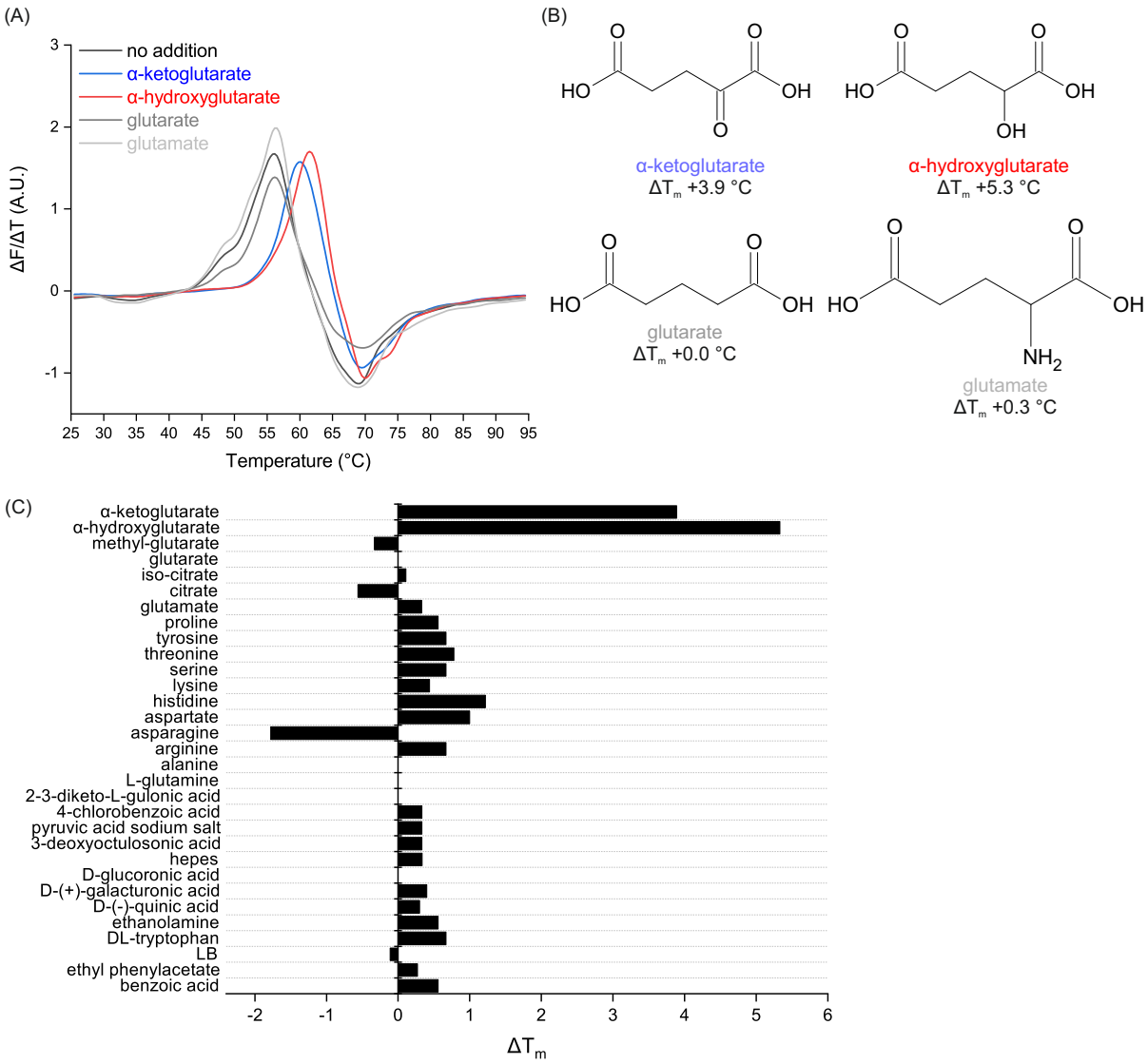


Figure 38 **DSF of the TAXI-SBP from *Proteus mirabilis*.** (A) DSF was used to identify potential substrates of the transporter TAXIPm-PQM from *Proteus mirabilis* using 10 µM protein and 40 µM potential substrate. (B) The chemical structures of α-hydroxyglutarate, α-ketoglutarate, glutamate and glutarate are shown. They bound to TAXIPm-P in DSF. (C) The bar chart shows all tested potential substrates with the corresponding thermal stabilization of the protein TAXIPm-P. Several ligands were screened by Melanie Engelin, Goethe University Frankfurt. The plots were created using OriginPro 2019b (OriginLab Corporation). The chemical structures were created using ACD/ChemSketch (Advanced Chemistry Development, Inc. 2021.2.1).

To gain insight into potential substrate binding sites, the binding of α-ketoglutarate and α-hydroxyglutarate to the SBP TAXIPm-P was simulated using Autodockvina and visualized using Ligplot+ (Figure 39) (Laskowski and Swindells 2011; Trott and Olson 2010). This was performed analogously to TAXIMh-P. Here, too, the potential ligands were docked in a single anionic and di-anionic form (Figure 39). In the simulated docking, α-ketoglutarate and α-hydroxyglutarate were similarly oriented and bound within the expected substrate binding cleft (Figure 39), analogous to other TRAP substrates that bind their corresponding SBP (Mulligan, Fischer, and Thomas 2011). As for TAXIMh-P (Figure 36), the binding of ligands to TAXIPm-P seems to mainly involve the backbone carbonyl groups (Figure 39).

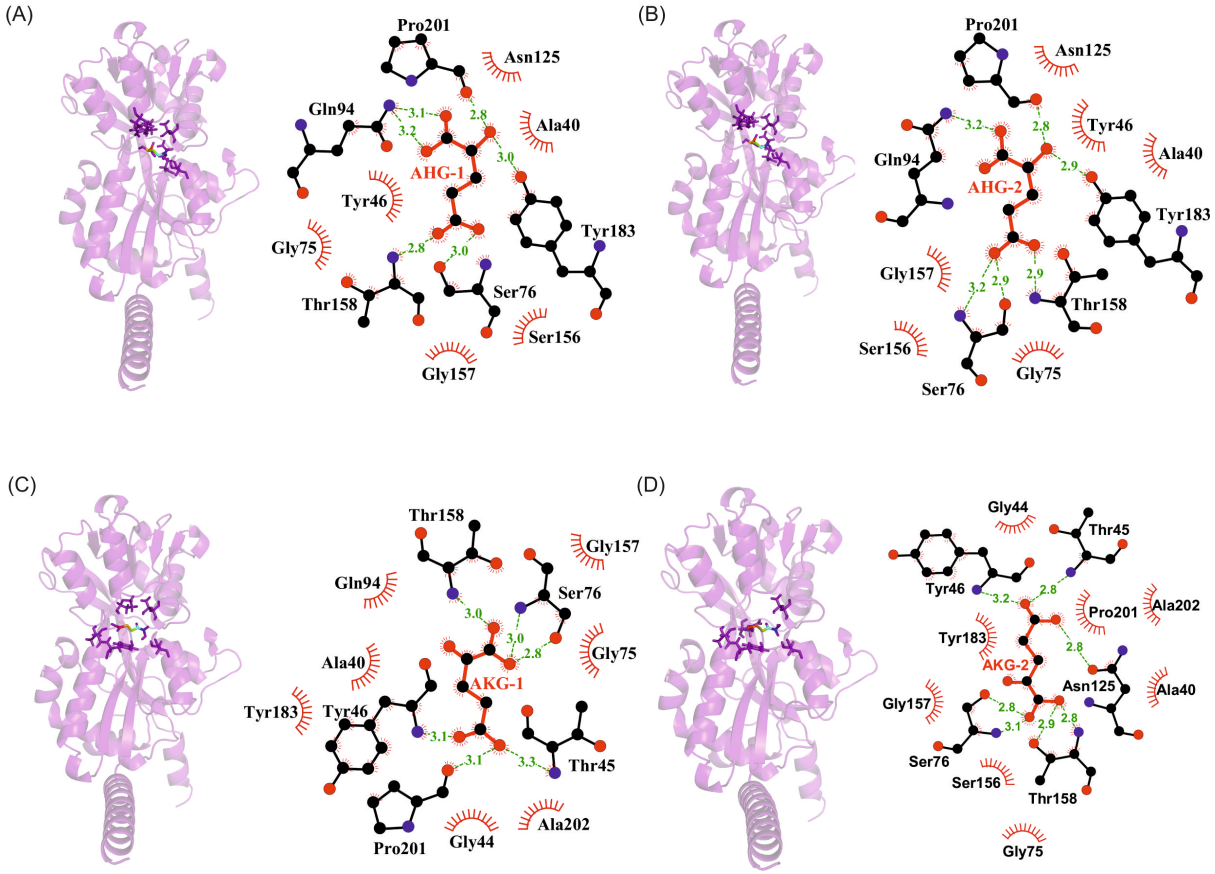


Figure 39 **Autodocking of TAXIPm-P with α-hydroxyglutarate and α-ketoglutarate.** To simulate the potential binding of α-hydroxyglutarate (AHG-1; AHG-2) and α-ketoglutarate (AKG-1; AKG-2) to TAXIPm-P and to get a closer look at the possible substrate binding and amino acids involved the in Chimera implemented tool Autodockvina was used (Trott and Olson 2010). The visualization was created using Ligplot^{+} (Laskowski and Swindells 2011). A and B show the docked single anionic ligands and C and D show the docked di-anionic ligands. The amino acids directly assumed to be involved in the binding are shown as molecules and depicted in green letters including their position in the SBP. Their distance to the docked ligand is indicated in Å. Neighboring hydrophobic amino acids are depicted in black letters next to red semicircles.

We showed for TAXIPm-P that α-ketoglutarate is a potential substrate. We then analyzed in more detail how protein concentrations or pH-values effect the binding. Conclusions about affinities can be drawn to a limited extent, as these are strongly temperature-dependent and can be assessed in DSF at the melting temperature only. The buffers used, protein concentrations, ionic strength and pH might lead to deviations. However, it can be assumed that under the same conditions, ligands that lead to a larger thermal shift also have a higher affinity for the protein. Changing the pH is interesting as the proportions of the substrate species change between single anionic and di-anionic which can give conclusions about the favored bound species. α-ketoglutarate has two pK_a-values 1.9 and 4.4 leading to the single anionic percentages of 7.4 % at pH 5.5, 2.5 % at pH 6.0 and 0.1 % at pH 7.5 (Figure 40). A change in the pH most likely additionally affects the protonation of the protein as the theoretical isoelectric point of TAXIPm-P is 6.11 (Expasy, ProtParam) which could influence the substrate binding. However, it is unlikely to affect the binding of α-hydroxyglutarate or α-ketoglutarate to TAXIPm-P, as it seems that rather the backbone carbonyl groups than charged residues of the protein seem to be involved in the binding of the substrate (Figure 39).

(A)

α-ketoglutarate
pKa1: 1.9
pKa2: 4.4

α-ketoglutarate	di-anionic	single anionic	neutral
pH 5.5	92.6 %	7.4 %	0.0 %
pH 6.0	97.5 %	2.5 %	0.0 %
pH 7.5	99.9 %	0.1 %	0.0 %

(B)

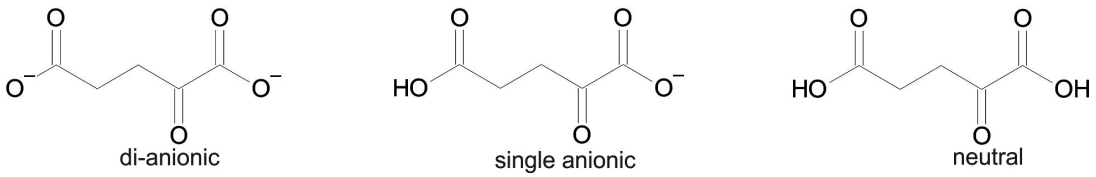


Figure 40 **α-ketoglutarate species at different pH-values** A) The C_5-dicarboxylate α-ketoglutarate has two pK_a-values: 1.9 and 4.4. The table shows the change in the species with the change of the pH. At pH 5.5 92.6 % of the di-anionic species are present and 7.4 % of die single anionic species, at pH 6.0 97.5 % of the di-anionic species and 2.5 % of die single anionic species and at pH 7.5 99.9 % of the di-anionic species and 0.1 % of die single anionic species. (B) The chemical structures of the different α-ketoglutarate species are shown: the di-anionic, the single anionic and the neutral form. The chemical structures were created using ACD/ChemSketch (Advanced Chemistry Development, Inc. 2021.2.1).

So, the SBP TAXIPm-P was purified by IMAC and Re-IMAC followed by SEC in different pH-values. The SEC profiles show that the protein is stable in all three pH conditions, so the purified proteins were used to perform DSF. Purified, unliganded TAXIPm-P showed an increase in the thermal stability of 2 ° C when lowering the applied pH-value from 7.5 to 5.5 (Figure 41).

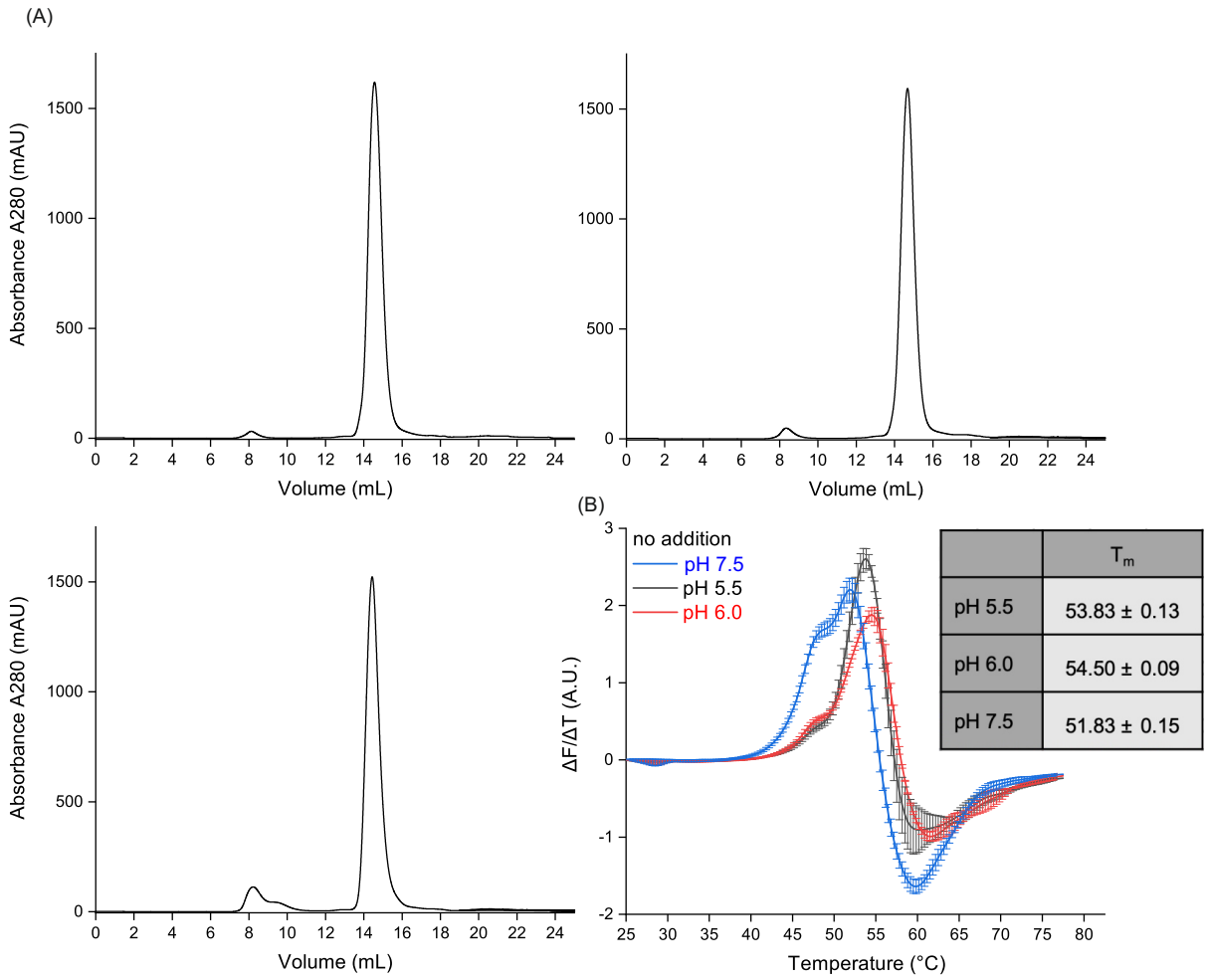
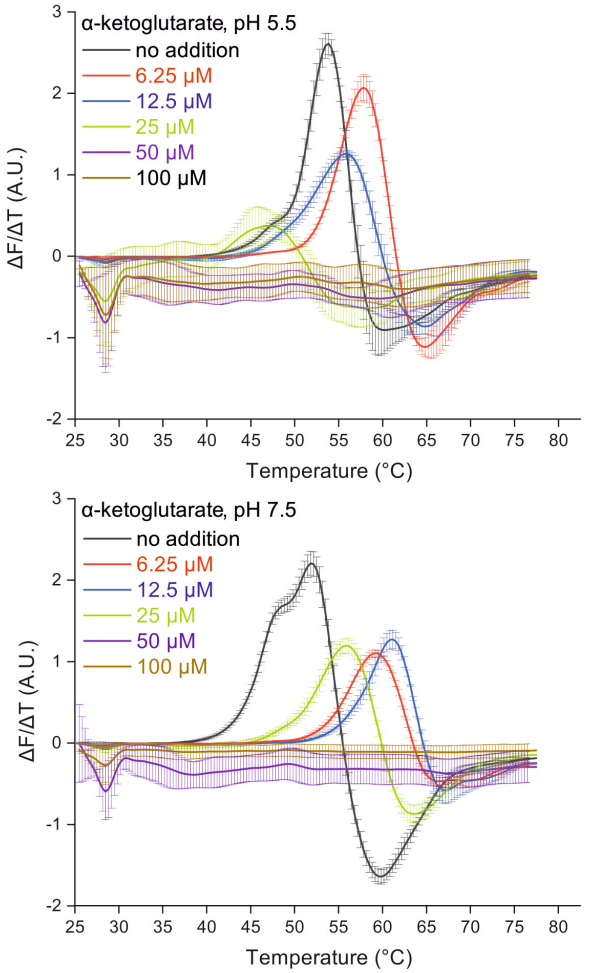


Figure 41 **Purification and DSF at different pH-values to test the effect of the protein's protonation state on its thermal stability.** The SBP TAXIPm-P was expressed in *E. coli* MC1061 and purified by IMAC followed by Re-IMAC and SEC. (A) The size exclusion chromatography profiles at pH 5.5, 6.0 and 7.5 show monodisperse peaks. The SEC runs were carried out on a Superdex 200 30/100 GL Increase column with a running buffer containing 50 mM KPi pH 5.5, 6.0 or 7.5 and 150 mM NaCl. (B) DSF was performed using the SBP TAXIPm-P at a concentration of 10 µM. No substrate was added to evaluate the effect of the pH only. The plot shows the thermal stabilization and destabilization effect in different pH-values: 5.5, 6.0 and 7.5. A change in the pH leads to different protonation states of protein. The theoretical isoelectric point of the SBP TAXIPm-P is 6.11 (calculated using Expasy ProtParam). Curves represent average values of triplicates with standard errors. The plots were created using OriginPro 2019b (OriginLab Corporation).

Not only the different proportions of single anionic and di-anionic substrate, but also the total concentration seems to impact the thermal stability of the protein. We compared the protein stability in the presence of substrate concentrations ranging from 6.25 µM to 100 µM at pH-values of 5.5, 6.0 and 7.5 (Figure 42). In all different substrate scenarios, the highest increase in thermal stability, generated by the ligand, was observed in pH 7.5 at a substrate concentration of 12.5 µM where almost only the di-anionic species of the substrate is present (Figure 40). At α-ketoglutarate concentrations of 50 µM and above no melting curve could be observed anymore, suggesting destabilization of the protein (Figure 42). The degree of thermal stabilization of TAXIPm-P thus depends on the substrate concentration used.

(A)



(B)

	ΔT_m 6.25 µM	ΔT_m 12.5 µM	ΔT_m 25 µM
pH 5.5	4.00 ± 0.16	2.00 ± 0.04	- 6.67 ± 0.19
pH 6.0	5.00 ± 0.03	1.33 ± 0.13	- 6.00 ± 0.19
pH 7.5	7.33 ± 0.05	9.33 ± 0.11	4.00 ± 0.09

Figure 42 **DSF at different pH-values to test stabilization and destabilization of the SBP TAXIPm-P by different concentrations of α-ketoglutarate.** (A) DSF was performed using the SBP TAXIPm-P at a concentration of 10 µM. The concentration of the substrate α-ketoglutarate was varied from 6.25 to 100 µM. The three plots show the thermal stabilization and destabilization effects of the substrate concentrations in different pH-values: 5.5, 6.0 and 7.5. A change in the pH leads to different protonation states of substrate and protein. The theoretical isoelectric point of the SBP TAXIPm-P is 6.11 (calculated using Expasy ProtParam). Curves represent average values of triplicates with standard errors. The bar chart visualizes that the highest protein stability within the tested scenarios is achieved at pH 7.5 having a substrate concentration of 12.5 µM when 99.9 % of the available substrate is in a di-anionic state. The plots were created using OriginPro 2019b (OriginLab Corporation).

4.6 Deorphanization of further homologs

Next to the successful deorphanization of two representatives from the TAXI family TAXIMh-P from *Marinobacter hydrocarbonoclasticus* and TAXIPm-P from *Proteus mirabilis*, two further proteins selected in the expression screenings, TAXIDc-P from *Desulfotomaculum carboxydivorans* and TAXISb-P from *Shimwellia blattae,* were tested in DSF supplying different substrates (Figure 43). The proteins melt at considerably different temperatures, TAXIDc-P from *Desulfotomaculum carboxydivorans* at approximately 78 °C and TAXISb-P from *Shimwellia blattae* at approximately 44 °C. This could originate from the different habitats of the organisms, to which the proteins must be adapted. *Desulfotomaculum carboxydivorans* is found for example in compost heaps and thermal spring water and is a moderately thermophilic Gram-positive bacterium with optimum growth at approximately 55 °C (Parshina et al. 2005; Visser et al. 2015). In contrast *Shimwellia blattae* was

first isolated from cockroaches and is described as a mesophilic Gram-negative bacterium with optimum growth at approximately 30 °C (Priest and Barker 2010). As none of the tested substrates bound to the respective SBPs, the following work focused on the two transporters TAXIMh-PQM from *Marinobacter hydrocarbonoclasticus* and TAXIPm-PQM from *Proteus mirabilis*.

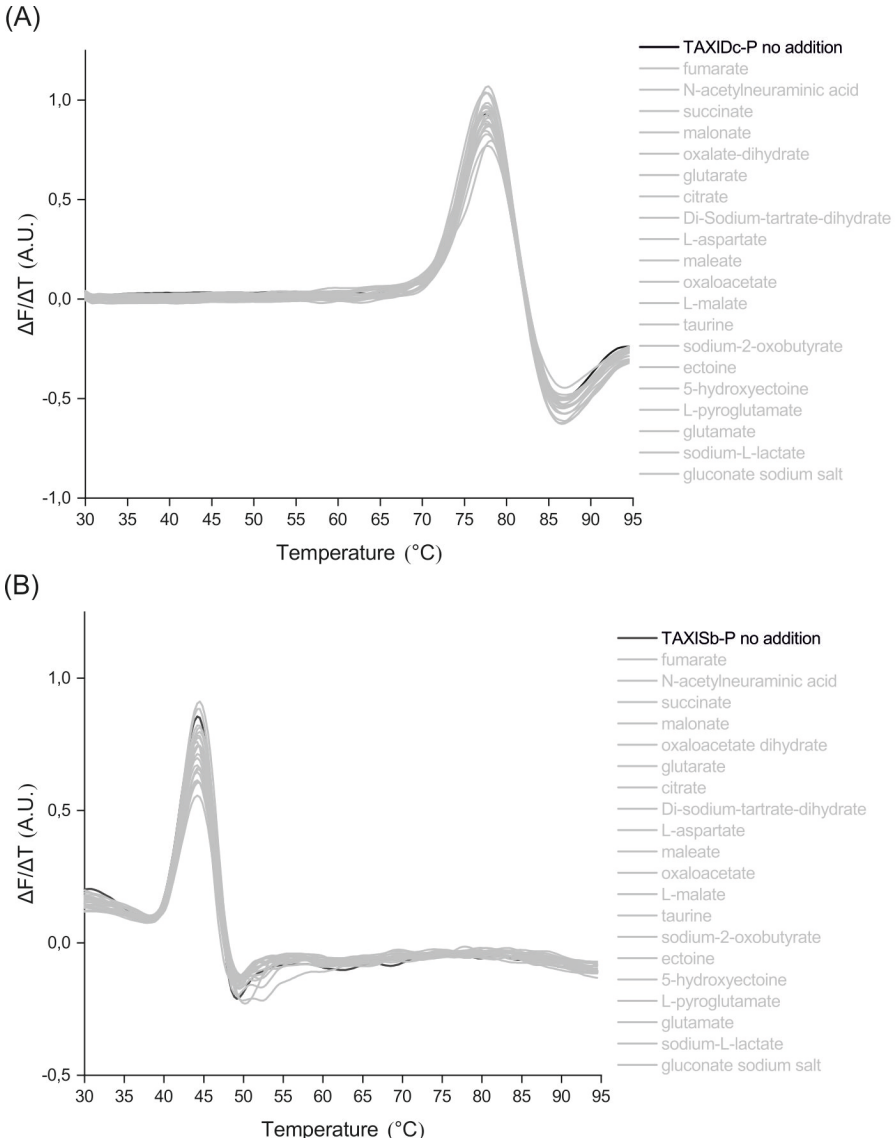


Figure 43 **DSF of the TAXI-SBP from *Desulfotomaculum carboxydivorans and Shimwellia blattae*.** (A) DSF was used to screen for potential substrates of the transporter TAXIDc-PQM from *Desulfotomaculum carboxydivorans* and (B) TAXISb-PQM from *Shimwellia blattae*. Both screenings contained 10 µM protein and 40 µM potential substrate in each sample. The plots were created using OriginPro 2019b (OriginLab Corporation).

4.7 Thermodynamic characterization using ITC

In previous experiments using DSF, ligands were found for the TAXI transport systems TAXIPm-PQM from *Proteus mirabilis* and TAXIMh-PQM from *Marinobacter hydrocarbonoclasticus*, which bound the SBPs and qualified as potential candidates for transport. To confirm the ligand binding and determine the affinities of the SBPs TAXIMh-P and TAXIPm-P for the ligands, we performed ITC. To additionally determine which species preferably binds SBPs of TAXI transport systems, we performed the ITC exemplarily for TAXIPm-P at pH 5.5, 6.0 and 7.5, thereby titrating α-ketoglutarate.

4.7.1 Thermodynamic characterization of TAXIMh-P

In DSF experiments fumarate, succinate and L-malate bound to the SBP of TAXIMh-PQM (section 4.5.2). This was recognized by their increased thermostability, which was highest for L-malate and lowest for fumarate, but significant for all three. The binding of L-malate and fumarate to TAXImh-P was confirmed by ITC, but the titration of succinate could not confirm its binding (Figure 44). This could originate from little to no heat effect with the binding of succinate to the SBP TAXIMh-P or from destabilization of the protein by succinate under the conditions present in the ITC (Boudker and Oh 2015). However, potential destabilization should not result from the supplied substrate concentration as 4.3 µM to 111.4 µM were titrated into the protein solution in the ITC and 40 µM resulted in successful stabilization in DSF (Figure 38). Also, the protonation states should be the same as both assays were performed at pH 7.5. For L-malate and fumarate however K_D-values of 10.9 µM and 31.2 µM with an approximate reaction stoichiometry of 1 were determined, respectively. The results are approximate, as no complete sigmoidal curves were obtained. Consequently, the y-axis intercept for the determination of the enthalpy was approximated. With these assumptions, the results are in line with what was observed in the previous DSF experiments where L-malate was favored over fumarate, but both can bind to TAXIMh-P. The change in enthalpy ΔH, obtained from the ITC, is negative for the binding of L-malate which signifies an exothermic reaction, while that of fumarate shows a positive ΔH and is thereby endothermic (Figure 44) Negative enthalpies often result from enhanced formation of hydrogen bonds, dipole-dipole or Coulomb interactions (Winter, Noll, and Czeslik 2011; Dorman et al. 2001). All three interactions are however presumably relevant for the binding of L-malate as well.

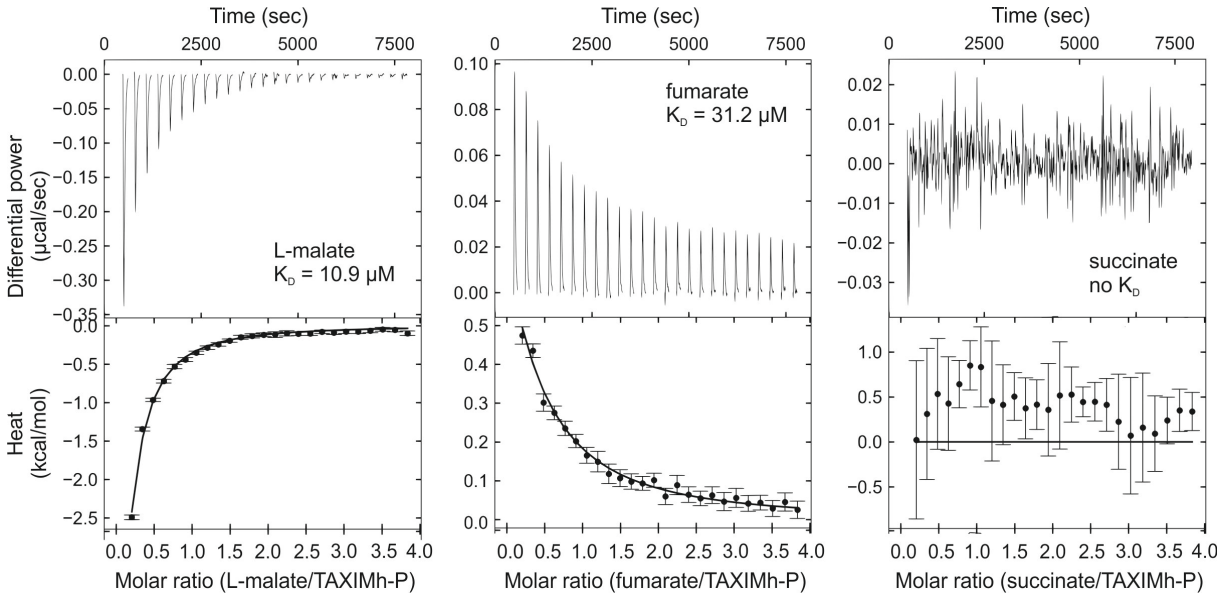


Figure 44 **Binding affinities of L-malate and fumarate to the SBP TAXIMh-P examined by ITC.** The top panels show the background-corrected raw heat exchange data, displaying the titration of L-malate, fumarate or succinate into the protein solution, respectively. The bottom panels show the binding isotherm, obtained from the integration of the heat pulses, normalized per mole of injection and fitted by a one-site binding model. Individual K_D-values are indicated. The software packages NITPIC, SEDPHAT and GUSSI were used to analyze and plot the data (Brautigam et al. 2016).

In addition to the affinities and enthalpies, further conclusions about the thermodynamic behavior of the proteins can be drawn from the ITC as most intermolecular interactions involve heat generation or heat absorption (Archer and Schulz 2020). The Gibbs energy ΔG change and also the entropy change ΔS during the binding process can be calculated from the available data giving the possibility to reveal differences in the binding behavior of the two ligands. ΔG is proportional to the logarithm of the equilibrium dissociation constant, with which the entropy can be calculated due to the isothermal measurement in the ITC (Figure 45) (Winter, Noll, and Czeslik 2011; Boudker and Oh 2015).

$\Delta G = -RT\ln(1/K_D)$ where R= 1.98 cal mol^{-1} K^{-1} and T= 293.15 K

$\Delta S = (\Delta G - \Delta H)/T$

Figure 45 **Equations to calculate the free Gibbs energy (ΔG) and the entropy (ΔS).** Affinity ($1/K_D$) and enthalpy ΔH are obtained from the ITC measurement. The cell of the ITC unit in which the binding process takes place was heated to 20 °C, so that the temperature T is known. R displays the universal gas constant (=1.98 cal/(mol*K). From these values, the Gibbs free energy (ΔG) and the entropy (ΔS) can be calculated.

For fumarate a thermodynamically favorable exergonic reaction was determined, represented by a negative Gibbs free energy (ΔG <0) (Figure 46). The -TΔS value considers the temperature dependence and is negative within the binding reaction of fumarate to TAXIMh-P which equals a positive entropy ΔS value (Figure 46). The increase in the entropy ΔS is thereby beneficial to the total reaction and thereby aids the negative ΔG. An increasing entropy can result from hydrophobic interactions and might indicate that water molecules are displaced from the binding pocket when the ligand binds (Winter, Noll, and Czeslik 2011). So, the overall binding reaction for the binding of fumarate to TAXIMh-P is exergonic, but ΔH is positive, indicating that the heat absorption at binding is compensated by the entropy gain. For the binding of L-malate to TAXIMh-P, no further thermodynamic parameters could be reliably calculated, since no clear value for the enthalpy ΔH

was obtained from the experimental data. It is nevertheless important to note that it was shown that for the binding of L-malate to TAXIMh-P the enthalpy ΔH is negative, whereas for the binding of fumarate the enthalpy ΔH is positive (Figure 46). This could suggest that for L-malate, an additional binding energy is present that might lead to the release of energy (Winter, Noll, and Czeslik 2011; Boudker and Oh 2015).

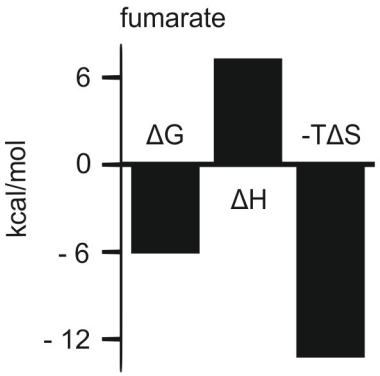


Figure 46 **Thermodynamic parameters of fumarate derived from ITC.** The bar chart displays the free Gibbs energy (ΔG), enthalpy (ΔH) and negative value of the entropy multiplied with the applied temperature (-TΔS). For comparability of the thermodynamic parameters, the negative value of entropy ΔS is shown since an increase in entropy is beneficial to the overall reaction, which is defined as negative ΔG. The bar chart was created using OriginPro 2019b (OriginLab Corporation).

4.7.2 Thermodynamic characterization of TAXIPm-P

In DSF experiments, α-ketoglutarate and α-hydroxyglutarate bound to the SBP TAXIPm-P (section 4.5.3). This was recognized by their increased thermostability, which was highest for α-hydroxyglutarate, but significant for both. Since α-ketoglutarate is commercially available in a radioactive form and can thus be used later for transport assays, the ITC focused on this ligand. The binding of α-ketoglutarate to TAXIPm-P was confirmed by ITC (Figure 47), showing a K_D of 8.4 µM. The binding proved exothermic, signified by a negative ΔH value (Figure 47). This suggests the formation of hydrogen bonds, dipole-dipole or Coulomb interactions (Winter, Noll, and Czeslik 2011; Dorman et al. 2001). From this, the Gibbs free energy and entropy can be calculated, resulting in a negative ΔG and positive ΔS (Figure 45 and Figure 47). The combination of the positive ΔS and negative ΔH implies that the binding is both entropy and enthalpy driven. An increasing entropy can result from hydrophobic interactions and might indicate that water molecules are displaced from the binding pocket when the ligand binds (Winter, Noll, and Czeslik 2011).

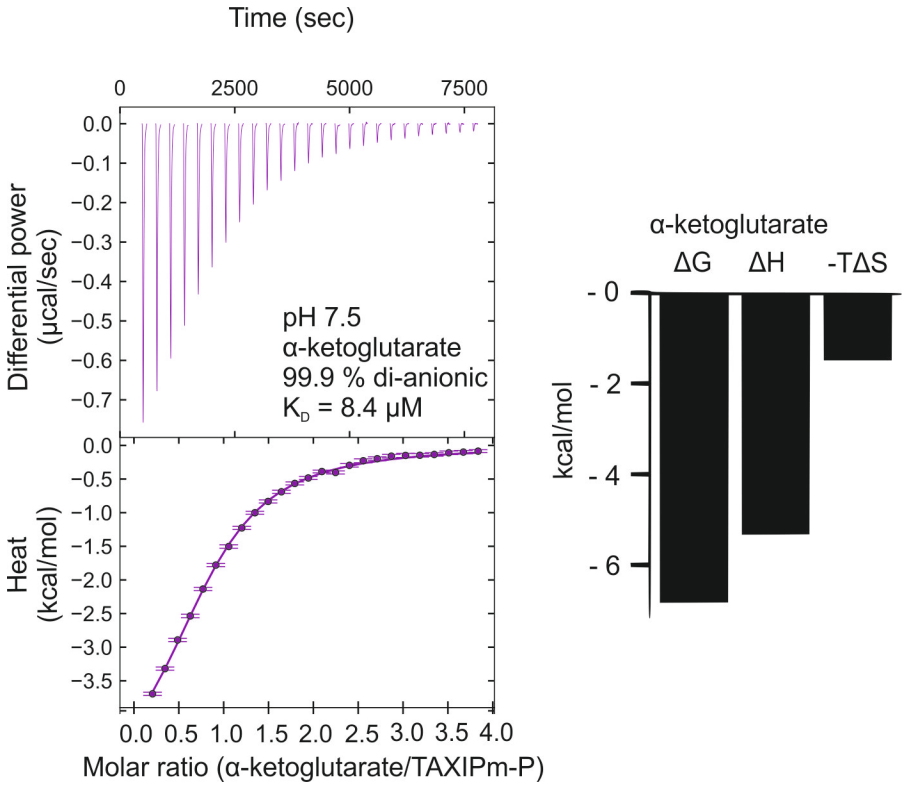
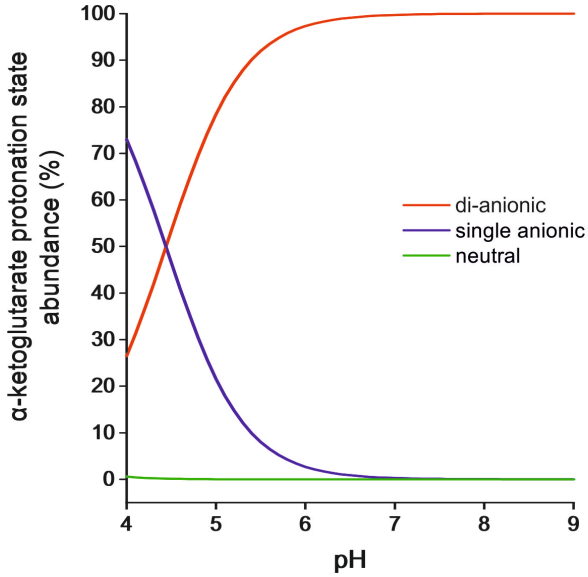


Figure 47 **Binding affinity of α-ketoglutarate to the SBP TAXIPm-P examined by ITC.** The top panel shows the background-corrected raw heat exchange data, displaying the titration of α-ketoglutarate into the protein solution. The bottom panel shows the binding isotherm, obtained from the integration of the heat pulses, normalized per mole of injection and fitted by a one-site binding model. The K_D-value is indicated. The bar chart displays the free Gibbs energy (ΔG), enthalpy (ΔH) and negative value of the entropy multiplied with the applied temperature (-TΔS). For comparability of the thermodynamic parameters, the negative value of entropy ΔS is shown since an increase in entropy is beneficial to the overall reaction, which is defined as negative ΔG. The software packages NITPIC, SEDPHAT and GUSSI were used to analyze and plot the data (left) (Brautigam et al. 2016). The bar chart was created using OriginPro 2019b (right) (OriginLab Corporation).

To further explore the binding characteristics of α-ketoglutarate to TAXIPm-P and investigate which species of the substrate might be favored, a series of ITC measurements was performed at pH 5.5, 6.0 and 7.5. In general, α-ketoglutarate occurs in three charge states, neutral, single anionic and di-anionic, although the neutral form is negligible at these pH-values given its pK_a-values of 1.9 and 4.4 (Figure 48).



α-ketoglutarate	di-anionic	single anionic	neutral
pH 5.5	92.6 %	7.4 %	0.0 %
pH 6.0	97.5 %	2.5 %	0.0 %
pH 7.5	99.9 %	0.1 %	0.0 %

Figure 48 **Protonation states of α-ketoglutarate at different pH-values.** The abundance of the three α-ketoglutarate forms, the di-anionic, the single anionic and the neutral form is illustrated in percent. The table highlights the exact values at the pH-values 5.5, 6.0 and 7.5. The plot was created using OriginPro 2019b (OriginLab Corporation).

The titrations resulted in exothermic binding of α-ketoglutarate to TAXIPm-P and K_D-values of 16.6 µM at pH 5.5, 15.4 µM at pH 6.0 and 8.4 µM at pH 7.5 when 92.6 %, 97.5 % or 99.9 % of α-ketoglutarate occurred in the di-anionic and 7.4 %, 2.5 % or 0.1 % in the single anionic form, respectively (Figure 49). Thus, the modest change in the K_D-value did not reflect the 70-fold decrease in concentration of the single anionic form. This implies that the single anionic species is not the exclusive substrate. This is further supported by the molar ratios at which binding is saturated as they barely change in the pH range tested. Therefore, we suggest that the most dominant, the di-anionic form, which is abundant in all scenarios, is either exclusively the substrate or that the TAXIPm-P does not discriminate between the single anionic and the di-anionic species.

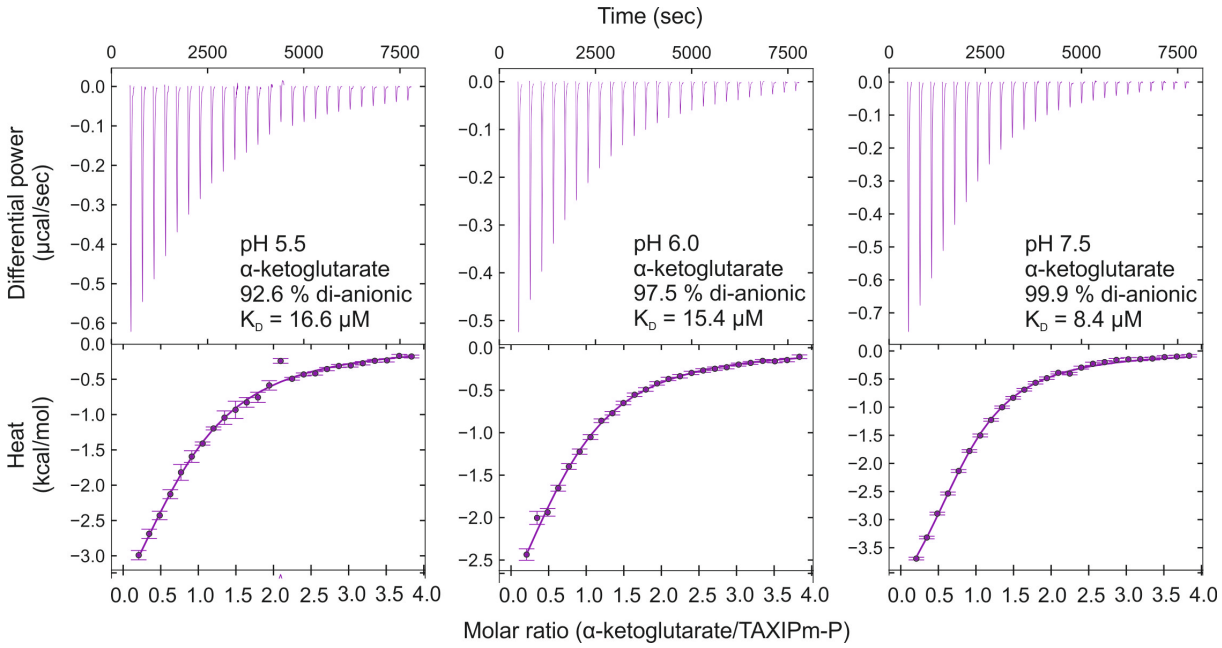


Figure 49 **pH-dependence in binding affinitiy of α-ketoglutarate to the SBP TAXIPm-P examined by ITC.** The top panel shows the background-corrected raw heat exchange data, displaying the titration of α-ketoglutarate into the protein solution. The bottom panel shows the binding isotherm, obtained from the integration of the heat pulses, normalized per mole of injection and fitted by a one-site binding model. Individual K_D-values are indicated. The software packages NITPIC, SEDPHAT and GUSSI were used to analyze and plot the data (Brautigam et al. 2016).

4.8 N-acetylneuraminic acid transport by TRAPHi-SiaPQM

It was previously shown that the TRAP transporter TRAPHi-SiaPQM from *Haemophilus influenzae* transports N-acetylneuraminic acid (Mulligan et al. 2009). This was confirmed within this work as a robust transport system was essential for our following experiments. For this purpose, we amplified the complete operon of the TRAP transport system from the plasmid pES7_SiaPQM (Wang and Kushner 1991; Mulligan et al. 2009) and cloned it into pEXC3sfGH that allows expression with the native signal sequence preceding the SBP and appends a carboxy-terminal fusion to QM composed of sfGFP and a decaHis tag (section 3.3.4). We chose to use the sfGFP instead of a commonly used GFP variant as the sfGFP can be used later for detection of various TRAP proteins, even if their C-terminus is located in the periplasm (section 1.3) (Dinh and Bernhardt 2011; Drew et al. 2003). Protein production was successful using the *lac* promoter (Figure 50, A). Expression quality was determined using the intensities of protein bands from before and after ultracentrifugation to assess how much functional protein was present during the transport experiment (Figure 50, A). Radioactive N-acetylneuraminic acid uptake was measured in *E. coli* JW3193-1 cells, which lack the *nanT* gene. NanT is a permease in *E. coli* that is responsible for transport of N-acetylneuraminic acid into the cytoplasm of the cells (Hopkins, Hawkhead, and Thomas 2013). We observed robust TRAPHi-SiaPQM dependent uptake of N-acetylneuraminic acid in glucose-energized cells (Figure 50, B).

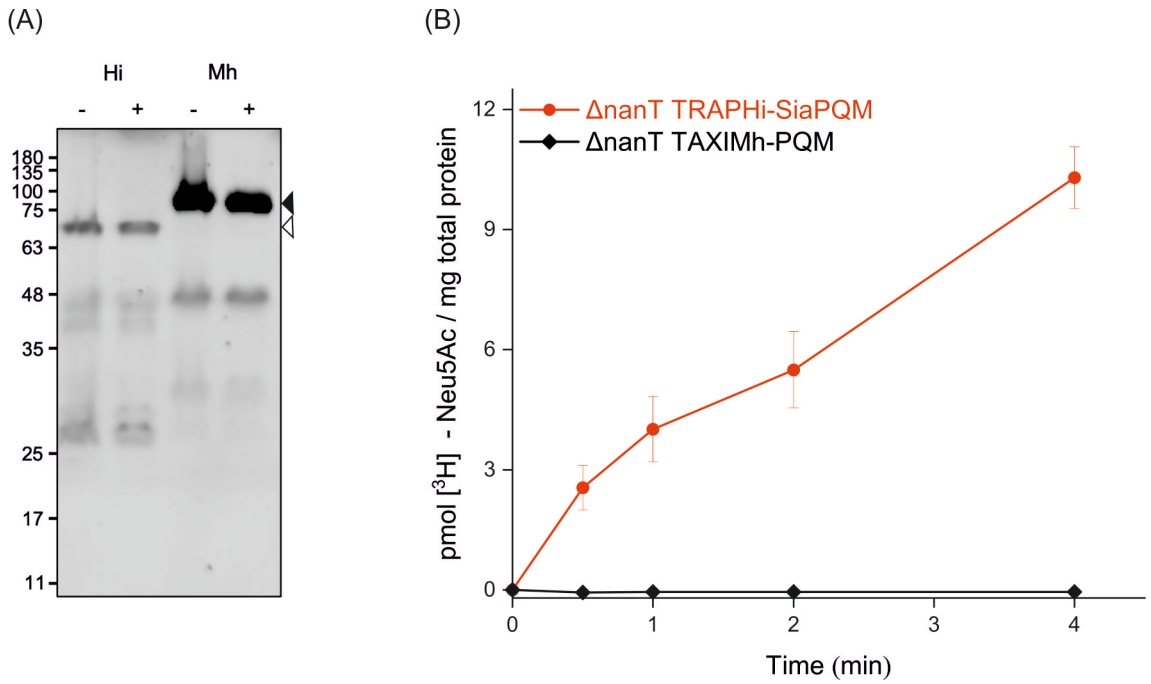


Figure 50 **Whole cell N-acetylneuraminic acid uptake of TRAPHi-SiaPQM from *Haemophilus influenzae*.** (A) Overexpression of TRAPHi-SiaPQM and the negative control TAXIMh-PQM from *Marinobacter hydrocarbonoclasticus*. Whole cell lysates of cells expressing TRAPHi-SiaPQM and TAXiMh-PQM were solubilized in 1 % DDM and submitted to ultracentrifugation to pellet aggregated protein. Target proteins were detected by *in gel* GFP fluorescence. Black and white arrows indicate the positions of TAXIMh-QM and TRAPHi-QM, respectively. (B) N-acetylneuraminic acid uptake by TRAPHi-SiaPQM in whole *E. coli* JW3193-1 cells (*ΔnanT*). Transport was assayed at 30 °C upon the addition of 11.2 nM [^{3}H] labeled N-acetylneuraminic acid using an inward-directed Na^+-gradient following a 2 min pre-energization in the presence of 0.5% (w/v) glucose. The sample TRAPHi-SiaPQM was measured in duplicate, the shown error bars equal the standard deviations. The negative control TAXIMh-PQM was measured in a single measurement. The plot was created using OriginPro 2019b (OriginLab Corporation).

4.8.1 Structure-function studies of the small membrane domain TRAPHi-SiaQ

For structure-function studies, we examined the small membrane domain Q of TRAPHi-SiaPQM from *Haemophilus influenzae* in detail. We performed the investigation before the TRAP QM structures TRAPPp-SiaQM from *Photobacterium profundum* and TRAPHi-SiaQM from *Haemophilus influenzae* were determined (Davies, Currie, and North 2023; Peter et al. 2022). Based on the two structures, it was recently suggested that TRAPHi-SiaQ forms an extension of the TRAPHi-SiaM scaffold domain and functions to provide stability for the protein complex in the membrane (Davies, Currie, and North 2023; Peter et al. 2022). However, it is still unclear whether Q has additional functional tasks and where exactly the interaction with the SBP occurs. To investigate the relevance of the small membrane domain Q for N-acetylneuraminic acid uptake, we created variants with TRAPHi-SiaQ deletion and truncations. Since TRAPHi-SiaQ and TRAPHi-SiaM occur fused in TRAPHi-SiaPQM, we first identified the part that presumably belongs to TRAPHi-SiaQ and where TRAPHi-SiaM begins. For this purpose, we performed protein sequence alignments with homologous proteins in which Q and M occur either separated or fused in the operon. For example, the first discovered TRAP transport system TRAPRc-DctQM from *Rhodobacter capsulatus* contains separated Q and M domains (Wyborn et al. 2001). Also, the protein sequences of TRAPHi-SiaQM from *Haemophilus influenzae* and TRAPVc-SiaQM from *Vibrio cholerae* were included as the N-acetylneuraminic acid transporter occurs in *Haemophilus influenzae* as a QM fusion protein but in *Vibrio cholerae* with Q and M separately expressed (Mulligan et al. 2012). For the variants, we either exchanged the first

periplasmic loop of TRAPHi-SiaQ by a Glycine Serine linker, deleted TRAPHi-SiaQ completely or partially. In - at that time - the absence of an experimental structure or alphafold model, we used a Topcon prediction (Tsirigos et al. 2015) and Phyre2 model (Kelley et al. 2015) which both suggested four transmembrane helices in TRAPHi-SiaQ and 12 transmembrane helices as well as two helical hairpins in TRAPHi-SiaM. The helices of TRAPHi-SiaM were assigned to the scaffold and transport domains one and two, respectively. This model agreed in most parts with the architectural design of the recently published structure of TRAPHi-SiaQM from *Haemophilus influenzae* which contained four transmembrane helices in TRAPHi-SiaQ, 11 transmembrane helices in TRAPHi-SiaM and two additional helical hairpins in TRAPHi-SiaM (Peter et al. 2022). In the analysis of the model used to construct the variants, one more transmembrane helix was incorrectly assigned to the connector unit, which did not affect the targeted region as in the constructed variants, the transmembrane helices assigned to the connector unit were maintained on TRAPHi-SiaM to ensure that the large transport unit remained complete (Figure 51). The recently published structure of TRAPHi-SiaQM reveals that TRAPHi-SiaQM adheres to the general design principles of elevator proteins as it consists of an inverted repeat and a scaffold and transport domain (Peter et al. 2022). Furthermore, the TRAPHi-SiaQM contains the two barrier helices that are a distinctive feature in elevator proteins (Figure 51) (Drew and Boudker 2016).

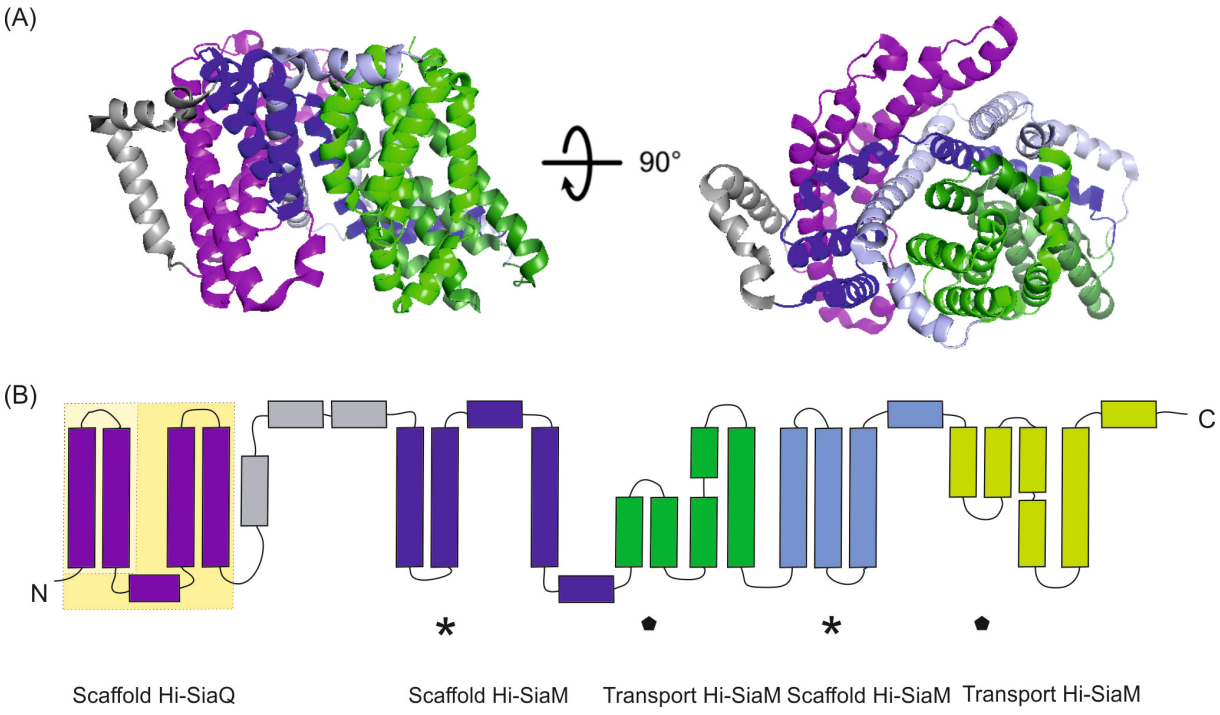


Figure 51 **Structure of TRAPHi-SiaQM from *Haemophilus influenzae*.** (A) Structure of TRAPHi-SiaQM showing the side view of TRAPHi-SiaQM approximately positioned as in the membrane and the top view. Q is depicted in purple, the connector unit between Q and M in grey, the scaffold domains one and two in dark and light blue and the transport domain one and two in dark and light green. The orientation of the protein presented is based on Topcon predictions. (B) Topology plot of TRAPHi-SiaQM. The same color scheme for the different domains is used as in A. The star marks the two barrier helices, which are a recurring feature in the scaffold domains of elevator proteins. The hairpins in the transport domains are marked with pentagons. The complete TRAP-SiaQ is shaded dark yellow and the partial TRAP-SiaQ light yellow (Peter et al. 2022).

The TRAPHi-SiaQM variants were designed such that the transport domain in TRAPHi-SiaM remained intact (Figure 52). A complete TRAPHi-SiaQ deletion variant was used to determine whether TRAPHi-SiaQ is essential for transport. As a further variant, a version was generated in which one half of Q was deleted, which results in a variant that lacks the first periplasmic loop. This should show whether the second longer periplasmic loop is alone relevant for interaction with the SBP. At that time, we assumed that TRAP transporter form oligomers similar to other elevator type transporters (Mulligan et al. 2014). It was not suggested yet that SiaQ forms a part of the scaffold (Davies, Currie, and North 2023; Peter et al. 2022). However, the investigation of SiaQ in this study should demonstrate whether half of TRAPHi-SiaQ as an extended scaffold domain provides enough stability in the membrane for TRAPHi-SiaM to enable transport. To further evaluate the importance of the periplasmic loops in TRAPHi-SiaQ in more detail, three additional variants were constructed in which either 10 or 15 amino acids of the first or second periplasmic loops were exchanged for two or three glycine serine linkers, respectively (Figure 52).

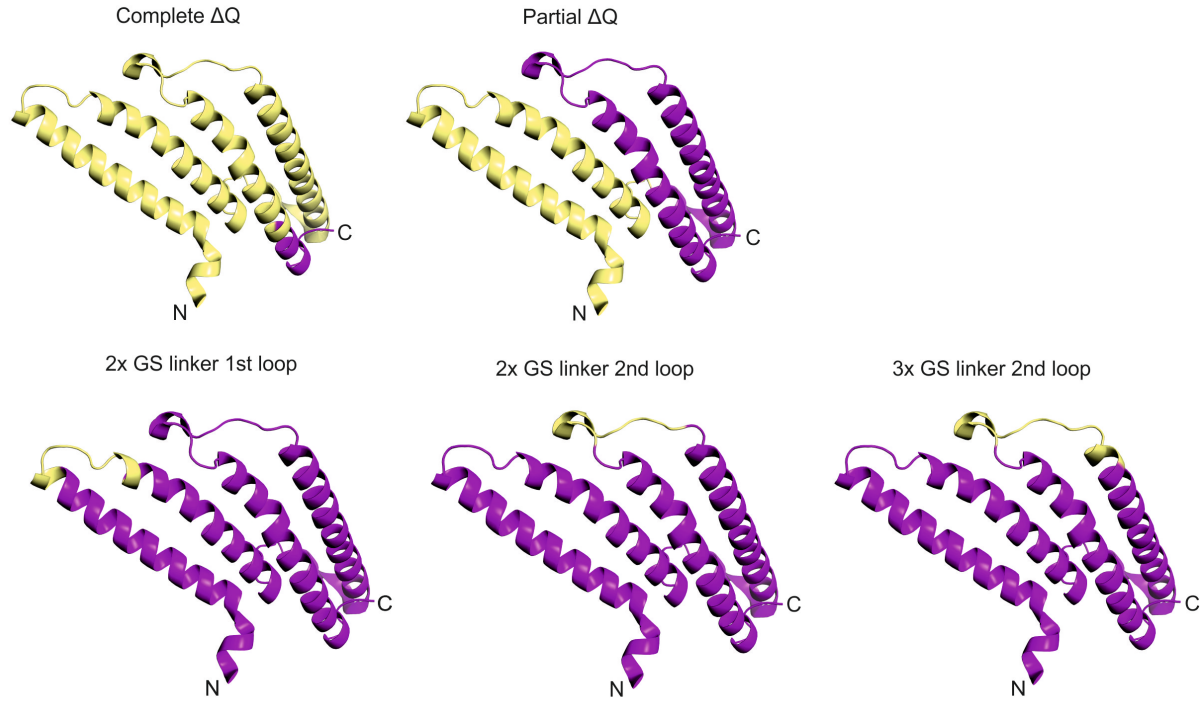


Figure 52 **Structures of TRAPHi-SiaQ variants from *Haemophilus influenzae*.** TRAPHi-SiaQ deletion variants in side view with Q in the front. In the variants TRAPHi-SiaQ was either completely (upper left) or partially deleted (upper right) or individual residues in the first or second periplasmic loop were exchanged by two or three glycine serine linkers which each consists of four glycine followed by one serine (lower left, miggle, right, respectively). The deleted parts are shown in yellow and the part remaining in purple. N- and C-terminus are indicated (Peter et al. 2022). The structures were created using Pymol 1.8.4.1 (Schrödinger, LLC).

As charged residues might be relevant for interactions to the substrate or SBP, it should be noted that the first periplasmic loop contains two negatively and two positively charged amino acids in the wild-type TRAPHi-SiaQ protein and the second periplasmic loop contains one negatively and one positively charged amino acid, all according to Topcon prediction (Figure 116, section 6) (Tsirigos et al. 2015). The constructed variants were expressed, and their transport activity was compared in

in vivo radioactive transport assays to the wild-type TRAPHi-SiaQM (Figure 53). Despite barely detectable expression levels of the TRAPHi-SiaQ deletion and truncation variants, they showed comparably high transport levels as the wildtype protein (Figure 53), B). In contrast, the negative control TRAPHi-SiaP(R147A)QM showed a high expression level but a lack of transport activity, as the SBP cannot bind the substrate with the consequence of a loss in transport (Figure 53), A). It was used as negative control as it cannot form the salt bridge that is usually built between the arginine and the carboxylic group of N-acetylneuraminic acid (Mulligan et al. 2009).

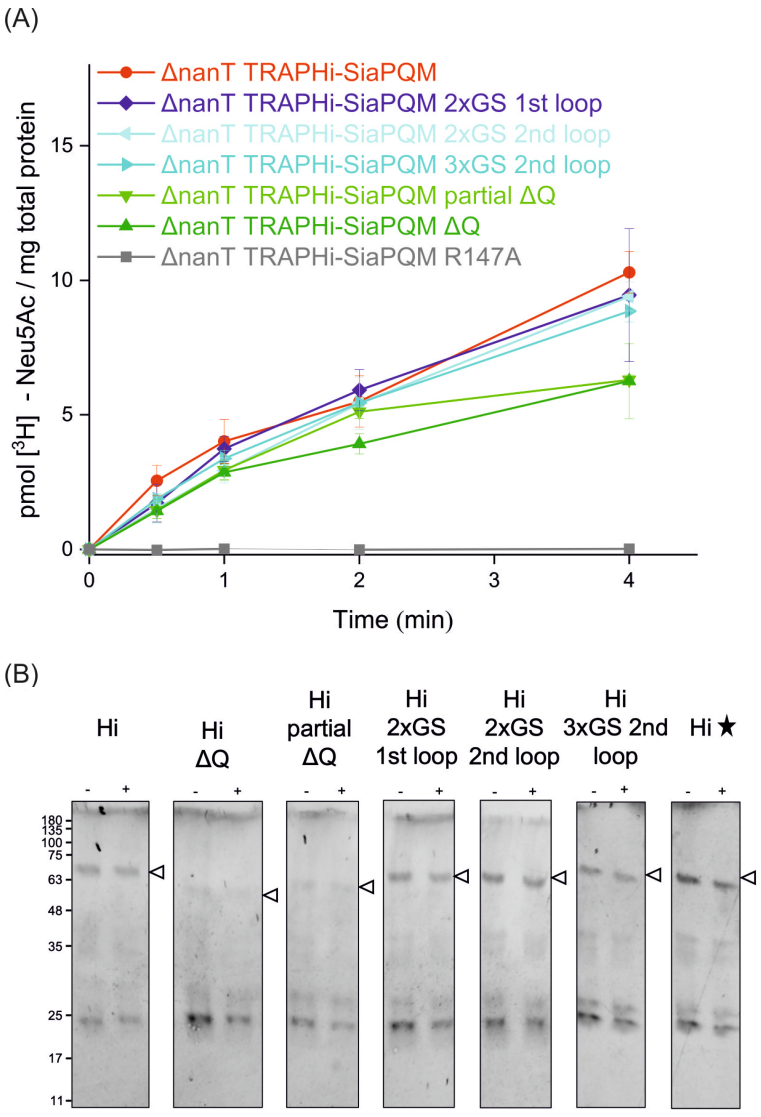


Figure 53 **N-acetylneuraminic acid uptake by TRAPHi-SiaPQM into *E. coli* whole cells and comparison of protein expression levels.** (A) N-acetylneuraminic acid uptake by TRAPHi-SiaPQM and TRAPHi-SiaPQM variants into *E. coli* whole cells JW3193-1 (*ΔnanT*), containing a SiaQ deletion, partial deletion or amino acid exchanges in the periplasmic loops. The transport was recorded in whole cells at 30 °C using 11.2 nM [^{3}H] labeled N-acetylneuraminic acid. All constructs were expressed with a lac promoter from the pEXC3sfGH plasmid which attaches a sfGFP-fusion protein. TRAPHi-SiaPQM with a mutation in the arginine 147 residue served as negative control. All samples were measured in duplicates, the error bars equal the standard deviations. The plot was created using OriginPro 2019b (OriginLab Corporation). (B) Whole cell lysates of samples from radioactive transport assay were solubilized in 1 % (w/v) DDM and ultracentrifuged to separate aggregated protein and assess how much functional protein was present during the transport experiment. All protein samples were visualized by *in gel* fluorescence. The TRAPHi-QM sample marked with an asterisk resembles the negative control in which the SBP carries a R147A mutation. White arrows indicate the positions of TRAPHi-QM wildtype and TRAPHi-QM variants.

Since we observed negligible difference in the transport performance between TRAPHi-PQM wildtype and TRAPHi-PQM variants, independent of which parts of TRAPHi-SiaQ were deleted and how much, whereas a mutation in the SBP completely abolished transport, we constructed a variant in which the SBP TRAPHi-SiaP was complete, but the membrane domains TRAPHi-SiaQ and TRAPHi-SiaM were deleted. This construct showed N-acetylneuraminic acid uptake (Figure 54, A) suggesting that the SBP SiaP alone suffices for transport, possibly involving an *E. coli* transporter that is not known to transport N-acetylneuraminic acid. The TRAPHi-SiaQM and TRAPHi-SiaP constructs used in this transport assay were well expressed in a high protein quality (Figure 54, B). In the samples where TRAPHi-SiaQM as well as TRAPHi-SiaP was present, we confirmed the quality of TRAPHi-SiaQM only, but assumed the same quality of TRAPHi-SiaP as both are in the same operon (Figure 54, B).

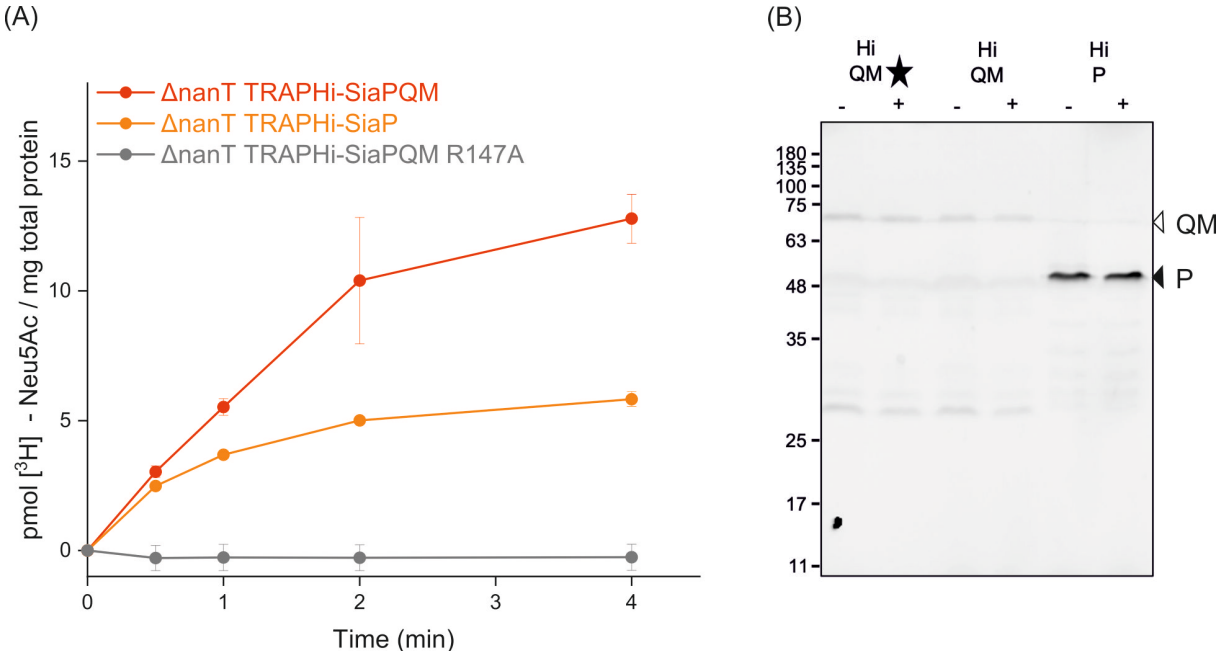


Figure 54 **N-acetylneuraminic acid transport by TRAPHi-SiaPQM and TRAPHi-SiaP into *E. coli* whole cells (*ΔnanT*).** (A) N-acetylneuraminic acid uptake was measured by TRAPHi-SiaPQM and TRAPHi-SiaP constructs. The transport was recorded in whole cells JW3193-1 (*ΔnanT*), at 30 °C using 11.2 nM [^{3}H] labeled N-acetylneuraminic acid. All constructs were expressed with a lac promoter from the pEXC3sfGH plasmid which attaches a sfGFP-fusion protein. TRAPHi-SiaPQM with a mutation in the arginine 147 residue was used as negative control. All samples were measured in triplicates, the shown error bars equal the standard deviations. The plot was created using OriginPro 2019b (OriginLab Corporation). (B) Whole cell lysates of the samples used in the radioactive transport assay were solubilized in 1 % DDM and submitted to ultracentrifugation to pellet aggregated protein. The proteins contain sfGFP fusion proteins as they were expressed in the pEXC3sfGH plasmid. All protein samples were visualized by *in gel* fluorescence. Expression levels were compared from before and after ultracentrifugation to assess how much functional protein was present during the transport experiment. White and black arrows indicate the positions of TRAPHi-QM and TRAPHi-P, respectively.

4.8.2 Investigation of TRAPHiSiaP interaction with *E. coli* TRAP YiaMNO

As we observed *in vivo* N-acetylneuraminic acid uptake in the absence of the membrane domain TRAPHi-SiaQM as long as the SBP SiaP was intact, we investigated the possible interaction of SiaP with other membrane domains than TRAPHi-SiaQM. It might possibly involve a SBP dependent *E. coli* transporter that has not been annotated as N-acetylneuraminic acid transporter. Within native *E. coli* cells ABC transporters are the largest protein group with 50 importers which most often depend

on periplasmic SBPs to deliver the substrate to the membrane (Moussatova et al. 2008). Next to the high number of ABC importers, only one TRAP transporter is annotated in *E. coli*, namely TRAPEc-YiaMNO (Benson et al. 2013; Plantinga et al. 2005). YiaMNO by itself cannot transport N-acetylneuraminic acid since it was present in our previous negative controls, which showed as expected no transport (Figure 50, Figure 53 and Figure 54). But it may be that only the SBP TRAPEc-YiaO cannot bind N-acetylneuraminic acid, but YiaMN can transport it. Therefore, we hypothesize that another SBP, which in our case would be TRAPHi-SiaP, could deliver N-acetylneuraminic acid to the transport domain TRAPEc-YiaMN, which subsequently transports the substrate. So, we investigated whether the SBP TRAPHi-SiaP can bind to the membrane domain of TRAPEc-YiaMNO and thereby enable N-acetylneuraminic acid transport. Next to 2,3-DKG, L-Xylulose was proposed as substrate for TRAPEc-YiaMNO which suggests that the transport system might have several substrates (Plantinga et al. 2005; Thomas et al. 2006). In a dedicated growth assay, the addition of L-Xylulose affected the OD_{600} of *E. coli* in the presence compared to the absence of TRAPEc-YiaMNO suggesting that it might be transported by TRAPEc-YiaMNO (Plantinga et al. 2005). Tryptophan fluorescence and mass spectrometry revealed 2,3-DKG as potential substrate for TRAPEc-YiaMNO leading to both a change in fluorescence emission and a mass change in mass spectrometry (Thomas et al. 2006). So TRAPEc-YiaMNO was assigned to these two substrates whereas N-acetylneuraminic acid was not tested. N-acetylneuraminic acid is larger than 2,3-DKG and L-Xylulose which implies that N-acetylneuraminic acid may not fit in the cavity of the SBP TRAPEc-YiaO and thus cannot be transported (Figure 55).

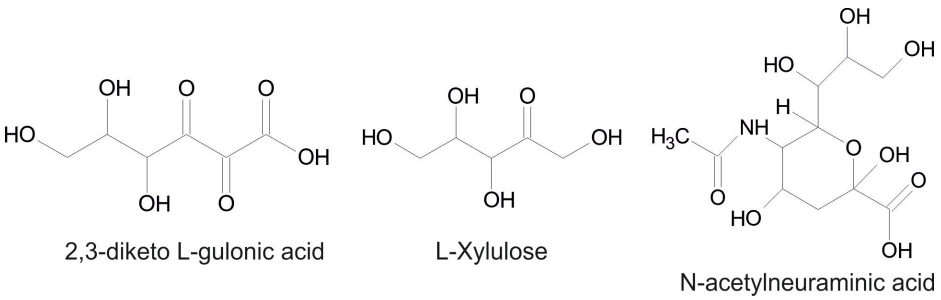


Figure 55 **Chemical structures of potentially transported substrates by TRAPEc-YiaMN from *E. coli*.** Chemical structures of 2,3-diketo L-gulonic acid, L-Xylulose and N-acetylneuraminic acid are shown. The chemical structures were created using ACD/ChemSketch (Advanced Chemistry Development, Inc. 2021.2.1).

Until today, it is not known whether the membrane domain of TRAP transport systems contributes to the substrate specificity. The results of our transport assays suggest that the SBP determines the substrate specificity as the transport is abolished as soon as the SBP cannot bind the substrate (Figure 54). However, N-acetylneuraminic acid might still be transported by the membrane domain TRAPEc-YiaMN using the SBP TRAPHi-SiaP, which can bind N-acetylneuraminic acid. As TRAPEc-YiaMNO is the only annotated TRAP transporter in *E. coli*, it is the most promising candidate (Benson et al. 2013; Plantinga et al. 2005). To investigate whether the SBP TRAPHi-SiaP can not only bind to the membrane domain TRAPHi-SiaQM but also to TRAPEc-YiaMN, we performed Clustal Ω alignments (Sievers et al. 2011). It is to date undetermined which amino acid residues are essential

for the transport of N-acetylneuraminic acid by TRAPHi-SiaQM. The comparison of the two small membrane domains TRAPEc-YiaM and TRAPHi-SiaQ exhibited an amino acid identity of 25.5 % and an amino acid similarity of 40.6 % (Figure 56). Several amino acid residues are conserved in the periplasmic loops in Q (Figure 56) where they are predicted to interact with TRAPHi-SiaP (Peter et al. 2022). Both first periplasmic loops contain identical charged amino acids at both ends of the predicted first periplasmic loop with arginine and glutamate (Figure 56), implying a conservation and relevance for the binding of TRAPHi-SiaP. The comparison of the two large membrane domains TRAPEc-YiaN and TRAPHi-SiaM resulted in a 51 % similarity with only a few conserved residues between TRAPHi-SiaM and TRAPEc-YiaN (Figure 115, section 6).

```
YiaM                1 MKKI--LEAIL--AINLAVLSCIVFINIILRYGFQTSILSVDELSRYLFV     46
                      ||.|  ||..|  |:.:|:.. |:...|:.|..|.:.::..:||::.|||
SiaQ                1 MKYINKLEEWLGGALFIAIFG-ILIAQILSRQVFHSPLIWSEELAKLLFV     49

YiaM               47 WLTFIGAIVAFMDNAHVQVTFLVEKLSPAWQRRVALVTHSLILFICGALA     96
                      ::..:|..||.....||.:.||. .|.|...|:.......|::|||..|.
SiaQ               50 YVGMLGISVAVRKQEHVFIDFLT-NLMPEKIRKFTNTFVQLLVFICIFLF     98

YiaM               97 WGATLKTIQDWSDYSPILGLPIGL----MYAACLPTSLVIAFFELRHLYQ    142
                      ....::|....|.....||   |:    ::|| ||...::..|.....
SiaQ               99 IHFGIRTFNGASFPIDALG---GISEKWIFAA-LPVVAILMMFRFIQA--    142

YiaM              143 LITRSNSLTSPPQGA    157

SiaQ              143 ---------------    142
```

Figure 56 **Alignment of the first periplasmic loops of the small membrane domains TRAPEc-YiaM and TRAPHi-SiaQ.** The two membrane domains TRAPEc-YiaM and TRAPHi-Q were aligned using Clustal Ω (Sievers et al. 2011). The amino acid residues that are highlighted in red resemble the positions of the two periplasmic loops of TRAPHi-SiaQ, predicted by Topcons (Tsirigos et al. 2015). Identical amino acids are labeled with solid lines, similar amino acids with double dots and dissimilar amino acids with single dots (Sievers et al. 2011).

As the SBP determines the specificity for the substrate (Mulligan et al. 2009) it is of particular interest that the SBP TRAPEc-YiaO also possesses the amino acid arginine that is in TRAPHi-SiaP at position 147 and was confirmed to be essential for the binding of N-acetylneuraminic acid to SiaP (Figure 54). This implies a general relevance of this amino acid in TRAP SBPs. Since one of the proposed substrates for TRAPEc-YiaO is 2,3 DKG (Thomas et al. 2006), the arginine might form a salt bridge to a carboxyl group of N-acetylneuraminic acid as in TRAPHi-SiaP. Further amino acids that were shown to be relevant in the binding of N-acetylneuraminic acid in TRAPHi-SiaP are Asn10, Glu67, Arg127 and Asn187 (Johnston et al. 2008) among which none can be found in TRAPEc-YiaO (Figure 57).

```
YiaO              1 -MKLRSVTYALFIAGLAAFSTSS--LAAQ-SLRFGYETSQTDSQHIAAKK     46
                    |||..    ||:|...:...||  |||. .|:||.....:.:::.||:.
SiaP              1 MMKLTK----LFLATAISLGVSSAVLAADYDLKFGMNAGTSSNEYKAAEM     46

YiaO             47 FNDLLQERTKGELKLKLFPDSTLGNAQAMISGVRGGTIDMEMSGSNNFAG     96
                    |...::|:::|::::.|:|.|.||:.:||:..::.|::|...:.|..|..
SiaP             47 FAKEVKEKSQGKIEISLYPSSQLGDDRAMLKQLKDGSLDFTFAESARFQL     96

YiaO             97 LSPVMNLLDVPFLFRDTAHAHKTL-DGKVGDDL-KASLEGKGLKVLAYWE    144
                    ..|...:..:|::..:...|.|.| |.:.|.|| |...:..|:.:|:...
SiaP             97 FYPEAAVFALPYVISNYNVAQKALFDTEFGKDLIKKMDKDLGVTLLSQAY    146

YiaO            145 NGWRDVTNSRAPVKTPADLKGLKIRTNNSPMNIAAFKVFGANPIPMPFAE    194
                    ||.|..|::|| :.:.||:||||:|..|:..|:|..|..||:|.||.|:|
SiaP            147 NGTRQTTSNRA-INSIADMKGLKLRVPNAATNLAYAKYVGASPTPMAFSE    195

YiaO            195 VYTGLETRTIDAQEHPINVVWSAKFFEVQKFLSLTHHAYSPLLVVINKAK    244
                    ||..|:|..:|.||:|:..|.:.||:|||||::|:|..:..|.:::...
SiaP            196 VYLALQTNAVDGQENPLAAVQAQKFYEVQKFLAMTNHILNDQLYLVSNET    245

YiaO            245 FDGLSPEFQQALVSSAQEAGNYQRKLVAEDQQKIIDGMKEAGVEVITDLD    294
                    :..|..:.|:.:..:|:.|..|..||..:.::.::...::.||: ||..|
SiaP            246 YKELPEDLQKVVKDAAENAAKYHTKLFVDGEKDLVTFFEKQGVK-ITHPD    294

YiaO            295 RKAFSDALGNQVRDMFVKDVPQ-GADLLKAVDEVQ-    328
                    ...|.::::....: |||...| |...||.::.:.
SiaP            295 LVPFKESMKPYYAE-FVKQTGQKGESALKQIEAINP    329
```

Figure 57 **Alignment of the SBP TRAPEc-YiaO and TRAPHi-SiaP**. The two SBP TRAPEc-YiaO and TRAPHi-SiaP were aligned using Clustal Ω (Sievers et al. 2011). The amino acids highlighted in yellow, Asn10, Glu67, Arg127, Arg147 and Asn187, are responsible for the binding of N-acetylneuraminic acid in TRAPHi-SiaP (Johnston et al. 2008). For the TRAPHi-SiaP the residues are moved three amino acids downwards when compared to their naming due to gaps. Identical amino acids are labeled with solid lines, similar amino acids with double dots and dissimilar amino acids with single dots (Sievers et al. 2011).

To further investigate whether the membrane domain of the TRAP transporter TRAPEc-YiaMNO in *E. coli* allows transport of N-acetylneuraminic acid when TRAPHi-SiaP delivers the substrate, we purified the transporter with the purpose of reconstitution to liposomes for *in vitro* tests. As YiaMN is not fused, it was uncertain whether TRAPEc-YiaMN was purified as a complex or in single proteins during the IMAC as the His tag is on M only (Figure 58). The theoretical molecular weight of YiaM corresponds to 17 kDa and of YiaN to 45 kDa (Gasteiger et al. 2005).

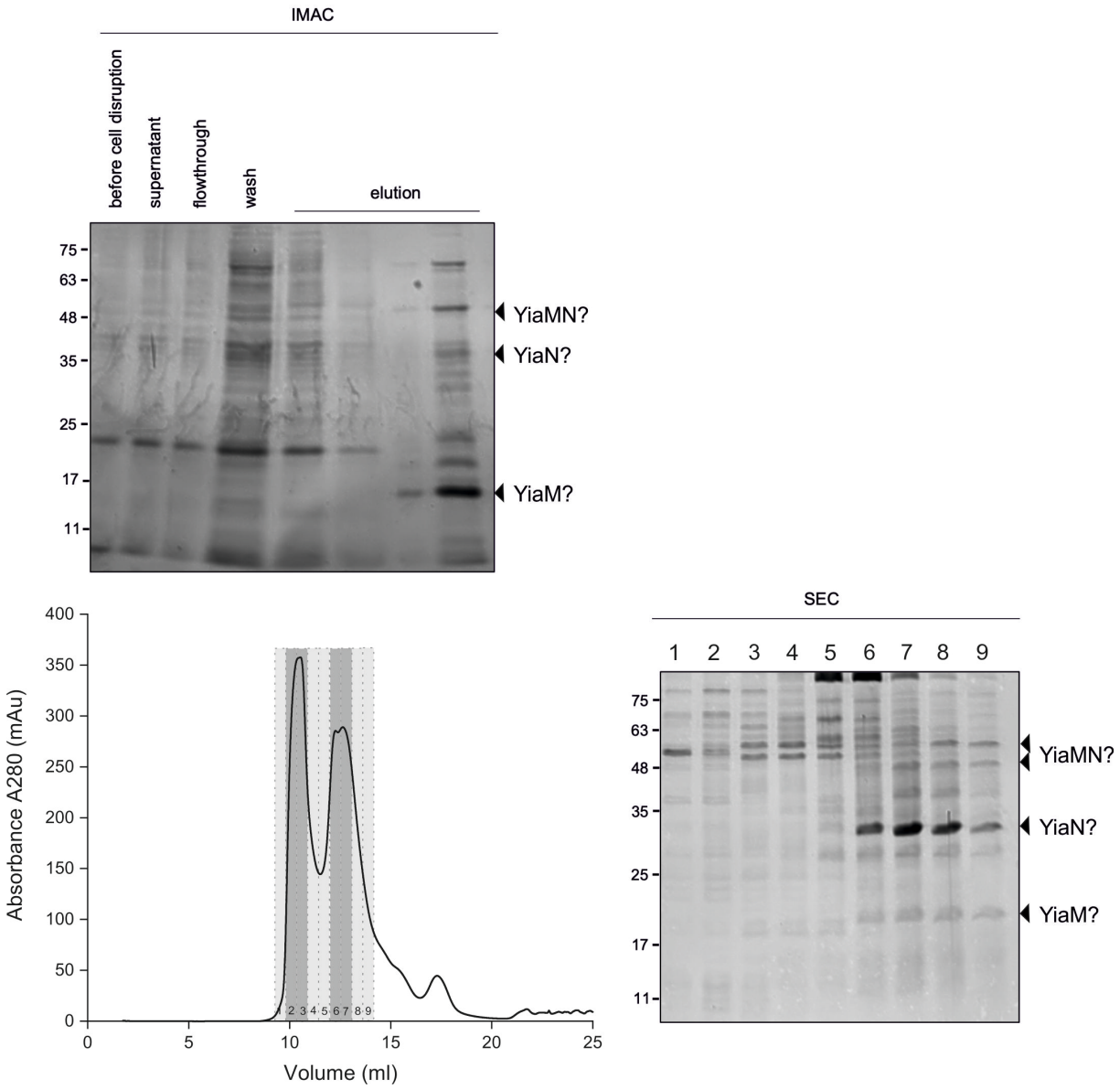


Figure 58 **Purification of TRAPEc-YiaMN.** DDM-solubilized TRAPEc-YiaMN was purified via IMAC, followed by SEC on a Superdex 200 30/100 GL Increase column. All buffers contained 0.05 % DDM. The protein samples from the purification steps were visualized using a Coomassie stain. The peak fractions from the SEC chromatogram were named accordingly in the corresponding gel. The plot was created using OriginPro 2019b (OriginLab Corporation).

Since it was unclear whether the transport domain was purified as a complex of TRAPEc-YiaM and TRAPEc-YiaN (Figure 58), the *in vitro* approach was not continued. Instead, an *E. coli* knockout strain with *nanT* and *yiaMNO* gene deletion was constructed to show whether TRAPHi-SiaP functions *in vivo* with TRAPEc-YiaMN. For this purpose, the *E. coli* strain JW5651-1 was used which lacks the endogenous *E. coli* TRAP transporter YiaMNO (Baba et al. 2006). We exchanged the nanT gene with a kanamycin cassette which allows for selection of *E. coli* colonies containing the previously created gene deletion (Datsenko and Wanner 2000). The following N-acetylneuraminic acid uptake still showed uptake when the SBP TRAPHi-SiaP was present in the absence of TRAPHi-SiaQM, TRAPEc-YiaMN and nanT (Figure 59). As before, the negative control with a mutation in the SBP TRAPHi-SiaP does not transport the substrate (Figure 59). This indicates that SiaP facilitates

the transport of N-acetylneuraminic acid with another SBP dependent transporter than TRAPEc-YiaMNO.

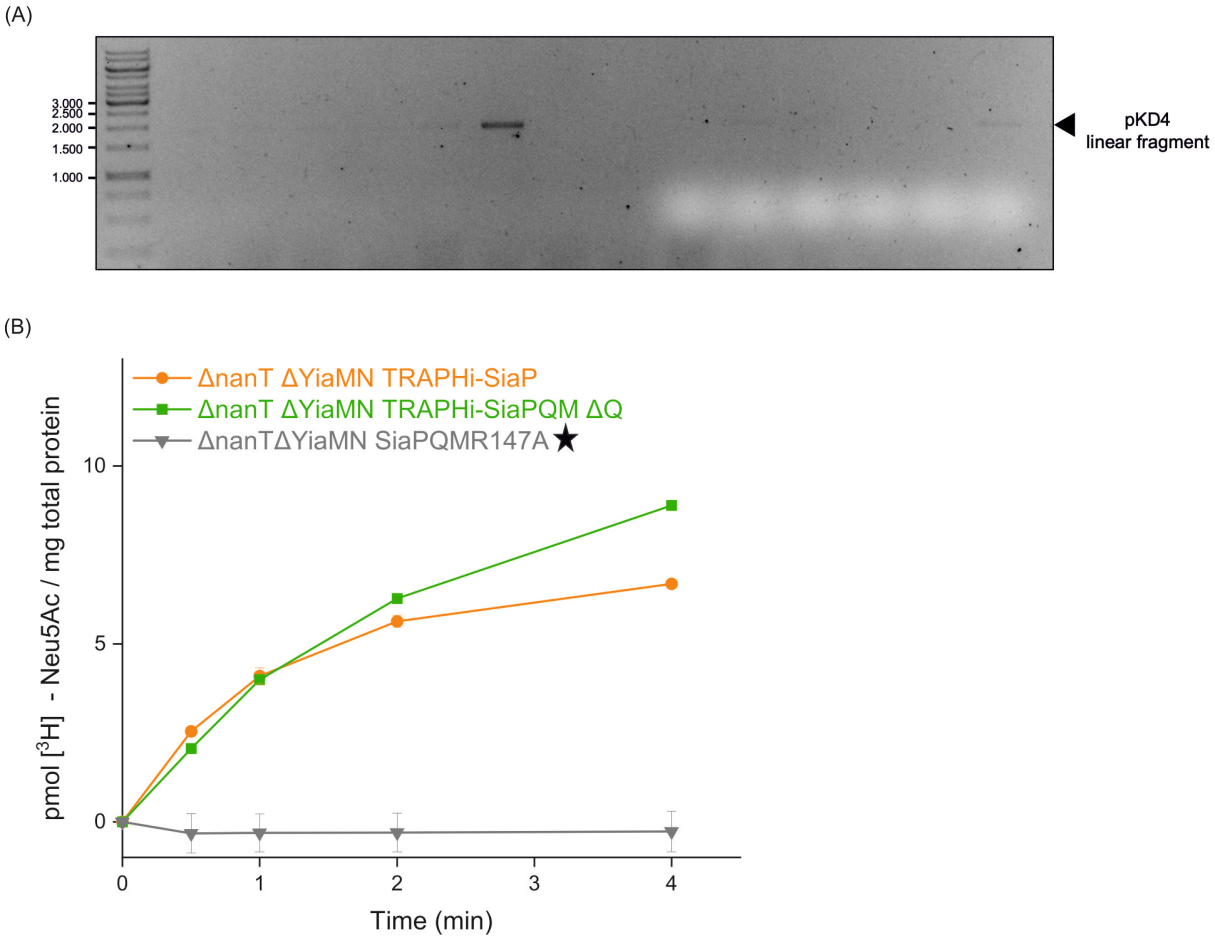


Figure 59 **N-acetylneuraminic acid transport into whole cells of a self-constructed YiaMNO/nanT knock out strain.** (A) PCR amplification with a following agarose gel electrophoresis and sequencing to show that YiaMNO was exchanged by a kanamycin resistance cassette in constructs 5, 6 and 10. The length of the gene encoding YiaMNO equals 2753 bp, whereas the PCR product equals 1852 bp consisting of a part of YiaM, the pKD4 insert including the FRT and kanamycin cassette and a part of YiaO. (B) N-acetylneuraminic acid uptake was measured by TRAPHi-SiaPQM with a deleted SiaQ and TRAPHi-SiaP constructs. The transport was recorded in whole cells at 30 °C using 11.2 nM [^{3}H] labeled N-acetylneuraminic acid. A self-constructed *E. coli* strain was used which lacks the endogenous *E. coli* TRAP transporter YiaMNO and the endogenous N-acetylneuraminic acid transporter nanT. All constructs were expressed with a lac promoter from the pEXC3sfGH plasmid which attaches a sfGFP-fusion protein. TRAPHi-SiaPQM with a mutation in the arginine 147 residue was used as negative control. TRAPHi-SiaP and the negative control were measured in triplicates, the shown error bars equal the standard deviations. TRAPHi-SiaPQM with deleted Q displays a single measurement. The plot was created using OriginPro 2019b (OriginLab Corporation).

4.9 Fumarate transport by TAXIMh-PQM

4.9.1 Fumarate uptake by TAXIMh-PQM into whole cells

Analysis of the SBP of the TAXI transporter TAXIMh-PQM from *Marinobacter hydrocarbonoclasticus* using DSF and ITC indicated that malate, succinate and fumarate are bound with affinity in the µM range (section 4.5.2 and 4.7.1). To verify whether these compounds are recognized and transported by the QM membrane domain as well, we established a transport assay. For this purpose, we amplified the complete TAXI operon from the genomic DNA of *Marinobacter hydrocarbonoclasticus* and expressed it with the native signal sequence preceding the SBP and a carboxy-terminal fusion

to QM composed of sfGFP and a decaHis tag (Figure 60, A). Protein production was successful both using the *araBAD* and *lac* promoters. Expression levels were compared using the intensities of protein bands from before and after ultracentrifugation to assess how much functional protein was present during the transport experiment (Figure 60, A). Radioactive fumarate uptake was measured in *E. coli* IMW424 cells (Figure 60, B), which lack the *dcuB* and *dcuA* genes (Kim 2006). *DcuB* codes for a fumarate/succinate antiporter and the *dcuA* gene was suggested to serve as a backup for DcuB by coding for an additional C_4-dicarboxylate transporter, thereby also enabling fumarate transport (Ingo G. Janausch, Garcia-Moreno, and Unden 2002; I. G. Janausch et al. 2001). We observed robust TAXIMh-PQM dependent uptake of fumarate in glucose-energized cells glucose using an inward-directed Na^+-gradient, indicating that fumarate is a substrate for the membrane-embedded transport domain as well (Figure 60, B).

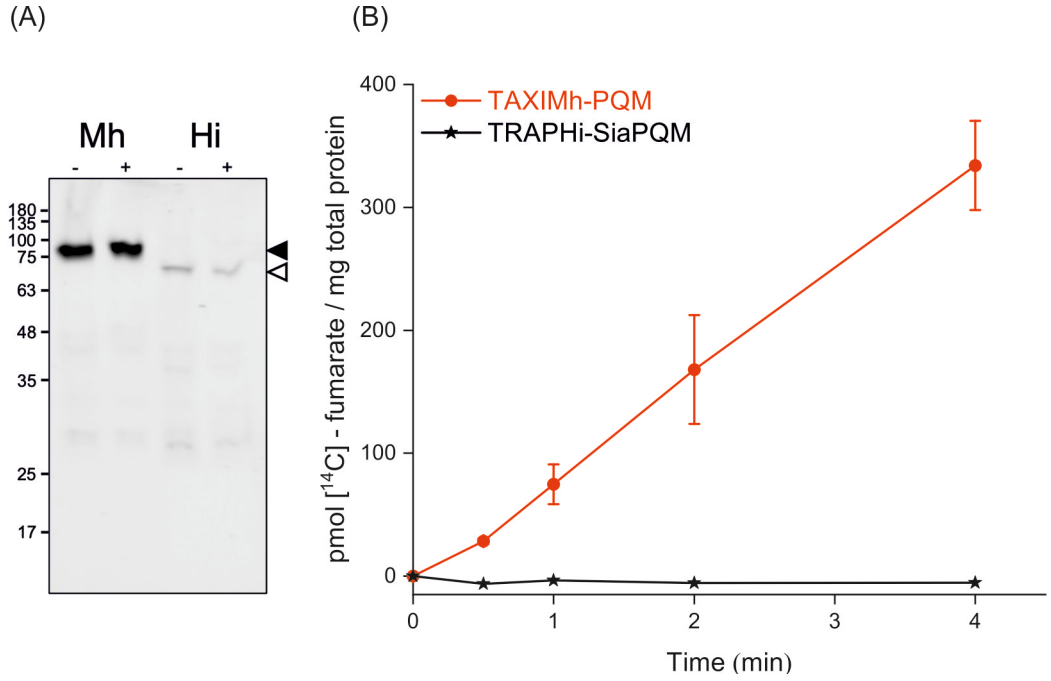


Figure 60 **Whole cell fumarate uptake by TAXIMh-PQM.** (A) Expression of TAXIMh-PQM in pBXC3sfGH and the negative control TRAPHi-SiaPQM from *Haemophilus influenzae* in pEXC3sfGH were induced by 0.01 % (w/v) arabinose and 1 mM IPTG, respectively. Whole cell lysates of TAXIMh-PQM and TRAPHi-SiaPQM cell suspensions were solubilized in 1% (w/v) DDM and submitted to ultracentrifugation to pellet aggregated protein. Target proteins were detected by *in gel* GFP fluorescence. (B) Fumarate uptake by TAXIMh-PQM into *E. coli* whole cells IMW424 **(***ΔDcuA/B*). Cells were pre-energized by a two-minute pre-incubation with 0.5% (w/v) glucose using an inward-directed Na^+-gradient. Transport was assayed at 30 °C and pH 7.5 in the presence of 4.1 µM [^{14}C] labeled fumarate. All samples but TRAPHi-SiaPQM were measured in duplicates, the error bars shown represent the standard deviation. Black and white arrows indicate the positions of TAXIMh-QM and TRAPHi-QM, respectively. The plot was created using OriginPro 2019b (OriginLab Corporation).

To determine the apparent affinity and maximum transport rate of the complete transport system, we determined the initial transport rate of a fumarate concentration from 1 to 102 µM (Figure 61, B). We calculated an apparent affinity of TAXIMh-PQM for fumarate of 14.6 µM (Figure 61, C). Whole cell lysates of samples from radioactive transport assay were solubilized in 1 % (w/v) DDM submitted to ultracentrifugation to pellet aggregated protein. Since the aggregated protein is collected in the pellet during ultracentrifugation and thereby separated, we consequently assume that the supernatant contains functional protein only. As we expressed the protein from the pEXC3sfGH plasmid, resulting in a carboxy-terminal fusion to sfGFP, we used the intensity of the *in gel* fluorescence of the GFP

calibration curve to estimate the concentration of TAXIMh-QM (Figure 61, A). Thus, all protein samples were separated in a 12 % gel and visualized by *in gel* GFP fluorescence in which the amount of detected sfGFP then corresponds to the amount of TAXIMh-QM protein. Expression levels were compared using the signal intensities of protein bands from before and after ultracentrifugation to assess how much functional protein was present during the transport experiment. We additionally applied a standard series of the sfGFP to the gel to generate a calibration curve and thereby determine the amount of functional TAXIMh-QM used in the transport assay (Figure 61, A). For this purpose, we assumed that the sfGFP which is either free or fused to the target protein have the same fluorescence strength. The signal intensities were measured, and the QM concentrations were calculated from the known GFP concentrations. This required significant extrapolation outside the range of the GFP calibration curve as the intensities of TAXIMh-QM exceeded the highest applied GFP concentration used. We calculated an approximate expression level of 2300 molecules QM per cell assuming that the OD_{600} of 1 corresponds to 8 X 10^8 cells/mL. This results in a V_{max} of 1236 nmol fumarate per nmol TRAPMh-PQM per minute (Figure 61, C). We assume that the SBP was not limiting in these experiments and expressed at least to a similar level as the membrane components, as TAXIMh-P precedes TAXIMh-QM in the operon.

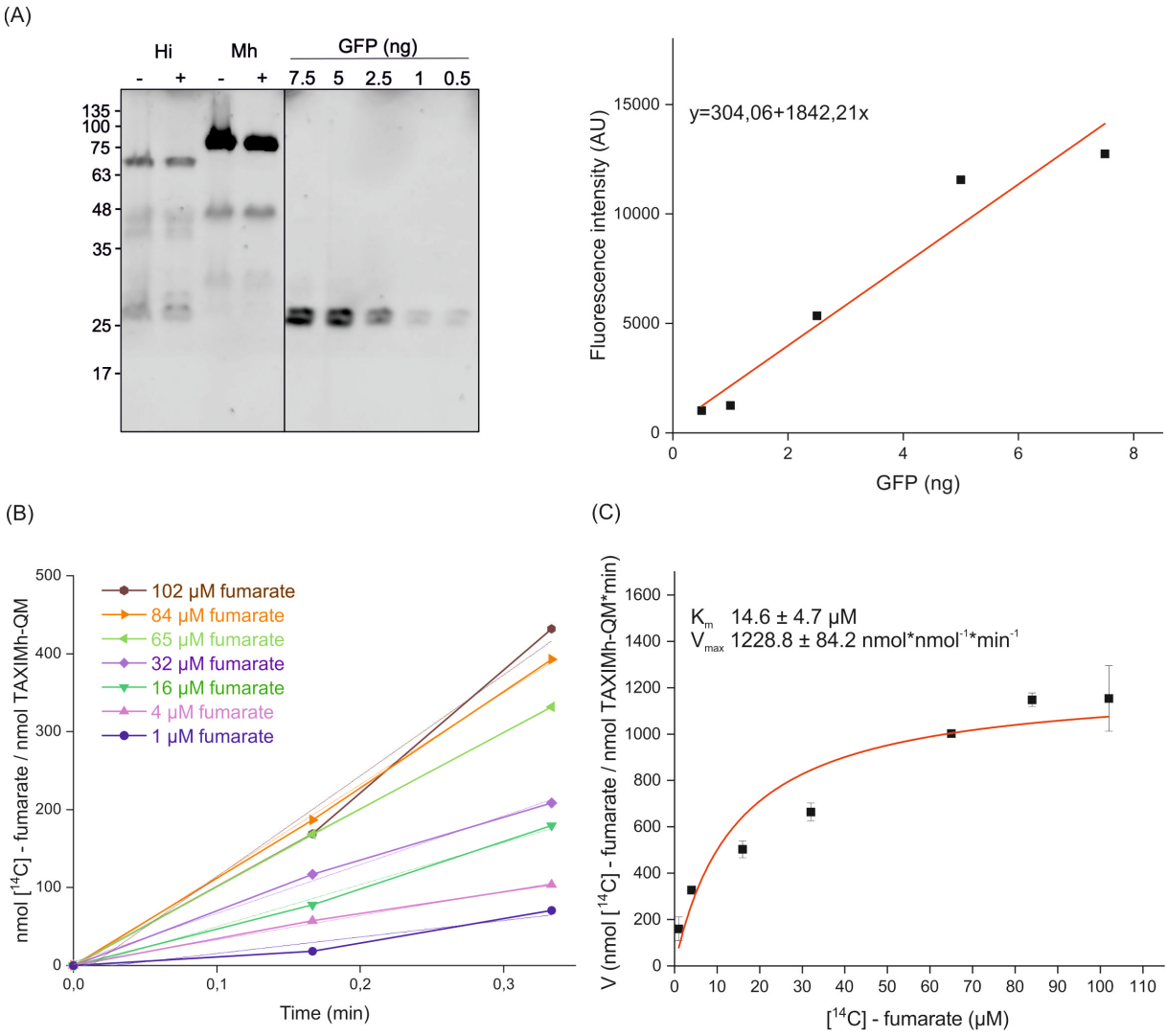


Figure 61 **Whole cell fumarate uptake by TAXIMh-PQM to determine K_m and Vmax.** (A) Whole cell lysates of TAXIMh-PQM and TRAPHi-SiaPQM cell suspensions used for radioactive transport assays were solubilized in 1% (w/v) DDM and submitted to ultracentrifugation to pellet aggregated protein. All protein samples were visualized by *in gel* GFP fluorescence. In addition, a GFP standard series was applied to create a GFP calibration curve and calculate the amount of protein in the gel. (B) Initial rate of fumarate uptake by TAXIMh-PQM in whole *E. coli* IMW424 cells *(ΔDcuA/B).* Cells were pre-energized by a 2 min pre-incubation with 0.5% (w/v) glucose. Transport was assayed at 30 °C and pH 7.5 in the presence of 125 nM to 102 µM [^{14}C] labeled fumarate using an inward-directed Na^+-gradient. Expression of TAXIMh-PQM in pEXC3sfGH was induced by 1 mM IPTG which attaches a sfGFP protein to QM. 0 sec, 10 sec and 20 sec time points were recorded in single measurements. The linear fit was created using Origin [ref|. (C) The initial uptake rates were plotted against the corresponding substrate concentrations to determine Km and V_{max} using the Michaelis Menten equation. The plots were created using OriginPro 2019b (OriginLab Corporation).

4.9.2 Membrane reconstitution of TAXIMh-QM

While whole cell uptake assays provided strong evidence that TAXIMh-PQM transports fumarate (Figure 60), this is most directly demonstrated *in vitro*. Furthermore, the defined nature of *in vitro* transport assays allows precise dissection of the driving forces. Therefore, the membrane domain TAXIMh-QM was expressed in the pBXC3sfGH plasmid attaching a carboxy-terminal composed of sfGFP and a decaHis tag. It was purified via IMAC whereby the target protein was eluted by HRV 3C protease cleavage which simultaneously removes the sfGFP and decaHis tag. The TAXIMh-QM was then further purified via SEC and reconstituted into liposomes composed of soy PC (Figure 62). The size-exclusion chromatogram of TAXIMh-QM shows a monodisperse peak (Figure 62), which overlaps with another broad peak eluting from the void volume. The corresponding gel shows that

the target protein was over 90 % pure. It migrates to an apparent size of approximately 60 kDa, which is below its predicted molecular weight of 78 kDa (Gasteiger et al. 2005). This increased electrophoretic mobility is more often observed for membrane proteins, for example also for TRAPHi-SiaQM and is accounted by the increased binding of SDS to the hydrophobic transmembrane segments (Mulligan et al. 2009). TAXIMh-QM was expressed with a sfGFP fusion which is visible in traces in the lower bands of the gel although cut off in the IMAC elution step by the HRV 3C protease. Additional contaminants originating from *E. coli* are possibly present in the gel. As soluble contaminants are not embedded in the membrane during reconstitution, the purity of the target protein TAXIMh-QM is further increased in this step (Figure 62). After reconstitution into soy PC liposomes, the efficiency was determined by resolubilizing the proteoliposomes in 1% (w/v) DDM followed by ultracentrifugation, separation of the proteins and visualization in a Coomassie stain. Assuming that well-reconstituted membrane protein remains folded, whereas membrane-associated proteins unfold during the reconstitution procedure, we expect that misfolded protein cannot be detergent-solubilized from the proteoliposomes and will consequently be cleared from the solution by ultracentrifugation. The reconstitution efficiency was inferred from the intensities of the protein bands on the SDS PAGE gel by quantifying the intensity of the TAXIMh-QM bands from before and after ultracentrifugation. We calculated a reconstitution efficiency of approximately 75 %, which is comparably high and should in principle suffice for *in vitro* transport assays (Figure 62).

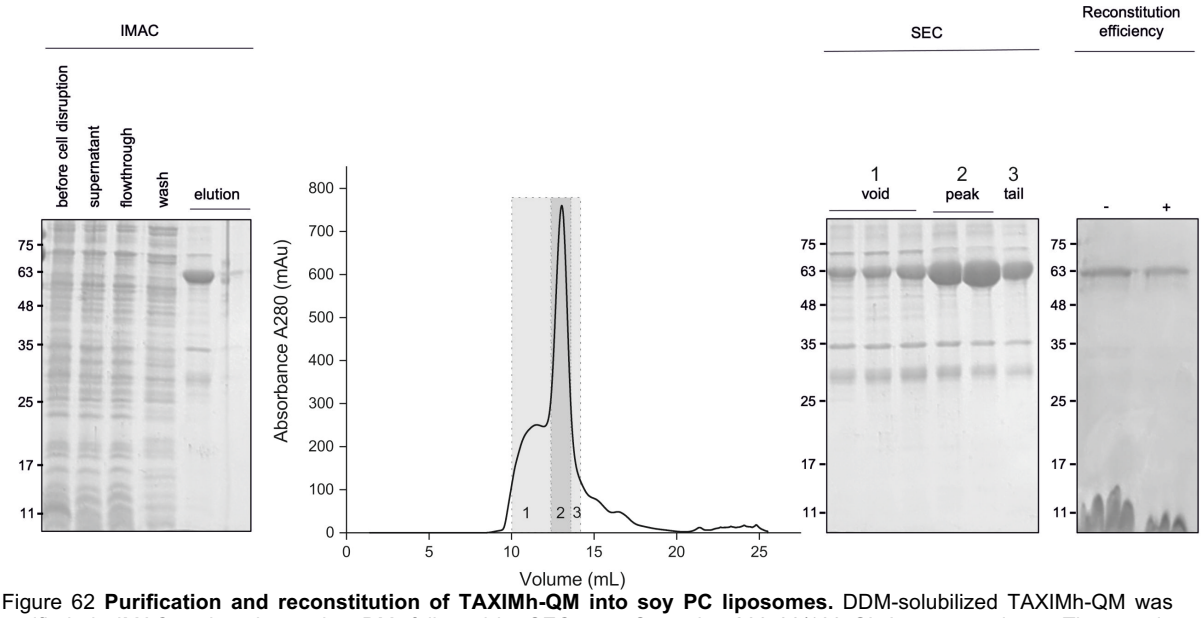


Figure 62 **Purification and reconstitution of TAXIMh-QM into soy PC liposomes.** DDM-solubilized TAXIMh-QM was purified via IMAC and exchanged to DM, followed by SEC on a Superdex 200 30/100 GL Increase column. The protein samples from the purification steps were separated in a 12 % SDS gel. Protein bands were visualized using a Coomassie stain. Pure protein fractions were mixed and reconstituted in a 1:50 ratio to liposomes. The reconstitution efficiency was determined by solubilizing the proteoliposomes in 1 % (w/v) DDM, followed by ultracentrifugation to pellet protein that is not reconstituted. The sample before and after ultracentrifugation were separated in a 12 % gel and visualised by Coomassie staining. The plot was created using OriginPro 2019b (OriginLab Corporation).

In addition, the SBP TAXIMh-P was expressed and purified. It was produced as a soluble protein using the pBXC3H vector that allows expression with a carboxy-terminal decaHis tag. TAXIMh-P was then purified by IMAC with imidazole elution, followed by Re-IMAC and SEC. The purified protein shows high homogeneity indicated by the monodisperse peak in the size-exclusion chromatogram

and corresponding protein bands in the gel. Pure TAXIMh-P migrates at approximately 35 kDa (Figure 63).

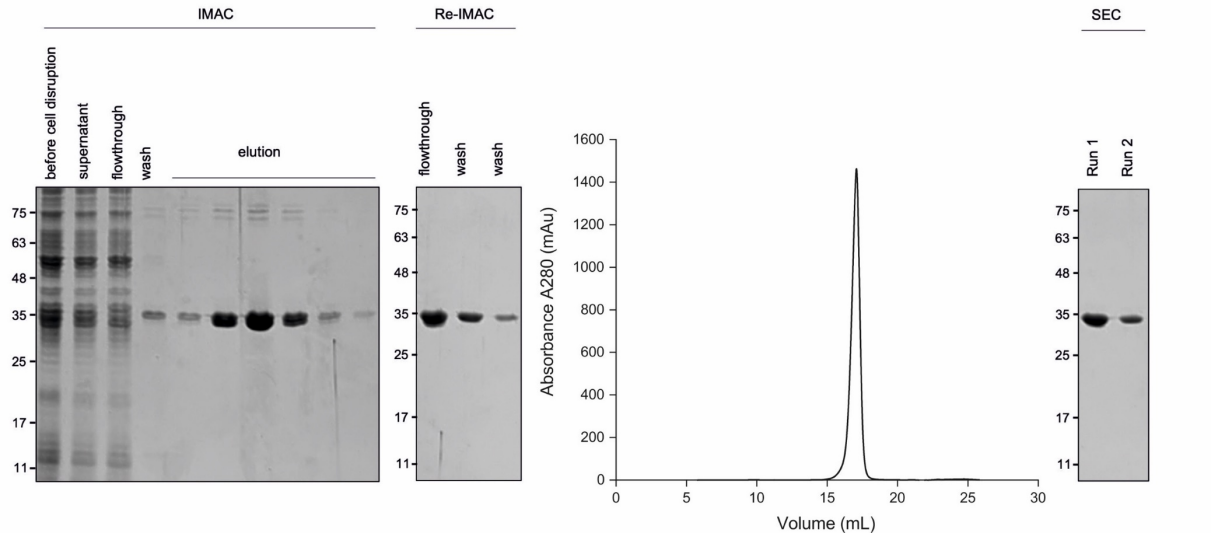


Figure 63 **Purification of TAXIMh-P.** TAXIMh-P was purified via IMAC, followed by Re-IMAC and SEC on a Superdex 200 30/100 GL Increase column. The protein samples from the purification steps were separated and visualized using a Coomassie stain. The plot was created using OriginPro 2019b (OriginLab Corporation).

4.9.3 *In vitro* fumarate transport by TAXIMh-PQM

Fumarate transport by TAXIMh-PQM was successfully determined in pre-energized *E. coli* cells (Figure 60). In these cells, three potential forces driving fumarate uptake by TAXIMh-PQM can be discerned: an inward-directed Na^+-gradient, an inward-directed proton-gradient, or an inward-directed substrate-gradient. Based on a transport assay previously established for the TRAP transporter TRAPHi-SiaPQM (Mulligan et al. 2009), we performed an *in-vitro* transport assay for TAXIMh-PQM proteoliposomes. For this purpose, the reconstituted TAXIMh-QM was previously stored in 10 % glycerol in liquid nitrogen and thawed prior to performing the transport assay. For each transport reaction, we used approximately 0.5 µM right-side-out reconstituted TAXIMh-QM. We then supplemented the external buffer with 5 µM soluble TAXIMh-P. To assure a high transport rate, a comparably high substrate concentration of 30 µM was used, which should approximate the v_{max} according to the K_m-value of 14.6 µM (Figure 61). For not only testing the dependence on an inward-directed Na^+- but also a proton gradient, we applied a combination of both. To additionally generate an inside-negative membrane potential 2 µM valinomycin was added. Despite the theoretically optimal conditions, no significant transport above the background signal was observed for reconstituted TAXIMh-PQM (Figure 64).

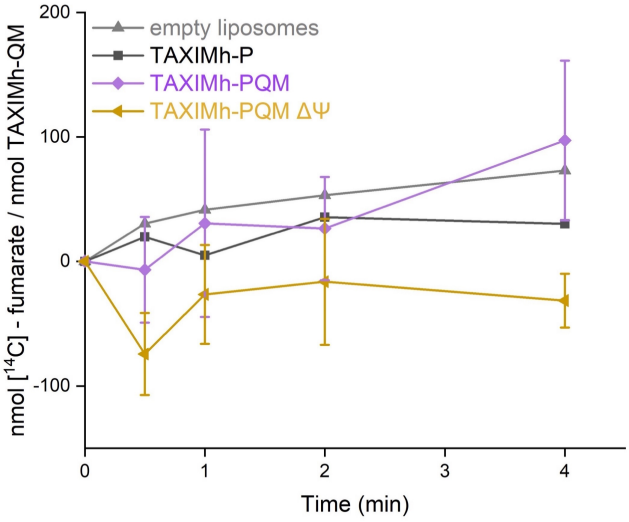


Figure 64 ***In vitro* uptake of fumarate in TAXIMh-QM proteoliposomes with soluble TAXIMh-P.** Fumarate uptake in TAXIMh-QM proteoliposomes was measured at 20 °C. 4.2 µg TAXIMh-QM was used per time point. The assay buffer in which the proteoliposomes were diluted contained 5 µM TAXIMh-P and 30 µM [^{14}C] labeled fumarate. A combination of an inward-directed Na^+- and proton gradient was applied which was in the ΔΨ sample supplied with 2 µM valinomycin to generate a negative membrane potential. As negative controls, empty liposomes and empty liposomes supplemented with soluble TAXIMh-P were used. The TAXIMh-PQM and TAXIMh-PQM ΔΨ samples were measured in duplicates, the negative controls represent single datapoints, error bars indicate standard deviations. The plot was created using OriginPro 2019b (OriginLab Corporation).

4.9.4 Structure-function studies of the small membrane domain TAXIMh-Q

As the fumarate transport by TAXIMh-PQM into *E. coli* whole cells could be measured well (Figure 60), it allowed us to further study structure-function relationships. To investigate the relevance of the small membrane domain TAXIMh-Q from *Marinobacter hydrocarbonoclasticus*, we created Q deletion and truncation variants. We first had to determine which part of the protein sequence belongs to the small membrane domain Q and which part to the large membrane domain M as they occur fused in TAXIMh-PQM. For this, we performed Clustal Ω alignments with protein sequences in which Q and M appeared separated or fused in the operon (Sievers et al. 2011). An example in which they occur separated is TRAPRc-DctQM from *Rhodobacter capsulatus* (Jacobs et al. 1996). In addition, the assumptions made for the TRAPHi-SiaQM and TRAPVc-SiaQM in the literature were compared as the N-acetylneuraminic acid transporter occurs in *Haemophilus influenzae* as a QM fusion protein but in *Vibrio cholerae* with Q and M separately expressed (Mulligan et al. 2012). In the absence of an experimental structure, we used a Topcon prediction and alphafold model (Tsirigos et al. 2015; Jumper et al. 2021) to create a topology plot of TAXIMh-QM. According to the predictions, the TAXIMh-Q domain contained four transmembrane helices and two periplasmic loops. For TAXIMh-M the helices were additionally assigned to the scaffold and transport domains one and two respectively (Figure 65). With this, the bioinformatic analysis suggested a TAXIMh-QM fusion protein with four transmembrane helices in TAXIMh-Q and 12 transmembrane helices as well as two helical hairpins in TAXIMh-M. Overall, the topology looked similar to what was shown in the structure of TRAPHi-SiaQM from *Haemophilus influenzae* with four transmembrane helices in TRAPHi-SiaQ, 11 transmembrane helices in TRAPHi-SiaQM as well as two additional helical hairpins in M (Peter et al. 2022). If we compare the TAXIMh-QM model with the TRAPHi-SiaQM structure, one additional

transmembrane helix was predicted for TAXIMh-QM in the second transport domain (Figure 65). Thereby, TAXIMh-QM adhered as TRAPHi-SiaQM to the general design principles of elevator proteins as it consists of an inverted repeat and a scaffold and transport domain (Drew and Boudker 2016). Furthermore, TAXIMh-QM was predicted with two barrier helices that are a distinctive feature in elevator type proteins (Figure 65) (Drew and Boudker 2016).

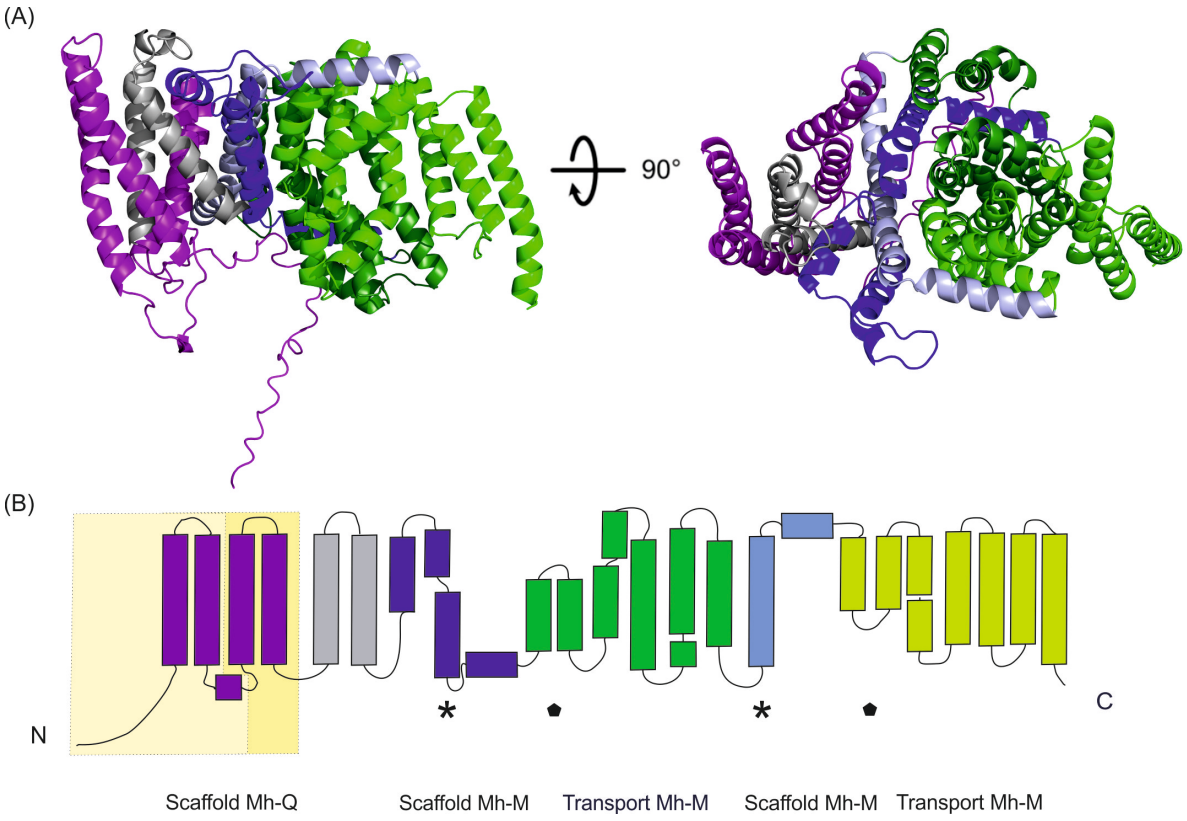


Figure 65 **Model of TAXIMh-QM using Topcon and alphafold predictions.** (A) Alphafold model of TAXIMh-QM showing the side view of QM approximately positioned as in the membrane and the top view. Q is depicted in purple, the connector unit between Q and M in grey, the scaffold domains 1 and 2 in dark and light blue and the transport domain 1 and 2 in dark and light green. The orientation of the protein presented is based on Topcon predictions. (B) Topology plot of TAXIMh-QM created based on alphafold predictions. The same color scheme for the different domains is used as in A. The star marks the two barrier helices, which are a recurring feature in the scaffold domains of elevator proteins. The potential hairpins in the transport domains are marked with pentagons. The complete TAXIMh-Q is shaded dark yellow and the partial TAXIMh-Q light yellow. (Tsirigos et al. 2015; Jumper et al. 2021) The structures were created using Pymol 1.8.4.1 (Schrödinger, LLC).

Based on the TRAPPp-SiaQM structure from *Photobacterium profundum*, it appears likely that the Q-domain forms an extension of the M scaffold domain and functions to provide stability for the protein complex in the membrane (Davies, Currie, and North 2023). However, it is still unclear whether additional functional tasks are involved and where exactly the interaction with the substrate-binding protein occurs. To investigate the relevance of the Q-domain for successful transport in detail, we cloned Q deletion and truncation variants, in which we either exchanged the first periplasmic loop of Q by a Glycine Serine linker, deleted Q completely or partially. The primers for the variant in which Q was completely removed were designed such that the transport domain in M remained intact (Figure 66) (Figure 117, section 6). The complete Q deletion variant was used to show whether Q is needed for successful transport. Furthermore, a variant was generated in which one half was deleted, resulting in a construct in which the first of the two periplasmic loops were

absent. This should show whether the second longer loop is sufficient for the binding of the SBP and whether a truncated Q as an extended scaffold provides enough stability in the membrane for M to enable successful transport. To examine this in detail, a third variant was created in which five amino acids of the first loop were exchanged for a glycine serine linker (Figure 66). The recently generated alphafold model of TAXIMh-Q supports the assumptions that the intended parts of Q were deleted or exchanged (Figure 65) (Figure 66).

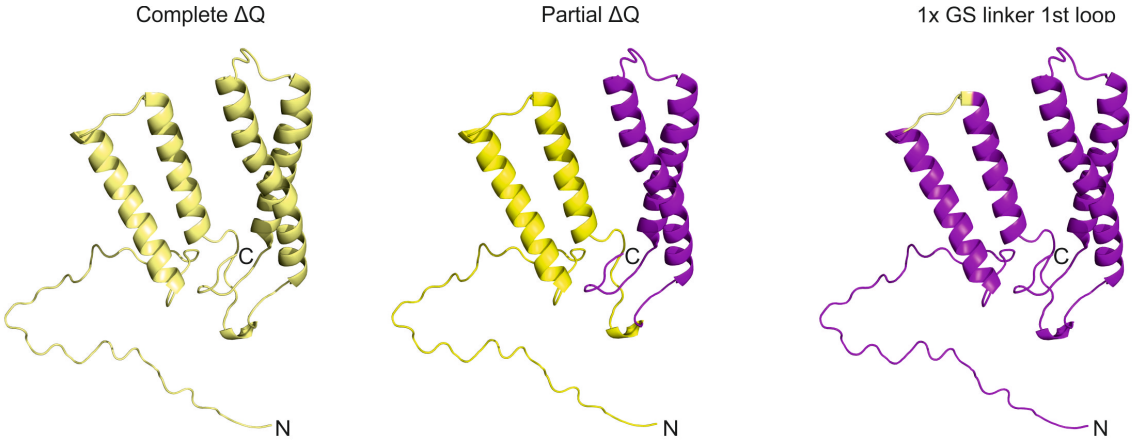


Figure 66 **Alphafold model of TAXIMh-Q variants.** TAXIMh-Q deletion variants in side view with Q in the front (Jumper et al. 2021). TAXIMh-Q was either completely (left) or partially deleted (middle) or five residues in the first periplasmic loop were exchanged by a glycine serine linker consisting of four glycine followed by one serine (right). The deleted parts are shown in yellow and the remaining parts in purple. N- and C-terminus are indicated. The structures were created using Pymol 1.8.4.1 (Schrödinger, LLC).

The variants were then compared in a radioactive transport assay to the wild-type protein. All three Q variants showed no transport (Figure 67, A). The full-length wildtype protein showed uptake of fumarate as demonstrated previously (Figure 67, A). To determine whether poor expression levels underlie this lack of transport activity, we quantified the protein levels of the samples used in the uptake experiment (Figure 67, B). Although there are very different expression levels in the different variants, protein is always present. The more of Q is truncated, the more the expression is reduced. In the variant in which five amino acids in the first periplasmic loop have been exchanged for a glycine serine linker, expression is higher compared to the wild-type protein (Figure 67, B). This underlines the significance of the loop, as this variant also led to a complete prevention of transport implying that the loss in activity cannot be explained by a reduced protein expression level.

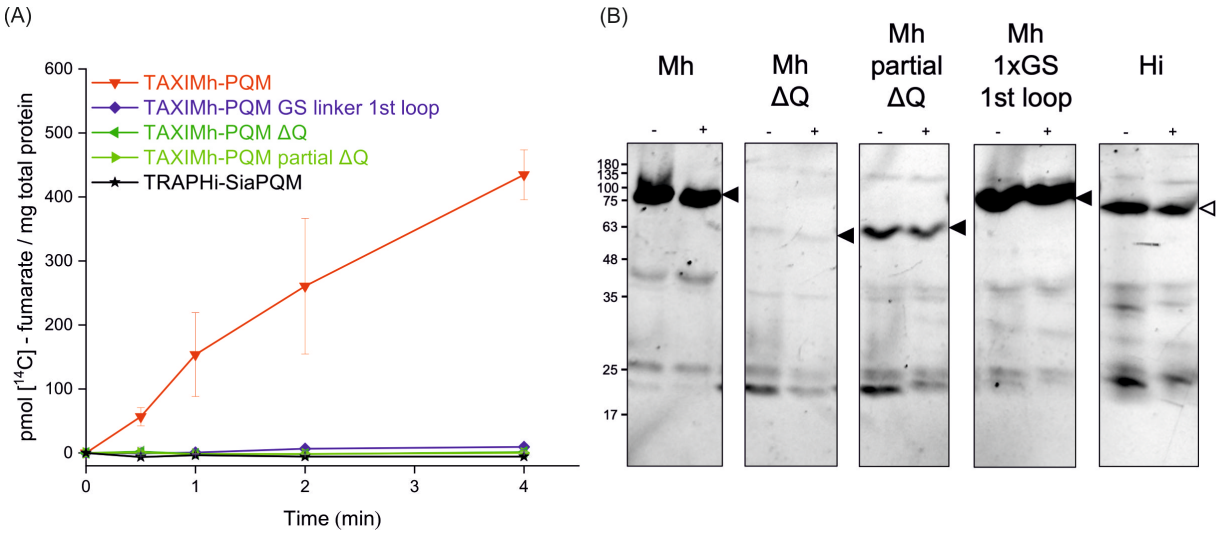


Figure 67 **Whole cell fumarate uptake by TAXIMh-PQM and comparison of protein expression levels.** (A) Fumarate uptake by TAXIMh-PQM and TAXIMh-PQM constructs, containing a Q deletion, partial deletion or amino acid exchanges in the periplasmic loops. The transport was recorded in *E. coli* whole cells IMW424 (*ΔDcuA/B*) at 30 °C using 4.09 µM [^{14}C] labeled fumaric acid and an inward-directed Na^+-gradient. TAXIMh-PQM was expressed with a lac promoter from the pEXC3sfGH plasmid which attaches a sfGFP protein to QM. The TRAPHi-SiaPQM from *Haemophilus influenzae* was used as negative control. All samples but the TRAPHi-SiaPQM were measured in duplicates, the shown error bars equal the standard deviations. (B) Whole cell lysates of samples from radioactive transport assay were solubilized in 1 % (w/v) DDM and ultracentrifuged in half to separate aggregated protein to assess how much functional protein was present during the transport experiment. All protein samples were visualized by *in gel* GFP fluorescence. Black and white arrows indicate the positions of TAXIMh-QM and variants and TRAPHi-QM, respectively. The plot was created using OriginPro 2019b (OriginLab Corporation).

So, the lack of transport with simultaneous expression of the deletion variants demonstrated the relevance of Q in the transport process. This was particularly evident in the minimal change in the protein, where the first periplasmic loop was replaced by a glycine serine linker. This abolished the transport completely but does not adversely affect expression (Figure 67). The significance of the specific amino acids in the first periplasmic loop of Q suggested that the SBP interacts with the first periplasmic loop of TAXIMh-Q and possibly with other positions in the protein to transfer the substrate to the wildtype QM protein which then facilitates transport.

4.10 α-Ketoglutarate transport by TAXIPm-PQM

4.10.1 α-Ketoglutarate transport by TAXIPm-PQM into whole cells

Analysis of the SBP of the TAXI transporter TAXIPm-PQM from *Proteus mirabilis* using DSF and ITC indicated that α-ketoglutarate and α-hydroxyglutarate are bound with µM affinity (4.5.3 and 4.7.2). To verify whether these compounds are recognized and transported by the QM membrane domain as well, we established a transport assay. For this purpose, we amplified the complete TAXI operon from the genomic DNA of *Proteus mirabilis* and expressed it in the pEXC3sfGH plasmid with the signal sequence of TRAP-SiaP from *Haemophilus influenzae* preceding the SBP. Additionally, a carboxy-terminal fusion composed of sfGFP and a decaHis tag was attached to the TAXIPm-QM domain. Protein production was successful using the *lac* promoter (Figure 68, A). As the SBP is on the same operon as TAXIPm-QM, we expect TAXIPm-P to be expressed as well. To assess the folding quality, we compared the intensities of protein bands from DDM-solubilized whole cell lysates before and after ultracentrifugation. Both proteins, TAXIMh-QM and TAXIPm-QM were expressed to high levels thereby showing good protein quality as 89 % of the TAXIMh-QM and 91 % of the TAXIPm-QM proteins were detected in the supernatant after ultracentrifugation (Figure 68, A). Radioactive α-ketoglutarate uptake was measured for TAXIPm-PQM in *E. coli* JW2571-1 cells, which lack the endogenous *E. coli* α-ketoglutarate permease encoded by the *kgtP* gene (Seol and Shatkin 1991). We observed robust TAXIPm-PQM dependent uptake of α-ketoglutarate in glucose-energized cells, indicating that α-ketoglutarate is a substrate for the membrane-embedded transport domain as well (Figure 68, B). A homolog of TAXIPm-PQM, named TAXIMh-PQM and for which we were able to show in DSF that the SBP does not bind α-ketoglutarate, serves as a negative control (section 4.5.2). All transport of α-ketoglutarate into whole cells containing the negative control is considered background activity (Figure 68, B).

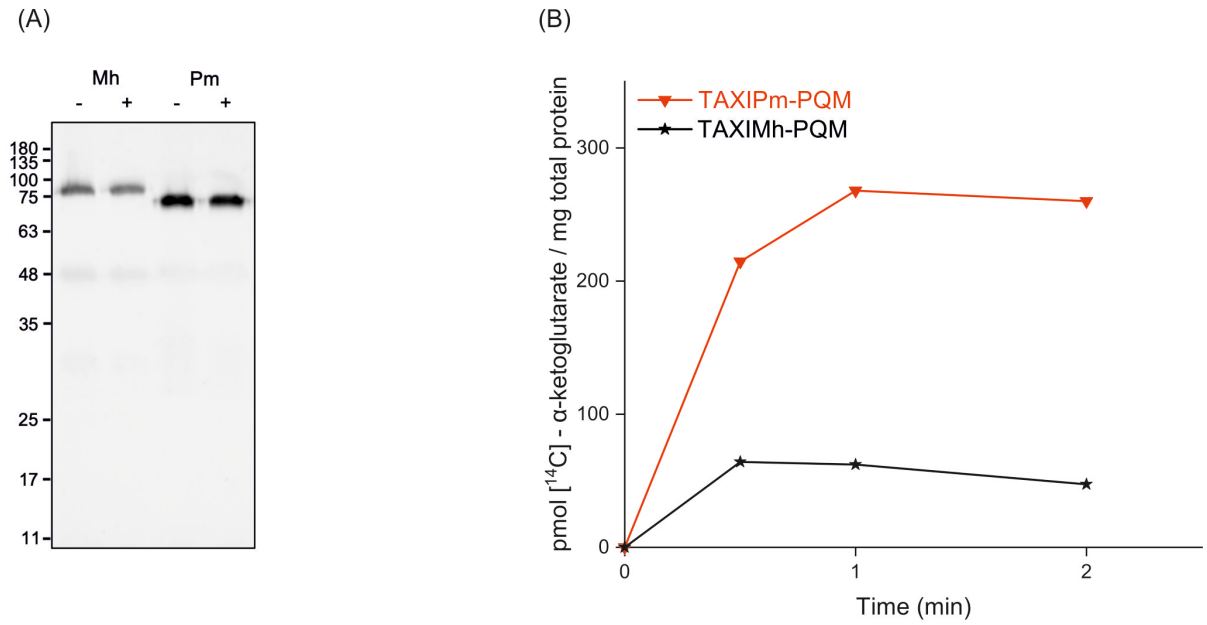


Figure 68 **Whole cell α-ketoglutarate uptake by TAXIPm-PQM.** (A) TAXIPm-PQM and the negative control TAXIMh-PQM from *Marinobacter hydrocarbonoclasticus* were expressed in pEXC3sfGH with the TRAPHi-SiaP signal peptide from *Haemophilus influenzae*, thereby induced using 1 mM IPTG. Whole cell lysates of TAXIMh-PQM and TAXIPm-PQM cell suspensions were solubilized in 1 % (w/v) DDM and submitted to ultracentrifugation to pellet aggregated protein. The target proteins were detected by *in gel* GFP fluorescence. (B) α-ketoglutarate uptake by TAXIPm-PQM in whole *E. coli* cells JW2571-1 (*ΔkgtP*). Cells were pre-energized by a 2 min pre-incubation with 0.5 % (w/v) glucose. Transport was assayed at 30 °C and pH 7.5 in the presence of 2 µM [^{14}C] labeled α-ketoglutarate using an inward-directed Na^+-gradient. All time points reflect single measurements. The plot was created using OriginPro 2019b (OriginLab Corporation).

4.10.2 Proton dependent transport by TAXIPm-PQM in whole cells

α-Ketoglutarate transport by TAXIPm-PQM was successfully determined in pre-energized *E. coli* cells (Figure 68) (Figure 120, section 6). In these cells, four potential forces driving α-ketoglutarate uptake by TAXIPm-PQM were investigated: the membrane potential, an inward-directed Na^+-gradient, proton-gradient and substrate gradient. In our *in vivo* transport assays, we supplied substrate from the outside which automatically generates the substrate gradient, that is *in vivo* most likely reinforced by cell metabolism. As the initial transport assays showed successful α-ketoglutarate uptake by TAXIPm-PQM in both cases in presence and absence of a sodium gradient, we assumed a potential proton coupling (Figure 120, section 6). Therefore α-ketoglutarate transport by TAXIPm-PQM was tested in presence and absence of the protonophore CCCP in the presence of a proton gradient. The CCCP diminished the proton gradient and the uptake, implying that the transport is coupled to the proton-motive force (Figure 69).

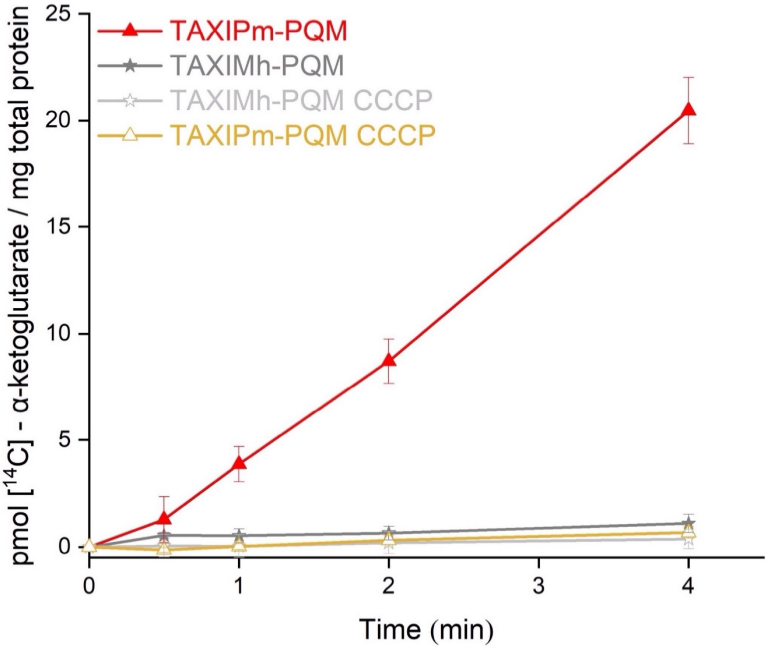


Figure 69 **Effect of the protonophore CCCP on α-ketoglutarate uptake by TAXIPm-PQM in whole cells.** α-ketoglutarate uptake by TAXIPm-PQM in whole *E. coli* JW2571(*ΔkgtP*) in the presence of a proton gradient. Cells were pre-energized by a 2 min pre-incubation with 0.5% (w/v) glucose, followed by the subsequent addition of 10 µM CCCP and a further two min pre-incubation. Transport was assayed at 30 °C and pH 6.0 in the presence of 2 µM [^{14}C] labeled α-ketoglutarate. Both, TAXIPm-PQM and the negative control TAXIMh-PQM from *Marinobacter hydrocarbonoclasticus* were expressed in pEXC3sfGH with the TRAPHi-SiaP signal peptide from *Haemophilus influenzae* and induced with 1 mM IPTG. All samples were measured in triplicates, the error bars shown represent the standard deviation. The plot was created using OriginPro 2019b (OriginLab Corporation).

4.10.3 Michaelis Menten kinetics in α-Ketoglutarate transport by TAXIPm-PQM

To determine the apparent affinity and maximum transport rate of the complete transport system, we determined the initial transport rate of a [^{14}C] labeled α-ketoglutarate concentration range from 0.25 to 102.4 µM (Figure 70). We calculated an apparent affinity of TAXIPm-PQM for α-ketoglutarate of 31 µM at pH 7.5 and of 84 µM when the pH has been shifted to 6.0, which affects the charge of the transporter and the C_5-dicarboxylic α-ketoglutarate (Figure 70). We used the intensity of the *in gel* fluorescence of the sfGFP fusion protein to evaluate the amount of total protein present during the transport (Figure 118, section 6). This results in a V_{max} of 2984 pmol α-ketoglutarate per mg total protein per minute at pH 7.5 and a V_{max} of 2611 pmol α-ketoglutarate per mg total protein per minute at pH 6.0. We assume that the SBP was not limiting in these experiments and expressed at least to a similar level as the membrane components, as TAXIPm-P preceeds TAXIPm-QM in the operon. With the change of the pH from 7.5 to 6.0, especially the proportion of the present substrate species changes simultaneously. As the charge of the C_5-dicarboxylate α-ketoglutarate is determined by the pK_a-values at 1.9 and 4.4 (Tokonami et al. 2013), we calculated an approximate 25-fold increase of the single negative form connected to the change of the pH being 0.1 % at pH 7.5 and 2.5 % at pH 6.0. So, if the single negative species would be the one that is transported with higher affinity, we assume a 25-fold change in the apparent K_m. As we observed a 2.7-fold difference with the change

in the K_m from 31 to 84 µM (Figure 70), we suggest that the di-anionic species is favorably transported.

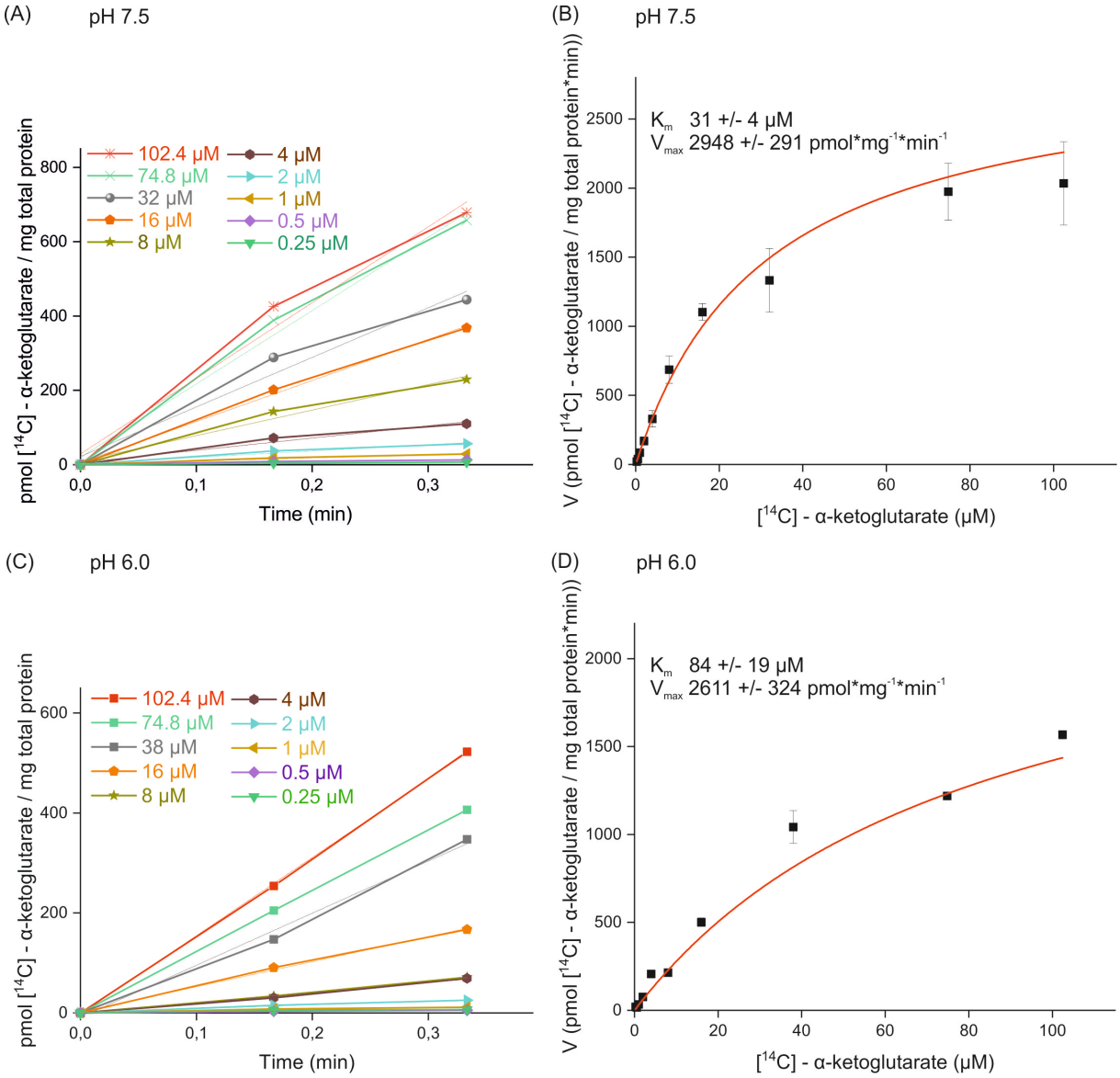


Figure 70 **Whole cell α-ketoglutarate uptake by TAXIPm-PQM to determine K_m and Vmax at pH 7.5 and pH 6.0.** (A) Initial rate of α-ketoglutarate uptake by TAXIPm-PQM in whole *E. coli* JW2571-1 cells (*ΔkgtP*), expressed in pEXC3sfGH with the TRAPHi-SiaP signal peptide from *Haemophilus influenzae* and induced with 1 mM IPTG. Cells were pre-energized by a 2 min pre-incubation with 0.5% (w/v) glucose. Transport was assayed at 30 °C and pH 7.5 in the presence 0.25 to 102.4 µM [^{14}C] labeled α-ketoglutarate. 0 sec, 10 sec and 20 sec time points were recorded in single measurements. The linear fit was created using Origin (OriginLab Corporation). (B) The initial uptake rates measured at pH 7.5 were plotted against the corresponding substrate concentrations to determine Km and V_{max} using the Michaelis Menten equation. (C) Initial rate of α-ketoglutarate uptake by TAXIPm-PQM in whole *E. coli* JW2571-1 cells (*ΔkgtP*), expressed in pEXC3sfGH with the TRAPHi-SiaP signal peptide from *Haemophilus influenzae* and induced with 1 mM IPTG. Cells were pre-energized by a 2 min pre-incubation with 0.5 % (w/v) glucose. Transport was assayed at 30 °C and pH 6.0 in the presence 0.25 to 102.4 µM [^{14}C] labeled α-ketoglutarate. 0 sec, 10 sec and 20 sec time points were recorded in single measurements. The linear fit was created using OriginPro 2019b (OriginLab Corporation). (D) The initial uptake rates measured at pH 6.0 were plotted against the corresponding substrate concentrations to determine K_m and Vmax using the Michaelis Menten equation. The plots were created using OriginPro 2019b (OriginLab Corporation).

4.10.4 Reconstitution of TAXIPm-QM to soybean L-α-phosphatidylcholine (soy PC)

To further characterize the transport of TAXIPm-PQM *in vitro*, the SBP TAXIPm-P and the transport unit TAXIPm-QM were both purified with a subsequent reconstitution of TAXIPm-QM into liposomes. While whole cell uptake assays provided strong evidence that TAXIPm-PQM transports α-ketoglutarate, this is most directly demonstrated *in vitro*. Furthermore, the defined nature of *in vitro* transport assays allows precise dissection of the driving forces. Therefore, the membrane domain TAXIPm-QM was expressed in the pBXC3sfGH plasmid attaching a carboxy-terminus composed of 3C cleavage site, sfGFP and decaHis tag. The TAXIPm-QM protein was purified by IMAC with a final elution step mediated by HRV 3C protease, which specifically cleaves the His tag and GFP fusion from the target (Figure 71, A). This was followed by SEC in which TAXIPm-QM elutes in two peaks (Figure 71, B). The first peak might correspond to a dimeric form, although recently published structures of TRAPPp-SiaQM and TRAPHi-SiaQM suggest that TRAP transporter do not oligomerize (Davies, Currie, and North 2023; Peter et al. 2022). Therefore, the first peak might resemble aggregates of the target protein and the second peak most likely the monomeric form of TAXIPm-QM. The associated gel shows that the target protein was over 90 % pure in both the first and second peak, whereby TAXIPm-QM migrates to an apparent size of approximately 50 kDa, which is below its predicted molecular weight of 70 kDa (Figure 71, C) (Gasteiger et al. 2005). This increased electrophoretic mobility is more often observed for membrane proteins, for example also for TRAPHi-SiaQM and is accounted by the increased binding of SDS to the hydrophobic transmembrane segments (Mulligan et al. 2009). The fractions from both peaks were mixed for the following reconstitution into liposomes. As soluble contaminants are not embedded in the membrane in the course of reconstitution, the purity of the target protein TAXIPm-QM is further increased in this step (Figure 71 D). Two different reconstitution techniques were compared to achieve the highest possible efficiencies. For this, the detergent was either rapidly diluted or absorbed by BioBeads, which in both cases leads to membrane embedment of the protein. The efficiency of the reconstitution into soy PC liposomes was determined by resolubilizing the proteoliposomes in 1% (w/v) DDM followed by ultracentrifugation and detection of the proteins in a Coomassie stained SDS PAGE gel (Figure 71, D). Assuming that membrane-inserted protein remains folded, whereas membrane-associated proteins unfold during the reconstitution procedure, we expect that misfolded protein cannot be detergent-solubilized from the proteoliposomes and will consequently be cleared from the solution by ultracentrifugation. The treatment with BioBeads to remove the detergent with simultaneous reconstitution was found to be very efficient compared to the rapid dilution. We calculated a reconstitution efficiency of approximately 91 % when using BioBeads and 36 % when the detergent was rapidly diluted to initiate reconstitution.

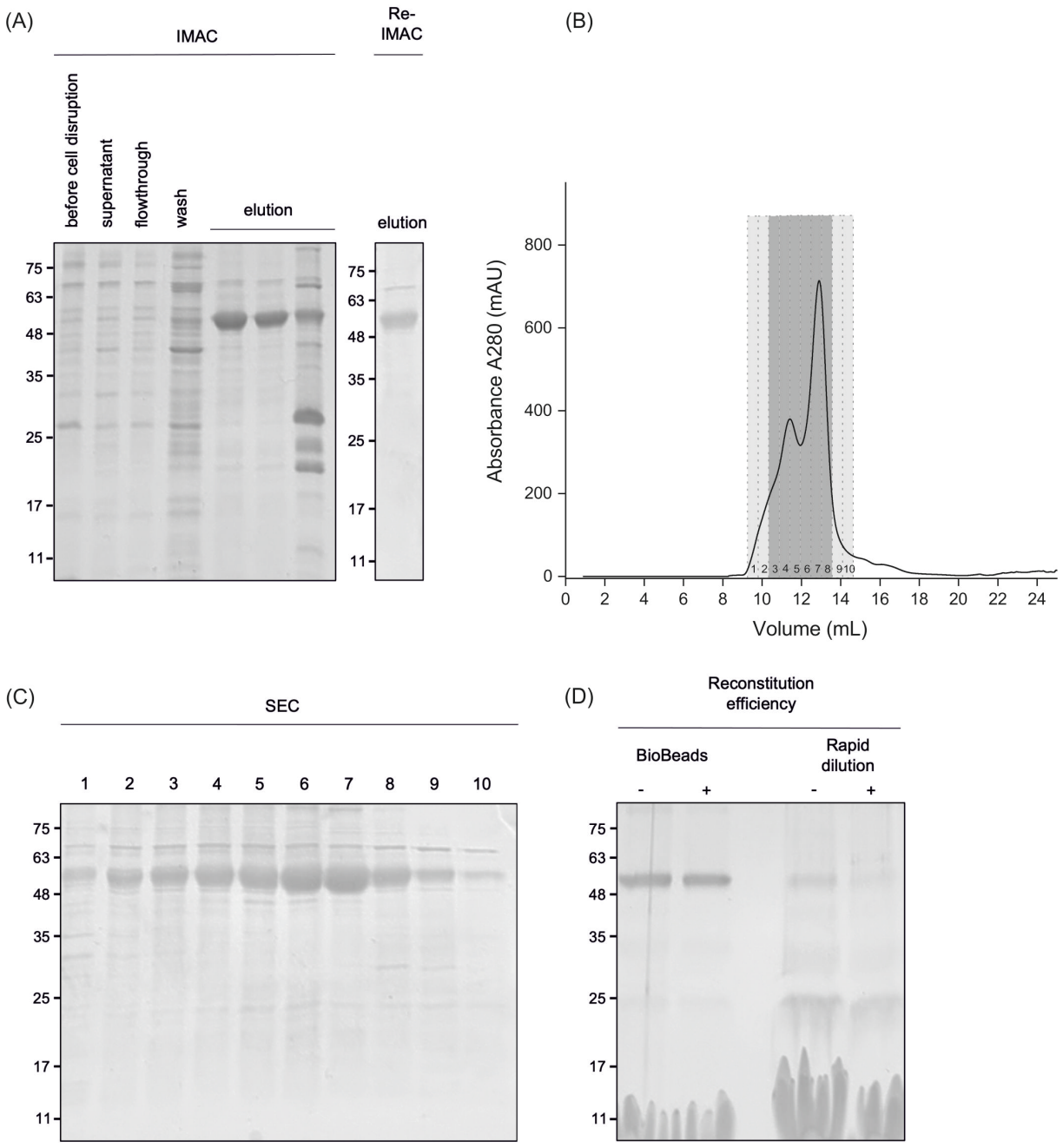


Figure 71 **Purification and reconstitution of TAXIPm-QM to soy PC liposomes.** (A) DDM-solubilized TAXIPm-QM was purified via IMAC and exchanged to DM with a final elution by 3C protease. (B) The IMAC was followed by SEC on a Superdex 200 30/100 GL Increase column. (C) The protein samples from the SEC were separated in a 12 % SDS gel. Protein bands were visualized using a Coomassie stain. (D) Pure protein fractions were mixed and reconstituted in a 1:50 ratio to liposomes. Either BioBeads were added, or a rapid dilution was performed to allow for reconstitution. The efficiencies were determined and compared by re-solubilizing the proteoliposomes in 1 % (w/v) DDM, followed by ultracentrifugation to pellet protein that is not reconstituted. The sample before and after ultracentrifugation were separated in a 12 % gel and visualized by Coomassie staining. Signal intensities were determined using ImageJ and the plot was created using OriginPro 2019b (OriginLab Corporation).

In addition, the SBP was expressed and purified. TAXIPm-P was produced as a soluble protein using the pBXNH3 vector that allows expression with an N-terminal decaHis tag and purification by IMAC (Geertsma and Dutzler 2011). The SBP was obtained in a high purity after IMAC, followed by Re-IMAC and SEC (Figure 72). The purified protein shows high homogeneity indicated by the monodisperse peak in the size-exclusion chromatogram and corresponding protein bands in the gel.

Pure TAXIPm-P migrates to approximately 35 kDa which equals its predicted molecular weight (Figure 72) (Gasteiger et al. 2005).

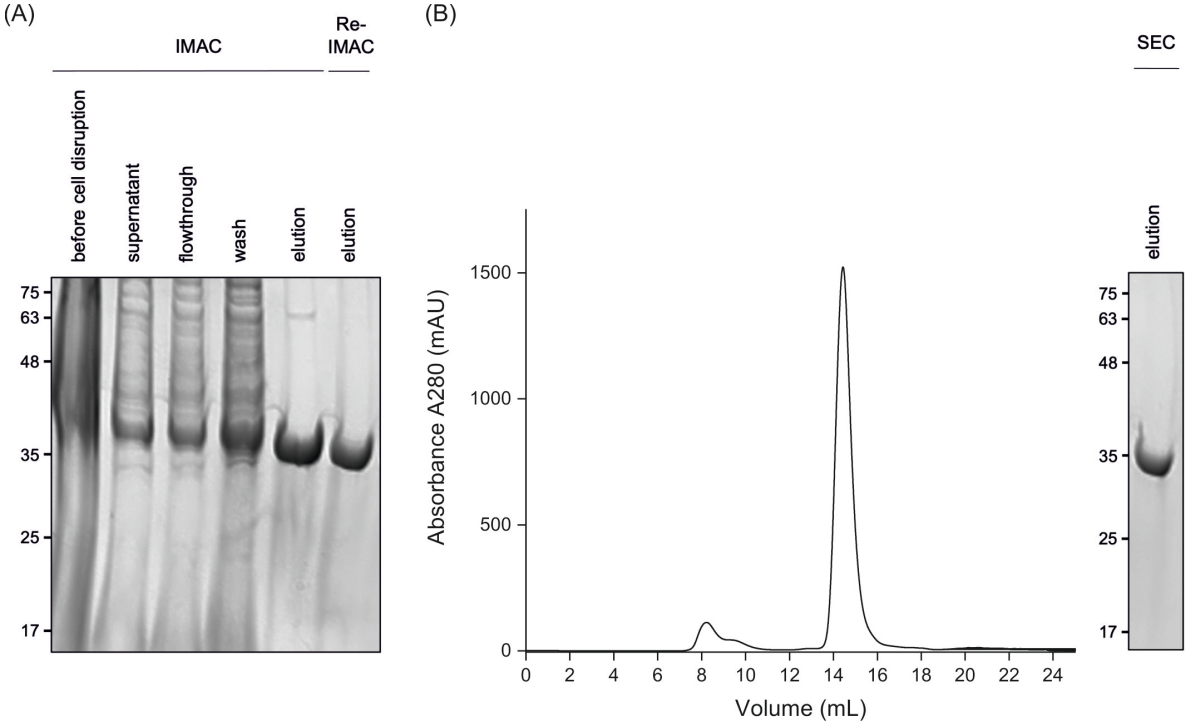


Figure 72 **Purification of TAXIPm-P**. (A) TAXIPm-P was purified via IMAC with a final imidazole elution step, followed by Re-IMAC. (B) The TAXIPm-P protein was further purified by SEC on a Superdex 200 30/100 GL Increase column. The protein samples from the purification steps were separated in a 12 % (w/v) SDS gel. Protein bands were visualized using a Coomassie stain. The plot was created using OriginPro 2019b (OriginLab Corporation).

4.10.5 *In vitro* transport assay of TAXIPm-QM and soluble TAXIPm-P

α-ketoglutarate transport by TAXIPm-PQM was successfully determined in pre-energized *E. coli* cells (Figure 68) (Figure 120, section 6). In these cells, four potential forces driving α-ketoglutarate uptake by TAXIPm-PQM can be discerned: the membrane potential, an inward-directed Na^+-gradient, proton-gradient and substrate gradient. The lack of transport upon the addition of the protonophore CCCP suggested that the α-ketoglutarate transport by TAXIPm-PQM is coupled to the proton-motive force. Based on this knowledge and a transport assay previously established for the TRAP transporter TRAPHi-SiaPQM (Mulligan et al. 2009), we performed *in vitro* transport assays for TAXIPm-PQM proteoliposomes. The reconstituted TAXIPm-QM was stored in 10 % glycerol in liquid nitrogen and thawed prior to performing the transport assay. For each transport reaction, we used 0.5 µM right-side-out reconstituted TAXIPm-QM. In our initial *in vitro* transport experiments, we supplemented the external buffer with 5, 10 or 50 µM soluble TAXIPm-P or no SBP as negative control. Neither an inward-directed Na^+-gradient with 5 µM substrate and 5, 10 or 50 µM soluble SBP, nor an inward-directed proton gradient with 5 µM substrate and 5 µM soluble TAXIPm-P facilitated transport (Figure 73).

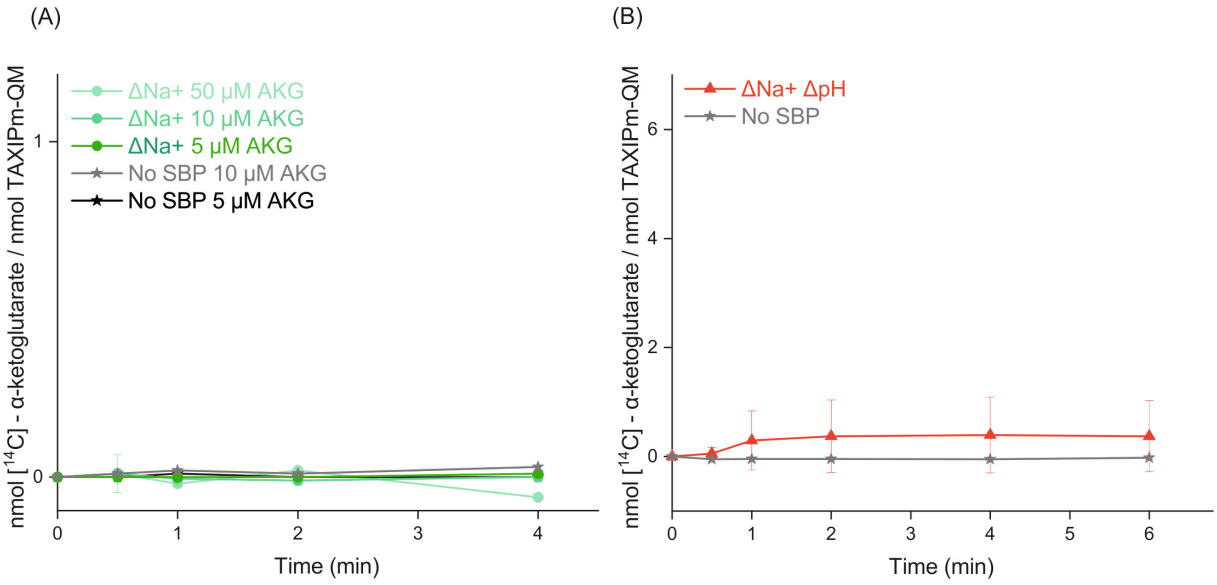


Figure 73 ***In vitro* transport of α-ketoglutarate in TAXIPm-PQM proteoliposomes with TAXIPm-P inside and outside proteoliposomes.** α-ketoglutarate uptake in TAXIPm-PQM proteoliposomes at 20 °C after reconstitution of TAXIPm-QM. (A) The assay buffer in which the proteoliposomes were diluted contained 5 µM soluble TAXIPm-P and 5, 10 or 50 µM [^{14}C] labeled α-ketoglutarate, respectively. An inward-directed Na^+-gradient was applied. The 5 µM sample was measured in duplicates, error bars indicate standard deviations. All other samples and the negative controls which contained no SBP TAXIPm-P were recorded in single measurements. (B) The assay buffer in which the proteoliposomes were diluted contained 5 µM soluble TAXIPm-P and 5 µM [^{14}C] labeled α-ketoglutarate. An inward-directed proton gradient was applied. All samples were measured in triplicates, error bars indicate standard deviations. The plots were created using OriginPro 2019b (OriginLab Corporation).

4.10.6 Analysis of the TAXIPm-P signal peptide

As we observed α-ketoglutarate transport by TAXIPm-PQM only *in vivo*, but not *in vitro*, we continued with detailed evaluation of the differences (Figure 68, Figure 73). The TAXIPm-QM membrane domain is located in the membrane *in vivo* and *in vitro*, but the SBP, on the other hand, is located in *E. coli* whole cells in the periplasmic space, which is a confined environment resulting in comparably high concentrations. Some SBPs appear anchored to the membrane through N-terminal lipid moieties or by incorporation of transmembrane helices in the membrane (van der Heide and Poolman 2002). Systems have also been found in which the SBP appear as a direct fusion at the TMDs, either individually or multiple times per TMD (van der Heide and Poolman 2002). In all scenarios, the anchor allows the SBP to be highly concentrated there in a two-dimensional space. We performed a bioinformatic analysis of the SBP TAXIPm-P (section 4.3). Additionally, we expressed the TAXIPm-P protein in *E. coli* MC1061 cells from the pBXC3H plasmid in two versions, one with and the second without its endogenous signal peptide from *Proteus mirabilis*. The *E. coli* cells were lysed, and the proteins were separated to compare their migration in the gel (Figure 30, section 4.4). The difference in electrophoretic mobility suggests that the predicted TAXIPm-P signal peptide is not cleaved, but functions as a membrane-anchor instead (Figure 30). These findings were additionally confirmed by mass spectrometry where in the sample with the soluble SBP a mass of 35527.92 Da was found whereas a mass of 38408.70 Da was found in the case of the membrane purified SBP (Figure 74). When comparing the detected masses to the theoretically calculated molecular weights of 35527.75 Da and 38539.57 Da (Gasteiger et al. 2005), the proteins can be assigned unequivocally. The difference in 0.15 Da for the construct without signal sequence is not significant and thus the

complete unmodified construct was detected. The difference of 130.7 Da in the construct with signal sequence is attributed to a truncation of the N-terminal methionine which reduces the molecular weight by 131.2 Da when theoretically calculated (Gasteiger et al. 2005).

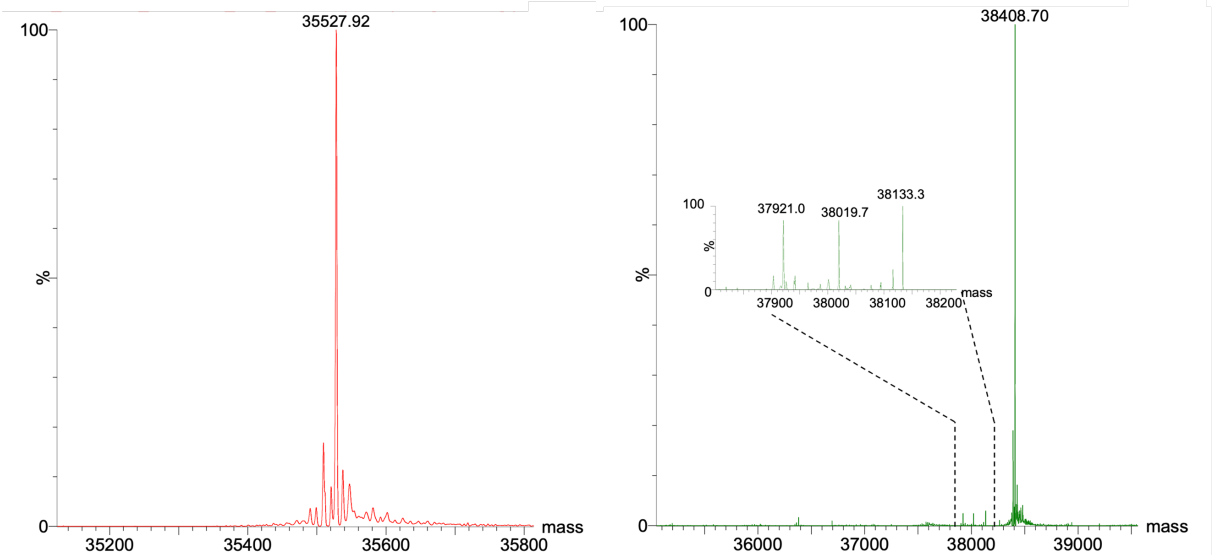


Figure 74 **Intact Mass spectrometry of TAXIPm-P with and without endogenous signal sequence.** Mass spectrometry revealed a mass of 35527.92 Da for the SBP without signal sequence and of 38408.70 Da for the SBP with signal sequence. The SBP without signal sequence is depicted in red and the SBP with signal sequence in green. The measurements were performed by the mass spectrometry facility of the Max Perutz labs, Vienna.

4.10.7 Purification and surface tethering of TAXIPm-P

In vivo, we showed that TAXIPm-P transports α-ketoglutarate, but this has not yet been verified *in vitro* using reconstituted TAXIPm-QM with soluble TAXIPm-P (Figure 73). While most SBP in gram negative bacteria occur in a soluble form in the periplasmic space, some appear surface tethered to the membrane which allows it to move in a two-dimensional space only. For TAXIPm-P we demonstrated that its predicted signal sequence originating from *Proteus mirabilis* is not cleaved in *E. coli*, suggesting it functions to anchor the protein to the membrane instead (Figure 30, section 4.4). Therefore, the full-length SBP TAXIPm-P including the predicted signal sequence was expressed overnight at 25 °C in the pBXC3H plasmid attaching a carboxy-terminal with decaHis tag (Geertsma and Dutzler 2011). The next day, the membrane was prepared and from this, the SBP was purified for the purpose of surface tethering to TAXIPm-QM proteoliposomes. For this, the SBP was solubilized in 1.6 % DDM and purified via IMAC in 0.05 % Triton X-100 with a final elution step by 3C protease. An additional imidazole elution step was performed which showed that 3C protease elutes a purer protein and should be used in the future (Figure 75, A). After the IMAC, the SBP was present in high purity already, which was further increased by a subsequent SEC (Figure 75, B). The size-exclusion chromatogram shows three peaks SEC (Figure 75, B). The first peak resembles the void peak which presumably contains aggregates, DNA and TAXIPm-P (fractions 1 and 2). The second peak contains the largest amount of Triton X-100 solubilized TAXIPm-P (fractions 3 to 6), but the third peak contains small fractions of it as well as can be seen on the corresponding SDS PAGE. Most likely the third peek contains additionally empty Triton X-100 micelles (fractions 7 to

10). After separation and analysis of the eluted target protein TAXIPm-P, a purity of over 90 % was determined (Figure 75, B). It migrates to an apparent size slightly above 35 kDa which is in agreement with its predicted molecular weight of 38 kDa (Gasteiger et al. 2005).

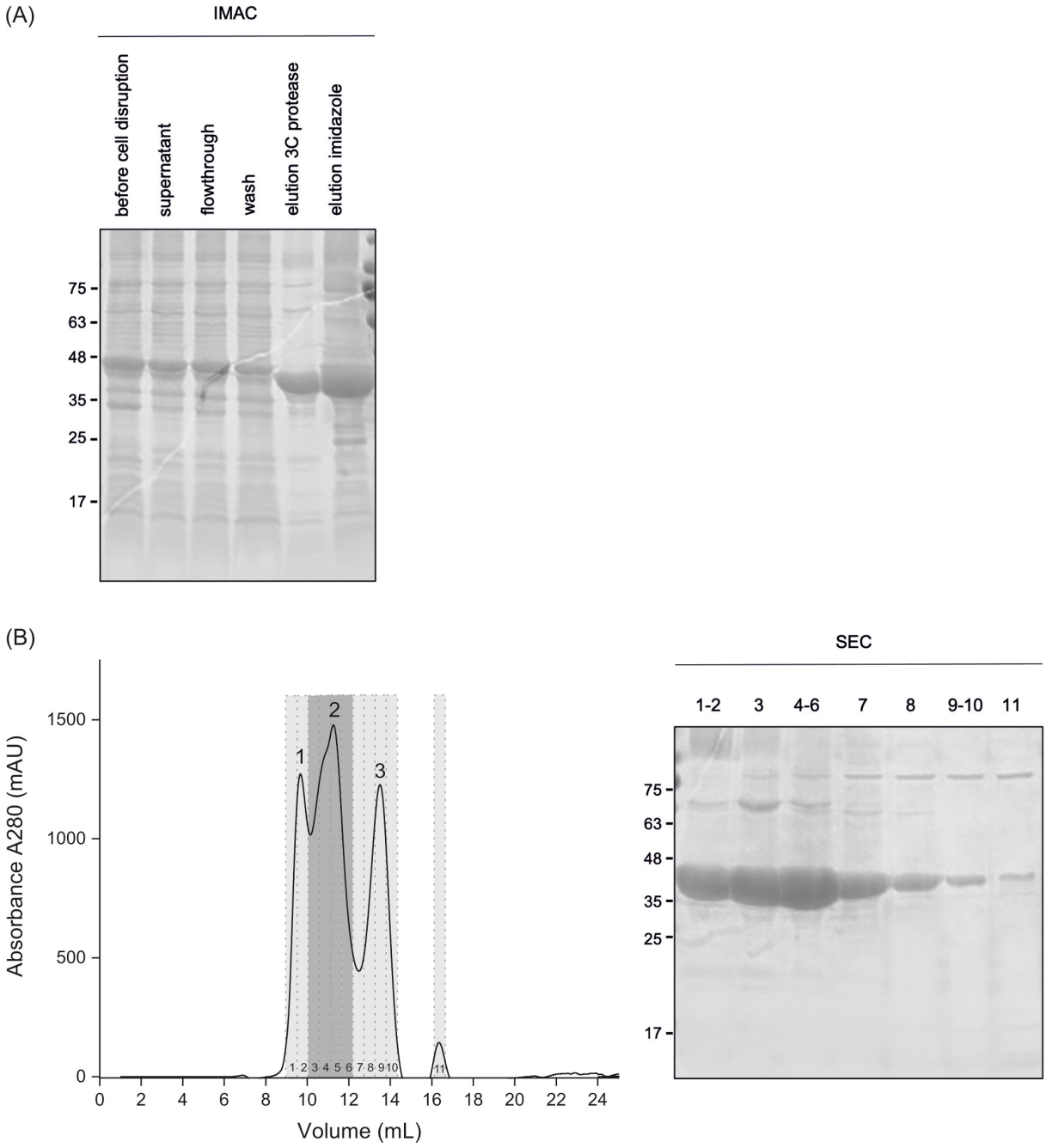


Figure 75 **Purification of TAXIPm-P from membrane by IMAC and SEC in detergent.** (A) DDM-solubilized TAXIPm-P was purified via IMAC in 0.05 % Triton X-100 with a final elution step by 3C protease. (B) The IMAC was followed by SEC on a Superdex 200 30/100 GL Increase column in 20 mM Hepes, 150 mM NaCl, 0.05 % Triton X-100. Individual fractions from the SEC were pooled and the proteins were visualized using a Coomassie stain. The plot was created using OriginPro 2019b (OriginLab Corporation).

While detergent is required to solubilize membrane proteins, high concentrations of detergent will also destabilize lipid membranes. In initial attempts to co-reconstitute full-length TAXIPm-P in

proteoliposomes containing TAXIPm-QM this resulted in the re-solubilization and removal of QM. In addition, destabilization of the liposomes may enable P to be embedded in an inside-out orientation, which may enhance substrate export by QM with a similar inside-out orientation. Since we aim to mimic the *in vivo* occurrence, it should be ensured that the SBP only attaches to the TAXIPm-QM proteoliposomes from the outside, therefore further destabilization should be prevented. For this purpose, we purified the SBP again by IMAC and SEC, but in the absence of detergent in the SEC, assuming that the membrane-anchors would form micelle-like structures and thereby prevent each other from denaturation. Without detergent, TAXIPm-P eluted in the void peak (Figure 76, A). Subsequent analysis of this material following the addition of 0.2 % DM showed an elution profile (Figure 76, B) mimicking the initial detergent-based elution volume (Figure 75, B), which suggest that the aggregation is reversible and primarily based on association of the membrane-anchors.

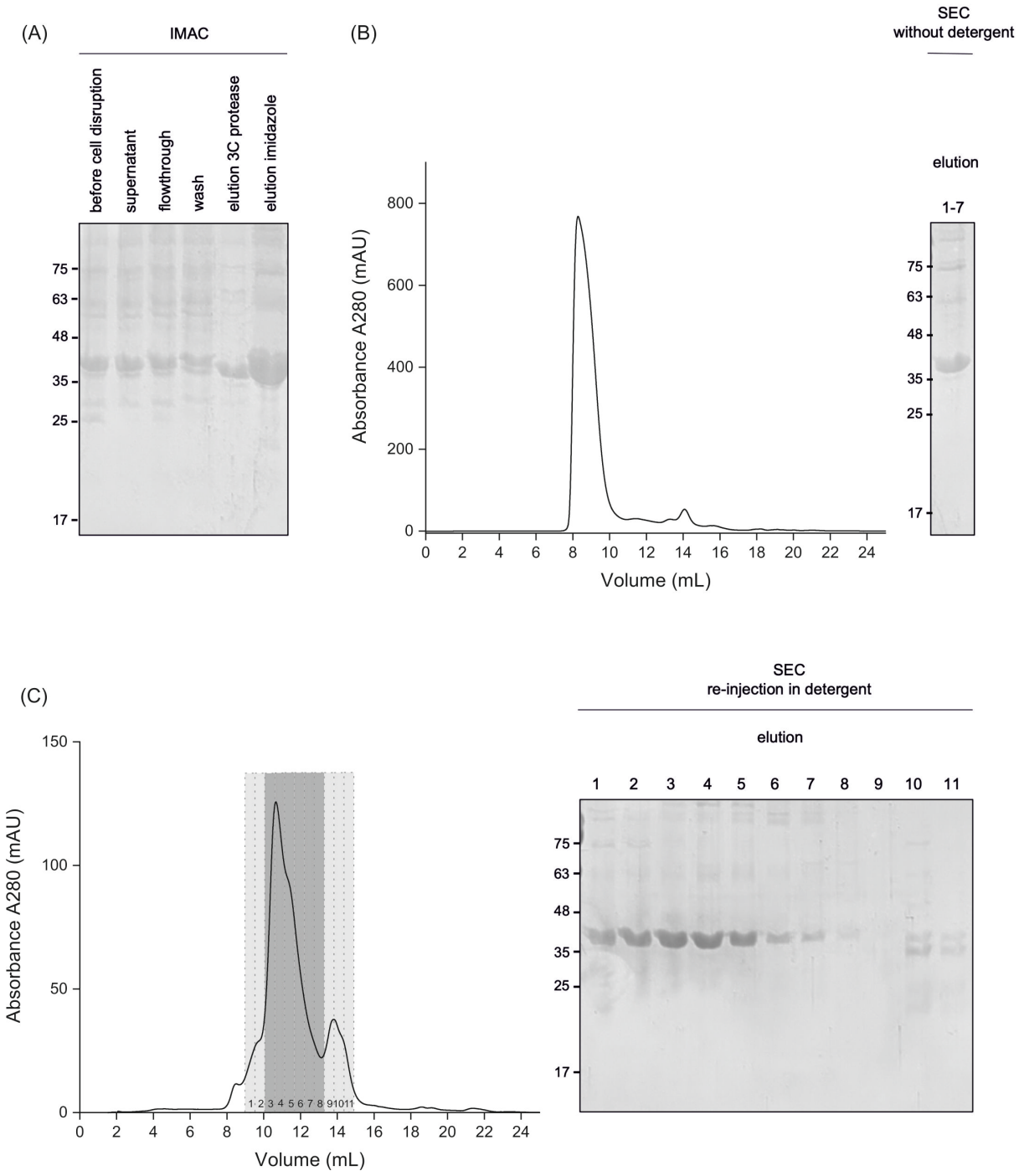


Figure 76 **Purification of TAXIPm-P without detergent and subsequent quality control.** (A) DDM-solubilized TAXIPm-P was purified via IMAC in 0.2 % DM. (B) The IMAC was followed by SEC on a Superdex 200 30/100 GL Increase column in 20 mM Hepes, 150 mM NaCl in absence of detergent. The target protein elutes with the void. (C) The SEC pure TAXIPm-P that eluted in the void was re-solubilized in DM and re-applied on a Superdex 200 30/100 GL Increase column in 20 mM Hepes, 150 mM NaCl, in presence of 0.2 % DM to demonstrate that the TAXIPm-P protein can be recovered from the void by supplying detergent. The plots were created using OriginPro 2019b (OriginLab Corporation).

During the reconstitution of TAXIPm-QM a random distribution towards inside-out and right-side-out occurs. Since it is unexplored whether one direction happens preferentially for TAXIPm-QM, the proteoliposomes were sonicated to assume a 50/50 distribution. The proteoliposomes were flash-frozen and thawed five times for membrane fusion. To generate unilamellar liposomes, they were extruded 11x and then mixed with the membrane-anchor containing TAXIPm-P in a 1:0.1 (TAXIPm-QM:TAXIPm-P) (w/w) ratio. The SBP was previously purified in the absence of detergent in the SEC

to allow surface tethering from the outside only. To verify that TAXIPm-P and TAXIPm-QM both remain in the membrane, we ultracentrifuged two TAXIPm-PQM proteoliposome samples from which one was treated beforehand with 1 % (w/v) DDM (Figure 77, A). As expected, the ultracentrifugation of the sample without detergent treatment verifies the tight embedment in the membrane as the SBP TAXIPm-P was detected in the pellet after ultracentrifugation next to TAXIPm-QM, but not in the supernatant (Figure 77, A). To determine the reconstitution efficiency the TAXIPm-PQM proteoliposomes were additionally re-solubilized. A subsequent separation in a 12 % gel shows high reconstitution efficiencies as the signal intensities from before and after ultracentrifugation remain at the same magnitude (Figure 77, A). The amount of TAXIPm-QM in the gel does not correspond to the amount of α-ketoglutarate importing TAXIPm-QM, since approximately 50 % is oriented inside-out and thus no substrate can be transferred from the TAXIPm-P attached to the outside. Additionally to the 0.1:1 (TAXIPm-QM:TAXIPm-P) (w/w) ratio the proteins were reconstituted in 1:1 and 1:10 ratios to later evaluate the ratio which allows highest transport rates (Figure 77, B).

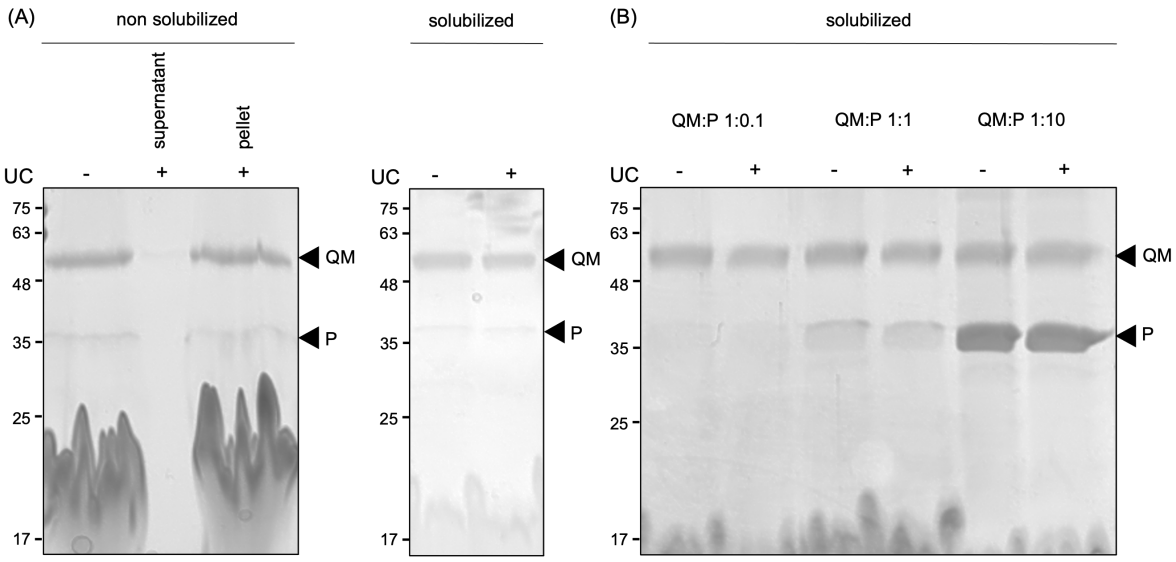


Figure 77 **Reconstitution of TAXIPm-P to TAXIPm-QM proteoliposomes.** TAXIPm-QM was reconstituted in a 1:50 (wt:wt) ratio to soy PC liposomes, flash frozen in liquid nitrogen and stored. (A) Thawed TAXIPm-QM proteoliposomes were extruded and mixed with TAXIPm-P in a 1:0.1 QM:P ratio and incubated for 15 min at room temperature to allow for surface tethering. To evaluate the efficiency the TAXIPm-PQM proteoliposomes, they were either re-solubilized in 1 % (w/v) DDM, followed by ultracentrifugation or subjected directly to ultracentrifugation without detergent treatment to compare the sample before ultracentrifugation with the supernatant and the pellet after ultracentrifugation. (B) Thawed TAXIPm-QM proteoliposomes were mixed with TAXIPm-P in a 1:0.1, 1:1 and 1:10 ratio and incubated for 15 min at room temperature each. To evaluate the efficiencies the TAXIPm-PQM proteoliposomes, they were re-solubilized in 1 % (w/v) DDM, followed by ultracentrifugation to compare the sample before and after ultracentrifugation. Signal intensities were determined using ImageJ (Schneider, Rasband, and Eliceiri 2012).

4.10.8 *In vitro* α-ketoglutarate transport by TAXIPm-PQM

α-ketoglutarate transport by TAXIPm-PQM was successfully determined in pre-energized *E. coli* cells (Figure 68), but could not be verified *in vitro* following a protocol which proved successful for the transport of N-acetylneuraminic acid by TRAPHi-SiaPQM from *Haemophilus influenzae* (Mulligan et al. 2009). The lack of transport into *E. coli* whole cells upon the addition of the protonophore CCCP suggested that the α-ketoglutarate transport by TAXIPm-PQM is coupled to the proton-motive force

(Figure 69, section 4.10.2). Based on the new discoveries that the SBP TAXIPm-P is possibly not cleaved from the signal sequence in *E. coli* and thus embedded in the membrane (section 4.3 and 4.4), we re-examined the *in vitro* uptake using 10 µM [^{14}C] labeled α-ketoglutarate, this time with reconstituted TAXIPm-P (section 4.10.7). For this purpose, TAXIPm-QM and TAXIPm-P were both purified from the membrane by IMAC and SEC in detergent with subsequent reconstitution. First, TAXIPm-QM was reconstituted, extruded and then mixed with TAXIPm-P. As both protein purifications contained detergent, they destabilized the liposomes leading to inside-out and right-side-out orientations for both proteins. In the following transport assay a significant efflux was observed next to import of α-ketoglutarate (Figure 78). This suggests that either the gradients were not maintained, or export is facilitated by inside-out reconstituted TAXIPm-P and TAXIPm-QM.

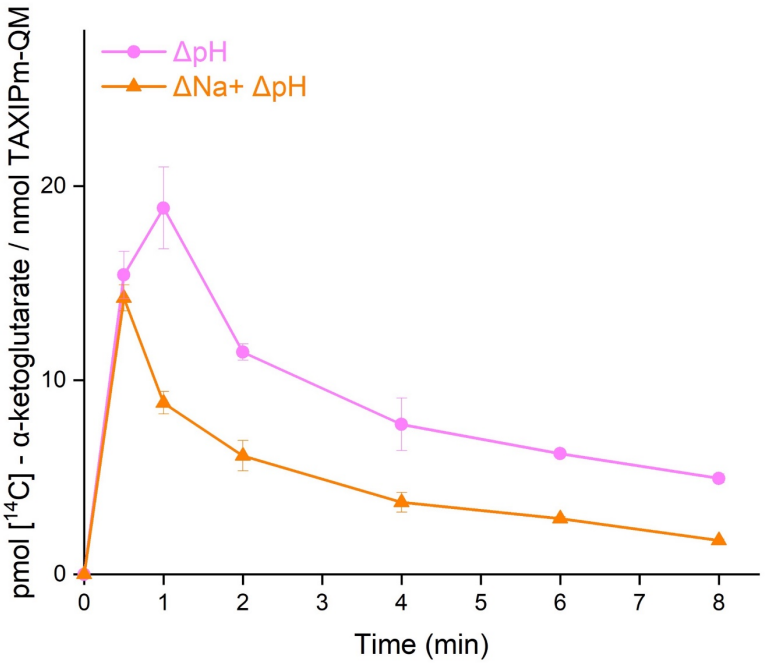


Figure 78 ***In vitro* transport of α-ketoglutarate in TAXIPm-PQM proteoliposomes with TAXIPm-P inside and outside proteoliposomes**. α-ketoglutarate uptake in TAXIPm-PQM proteoliposomes at 20 °C after purification of TAXIPm-QM and TAXIPm-P in detergent, followed by reconstitution. The assay buffer in which the proteoliposomes were diluted contained 10 µM [^{14}C] labeled α-ketoglutarate. An inward-directed proton-gradient and an inward-directed proton-gradient combined with and inward-directed Na^+-gradient were applied. All samples were measured in duplicates, error bars indicate standard deviations. The plot was created using OriginPro 2019b (OriginLab Corporation).

Since the SBP TAXIPm-P endogenously occurs in the periplasmatic space only, we optimized the purification and subsequent reconstitution to prevent the liposomes from being destabilized upon surface tethering of the TAXIPm-P to the TAXIPm-QM proteoliposomes. For this, we purified, reconstituted, and extruded TAXIPm-QM as before. However, when purifying the membrane-anchor containing TAXIPm-P, the detergent was omitted this time (Figure 76), so that no destabilization of the TAXIPm-QM proteoliposomes should occur during the additional reconstitution of TAXIPm-P resulting in TAXIPm-P present outside the TAXIPm-QM proteoliposomes only. The TAXIPm-QM proteoliposomes with surface-tethered TAXIPm-P were then concentrated, and immediately used for

the transport assay without further freezing to prevent fusions of the proteoliposomes. The following α-ketoglutarate uptake was thus increased, and the efflux occurred later (Figure 79). In addition to proton coupling, we tested the additional influence of Na^+-ions. An inward directed Na^+-gradient did not enable α-ketoglutarate transport by TAXIPm-PQM (Figure 79). The combination of an inward directed Na^+-gradient and proton gradient, however, seemed to earlier initiate the efflux than was observed with the inward directed proton gradient alone (Figure 79). One might speculate that Na^+-ions can compete with protons for their binding site but not enable transport themselves. Alternatively, the combination of an inward directed Na^+-gradient and proton gradient could also slightly shift the transport maximum to a slightly earlier time point and therefore result in an earlier starting decrease of transport (Figure 79).

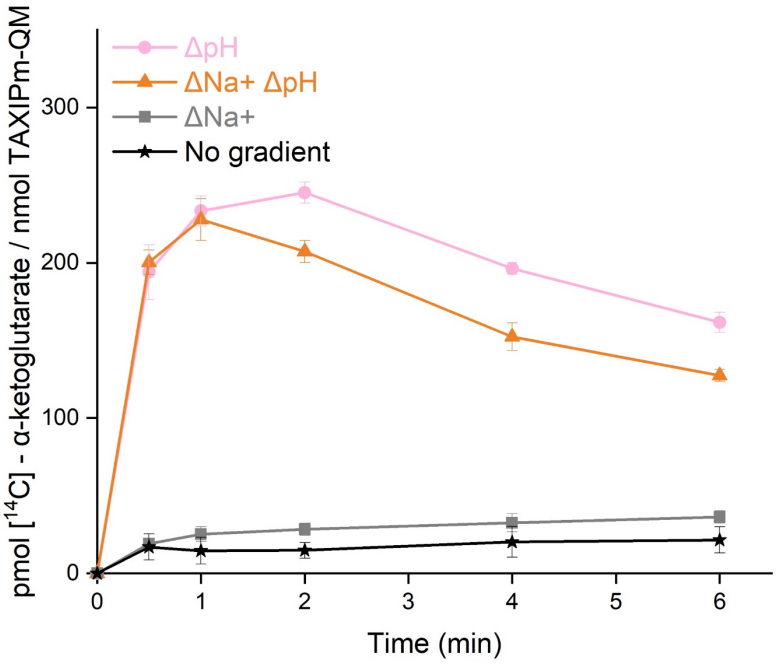


Figure 79 ***In vitro* transport of α-ketoglutarate in TAXIPm-PQM proteoliposomes with external surface tethered TAXIPm-P**. α-ketoglutarate uptake was performed using TAXIPm-PQM proteoliposomes at 20 °C. Prior to the transport assay, TAXIPm-QM was purified in detergent, whereas the membrane-anchor-containing TAXIPm-P was purified without detergent in the SEC. The assay buffer in which the proteoliposomes were diluted contained 10 µM [^{14}C] labeled α-ketoglutarate. An inward-directed proton-gradient and an inward-directed proton-gradient combined with and inward-directed Na^+-gradient were applied. All samples were measured in triplicates, error bars indicate standard deviations. The plot was created using OriginPro 2019b (OriginLab Corporation).

To further evaluate the influence of the membrane potential, we applied a negative membrane potential as potential driving force by supplying valinomycin combined with an inwards-directed Na^+-gradient and additionally an inwards-directed Na^+-gradient only (Figure 80). The negative membrane potential mimics the native potential of most procaryotes (Ramahi and Ruff 2014). Neither sodium coupling, nor the membrane potential enabled α-ketoglutarate transport by TAXIPm-PQM (Figure 80).

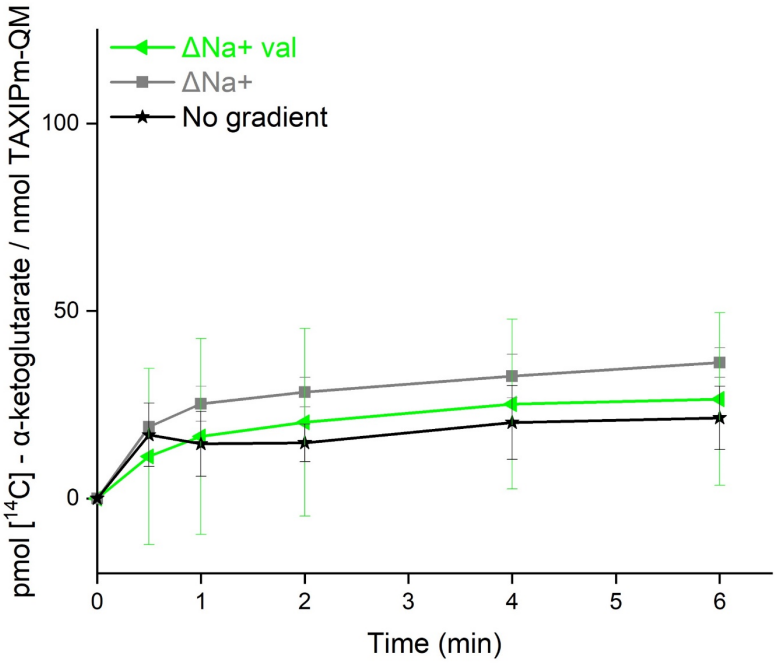


Figure 80 **Screen of potential forces driving α-ketoglutarate uptake by TAXIPm-PQM.** A negative membrane potential combined with an inward-directed Na^+-gradient, and an inward-directed Na^+-gradient alone were applied. The assay was performed at 20 °C, the buffer in which the proteoliposomes were diluted contained 10 µM [^{14}C] labeled α-ketoglutarate. To generate a negative membrane potential 500 nM valinomycin were added. As negative control, no gradient was applied. All transport experiments were performed in triplicates, error bars indicate standard deviations. The plot was created using OriginPro 2019b (OriginLab Corporation).

The *in vitro* transport assays indicated that proton coupling is essential to allow α-ketoglutarate transport by TAXIPm-PQM proteoliposomes. If we combine the inward-directed pH gradient with a negative membrane potential, the result is a slight reduction in α-ketoglutarate uptake (Figure 81). The slight reduction might be negligible as we did not observe a change in the uptake rate when combining the inward-directed pH gradient with a positive membrane potential (Figure 81). This suggests that the combined charge of the substrate and the proton(s) is neutral.

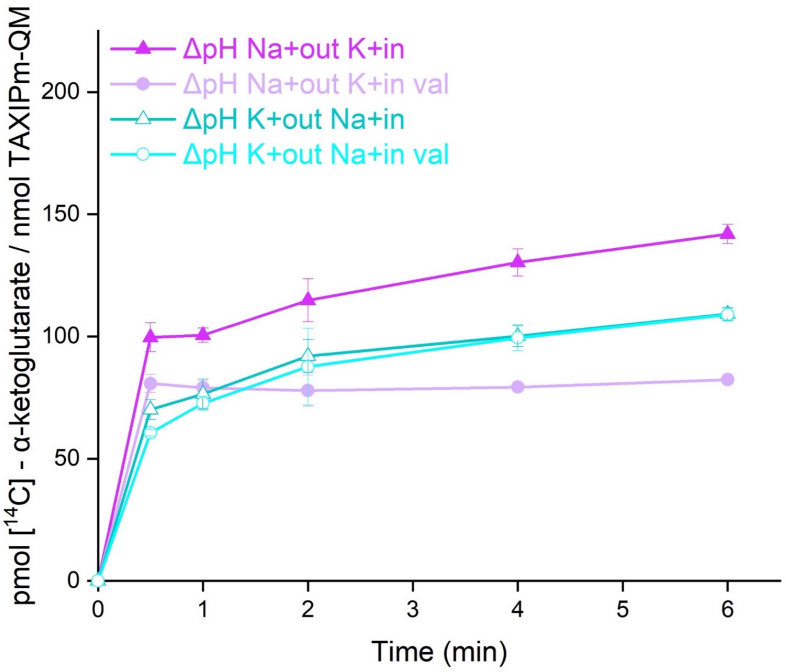


Figure 81 **Dependence of TAXIPm-PQM mediated α-ketoglutarate uptake on membrane potential.** A negative and positive membrane potential was generated in combination with an inward directed proton gradient. Na^+_{out} K^+_{in} val containing 500 nM valinomycin resembles the negative membrane potential, depicted in light purple. The inverted gradient K^+_{out} Na^+_{in} val also containing 500 nM valinomycin resembles the positive membrane potential, depicted in light blue. To demonstrate the effect of the applied membrane potential negative controls were added which contain no valinomycin, depicted in magenta and teal. The assay was performed at 20 °C, the buffer in which the proteoliposomes were diluted contained 10 µM [^{14}C] labeled α-ketoglutarate. All transport experiments were performed in triplicates, error bars indicate standard deviations. The plot was created using OriginPro 2019b (OriginLab Corporation).

Alongside sufficient substrate concentration and a stringent proton coupling, the number of functionally reconstituted TAXIPm-QM and TAXIPm-P molecules seems essential. The ratio of transporter, SBP and substrate in relation as well most likely influences the transport reaction. SBPs without attached substrate often bind to their respective transporter and thus block the binding site for substrate-bound SBPs. This leads to uptake combined with immediate efflux, initiated by the SBP as it might enable efflux when bound to the transporter in an unliganded form (Mulligan et al. 2009). To examine the influence of the number of TAXIPm-P molecules on the transport behavior of TAXIPm-PQM, we reconstituted TAXIPm-QM and TAXI-Pm-P in different ratios while keeping the α-ketoglutarate concentration at 10 µM. For this purpose, we adjusted the theoretical QM:P ratios to 1:0.1, 1:1 and 1:10 which approximately corresponds to 0.5 µM right-side-out oriented from 1 µM total TAXIPm-QM and varying TAXIPm-P concentrations. If supplied as soluble SBP, concentrations of TAXIPm-P would result in 0.1, 1 and 10 µM TAXIPm-P. As we reconstituted the SBP, however, the local TAXIPm-P concentrations on the membrane correspond to approximately 42 µM, 420 µM and 4.2 mM (section 5.3). As an increasing amount of TAXIPm-P resulted in an earlier starting efflux, we assume that next to liganded TAXIPm-P, unliganded TAXIPm-P binds to TAXIPM-QM as well (Figure 82).

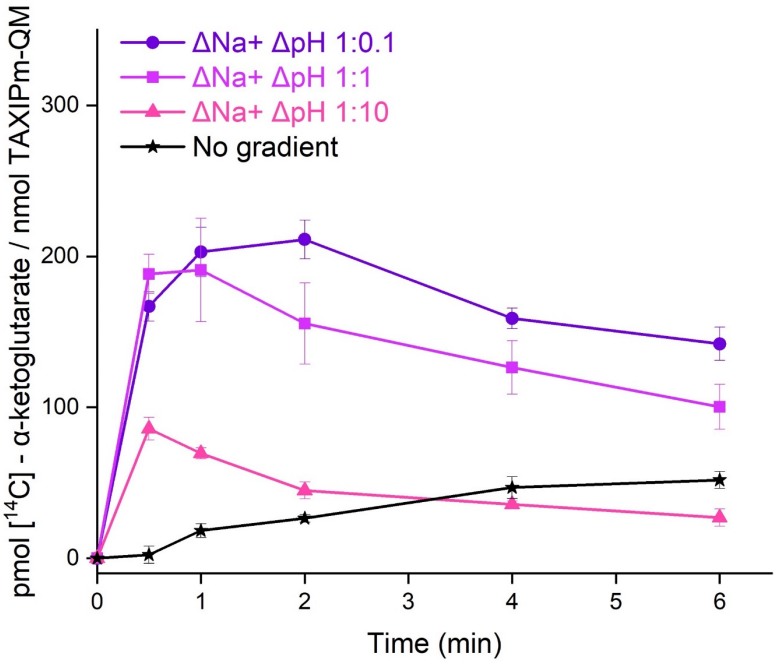


Figure 82 **Dependence of α-ketoglutarate transport in TAXIPm-PQM proteoliposomes on the TAXIPm-P concentration.** α-ketoglutarate uptake in TAXIPm-PQM proteoliposomes at 20 °C supplying 10 µM [^{14}C] labeled α-ketoglutarate. The concentration of the SBP TAXIPm-P was varied during reconstitution resulting in the theoretical QM:P ratios of 1:0.1, 1:1 and 1:10. The concentration of functional TAXIPm-QM is overestimated as it is not only right-side-out, but also inside-out oriented. An inward-directed proton-gradient was applied. As negative control, no gradient was applied. All samples were measured in triplicates, error bars indicate standard deviations. The plot was created using OriginPro 2019b (OriginLab Corporation).

It was shown for the TRAP SiaPQM from *Haemophilus influenzae* that it imports N-acetylneuraminic acid unidirectionally. This cannot be reversed to export by the addition of high substrate concentrations as it is the case for conventional secondary transporter. Instead, the uptake of N-acetylneuraminic acid by the TRAP SiaPQM stagnates when 1 mM substrate is provided (Mulligan et al. 2009). To investigate this for TAXIPm-PQM, we tested to what extent an excess of α-ketoglutarate influences its transport behavior. Again 10 µM [^{14}C] labeled α-ketoglutarate and an inward directed proton gradient were applied. At 90 sec, either 1 mM unlabeled α-ketoglutarate or the same amount of dH_2O was added. While the concentration of the SBP TAXIPm-P had a direct effect on α-ketoglutarate transport (Figure 82), the multiple abundance of substrate appears to have no effect (Figure 83).

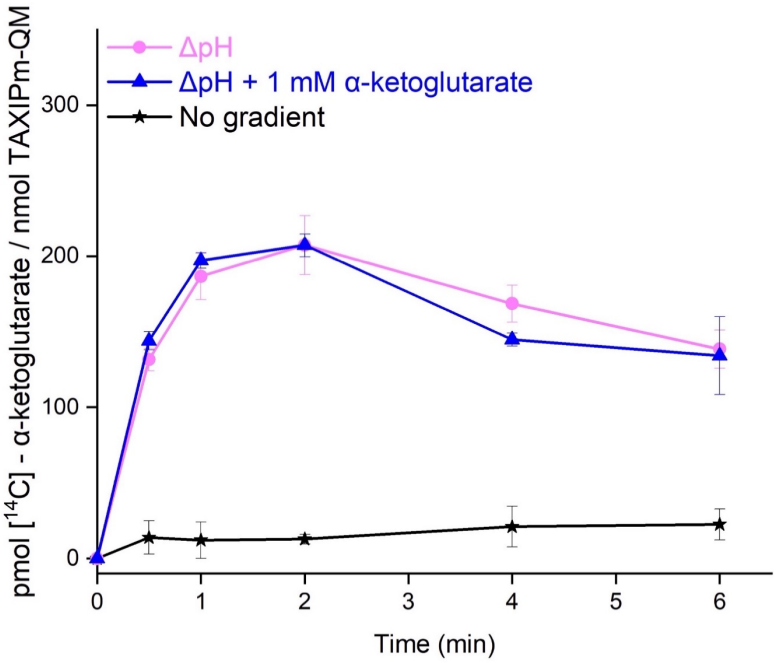


Figure 83 **Effect of α-ketoglutarate excess on TAXIPm-PQM mediated transport.** The assay was performed at 20 °C, the buffer in which the proteoliposomes were diluted contained 10 µM [^{14}C] labeled α-ketoglutarate. An inward directed proton gradient was applied. At 90 sec, either 1 mM unlabeled α-ketoglutarate or the same amount of dH_2O was added. As negative control, no gradient was applied. All transport experiments were performed in triplicates, error bars indicate standard deviations. The plot was created using OriginPro 2019b (OriginLab Corporation).

4.11 Screening of TAXIPm-QM conformation stabilizing nanobodies

After the general functional characterization of TAXIPm-PQM, structural characterization of the entire complex and in different conformations would help to further investigate the substrate and proton coordination and the interaction between the SBP and membrane domain. Although it is becoming increasingly straightforward to obtain initial structural statements about proteins using Alphafold, it remains a challenge to obtain entire complexes and different conformations to be able to draw functional conclusions. To initialize the work on this, nanobodies against the membrane domain of TAXIPm-QM were generated by immunization of alpacas and selected by observing binding of the nanobodies and inhibition of transport using ELISA and *E. coli* whole cell uptake assays. The selected TAXIPm-QM specific nanobodies should then aid to determine structures at the MPI CBG, Dresden (Eric Geertsma Group, www.mpi-cbg.de).

4.11.1 Immunization of two alpacas

In collaboration with the Nanobody Service facility from the University of Zurich two alpacas - Adonis and Blitz - were immunized (www.nsf.uzh.ch). The diversity of the resulting heavy chain only antibodies was increased by using immune responses from two different animals. For this purpose, detergent-solubilized TAXIPm-QM protein was purified by IMAC (Figure 84, A) and SEC (Figure 84, B) after which the protein from both peaks was mixed. During the purification the detergent n-decyl-β-D-maltoside (DM) was used as the unmodified wildtype TAXIPm-QM protein showed good stability and high functionality in this detergent and it leads to a smaller micelle when compared to n-dodecyl-

β-D-maltoside (DDM) which is more commonly used for the purification of membrane proteins (Bamber et al. 2006). Over a period of six weeks protein was injected once a week combined with an adjuvant to boost the immune response. To examine the protein quality after storage at -80 °C, a test sample was thawed and analyzed in a SEC and SDS PAGE (Figure 84, C, D). No changes were detected in the protein, so it was concluded that functional TAXIPm-QM protein was used for immunization.

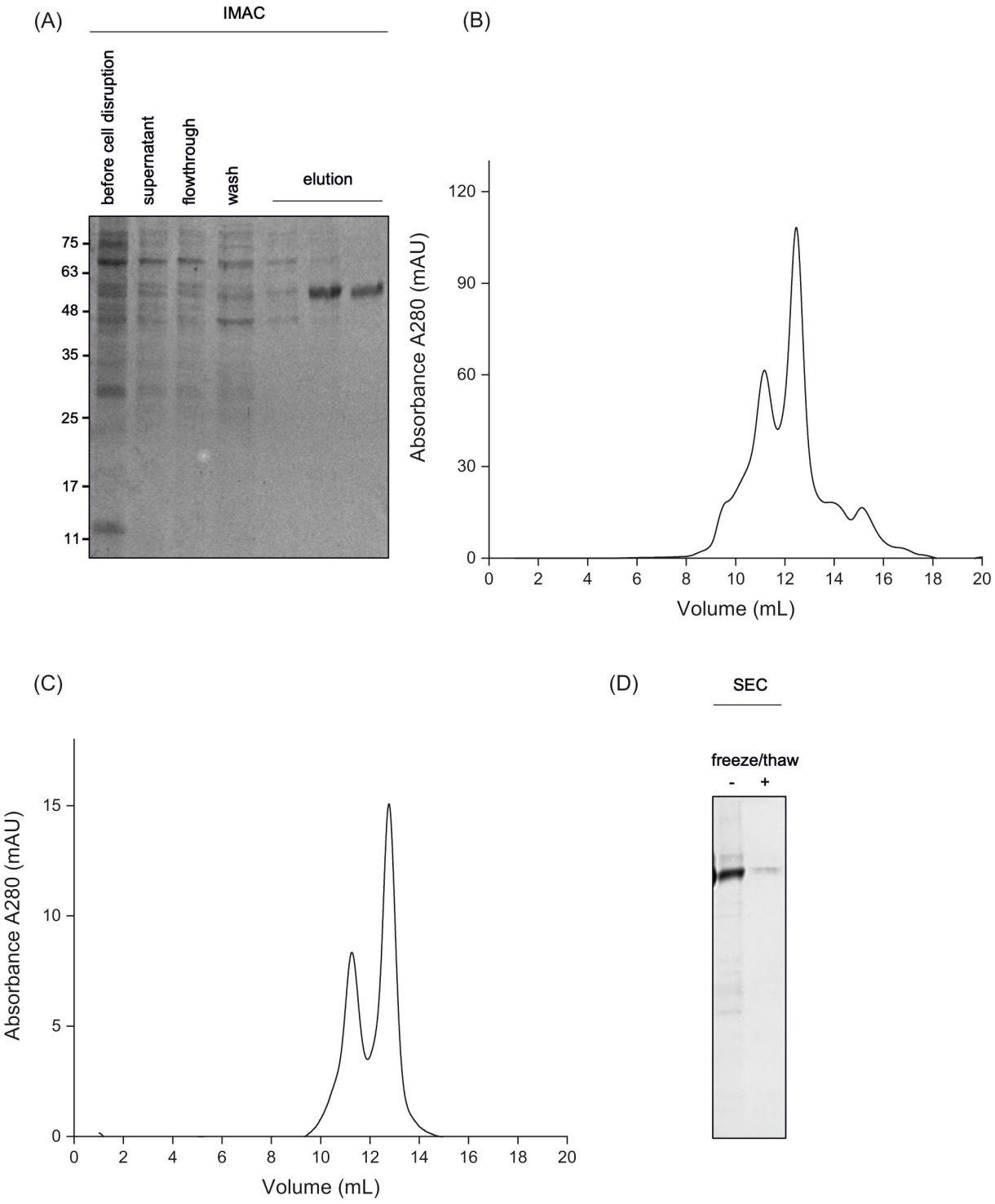


Figure 84 **Purification and quality control of TAXIPm-QM after storage at – 80 °C and thawing for immunization.** (A) DDM-solubilized TAXIPm-QM was purified via IMAC in a purification buffer containing DM. The protein samples from the purification steps were separated in a 12 % SDS PAGE. (B) The IMAC pure TAXIPm-QM protein was further purified by SEC on a Superdex 200 30/100 GL Increase column in 20 mM Hepes and 150 mM NaCl at pH 7.5 after which the protein from both peaks was mixed, supplemented with 10 % glycerol and frozen at -80 °C. (C) The protein previously stored at - 80 °C was re-applied on a Superdex 200 30/100 GL Increase column in the same buffer to verify the presence of well folded protein. Protein from both peaks was again mixed. (D) Protein samples from the two SEC runs, from before **(-)** and after **(+)** freezing were separated in a 12 % SDS PAGE and visualized using a Coomassie stain. The plots were created using OriginPro 2019b (OriginLab Corporation).

The immune response was quantified after performing the ELISA with secondary antibodies specific to IgG1a, IgG1b, IgG2, IgG3a and IgG3b. The titers of the conventional antibodies, abbreviated as

IgG1A and IgG1b, were only examined for comparison; they were not further processed. With increasing number of injections, the immune response increased indicating that an immune response was generated against the injected antigens (Figure 85). In contrast to the antibody IgG1a, the heavy chain only antibody yield was quantitatively lower in both alpacas, but sufficient for enrichment in subsequent steps (section 4.11.2).

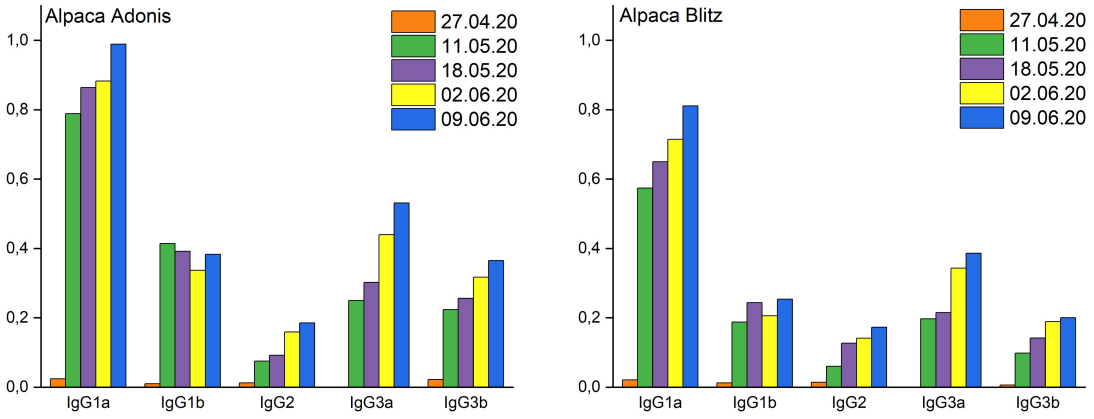


Figure 85 **Immune response of two alpacas during immunization with TAXIPm-QM.** In collaboration with the Nanobody Service facility from the University of Zurich two alpacas - Adonis and Blitz - were immunized (www.nsf.uzh.ch). The immune responses of Adonis and Blitz were evaluated by an ELISA. The immunization was conducted over a period of six weeks with protein a total of four injections (on 27.04.20, 11.05.20, 18.05.20, and 02.06.20). The order of the bars corresponds to the time sequence. The plots were created using OriginPro 2019b (OriginLab Corporation).

The following preparation of the two immune libraries from the blood samples was performed by Dr. Eric Geertsma at the MPI CBG in Dresden (Eric Geertsma Group, www.mpi-cbg.de) whereby the RNA of the heavy chain only antibodies was purified, followed by synthesis of the corresponding cDNA. From this the nanobody open reading frames were obtained by PCR and cloned into phage display vectors for enrichment.

4.11.2 Enrichment of TAXIPm-QM specific nanobodies by Phage display

For phage display based nanobody enrichment and selection, phages presenting nanobodies were applied on a plate which contained the target antigen TAXIPm-QM. Those phages that recognized the antigen were selected for transformation of *E. coli* TG1 cells and subsequent amplification. The bacteria were transformed in a 10-fold dilution series and at the same time the phages selected in round one were used for a second round of selection, which should lead to an enrichment of phages carrying the nanobody gene sequence specific for the TAXIPm-QM protein. During the evaluation, higher growth was observed in the samples containing the target antigen TAXIPm-QM compared to the samples with the arbitrarily selected antigen EctI, which thus showed non-specific enrichment. Furthermore, in round two, a distinct enrichment of the phages carrying the nanobody gene sequence specific for the TAXIPm-QM protein was detected from both alpacas, Adonis and Blitz, yet to a higher level for Adonis (Figure 86). The phage display resulted in four libraries to be used in the further selection process.

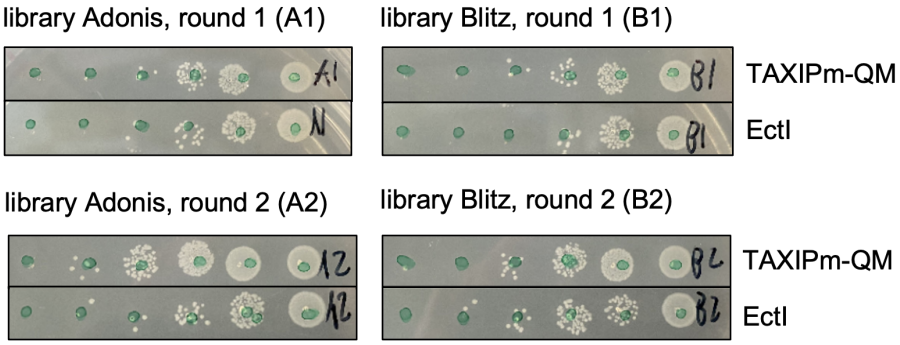


Figure 86 **Phage display to select for TAXIPm-QM specific nanobodies.** Two rounds of phage display were performed resulting in four libraries A1, B1, A2 and B2 of which A1 and A2 were derived from the alpaca Adonis and B1 and B2 from Blitz. The protein EctI was used as negative control to display unspecific binding, kindly provided by Prof. Inga Hänelt, Goethe Universität Frankfurt (Institute of Biochemistry, www.biochem.uni-frankfurt.de). The phage display selections were performed at the MPI CBG Dresden by Dr. Eric Geertsma (Eric Geertsma Group, www.mpi-cbg.de).

4.11.3 Production of biotinylated TAXIPm-QM for selections by ELISA

For the selection process of TAXIPm-QM specific nanobodies by the enzyme-linked immunosorbent assay, TAXIPm-QM was expressed and purified with an Avi-tag fusion specific for BirA-based *in vitro* biotinylation. Since a high protein quality provides the foundation for efficient binder selections, this was verified in several steps. SDS PAGE and SEC served well for this purpose, also to evaluate where to attach the Avi-tag while preserving the function of the protein. The expression and purification of both, the C-terminal and N-terminal Avi-tagged TAXIPm-QM resulted in the correct constructs, determined by evaluation of elution volumes in SEC profiles and size in SDS PAGE (Figure 87 and Figure 88). The SDS PAGE and SEC comparing unbiotinylated and biotinylated samples of the C-terminal and the N-terminal Avi-tagged protein show the correct size of the proteins and elution volumes (Figure 87 and Figure 88) similar to the wildtype TAXIPm-QM protein (Figure 71, section 4.10.4). In both versions, the biotinylation had no effect on the stability of the proteins (Figure 87 and Figure 88).

The *in vitro* biotinylation of the C-terminal Avi-tagged protein showed an efficiency of 70% after quantification in the SDS gel by the streptavidin induced mobility assay (Figure 87) (B. T. Kuhn et al. 2020).

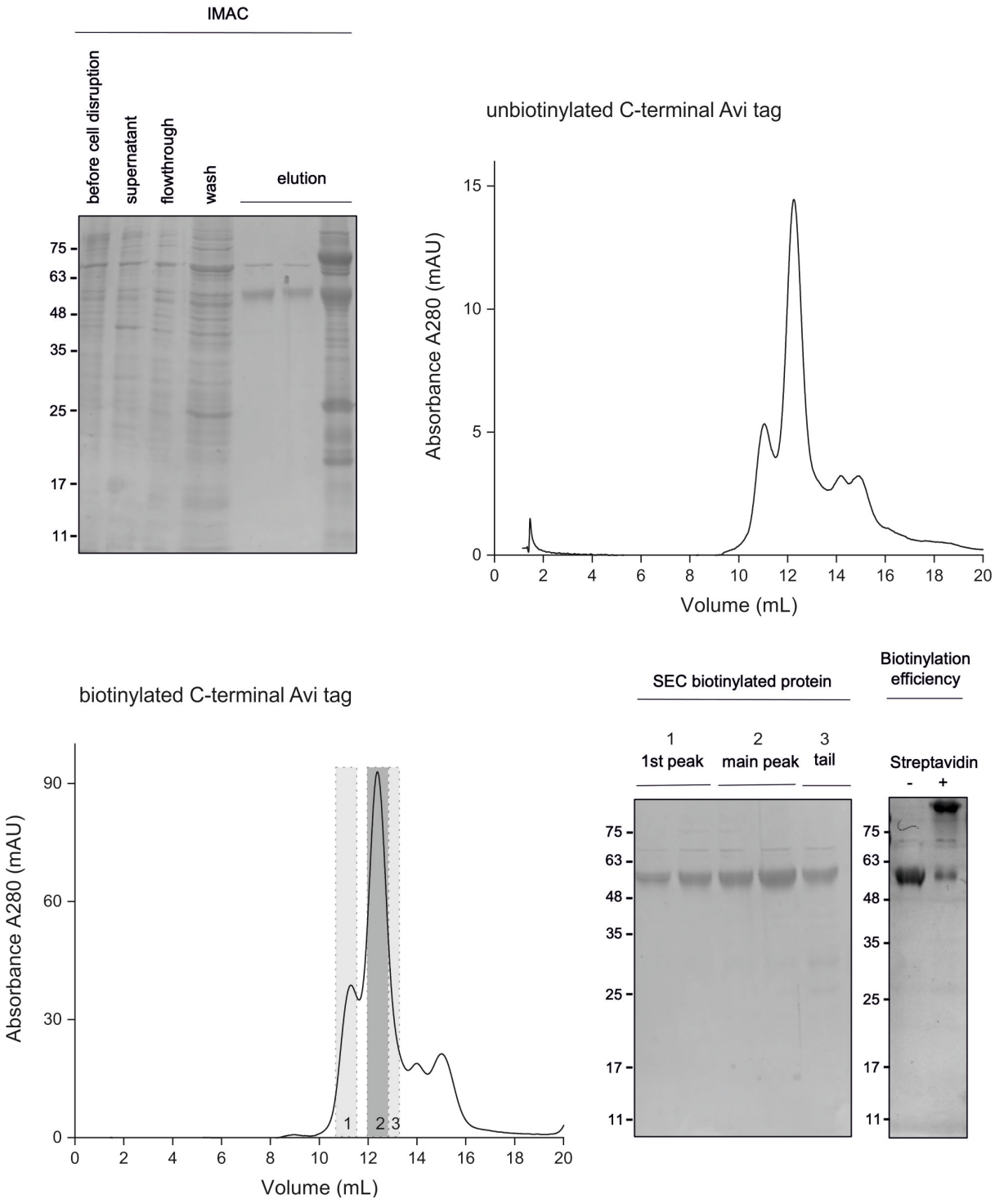


Figure 87 **Purification and biotinylation of TAXIPm-QM with C-terminal Avi tag fusion.** (A) DDM-solubilized TAXIPm-QM was purified via IMAC in a purification buffer containing DM. The protein samples from the purification steps were separated in a 12 % SDS gel. (B) The IMAC pure TAXIPm-QM protein with an Avi tag fusion was further purified by SEC on a Superdex 200 30/100 GL Increase column after which a BirA based *in vitro* biotinylation was performed. (C) The biotinylated TAXIPm-QM protein was re-applied on a Superdex 200 30/100 GL Increase column to ensure a good protein quality. (D) Protein samples from the peak fractions were separated in a 12 % SDS gel and visualized using a Coomassie stain. The biotinylation efficiency was determined by comparing signal intensities of the protein in a 12 % SDS gel with (+) and without (-) supplementing streptavidin. The analysis was performed using ImageJ (Schneider, Rasband, and Eliceiri 2012). The plots were created using OriginPro 2019b (OriginLab Corporation).

The efficiency of the *in vitro* biotinylation of the N-terminal Avi-tagged protein was also quantified in the SDS PAGE by the streptavidin induced mobility assay and resulted in 64 % (Figure 88) (B. T. Kuhn et al. 2020).

Overall, the SEC profiles demonstrated favorable results for the C-terminal Avi-tag (Figure 87) which was therefore used in the following panning and later evalulation via ELISA (section 4.11.4). Also, the biotinylation efficiency was higher for the C-terminal Avi-tagged protein when compared to the N-terminal Avi-tagged protein in a streptavidin-induced mobility shift assay (Figure 87 and Figure 88) (B. T. Kuhn et al. 2020).

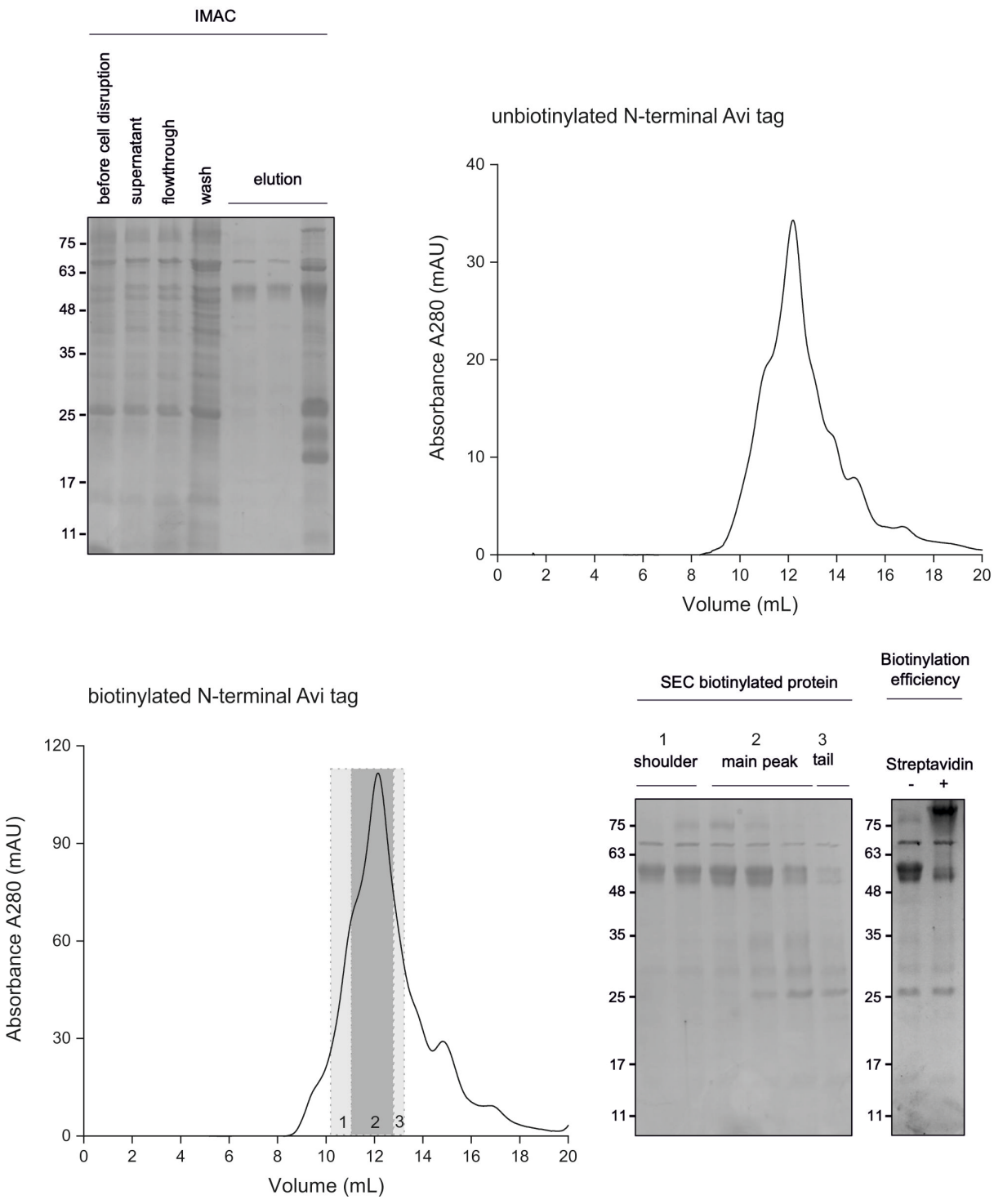


Figure 88 **Purification and biotinylation of TAXIPm-QM with N-terminal Avi tag fusion.** (A) DDM-solubilized TAXIPm-QM was purified via IMAC in a purification buffer containing DM. The protein samples from the purification steps were separated in a 12 % SDS gel. (B) The IMAC pure TAXIPm-QM protein with an Avi tag fusion was further purified by SEC on a Superdex 200 30/100 GL Increase column after which a BirA based *in vitro* biotinylation was performed. (C) The biotinylated TAXIPm-QM protein was re-applied on a Superdex 200 30/100 GL Increase column to ensure a good protein quality. (D) Protein samples from the peak fractions were separated in a 12 % SDS gel and visualized using a Coomassie stain. The biotinylation efficiency was determined by comparing signal intensities of the protein in a 12 % SDS gel with (+) and without (-) supplementing streptavidin. The analysis was performed using ImageJ (Schneider, Rasband, and Eliceiri 2012). The plots were created using OriginPro 2019b (OriginLab Corporation).

4.11.4 Screening for TAXIPm-QM specific nanobodies by ELISA

The selection of nanobodies against TAXIPm-QM was performed using ELISA. For this purpose, 95 clones were picked from each phage display library and expressed in 96 well plates resulting in four plates. Specific binding was detected for 163 samples that showed at least 1.3 fold higher signal when compared to the background samples from the plates without the target protein (Figure 89). Additionally, the nanobodies were tested against the uracil transporter UraA to further exclude nonspecific binding nanobodies. Preference was later given to the nanobodies that exceeded both signal thresholds, the one from the background of the plate and the one from nonspecific binding to UraA (section 4.11.5).

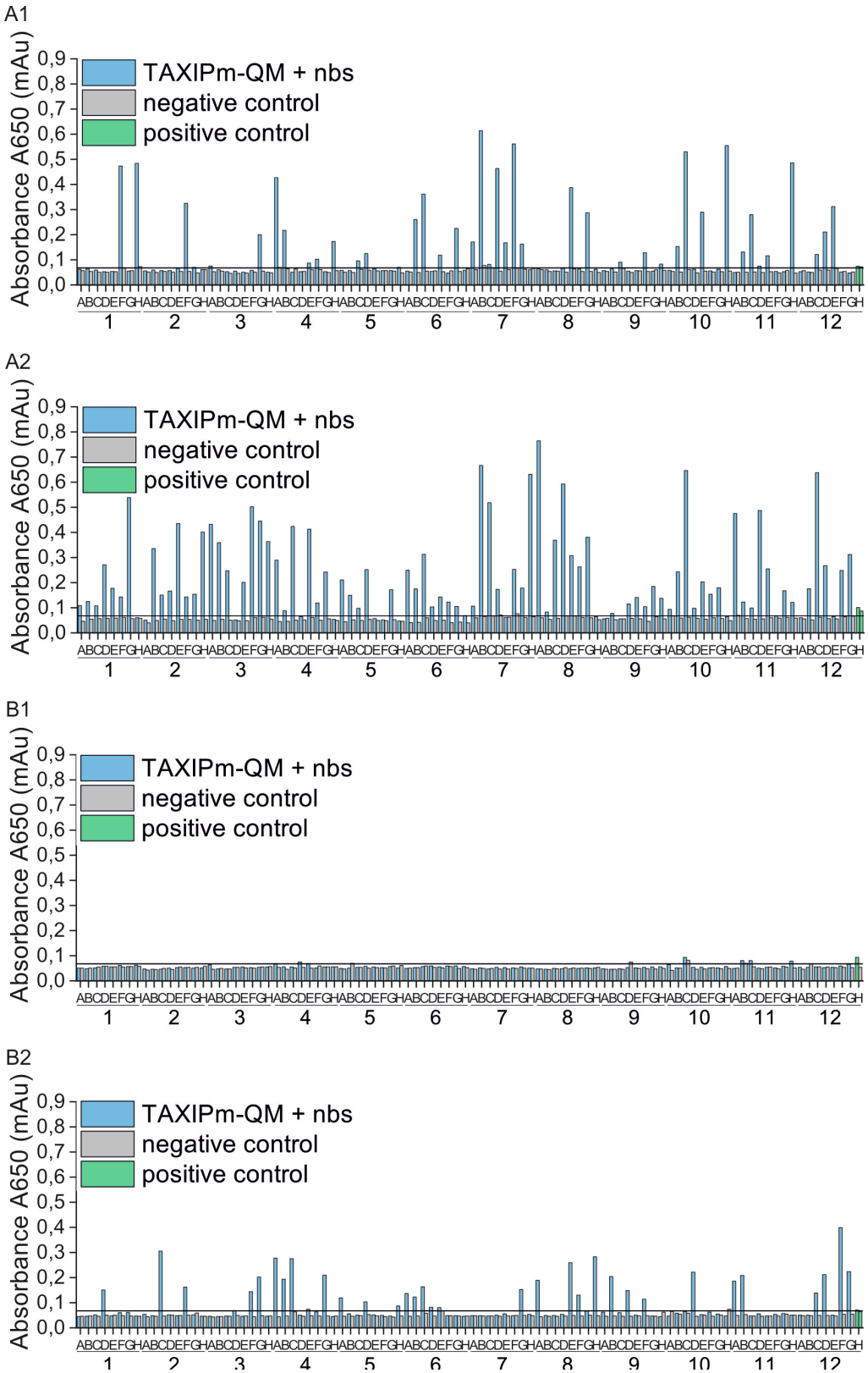


Figure 89 **Screening for TAXIPm-QM specific nanobodies by ELISA.** To identify TAXIPm-QM specific nanobodies, the absorbance was measured at 650 nm for the libraries A1 and A2 from Adonis and B1 and B2 from Blitz. Blue bars represent the nanobodies directed against their antigen TAXIPm-QM, grey bars show unspecific binding against UraA or the plate and as comparison green bars represent Sy45 directed against its target antigen UraA. The threshold line was set to define significant results with 1.3 fold higher signal than the background from the plate. The membrane protein UraA and sybody Sy45 were kindly provided by Dr. Benedikt Kuhn, MPI CBG Dresden (Eric Geertsma Group, www.mpi-cbg.de). The ELISA was carried out under my supervision by Christiane Ruse, Goethe University Frankfurt. The plots were created using OriginPro 2019b (OriginLab Corporation).

However, within the specific binding nanobodies, it is important to select for a broad variety to find candidates that bind different epitopes. Differently high signals can result from different affinities, but also from different expression levels. Higher signals are more likely to represent nanobodies specifically directed against TAXIPm-QM, but those with slightly lower signals may also be relevant as they might stabilize different conformations of TAXIPm-QM. We therefore analyzed all ELISA positive nanobodies which showed at least a 1.3 fold higher signal than the background (Figure 89). In the analysis, we especially focused on the three CDR regions as they are essential for the binding of different epitopes. From protein sequence alignments we identified 39 nanobodies that comprised the typical size of a nanobody and distinct CDR regions. A comparison of the sequences gives a strong indication that the selected nanobodies are highly heterogeneous, with a high diversity in the CDR regions (Figure 90) thereby likely to bind different epitopes. In total, 23 families were identified that exhibit different amino acid residues in the CDR3 regions (Figure 90).

```
A1A7       TASGTDFSVFA--MGWVATVF---SGSSTRYANRVGAPVSW--------------
B2A4       AASGSAFSIRS--LGLVAEISSSISGGTTNYARDLLSG-SW--------------
B2A11      AASGRTFSSDT--MALVAHIT---SGGSTNYAAGRRV----------IGTNYVYW
A2F11      AASINIFSINA--MGMVVGIA---SDGSTNYANALFRDSGS--------GYRSYW
B1D4       TASGSSFIIND--MGLVAAIT---SGGSTNYVAAGIGLPGWTEVAQAMGDPDGYW
B2H5       AASGSFFSIND--MGLVASIT---SRGSTNYANAKIQRVS------DPWREYDYW
B2E6       AASGSISLITA--MGLVATIT---NGGSTSYANAPTFR-------------DDYW
B2G12      AASGSIFLITA--MGLVATIT---NGGSTSYANAPTFR-------------DDYW
B2G4       AASGSIFLITA--MGLVATIT---NGGSTSYANAPTFR-------------DDYW
B2G7       AASGSIFLITA--MGLVATIT---SGGSTNYANAPTWR-------------DDYW
A2C1       AASGSIFLIQA--MGLVATIT---NGGSTNYANAPTWR-------------DDYW
B2D6       AASGFTFDDYA--MSWLSTIS--WNGGTITYACSRPPSF--------NYYAMDYW
A2D2       TVSGRTSSEYR--MGFVAAIT--WHSNIY-YTAADPDRS----NYFLRDLNYYYW
A2A4       SASGRTFSTLA--VAFVATVG--LSAGSSDFAATTYASR------ARTAEEYRNW
B2B11      ASSERTFSTYAMTMGFVAGIG--RFSDNTNYAAAAVPL----------TTTYKYW
A2G1       APSGRTNSPFP--TGFVAVVD--WSGSRTYYAAARRIIS---SAPTTSPDDYAYW
A2G3       ATSERTFSTYT--MGFVAAIN--YNGDSTNYVAVGPQYG---SAYHRAETAYAYW
A2E3       VDSGRTFSTYA--MGFVAAIK--YDGGSTSYVAVGPRYG---SAYHRAETAYDYW
B2C12      AASDRTFSTYV--MAFVAAIN--YNGGSTNYEAVGPRYG---SAYHRGETAYGYW
A1F2       AASERTFSTYT--MGFVAAIN--YNGASTDYVAVGPRYG---SAYHRGKTAYDYW
A2B2       AASERTFSTYT--MGFVAAIN--YNGASTDYVAVGPRYG---SAYHRGKTAYDYW
A1C6       ADTERTFSTYT--MGFVAAIN--YNGGSTSYVAVGPRYG---SAYHRGETAYAYW
A1C11      TGSGRTFSTYA--MGFVAAIN--YNGGSTSYVAVGPRYG---SAYHRGERAYDYW
A2G12      AASERTFSTYA--MGFVAAIN--YNGGSTSYVAVGPRYG---SAYHRGETAYDYW
A1C9       AASGSIFSSYA--MGFVATIS--RSGKNTYYSASRAY------FLNSSDREYDYW
A1F7       AASGRGFSDHA--MGIVARVS--WSGGSTYYAAAEWDSDYVGG----------YW
A1D7       ATSERTFNSYA--VGFVAAIS--WSGGSTYYAATTASALMYSDYVALETDEYEYW
A2E10      AASERTFSSYA--MGFVAAIS--WSGGSTYYAATTASALMYSDYVALETDEYEYW
A2F10      SASRRTLSSDA--VGFVAGIT--QSGGSTYYGAAQVQYRER-LSDGTNWAAYHYW
A2D5       AASGRTFRGAA--LAFVAVIS--EDGGSTLYSAGNVITN-----RGTKPQGYDYW
A2C6       AASERAFTYYT--KGFVATIN--WNGESTYYAAADSSAL---WLDTKRATRYDYW
B2D9       AASGRTFRSYD--MAFVASIT--WRGRLTDYAAATSS---------YLARSYDYW
A1B10      TASGRTFSGYA--MGFVAVIS--DDGGFTNYAAANGVSS------STNPASYWYW
A2E11      AASGRTFSSYA--MGFVAAIG--ESGDITYYAAANRVYP------SSNSVGYDYW
A1B4       ADSGRTFSSYA--MGFVAAI-----SGAYTVYAAANSVAA------SMRAGAYDYW
A2D7       VASGLSFDEYA--VGFVAAIA--ESSGYTVYAGANSVAA------STRAGAYDYW
A2B3       AASGRTFSSYA--MGFVAAMA--DSGSYTVYAAANSVAA------SIRAGAYDYW
A2C2       AASGRTFSSYA--MGFVAAIA--ESSGYTVYAGANPVAA------STRAGAYDYW
A2E4       AASGRTFSSYA--MGFVAAIA--ESSGYTVYAGANSVAA------STRAGAYDYW
             :          . :  :         :                    . .
```

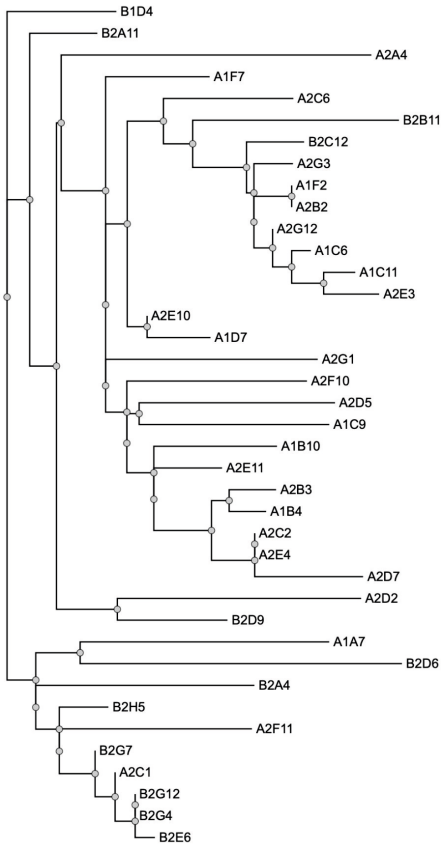


Figure 90 **Alignment of CDR regions from 39 unique TAXIPm-QM specific nanobodies**. The three CDR regions of 39 TAXIPm-QM specific nanobodies were aligned using Clustal Ω. Amino acids present in all sequences are labelled with double dots and amino acids present in most sequences are labelled with one dot, additionally both are highlighted in yellow (Sievers et al. 2011). The evaluation of the data was supported by Dr. Eric Geertsma, MPI CBG Dresden (Eric Geertsma Group, www.mpi-cbg.de). The diversity of the selected nanobodies was visualized in a phylogenetic tree using NGPhylogeny (Lemoine et al. 2019).

4.11.5 Co-expression of TAXIPm-QM and selected nanobodies

To test whether the nanobodies can be co-expressed with the transporter TAXIPm-QM in *E. coli* whole cells and to evaluate whether expression of the nanobodies alters the expression level of the transporter TAXIPm-QM, an expression screening was performed (Figure 91). For this purpose, the transporter was expressed in 15 mL medium in the presence and absence of TAXIPm-QM specific nanobodies selected previously in the ELISA (section 4.11.4). The proteins were expressed from different platforms, TAXIPm-QM from the pEXC3sfGH plasmid and the nanobodies from the pSbInit plasmid (Zimmermann et al. 2018). The use of these two plasmids allows independent control of expression levels on the one hand and independent visualization on the other. The promoters are induced differently, from the pEXC3sfGH plasmid with IPTG and from the pSbInit plasmid with arabinose. In addition, the pEXC3sfGH plasmid contains an amp resistance for selection, while the pSbInit plasmid contains a chloramphenicol resistance. Following disruption of the cells by bead-beating, the proteins were immediately solubilized in 1 % (w/v) DDM, followed by an incubation for 1 h at 4 °C and ultracentrifugation to pellet aggregated protein. All protein samples were separated in a 12 % SDS PAGE and visualized by *in gel* GFP fluorescence showing TAXIPm-QM with sfGFP fusion. Furthermore, an *anti*-His Western blot was performed to visualize the TAXIPm-QM-sfGFP protein and the nanobodies (Figure 91). As TAXIPm-QM and TAXIPm-P are on one operon we expect expression for P where we observe expression for QM. The co-expression of nanobodies and TAXIPm-QM was found in all samples, whereby the expression levels of the nanobodies exhibited different yields. A1B4, B2G4, B2H5 and B2G7 were significantly less expressed than A1C6, A1D7, A1B10, A1C11 and B2D6, though all in multiple excess to the transporter (Figure 91). In general, all nanobodies were highly overexpressed, which most likely causes a reduction in the solubility of A1D7, A1C11, B2G4, B2G7 and especially B2D6 (Figure 91).

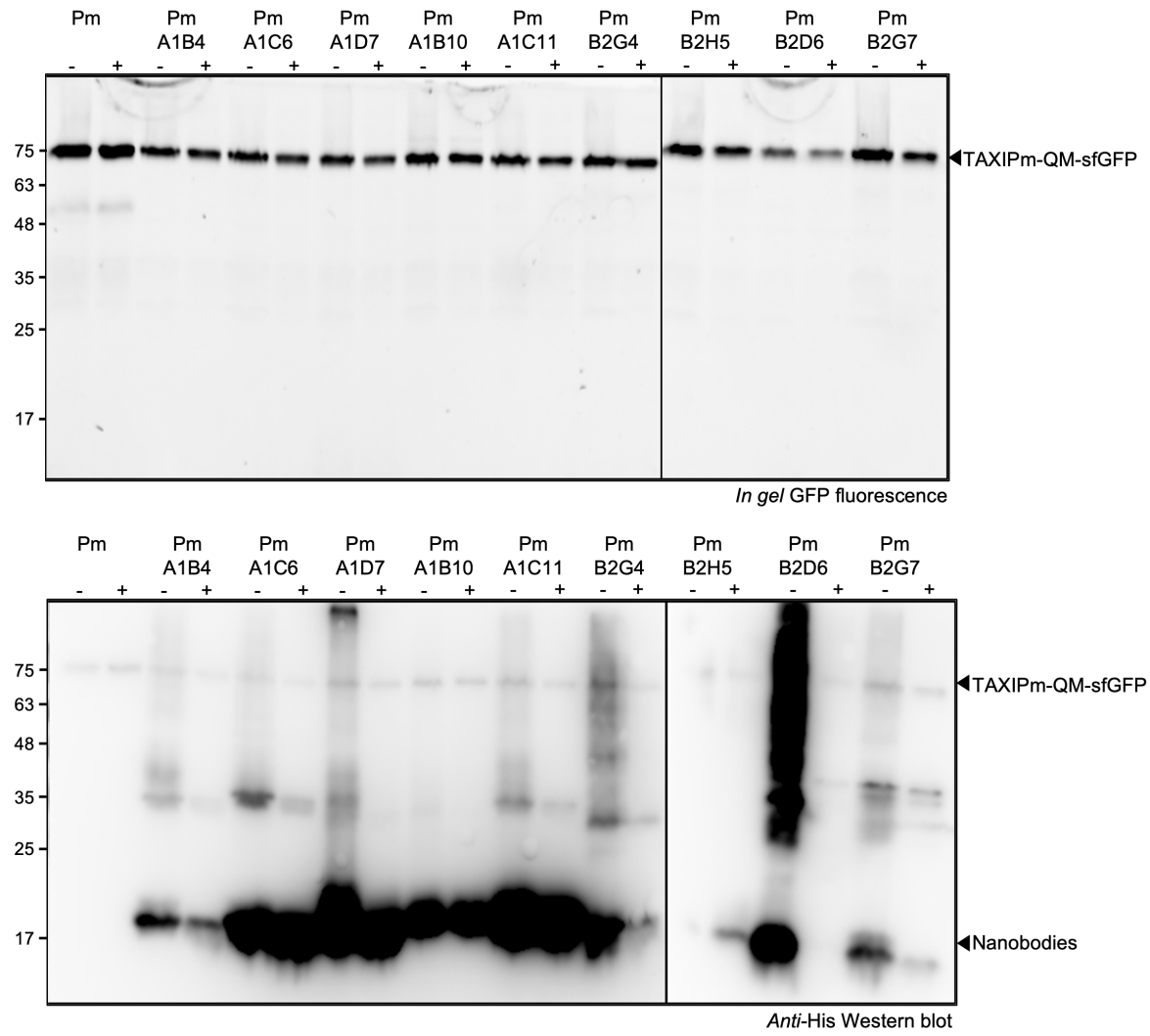


Figure 91 **Overexpression of TAXIPm-PQM and nanobodies.** Whole cell lysates of cell suspensions containing TAXIPm-PQM and nanobodies which were used for radioactive transport assays were solubilized in 1 % (w/v) DDM and submitted to ultracentrifugation to pellet aggregated protein. All protein samples were separated in a 12 % PAGE and visualized by *in gel* GFP fluorescence showing TAXIPm-QM with sfGFP fusion and *anti*-His Western blot showing next to the TAXIPm-QM-GFP protein additionally the nanobodies. TAXIPm-PQM is abbreviated with Pm and the nanobodies with their corresponding numbering on the 96 well plate. The data was generated under my supervision by Christiane Ruse, Goethe University Frankfurt.

We determined the signal intensities of the TAXIPm-QM-sfGFP expressed protein in the 12 % SDS PAGE from after the ultracentrifugation (Figure 91) which showed in a relative comparison the influence of the nanobodies on the expression of the transporter (Figure 92). The comparison of the TAXIPm-QM-sfGFP signal intensities indicated that all nanobodies influenced the expression of TAXIPm-QM, whereby one nanobody, abbreviated with B2D6, noticeably reduced the expression level of TAXIPm-QM (Figure 92).

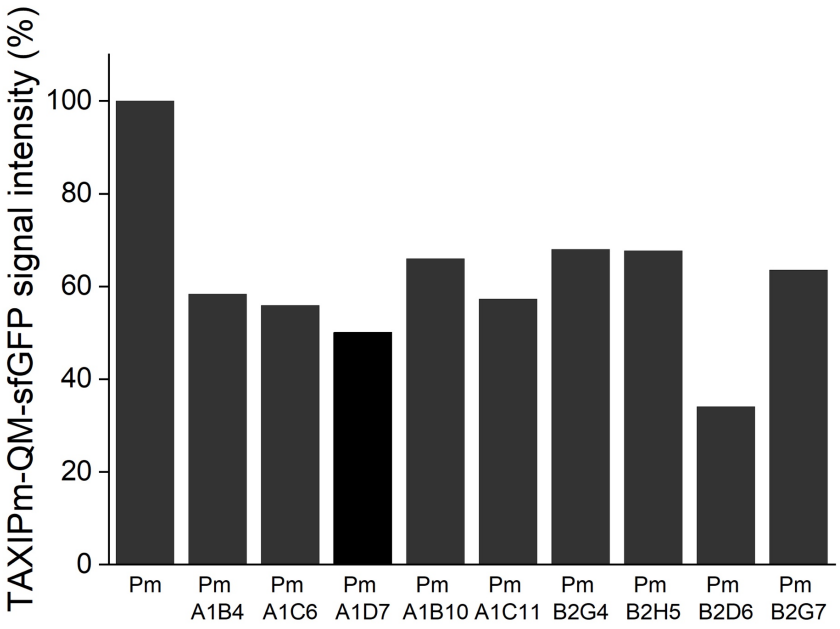


Figure 92 **Evaluation of TAXIPm-QM-sfGFP expression levels in co-expression with nanobodies.** Signal intensities of the TAXIPm-QM-sfGFP expressed protein in a 12 % SDS gel from after the ultracentrifugation were measured. The protein samples from the co-expressions with different nanobodies were compared relative to the expression of the transporter without nanobody addition. TAXIPm-PQM is abbreviated with Pm and the nanobodies with their corresponding numbering on the 96 well plate. The analysis was performed using ImageJ (Schneider, Rasband, and Eliceiri 2012). The data was generated under my supervision by Christiane Ruse, Goethe University Frankfurt. The bar chart was created using OriginPro 2019b (OriginLab Corporation).

4.11.6 *In vivo* transport of α-ketoglutarate by TAXIPm-PQM in presence of nanobodies

In a further selection process of conformationally-selective nanobodies after the ELISA (section 4.11.4), we screened for nanobodies that inhibit the α-ketoglutarate uptake by TAXIPm-PQM in whole *E. coli* cells JW2571-1 lacking the endogenous α-ketoglutarate transport of kgtP. We assumed that a reduced transport activity is then caused by binding of a nanobody to TAXIPm-QM from the periplasmic site which leads to conformational trapping or competition with TAXIPm-P for binding. In the evaluation, however, the influence on expression levels of TAXIPm-QM should be taken into account (Figure 92, section 4.11.5). As TAXIPm-QM and TAXIPm-P are on one operon we expect expression for P where we observe expression for QM. For this purpose, the α-ketoglutarate deficient *E. coli* strain JW2571-1 was co-transformed with TAXIPm-PQM on a pEXC3sfGH plasmid and the nanobodies on the pSBInit plasmid so that both can be simultaneously expressed in the same cell (Zimmermann et al. 2018). IPTG was used to induce the expression of TAXIPm-PQM and arabinose for the expression of the nanobodies. Having the arabinose promoter on the pSBInit plasmid was especially beneficial so that the amount of nanobody could be tuned thereby ensuring an excess of binder. The comparison of transport inhibition in the presence of the non-specific sybody Sy45 and previously selected TAXIPm-QM specific nanobodies showed a reduction initiated by both, the unspecific and specific binders (Figure 93). However, many specific nanobodies led to a stronger reduction of α-ketoglutarate uptake by TAXIPm-PQM than the non-specific sybody Sy45. The strongly reduced transport caused by B2D6 is most likely a consequence of the reduced expression of the TAXIPm-QM caused by the co-expression of this nanobody (Figure 92, section 4.11.5). Other

nanobodies also showed reduced expression of the transporter. However, in a direct comparison, the reductions in expression did not always correlate with the reductions in transport. For example, co-expression of the nanobodies B2G4 and B2H5 reduced the expression of TAXIPm-QM only slightly but reduced the transport strongly (Figure 92 and Figure 93). Especially the latter, B2H5 seemed a promising candidate as additionally, it was not as highly overexpressed when compared to other nanobodies and overall showed well folding (Figure 91).

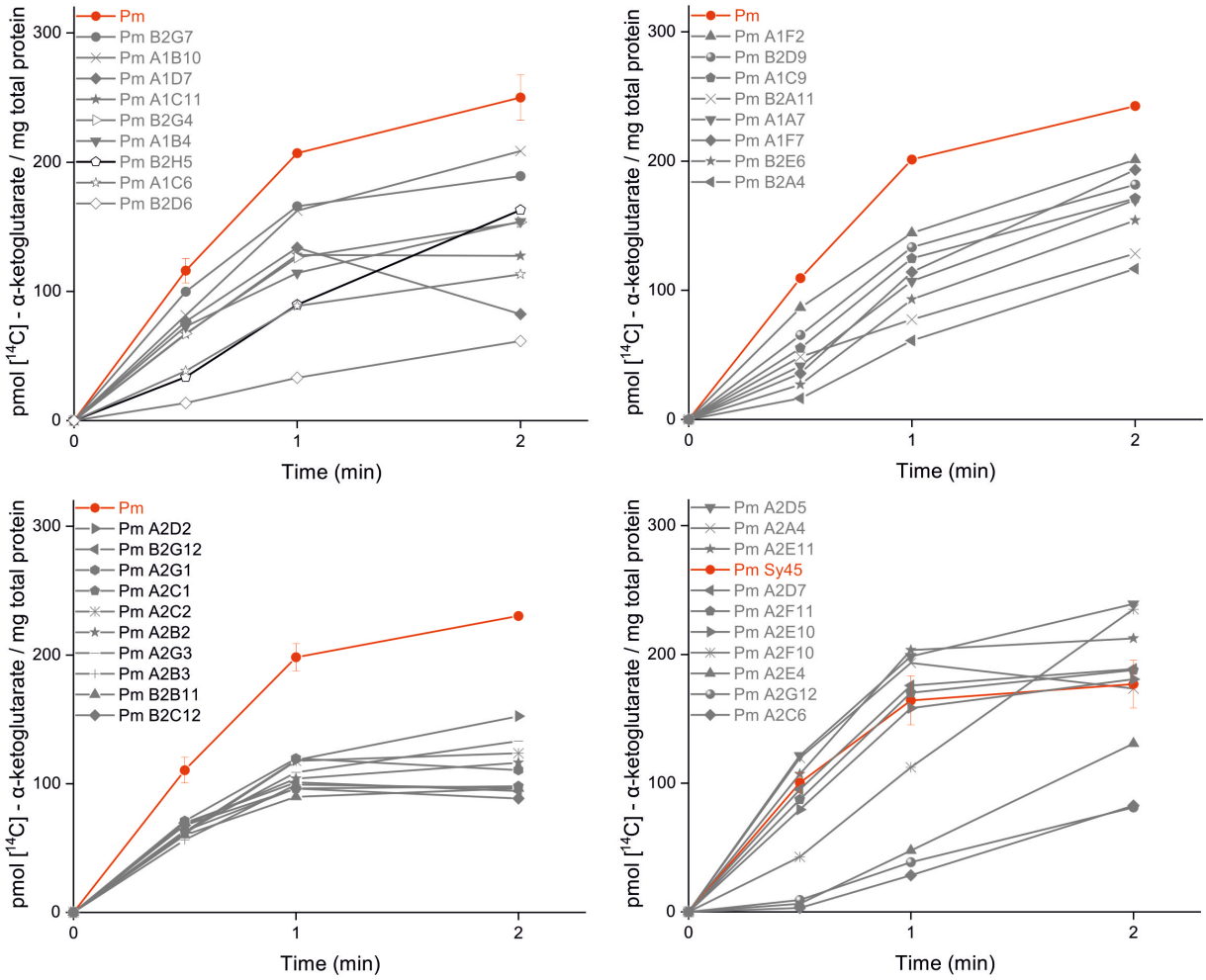


Figure 93 **α-ketoglutarate transport inhibition of TAXIPm-PQM by nanobodies.** α-ketoglutarate uptake by TAXIPm-PQM in whole *E. coli* JW2571-1 lacking the endogenous α-ketoglutarate transport of kgtP in presence or absence of previously selected TAXIPm-QM specific nanobodies. Cells were pre-energized by a 2 min pre-incubation with 0.5 % (w/v) glucose. Transport was assayed at 30 °C and pH 7.5 in the presence of 2 µM [^{14}C] labelled α-ketoglutarate. Expression of TAXIPm-PQM in pEXC3sfGH and nanobodies in pSbInit were induced by 1 mM IPTG and 0.01 % arabinose, respectively. The data points that contain error bars are duplicates, all other transport data derives from single time point measurements. TAXIPm-PQM is abbreviated with Pm and the nanobodies with their corresponding position on the 96 well plate. Sy45 is a UraA specific sybody which serves as negative control. The data was partly generated by me and partly under my supervision by Christiane Ruse, Goethe University Frankfurt. The plots were created using OriginPro 2019b (OriginLab Corporation).

4.11.7 Purification of transport inhibiting nanobodies

The promising nanobody B2H5 nanobody that showed strong inhibition in *E. coli* whole cell uptake assays next to a slight influence on the expression level of the transporter was expressed in a 1 L scale and purified (Figure 94). To increase the variety, further nanobodies, abbreviated with A1C6, B2G4 and B2D6 were equally expressed in a 1 L scale and purified. For B2D6 a low yield was

obtained (Figure 94) which confirms that this nanobody is too high overexpressed which most likely causes a decrease in solubility (section 4.11.5). The SEC chromatograms show peaks at the expected elution volume of approximately 13 mL for each nanobody indicating high purity and functionality of the proteins within these fractions (Figure 94). Additional peaks presumably comprise aggregated protein and impurities (Figure 94).

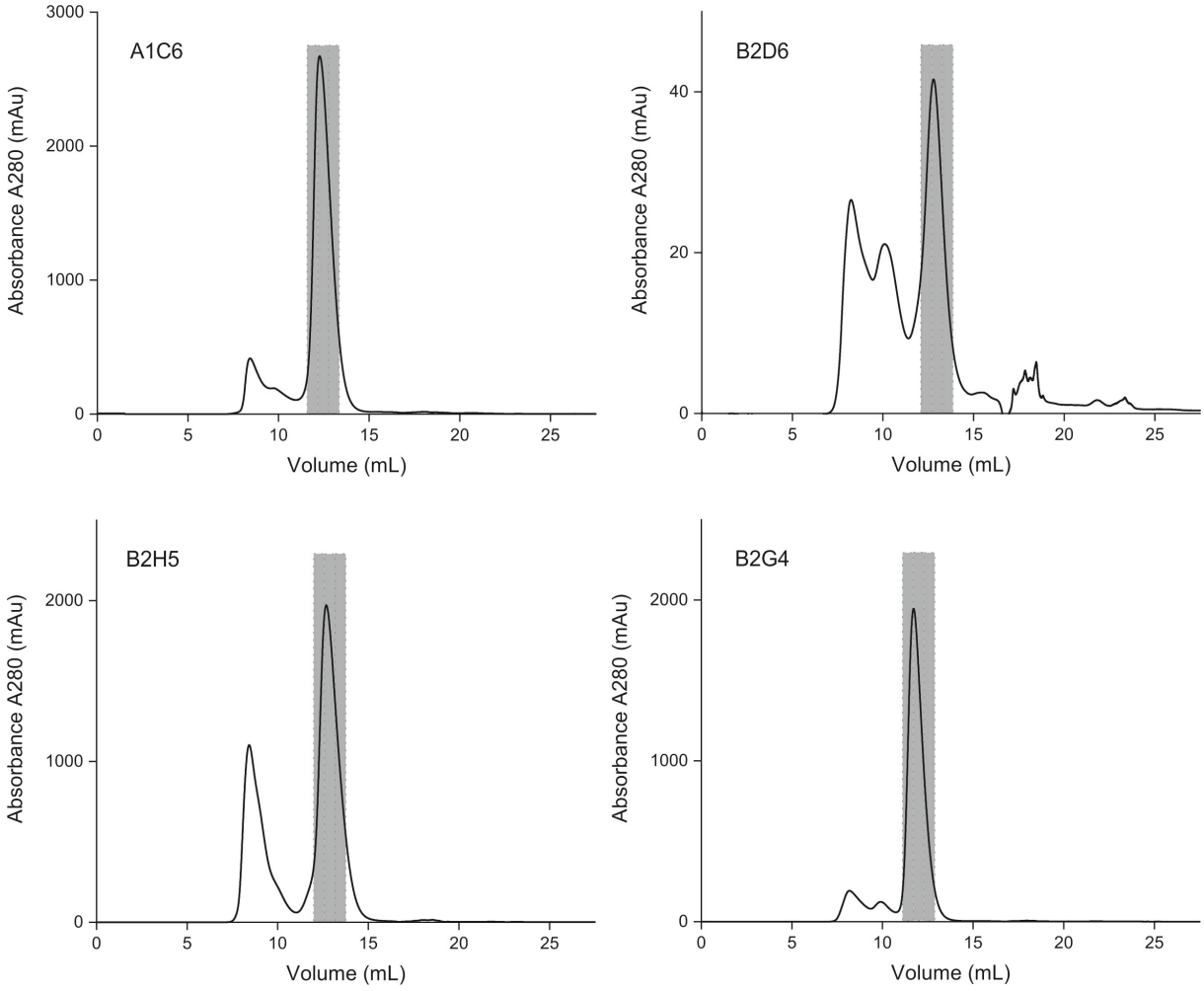


Figure 94 **Purification TAXIPm-QM specific nanobodies for structural analysis**. The previously selected TAXIPm-QM specific nanobodies A1C6, B2D6, B2H5 and B2G4 were purified via IMAC and further purified by SEC on a Superdex 75 30/100 GL Increase column in 20 mM Hepes, 150 mM NaCl at pH 7.5. The nanobodies eluted at approximately 13 mL, the corresponding fractions are depicted in grey. The data was generated under my supervision by Christiane Ruse, Goethe University Frankfurt. The plots were created using OriginPro 2019b (OriginLab Corporation).

The SEC fractions which contained the pure target protein, approximately at 13 mL, were sent to the MPI CBG in Dresden (Eric Geertsma Group, www.mpi-cbg.de) for conformational stabilization and structural analysis in cryo-EM studies with the aim of different TAXIPm-QM conformations stabilized by the selected nanobodies. This should additionally provide new insights in structure-function relationships in TAXI proteins.

5 Discussion

In contrast to the mechanism of primary transport systems, the mechanism of SBP dependent secondary transport is not well understood. Detailed *in vitro* characterization has thus far been limited to three N-acetylneuraminic acid TRAP transporters, derived from *Haemophilus influenzae, Vibrio cholerae* and *Photobacterium profundum* (Mulligan et al. 2012, 2009; Davies, Currie, and North 2023). As the TRAP-TAXI family is the least studied group, we selected homologs of this group for functional characterization.

5.1 Deorphanization of TAXIPm-PQM

To identify unknown proteins with superior biochemical properties a homolog screening of TRAP transport systems was performed. We specifically focused on those proteins that have fused QM domains as we thereby avoid losing of one of the subunits during the purification procedure. The number of candidates was additionally reduced by selecting TRAP-QM domains in our homolog search for which we could unambiguously identify the associated SBP in the same operon (section 4.1). One such candidate for which we identified the substrate, was TAXIPm-PQM from *Proteus mirabilis* which is an enteric bacterium causing urinary tract infections (Pearson et al. 2011).

In our study, thermal melting of the SBP TAXIPm-P was increased by the addition of α-hydroxyglutarate and α-ketoglutarate (Figure 38, section 4.5.3), from which transport by TAXIPm-PQM could be verified for α-ketoglutarate (Figure 68, section 4.10.1 and Figure 79, section 4.10.8). Structurally similar substrates like glutamate and glutarate did not bind (Figure 38, section 4.5.3). It was previously suggested that some TRAP SBPs are very specific while others bind structurally similar ligands (Vetting et al. 2015). The TRAP SBP B2673 from *Bacillus haludorans* showed significant thermal stabilization when D-gluconate, D-allonate, D-galactonate, D-xylonate or D-ribonate bound, while no stereoisomers were recognized (Vetting et al. 2015). It was therefore suggested that the carboxylates and stereochemistry determine whether ligands can bind (Vetting et al. 2015). With the binding of α-hydroxyglutarate and α-ketoglutarate to TAXIPm-PQM, we report the first C_5-dicarboxylate TRAP transporter. Previously identified substrates for TRAP systems involved dicarboxylates but was limited to C_3- and C_4-dicarboxylates only (Vetting et al. 2015).

Different strategies can be followed to identify ligands for proteins, such as surface plasmon resonance, mass spectrometry (Stigter, de Jong, and van Bennekom 2013), crystallization, thermal shift assaying (Vetting et al. 2015) or simulated molecular docking (Bender et al. 2021). We used a combination of a thermal shift assay and a transport assay in order to screen a broad range of ligands for binding and only a small subset for transport. In our compound library, which we screened in DSF, we predominantly included substrates that had been assigned to other TRAP transport systems (Vetting et al. 2015; Mulligan, Fischer, and Thomas 2011). Our inclusion of α-ketoglutarate and its direct precursor α-hydroxyglutarate was motivated by the presence of the *lhgO* gene (PMI1053), coding for a putative L-2-hydroxyglutarate oxidase, located downstream of the TAXIPm-QM (PMI1055) and TAXIPm-P (PMI1056) genes in the same operon (Figure 95).

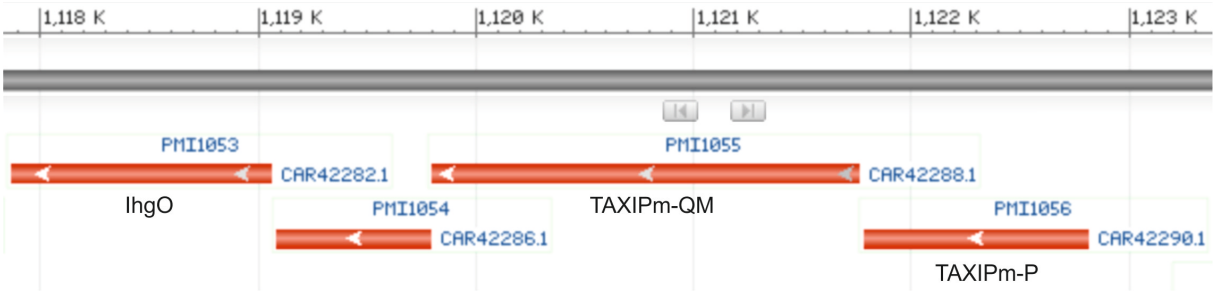


Figure 95 **Gene neighborhood of TAXIPm-PQM.** lhgO, the gene encoding α-hydroxyglutarate is in the direct gene neighborhood of TAXIPm-PQM. Figure was adapted from the NCBI BLAST database (Madden 2002). The initial genomic DNA analysis of Proteus mirabilis was carried out by Melanie Engelin, Goethe University Frankfurt.

The *lhgO* gene encodes L-2-hydroxyglutarate oxidase (PMI1053) (Price et al. 2005) which converts α-hydroxyglutarate into α-ketoglutarate (Figure 96) (Kalliri, Mulrooney, and Hausinger 2008).

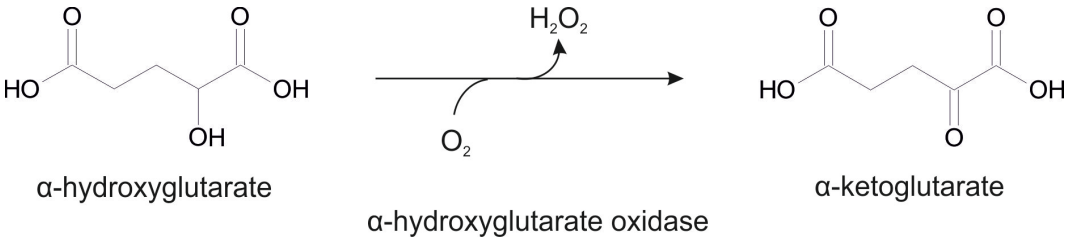


Figure 96 **Reaction catalyzed by α-hydroxyglutarate oxidase.** α-hydroxyglutarate is converted to α-ketoglutarate upon the presence of oxygen by α-hydroxyglutarate oxidase. The chemical structures were created using ACD/ChemSketch (Advanced Chemistry Development, Inc. 2021.2.1).

The *lhgO* gene might be essential in *Proteus mirabilis* to replenish α-ketoglutarate when it is rare. This behavior implies a strategy to ensure sufficient amounts of α-ketoglutarate by creating an α-ketoglutarate uptake and conversion pipeline. The particular relevance of providing higher amounts of α-ketoglutarate in *Proteus mirabilis* when compared to *E. coli* was previously shown (Pearson et al. 2011). Carbon sources are limited in urine with the absence of glucose but presence of citrate. Most *Proteus mirabilis* strains can use citrate as sole carbon source whereas most *E. coli* strains cannot. This ability explains the pathogenic potential of *Proteus mirabilis* to cause urinary tract infections. However, the production of sufficient amino acids is complicated if carbon sources are rare. Therefore, the carbon-nitrogen balance is opposite to *E. coli*. *Proteus mirabilis* facilitates upregulation of the glutamate dehydrogenase to produce glutamate by using α-ketoglutarate from the tricarboxylic acid (TCA) cycle (Figure 97). For this purpose, it possesses an efficient urease for hydrolysis of urea into ammonia and carbon dioxide to ensure accumulation of high nitrogen amounts. As a consequence, in addition to the conversion of α-ketoglutarate to succinate in the TCA cycle, α-ketoglutarate is also extensively converted to L-glutamate in *Proteus mirabilis* (Figure 97). To provide sufficient amounts of predecessor of α-ketoglutarate and thereby sustain the full TCA cycle next to its conversion to L-glutamate, all enzymes upstream α-ketoglutarate need upregulation (Pearson et al. 2011).

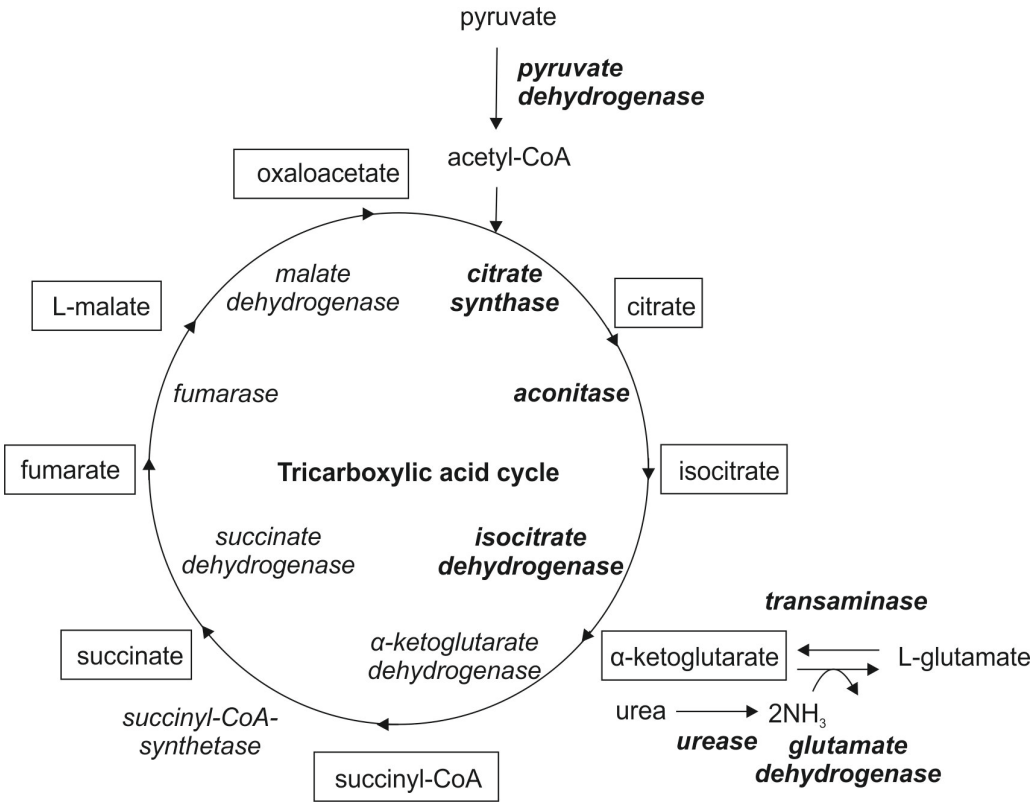


Figure 97 **Interplay of TCA cycle and carbon-nitrogen balance.** Pyruvate conversion and ammonia (NH3) production are high to ensure L-glutamate production while maintaining the full TCA cycle. Enzymes in bold are upregulated at least two-fold in *Proteus mirabilis* when compared to *E. coli.* Figure was adapted from (Pearson et al. 2011).

To initially deorphanize TAXIPm-PQM, we performed the thermal shift assay with TAXIPm-P alone as in tripartite systems the SBP determines the substrate specificity. We used 40 µM substrate, which resulted in increases in the melting temperature of + 5.3 °C for α-hydroxyglutarate and + 3.9 °C for α-ketoglutarate relative to the apo-protein. Upon titration of the substrate we no longer observed a melting curve at α-ketoglutarate concentrations above 50 µM (Figure 42, section 4.5.3), suggesting that high substrate concentrations destabilize the protein (Scott, Spry, and Abell 2016). The maximum temperature shift of + 9.3 °C was detected at an α-ketoglutarate concentration of 12.5 µM showing that lower substrate concentrations might yield a higher stabilization of the protein. The opposite was expected as higher substrate concentrations lead to a higher degree of saturation of the binding protein which typically results in increased stabilization. The observed destabilization can have many causes. High substrate concentrations might induce non-specific effects resulting in the substrate not only bound to the native but also the unfolded protein. This indiscriminate binding might affect the binding of the substrate to the native protein, and it is thereby likely that multiple substrate molecules bind to the protein which can result in both positive and negative thermal shifts. Furthermore, high substrate concentrations could stabilize a conformation of the protein with a lower melting temperature or a conformation with a significantly higher hydrophobic surface which enables binding of the fluorescent dye. Both would decrease the melting temperature compared to the main native protein (Scott, Spry, and Abell 2016). A previous screening of 82 TRAP SBPs by DSF resulted in an increased stability of 61 TRAP SBPs by >5 °C and seven TRAP SBPs by >10 °C upon the

addition of the ligands. An increase of the protein stability was reported only, thereby using 1 mM ligand and 10 µM protein (Vetting et al. 2015). In the case of ABC transporters some were stabilized by their substrates and others were not. The respective behavior was observed to also vary even with same substrate specificities so that it was concluded to correlate with different folds of the transmembrane domains (Vigonsky, Ovcharenko, and Lewinson 2013). Destabilization was observed for the ABC transport systems BtuCD-F and hiMolBC–A as higher substrate concentrations reduced the transport rate (Lewinson et al. 2010). As both belong to the ABC transport systems type II, it is potentially characteristic for this group. *In vitro* experiments using reconstituted liposomes showed a decreased affinity between BtuCD and BtuF at substrate concentrations above 50 µM (Lewinson et al. 2010). The same was observed for hiMolBC–A as more substrate resulted in less complex formation of the membrane domain hiMolBC with the correlating SBP hiMolA (Vigonsky, Ovcharenko, and Lewinson 2013). This reduced interaction was explained by a lower affinity of the SBP for the membrane domain resulting in a lower K_{on} upon the presence of substrate. This might result from an interaction of the membrane domain with the substrate and associated allosteric regulation (Vigonsky, Ovcharenko, and Lewinson 2013), potentially leading to conformational changes and prevention of transport. As ABC type II systems generally transport substrates that are available only at low concentrations (Locher 2016), the destabilization might be physiologically irrelevant.

For the TRAP-TAXI representative TAXIPm-P, we showed that high substrate concentrations significantly reduce protein stability (Figure 42, section 4.5.3). Transport of α-ketoglutarate by TAXIPm-PQM was initially shown in *E. coli* whole cells JW2571-1 using 2 µM substrate (Figure 68, section 4.10.1) and then confirmed *in vitro* in proteoliposomes using 10 µM substrate (Figure 79, section 4.10.8) which are both concentrations that do not destabilize the SBP. This may indicate a control-mechanism allowing TAXIPm-PQM-mediated α-ketoglutarate uptake only when the substrate is below a certain threshold concentration. However, there is no obvious reason why *P. mirabilis* would limit the α-ketoglutarate uptake. The observed destabilization might be physiologically irrelevant as *Proteus mirabilis* is primarily associated with causing infections in the urinary tract where α-ketoglutarate is rare (Pearson et al. 2011). Therefore, it is likely that the destabilization does not occur *in vivo*.

For the SBP TAXIPm-P, we analyzed the substrate specificity for α-ketoglutarate in detail by ITC using 30 µM protein and ligand concentrations from 4.3 µM to 111.4 µM. As the protein was saturated before the substrate titration reached concentrations above 50 µM in the reaction cell, a possible (de)-stabilization of the protein at higher substrate concentrations could not be investigated in ITC. To determine which species of α-ketoglutarate binds to TAXIPm-P, we purified the SBP and titrated the substrate at pH 5.5, 6.0 and 7.5 (Figure 49, section 4.7.2). As dicarboxylates can exist in three charged states, the neutral, single anionic and di-anionic form, equilibria i.e., pK_a-values are associated with these states. The pK_a-values of α-ketoglutarate are 1.9 and 4.4, so at pH 5.5, 6.0 and 7.5 always deprotonated, either single anionic or di-anionic. In general, the di-anionic form predominates at all pH-values and the percentage of the single anionic form increases towards lower pH-values (Figure 48, section 4.7.2). The relative amount of the single anionic form changes its

percentage >70-fold from 7.4 % at pH 5.5 down to 0.1 % at pH 7.5. We observed α-ketoglutarate binding affinities for TAXIPm-P of 16.6 µM, 15.4 µM and 8.4 µM at pH 5.5, 6.0 and 7.5, respectively (Figure 49, section 4.7.2). This modest change in binding affinity does not agree with the 70-fold reduction that the single anionic species undergoes in this pH-range. In addition, the molar ratios at which the binding is saturated are almost identical. Together, these findings imply that the concentration of the species that is able to bind does not vary significantly in the pH-range tested. The results point to slightly altered binding properties or destabilization of the protein under the changed conditions (Broom et al. 2015). We thus suggest that TAXIPm-P either exclusively or favorably transports the di-anionic species of α-ketoglutarate.

5.2 Deorphanization of TAXIMh-PQM

For the SBP TAXIMh-P from *Marinobacter hydrocarbonoclasticus* we identified three structurally similar ligands, fumarate, succinate, and L-malate (Figure 35, section 4.5.2). The binding of fumarate and L-malate to the SBP was confirmed by ITC resulting in K_D-values of 31.2 µM and 10.9 µM, respectively (Figure 44, section 4.7.1). The binding could not be confirmed for succinate (Figure 44). In addition to binding, we could confirm the ability of TAXIMh-PQM to transport fumarate as well using a whole cell transport assay in *E. coli.* (Figure 60, section 4.9.1). In our compound library, which we screened in DSF, we predominantly included substrates that had been assigned to other TRAP transport systems (Vetting et al. 2015; Mulligan, Fischer, and Thomas 2011). Among those were fumarate, succinate, and L-malate which proved to also bind TAXIMh-P. In retrospect, the binding is not surprising given the close gene proximity of TRAPMh-P (MARHY0896) and TRAPMh-QM (MARHY0897) to the *DcuB* and *DcuD* gene which are both associated with anaerobic C_4-dicarboxylate metabolism and fumarate respiration (Figure 98) (Kim 2006).

Figure 98 **Gene neighborhood of TAXIMh-PQM.** DctB and DctD are in the direct gene neighborhood of TAXIPm-PQM which both of which are associated with anaerobic C_4-dicarboxylate metabolism and fumarate respiration. Figure was adapted from the NCBI BLAST database (Madden 2002).

Similar as for TAXIPm-PQM (section 5.1), genes located in close proximity to the transporter provided information about the transporting substrate for TAXIMh-PQM. Therefore, it might be the best approach in the deorphanization of TRAP transport systems to examine the genes near the genes encoding the P- and QM-domains before large compound libraries are considered. In contrast to TAXIPm-PQM, to which only α-hydroxyglutarate and α-ketoglutarate bind, structurally similar

substrates were found to bind to TAXIMh-PQM. However, L-malate, fumarate, and succinate are also metabolically connected (Kim 2006) simultaneously to α-hydoxyglutarate and α-ketoglutarate (Brunengraber 2007). *DcuB* encodes an antiporter which imports the C_4-dicarboxylates fumarate, malate, aspartate and D-tartrate in exchange for succinate. After the substrates are imported, fumarase and fumarate reductase convert L-malate to fumarate and fumarate to succinate (Figure 99). The aim of these conversion steps is to enhance anaerobic fumarate respiration which allows the organism to gain energy (Nguyen et al. 2020). As *DcuB* does not only export succinate but also imports fumarate and L-malate, it can contribute to maintaining metabolism under anaerobic conditions as the TCA cycle operates under aerobic conditions. *DcuD* is a silent gene which might be a potential C_4-dicarboxylate transporter and therefore serve as a backup for *DcuB* (Ingo G. Janausch, Garcia-Moreno, and Unden 2002; Kim 2006). The importance of fumarate reductase and succinate dehydrogenase to maintain metabolism was demonstrated in mutagenesis studies, i.e. mutations led to diminished bacterial growth (Iverson et al. 1999).

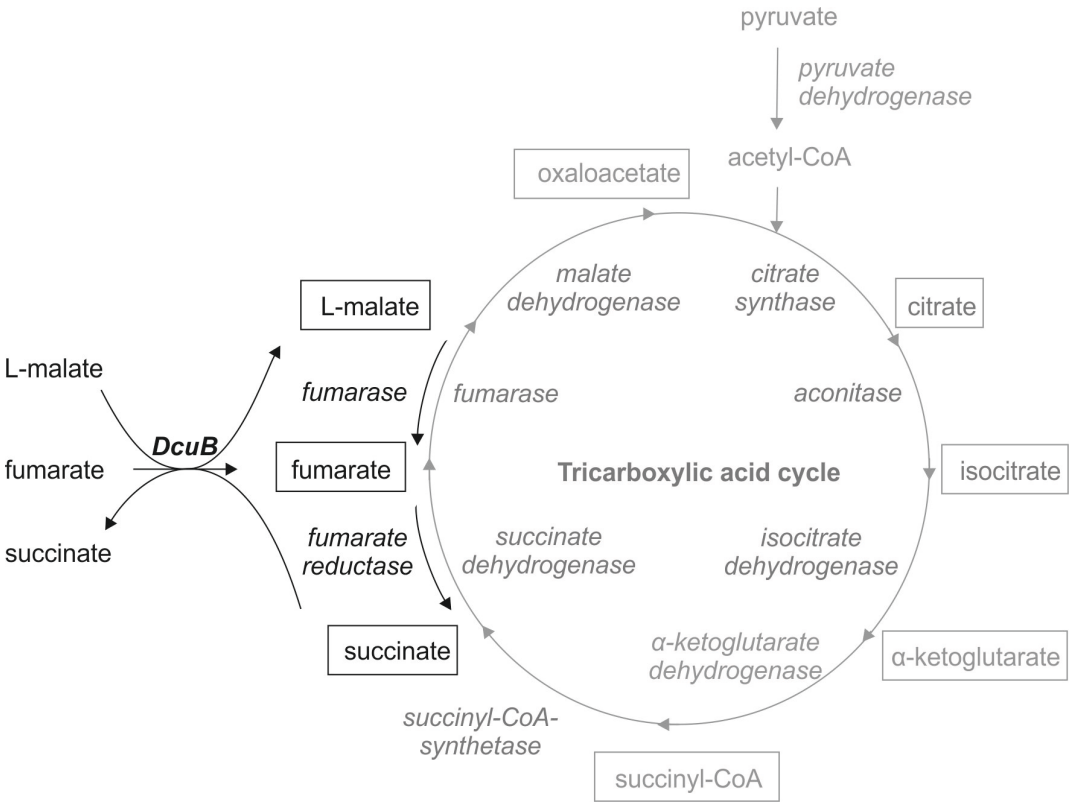


Figure 99 **Interplay of TCA cycle and DcuB.** DcuB is an antiporter which imports L-malate and fumarate in exchange for succinate. L-malate is converted to fumarate by fumarase and subsequently to succinate by fumarate reductase. Figure was adapted from (Nguyen et al. 2020).

TAXIMh-PQM might be essential in *Marinobacter hydrocarbonoclasticus* to replenish fumarate and most likely L-malate and succinate when these substrates are rare, and when the other sources are diminished. The extreme habitat conditions make this likely, as the Gram-negative bacterium naturally occurs in marine waters. TAXIMh-PQM has the advantage of being a high affinity SBP dependent uptake system which might be needed to accumulate sufficient substrate in the marine environment. It was previously shown that the organism can use fumarate next to others as the sole

carbon source to gain energy, but not carbohydrates (Gauthier et al. 1992; Handley and Lloyd 2013). It was furthermore shown in growth assays that *Marinobacter hydrocarbonoclasticus* has an obligate requirement for sodium. As it occurs in the marine environment, which is rich in Na^+ with approximately 500 mM (Handley and Lloyd 2013; Milo and Phillips 2015), one might speculate that TAXIMh-PQM might be a sodium coupled transport system. Overall, with the transport of fumarate by TAXIMh-PQM, we confirm the paradigm that a ligand able to bind to the SBP is likely transported by the corresponding membrane transport domains. This implies that the SBP determines the substrate specificity, and the membrane domain transports all substrates delivered by the corresponding SBP. Indeed, the ABC transporter OppBCDF from *Lactococcus lactis*, can transport all peptides delivered by the SBP OppA, with which it is able to interact (Doeven 2006). Although transport rates were highest for peptides with nine residues, peptides with up to 35 residues were transported, as were those with modified N- or C-termini. As long as the interaction between the SBP and membrane domain is not disturbed, all peptides delivered by the SBP are transported (Doeven 2006).

To further investigate the fumarate transport of TAXIMh-PQM in detail, we continued with *in vitro* transport assays. For this purpose, the wildtype TAXIMh-QM was reconstituted into liposomes and transport assays were performed in which soluble SBP TAXIMh-P was supplied using an established protocol of *in vitro* assays previously performed for the TRAPHi-SiaPQM from *Haemophilus influenzae* (Mulligan et al. 2009). Different samples were prepared with varying substrate concentrations and the application of either a proton or sodium gradient was tested. No activity was measured although fumarate can bind to the SBP TAXIMh-P, as demonstrated in both DSF and ITC (section 4.5.2 and 4.7.1) and can be transported into *E. coli* whole cells (section 4.9.1). The reconstitution efficiency of TAXIMh-QM yielded 75.2 % (Figure 62, section 4.9.2), although it is unclear whether the membrane domains are reconstituted inside-out or right-side-out. If TAXIMh-QM is reconstituted inside-out only, the SBP cannot transfer substrate to the membrane domain. Also, other lipids in the liposomes might be needed to achieve another conformational state that may be needed for successful transport. Reasons for non-activity may also originate from the SBP which must not necessarily occur in a soluble state endogenously but rather anchored to the membrane as for TAXIPm-PQM and suggested for other proteins belonging to the TRAP-TPAT family (Brautigam et al. 2012). An analysis of the signal peptide by SignalP suggests a transmembrane-anchor with a likelihood of 67 %, labeled as ‘other’ as the signal peptide cannot be assigned to a known signal peptide. The likelihood for a secretory pathway with signal peptidase I cleavage was predicted with 22 % and a secretory pathway with signal peptidase II cleavage which results in a lipid anchor with 10 % (Figure 100) (Teufel et al. 2022). Overall, we conclude that the SBP most likely needs to be anchored in the membrane to allow transport. This anchor dependent transport mechanism is common for several SBP dependent primary transport systems but it has not been reported for any SBP dependent secondary transport systems so far (van der Heide and Poolman 2002).

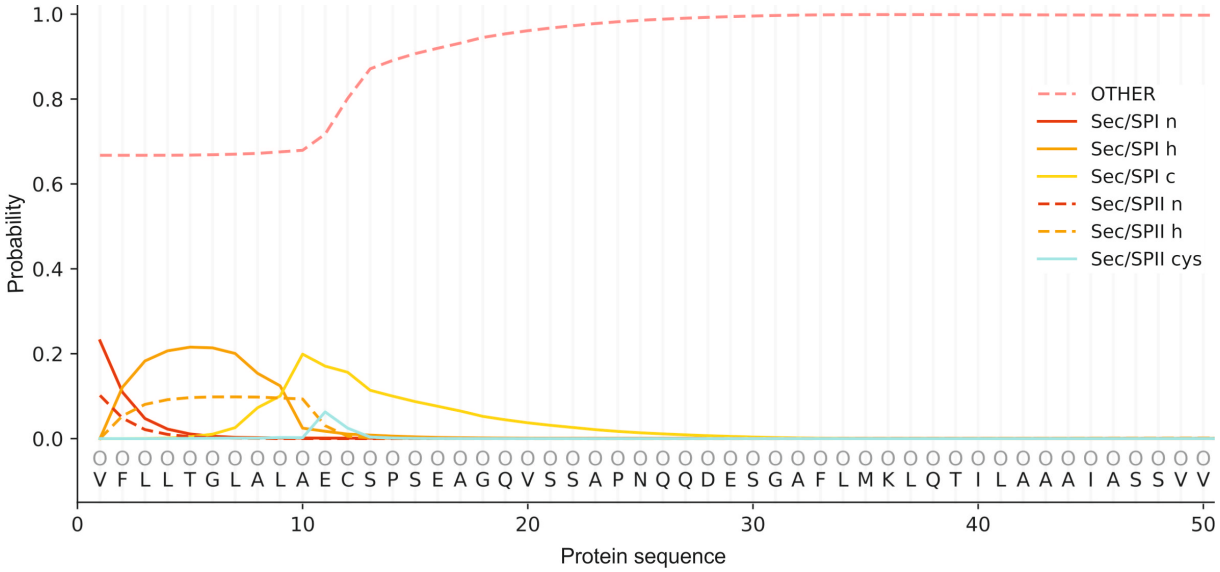


Figure 100 **Analysis of TAXIMh-P signal sequence.** The TAXIMh-P signal sequence from *Marinobacter hydrocarbonoclasticus* was analyzed using SignalP with its endogenous signal sequence. The light red dashed line indicates the probability for "other", which refers to no prediction of a signal sequence, but an N-terminal transmembrane segment. The red, orange and yellow lines show the probability for a secretory pathway with Signal Peptidase I cleavage, respectively. The dashed red, dashed orange and turquoise lines show the probability for a secretory pathway with Signal Peptidase II cleavage, respectively. The predictions are detailed for the N-region (n), middle region (h) and C-region (c/cys) of the sequence (Teufel et al. 2022).

5.3 Relevance of a membrane-anchored SBP

Thus far *in vitro* characterization of SBP dependent secondary transporters is limited to three N-acetylneuraminic acid transporters from *Haemophilus influenzae, Vibrio cholerae* and *Photobacterium profundum* which all use soluble SBPs (Mulligan et al. 2009, 2012; Davies, Currie, and North 2023). We demonstrated the dependency on a membrane-tethered SBP for TAXIPm-QM to facilitate α-ketoglutarate uptake, the first TRAP for which this dependence has been experimentally demonstrated. Membrane-anchors are commonly used to prevent diffusion of the SBP out of the cell in Gram-positive prokaryotes or archaea. They also appear more rarely in Gram-negative bacteria (Berntsson et al. 2010; Detmers 2001). The anchoring is realized either through N-terminal lipid moieties (Sutcliffe and Russell 1995) or by incorporation of transmembrane helices in the membrane (Albers et al. 1999). Systems have also been found in which the SBP appear as a direct fusion to the translocator protein, either individually or multiple times per TMD (van der Heide and Poolman 2002).

However, as TAXIPm-PQM is derived from the Gram-negative organism *Proteus mirabilis* (Schaffer and Pearson 2015), the membrane-anchor must serve a different role than preventing the loss of the SBP. For example, membrane-anchors can efficiently increase the local concentration of the SBP near the membrane-transport domain as diffusion is possible in a two-dimensional space only (Doeven 2006). In our transport assay, we initially supplied total 5 µM soluble SBP and 10 µM α-ketoglutarate which, given a K_D-value between 8 and 16 µM, provides approximately 2.5 µM liganded SBP. This concentration range of the liganded SBP did suffice for the N-acetylneuraminic acid transporter TRAPHi-SiaPQM from *Haemophilus influenzae* (Figure 119, section 6) and TRAPVc-SiaPQM from *Vibrio cholerae* when using 5 µM N-acetylneuraminic acid combined with 5 µM TRAPHi-SiaP or TRAPVc-SiaP (Mulligan et al. 2012, 2009). For TAXIPm-PQM, however, these

conditions failed. Presumably, the binding affinity of TAXIPm-QM or TAXIPm-P is too low. Only after providing the SBP TAXIPm-P membrane-tethered, we did observe α-ketoglutarate transport *in vitro* (Figure 79, section 4.10.8). In our transport assays we used 0.01, 0.1 or 1 nmol SBP per 200 µg lipids to record each data point depending on the TAXIPm-QM to TAXIPm-P ratio of 1:0.1, 1:1 or 1:10, respectively. The effective concentration of the SBP can be estimated assuming that: *i)* all proteoliposomes were unilamellar; *ii)* an average molecular weight and surface per lipid of 750 Da and 0.5 nm^2, respectively (Phillips 2018); and *iii)* a maximum distance of 6.31 nm between the SBP and the membrane (Figure 101). Based on this data, we calculated a local TAXIPm-P concentration of 42 µM, 420 µM or 4.2 mM, which greatly exceeds the concentration of 0.1, 1 or 10 µM if the same amount of SBP is added without membrane-anchor.

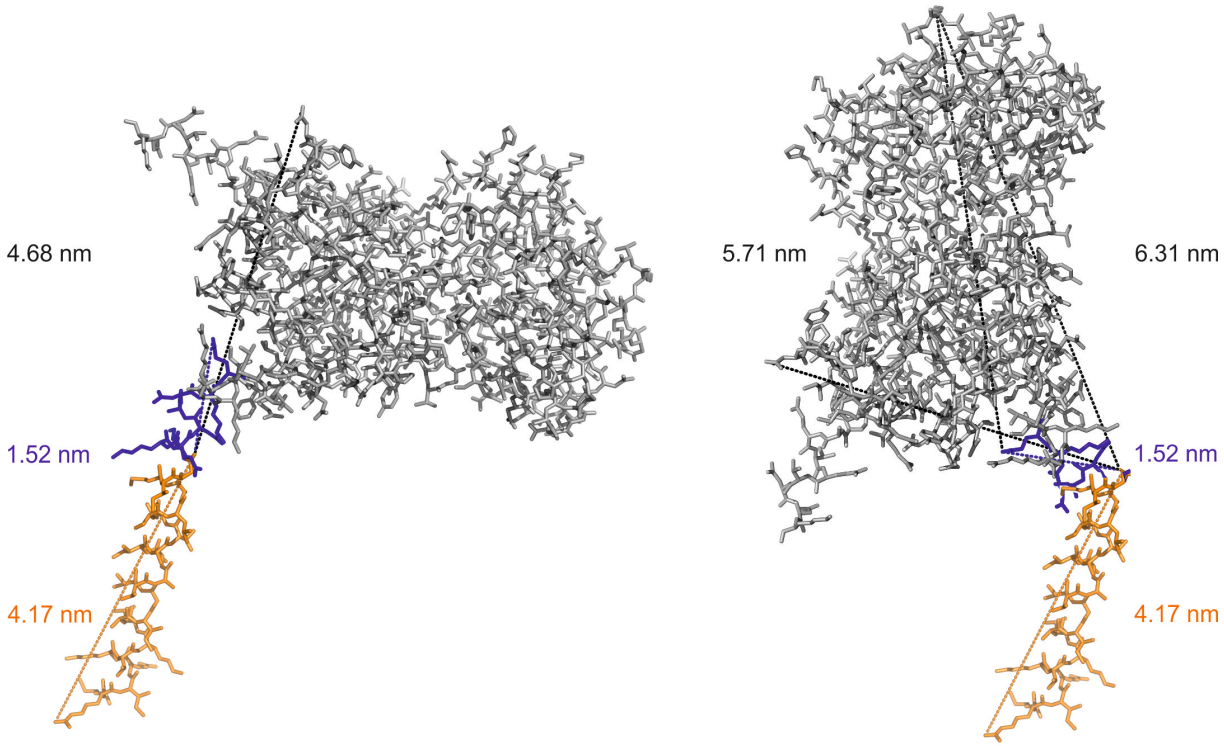


Figure 101 **Distance of TAXIPm-P from membrane.** The hydrophobic stretch which presumably incorporates into the membrane is displayed in orange, the flexible part in blue and the part of the protein which locates in the periplasmic space is displayed in grey. Distances were measured and structures created using Pymol 1.8.4.1 (Schrödinger, LLC).

To investigate the occurrence of membrane-bound SBPs, we analyzed 215 TAXI and 636 TRAP SBP sequences from the TRAP database (Mulligan, Kelly, and Thomas 2007) in SignalP 6.0 (Teufel et al. 2022). We observed that full-length TAXIPm-P is not cleaved by *E. coli* Signal Peptidase I, in line with no predicted Signal Peptidase I cleavage site upon analysis using SignalP 6.0 (section 4.3) (Teufel et al. 2022).

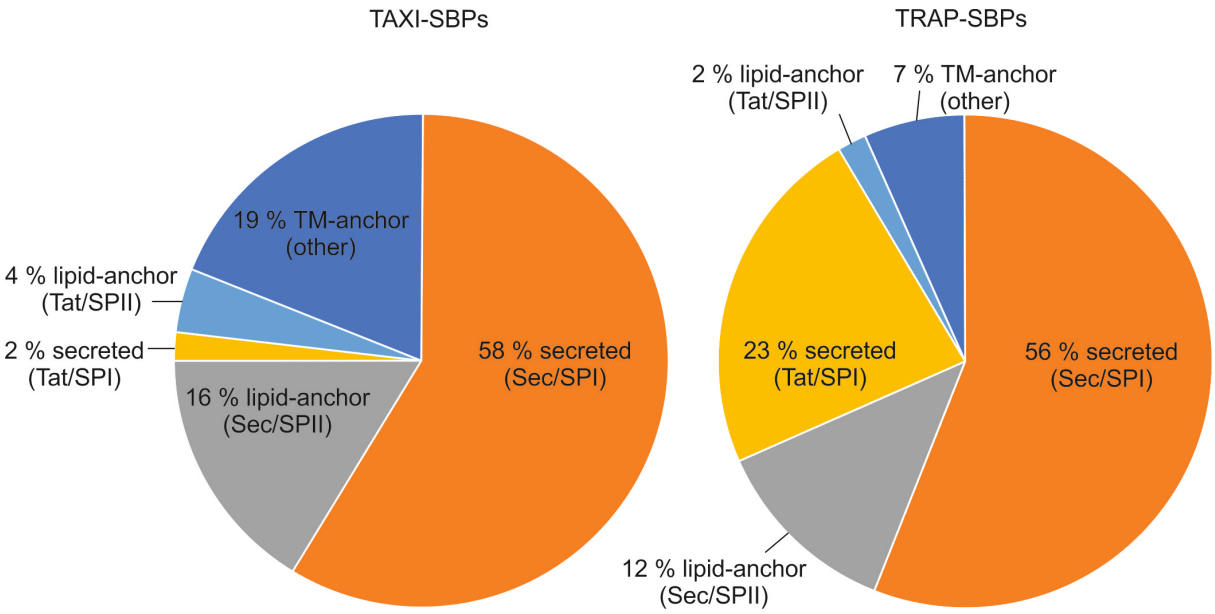


Figure 102 **Analysis of signal sequences in TAXI- and TRAP-SBPs.** Signal sequences from 215 TAXI and 636 TRAP SBP sequences were obtained from the TRAP database (Mulligan, Kelly, and Thomas 2007) and analyzed using SignalP (Teufel et al. 2022). The TAXI-SBP dataset contained protein sequences with a maximum, minimum, median, and average identity of 99.7, 8.3, 23.4, and 25.9 % and the TRAP-SBP dataset of 99.7, 7.0, 20.2, and 21.6 %, respectively. SPI, SPII, Sec, and Tat refer to Signal Peptidase I, Signal Peptidase II, Sec translocon, and the Tat translocon, respectively. Sequences that have no predicted signal sequence in SignalP but have N-terminal hydrophobic transmembrane segments are designated as "other". The pie charts were created using Microsoft Excel 16.71 (Microsoft).

We then performed protein sequence alignments derived from the *E. coli* signal peptidase I and putative *Proteus mirabilis'* signal peptidases I. The protein sequence alignments showed that the catalytic residues of the *E. coli* signal peptidase I are conserved in the putative *Proteus mirabilis'* signal peptidases I as well (Figure 103 and Figure 111) (Mark Paetzel 2013). The alignments showed an overall sequence identity of approximately 59 % (Figure 103 and Figure 111). The conservation of the catalytic residues confirms that SignalP should be able to identify Signal Peptidase I cleavage sites in TAXIPm-P if present. As SignalP did not predict a cleavage site, but instead assigned the TAXIPm-P signal sequence to an N-terminal hydrophobic transmembrane anchor (Figure 100), we strongly assume that the membrane-tethered TAXIPm-P is not only present after expression in *E. coli* but occurs in *Proteus mirabilis* as well.

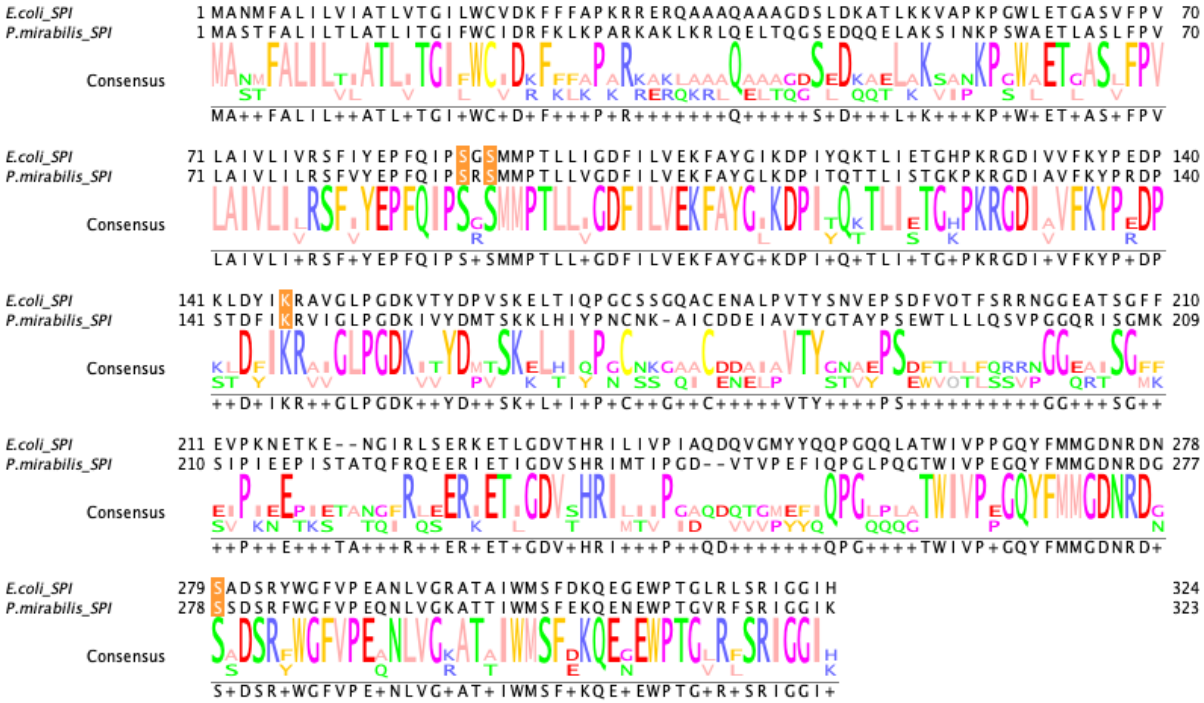


Figure 103 **Protein sequence alignment of the *E. coli* and the putative *P. mirabilis* signal peptidases I.** The catalytic residues, known from the signal peptidase I of *E. coli,* are depicted in orange (Mark Paetzel 2013). Protein sequences were aligned using Clustal Ω (Sievers et al. 2011). An extended protein sequence alignment of the *E. coli* signal peptidase I and *P. mirabilis* signal peptidases I from different strains can be found in the appendix (Figure 111, section 6). The alignment was colored and the consensus logo was created using Jalview (Waterhouse et al. 2009).

To further investigate the TAXIPm-P signal sequence and determine in detail the differences between the TAXIPm-P and a Signal Peptidase I signal sequence, such as in TRAPHi-SiaP, we analyzed the potential signal sequences of TRAPHi-SiaP and TAXIPm-P in more detail (Figure 104). SignalP predicts a Signal Peptidase I signal sequence in TRAPHi-SiaP with a probability of 66 % and cleavage between residue 23 and 24 (Figure 28, section 4.3). An initial comparison of the amino acid sequences did not argue against a cleavage of the signal sequence of TAXIPm-P. Both sequences from TAXIPm-P and the TRAPHi-SiaP appear to contain the hallmarks of a signal peptide, that is: a net positive charge in the first amino acids followed by a long hydrophobic strand that is terminated by small aliphatic amino acids (Figure 104 (Mark Paetzel 2013; M Paetzel 2019; Auclair, Bhanu, and Kendall 2012).

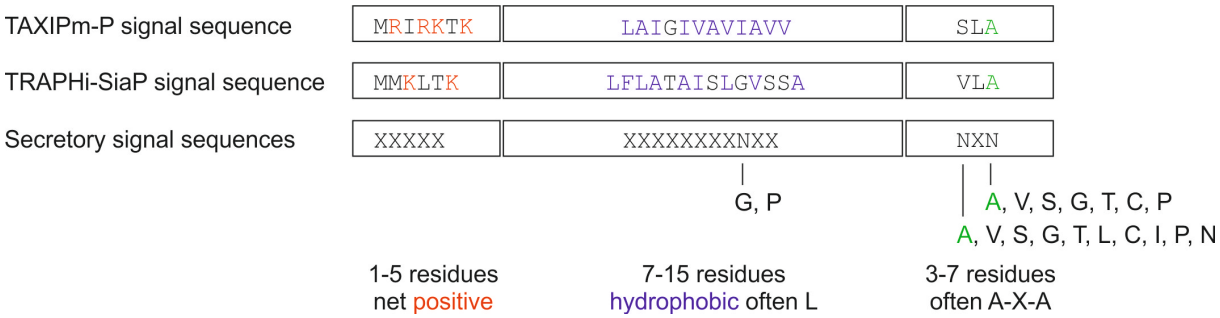


Figure 104 **Analysis of TAXIPm-P and TRAPHi-SiaP signal sequence.** The TAXIPm-P and TRAPHi-SiaP signal sequences are compared to secretory signal sequences from *E. coli*. Residues which are pre-dominant in the *E. coli* signal sequence with positive residues in the first part are colored in red, followed by a long strand of hydrophobic residues colored in purple and terminated with alanine as the most common small aliphatic residue in green (Mark Paetzel 2013).

At the terminal position in the signal sequence of TAXIPm-P is alanine, the amino acid that occurs most frequently in signal sequences for secretory SBPs (Mark Paetzel 2013; Auclair, Bhanu, and Kendall 2012; M Paetzel 2019). Serine is found at two positions before this in the signal sequence of TAXIPm-P, a position at which secreted proteins most often contain another alanine. Exchanging this serine for an alanine results in no significant change in the predicted cleavage probability (Figure 112, section 6). Glycine and proline are commonly found in the hydrophobic segment; where the signal peptide of TAXIPm-P contains an alanine at position 17 (Mark Paetzel 2013). If we substitute alanine at position 17 with glycine, no increase in the cleavage probability is observed (Figure 112, section 6). Only upon exchange of alanine at position 17 with proline is the cleavage probability for the signal peptidase I significant with 39 % (Figure 105). The cleavage probability is not as high as the 66 % probability after combination of the TRAPHi-SiaP signal peptide with the protein sequence of the TAXIPm-P (Figure 112, section 6). Nevertheless, using the *P. mirabilis* signal sequence combined with the TAXIPm-P sequence, the cleavage probability for the signal peptidase I was increased from 0 % to 39 % by one mutation in the signal peptide (A17P) (Figure 105) implying a difference between a membrane-anchored and soluble SBP might be only one amino acid. In general, the existence of both membrane-anchored and soluble forms of TAXIPm-P is likely, for example in different organisms or strains, or in evolved forms. The uptake mechanism may have been optimized by evolving from a soluble to an anchored form of the SBP for improved α-ketoglutarate uptake.

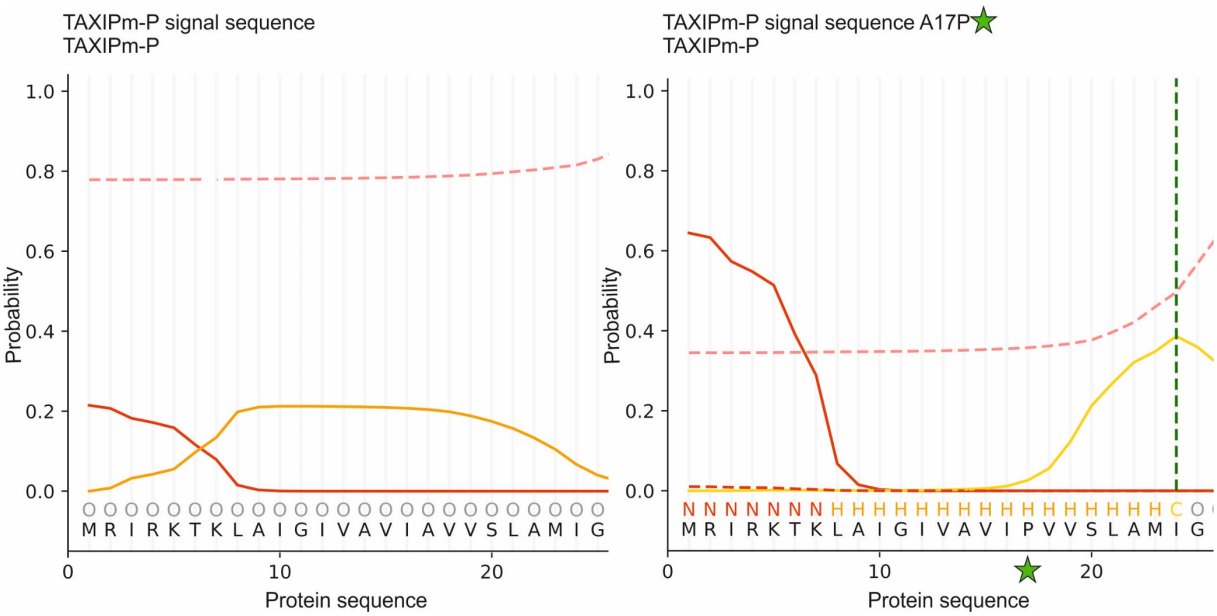


Figure 105 **Analysis of TAXIPm-P signal sequence.** The TAXIPm-P signal sequence from *Proteus mirabilis* was analyzed using SignalP with its endogenous sequence (left) and a mutation at position 17 (right). With the mutation a cleavage site probability of 39 % is predicted. The red, orange and yellow lines show the probability for a secretory pathway with Signal Peptidase I cleavage based on the N-region (n), middle region (h) and C-region (c) of the sequence, respectively. The green dashed line shows the predicted cleavage site for signal peptidase I. The light red dashed line indicates the probability for "other", which refers to no prediction of a signal sequence, but an N-terminal transmembrane segment. The green asterisk labels the position 17 where the alanine was exchanged by a proline (Teufel et al. 2022).

5.4 Coupling ion coordination

In vitro characterization has so far been performed for the N-acetylneuraminic acid transporting systems TRAPHi-SiaPQM from *Haemophilus influenzae*, TRAPPp-SiaPQM *Photobacterium profundum* and TRAPVc-SiaPQM from *Vibrio cholerae* (Davies, Currie, and North 2023; Mulligan et al. 2012). All three tripartite systems require Na^+-gradients whereby at least two but most likely three Na^+-ions are co-transported next to one substrate molecule which binds the SBP in a 1:1 stoichiometry (Mulligan et al. 2009; Müller et al. 2006; Davies, Currie, and North 2023; Mulligan et al. 2012). It was previously shown that the membrane domain is electrogenic as a negative membrane potential increased the uptake of N-acetylneuraminic acid (Mulligan et al. 2009; Davies, Currie, and North 2023). To ensure a net uptake of the single negatively charged N-acetylneuraminic acid at least two Na^+-ions are co-transported. For TRAPPp-SiaPQM from *Photobacterium profundum* it was demonstrated that most likely three Na^+-ions are co-transported as it showed a Hill coefficient of 2.7 upon variation of external Na^+ at a constant substrate concentration (Davies, Currie, and North 2023). In contrast to this, we demonstrated in this study in *in vitro* transport assays that TAXIPm-PQM from *Proteus mirabilis* is a proton-symporter (Figure 78, section 4.10.8). The α-ketoglutarate uptake by TAXIPm-PQM in *E. coli* JW2571-1 whole cells was completely abolished upon the addition of the protonophore CCCP (Figure 68, section 4.10.2). While CCCP does not only dissipate the proton gradient, but also the membrane potential and sodium gradient (Holzhüter and Geertsma 2022), we verified the strict proton coupling in *in vitro* transport assays (Figure 78, section 4.10.8). When we combined the inward-directed pH gradient with a negative membrane potential, the α-ketoglutarate uptake was slightly reduced (Figure 81, section 4.10.8). The slight reduction might be

negligible, however, as we did not observe a change in the uptake rate when combining the inward-directed pH gradient with a positive membrane potential (Figure 81, section 4.10.8). The lack of an effect of the membrane potential on the transport rates suggests that the combined charge of the substrate and the proton(s) is neutral.

To obtain insights in the differences between sodium- and proton-coupled transport systems and about their ion coordination sites, we compared protein sequences of the three sodium-coupled TRAP-SiaM transport systems, the proton coupled TAXIPm-M, and other TRAP and TAXI transport systems with unknown coupling ion (Figure 106, Figure 121 section 6). It is likely that proton ions bind at similar sites as sodium ions in homologous proteins. However, they could be coordinated by other residues or be coordinated by backbone oxygens which would decrease the relevance of the side chains (Davies, Currie, and North 2023). For the N-acetylneuraminic acid transport systems TRAP-SiaM at least two sodium ions are assumed to be co-transported as the substrate contains one negative charge. For TRAPPp-SiaQM from *Photobacterium profundum* the first sodium site is assumed in the transmembrane helix 5 in between 5a and 5b (Figure 106) (Davies, Currie, and North 2023). This Na1 site is coordinated by the five backbone carbonyl groups of S103, S106, G145, V148 and P150. A potential second sodium site is located in the transmembrane helix 11, in between 11a and 11b, coordinated by the backbone carbonyl groups of G325, G366, T369 and M372 and the sidechain hydroxyl group of T369 (Figure 106) (Davies, Currie, and North 2023). We aligned these sections of the protein sequences and found the amino acids predicted for Na^+ binding conserved within the three N-acetylneuraminic acid transporting systems, but not in the other TRAP or TAXI (Figure 106). For TRAPHi-SiaQM from *Haemophilus influenzae* two potential sodium coordinating residues were proposed in close proximity to those for TRAPPp-SiaPQM from *Photobacterium profundum*, in TM4b and TM10b (Peter et al. 2022), which are equivalently not conserved in the other TRAP or TAXI systems (Figure 106). However, it might not be necessary to have the exact amino acids conserved as mostly backbone carbonyl groups seem involved in the ion coordination. Furthermore, both potential Na^+-binding regions show a twin-proline motif between TM5a and 5b and between TM11a and 11b which was shown for other transporting systems to be essential for sodium binding (Davies, Currie, and North 2023). As the twin-proline motif is highly conserved in both TRAP and TAXI systems, independent of whether sodium- or proton-coupled, this motif might also be relevant for the coordination of protons (Figure 106).

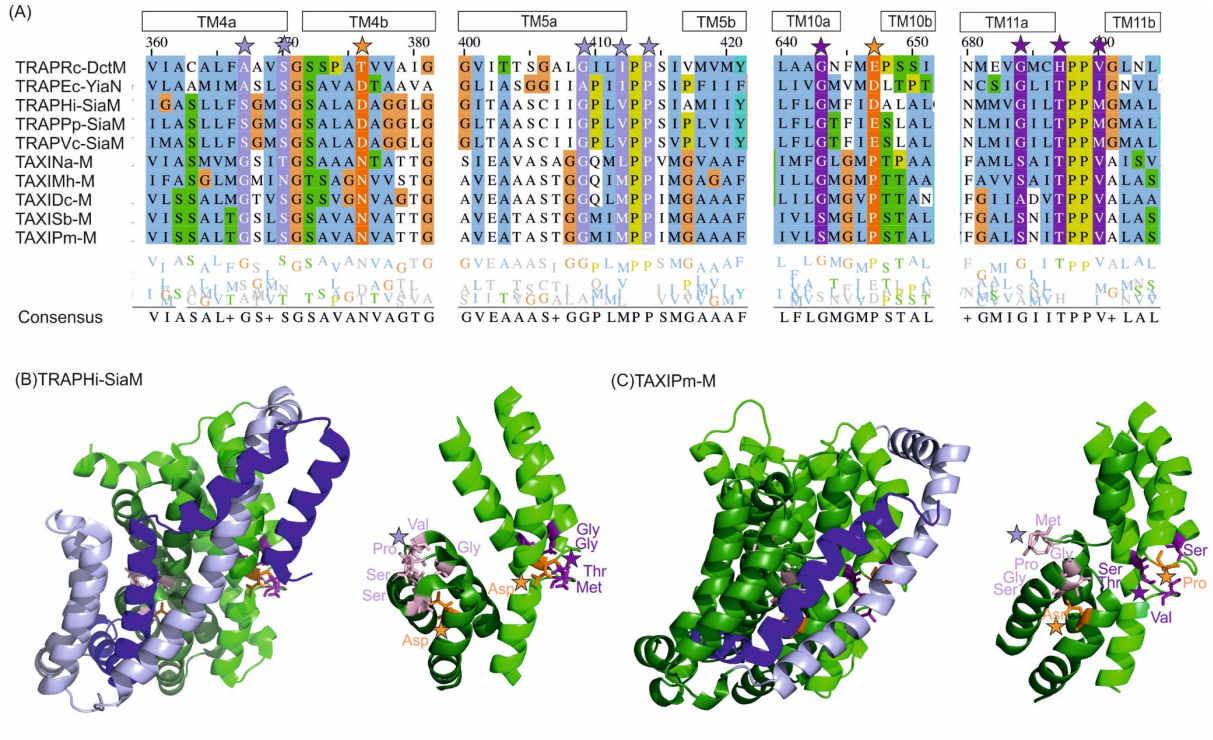


Figure 106 **Potential sodium coordinating residues in M-domains.** (A) Protein sequence alignment of TRAP and TAXI TM helices that include potential sodium coordinating residues, labeled with light and dark purple asterisks (assigned by (Davies, Currie, and North 2023)) and orange asterisks (assigned by (Peter et al. 2022)). The twin-proline motif between TM5a and 5b as well as TM11a and TM11b is displayed in yellow and purple. The aligned sequences are shown in clustalx coloring using Jalview (Waterhouse et al. 2009). (B) Structure of TRAPHi-SiaM with the scaffold region in the front (left) and the TM helices with potential sodium coordinating residues (right). (C) Alphafold model of TAXIPm-M with the scaffold region in the front (left) and the TM helices with highlighted residues that correspond to the positions that are responsible for sodium coordination in TRAPHi-SiaM (right). Transport domains 1 and 2 are colored in light and dark green and scaffold domains 1 and 2 in light and dark blue, respectively. The structures were created using Pymol 1.8.4.1 (Schrödinger, LLC).

To further investigate the coupling ion coordination and to compare data on coordination sites which originates from another elevator type, but not an N-acetylneuraminic acid transporting system, we compared the amino acid residues of the sodium coupling sites in the elevator type dicarboxylate transporter VcINDY from *Vibrio cholerae* with those of TRAPHiSiaQM, TRAPPp-SiaQM and TRAPVc-SiaM and TAXIPm-QM. Based on their topology similarities, all contain the scaffold as well as the transport domains 1 and 2 with two hairpins. The dicarboxylate import by VcINDY is coupled to three sodium ions whereby Na1 and Na2 bind to the tips of the two hairpins (Sauer et al. 2022). After the sodium ions are bound, VcINDY undergoes a conformational change that allows the substrate to bind as well. For this, both hairpin tips contain the same residues Ser-Asn-Thr which bind the carboxylate groups of the substrate (Sauer et al. 2022). The residues Ser-Asn-Thr (SNT) responsible for Na^+-coupling in VcINDY remain absent in the three Na^+ coupled N-acetylneuraminic acid transporter TRAPHiSiaQM, TRAPPp-SiaQM and TRAPVc-SiaM (Figure 107). Nevertheless, it was suggested that Na1 and Na2 bind at the tips of the two hairpins as in VcINDY (Davies, Currie, and North 2023; Peter et al. 2022).

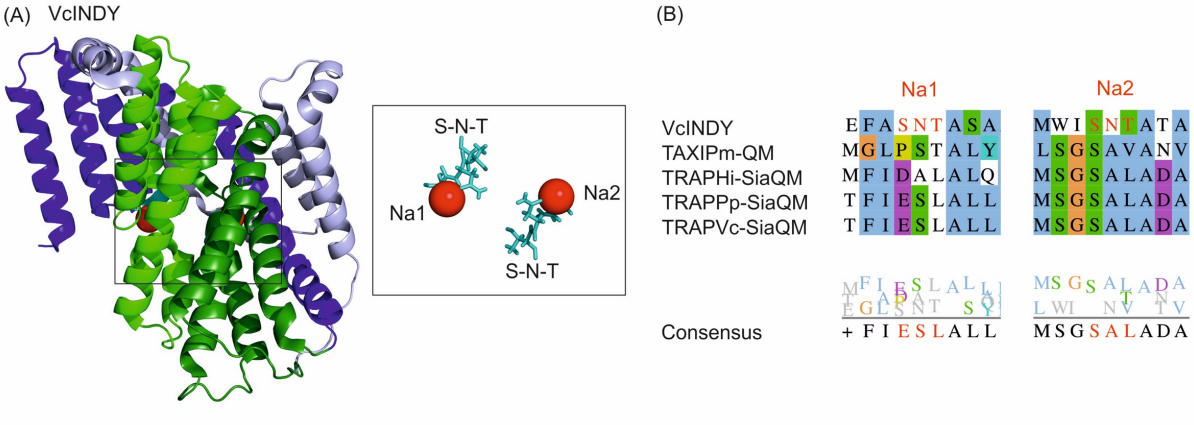


Figure 107 **Sodium coordinating residues in VcINDY from *Vibrio cholerae*.** (A) Structure of VcINDY with the scaffold region in the front (left) and the sodium binding residues SNT highlighted (right) (PDB code: 7T9G). Transport domains 1 and 2 are colored in light and dark green and scaffold domains 1 and 2 in light and dark purple, respectively. (B) Section of protein sequence alignment from VcINDY and TRAP QM domains where VcINDY binds two sodium ions. Aligned sequences are shown in clustalx coloring using Jalview (Waterhouse et al. 2009). The structure was created using Pymol 1.8.4.1 (Schrödinger, LLC).

A general conclusion regarding amino acid residues responsible for coupling ion coordination in TRAP and TAXI transport systems and more specifically in TAXIPm-QM from *Proteus mirabilis* cannot be provided yet. To unambiguously identify the proton binding sites in TAXIPm-QM, a combined structural approach with all-atom MD simulations and Brownian dynamics models might be necessary (Erban 2016). One might speculate, however, that the proton ion(s) bind to the tips of the hairpins in TAXIPm-QM simultaneously as the sodium ions bind in VcINDY from *Vibrio cholerae*. For this reason, we superimposed the structure of VcINDY from *Vibrio cholerae* which contains two bound Na^+-ions on the alphafold model of TAXIPm-QM (Figure 108). The hairpins in TAXIPm-QM are predicted to be architecturally similar and in close proximity to those in VcINDY, but they do not align in the superimposition of the complete transporters (Figure 108). Therefore, it remains a speculation that the proton-ions in TAXIPm-QM might bind to the tips of the hairpins similarly to the Na^+-ions in VcINDY.

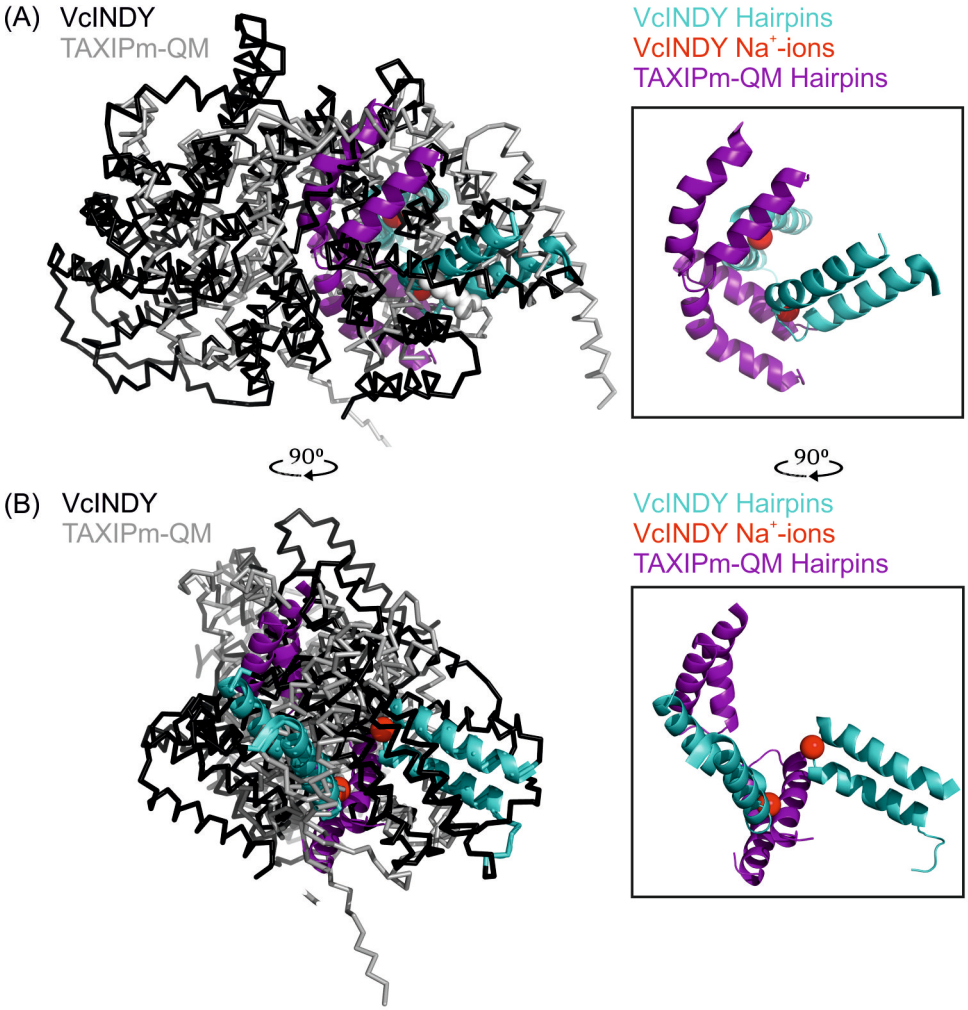


Figure 108 **Superimposition of VcINDY on the alphafold model of TAXIPm-QM.** (A) Front view of the alphafold model of TAXIPm-QM from *Proteus mirabilis* with superimposition of the dimeric structure of VcINDY from *Vibrio cholerae* (PDB code: 7T9G). (B) Side view of the alphafold model of TAXIPm-QM from *Proteus mirabilis* with superimposition of the dimeric structure of VcINDY from *Vibrio cholerae* (PDB code: 7T9G). The hairpins from VcINDY are colored teal, those of TAXIPm-QM in purple. The structure of VcINDY contains Na^+-ions which are colored in red. The structures and superimposition were created using Pymol 1.8.4.1 (Schrödinger, LLC).

5.5 Unidirectionality in secondary transporters

Our transport assays showed that TAXIPm-PQM facilitates unidirectional import of α-ketoglutarate as the transport behavior was not affected upon the addition 100-fold excess of unlabeled substrate. Such high substrate concentrations would reverse the transport in conventional secondary transporters, thereby enabling efflux (Mulligan et al. 2009).

Unidirectional transport in general describes import or export only, which is realized through different strategies. Conventional secondary transporters operate bidirectional whereby the direction of transport is oriented towards the direction of the sum of all gradients. Typical for secondary transporters is a bidirectional alternating access, which is well-described for LeuT folded proteins. Their outward-facing structure is almost identical to the inward-facing one. The flexibility gives the substrates the ability to easily pass through these secondary transporters in both directions, based on the direction of the proton gradient (Bianchi et al. 2016).

To facilitate unidirectional substrate uptake only, TRAP transporting systems use a SBP for high specificity and affinity. The SBPs are implemented in the periplasm of Gram-negative or outside the membrane of Gram-positive cells and archaea, thereby restricting TRAP transport systems to import only (Figure 3) (van der Heide and Poolman 2002; Hofmann et al. 2019). In *in vitro* transport assays unidirectional import was shown for the transport of N-acetylneuraminic acid by TRAPHi-SiaPQM from *Haemophilus influenzae*. When an excess of the substrate was supplied to the outside milieu to generate a strong inwards directed substrate gradient no exchange was observed. No radioactive substrate was effluxed, as would be expected if TRAPHi-SiaPQM were a bidirectional transporter. In contrast to that, in the presence of the conventional SBP independent N-acetylneuraminic acid secondary transporter nanT these high substrate concentrations on the outside led to efflux (Mulligan et al. 2009). Furthermore, it was shown that the membrane domain of TRAPHi-SiaPQM exhibits an absolute dependence on the SBP. No transport was observed across the membrane domain in the absence of the SBP, indicating that the affinity of the membrane domain for the substrate itself is very low or non-existent (Mulligan et al. 2009, 2012). This observation might apply in general for all SBP dependent transport systems as it was previously also shown for primary SBP dependent transporters (Shuman 1982). Radioactive *in vitro* uptake assays showed that due to the dependency on the SBP, only import is supported. The SBP was added from the outside as it would also occur in the cell only in the periplasm. Surprisingly, an efflux of the substrate from the proteoliposomes could be generated after a large excess of unliganded SBP was supplied, thus showing that the SBP is responsible for the transport direction of the substrate. Additionally, this demonstrated that the unliganded version can adapt the same conformational state as the liganded SBP thereby being able to bind to the membrane domain (Mulligan et al. 2009). However, such high SBP concentrations are unlikely under physiological conditions (Mulligan et al. 2009). In our transport assays with TAXIPm-PQM, we used a membrane tethered SBP through which we automatically achieved high local SBP concentrations on the membrane. As for TRAPHi-SiaPQM we did not observe an effect of the transport behavior upon the addition of a large excess of unlabeled substrate which demonstrates that transport systems with membrane tethered SBPs operate unidirectionally as well (Figure 82, section 4.10).

Next to TRAP transporting systems, it was shown for the ABC transporter BtuCD-BtuF that unidirectionality is achieved by using a SBP which delivers vitamin B12 to the membrane domain with high affinity which implies that all SBP dependent transporting systems operate unidirectionally. For BtuCD-BtuF the transport mechanism is well described as a sequence of conformational changes that allow import only. The binding of BtuF to BtuCD happens through closure followed by partial opening of the NBDs to lower the affinity for ATP which can then bind; subsequent ATP hydrolyzation leads to an opening of NBDs and a release of BtuF (Yang et al. 2018).

For secondary transporters unidirectional transport is not typical, but it is important for slow transporters like the eucaryotic Lyp1 to enable accumulation when substrates are rare (Bianchi et al. 2016). The uptake of basic amino acids by Lyp1 in *S. cerevisiae* could in essence operate bidirectionally but proceeds unidirectionally as efflux is limited through a 3 to 4 magnitude higher K_M

when compared to import. In contrast the procaryotic homolog LysP shows the same K_M for import and export and thus operates bidirectionally (Bianchi et al. 2016).

Unidirectionality is not limited to import, as export is important when it would be lethal for the organism if certain toxins were imported. It was suggested that the secondary RND transporter mdfA exports toxic compounds unidirectionally only (Sigal et al. 2006). Unidirectional export is already well described for ATP dependent P-glycoprotein homologs in eukaryotes and prokaryotes, also to protect the cells from toxins. For this, unidirectionality is mechanistically enforced by allowing for transport only when the NBDs are asymmetrically occupied (Xu, Seelig, and Bernèche 2017). Structural parts like inverted repeats or additional loops might involve rectification. It was shown for heavy-metal detoxification by the ABC transporter NaAtm1 from *Novosphingobium aromaticivorans* that toxification through high concentrations is avoided by realizing unidirectionality thermodynamically (Fan, Kaiser, and Rees 2020). To save additional conformational changes and to prevent either uptake or efflux, it was shown that MgADPVO4 stabilizes kinking of TM6 in NaAtm1 to eliminate the substrate-binding cavity. This thermodynamic stabilization of a particular conformation seems to be a common phenomenon, as parallels can be seen in TmrAB from *Thermus thermophilus* (Fan, Kaiser, and Rees 2020).

Furthermore, unidirectional transport can be found in channels and pumps. In general, P-type ATPases facilitate cation transport to maintain ion gradients. One special combination has formed with the procaryotic ATP-driven K^+-pump KdpFABC to allow K^+ import even when external concentrations are very low. Within this multi-subunit complex, the P-type ATPase KdpB merged with the channel-like protein KdpA forming a pair of two half channels to enable highly effective transport. K^+ ions move from KdpA to KdpB through an intersubunit tunnel which triggers ATP hydrolysis by KdpB (Stock et al. 2018; Silberberg et al. 2021).

However, unidirectionality may be reversible through evolution for all the systems mentioned. This reversibility has been shown for the human peptide ABC transporter TAP which is a strict unidirectional importer. Efflux is still not seen when the outward substrate concentration gradient is generated by preloading proteoliposomes with substrate. Nevertheless, TAP can be modified to a bidirectional transporter by a single mutation which prevents ATP-binding thereby demonstrating that ATP hydrolysis is needed to allow active unidirectional transport but not for the transport in general (Grossmann et al. 2014). For SBP dependent systems like TRAP transporting systems, however, it remains unlikely that the transport mechanism can be reversed, since the SBP occurs in the periplasm only. We indeed observed *in vitro* import and export of α-ketoglutarate by TAXIPm-PQM when the SBP was present at the inside and outside of the QM proteoliposomes (Figure 78, section 4.10.8). Therefore, we would assume the membrane domain to operate bidirectional if it would be a SBP independent transporter. As the membrane domain TAXIPm-QM is SBP dependent, however, the periplasmic SBP TAXIPm-P determines the unidirectional transport of α-ketoglutarate to import only.

5.6 Relevance of the membrane domain Q

The structural relevance of the small membrane domain Q in TRAP transporter is evident as it extends to the large membrane domain M to ensure a firm anchoring within the membrane (Peter et al. 2022). Therefore, TRAP transporters do not need to oligomerize. Until recently, it was believed that all elevator type proteins appear in oligomeric states, more specifically as dimers or trimers, for example GltPh from the SLC1 family or hCNT3 proteins from the SLC28 family (Holzhueter and Geertsma 2020; Grewer, Gameiro, and Rauen 2014; Stecula et al. 2017). However, TRAP transport systems have been assigned to the group of elevator type proteins as well but appear in a monomeric state (Peter et al. 2022). The structural relevance of the oligomerization in general is evident as it ensures a firm anchoring within the membrane by forming a rigid scaffold as it was shown for the SLC26 family, for example (Chang et al. 2019). The assignment of TRAP transport systems to the elevator type proteins is based on structural data of the membrane domain TRAPHi-SiaQM as its transport domain TRAPHi-SiaM shows a high similarity to the sodium-coupled dicarboxylate transporter VcINDY (Peter et al. 2022). Structurally the oligomerization is not necessary in TRAP transporter as they contain Q which serves as extension of the scaffold domain in M thereby providing stability for the whole protein complex and allowing this protein group to operate as monomers (Davies, Currie, and North 2023). To date, there is no explanation why the membrane domain is composed of the two units Q and M rather than having one unit and assembling as oligomers, as observed in other elevator type proteins (Peter et al. 2022; Drew and Boudker 2016). The membrane oxidoreductase CcdA from *Thermus thermophilus* was also suggested to present a monomeric protein with an elevator type transport mechanism (Garaeva and Slotboom 2020; U. Zhou and Bushweller 2018). Furthermore, the bile acid transporter ABST was shown to have a monomeric structure, but to combine a moving barrier with the elevator type mechanism as it has no fixed barrier (X. Zhou et al. 2014; Garaeva and Slotboom 2020). Biochemical data, however, indicate a dimeric structure (Bocian-Ostrzycka et al. 2017). Besides the structural relevance of oligomerization or the additional Q domain, both most likely have functional relevance as well. At the first glance NhaA seems to be an exception which is active as a monomer under most tested conditions. However the functional relevance of the monomer remains unclear, as it was shown that the dimeric NhaA enables higher transport rates when exposed to high salt and pH (Rimon, Tzubery, and Padan 2007). The higher transport rates which occur only after dimerization of NhaA suggest that there is indeed functional significance in the oligomerization for the successful transport mechanism.

Previously, an additional chaperon function of the Q- for the M-domain was suggested, an unlikely characteristic, since there are also TRAP systems, for example in *Roseovarius* sp. 217, in which the gene encoding Q is located in the operon after the gene encoding M (Mulligan, Kelly, and Thomas 2007). For elevator type proteins in general, different transport mechanisms are described containing symport, exchange, and uniport systems. Also, within protein families, for example the SLC26 family, the transport mechanism varies as they function as exchangers or uniporters, some enable fast uniport and others, such as ion channels, operate more slowly. SLC26A5 is an exception as it works as a motor protein (Rapp, Bai, and Reithmeier 2017; Walter, Sawicka, and Dutzler 2019; Dallos and

Fakler 2002). The TRAP transporter TRAPHi-SiaPQM from *Haemophilus influenzae* is a symporter which imports N-acetylneuraminic acid unidirectionally due to its SBP dependency (Mulligan et al. 2009).

5.7 Characterization of the SBP-membrane domain interaction

So far, it remains elusive how the SBP in TRAP transport systems interacts with the membrane domain. Numerous structures were solved for TRAP SBPs (Vetting et al. 2015) and recently, structures have been solved for the membrane domains of TRAPHi-SiaQM from *Haemophilus influenzae* and TRAPPp-SiaPQM from *Photobacterium profundum* (Peter et al. 2022; Davies, Currie, and North 2023). So far, however, only structures of the membrane domain and SBP could be successfully resolved separately and not in complex (Peter et al. 2022; Davies, Currie, and North 2023). In the absence of a resolved complex, only speculation about the interface and the docking site for the SBP is possible. Bioinformatic and biochemical studies have suggested that TRAP SBPs interact with both membrane domains, Q and M, during the transport cycle. The structure of TRAPPp-SiaP from *Photobacterium profundum* with bound N-acetylneuraminic acid was used to model the complex with the corresponding TRAPPp-SiaQM. For this modeling the algorithms RaptorX, Gremlin and Alphafold predicted similar residues to be involved in the interaction which mainly included surface residues of the scaffold domain. These surface residues show a high degree of sequence conservation. It was predicted that the two periplasmic loops of the Q domain also interact with the SBP (Davies, Currie, and North 2023). Previously performed superpositions of TRAPHi-SiaQM from *Haemophilus influenzae* and VcINDY from *Vibrio cholerae* suggest that the N-terminal part of TRAPHi-SiaP binds to the scaffold and the C-terminal part to the transport domain of TRAPHi-SiaQM (Peter et al. 2022). To verify this binding pattern, growth assays were performed for which TRAPHi-SiaP and TRAPHi-SiaQM variants were constructed with mutations in the amino acid residues most likely involved in the interaction. Ten variants significantly affected the growth which implies an importance for interaction but could also result from a changed protein conformation (Peter et al. 2022). The importance of especially the periplasmic loops was confirmed in this study. To investigate the interaction between the SBP and membrane domain, we constructed variants of TAXIMh-PQM from *Marinobacter hydrocarbonoclasticus* and TRAPHi-SiaPQM from *Haemophilus influenzae*. The exchange of five residues in the first periplasmic loop of the Q membrane domain of TAXIMh-QM abolished the transport of fumarate into *E. coli* whole cells completely (Figure 67, A section 4.9.4) It was shown that this abolished transport can neither be attributed to decreased expression levels nor to decreases in functional expression of the membrane domain due to a lack of folding (Figure 67, B section 4.9.4). As the SBP is on the same operon, we expect TAXIMh-P to be expressed as well. It is possible that the folding of the membrane domain is altered due to the mutations or that the membrane domain cannot adopt a required conformational change. However, the abolished transport suggests that the mutations in the docking site for the SBP TAXIMh-P are responsible for the abolished transfer of the substrate to the membrane domain.

While mutations in TAXIMh-QM abolished the fumarate transport by TAXIMh-PQM, we observed negligible differences in the N-acetylneuraminic acid transport by TRAPHi-SiaPQM upon inserting

equivalent mutations (Figure 53, A and Figure 54, section 4.8.1). As the transport activity of the TRAPHi-SiaPQM variant did not correlate with expression levels of TRAPHi-SiaPQM (Figure 53, B, section 4.8.1), we concluded that TRAPHi-SiaP delivers the substrate to another *E. coli* transport system. Candidates for the interaction with TRAPHi-SiaP are different TRAP or SBP dependent ABC transport systems. Therefore, we constructed a knock-out for the only endogenous *E. coli* TRAP transporter, YiaMNO. Subsequent N-acetylneuraminic acid uptake still showed uptake when the SBP TRAPHi-SiaP was present, indicating that SiaP facilitates the transport of N-acetylneuraminic acid with a SBP dependent transporter other than TRAPEc-YiaMNO (Figure 59, section 4.8.2).

In TRAP transport systems the interaction between membrane domains and their corresponding SBP was assumed to be specific as it was previously shown that the TRAPHi-SiaQM membrane domain from *Haemophilus influenzae* cannot facilitate transport using the SBP of TRAPVc-SiaPQM from *Vibrio cholerae* (Mulligan et al. 2009). This is surprising given that they share identity of 49.83 % in the SBPs, 30.91 % in the Q and 55.42 % in the M domain (Figure 114, section 6) (Sievers et al. 2011). It was previously concluded that for the interaction of TRAPHi-SiaP and TRAPHi-SiaQM from *Haemophilus influenzae*, certain residues are essential (Peter et al. 2022). Possibly a transporter that contains the same amino acids can interact with the same SBP. Since there is only partial conservation of the described amino acids between TRAPHi-SiaQM and TRAPVc-SiaQM, this explains most likely why TRAPHi-SiaP cannot interact with TRAPVc-SiaQM (Figure 113 and Figure 114, section 6).

However, membrane domains of TTT transporters have been suggested to allow binding of multiple different SBPs as the gene ratio of SBPs to membrane domains in α- and β-proteobacteria exceeds multiple times. In *B. pertussis* two sets of genes encoding homologs to the large membrane domain *TctA* from the *TctCBA* citrate uptake system of *Salmonella typhimurium* were found next to 81 homologs to the corresponding SBP *TctC* (L. T. Rosa, Bianconi, et al. 2018). It was assumed that a range of these SBPs can interact with the same membrane domain which can as a result transport different substrates (Hosaka et al. 2013). This would not be the first assumption that transporter in general might interact with different SBPs (Antoine et al. 2003). So, the specific interaction is the key for transport, which is carried out by the membrane domain. Furthermore, in *Mageeibacillus indolicus* a gene encoding the large membrane domain of a TTT transporter was found, but no corresponding gene encoding the SBP was identified. This suggests that either the membrane domain can facilitate transport alone (Winnen, Hvorup, and Saier 2003) or interact with a SBP of another protein family (Antoine et al. 2003). Similar observations were made for the ECF ABC transporters. These interact with an S-component responsible for substrate specificity to enable import instead of an SBP. ECF transporters are divided into two groups, of which proteins in group I receive substrates from a single S-component and proteins in group II receive substrates from multiple S-components. In a bioinformatic analysis, 1,927 ECF transporters could be assigned to group I and 787 ECF transporters to group II. The latter includes 4,387 S-components which are unrelated in sequence and deliver a diversity of substrates (Erkens et al. 2012; Rempel, Stanek, and Slotboom 2019).

Considering all data, we conclude, that TRAPHi-SiaP may interact with a membrane domain that normally transports a substrate other than N-acetylneuraminic acid. This domain most likely belongs

to a SBP dependent ABC importer as we showed in transport assays that the only endogenous TRAP membrane domain in *E. coli*, YiaMN, is not responsible for the N-acetylneuraminic acid transport (Figure 59, section 4.8.2). ABC transporters comprise the largest protein group in *E. coli* with 50 importers out of a total of 65 functional ABC transport systems (Moussatova et al. 2008). Therefore, it is likely that one of these functions with TRAPHi-SiaP.

5.8 Conformational stabilization by nanobodies

Nanobodies were generated against TAXIPm-QM and a wide range of nanobodies specific for TAXIPm-QM were selected through the ELISA (section 3.5.3). Since higher signal intensities in the ELISA do not necessarily correspond to higher affinities for TAXIPm-QM but are more likely the result of different expression levels, a variety of nanobodies was chosen from those. In general, we demonstrated high expression levels for the nanobodies which reduced but still allowed high expression levels of the co-expressed TAXIPm-QM protein (Figure 91, section 4.11.5). Several of the selected nanobodies inhibited α-ketoglutarate uptake into *E. coli* whole cells which demonstrates their binding to periplasmic epitopes of TAXIPm-QM (Figure 93, section 4.11.6). Some nanobodies, such as B2H5 significantly reduced the transport rate next to well expression of both nanobody and TAXIPm-QM (sections 4.11.5 and 4.11.6). These nanobodies will be promising candidates to aid structural elucidation of TRAP-TAXI proteins at the MPI CBG, Dresden (Eric Geertsma Group, www.mpi-cbg.de). We assume that the binding of the nanobodies could either facilitate conformational trapping of TAXIPm-QM or reduce the interaction between the SBP and the QM-domain.

5.9 Concluding remarks and outlook

It can be concluded from this study that TAXIPm-PQM from *Proteus mirabilis* is a proton coupled α-ketoglutarate transporter that requires a membrane-tethered SBP (sections 4.3, 4.4 and 4.10). *In vitro* transport assays showed the dependence of TAXIPm-QM on a membrane-tethered SBP, likely caused by an enriched local concentration of SBP at the membrane (Figure 79, section 4.10.8). However, the mechanistic differences between a soluble and a membrane tethered SBP should be investigated in more detail. For this and in general to further characterize especially TRAP-TAXI proteins, structures should be solved. So far, structures of TRAP membrane domains have been solved for two N-acetylneuraminic acid transporting systems only (Peter et al. 2022; Davies, Currie, and North 2023) and the information gained from these structures might not necessarily apply to TRAP-TAXI transporting systems. For both TRAP membrane domains, the structures were solved without SBP, hence the interface can only be speculated at this stage. It has not yet been achieved to structurally resolve a complex of membrane domain and SBP, likely due to low affinities as a K_D-value of approximately 400 µM was determined for the binding of TRAPPp-SiaP to TRAPPp-SiaQM in the presence of N-acetylneuraminic acid (Davies, Currie, and North 2023). Next to low affinities it is possible that TRAPPp-SiaM preferentially occurs in the inward-facing conformation which would then not be able to interact with TRAPPp-SiaP (Davies, Currie, and North 2023). For the purpose to also generate outward-facing conformations, nanobodies could aid to stabilize different

conformations (section 4.11). Should the variety of conformation stabilizing nanobodies be increased, nanobody selections can be performed in the presence of different substrate concentrations (Arkhipova, Guskov, and Slotboom 2020). In the case of our tripartite system, the SBP should additionally be supplied with the varying substrate concentrations as affinities of the membrane domain for the substrate are non-existent or poor (Davies, Currie, and North 2023). Since TAXI SBPs belong to a different class than TRAP SBPs, speculations about the interface cannot be reliably applied to TRAP-TAXI systems. Therefore, it is necessary to solve structures not only from conventional TRAP but also from TRAP-TAXI systems and to solve them in complex of the membrane domain with the SBP. Furthermore, a complex should reveal whether, in addition to liganded SBP, unliganded SBP can also bind to the membrane domain, which has so far only been assumed. Furthermore, ion coupling should be investigated in more detail, since so far only assumptions are possible about both sodium coupling and proton coupling in TRAP transport systems. Structures may provide indications, but ion binding sites, especially for protons can often not be precisely assigned. Therefore, it may be advantageous to combine a structural approach with all-atom MD simulations and Brownian dynamics models (Erban 2016). Furthermore, it was shown in this study that TRAPHi-SiaP facilitates N-acetylneuraminic acid uptake into *E. coli* whole cells in absence of the corresponding membrane domain TRAPHi-SiaQM (Figure 54, section 4.8.1). This observation suggests that TRAPHi-SiaP and SBPs in general might be able to interact with different membrane domains and thereby determine the substrate specificity. It remained undetermined with which membrane domain TRAPHi-SiaP interacts. However, it is most likely a SBP dependent transporter, such as an ABC transporter since the only endogenously occurring TRAP transporter in *E. Coli*, YiaMNO, was knocked out and can thus be excluded (Figure 59, section 4.8.2). We conclude that in general further tripartite systems should be characterized structurally and functionally to understand which properties are common to individual proteins and which apply to the whole family.

6 Supplementary data

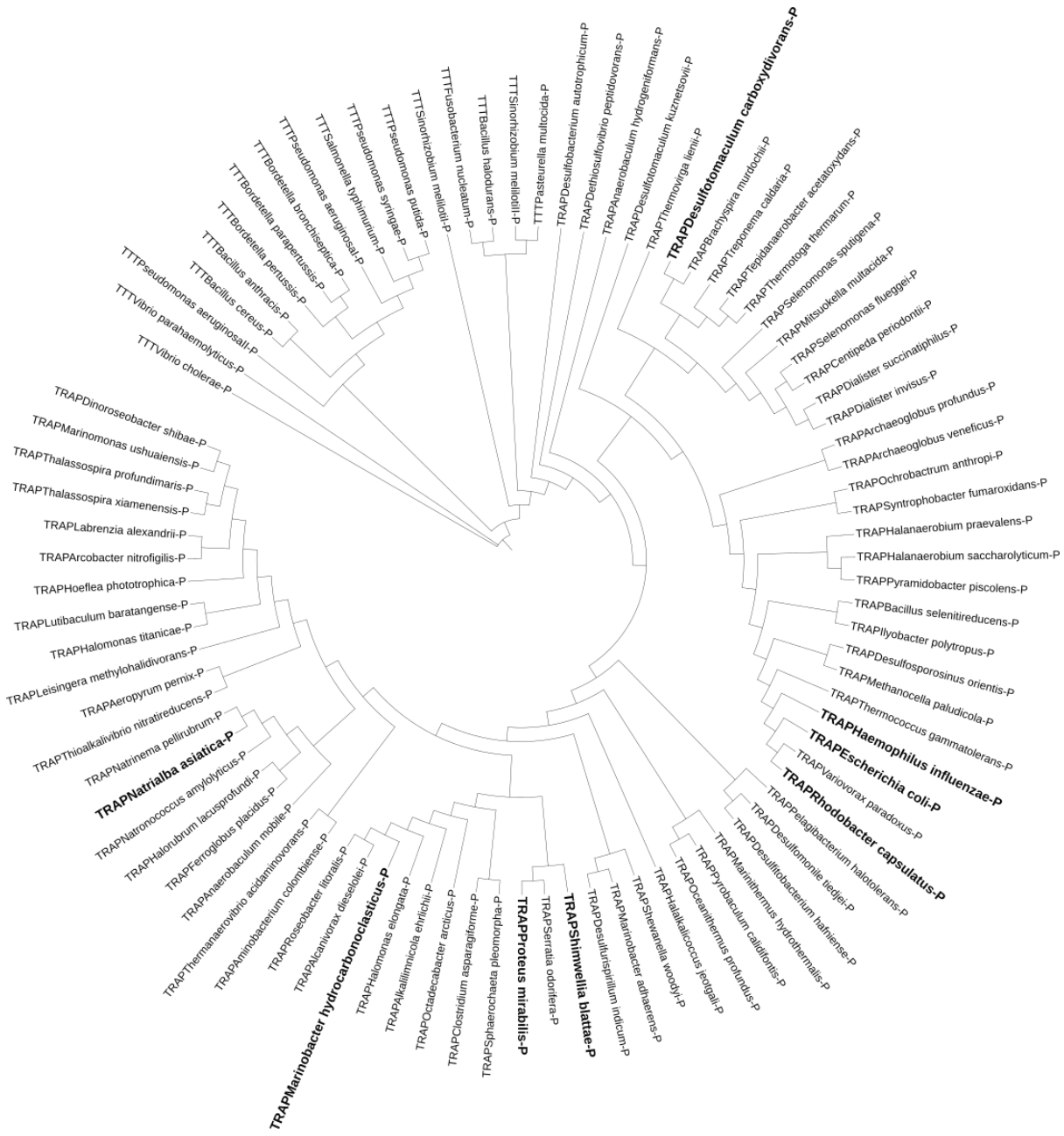


Figure 109 **Phylogenetic tree of DctP homologs from *Rhodobacter capsulatus*.** To evaluate similarities between TRAP-SBP and illustrate the distance from TTT SBP, sequence alignments of DctP homologs were carried out using Clustal Ω (Sievers et al. 2011) and a phylogenetic tree was created using NGPhylogeny (Lemoine et al. 2019).

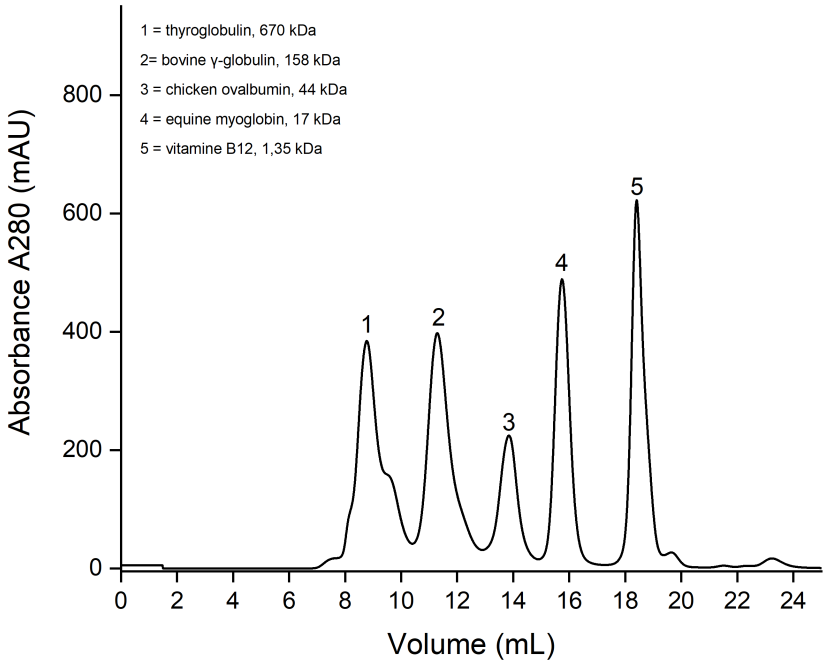
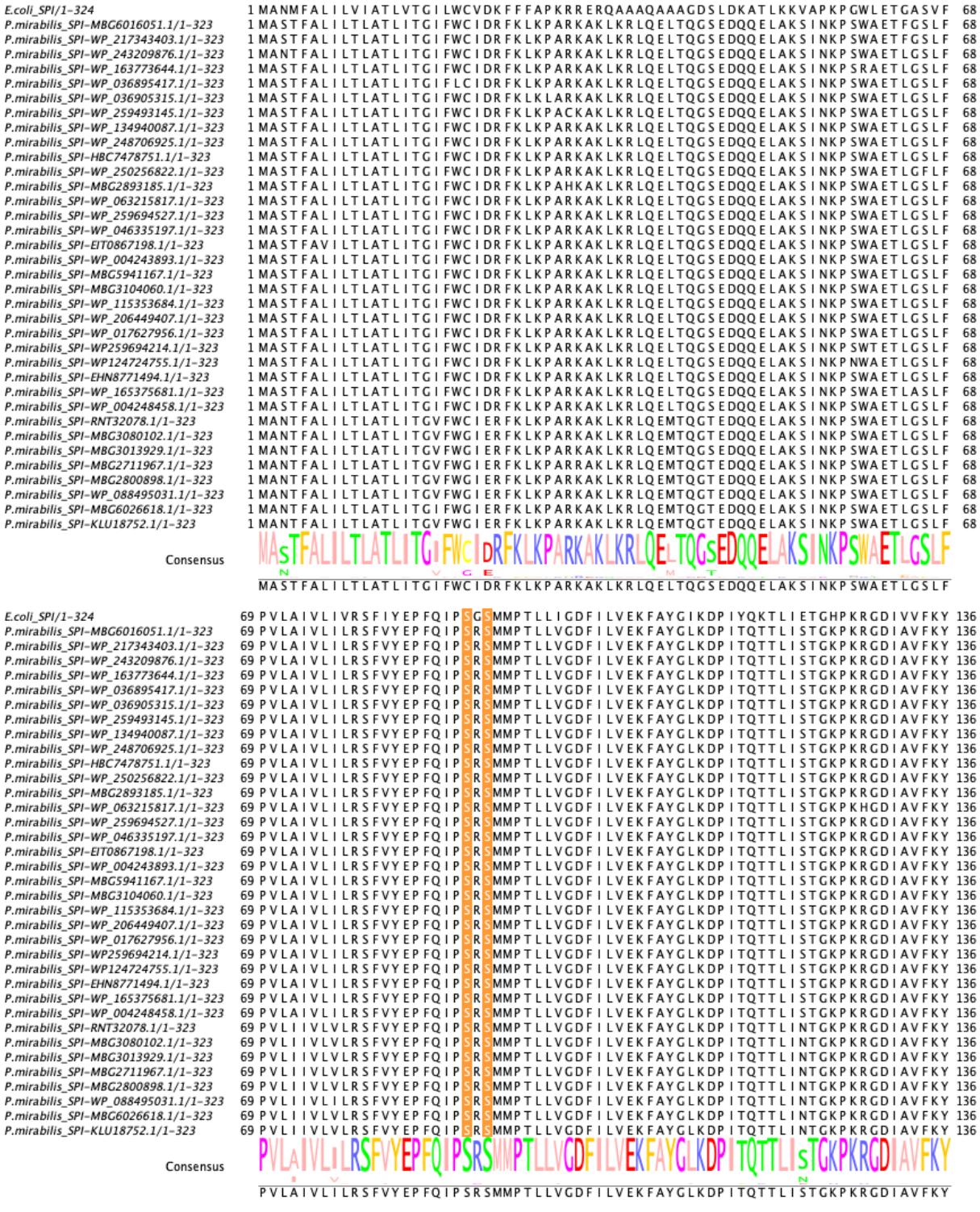
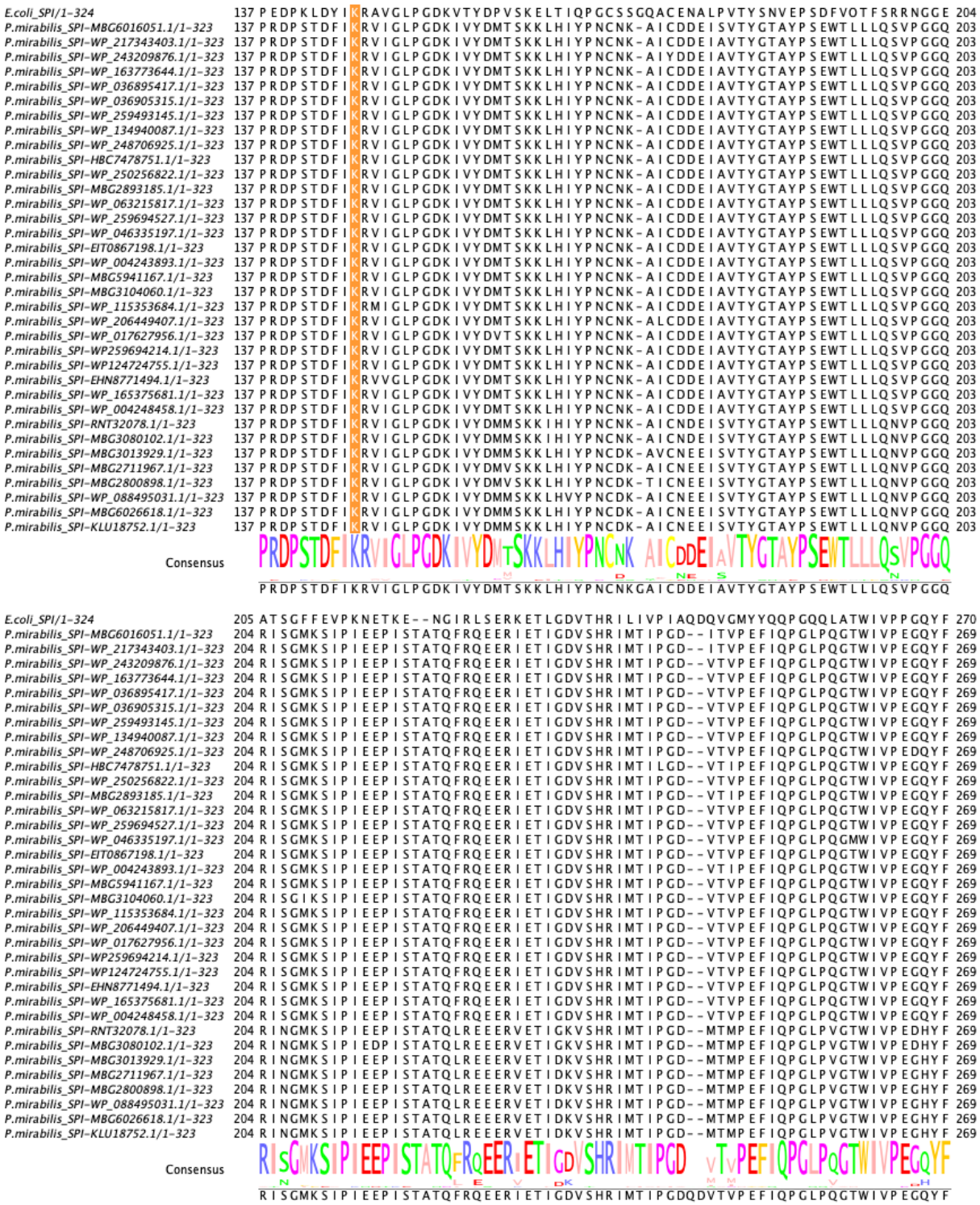


Figure 110 **SEC chromatogram of protein solution standard.** A SEC run was performed on a Superdex 200 30/100 GL Increase column with a gel filtration protein solution standard (Bio-Rad 1511901). The plot was created using OriginPro 2019b (OriginLab Corporation).


E.coli_SPI/1-324 1 MANMFALILVIATLVTGILWCVDKFFFAPKRRERQAAAQAAAGDSLDKATLKKVAPKPGWLETGASVF 68
P.mirabilis_SPI-MBG6016051.1/1-323 1 MASTFALILTLATLITGIFWCIDRFKLKPARKAKLKRLQELTQGSEDQQELAKSINKPSWAETFGSLF 68
P.mirabilis_SPI-WP_217343403.1/1-323 1 MASTFALILTLATLITGIFWCIDRFKLKPARKAKLKRLQELTQGSEDQQELAKSINKPSWAETFGSLF 68
P.mirabilis_SPI-WP_243209876.1/1-323 1 MANTFALILTLATLITGIFWCIDRFKLKPARKAKLKRLQELTQGSEDQQELAKSINKPSWAETLGSLF 68
P.mirabilis_SPI-WP_163773644.1/1-323 1 MASTFALILTLATLITGIFWCIDRFKLKPARKAKLKRLQELTQGSEDQQELAKSINKPSRAETLGSLF 68
P.mirabilis_SPI-WP_036895417.1/1-323 1 MASTFALILTLATLITGIFLCIDRFKLKPARKAKLKRLQELTQGSEDQQELAKSINKPSWAETLGSLF 68
P.mirabilis_SPI-WP_036905315.1/1-323 1 MASTFALILTLATLITGIFWCIDRFKLKLARKAKLKRLQELTQGSEDQQELAKSINKPSWAETLGSLF 68
P.mirabilis_SPI-WP_259493145.1/1-323 1 MASTFALILTLATLITGIFWCIDRFKLKPACKAKLKRLQELTQGSEDQQELAKSINKPSWAETLGSLF 68
P.mirabilis_SPI-WP_134940087.1/1-323 1 MASTFALILTLATLITGIFWCIDRFKLKPARKAKLKRLQELTQGSEDQQELAKSINKPSWAETLGSLF 68
P.mirabilis_SPI-WP_248706925.1/1-323 1 MASTFALILTLATLITGIFWCIDRFKLKPARKAKLKRLQELTQGSEDQQELAKSINKPSWAETLGSLF 68
P.mirabilis_SPI-HBC7478751.1/1-323 1 MASTFALILTLATLITGIFWCIDRFKLKPARKAKLKRLQELTQGSEDQQELAKSINKPSWAETLGSLF 68
P.mirabilis_SPI-WP_250256822.1/1-323 1 MASTFALILTLATLITGIFWCIDRFKLKPARKAKLKRLQELTQGSEDQQELAKSINKPSWAETLGFLF 68
P.mirabilis_SPI-MBG2893185.1/1-323 1 MASTFALILTLATLITGIFWCIDRFKLKPAHKAKLKRLQELTQGSEDQQELAKSINKPSWAETLGSLF 68
P.mirabilis_SPI-WP_063215817.1/1-323 1 MASTFALILTLATLITGIFWCIDRFKLKPARKAKLKRLQELTQGSEDQQELAKSINKPSWAETLGSLF 68
P.mirabilis_SPI-WP_259694527.1/1-323 1 MASTFALILTLATLITGIFWCIDRFKLKPARKAKLKRLQELTQGSEDQQELAKSINKPSWAETLGSLF 68
P.mirabilis_SPI-WP_046335197.1/1-323 1 MASTFALILTLATLITGIFWCIDRFKLKPARKAKLKRLQELTQGSEDQQELAKSINKPSWAETLGSLF 68
P.mirabilis_SPI-EIT0867198.1/1-323 1 MASTFAVILTLATLITGIFWCIDRFKLKPARKAKLKRLQELTQGSEDQQELAKSINKPSWAETLGSLF 68
P.mirabilis_SPI-WP_004243893.1/1-323 1 MASTFALILTLATLITGIFWCIDRFKLKPARKAKLKRLQELTQGSEDQQELAKSINKPSWAETLGSLF 68
P.mirabilis_SPI-MBG5941167.1/1-323 1 MASTFALILTLATLITGIFWCIDRFKLKPARKAKLKRLQELTQGSEDQQELAKSINKPSWAETLGSLF 68
P.mirabilis_SPI-MBG3104060.1/1-323 1 MASTFALILTLATLITGIFWCIDRFKLKPARKAKLKRLQELTQGSEDQQELAKSINKPSWAETLGSLF 68
P.mirabilis_SPI-WP_115353684.1/1-323 1 MASTFALILTLATLITGIFWCIDRFKLKPARKAKLKRLQELTQGSEDQQELAKSINKPSWAETLGSLF 68
P.mirabilis_SPI-WP_206449407.1/1-323 1 MASTFALILTLATLITGIFWCIDRFKLKPARKAKLKRLQELTQGSEDQQELAKSINKPSWAETLGSLF 68
P.mirabilis_SPI-WP_017627956.1/1-323 1 MASTFALILTLATLITGIFWCIDRFKLKPARKAKLKRLQELTQGSEDQQELAKSINKPSWAETLGSLF 68
P.mirabilis_SPI-WP259694214.1/1-323 1 MASTFALILTLATLITGIFWCIDRFKLKPARKAKLKRLQELTQGSEDQQELAKSINKPSWTETLGSLF 68
P.mirabilis_SPI-WP124724755.1/1-323 1 MASTFALILTLATLITGIFWCIDRFKLKPARKAKLKRLQELTQGSEDQQELAKSINKPNWAETLGSLF 68
P.mirabilis_SPI-EHN8771494.1/1-323 1 MASTFALILTLATLITGIFWCIDRFKLKPARKAKLKRLQELTQGSEDQQELAKSINKPSWAETLGSLF 68
P.mirabilis_SPI-WP_165375681.1/1-323 1 MASTFALILTLATLITGIFWCIDRFKLKPARKAKLKRLQELTQGSEDQQELAKSINKPSWAETLASLF 68
P.mirabilis_SPI-WP_004248458.1/1-323 1 MASTFALILTLATLITGIFWCIDRFKLKPARKAKLKRLQELTQGSEDQQELAKSINKPSWAETLGSLF 68
P.mirabilis_SPI-RNT32078.1/1-323 1 MANTFALILTLATLITGVFWCIERFKLKPARKAKLKRLQEMTQGTEDQQELAKSINKPSWAETLGSLF 68
P.mirabilis_SPI-MBG3080102.1/1-323 1 MANTFALILTLATLITGVFWGIERFKLKPARKAKLKRLQEMTQGTEDQQELAKSINKPSWAETLGSLF 68
P.mirabilis_SPI-MBG3013929.1/1-323 1 MANTFALILTLATLITGVFWGIERFKLKPARKAKLKRLQEMTQGSEDQQELAKSINKPSWAETLGSLF 68
P.mirabilis_SPI-MBG2711967.1/1-323 1 MANTFALILTLATLITGVFWGIERFKLKPARRAKLKRLQEMTQGTEDQQELAKSINKPSWAETLGSLF 68
P.mirabilis_SPI-MBG2800898.1/1-323 1 MANTFALILTLATLITGVFWGIERFKLKPARKAKLKRLQEMTQGTEDQQELAKSINKPSWAETLGSLF 68
P.mirabilis_SPI-WP_088495031.1/1-323 1 MANTFALILTLATLITGVFWGIERFKLKPARKAKLKRLQEMTQGTEDQQELAKSINKPSWAETLGSLF 68
P.mirabilis_SPI-MBG6026618.1/1-323 1 MANTFALILTLATLITGIFWGIERFKLKPARKAKLKRLQEMTQGTEDQQELAKSINKPSWAETLGSLF 68
P.mirabilis_SPI-KLU18752.1/1-323 1 MANTFALILTLATLITGVFWGIERFKLKPARKAKLKRLQEMTQGTEDQQELAKSINKPSWAETLGSLF 68
Consensus
MASTFALILTLATLITGIFWCIDRFKLKPARKAKLKRLQELTQGSEDQQELAKSINKPSWAETLGSLF
E.coli_SPI/1-324 69 PVLAIVLIVRSFIYEPFQIPSGSMMPTLLIGDFILVEKFAYGIKDPIYQKTLIETGHPKRGDIVVFKY 136
P.mirabilis_SPI-MBG6016051.1/1-323 69 PVLAIVLILRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLISTGKPKRGDIAVFKY 136
P.mirabilis_SPI-WP_217343403.1/1-323 69 PVLAIVLILRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLISTGKPKRGDIAVFKY 136
P.mirabilis_SPI-WP_243209876.1/1-323 69 PVLAIVLILRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLISTGKPKRGDIAVFKY 136
P.mirabilis_SPI-WP_163773644.1/1-323 69 PVLAIVLILRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLISTGKPKRGDIAVFKY 136
P.mirabilis_SPI-WP_036895417.1/1-323 69 PVLAIVLILRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLISTGKPKRGDIAVFKY 136
P.mirabilis_SPI-WP_036905315.1/1-323 69 PVLAIVLILRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLISTGKPKRGDIAVFKY 136
P.mirabilis_SPI-WP_259493145.1/1-323 69 PVLAIVLILRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLISTGKPKRGDIAVFKY 136
P.mirabilis_SPI-WP_134940087.1/1-323 69 PVLAIVLILRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLISTGKPKRGDIAVFKY 136
P.mirabilis_SPI-WP_248706925.1/1-323 69 PVLAIVLILRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLISTGKPKRGDIAVFKY 136
P.mirabilis_SPI-HBC7478751.1/1-323 69 PVLAIVLILRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLISTGKPKRGDIAVFKY 136
P.mirabilis_SPI-WP_250256822.1/1-323 69 PVLAIVLILRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLISTGKPKRGDIAVFKY 136
P.mirabilis_SPI-MBG2893185.1/1-323 69 PVLAIVLILRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLISTGKPKRGDIAVFKY 136
P.mirabilis_SPI-WP_063215817.1/1-323 69 PVLAIVLILRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLISTGKPKHGDIAVFKY 136
P.mirabilis_SPI-WP_259694527.1/1-323 69 PVLAIVLILRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLISTGKPKRGDIAVFKY 136
P.mirabilis_SPI-WP_046335197.1/1-323 69 PVLAIVLILRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLISTGKPKRGDIAVFKY 136
P.mirabilis_SPI-EIT0867198.1/1-323 69 PVLAIVLILRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLISTGKPKRGDIAVFKY 136
P.mirabilis_SPI-WP_004243893.1/1-323 69 PVLAIVLILRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLISTGKPKRGDIAVFKY 136
P.mirabilis_SPI-MBG5941167.1/1-323 69 PVLAIVLILRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLISTGKPKRGDIAVFKY 136
P.mirabilis_SPI-MBG3104060.1/1-323 69 PVLAIVLILRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLISTGKPKRGDIAVFKY 136
P.mirabilis_SPI-WP_115353684.1/1-323 69 PVLAIVLILRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLISTGKPKRGDIAVFKY 136
P.mirabilis_SPI-WP_206449407.1/1-323 69 PVLAIVLILRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLISTGKPKRGDIAVFKY 136
P.mirabilis_SPI-WP_017627956.1/1-323 69 PVLAIVLILRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLISTGKPKRGDIAVFKY 136
P.mirabilis_SPI-WP259694214.1/1-323 69 PVLAIVLILRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLISTGKPKRGDIAVFKY 136
P.mirabilis_SPI-WP124724755.1/1-323 69 PVLAIVLILRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLISTGKPKRGDIAVFKY 136
P.mirabilis_SPI-EHN8771494.1/1-323 69 PVLAIVLILRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLISTGKPKRGDIAVFKY 136
P.mirabilis_SPI-WP_165375681.1/1-323 69 PVLAIVLILRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLISTGKPKRGDIAVFKY 136
P.mirabilis_SPI-WP_004248458.1/1-323 69 PVLAIVLILRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLISTGKPKRGDIAVFKY 136
P.mirabilis_SPI-RNT32078.1/1-323 69 PVLIIVLVLRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLINTGKPKRGDIAVFKY 136
P.mirabilis_SPI-MBG3080102.1/1-323 69 PVLIIVLVLRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLINTGKPKRGDIAVFKY 136
P.mirabilis_SPI-MBG3013929.1/1-323 69 PVLIIVLVLRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLINTGKPKRGDIAVFKY 136
P.mirabilis_SPI-MBG2711967.1/1-323 69 PVLIIVLVLRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLINTGKPKRGDIAVFKY 136
P.mirabilis_SPI-MBG2800898.1/1-323 69 PVLIIVLVLRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLINTGKPKRGDIAVFKY 136
P.mirabilis_SPI-WP_088495031.1/1-323 69 PVLIIVLVLRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLINTGKPKRGDIAVFKY 136
P.mirabilis_SPI-MBG6026618.1/1-323 69 PVLIIVLVLRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLINTGKPKRGDIAVFKY 136
P.mirabilis_SPI-KLU18752.1/1-323 69 PVLIIVLVLRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLINTGKPKRGDIAVFKY 136
Consensus
PVLAIVLILRSFVYEPFQIPSRSMMPTLLVGDFILVEKFAYGLKDPITQTTLISTGKPKRGDIAVFKY


E.coli_SPI/1-324 137 P E D P K L D Y I K R A V G L P G D K V T Y D P V S K E L T I Q P G C S S G Q A C E N A L P V T Y S N V E P S D F V O T F S R R N G G E 204
P.mirabilis_SPI-MBG6016051.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M T S K K L H I Y P N C N K - A I C D D E I A V T Y G T A Y P S E W T L L L Q S V P G G Q 203
P.mirabilis_SPI-WP_217343403.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M T S K K L H I Y P N C N K - A I C D D E I A V T Y G T A Y P S E W T L L L Q S V P G G Q 203
P.mirabilis_SPI-WP_243209876.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M T S K K L H I Y P N C N K - A I Y D D E I A V T Y G T A Y P S E W T L L L Q S V P G G Q 203
P.mirabilis_SPI-WP_163773644.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M T S K K L H I Y P N C N K - A I C D D E I A V T Y G T A Y P S E W T L L L Q S V P G G Q 203
P.mirabilis_SPI-WP_036895417.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M T S K K L H I Y P N C N K - A I C D D E I A V T Y G T A Y P S E W T L L L Q S V P G G Q 203
P.mirabilis_SPI-WP_036905315.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M T S K K L H I Y P N C N K - A I C D D E I A V T Y G T A Y P S E W T L L L Q S V P G G Q 203
P.mirabilis_SPI-WP_259493145.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M T S K K L H I Y P N C N K - A I C D D E I A V T Y G T A Y P S E W T L L L Q S V P G G Q 203
P.mirabilis_SPI-WP_134940087.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M T S K K L H I Y P N C N K - A I C D D E I A V T Y G T A Y P S E W T L L L Q S V P G G Q 203
P.mirabilis_SPI-WP_248706925.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M T S K K L H I Y P N C N K - A I C D D E I A V T Y G T A Y P S E W T L L L Q S V P G G Q 203
P.mirabilis_SPI-HBC7478751.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M T S K K L H I Y P N C N K - A I C D D E I A V T Y G T A Y P S E W T L L L Q S V P G G Q 203
P.mirabilis_SPI-WP_250256822.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M T S K K L H I Y P N C N K - A I C D D E I A V T Y G T A Y P S E W T L L L Q S V P G G Q 203
P.mirabilis_SPI-MBG2893185.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M T S K K L H I Y P N C N K - A I C D D E I A V T Y G T A Y P S E W T L L L Q S V P G G Q 203
P.mirabilis_SPI-WP_063215817.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M T S K K L H I Y P N C N K - A I C D D E I A V T Y G T A Y P S E W T L L L Q S V P G G Q 203
P.mirabilis_SPI-WP_259694527.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M T S K K L H I Y P N C N K - A I C D D E I A V T Y G T A Y P S E W T L L L Q S V P G G Q 203
P.mirabilis_SPI-WP_046335197.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M T S K K L H I Y P N C N K - A I C D D E I A V T Y G T A Y P S E W T L L L Q S V P G G Q 203
P.mirabilis_SPI-EIT0867198.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M T S K K L H I Y P N C N K - A I C D D E I A V T Y G T A Y P S E W T L L L Q S V P G G Q 203
P.mirabilis_SPI-WP_004243893.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M T S K K L H I Y P N C N K - A I C D D E I A V T Y G T A Y P S E W T L L L Q S V P G G Q 203
P.mirabilis_SPI-MBG5941167.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M T S K K L H I Y P N C N K - A I C D D E I A V T Y G T A Y P S E W T L L L Q S V P G G Q 203
P.mirabilis_SPI-MBG3104060.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M T S K K L H I Y P N C N K - A I C D D E I A V T Y G T A Y P S E W T L L L Q S V P G G Q 203
P.mirabilis_SPI-WP_115353684.1/1-323 137 P R D P S T D F I K R M I G L P G D K I V Y D M T S K K L H I Y P N C N K - A I C D D E I A V T Y G T A Y P S E W T L L L Q S V P G G Q 203
P.mirabilis_SPI-WP_206449407.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M T S K K L H I Y P N C N K - A L C D D E I A V T Y G T A Y P S E W T L L L Q S V P G G Q 203
P.mirabilis_SPI-WP_017627956.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D V T S K K L H I Y P N C N K - A I C D D E I A V T Y G T A Y P S E W T L L L Q S V P G G Q 203
P.mirabilis_SPI-WP259694214.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M T S K K L H I Y P N C N K - A I C D D E I A V T Y G T A Y P S E W T L L L Q S V P G G Q 203
P.mirabilis_SPI-WP124724755.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M T S K K L H I Y P N C N K - A I C D D E I A V T Y G T A Y P S E W T L L L Q S V P G G Q 203
P.mirabilis_SPI-EHN8771494.1/1-323 137 P R D P S T D F I K R V V G L P G D K I V Y D M T S K K L H I Y P N C N K - A I C D D E I A V T Y G T A Y P S E W T L L L Q S V P G G Q 203
P.mirabilis_SPI-WP_165375681.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M T S K K L H I Y P N C N K - A I C D D E I A V T Y G T A Y P S E W T L L L Q S V P G G Q 203
P.mirabilis_SPI-WP_004248458.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M T S K K L H I Y P N C N K - A I C D D E I A V T Y G T A Y P S E W T L L L Q S V P G G Q 203
P.mirabilis_SPI-RNT32078.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M M S K K I H I Y P N C N K - A I C N D E I S V T Y G T A Y P S E W T L L L Q N V P G G Q 203
P.mirabilis_SPI-MBG3080102.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M M S K K I H I Y P N C N K - A I C N D E I S V T Y G T A Y P S E W T L L L Q N V P G G Q 203
P.mirabilis_SPI-MBG3013929.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M M S K K L H I Y P N C D K - A V C N E E I S V T Y G T A Y P S E W T L L L Q N V P G G Q 203
P.mirabilis_SPI-MBG2711967.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M V S K K L H I Y P N C D K - A I C N E E I S V T Y G T A Y P S E W T L L L Q N V P G G Q 203
P.mirabilis_SPI-MBG2800898.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M V S K K L H I Y P N C D K - T I C N E E I S V T Y G T A Y P S E W T L L L Q N V P G G Q 203
P.mirabilis_SPI-WP_088495031.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M M S K K L H V Y P N C D K - A I C N E E I S V T Y G T A Y P S E W T L L L Q N V P G G Q 203
P.mirabilis_SPI-MBG6026618.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M M S K K L H I Y P N C D K - A I C N E E I S V T Y G T A Y P S E W T L L L Q N V P G G Q 203
P.mirabilis_SPI-KLU18752.1/1-323 137 P R D P S T D F I K R V I G L P G D K I V Y D M M S K K L H I Y P N C D K - A I C N E E I S V T Y G T A Y P S E W T L L L Q N V P G G Q 203
Consensus
P R D P S T D F I K R V I G L P G D K I V Y D M T S K K L H I Y P N C N K G A I C D D E I A V T Y G T A Y P S E W T L L L Q S V P G G Q
E.coli_SPI/1-324 205 A T S G F F E V P K N E T K E - - N G I R L S E R K E T L G D V T H R I L I V P I A Q D Q V G M Y Y Q Q P G Q Q L A T W I V P P G Q Y F 270
P.mirabilis_SPI-MBG6016051.1/1-323 204 R I S G M K S I P I E E P I S T A T Q F R Q E E R I E T I G D V S H R I M T I P G D - - I T V P E F I Q P G L P Q G T W I V P E G Q Y F 269
P.mirabilis_SPI-WP_217343403.1/1-323 204 R I S G M K S I P I E E P I S T A T Q F R Q E E R I E T I G D V S H R I M T I P G D - - I T V P E F I Q P G L P Q G T W I V P E G Q Y F 269
P.mirabilis_SPI-WP_243209876.1/1-323 204 R I S G M K S I P I E E P I S T A T Q F R Q E E R I E T I G D V S H R I M T I P G D - - V T V P E F I Q P G L P Q G T W I V P E G Q Y F 269
P.mirabilis_SPI-WP_163773644.1/1-323 204 R I S G M K S I P I E E P I S T A T Q F R Q E E R I E T I G D V S H R I M T I P G D - - V T V P E F I Q P G L P Q G T W I V P E G Q Y F 269
P.mirabilis_SPI-WP_036895417.1/1-323 204 R I S G M K S I P I E E P I S T A T Q F R Q E E R I E T I G D V S H R I M T I P G D - - V T V P E F I Q P G L P Q G T W I V P E G Q Y F 269
P.mirabilis_SPI-WP_036905315.1/1-323 204 R I S G M K S I P I E E P I S T A T Q F R Q E E R I E T I G D V S H R I M T I P G D - - V T V P E F I Q P G L P Q G T W I V P E G Q Y F 269
P.mirabilis_SPI-WP_259493145.1/1-323 204 R I S G M K S I P I E E P I S T A T Q F R Q E E R I E T I G D V S H R I M T I P G D - - V T V P E F I Q P G L P Q G T W I V P E G Q Y F 269
P.mirabilis_SPI-WP_134940087.1/1-323 204 R I S G M K S I P I E E P I S T A T Q F R Q E E R I E T I G D V S H R I M T I P G D - - V T V P E F I Q P G L P Q G T W I V P E G Q Y F 269
P.mirabilis_SPI-WP_248706925.1/1-323 204 R I S G M K S I P I E E P I S T A T Q F R Q E E R I E T I G D V S H R I M T I P G D - - V T V P E F I Q P G L P Q G T W I V P E D Q Y F 269
P.mirabilis_SPI-HBC7478751.1/1-323 204 R I S G M K S I P I E E P I S T A T Q F R Q E E R I E T I G D V S H R I M T I L G D - - V T I P E F I Q P G L P Q G T W I V P E G Q Y F 269
P.mirabilis_SPI-WP_250256822.1/1-323 204 R I S G M K S I P I E E P I S T A T Q F R Q E E R I E T I G D V S H R I M T I P G D - - V T V P E F I Q P G L P Q G T W I V P E G Q Y F 269
P.mirabilis_SPI-MBG2893185.1/1-323 204 R I S G M K S I P I E E P I S T A T Q F R Q E E R I E T I G D V S H R I M T I P G D - - V T I P E F I Q P G L P Q G T W I V P E G Q Y F 269
P.mirabilis_SPI-WP_063215817.1/1-323 204 R I S G M K S I P I E E P I S T A T Q F R Q E E R I E T I G D V S H R I M T I P G D - - V T V P E F I Q P G L P Q G T W I V P E G Q Y F 269
P.mirabilis_SPI-WP_259694527.1/1-323 204 R I S G M K S I P I E E P I S T A T Q F R Q E E R I E T I G D V S H R I M T I P G D - - V T V P E F I Q P G L P Q G T W I V P E G Q Y F 269
P.mirabilis_SPI-WP_046335197.1/1-323 204 R I S G M K S I P I E E P I S T A T Q F R Q E E R I E T I G D V S H R I M T I P G D - - V T V P E F I Q P G L P Q G M W I V P E G Q Y F 269
P.mirabilis_SPI-EIT0867198.1/1-323 204 R I S G M K S I P I E E P I S T A T Q F R Q E E R I E T I G D V S H R I M T I P G D - - V T V P E F I Q P G L P Q G T W I V P E G Q Y F 269
P.mirabilis_SPI-WP_004243893.1/1-323 204 R I S G M K S I P I E E P I S T A T Q F R Q E E R I E T I G D V S H R I M T I P G D - - V T I P E F I Q P G L P Q G T W I V P E G Q Y F 269
P.mirabilis_SPI-MBG5941167.1/1-323 204 R I S G M K S I P I E E P I S T A T Q F R Q E E R I E T I G D V S H R I M T I P G D - - V T V P E F I Q P G L P Q G T W I V P E G Q Y F 269
P.mirabilis_SPI-MBG3104060.1/1-323 204 R I S G I K S I P I E E P I S T A T Q F R Q E E R I E T I G D V S H R I M T I P G D - - V T V P E F I Q P G L P Q G T W I V P E G Q Y F 269
P.mirabilis_SPI-WP_115353684.1/1-323 204 R I S G M K S I P I E E P I S T A T Q F R Q E E R I E T I G D V S H R I M T I P G D - - V T V P E F I Q P G L P Q G T W I V P E G Q Y F 269
P.mirabilis_SPI-WP_206449407.1/1-323 204 R I S G M K S I P I E E P I S T A T Q F R Q E E R I E T I G D V S H R I M T I P G D - - V T V P E F I Q P G L P Q G T W I V P E G Q Y F 269
P.mirabilis_SPI-WP_017627956.1/1-323 204 R I S G M K S I P I E E P I S T A T Q F R Q E E R I E T I G D V S H R I M T I P G D - - V T V P E F I Q P G L P Q G T W I V P E G Q Y F 269
P.mirabilis_SPI-WP259694214.1/1-323 204 R I S G M K S I P I E E P I S T A T Q F R Q E E R I E T I G D V S H R I M T I P G D - - V T V P E F I Q P G L P Q G T W I V P E G Q Y F 269
P.mirabilis_SPI-WP124724755.1/1-323 204 R I S G M K S I P I E E P I S T A T Q F R Q E E R I E T I G D V S H R I M T I P G D - - V T V P E F I Q P G L P Q G T W I V P E G Q Y F 269
P.mirabilis_SPI-EHN8771494.1/1-323 204 R I S G M K S I P I E E P I S T A T Q F R Q E E R I E T I G D V S H R I M T I P G D - - V T V P E F I Q P G L P Q G T W I V P E G Q Y F 269
P.mirabilis_SPI-WP_165375681.1/1-323 204 R I S G M K S I P I E E P I S T A T Q F R Q E E R I E T I G D V S H R I M T I P G D - - V T V P E F I Q P G L P Q G T W I V P E G Q Y F 269
P.mirabilis_SPI-WP_004248458.1/1-323 204 R I S G M K S I P I E E P I S T A T Q F R Q E E R I E T I G D V S H R I M T I P G D - - V T V P E F I Q P G L P Q G T W I V P E G Q Y F 269
P.mirabilis_SPI-RNT32078.1/1-323 204 R I N G M K S I P I E E P I S T A T Q L R E E E R V E T I G K V S H R I M T I P G D - - M T M P E F I Q P G L P V G T W I V P E D H Y F 269
P.mirabilis_SPI-MBG3080102.1/1-323 204 R I N G M K S I P I E D P I S T A T Q L R E E E R V E T I G K V S H R I M T I P G D - - M T M P E F I Q P G L P V G T W I V P E D H Y F 269
P.mirabilis_SPI-MBG3013929.1/1-323 204 R I N G M K S I P I E E P I S T A T Q L R E E E R V E T I D K V S H R I M T I P G D - - M T M P E F I Q P G L P V G T W I V P E G H Y F 269
P.mirabilis_SPI-MBG2711967.1/1-323 204 R I N G M K S I P I E E P I S T A T Q L R E E E R V E T I D K V S H R I M T I P G D - - M T M P E F I Q P G L P V G T W I V P E G H Y F 269
P.mirabilis_SPI-MBG2800898.1/1-323 204 R I N G M K S I P I E E P I S T A T Q L R E E E R V E T I D K V S H R I M T I P G D - - M T M P E F I Q P G L P V G T W I V P E G H Y F 269
P.mirabilis_SPI-WP_088495031.1/1-323 204 R I N G M K S I P I E E P I S T A T Q L R E E E R V E T I D K V S H R I M T I P G D - - M T M P E F I Q P G L P V G T W I V P E G H Y F 269
P.mirabilis_SPI-MBG6026618.1/1-323 204 R I N G M K S I P I E E P I S T A T Q L R E E E R V E T I D K V S H R I M T I P G D - - M T M P E F I Q P G L P V G T W I V P E G H Y F 269
P.mirabilis_SPI-KLU18752.1/1-323 204 R I N G M K S I P I E E P I S T A T Q L R E E E R V E T I D K V S H R I M T I P G D - - M T M P E F I Q P G L P V G T W I V P E G H Y F 269
Consensus
R I S G M K S I P I E E P I S T A T Q F R Q E E R I E T I G D V S H R I M T I P G D Q D V T V P E F I Q P G L P Q G T W I V P E G Q Y F

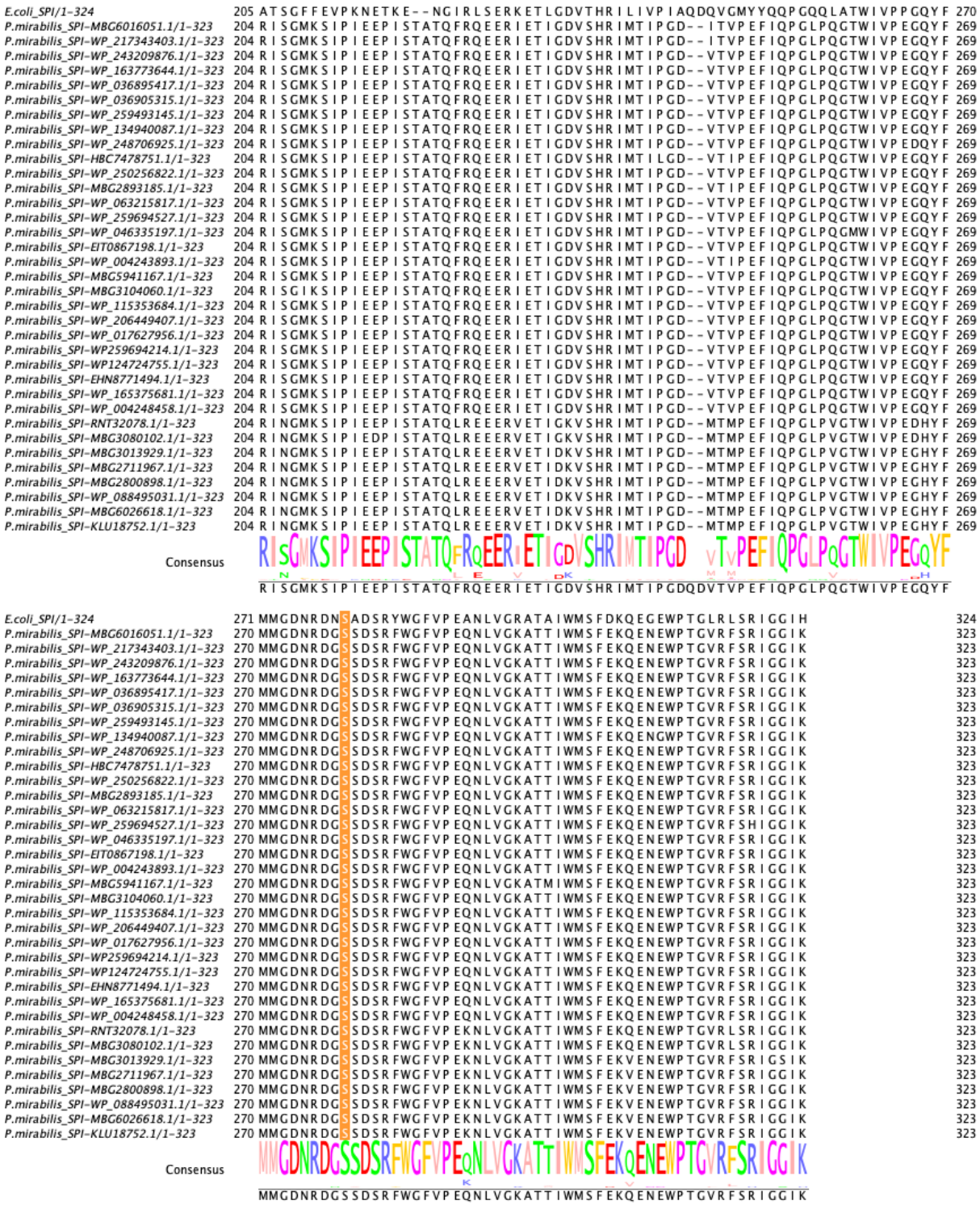
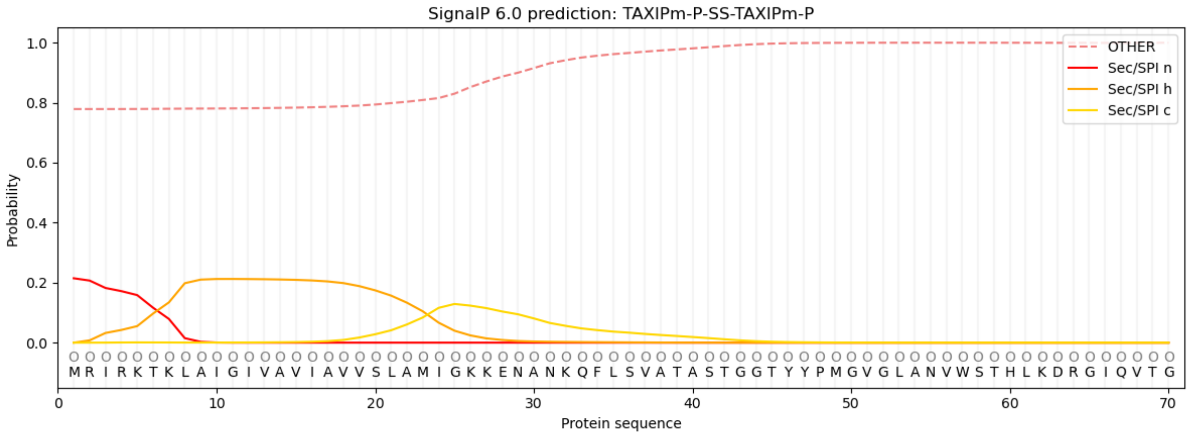
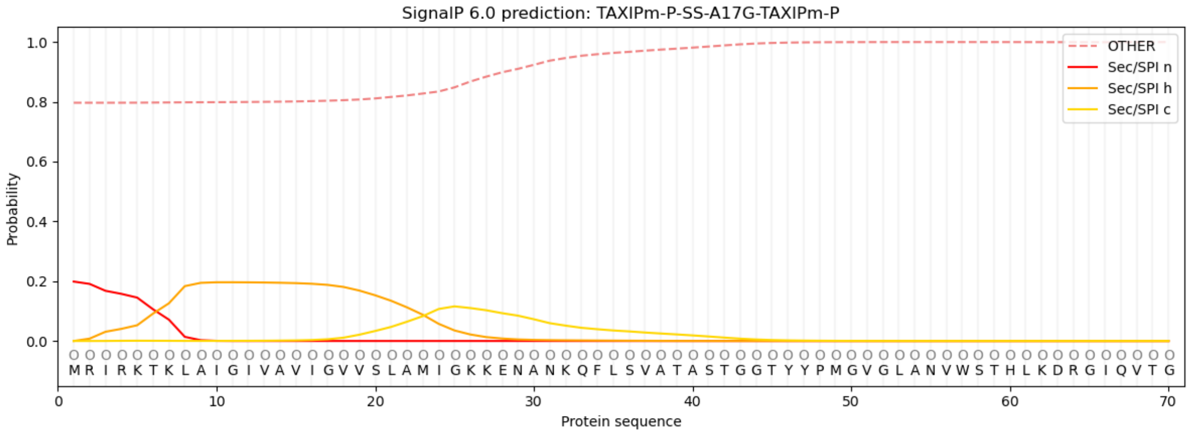
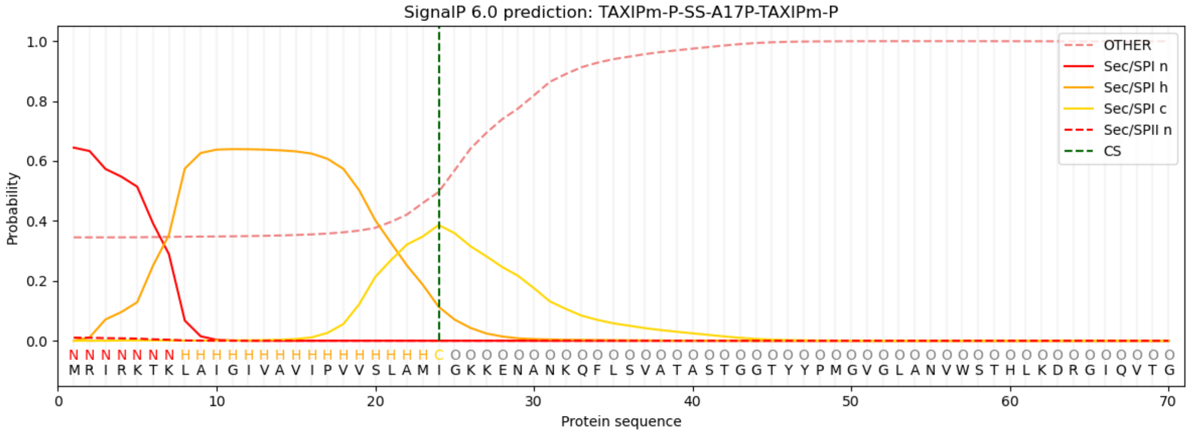


Figure 111 **Protein sequence alignment of the *E. coli* and *P. mirabilis'* signal peptidases I.** The catalytic residues, known from the signal peptidase I of *E. coli,* are depicted in orange (Mark Paetzel 2013). Protein sequences were aligned using Clustal Ω (Sievers et al. 2011). The alignment was colored and the consensus logo was created using Jalview (Waterhouse et al. 2009).


SignalP 6.0 prediction: TAXIPm-P-SS-TAXIPm-P


SignalP 6.0 prediction: TAXIPm-P-SS-A17G-TAXIPm-P


SignalP 6.0 prediction: TAXIPm-P-SS-A17P-TAXIPm-P

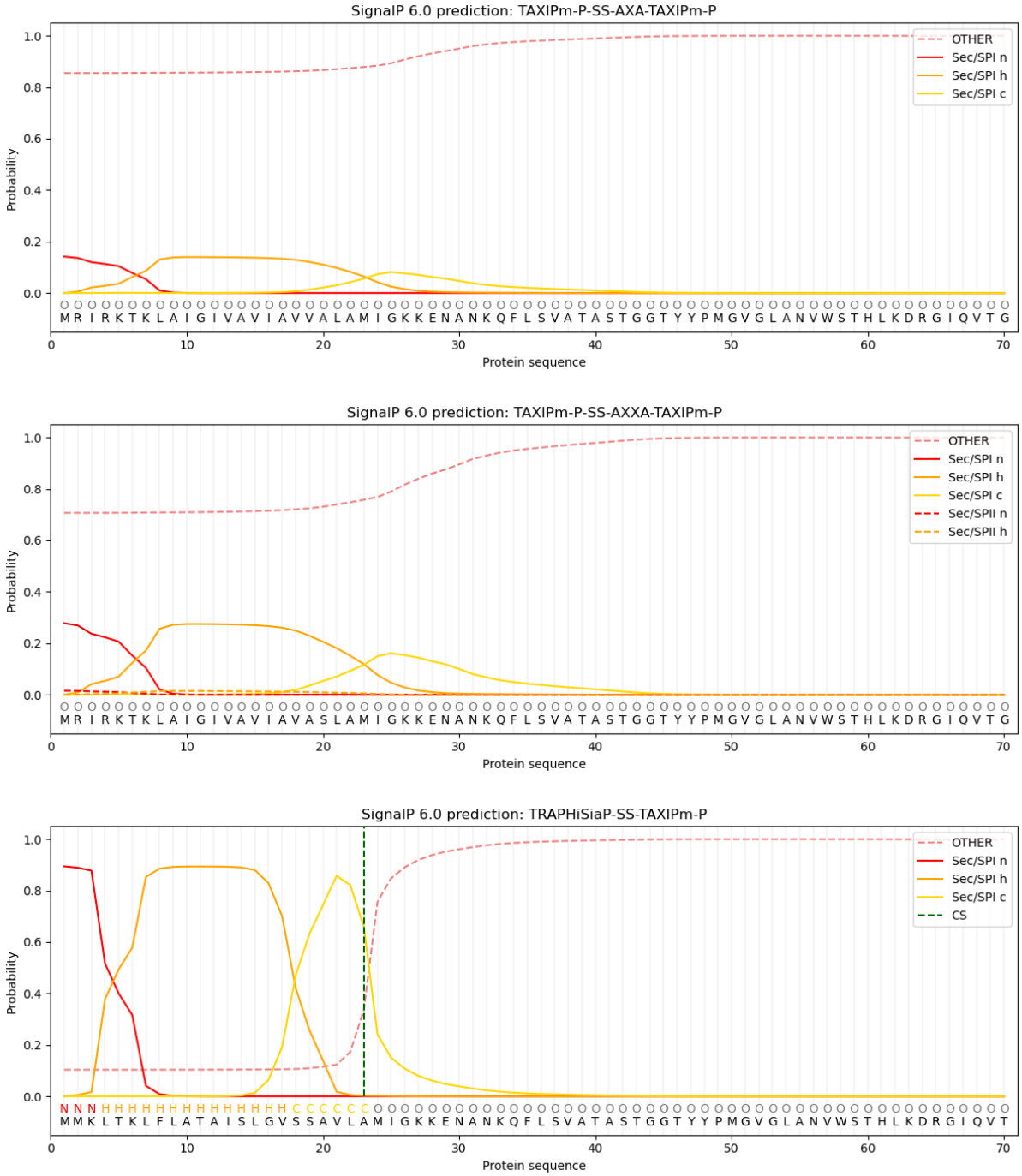


Figure 112 **Analysis of TAXIPm-P signal sequence.** The TAXIPm-P signal sequence from Proteus mirabilis was analyzed using SignalP with its endogenous sequence and amino acid substitution at position 17 from Alanine to either Glycine or Proline. Cleavage site probability for the latter is 39 %. Additionally, the Serine and Valine at -3 and -4 position were substituted respectively by Alanine. Then, TAXIPm-P with TRAPHi-SiaP signal sequence was analyzed, resulting in a cleavage site probability of 66 %. The red, orange and yellow line show the probability for a secretory pathway with Signal Peptidase I cleavage based on the N-, middle and C-region of the sequence, respectively. The green dashed line shows the predicted cleavage site for signal peptidase I. The red dashed line indicates the probability for "other", which refers to no prediction of a signal sequence, but an N-terminal transmembrane segment. (Teufel et al. 2022).

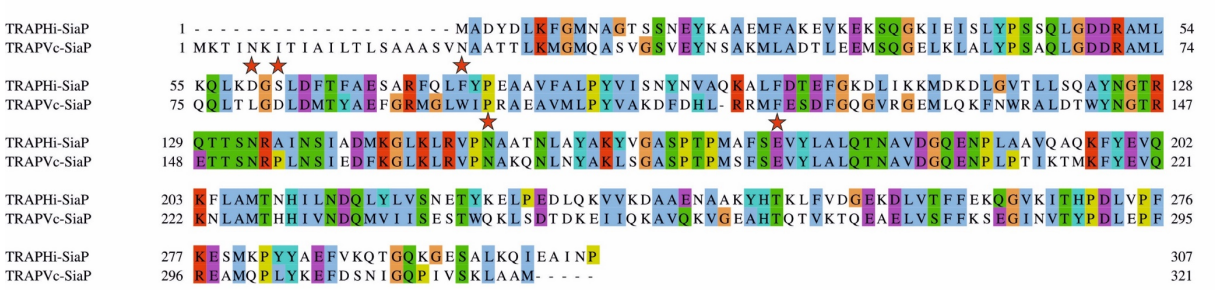


Figure 113 **Protein sequence alignment of TRAPVc-SiaP and TRAPHi-SiaP.** Red asterisks label residues predicted to be involved by Peter 2022 in the interaction between TRAPHi-SiaP and TRAPHi-SiaQM (Peter et al. 2022). Aligned sequences are shown in clustalx coloring using Jalview (Waterhouse et al. 2009).

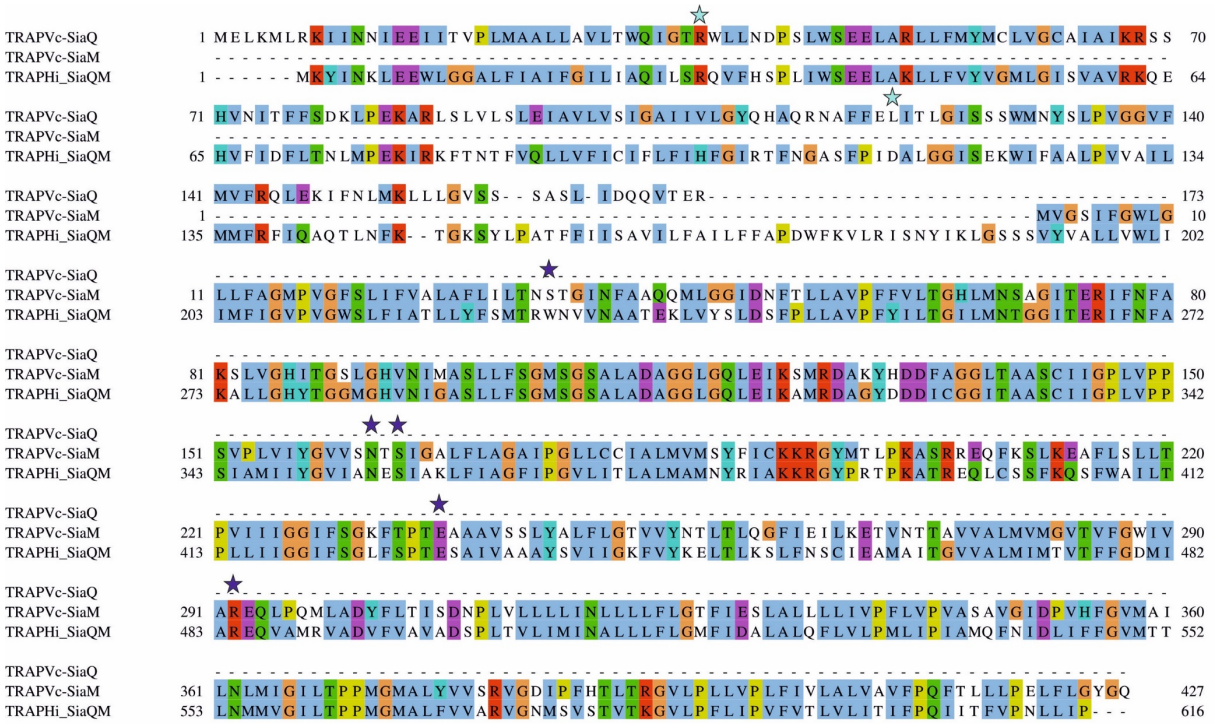


Figure 114 **Protein sequence alignment of TRAPVc-SiaQM and TRAPHi-SiaQM.** Light and dark blue asterisks label residues predicted by Peter 2022 to be involved in the interaction between TRAPHi-SiaP and TRAPHi-SiaQ and TRAPHi-SiaM, respectively (Peter et al. 2022). Aligned sequences are shown in clustalx coloring using Jalview (Waterhouse et al. 2009).

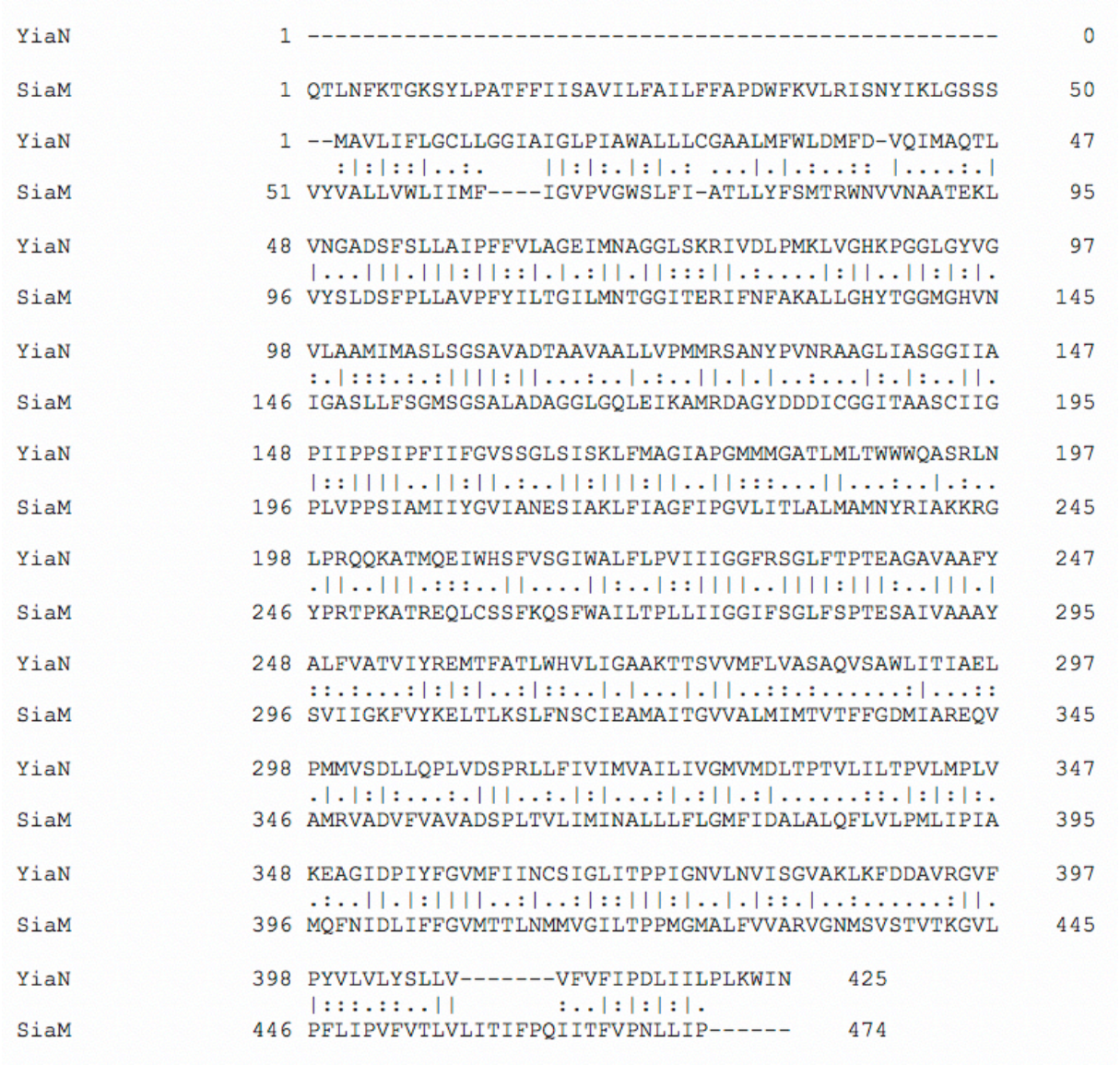
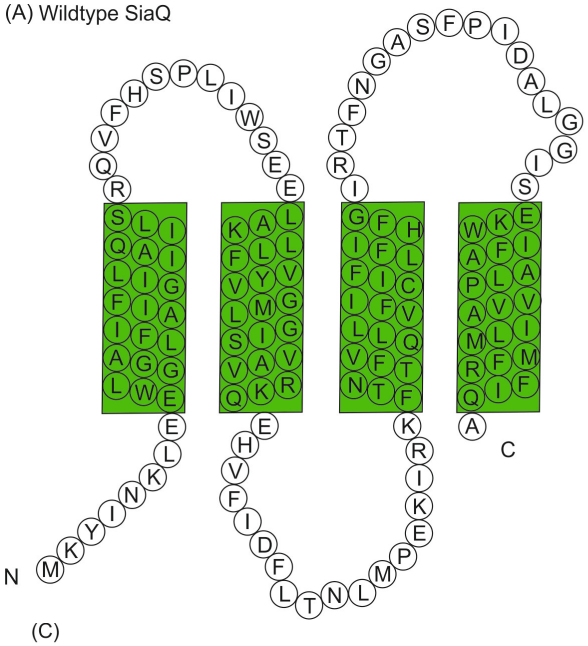
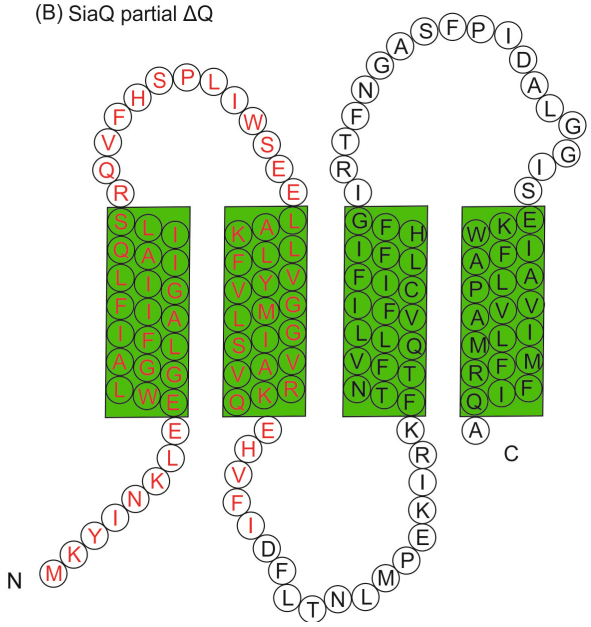


Figure 115 **Alignment of the two transport domains TRAPEc-YiaN and TRAPHi-SiaM**. The two transport domains TRAPEc-YiaN and TRAPHi-SiaM were aligned using Clustal Ω. Identical amino acids are labeled with solid lines and similar amino acids are labeled with double dots (Sievers et al. 2011).

(A) Wildtype SiaQ

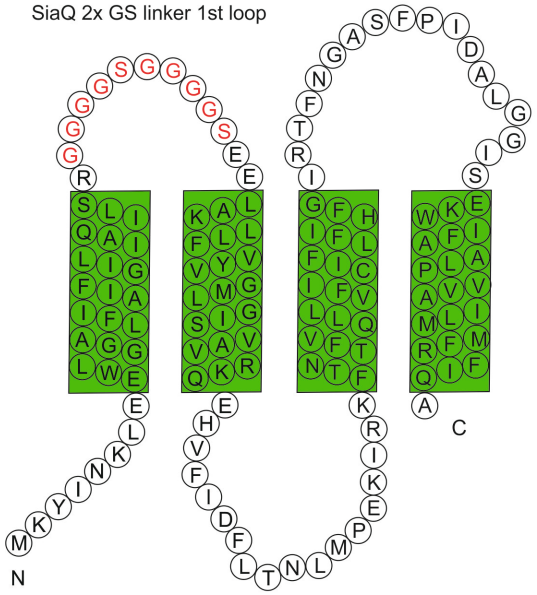


(B) SiaQ partial ΔQ

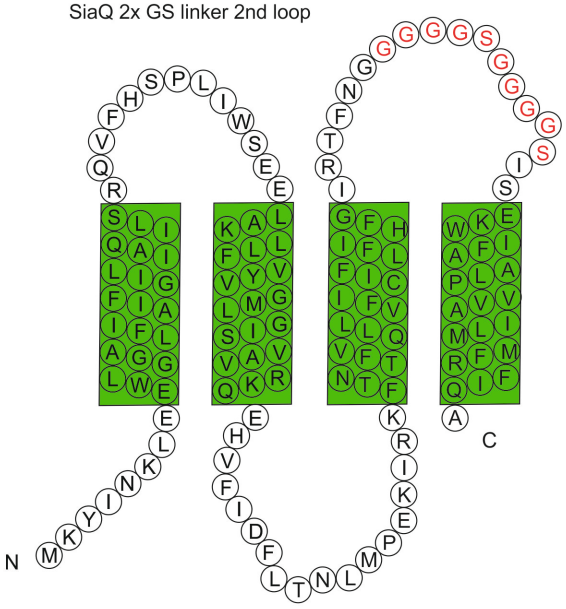


(C)

SiaQ 2x GS linker 1st loop



SiaQ 2x GS linker 2nd loop



SiaQ 3x GS linker 2nd loop

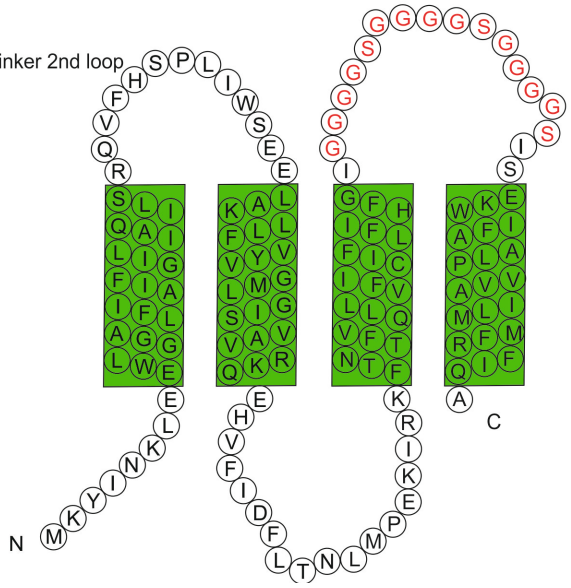


Figure 116 **Topology plot of TRAPHi-Q using Topcon and alphafold predictions.** Bioinformatic predictions of the TRAPHi-Q structure were used to design variants in which TRAPHi-Q has been either completely or partially deleted. In a third variant, five amino acids were exchanged by a glycine serine linker. The targeted amino acids are shown. (A) The wildtype TRAPHi-Q is shown, which is completely removed in the Q deletion variant. (B) The amino acids shown in red are removed in the partial deletion of TRAPHi-Q. (C) Either ten or fifteen amino acids were exchanged by two or three glycine/serine linkers in the first or second periplasmic loop of TRAPHi-SiaQ. The topology plots were created using Topcon and alphafold predictions (Tsirigos et al. 2015; Jumper et al. 2021).

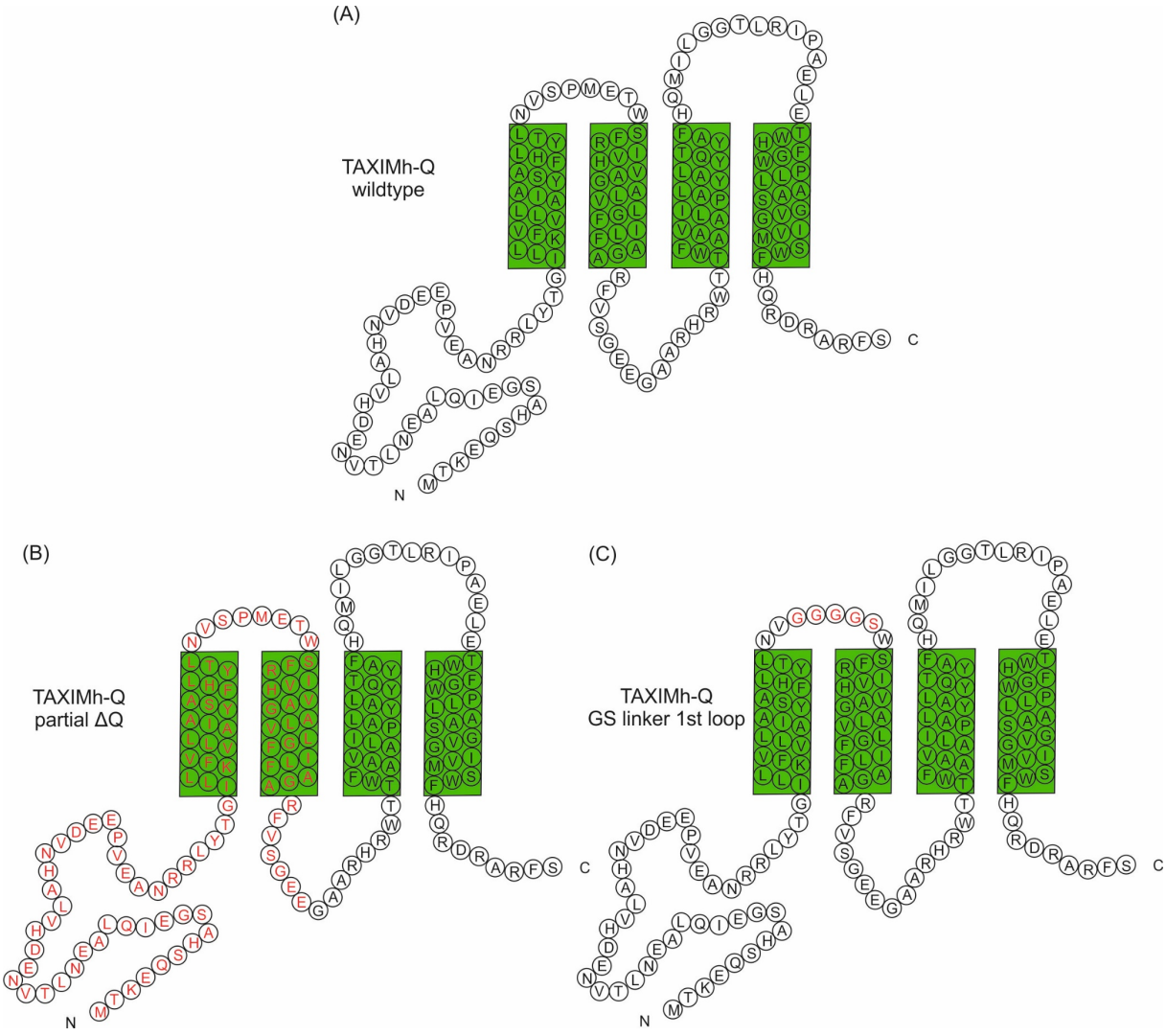


Figure 117 **Topology plot of TAXIMh-Q using Topcon and alphafold predictions.** Bioinformatic predictions of the TAXIMh-Q structure were used to design mutants in which TAXIMh-Q has been either completely or partially deleted. In a third variant, five amino acids were exchanged by a glycine serine linker. The targeted amino acids are shown. (A) The wildtype TAXIMh-Q is shown, which is completely removed in the Q deletion variant. (B) The amino acids shown in red are removed in the partial deletion of TAXIMh-Q. (C) Five amino acids in the first periplasmic loop were exchanged by four glycins and one serine. (Tsirigos et al. 2015; Jumper et al. 2021).

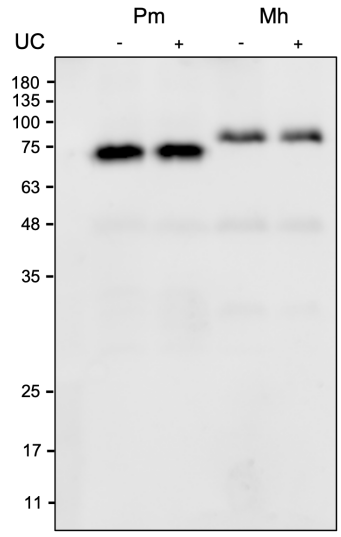
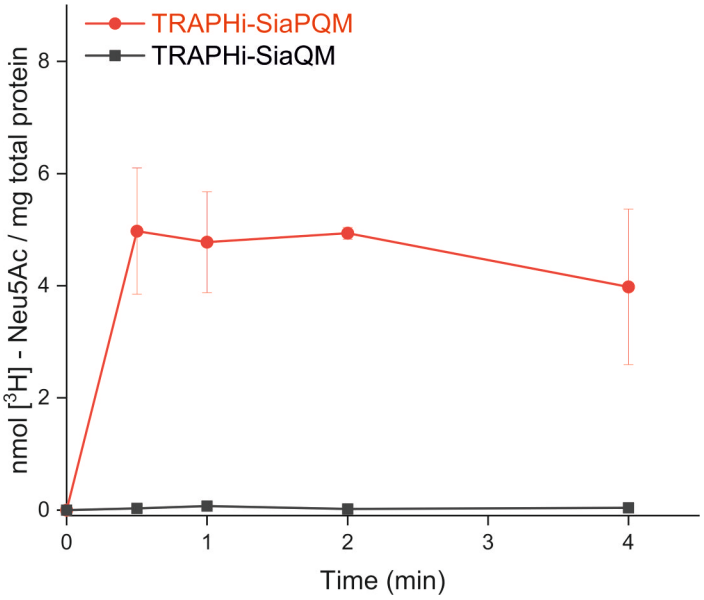


Figure 118 **Expression levels of TAXIPm-QM and TAXIMh-QM from expression for *in vivo* K_M/v_{max} transport assays.** TAXIPm-PQM from *Proteus mirabilis* and the negative control TAXIMh-PQM from *Marinobacter hydrocarbonoclasticus* were expressed in pEXC3sfGH with the TRAPHi-SiaP signal peptide from *Haemophilus influenzae*, thereby induced using 1 mM IPTG. Whole cell lysates of TAXIMh-PQM and TAXIPm-PQM cell suspensions were solubilized in 1 % (w/v) DDM and submitted to ultracentrifugation to pellet aggregated protein. Samples were applied from before (-) and after (+) ultracentrifugation (UC). The target proteins were detected by measuring in gel GFP fluorescence.



.

Figure 119 ***In vitro* uptake of N-acetylneuraminic acid by TRAPHi-SiaPQM in TRAP-SiaQM proteoliposomes, supplemented with TRAPHi-SiaP.** The N-acetylneuraminic acid uptake was measured in TRAPHi-SiaQM proteoliposomes at 20 °C. 5 µg TRAPHi-SiaQM was used per time point. The assay buffer in which the proteoliposomes were diluted contained 5 µM TRAPHi-SiaP and 5 µM [^{3}H] labelled N-acetylneuraminic acid. An inward-directed Na^+-gradient was applied. As negative control, SiaQM proteoliposomes in the absence of the SBP TRAPHi-SiaP were used. The negative control was measured in a single measurement and the TRAPHi-SiaPQM sample in duplicates, error bars indicate standard deviations. The plot was created using OriginPro 2019b (OriginLab Corporation).

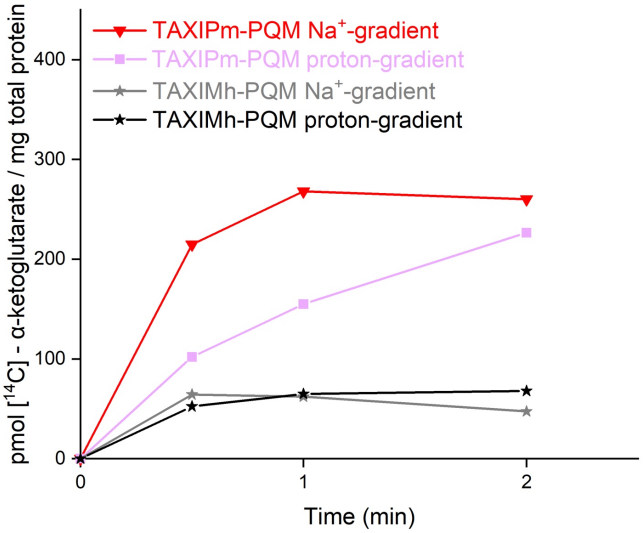


Figure 120 **Whole cell α-ketoglutarate uptake by TAXIPm-PQM. α-ketoglutarate uptake by TAXIPm-PQM in whole *E. coli* JW2571-1 (*ΔkgtP*).** Cells were pre-energized by a 2 min pre-incubation with 0.5% (w/v) glucose. Transport was assayed at 30 °C and pH 7.5 in the presence of 2 µM [^{14}C] labeled α-ketoglutarate using an inwards directed Na^+-gradient or proton gradient. Both, TAXIPm-PQM and the negative control TAXIMh-PQM from *Marinobacter hydrocarbonoclasticus* were expressed in pEXC3sfGH with the TRAPHi-SiaP signal peptide from *Haemophilus influenzae* and induced with 1 mM IPTG. All time points reflect single measurements. The plot was created using OriginPro 2019b (OriginLab Corporation).

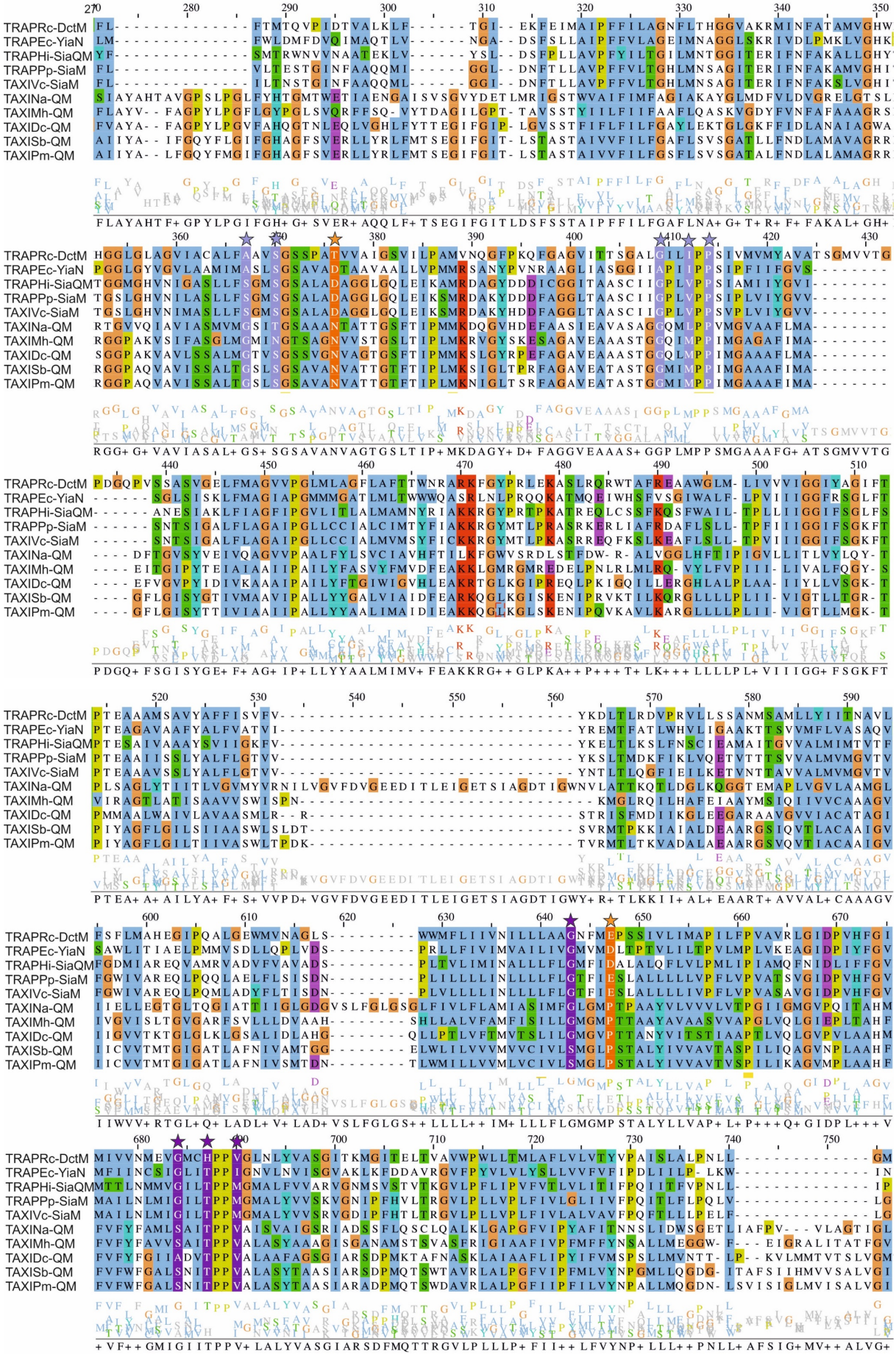

TRAPRc-DctM
TRAPEc-YiaN
TRAPHi-SiaQM
TRAPPp-SiaM
TAXIVc-SiaM
TAXINa-QM
TAXIMh-QM
TAXIDc-QM
TAXISb-QM
TAXIPm-QM

Figure 121 **Complete protein sequence alignment of TRAP and TAXI sequences.** Protein sequence alignment to evaluate TM helices that include potential sodium coordinating residues Potential sodium coordinating residues in M-domains. Residues assigned for sodium coordination by Davies, 2023 for TRAPPp-SiaM are displayed in light and dark purple and residues assigned for sodium coordination by Peter, 2022 for TRAPHi-SiaM are displayed in orange (Peter et al. 2022; Davies, Currie, and North 2023). Aligned sequences are shown in clustalx coloring using Jalview (Waterhouse et al. 2009).

References

Albers, Sonja V., Marieke G.L. Elferink, Robert L. Charlebois, Christoph W. Sensen, Arnold J.M. Driessen, and Wil N. Konings. 1999. 'Glucose Transport in the Extremely Thermoacidophilic Sulfolobus Solfataricus Involves a High-Affinity Membrane-Integrated Binding Protein'. *Journal of Bacteriology* 181 (14): 4285–91. https://doi.org/10.1128/jb.181.14.4285-4291.1999.

Almagro Armenteros, José Juan, Konstantinos D. Tsirigos, Casper Kaae Sønderby, Thomas Nordahl Petersen, Ole Winther, Søren Brunak, Gunnar von Heijne, and Henrik Nielsen. 2019. 'SignalP 5.0 Improves Signal Peptide Predictions Using Deep Neural Networks'. *Nature Biotechnology* 37 (4): 420–23. https://doi.org/10.1038/s41587-019-0036-z.

Antoine, Rudy, Françoise Jacob-Dubuisson, Hervé Drobecq, Eve Willery, Sarah Lesjean, and Camille Locht. 2003. 'Overrepresentation of a Gene Family Encoding Extracytoplasmic Solute Receptors in Bordetella'. *Journal of Bacteriology* 185 (4): 1470–74. https://doi.org/10.1128/JB.185.4.1470-1474.2003.

Archer, William R., and Michael D. Schulz. 2020. 'Isothermal Titration Calorimetry: Practical Approaches and Current Applications in Soft Matter'. *Soft Matter* 16 (38): 8760–74. https://doi.org/10.1039/d0sm01345e.

Arkhipova, Valentina, Albert Guskov, and Dirk J. Slotboom. 2020. 'Structural Ensemble of a Glutamate Transporter Homologue in Lipid Nanodisc Environment'. *Nature Communications* 11 (1): 1–9. https://doi.org/10.1038/s41467-020-14834-8.

Auclair, Sarah M., Meera K. Bhanu, and Debra A. Kendall. 2012. 'Signal Peptidase I: Cleaving the Way to Mature Proteins'. *Protein Science* 21 (1): 13–25. https://doi.org/10.1002/pro.757.

Baba, Tomoya, Takeshi Ara, Miki Hasegawa, Yuki Takai, Yoshiko Okumura, Miki Baba, Kirill A. Datsenko, Masaru Tomita, Barry L. Wanner, and Hirotada Mori. 2006. 'Construction of Escherichia Coli K-12 in-Frame, Single-Gene Knockout Mutants: The Keio Collection'. *Molecular Systems Biology* 2. https://doi.org/10.1038/msb4100050.

Bamber, Lisa, Marilyn Harding, G. Jonathan P. Butler, and Edmund R.S. Kunji. 2006. 'Yeast Mitochondrial ADP/ATP Carriers Are Monomeric in Detergents'. *Proceedings of the National Academy of Sciences of the United States of America* 103 (44): 16224–29. https://doi.org/10.1073/pnas.0607640103.

Bannas, Peter, Julia Hambach, and Friedrich Koch-Nolte. 2017. 'Nanobodies and Nanobody-Based Human Heavy Chain Antibodies as Antitumor Therapeutics'. *Frontiers in Immunology* 8 (NOV): 1–13. https://doi.org/10.3389/fimmu.2017.01603.

Batista-García, Ramón Alberto, Ayixon Sánchez-Reyes, César Millán-Pacheco, Víctor Manuel González-Zuñiga, Soledad Juárez, Jorge Luis Folch-Mallol, and Nina Pastor. 2014. 'A Novel TctA Citrate Transporter from an Activated Sludge Metagenome: Structural and Mechanistic Predictions for the TTT Family'. *Proteins: Structure, Function and Bioinformatics* 82 (9): 1756–64. https://doi.org/10.1002/prot.24529.

Bender, Brian J., Stefan Gahbauer, Andreas Luttens, Jiankun Lyu, Chase M. Webb, Reed M. Stein, Elissa A. Fink, et al. 2021. 'A Practical Guide to Large-Scale Docking'. *Nature Protocols* 16 (10): 4799–4832. https://doi.org/10.1038/s41596-021-00597-z.

Benson, Dennis A., Mark Cavanaugh, Karen Clark, Ilene Karsch-Mizrachi, David J. Lipman, James Ostell, and Eric W. Sayers. 2013. 'GenBank'. *Nucleic Acids Research* 41 (D1): 36–42. https://doi.org/10.1093/nar/gks1195.

Berntsson, Ronnie P.A., Sander H.J. Smits, Lutz Schmitt, Dirk Jan Slotboom, and Bert Poolman. 2010. 'A Structural Classification of Substrate-Binding Proteins'. *FEBS Letters* 584 (12): 2606–17. https://doi.org/10.1016/j.febslet.2010.04.043.

Bianchi, Frans, Joury S. Van 'T Klooster, Stephanie J. Ruiz, Katja Luck, Tjeerd Pols, Ina L. Urbatsch, and Bert Poolman. 2016. 'Asymmetry in Inward- and Outward-Affinity Constant of Transport Explain Unidirectional Lysine Flux in Saccharomyces Cerevisiae'. *Scientific Reports* 6: 1–13. https://doi.org/10.1038/srep31443.

Bloch, Joël S., Somnath Mukherjee, Julia Kowal, Ekaterina V. Filippova, Martina Niederer, Els Pardon, Jan Steyaert, Anthony A. Kossiakoff, and Kaspar P. Locher. 2021. 'Development of a Universal Nanobody-Binding Fab Module for Fiducial-Assisted Cryo-EM Studies of Membrane Proteins'. *Proceedings of the National Academy of Sciences of the United States of America* 118 (47). https://doi.org/10.1073/pnas.2115435118.

Bocian-Ostrzycka, Katarzyna M., Magdalena J. Grzeszczuk, Anna M. Banaś, and Elżbieta Katarzyna Jagusztyn-Krynicka. 2017. 'Bacterial Thiol Oxidoreductases — from Basic Research to New Antibacterial Strategies'. *Applied Microbiology and Biotechnology* 101 (10): 3977–89. https://doi.org/10.1007/s00253-017-8291-8.

Bosdriesz, Evert, Stefanía Magnúsdõttir, Frank J. Bruggeman, Bas Teusink, and Douwe Molenaar. 2015. 'Binding Proteins Enhance Specific Uptake Rate by Increasing the Substrate-Transporter Encounter Rate'. *FEBS Journal* 282 (12): 2394–2407. https://doi.org/10.1111/febs.13289.

Boudker, Olga, and Se Cheol Oh. 2015. 'Isothermal Titration Calorimetry of Ion-Coupled Membrane Transporters'. *Methods* 76: 171–82. https://doi.org/10.1016/j.ymeth.2015.01.012.

Brautigam, Chad A., Ranjit K. Deka, Peter Schuck, Diana R. Tomchick, and Michael V. Norgard. 2012. 'Structural and Thermodynamic Characterization of the Interaction between Two Periplasmic Treponema Pallidum Lipoproteins That Are Components of a TPR-Protein-Associated TRAP Transporter (TPAT)'. *Journal of Molecular Biology* 420 (1–2): 70–86. https://doi.org/10.1016/j.jmb.2012.04.001.

Brautigam, Chad A., Huaying Zhao, Carolyn Vargas, Sandro Keller, and Peter Schuck. 2016. 'Integration and Global Analysis of Isothermal Titration Calorimetry Data for Studying Macromolecular Interactions'. *Nat Protoc* 11: 882–94. https://doi.org/10.1038/nprot.2016.044.Integration.

Broom, Helen R., Jessica A.O. Rumfeldt, Kenrick A. Vassall, and Elizabeth M. Meiering. 2015. 'Destabilization of the Dimer Interface Is a Common Consequence of Diverse ALS-Associated Mutations in Metal Free SOD1'. *Protein Science* 24 (12): 2081–89. https://doi.org/10.1002/pro.2803.

Brunengraber, Henri. 2007. 'α-Ketoglutarate Is the Precursor of L-2-Hydroxyglutarate'. *Journal of Inherited Metabolic Disease* 30 (5): 628. https://doi.org/10.1007/s10545-007-9984-4.

Chang, Yung Ning, Eva A. Jaumann, Katrin Reichel, Julia Hartmann, Dominik Oliver, Gerhard

Hummer, Benesh Joseph, and Eric R. Geertsma. 2019. 'Structural Basis for Functional Interactions in Dimers of SLC26 Transporters'. *Nature Communications* 10 (1). https://doi.org/10.1038/s41467-019-10001-w.

Dallos, Peter, and Bernd Fakler. 2002. 'Prestin, a New Type of Motor Protein'. *Nature Reviews Molecular Cell Biology* 3 (2): 104–11. https://doi.org/10.1038/nrm730.

Datsenko, Kirill A., and Barry L. Wanner. 2000. 'One-Step Inactivation of Chromosomal Genes in Escherichia Coli K-12 Using PCR Products'. *Proceedings of the National Academy of Sciences of the United States of America* 97 (12): 6640–45. https://doi.org/10.1073/pnas.120163297.

Davidson, Amy L., and Peter C. Maloney. 2007. 'ABC Transporters: How Small Machines Do a Big Job'. *Trends in Microbiology* 15 (10): 448–55. https://doi.org/10.1016/j.tim.2007.09.005.

Davies, James S., Michael J. Currie, and Rachel A. North. 2023. 'Structure and Mechanism of a Tripartite ATP- Independent Periplasmic TRAP Transporter'. *Nature Communications*, 1–41. https://doi.org/https://doi.org/10.1038/s41467-023-36590-1.

Deamer, David. 2016. 'Membranes and the Origin of Life: A Century of Conjecture'. *Journal of Molecular Evolution* 83 (5–6): 159–68. https://doi.org/10.1007/s00239-016-9770-8.

Deka, Ranjit K., Chad A. Brautigam, Martin Goldberga, Peter Schuck, Diana R. Tomchick, and Michael V. Norgard. 2012. 'Structural, Bioinformatic, and in Vivo Analyses of Two Treponema Pallidum Lipoproteins Reveal a Unique TRAP Transporter'. *J Mol Biol.*, no. 416: 678–96. https://doi.org/10.1016/j.jmb.2012.01.015.Structural.

Detmers, Jan. 2001. *The Oligopeptide Transport System of Lactococcus Lactis*.

Dinh, Thuy, and Thomas G. Bernhardt. 2011. 'Using Superfolder Green Fluorescent Protein for Periplasmic Protein Localization Studies'. *Journal of Bacteriology* 193 (18): 4984–87. https://doi.org/10.1128/JB.00315-11.

Doeven, Mark K., Rupert Abele, Robert Tampé, and Bert Poolman. 2004. 'The Binding Specificity of OppA Determines the Selectivity of the Oligopeptide ATP-Finding Cassette Transporter'. *Journal of Biological Chemistry* 279 (31): 32301–7. https://doi.org/10.1074/jbc.M404343200.

Doeven, Mark K. 2006. 'ABC Transporters : Lessons from a Bacterial Oligopeptide Uptake System'.

Don, R. H., P. T. Cox, B. J. Wainwright, K. Baker, and J. S. Mattick. 1991. '"Touchdown" PCR to Circumvent Spurious Priming during Gene Amplification'. *Nucleic Acids Research* 19 (14): 4008. https://doi.org/10.1093/nar/19.14.4008.

Dorman, G., P. J. Hergenrother, C. Herrmann, J. Kuhlmann, A. Levitzki, S. F. Martin, G. Miiller, M. Thutewohl, and H Waldmann. 2001. *Topics in Current Chemistry: Bioorganic Chemistry of Biological Signal Transduction*. Edited by Herbert Waldmann. *Topics in Current Chemistry*. Vol. 211.

Drew, David, and Olga Boudker. 2016. 'Shared Molecular Mechanisms of Membrane Transporters'. *Annual Review of Biochemistry* 85: 543–72. https://doi.org/10.1146/annurev-biochem-060815-014520.

Drew, David, Linda Fröderberg, Louise Baars, and Jan Willem L. De Gier. 2003. 'Assembly and Overexpression of Membrane Proteins in Escherichia Coli'. *Biochimica et Biophysica Acta - Biomembranes* 1610 (1): 3–10. https://doi.org/10.1016/S0005-2736(02)00707-1.

Erban, Radek. 2016. ‘Coupling All-Atom Molecular Dynamics Simulations of Ions in Water with Brownian Dynamics’. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 472 (2186). https://doi.org/10.1098/rspa.2015.0556.

Erkens, Guus B., Maria Majsnerowska, Josy Ter Beek, and Dirk Jan Slotboom. 2012. ‘Energy Coupling Factor-Type ABC Transporters for Vitamin Uptake in Prokaryotes’. *Biochemistry* 51 (22): 4390–96. https://doi.org/10.1021/bi300504v.

Fan, Chengcheng, Jens T. Kaiser, and Douglas C. Rees. 2020. ‘A Structural Framework for Unidirectional Transport by a Bacterial ABC Exporter’. *Proceedings of the National Academy of Sciences of the United States of America* 117 (32): 19228–36. https://doi.org/10.1073/pnas.2006526117.

Fischer, Marcus, Qian Yi Zhang, Roderick E. Hubbard, and Gavin H. Thomas. 2010. ‘Caught in a TRAP: Substrate-Binding Proteins in Secondary Transport’. *Trends in Microbiology* 18 (10): 471–78. https://doi.org/10.1016/j.tim.2010.06.009.

Forloni, Matteo, Alex Y. Liu, and Narendra Wajapeyee. 2019. ‘Megaprimer Polymerase Chain Reaction (PCR)-Based Mutagenesis’. *Cold Spring Harbor Protocols* 2019 (6): 513–15. https://doi.org/10.1101/pdb.prot097824.

Forrest, Lucy R. 2015. ‘STRUCTURAL SYMMETRY IN MEMBRANE PROTEINS’. *Annu Rev Biophys.*, no. 44: 311–37. https://doi.org/10.1146/annurev-biophys-051013-023008.

Forsström, Björn, Barbara Bisławska Axnäs, Johan Rockberg, Hanna Danielsson, Anna Bohlin, and Mathias Uhlen. 2015. ‘Dissecting Antibodies with Regards to Linear and Conformational Epitopes’. *PLoS ONE* 10 (3): 1–11. https://doi.org/10.1371/journal.pone.0121673.

Forward, J A, M C Behrendt, N R Wyborn, R Cross, and D J Kelly. 1997. ‘TRAP Transporters: A New Family of Periplasmic Solute Transport Systems Encoded by the [i]DctPQM[/i] Genes of [i]Rhodobacter Capsulatus[/i] and by Homologs in Diverse Gram-Negative Bacteria’. *J Bacteriol* 179 (17): 5482–93. https://doi.org/10.1128/JB.179.17.5482-5493.1997.

Garaeva, Alisa A., and Dirk J. Slotboom. 2020. ‘Elevator-Type Mechanisms of Membrane Transport’. *Biochemical Society Transactions* 48 (3): 1227–41. https://doi.org/10.1042/BST20200290.

Gasteiger, E., C. Hoogland, A. Gattiker, S. Duvaud, M.R. Wilkins, R.D. Appel, and A. Bairoch. 2005. ‘Protein Identification and Analysis Tools in the ExPASy Server.’ In *The Proteomics Protocols Handbook*, 571–607.

Gauthier, M.J., R Christen, L Fernandez, and M Acquaviva. 1992. ‘Maninobacter Hydrocarbonoclasticus Gen. Nov,, Sp. Nov., a New, Extremely Halotolerant, Hydrocarbon-Degrading Marine Bacterium’. *Growth Lakeland* 42 (4 mM): 568–76. http://ijs.sgmjournals.org/content/42/4/568.short.

Gautom, Trishnamoni, Dharmendra Dheeman, Colin Levy, Thomas Butterfield, Guadalupe Alvarez Gonzalez, Philip Le Roy, Lewis Caiger, Karl Fisher, Linus Johannissen, and Neil Dixon. 2021. ‘Structural Basis of Terephthalate Recognition by Solute Binding Protein TphC’. *Nature Communications* 12 (1): 1–12. https://doi.org/10.1038/s41467-021-26508-0.

Geertsma, Eric R., and Raimund Dutzler. 2011. ‘A Versatile and Efficient High-Throughput Cloning Tool for Structural Biology’. *Biochemistry* 50 (15): 3272–78. https://doi.org/10.1021/bi200178z.

Geertsma, Eric R., N. A.B. Nik Mahmood, Gea K. Schuurman-Wolters, and Bert Poolman. 2008. 'Membrane Reconstitution of ABC Transporters and Assays of Translocator Function'. *Nature Protocols* 3 (2): 256–66. https://doi.org/10.1038/nprot.2007.519.

Geertsma, Eric R, Maarten Groeneveld, Dirk-jan Slotboom, and Bert Poolman. 2008. 'Quality Control of Overexpressed Membrane Proteins' 105 (15).

Grewer, Christof, Armanda Gameiro, and Thomas Rauen. 2014. 'SLC1 Glutamate Transporters'. *Pflugers Arch.* 466 (1): 3–24. https://doi.org/doi:10.1007/s00424-013-1397-7.

Grossmann, Nina, Ahmet S. Vakkasoglu, Sabine Hulpke, Rupert Abele, Rachelle Gaudet, and Robert Tampé. 2014. 'Mechanistic Determinants of the Directionality and Energetics of Active Export by a Heterodimeric ABC Transporter'. *Nature Communications* 5. https://doi.org/10.1038/ncomms6419.

Guan, Lan, and H. Ronald Kaback. 2006. 'Lessons from Lactose Permease'. *Annu Rev Biophys Biomol Struct.* 35: 67–91. https://doi.org/110.1146/annurev.biophys.35.040405.102005.

Guidotti, Guido. 1996. 'ATP Transport and ABC Proteins'. *Chemistry and Biology* 3 (9): 703–6. https://doi.org/10.1016/S1074-5521(96)90244-6.

Gul, Nadia, and Bert Poolman. 2012. 'Functional Reconstitution and Osmoregulatory Properties of the ProU ABC Transporter from Escherichia Coli.' *Molecular Membrane Biology* 30 (2): 138–48. https://doi.org/10.3109/09687688.2012.754060.

Guzman, L. M., D. Belin, M. J. Carson, and J. Beckwith. 1995. 'Tight Regulation, Modulation, and High-Level Expression by Vectors Containing the Arabinose P(BAD) Promoter'. *Journal of Bacteriology* 177 (14): 4121–30. https://doi.org/10.1128/jb.177.14.4121-4130.1995.

Hall, Barry G. 2013. 'Building Phylogenetic Trees from Molecular Data with MEGA'. *Molecular Biology and Evolution* 30 (5): 1229–35. https://doi.org/10.1093/molbev/mst012.

Handley, Kim M., and Jonathan R. Lloyd. 2013. 'Biogeochemical Implications of the Ubiquitous Colonization of Marine Habitats and Redox Gradients by Marinobacter Species'. *Frontiers in Microbiology* 4 (MAY): 1–10. https://doi.org/10.3389/fmicb.2013.00136.

Heide, Tiemen van der, and Bert Poolman. 2002. 'ABC Transporters: One, Two or Four Extracytoplasmic Substrate-Binding Sites?' *EMBO Reports* 3 (10): 938–43. http://www.nature.com/embor/journal/v3/n10/abs/embor057.html.

Hofmann, Susanne, Dovile Januliene, Ahmad R. Mehdipour, Christoph Thomas, Erich Stefan, Stefan Brüchert, Benedikt T. Kuhn, et al. 2019. *Conformation Space of a Heterodimeric ABC Exporter under Turnover Conditions*. *Nature*. Vol. 571. https://doi.org/10.1038/s41586-019-1391-0.

Holzhueter, Katharina, and Eric R Geertsma. 2020. 'Functional (Un)Cooperativity in Elevator Transport Proteins'. *Biochem Soc*, no. 48: 1047–55.

Holzhüter, Katharina, and Eric R. Geertsma. 2022. 'Uniport, Not Proton-Symport, in a Non-Mammalian SLC23 Transporter'. *Journal of Molecular Biology* 434 (2). https://doi.org/10.1016/j.jmb.2021.167393.

Hopkins, Adam P., Judith A. Hawkhead, and Gavin H. Thomas. 2013. 'Transport and Catabolism of the Sialic Acids N-Glycolylneuraminic Acid and 3-Keto-3-Deoxy-d-Glycero-d-Galactonononic Acid by Escherichia Coli K-12'. *FEMS Microbiology Letters* 347 (1): 14–22.

https://doi.org/10.1111/1574-6968.12213.

Hosaka, Masaru, Naofumi Kamimura, Shotaro Toribami, Kosuke Mori, Daisuke Kasai, Masao Fukuda, and Eiji Masai. 2013. 'Novel Tripartite Aromatic Acid Transporter Essential for Terephthalate Uptake in Comamonas Sp. Strain E6'. *Applied and Environmental Microbiology* 79 (19): 6148–55. https://doi.org/10.1128/AEM.01600-13.

Hosie, A. H.F., D. Allaway, M. A. Jones, D. L. Walshaw, A. W.B. Johnston, and P. S. Poole. 2001. 'Solute-Binding Protein-Dependent ABC Transporters Are Responsible for Solute Efflux in Addition to Solute Uptake'. *Molecular Microbiology* 40 (6): 1449–59. https://doi.org/10.1046/j.1365-2958.2001.02497.x.

Hvorup, Rikki N, Birke A Goetz, Martina Niederer, Kaspar Hollenstein, Eduardo Perozo, and Kaspar P Locher. 2007. 'Asymmetry in the Structure of the ABC Transporter–Binding Protein Complex BtuCD-BtuF'. *Science* 317 (September): 1387–91.

Iverson, Tina M., César Luna-Chavez, Gary Cecchini, and Douglas C. Rees. 1999. 'Structure of the Escherichia Coli Fumarate Reductase Respiratory Complex'. *Science* 284 (5422): 1961–66. https://doi.org/10.1126/science.284.5422.1961.

Jacobs, Mariken H.J., Tiemen Van Der Heide, Arnold J.M. Driessen, and Wil N. Konings. 1996. 'Glutamate Transport in Rhodobacter Sphaeroides Is Mediated by a Novel Binding Protein-Dependent Secondary Transport System'. *Proceedings of the National Academy of Sciences of the United States of America* 93 (23): 12786–90. https://doi.org/10.1073/pnas.93.23.12786.

Janausch, I. G., E. Zientz, Q. H. Tran, A. Kröger, and G. Unden. 2001. 'C4-Dicarboxylate Carriers and Sensors in Bacteria'. *Biochimica et Biophysica Acta - Bioenergetics* 1553 (1–2): 39–56. https://doi.org/10.1016/S0005-2728(01)00233-X.

Janausch, Ingo G., Inma Garcia-Moreno, and Gottfried Unden. 2002. 'Function of DcuS from Escherichia Coli as a Fumarate-Stimulated Histidine Protein Kinase in Vitro'. *Journal of Biological Chemistry* 277 (42): 39809–14. https://doi.org/10.1074/jbc.M204482200.

Johnston, Jason W., Nathan P. Coussens, Simon Allen, Jon C.D. Houtman, Keith H. Turner, Anthony Zaleski, S. Ramaswamy, Bradford W. Gibson, and Michael A. Apicella. 2008. '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* 283 (2): 855–65. https://doi.org/10.1074/jbc.M706603200.

Jumper, John, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, et al. 2021. 'Highly Accurate Protein Structure Prediction with AlphaFold'. *Nature* 596 (7873): 583–89. https://doi.org/10.1038/s41586-021-03819-2.

Kalliri, Efthalia, Scott B. Mulrooney, and Robert P. Hausinger. 2008. 'Identification of Escherichia Coli YgaF as an L-2-Hydroxyglutarate Oxidase'. *Journal of Bacteriology* 190 (11): 3793–98. https://doi.org/10.1128/JB.01977-07.

Kelley, Lawrence A., Stefans Mezulis, Christopher M. Yates, Mark N. Wass, and Michael J.E. Sternberg. 2015. 'The Phyre2 Web Portal for Protein Modeling, Prediction and Analysis'. *Nature Protocols* 10 (6): 845–58. https://doi.org/10.1038/nprot.2015.053.

Kelly, David J., and Gavin H. Thomas. 2001. 'The Tripartite ATP-Independent Periplasmic (TRAP)

Transporters of Bacteria and Archaea'. *FEMS Microbiology Reviews* 25 (4): 405–24. https://doi.org/10.1016/S0168-6445(01)00061-4.

Kim, Ok Bin. 2006. 'Carrier Und Regulatoren Des Tartrat- Und C 4 -Dicarboxylatstoffwechsels von Escherichia Coli'. Johannes Gutenberg-Universität Mainz.

Kranz, James K., and Celine Schalk-Hihi. 2011. *Protein Thermal Shifts to Identify Low Molecular Weight Fragments*. *Methods in Enzymology*. 1st ed. Vol. 493. Elsevier Inc. https://doi.org/10.1016/B978-0-12-381274-2.00011-X.

Kuhn, Benedikt T. 2020. 'Tuning an Elevator'. Goethe Universität Frankfurt.

Kuhn, Benedikt T., Iwan Zimmermann, Roger J.P. Dawson, Markus A. Seeger, and Eric R. Geertsma. 2018. 'Kamel-Antikörper Aus Dem Reagenzglas'. *BioSpektrum* 24 (7): 688–90. https://doi.org/10.1007/s12268-018-0982-x.

Kuhn, Benedikt T., Iwan Zimmermann, Pascal Egloff, Lea M. Hürlimann, Cedric A. J. Hutter, Christian Miscenic, Roger J. P. Dawson, and and Eric R. Geertsma Markus A. Seeger. 2020. 'Biotinylation of Membrane Proteins for Binder Selections'. In *Expression, Purification, and Structural Biology of Membrane Proteins, Methods in Molecular Biology*, 2127:151–65. https://doi.org/doi.org/10.1007/978-1-0716-0373-4_11.

Laskowski, Roman A., and Mark B. Swindells. 2011. 'LigPlot+: Multiple Ligand-Protein Interaction Diagrams for Drug Discovery'. *Journal of Chemical Information and Modeling* 51 (10): 2778–86. https://doi.org/10.1021/ci200227u.

Lemoine, Frédéric, Damien Correia, Vincent Lefort, Olivia Doppelt-Azeroual, Fabien Mareuil, Sarah Cohen-Boulakia, and Olivier Gascuel. 2019. 'NGPhylogeny.Fr: New Generation Phylogenetic Services for Non-Specialists'. *Nucleic Acids Research* 47 (W1): W260–65. https://doi.org/10.1093/nar/gkz303.

Lewinson, Oded, Allen T Lee, Kaspar P Locher, and Douglas C Rees. 2010. 'By the Dynamics of Complex Formation'. *Technology* 17 (3): 332–38. https://doi.org/10.1038/nsmb.1770.A.

Lewinson, Oded, and Nurit Livnat-Levanon. 2017. 'Mechanism of Action of ABC Importers: Conservation, Divergence, and Physiological Adaptations'. *Journal of Molecular Biology* 429 (5): 606–19. https://doi.org/10.1016/j.jmb.2017.01.010.

Lipowsky, Reinhard. 1991. 'The Conformation of Membranes'. *Nature* 349 (6309): 475–81. https://doi.org/10.1038/349475a0.

Locher, Kaspar P. 2016. 'Mechanistic Diversity in ATP-Binding Cassette (ABC) Transporters'. *Nature Structural and Molecular Biology* 23 (6): 487–93. https://doi.org/10.1038/nsmb.3216.

Madden, Tom. 2002. 'Chapter 16 : The BLAST Sequence Analysis Tool'. *The NCBI Handbook[Internet]*, no. Md: 1–15.

Marino, Jacopo, Katharina Holzhüter, Benedikt Kuhn, and Eric R. Geertsma. 2017. 'Efficient Screening and Optimization of Membrane Protein Production in Escherichia Coli'. *Methods in Enzymology* 594: 139–64. https://doi.org/10.1016/bs.mie.2017.05.011.

Mayfield, John E., Betsy J. Bricker, Holly Godfrey, Renae M. Crosby, David J. Knight, Shirley M. Hailing, Doris Balinsky, and Louisa B. Tabatabai. 1988. 'The Cloning, Expression, and Nucleotide Sequence of a Gene Coding for an Immunogenic Bruceh Abortus Protein'. *Elsevier*

63.

Milo, Ron, and Rob Phillips. 2015. 'Cell Biology by the Numbers'. *Cell Biology by the Numbers*. https://doi.org/10.1201/9780429258770.

Moussatova, Anastassiia, Christian Kandt, Megan L. O'Mara, and D. Peter Tieleman. 2008. 'ATP-Binding Cassette Transporters in Escherichia Coli'. *Biochimica et Biophysica Acta - Biomembranes* 1778 (9): 1757–71. https://doi.org/10.1016/j.bbamem.2008.06.009.

Müller, Axel, Emmanuele Severi, Christopher Mulligan, Andrew G. Watts, David J. Kelly, Keith S. Wilson, Anthony J. Wilkinson, and Gavin H. Thomas. 2006. '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* 281 (31): 22212–22. https://doi.org/10.1074/jbc.M603463200.

Mulligan, Christopher, Cristina Fenollar-Ferrer, Gabriel A. Fitzgerald, Ariela Vergara- Jaque, Desirée Kaufmann, Yan Li, Lucy R. Forrest, and Mindell Joseph A. 2017. 'The Bacterial Dicarboxylate Transporter, VcINDY, Uses a Two- Domain Elevator-Type Mechanism'. *Nat Struct Mol Biol.* 23 (3): 256–63. https://doi.org/10.1038/nsmb.3166.

Mulligan, Christopher, Marcus Fischer, and Gavin H. Thomas. 2011. 'Tripartite ATP-Independent Periplasmic (TRAP) Transporters in Bacteria and Archaea'. *FEMS Microbiology Reviews* 35 (1): 68–86. https://doi.org/10.1111/j.1574-6976.2010.00236.x.

Mulligan, Christopher, Gabriel A. Fitzgerald, Da Neng Wang, and Joseph A. Mindell. 2014. 'Functional Characterization of a Na+-Dependent Dicarboxylate Transporter from Vibrio Cholerae'. *Journal of General Physiology* 143 (6): 745–59. https://doi.org/10.1085/jgp.201311141.

Mulligan, Christopher, Eric R. Geertsma, Emmanuele Severi, David J. Kelly, Bert Poolman, and Gavin H. Thomas. 2009. 'The Substrate-Binding Protein Imposes Directionality on an Electrochemical Sodium Gradient-Driven TRAP Transporter'. *Proceedings of the National Academy of Sciences of the United States of America* 106 (6): 1778–83. https://doi.org/10.1073/pnas.0809979106.

Mulligan, Christopher, David J. Kelly, and Gavin H. Thomas. 2007. 'Tripartite ATP-Independent Periplasmic Transporters: Application of a Relational Database for Genome-Wide Analysis of Transporter Gene Frequency and Organization'. *Journal of Molecular Microbiology and Biotechnology* 12 (3–4): 218–26. https://doi.org/10.1159/000099643.

Mulligan, Christopher, Andrew P. Leech, David J. Kelly, and Gavin H. Thomas. 2012. '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'. *Journal of Biological Chemistry* 287 (5): 3598–3608. https://doi.org/10.1074/jbc.M111.281030.

Nguyen, Bidong D., Miguelangel Cuenca V., Johannes Hartl, Ersin Gül, Rebekka Bauer, Susanne Meile, Joel Rüthi, et al. 2020. 'Import of Aspartate and Malate by DcuABC Drives H2/Fumarate Respiration to Promote Initial Salmonella Gut-Lumen Colonization in Mice'. *Cell Host and Microbe* 27 (6): 922-936.e6. https://doi.org/10.1016/j.chom.2020.04.013.

Niesen, Frank H., Helena Berglund, and Masoud Vedadi. 2007. 'The Use of Differential Scanning Fluorimetry to Detect Ligand Interactions That Promote Protein Stability'. *Nature Protocols* 2 (9): 2212–21. https://doi.org/10.1038/nprot.2007.321.

Oldham, Michael L., Shanshuang Chenb, and Jue Chen. 2013. 'Structural Basis for Substrate Specificity in the Escherichia Coli Maltose Transport System'. *PNAS*.

Oldham, Michael L., Dheeraj Khare, Florante A. Quiocho, Amy L. Davidson, and Jue Chen. 2007. 'Crystal Structure of a Catalytic Intermediate of the Maltose Transporter'. *Nature* 450 (7169): 515–21. https://doi.org/10.1038/nature06264.

Paetzel, M. 2019. *Bacterial Cell Walls and Membranes, Bacterial Signal Peptides*. Edited by Andreas Kuhn. *Subcellular Biochemistry*. https://doi.org/https://doi.org/10.1007/978-3-030-18768-2 ©.

Paetzel, Mark. 2013. 'Structure and Mechanism of Escherichia Coli Type I Signal Peptidase'. *Biochimica et Biophysica Acta - Molecular Cell Research* 1843 (8): 1497–1508. https://doi.org/10.1016/j.bbamcr.2013.12.003.

Pardon, Els, Toon Laeremans, Sarah Triest, Søren G.F. Rasmussen, Alexandre Wohlkönig, Armin Ruf, Serge Muyldermans, Wim G.J. Hol, Brian K. Kobilka, and Jan Steyaert. 2014. 'A General Protocol for the Generation of Nanobodies for Structural Biology'. *Nature Protocols* 9 (3): 674–93. https://doi.org/10.1038/nprot.2014.039.

Parshina, Sofiya N., Jan Sipma, Yutaka Nakashimada, Anne Meint Henstra, Hauke Smidt, Anatoly M. Lysenko, Piet N.L. Lens, Gatze Lettinga, and Alfons J.M. Stams. 2005. 'Desulfotomaculum Carboxydivorans Sp. Nov., a Novel Sulfate-Reducing Bacterium Capable of Growth at 100% CO'. *International Journal of Systematic and Evolutionary Microbiology* 55 (5): 2159–65. https://doi.org/10.1099/ijs.0.63780-0.

Pearson, Melanie M., Alejandra Yep, Sara N. Smith, and Harry L.T. Mobley. 2011. 'Transcriptome of Proteus Mirabilis in the Murine Urinary Tract: Virulence and Nitrogen Assimilation Gene Expression'. *Infection and Immunity* 79 (7): 2619–31. https://doi.org/10.1128/IAI.05152-11.

Peter, Martin F., Jan A. Ruland, Peer Depping, Niels Schneberger, Emmanuele Severi, Jonas Moecking, Karl Gatterdam, et al. 2022. 'Structural and Mechanistic Analysis of a Tripartite ATP-Independent Periplasmic TRAP Transporter'. *Nature Communications* 13 (1): 1–15. https://doi.org/10.1038/s41467-022-31907-y.

Peter, Martin F, Jan A Ruland, Yeojin Kim, Philipp Hendricks, and Jan Peter Siebrasse. 2023. 'Conformational Coupling of the Sialic Acid TRAP Transporter HiSiaQM with Its Substrate Binding Protein HiSiaP'. *BioRxiv*, 1–28. https://doi.org/https://doi.org/10.1101/2023.03.04.531103.

Phillips, Rob. 2018. *Membranes by the Numbers, Physics of Biological Membranes*.

Plantinga, Titia H., Chris van der Does, Danuta Tomkiewicz, Geertje van Keulen, Wil N. Konings, and Arnold J.M. Driessen. 2005. 'Deletion of the YiaMNO Transporter Genes Affects the Growth Characteristics of Escherichia Coli K-12'. *Microbiology* 151 (5): 1683–89. https://doi.org/10.1099/mic.0.27851-0.

Price, Morgan N., Katherine H. Huang, Eric J. Alm, and Adam P. Arkin. 2005. 'A Novel Method for Accurate Operon Predictions in All Sequenced Prokaryotes'. *Nucleic Acids Research* 33 (3):

880–92. https://doi.org/10.1093/nar/gki232.

Priest, Fergus G., and Margaret Barker. 2010. 'Gram-Negative Bacteria Associated with Brewery Yeasts: Reclassification of Obesumbacterium Proteus Biogroup 2 as Shimwellia Pseudoproteus Gen. Nov., Sp. Nov., and Transfer of Escherichia Blattae to Shimwellia Blattae Comb. Nov.' *International Journal of Systematic and Evolutionary Microbiology* 60 (4): 828–33. https://doi.org/10.1099/ijs.0.013458-0.

Ramahi, A. A., and R. L. Ruff. 2014. 'Membrane Potential'. *Encyclopedia of the Neurological Sciences* 3: 1034–35. https://doi.org/10.1016/B978-0-12-385157-4.00062-2.

Rapp, Chloe, Xiaoyun Bai, and Reinhart A.F. Reithmeier. 2017. 'Molecular Analysis of Human Solute Carrier SLC26 Anion Transporter Disease-Causing Mutations Using 3-Dimensional Homology Modeling'. *Biochimica et Biophysica Acta - Biomembranes* 1859 (12): 2420–34. https://doi.org/10.1016/j.bbamem.2017.09.016.

Rempel, S., W. K. Stanek, and D. J. Slotboom. 2019. 'ECF-Type ATP-Binding Cassette Transporters'. *Annual Review of Biochemistry* 88: 551–76. https://doi.org/10.1146/annurev-biochem-013118-111705.

Reuss, Luis, and Guillermo A. Altenberg. 2008. 'Mechanisms of Ion Transport across Cell Membranes and Epithelia'. *Seldin and Giebisch's The Kidney: Physiology and Pathophysiology*, 45–66. https://doi.org/10.1016/B978-0-12-381462-3.00002-1.

Rimon, Abraham, Tzvi Tzubery, and Etana Padan. 2007. 'Monomers of the NhaA Na+/H+ Antiporter of Escherichia Coli Are Fully Functional yet Dimers Are Beneficial under Extreme Stress Conditions at Alkaline PH in the Presence of Na+ or Li+'. *Journal of Biological Chemistry* 282 (37): 26810–21. https://doi.org/10.1074/jbc.M704469200.

Rosa, Leonardo, Samuel R. Dix, John B. Rafferty, and David J. Kelly. 2019. 'A New Mechanism for High-Affinity Uptake of C4-Dicarboxylates in Bacteria Revealed by the Structure of Rhodopseudomonas Palustris MatC (RPA3494), a Periplasmic Binding Protein of the Tripartite Tricarboxylate Transporter (TTT) Family'. *Journal of Molecular Biology* 431 (2): 351–67. https://doi.org/10.1016/j.jmb.2018.11.016.

Rosa, Leonardo T., Matheus E. Bianconi, Gavin H. Thomas, and David J. Kelly. 2018. 'Tripartite ATP-Independent Periplasmic (TRAP) Transporters and Tripartite Tricarboxylate Transporters (TTT): From Uptake to Pathogenicity'. *Frontiers in Cellular and Infection Microbiology* 8 (FEB). https://doi.org/10.3389/fcimb.2018.00033.

Rosa, Leonardo T., Samuel R. Dix, John B. Rafferty, and David J. Kelly. 2017. 'Structural Basis for High-Affinity Adipate Binding to AdpC (RPA4515), an Orphan Periplasmic-Binding Protein from the Tripartite Tricarboxylate Transporter (TTT) Family in Rhodopseudomonas Palustris'. *FEBS Journal* 284 (24): 4262–77. https://doi.org/10.1111/febs.14304.

Rosa, Leonardo T., Vicki Springthorpe, Matheus E. Bianconi, Gavin H. Thomas, and David J. Kelly. 2018. 'Massive Over-Representation of Solute-Binding Proteins (SBPs) from the Tripartite Tricarboxylate Transporter (TTT) Family in the Genome of the α-Proteobacterium Rhodoplanes Sp. Z2-YC6860'. *Microbial Genomics* 4 (5). https://doi.org/10.1099/mgen.0.000176.

Rosano, Germán L., and Eduardo A. Ceccarelli. 2014. 'Recombinant Protein Expression in

Escherichia Coli: Advances and Challenges'. *Frontiers in Microbiology* 5 (APR): 1–17. https://doi.org/10.3389/fmicb.2014.00172.

Saier, Milton H., Vamsee S. Reddy, Gabriel Moreno-Hagelsieb, Kevin J. Hendargo, Yichi Zhang, Vasu Iddamsetty, Katie Jing Kay Lam, et al. 2021. 'The Transporter Classification Database (TCDB): 2021 Update'. *Nucleic Acids Research* 49 (D1): D461–67. https://doi.org/10.1093/nar/gkaa1004.

Saier, Milton H., Can V. Tran, and Ravi D. Barabote. 2006. 'TCDB: The Transporter Classification Database for Membrane Transport Protein Analyses and Information.' *Nucleic Acids Research* 34 (Database issue): 181–86. https://doi.org/10.1093/nar/gkj001.

Salvador, J. Pablo, Lluïsa Vilaplana, and M. Pilar Marco. 2019. 'Nanobody: Outstanding Features for Diagnostic and Therapeutic Applications'. *Analytical and Bioanalytical Chemistry* 411 (9): 1703–13. https://doi.org/10.1007/s00216-019-01633-4.

Santos, Rita, Oleg Ursu, Anna Gaulton, A Patrícia Bento, Ramesh S Donadi, G Bologa, Anneli Karlsson, Bissan Al-lazikani, Anne Hersey, and Tudor I Oprea. 2017. 'A Comprehensive Map of Molecular Drug Targets'. *Nature Reviews Drug Discovery* 16 (1): 19–34. https://doi.org/10.1038/nrd.2016.230.A.

Sauer, David B., Jennifer J. Marden, Joseph C. Sudar, Jinmei Song, Christopher Mulligan, and Da Neng Wang. 2022. 'Structural Basis of Ion – Substrate Coupling in the Na+-Dependent Dicarboxylate Transporter VcINDY'. *Nature Communications* 13 (1). https://doi.org/10.1038/s41467-022-30406-4.

Saurin, William, Maurice Hofnung, and Elie Dassa. 1999. 'Getting in or out: Early Segregation between Importers and Exporters in the Evolution of ATP-Binding Cassette (ABC) Transporters'. *Journal of Molecular Evolution* 48 (1): 22–41. https://doi.org/10.1007/PL00006442.

Schaffer, Jessica N., and Melanie M. Pearson. 2015. 'Proteus Mirabilis and Urinary Tract Infections Jessica'. *Microbiol Spectr.* 3. https://doi.org/10.1128/microbiolspec.UTI-0017-2013.Proteus.

Scheepers, Giel H., Jelger A. Lycklama a Nijeholt, and Bert Poolman. 2016. 'An Updated Structural Classification of Substrate-Binding Proteins'. *FEBS Letters* 590 (23): 4393–4401. https://doi.org/10.1002/1873-3468.12445.

Schneider, Caroline A., Wayne S. Rasband, and Kevin W. Eliceiri. 2012. 'NIH Image to ImageJ: 25 Years of Image Analysis'. *Nature Methods* 9 (7): 671–75. https://doi.org/10.1038/nmeth.2089.

Scott, Duncan E., Christina Spry, and Chris Abell. 2016. 'Differential Scanning Fluorimetry as Part of a Biophysical Screening Cascade', 139–72. https://doi.org/10.1002/9783527683604.ch07.

Seol, Wongi, and Aaron J. Shatkin. 1991. 'Escherichia Coli KgtP Encodes an α-Ketoglutarate Transporter'. *Proceedings of the National Academy of Sciences of the United States of America* 88 (9): 3802–6. https://doi.org/10.1073/pnas.88.9.3802.

Severi, Emmanuele, Gaynor Randle, Polly Kivlin, Kate Whitfield, Rosie Young, Richard Moxon, David Kelly, Derek Hood, and Gavin H. Thomas. 2005. 'Sialic Acid Transport in Haemophilus Influenzae Is Essential for Lipopolysaccharide Sialylation and Serum Resistance and Is Dependent on a Novel Tripartite ATP-Independent Periplasmic Transporter'. *Molecular*

Microbiology 58 (4): 1173–85. https://doi.org/10.1111/j.1365-2958.2005.04901.x.

Shi, Yigong. 2013. 'Common Folds and Transport Mechanisms of Secondary Active Transporters'. *Annual Review of Biophysics* 42 (1): 51–72. https://doi.org/10.1146/annurev-biophys-083012-130429.

Shilton, Brian H. 2008. 'The Dynamics of the MBP-MalFGK2 Interaction: A Prototype for Binding Protein Dependent ABC-Transporter Systems'. *Biochimica et Biophysica Acta - Biomembranes* 1778 (9): 1772–80. https://doi.org/10.1016/j.bbamem.2007.09.005.

Shohei Koide. 2009. 'Engineering of Recombinant Crystallization Chaperones'. *Curr Opin Struct Biol.* 19 (4): 449–57. https://doi.org/10.1016/j.sbi.2009.04.008.

Shuman, H. A. 1982. 'Active Transport of Maltose in Escherichia Coli K12. Role of the Periplasmic Maltose-Binding Protein and Evidence for a Substrate Recognition Site in the Cytoplasmic Membrane'. *Journal of Biological Chemistry* 257 (10): 5455–61. https://doi.org/10.1016/s0021-9258(19)83799-7.

Sievers, Fabian, Andreas Wilm, David Dineen, Toby J. Gibson, Kevin Karplus, Weizhong Li, Rodrigo Lopez, et al. 2011. 'Fast, Scalable Generation of High-Quality Protein Multiple Sequence Alignments Using Clustal Omega'. *Molecular Systems Biology* 7 (539). https://doi.org/10.1038/msb.2011.75.

Sigal, Nadejda, Devora Cohen-Karni, Shira Siemion, and Eitan Bibi. 2006. 'MdfA from Escherichia Coli, a Model Protein for Studying Secondary Multidrug Transport'. *Journal of Molecular Microbiology and Biotechnology* 11 (6): 308–17. https://doi.org/10.1159/000095633.

Silberberg, Jakob M., Robin A. Corey, Lisa Hielkema, Charlott Stock, Phillip J. Stansfeld, Cristina Paulino, and Inga Hänelt. 2021. 'Deciphering Ion Transport and ATPase Coupling in the Intersubunit Tunnel of KdpFABC'. *Nature Communications* 12 (1): 1–12. https://doi.org/10.1038/s41467-021-25242-x.

Skoczyńska, Anna, Magdalena Lewandowska, Anna Klarowicz, and Waleria Hryniewicz. 2005. 'Prevalence and Serotype Distribution of Encapsulated Haemophilus Influenzae Isolates from Patients with Lower Respiratory Tract Infections in Poland'. *Journal of Clinical Microbiology* 43 (2): 938–41. https://doi.org/10.1128/JCM.43.2.938-941.2005.

Stecula, Adrian, Avner Schlessinger, Kathleen M Giacomini, Andrej Sali, Therapeutic Sciences, San Francisco, San Francisco, and San Francisco. 2017. 'Human Concentrative Nucleoside Transporter 3 (HCNT3, SLC28A3) Forms a Cyclic Homo-Trimer' 56 (27): 3475–83. https://doi.org/doi:10.1021/acs.biochem.7b00339.

Steeland, Sophie, Roosmarijn E. Vandenbroucke, and Claude Libert. 2016. 'Nanobodies as Therapeutics: Big Opportunities for Small Antibodies'. *Drug Discovery Today* 21 (7): 1076–1113. https://doi.org/10.1016/j.drudis.2016.04.003.

Steyaert, Jan, and Brian K Kobilka. 2011. 'Nanobody Stabilization of G Protein Coupled Receptor Conformational States'. *Curr Opin Struct Biol.* 21 (4): 567–72. https://doi.org/110.1016/j.sbi.2011.06.011.

Stigter, E. C.A., G. J. de Jong, and W. P. van Bennekom. 2013. 'Coupling Surface-Plasmon Resonance and Mass Spectrometry to Quantify and to Identify Ligands'. *TrAC - Trends in*

Analytical Chemistry 45: 107–20. https://doi.org/10.1016/j.trac.2012.09.004.

Stock, C., L. Hielkema, I. Tascón, D. Wunnicke, G. T. Oostergetel, M. Azkargorta, C. Paulino, and I. Hänelt. 2018. 'Cryo-EM Structures of KdpFABC Suggest a K+ Transport Mechanism via Two Inter-Subunit Half-Channels'. *Nature Communications* 9 (1): 1–10. https://doi.org/10.1038/s41467-018-07319-2.

Sutcliffe, I. C., and R. R.B. Russell. 1995. 'Lipoproteins of Gram-Positive Bacteria'. *Journal of Bacteriology* 177 (5): 1123–28. https://doi.org/10.1128/jb.177.5.1123-1128.1995.

Takahashi, Hitomi, Eiji Inagaki, Chizu Kuroishi, and Tahir H. Tahirov. 2004. 'Structure of the Thermus Thermophilus Putative Periplasmic Glutamate/Glutamine-Binding Protein'. *Acta Crystallographica Section D: Biological Crystallography* 60 (10): 1846–54. https://doi.org/10.1107/S0907444904019420.

Teufel, Felix, José Juan Almagro Armenteros, Alexander Rosenberg Johansen, Magnús Halldór Gíslason, Silas Irby Pihl, Konstantinos D. Tsirigos, Ole Winther, Søren Brunak, Gunnar von Heijne, and Henrik Nielsen. 2022. 'SignalP 6.0 Predicts All Five Types of Signal Peptides Using Protein Language Models'. *Nature Biotechnology* 40 (7): 1023–25. https://doi.org/10.1038/s41587-021-01156-3.

Thomas, Gavin H., Thomas Southworth, Maria Rocio León-Kempis, and Andrew Leech. 2006. 'Novel Ligands for the Extracellular Solute Receptors of Two Bacterial TRAP Transporters'. *Microbiology* 152 (1): 187–98. https://doi.org/10.1099/mic.0.28334-0.

Tokonami, Natsuko, Luciana Morla, Gabriel Centeno, David Mordasini, Suresh Krishna Ramakrishnan, Svetlana Nikolaeva, Carsten A. Wagner, et al. 2013. 'α-Ketoglutarate Regulates Acid-Base Balance through an Intrarenal Paracrine Mechanism'. *Journal of Clinical Investigation* 123 (7): 3166–71. https://doi.org/10.1172/JCI67562.

Traenkle, Bjoern, and Ulrich Rothbauer. 2017. 'Under the Microscope: Single-Domain Antibodies for Live-Cell Imaging and Super-Resolution Microscopy'. *Frontiers in Immunology* 8 (AUG): 1–8. https://doi.org/10.3389/fimmu.2017.01030.

Trott, Oleg, and Arthur J. Olson. 2010. 'AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization and Multithreading'. *J Comput Chem.* 31 (2): 455–61. https://doi.org/10.1002/jcc.21334.

Tsirigos, Konstantinos D., Christoph Peters, Nanjiang Shu, Lukas Käll, and Arne Elofsson. 2015. 'The TOPCONS Web Server for Consensus Prediction of Membrane Protein Topology and Signal Peptides'. *Nucleic Acids Research* 43 (W1): W401–7. https://doi.org/10.1093/nar/gkv485.

Uchański, Tomasz, Simonas Masiulis, Baptiste Fischer, Valentina Kalichuk, Uriel López-Sánchez, Eleftherios Zarkadas, Miriam Weckener, et al. 2021. 'Megabodies Expand the Nanobody Toolkit for Protein Structure Determination by Single-Particle Cryo-EM'. *Nature Methods* 18 (1): 60–68. https://doi.org/10.1038/s41592-020-01001-6.

Velazquez-Campoy, Adrian, Stephanie A. Leavitt, and Ernesto Freire. 2010. 'Characterization of Protein–Protein Interactions by Isothermal Titration Calorimetry'. In *Methods in Molecular Biology*, 261:35–54. http://files/2933/Ross and Freire - Enthalpy Screening by Isothermal

Titration Calorim.pdf.

Vetting, Matthew W., Nawar Al-Obaidi, Suwen Zhao, Brian San Francisco, Jungwook Kim, Daniel J. Wichelecki, Jason T. Bouvier, et al. 2015. 'Experimental Strategies for Functional Annotation and Metabolism Discovery: Targeted Screening of Solute Binding Proteins and Unbiased Panning of Metabolomes'. *Biochemistry* 54 (3): 909–31. https://doi.org/10.1021/bi501388y.

Vigonsky, Elena, Elena Ovcharenko, and Oded Lewinson. 2013. 'Two Molybdate/Tungstate ABC Transporters That Interact Very Differently with Their Substrate Binding Proteins'. *Proceedings of the National Academy of Sciences of the United States of America* 110 (14): 5440–45. https://doi.org/10.1073/pnas.1213598110.

Visser, Michael, Sofiya N. Parshina, Joana I. Alves, Diana Z. Sousa, Inês A.C. Pereira, Gerard Muyzer, Jan Kuever, et al. 2015. 'Genome Analyses of the Carboxydotrophic Sulfate-Reducers Desulfotomaculum Nigrificans and Desulfotomaculum Carboxydivorans and Reclassification of Desulfotomaculum Caboxydivorans as a Later Synonym of Desulfotomaculum Nigrificans'. *Standards in Genomic Sciences* 9 (3): 655–75. https://doi.org/10.4056/sigs.4718645.

Walter, Justin D., Marta Sawicka, and Raimund Dutzler. 2019. 'Cryo-EM Structures and Functional Characterization of Murine Slc26a9 Reveal Mechanism of Uncoupled Chloride Transport'. *ELife* 8: 1–31. https://doi.org/10.7554/eLife.46986.

Wang, Rong Fu, and Sidney R. Kushner. 1991. 'Construction of Versatile Low-Copy-Number Vectors for Cloning, Sequencing and Gene Expression in Escherichia Coli'. *Gene* 100 (C): 195–99. https://doi.org/10.1016/0378-1119(91)90366-J.

Waterhouse, Andrew M., James B. Procter, David M.A. Martin, Michèle Clamp, and Geoffrey J. Barton. 2009. 'Jalview Version 2-A Multiple Sequence Alignment Editor and Analysis Workbench'. *Bioinformatics* 25 (9): 1189–91. https://doi.org/10.1093/bioinformatics/btp033.

Widenhorn, K. A., J. M. Somers, and W. W. Kay. 1989. 'Genetic Regulation of the Tricarboxylate Transport Operon (TctI) of Salmonella Typhimurium.' *Journal of Bacteriology* 171 (8): 4436–41. https://doi.org/10.1128/jb.171.8.4436-4441.1989.

Winnen, Brit, Rikki N. Hvorup, and Milton H. Saier. 2003. 'The Tripartite Tricarboxylate Transporter (TTT) Family'. *Research in Microbiology* 154 (7): 457–65. https://doi.org/10.1016/S0923-2508(03)00126-8.

Winter, Roland, Frank Noll, and Claus Czeslik. 2011. *Methoden Der Biophysikalischen Chemie*. Vieweg & Teubner.

Wyborn, Neil R., Jesse Alderson, Simon C. Andrews, and David J. Kelly. 2001. 'Topological Analysis of DctQ, the Small Integral Membrane Protein of the C4-Dicarboxylate TRAP Transporter of Rhodobacter Capsulatus'. *FEMS Microbiology Letters* 194 (1): 13–17. https://doi.org/10.1016/S0378-1097(00)00492-4.

Xu, Yanyan, Anna Seelig, and Simon Bernèche. 2017. 'Unidirectional Transport Mechanism in an ATP Dependent Exporter'. *ACS Central Science* 3 (3): 250–58. https://doi.org/10.1021/acscentsci.7b00068.

Yang, Min, Nurit Livnat Levanon, Burçin Acar, Burcu Aykac Fas, Gal Masrati, Jessica Rose, Nir Ben-Tal, Turkan Haliloglu, Yongfang Zhao, and Oded Lewinson. 2018. 'Single-Molecule Probing of

the Conformational Homogeneity of the ABC Transporter BtuCD'. *Nature Chemical Biology* 14 (7): 715–22. https://doi.org/10.1038/s41589-018-0088-2.

Zhou, Unpeng, and John H. Bushweller. 2018. 'Solution Structure and Elevator Mechanism of the Membrane Electron Transporter CcdA'. *Nat Struct Mol Biol.* 176 (1): 139–48. https://doi.org/10.1038/s41594-018-0022-z.Solution.

Zhou, Xiaoming, Elena J Levin, Yaping Pan, Jason G Mccoy, Ruchika Sharma, Brian Kloss, Renato Bruni, Matthias Quick, and Ming Zhou. 2014. 'Structural Basis of the Alternating-Access Mechanism in a Bile Acid Transporter'. *Nature* 505 (7484): 569–73. https://doi.org/10.1038/nature12811.

Zimmermann, Iwan, Pascal Egloff, Cedric A.J. Hutter, Fabian M. Arnold, Peter Stohler, Nicolas Bocquet, Melanie N. Hug, et al. 2018. 'Synthetic Single Domain Antibodies for the Conformational Trapping of Membrane Proteins'. *ELife* 7: 1–32. https://doi.org/10.7554/eLife.34317.

List of abbreviations

aa	amino acids
ABC transporter	ATP binding cassette transporter
ADP	adenosine diphosphate
amp	amp
APS	ammonium persulfate
ara	arabinose
ATP	adenosine triphosphate
CCCP	carbonyl cyanide m-chlorophenyl hydrazone
CAM	chloramphenicol
CDR	complementarity determining region
DDM	n-Dodecyl-ß-maltoside
DM	n-Decyl-ß-maltoside (DM)
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleoside Triphosphate
DSF	differential scanning fluorimetry
E. coli	*Escherichia coli*
ELISA	enzyme linked immunosorbent assay
Fab fragment	antigen binding fragment
For	foward
GFP	green fluorescent protein
GS-linker	glycine-serine-linker
HcAb	heavy-chain antibody
Hepes	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
HP	helical hairpin
HPLC	high-pressure liquid chromatography
HRP	horseradish peroxidase
HRV	human rhinovirus
IMAC	immobilized metal affinity chromatography
IgG	Immunoglobulin G
IPTG	isopropyl β-D-1-thiogalactopyranoside
ITC	isothermal titration calorimetry

K_D	dissociation constant
KPi	potassium phosphate
LB	lysogeny broth
MS	mass spectrometry
MWCO	molecular weight cutoff
NaPi	sodium phosphate
Ni-NTA	nickel nitrilotriacetic acid
OD_{600}	optical density at a wavelength of 600 nm
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDB	protein data bank
PDVF	polyvinylidene fluoride
PMSF	phenylmethylsulfonyl fluoride
Rev	reverse
RMSD	root mean square deviation
RNA	ribonucleic acid
SBP	SBP
scFv	single-chain variable fragment
SDS	sodium dodecyl sulfate
SEC	size-exlcusion chromatography
sfGFP	superfolder green fluorescent protein
SLC	solute carrier
Soy PC	L-α-phosphatidylcholine from soybean
T	temperature
TAE	tris-acetat-EDTA
TAXI	TRAP-associated extracytoplasmic immunogenic
TB	terrific broth
TCA cycle	tricarboxylic acid cycle
TCDB	transporter classification database
TEMED	tetramethylethylenediamine
TM	transmembrane
T_m	melting temperature

TRAP	Tripartite ATP independent perplasmic
Tris	tris(hydroxymethyl)aminomethane
TTT	Tripartite tricarboxylate transporter
TPAT	Tetratricopeptide repeat protein associated transport
UC	ultracentrifugation
WT	wildtype
ΔG	Gibbs free energy
ΔH	enthalpy
ΔS	entropy

Declaration of collaborative work

Except where stated otherwise by reference or acknowledgment, the work presented was generated by myself under the supervision of my advisors during my doctoral studies. All contributions from colleagues are explicitly referenced in the thesis. The material listed below was obtained in the context of collaborative research:

Figure 25 **Purification of the QM domains of selected transporter.** The QM domains of selected proteins from *Desulfotomaculum carboxydivorans* (Dc-QM), *Natrialba asiatica* (Na-QM), *Proteus mirabilis* (Pm-QM) and *Marinobacter hydrocarbonoclasticus* (Mh-QM) were expressed in *E. coli* MC1061, solubilized in 1.6 % DM and purified by IMAC followed by SEC as a preliminary test. The SEC were carried out on a Superdex 200 30/100 GL Increase column with a running buffer containing 20 mM Hepes pH 7.5, 150 mM NaCl, 0.2 % DM. The initial test purifications were carried out by Melanie Engelin, Goethe University Frankfurt. The plots were created using OriginPro 2019b (OriginLab Corporation).

Figure 38 **DSF of the TAXI-SBP from *Proteus mirabilis*.** (A) DSF was used to identify potential substrates of the transporter TAXIPm-PQM from *Proteus mirabilis* using 10 µM protein and 40 µM potential substrate. (B) The chemical structures of α-hydroxyglutarate, α-ketoglutarate, glutamate and glutarate are shown. They bound to TAXIPm-P in DSF. (C) The bar chart shows all tested potential substrates with the corresponding thermal stabilization of the protein TAXIPm-P. Several ligands were screened by Melanie Engelin, Goethe University Frankfurt. The plots were created using OriginPro 2019b (OriginLab Corporation). The chemical structures were created using ACD/ChemSketch (Advanced Chemistry Development, Inc. 2021.2.1).

Figure 74 **Intact Mass spectrometry of TAXIPm-P with and without endogenous signal sequence.** Mass spectrometry revealed a mass of 35527.92 Da for the SBP without signal sequence and of 38408.70 Da for the SBP with signal sequence. The SBP without signal sequence is depicted in red and the SBP with signal sequence in green. The measurements were performed by the mass spectrometry facility of the Max Perutz labs, Vienna.

Figure 85 **Immune response of two alpacas during immunization with TAXIPm-QM.** In collaboration with the Nanobody Service facility from the University of Zurich two alpacas - Adonis and Blitz - were immunized (www.nsf.uzh.ch). The immune responses of Adonis and Blitz were evaluated by an ELISA. The immunization was conducted over a period of six weeks with protein a total of four injections (on 27.04.20, 11.05.20, 18.05.20, and 02.06.20). The order of the bars corresponds to the time sequence. The plots were created using OriginPro 2019b (OriginLab Corporation).

Figure 86 **Phage display to select for TAXIPm-QM specific nanobodies.** Two rounds of phage display were performed resulting in four libraries A1, B1, A2 and B2 of which A1 and A2 were derived

from the alpaca Adonis and B1 and B2 from Blitz. The protein EctI was used as negative control to display unspecific binding, kindly provided by Prof. Inga Hänelt, Goethe Universität Frankfurt (Institute of Biochemistry, www.biochem.uni-frankfurt.de). The phage display selections were performed at the MPI CBG Dresden by Dr. Eric Geertsma (Eric Geertsma Group, www.mpi-cbg.de).

Figure 89 **Screening for TAXIPm-QM specific nanobodies by ELISA.** To identify TAXIPm-QM specific nanobodies, the absorbance was measured at 650 nm for the libraries A1 and A2 from Adonis and B1 and B2 from Blitz. Blue bars represent the nanobodies directed against their antigen TAXIPm-QM, grey bars show unspecific binding against UraA or the plate and as comparison green bars represent Sy45 directed against its target antigen UraA. The threshold line was set to define significant results with 1.3 fold higher signal than the background from the plate. The membrane protein UraA and sybody Sy45 were kindly provided by Dr. Benedikt Kuhn, MPI CBG Dresden (Eric Geertsma Group, www.mpi-cbg.de). The ELISA was carried out under my supervision by Christiane Ruse, Goethe University Frankfurt. The plots were created using OriginPro 2019b (OriginLab Corporation).

Figure 90 **Alignment of CDR regions from 39 unique TAXIPm-QM specific nanobodies**. The three CDR regions of 39 TAXIPm-QM specific nanobodies were aligned using Clustal Ω. Amino acids present in all sequences are labelled with double dots and amino acids present in most sequences are labelled with one dot, additionally both are highlighted in yellow (Sievers et al. 2011). The evaluation of the data was supported by Dr. Eric Geertsma, MPI CBG Dresden (Eric Geertsma Group, www.mpi-cbg.de). The diversity of the selected nanobodies was visualized in a phylogenetic tree using NGPhylogeny (Lemoine et al. 2019).

Figure 91 **Overexpression of TAXIPm-PQM and nanobodies.** Whole cell lysates of cell suspensions containing TAXIPm-PQM and nanobodies which were used for radioactive transport assays were solubilized in 1 % (w/v) DDM and submitted to ultracentrifugation to pellet aggregated protein. All protein samples were separated in a 12 % PAGE and visualized by *in gel* GFP fluorescence showing TAXIPm-QM with sfGFP fusion and *anti*-His Western blot showing next to the TAXIPm-QM-GFP protein additionally the nanobodies. TAXIPm-PQM is abbreviated with Pm and the nanobodies with their corresponding numbering on the 96 well plate.

Figure 92 **Evaluation of TAXIPm-QM-sfGFP expression levels in co-expression with nanobodies.** Signal intensities of the TAXIPm-QM-sfGFP expressed protein in a 12 % SDS gel from after the ultracentrifugation were measured. The protein samples from the co-expressions with different nanobodies were compared relative to the expression of the transporter without nanobody addition. TAXIPm-PQM is abbreviated with Pm and the nanobodies with their corresponding numbering on the 96 well plate. The analysis was performed using ImageJ (Schneider, Rasband, and Eliceiri 2012). The data was generated under my supervision by Christiane Ruse, Goethe University Frankfurt. The bar chart was created using OriginPro 2019b (OriginLab Corporation).

Figure 93 **α-ketoglutarate transport inhibition of TAXIPm-PQM by nanobodies.** α-ketoglutarate uptake by TAXIPm-PQM in whole *E. coli* JW2571-1 lacking the endogenous α-ketoglutarate transport of kgtP in presence or absence of previously selected TAXIPm-QM specific nanobodies. Cells were pre-energized by a 2 min pre-incubation with 0.5 % (w/v) glucose. Transport was assayed at 30 °C and pH 7.5 in the presence of 2 µM [^{14}C] labelled α-ketoglutarate. Expression of TAXIPm-PQM in pEXC3sfGH and nanobodies in pSbInit were induced by 1 mM IPTG and 0.01 % arabinose, respectively. The data points that contain error bars are duplicates, all other transport data derives from single time point measurements. TAXIPm-PQM is abbreviated with Pm and the nanobodies with their corresponding position on the 96 well plate. Sy45 is a UraA specific sybody which serves as negative control. The data was partly generated by me and partly under my supervision by Christiane Ruse, Goethe University Frankfurt. The plots were created using OriginPro 2019b (OriginLab Corporation).

Figure 94 **Purification TAXIPm-QM specific nanobodies for structural analysis**. The previously selected TAXIPm-QM specific nanobodies A1C6, B2D6, B2H5 and B2G4 were purified via IMAC and further purified by SEC on a Superdex 75 30/100 GL Increase column in 20 mM Hepes, 150 mM NaCl at pH 7.5. The nanobodies eluted at approximately 13 mL, the corresponding fractions are depicted in grey. The data was generated under my supervision by Christiane Ruse, Goethe University Frankfurt. The plots were created using OriginPro 2019b (OriginLab Corporation).

Figure 95 **Gene neighborhood of TAXIPm-PQM.** IhgO, the gene encoding α-hydroxyglutarate is in the direct gene neighborhood of TAXIPm-PQM. Figure was adapted from the NCBI BLAST database (Madden 2002). The initial genomic DNA analysis of Proteus mirabilis was carried out by Melanie Engelin, Goethe University Frankfurt.

Publication

Authors: Roden, Anja; Engelin, Melanie; Pos, Klaas; Geertsma, Eric

Title: "Membrane-anchored substrate binding proteins are deployed in secondary TAXI transporters"

https://doi.org/10.1515/hsz-2022-0337

Published online March 15, 2023, Biological Chemistry 2023

Acknowledgements

First and most importantly, I would like to thank Dr. Eric Geertsma for the supervision during my PhD study. You trusted that I could manage to do a PhD part-time and gave me every flexibility I needed. I was happy to get valuable knowledge and experimental skills with tips and tricks from you. This is what makes the difference at the end of the day and makes me proud to have been part of your group.

Secondly, I would like to thank Prof. Dr. Klaas Martinus Pos for sharing not only the lab working space and equipment, but also sharing his enormous knowledge and ideas during our group seminar. There is nothing better than someone with deep experience looking at the project from a different and broad perspective. Additional thanks for being my second thesis reviewer.

I also would like to thank Prof. Dr. Inga Hänelt. I think, you are a role model in terms of team leadership. Thank you for sharing your experience and for integrating me in your group.

Additional thanks to all my thesis committees, for evaluating my work and being my examiners.

Furthermore, I would like to thank the KNAUER Wissenschaftliche Geräte GmbH, especially the CEO Carsten Losch, for the great support and giving me the flexibility in my job that permitted this PhD thesis to be realized.

Thank you to my microscopy friends Chris Spahn and Mathilda Gläsmann. It was great to exchange knowledge. I learned much and enjoyed the coffee and ice cream breaks with you.

Furthermore, I would like to thank Prof. Dr. Volker Dötsch for his permission to use the ITC equipment in his laboratories and Christian Osterburg for the instruction in the device. I also would like to thank the Max Perutz Labs Support GmbH for the excellent measurement of our protein samples in the LC-MS and the Nanobody Service facility from the University of Zurich for the immunization of Adonis and Blitz.

I further would like to thank Christiane Ruse. I did not regret it for a second that I took you on board as a bachelor student. Your motivation and purposefulness were outstanding which will take you far.

To all the former and current members of the Geertsma, Pos and Hänelt lab, thank you for creating such a nice working environment. To name the people who were part of this venture - Melanie, Bene, Paula, Selena, Miriam, Jasmin, Jessie, Christian, Andrea, Mariya, Reinke, Michael, Jakob, Janina, Celina, David, Yvonne, Dorith, Olli.

A special thanks goes to Michael. During the last 1.5 years of my PhD thesis, you became my partner and my best friend. Thank you for completing my happiness. I am looking forward to everything that awaits us.

Ultimately, I want to thank my family. Thank you for your constant interest and belief in me and my work, thank you for unconditional unity. Especially I would like to thank my parents for ongoing support throughout my private and scientific career. We can always count on each other.

Curriculum Vitae

Anja Roden

[redacted]

E-Mail: anjaroden@web.de
https://orcid.org/0009- 0003-1725-1375

Personal details

Name:	Anja Roden
[redacted]	
Place of birth:	Wittlich
Nationality:	German

Work experience

01/2022 – now	Head of Chromatographic Purification & Customized Solutions, KNAUER Wissenschaftliche Geräte GmbH, Lead of 2 departments, Product & Project management for R&D and GMP, preparative HPLC, FPLC, Encapsulation of APIs in lipid nanoparticles, Customized Hard-/Software, International competence
01/2020 – 12/2021	Head of Sales DACH, KNAUER Wissenschaftliche Geräte GmbH, Team lead DACH region for HPLC, FPLC, SMB
01/2018 – now	PhD candidate with Dr. E. R. Geertsma and Prof. Dr. K. M. Pos Institute of Biochemistry, Goethe University Frankfurt am Main Molecular Cloning, Protein expression, purification & reconstitution, Radioactive membrane transport assays, Differential scanning fluorimetry, Isothermal titration calorimetry, Nanobody selections
03/2014 – 12/2019	Field Sales Manager, KNAUER Wissenschaftliche Geräte GmbH, Customer visits and consultation for HPLC, FPLC, SMB, tender preparation, installation and instruction in hardware and software, organisation of trade fairs and seminars, talks about HPLC/FPLC

Publication

02/2023	Roden, Anja, Melanie K. Engelin, Klaas M. Pos, Eric R. Geertsma "Membrane-anchored substrate binding proteins are deployed in secondary TAXI transporters" *Biological Chemistry* 2023

Further education

09/2022	Haufe Academy: Technical product management compact
02/2020 + 01/2021	Haufe Academy: From colleague to supervisor I + II
03/2016 - 05/2017	Chamber of Industry and Commerce (IHK) Course certificate: Technical sales manager (IHK)
12/2015	Provadis Kromidas: HPLC advanced course
11/2015	Bichlmeier LifeScience Academy: HPLC course method development
03/2014 - 06/2017	Communication Optimizer Zimmermann: Conversational skills

Third level education

10/2012 - 03/2014	Master of Science: Applied and molecular biotechnology, RWTH Aachen (Grade: 1.6) Master's thesis (09/2013 - 03/2014): Clariant Produkte (Deutschland) GmbH, Munich: Process development: fermentation, electrodialysis, FPLC, spray drying, lyophilisation (Grade: 1.5) Supervised by Prof. Dr.-Ing. Jochen Büchs, Aachener Verfahrenstechnik (AVT) RWTH Aachen
10/2009 - 09/2012	Bachelor of Science: Applied and molecular biotechnology, RWTH Aachen, (Grade: 2.3) Bachelor's thesis: Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Aachen: Cloning and production of antibodies in eukaryotic and plant cells, Tests with *Plasmodium falciperum*, (Grade: 1.0) Supervised by Prof. Dr. Rainer Fischer, Institute for Molecular Biotechnology RWTH Aachen & Senior Executive Director Fraunhofer Institute for Molecular Biology and Applied Ecology (IME)
10/2009 - 08/2013	Internships and research assistant at RWTH Aachen: Biomolecular techniques, fermentation, product purification, data evaluation, general laboratory work, literature research

Second level education

03/2009	Abitur, recipient of the Karl-von-Frisch award for outstanding student performance in Biology, (Grade: 1.3)
08/2001 - 03/2009	Secondary school: Cusanus Gymnasium Wittlich

Stay abroad

10/2012 - 12/2012	Internship at the University of New South Wales, Sydney/Australia: Planning of a biofilm reactor; simulation of biofilm production
04/2011 - 05/2011	Internship at the Malaria Research Centre of the University of Melbourne/Australia: Preparation of mosquitoes, Cloning of antibodies
05/2009 - 09/2009	Au Pair in France