Integrated structural biology investigation of the bacterial TRAP transporter SiaPQM from pathogenic bacteria

Dissertation

zur Erlangung des Doktorgrades (Dr. rer. nat.) der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Janin Glänzer

aus

Bagdad

Bonn, 2022

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn.

1. Gutachter:	Dr. Gregor Hagelüken
2. Gutachter:	Prof. Dr. Ulrich Kubitschek
Tag der Promotion:	22.02.2023
Erscheinungsjahr:	2023

Dedicated to my mother and my father

Abstract

Sialic acid is utilized by human pathogens like *Vibrio cholerae* in various ways to colonize the host. Upon contamination it scavenges the sialic acid from mucus-rich parts of the body such as the gut and causes the disease Cholera - for many countries still a reality of everyday life. Sialic acid acts as a carbon source, cell surface decoration and even as factor enhancing the virulence of this bacteria.
Vibrio cholerae uses the Tripartite ATP-independent periplasmic (TRAP) transporter VcSiaPQM to catch the scarce substrate and to translocate it within the inside of the cell. Also, it is the sole sialic acid transporter in *Vibrio cholerae* and interestingly is absent in humans which further makes it an interesting drug target.
This thesis investigates the substrate binding protein VcSiaP in an integrated structural biology approach using x-ray crystallography, electron paramagnetic resonance and microbiological techniques. VcSiaP was found to be exclusively in open conformation in solution. The closing is triggered by the binding substrate and therefore VcSiaP can function like a switch. Further, the hypothesis that the three binding cleft amino acids R125, E184 and H207 form an ionic triade which is responsible for protein closure was falsified.
For the first time, the binding of an artificial peptide by VcSiaP in crystal structure was observed. Moreover, the binding of this artificial substrate hints towards the idea that closing occurs through bringing both protein lobes closer together.
Other possible short peptide binders were identified with a peptide scan. However, the attempts to determine the binding affinity failed.
Further, EPR experiments were conducted to detect the interaction between the periplasmic domain VcSiaP and the transmembrane part VcSiaQM. This was also not possible.
In a combined action of peptide scan use, uptake assays and bacterial growth assays the attempt was made to find a possible scooping loop of VcSiaP which would lead to opening of the protein and to substrate release.
Overall, the presented work contributes to a deeper understanding of the transporter and develops first routes to possible peptide inhibition in future.

Zusammenfassung

Sialinsäure wird von menschlichen Pathogenen wie *Vibrio cholerae* in unterschiedlichsten Weisen zur Kolonisierung des Wirts benutzt. Nach der Kontamination mit diesem Erreger sammelt dieser Sialinsäure aus schleimhaltigen Teilen des Körpers, des Darms zum Beispiel, und verursacht die Krankheit Cholera – eine Krankheit, die in vielen Staaten der Welt immer noch Teil des Alltags ist. Sialinsäure wirkt als Kohlenstoffquelle, Zelloberflächenverkleidung und sogar als virulenzsteigernder Faktor für das Bakterium.

Vibrio cholerae benutzt den Tripartite ATP-independent periplasmic (TRAP) Transporter VcSiaPQM um das Substrat einzufangen und es in das Innere der Zelle zu bringen. Außerdem ist es der einzige Sialinsäurentransporter in *Vibrio cholerae*, welcher interessanterweise in Menschen abwesend ist. Dies macht den Transporter zu einem potentiellen drug target.

Diese Dissertation untersucht das Substratbindende Protein VcSiaP in einem integrierten strukturbiologischen Ansatz, der Röntgenkristallographie, Elektronenspinresonanz (ESR) und mikrobiologische Methoden vereinigt. VcSiaP wurde bisher in Lösung ausschließlich in offener Konformation gefunden. Das Schließen des Proteins wird durch das Substrat ausgelöst, weshalb VcSiaP als Schalter fungieren kann. Weiterhin konnte die Hypothese falsifiziert werden, dass die drei in der Bindungstasche lokalisierten Aminosäuren R125, E184 und H207 eine für die Proteinschließung verantwortliche ionische Triade bilden.

Die Bindung eines künstlichen Peptids durch VcSiaP wurde zum ersten Mal in einer Kristallstruktur beobachtet. Diese Bindung deutet darauf, dass die Schließung durch Annäherung der beiden Proteinhälften stattfindet.

Andere als mögliche Binder fungiernde kurze Peptide wurden mit einem Peptidscan identifiziert. Es war jedoch nicht möglich, die Bindungsaffinitäten dieser Peptide zu bestimmen.

Weiterhin wurden ESR-Experimente durchgeführt um die Wechselwirkung zwischen der periplasmatischen Domäne VcSiaP und dem Transmembranprotein VcSiaQM zu detektieren. Auch dies war nicht möglich.

Mit einer konzertierten Anwendung von Peptidscans, Aufnahmeassays und Bakte-

rienwachstumsexperimenten wurde versucht, einen möglichen “scooping loop” von VcSiaP zu finden, welcher zu einer Öffnung des Proteins und einer Loslösung des Substrats führen würde.
Insgesamt leistet die vorgestellte Arbeit einen Beitrag zum tieferen Verständnis des Transporters und entwickelt erste Ansätze zur möglichen zukünftigen Inhibtion mittels kurzer Peptide.

Acknowledgements

No dissertation is a one-man-show, therefore I would like to express my gratitude in this section.

First of all I would like to say a special thank you to my supervisor, Gregor Hagelüken. Thank you for this very interesting project, your patience and support even in difficult times.
I also want to thank Olav Schiemann for hosting me into the group and for letting me to use the EPR machines.
Further I want to thank Prof. Kubitscheck, Prof. Lützen and Prof. Schmid-Burgk for being part my examination board.

I am also indebted for the technical support of Rolf Paulig and Knut Hintzen: without you this thesis would have to be written with a pen and the plots be made with a pencil.
I also want to say my warmest thank you to the Hagelüken and Schiemann groups, especially: Nicole Florin and Frasier Duthie for all your technical help in the lab; Martin Peter and Christine Wübben for being great office mates; also Jean Jacques Jassoy, Andreas Berndhäuser, Andreas Meyer, Maria Vicino, Tobias Hett and Sebastian Kühn for being amazing coworkers.

A section also needs to be dedicated to the friends and family who have been by my side all along and who believed in me even when I did not:
Lorenz Mayer, Felipe "Momo" Montealegre Mora and Nikonas Spiliotakis helped me survive the pandemic. Your company made Mordor a home. (And then the flooding made it into Moldor...) Shreyasi Bhattacharya, Robert Rampmaier, Malwin Niehus and Moritz Breyer for support and inspiration.
Last but not at least I want to thank my family: my parents and brothers who have supported me in every way possible.
Lorenz, your unconditional love and support are beyond everything I have ever experienced. Thank you for everything.

Contents

1. Introduction

Until recently, in the mind of modern man, pandemics and plagues belonged to the realm of history. In the developed countries, as science and society progressed, the threat of plagues has been mitigated through technological progress and a higher standard of living. Still, as has been warned by the World Health Organization [1, 2] and illustrated by the recent COVID-19 outbreak, even the most advanced industrial nations are not safe from that looming specter of disease.
Even more unsettling: for developing countries epidemics never went away. Outbreaks accompany war, natural catastrophes and poverty.
Humankind has withstood several severe outbreaks of disease [3]. The bacterium *Vibrio cholerae*has scourged the world seven times with the disease Cholera [4]. The first was recorded in 1817, while the current started in 1961 [5], making it the world's longest running pandemic [6]. Annually, $1.3 - 4.0$ million infections are estimated [7].

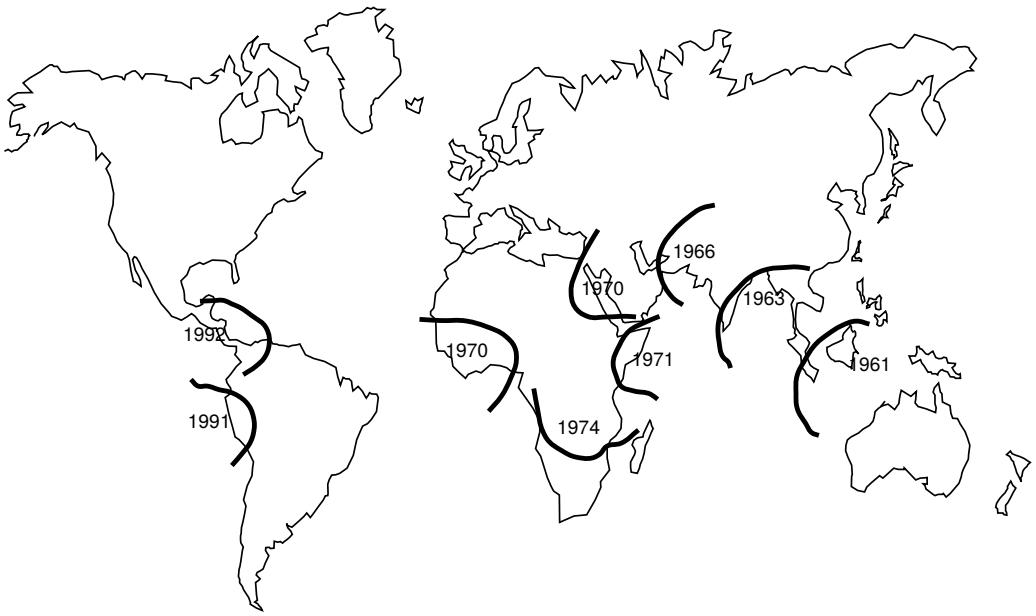


Figure 1.1. – The spread of the current (seventh) Cholera pandemic. The outbreak appeared first in Indonesia 1961 and from there expanded further into the world. Figure adapted from Beeching et al. [8].

Figure 1.1 shows the spread of the disease in the current pandemic across the world. The pathogen surfaced first in 1961 in Indonesia, spreading through Asia and the Middle East to reach the African continent in 1970, causing great havoc. By 1990 nearly all cholera cases reported to the WHO stemmed from Africa (90%) [9]. In 1991 the bacteria crossed the ocean, arriving first in Peru and from there spreading to Central America and Mexico.
In current years there have been numerous further outbreaks, the most current one in Yemen [10].

The disease is caused by ingesting water contaminated by the bacterium *Vibrio cholerae*. It is a gram-negative bacterium which travels through the stomach and finally colonizes the small intestine. For this, it travels through the mucus layer and then attaches to the cell walls of the lumen. Here, it finds a sialic acid rich surface, which helps in developing the virulence of this pathogen [11]. *Vibrio cholerae* exhibits a 57 kb virulence gene cluster called *Vibrio* pathogenicity Island-2 (VPI-2). Alongside with the protein cholera toxin (CT), which causes the explosive watery diarrhea, this region encodes for proteins that are involved in the scavenging, transport and catabolism of sialic acid. This region is called the sialic acid catabolism (SAC) gene cluster. This cluster is only present in pathogenic strains of *Vibrio cholerae* and permits these strains to grow on sialic acid as sole carbon source [12][13].
The uptake of the important virulence factor sialic acid in *Vibrio cholerae* is mediated by a transmembrane transporter, called VcSiaPQM. The role of this transporter in the uptake is the heart of this thesis.
In this introductory chapter, the fundamentals for understanding this thesis are explained. The architecture of the cell envelope is described and an overview of the most important bacterial transporters is given. Focus is also put on the most important sialic acid, Neu5Ac, which is the substrate transported by VcSiaPQM in this thesis.

1.1. The bacterial cell envelope - barriers of the cell

The bacterial cell envelope is a complex structure composed of many layers. It protects the cytoplasm of the bacteria from the harmful environment and thus ensures the cell's survival. But it has also the function of maintaining the structure of the cell. Depending on the architecture of the cell envelope, bacteria are divided into two major groups: gram-positive and gram-negative.
Gram-positive cell envelopes are made of the cytoplasmic membrane and a cell wall of

multi-layered peptidoglycan. Gram-negative cell envelopes on the other hand consist of three "organs": the cytoplasmic membrane, a peptidoglycan layer and the outer membrane. The space with all the cell organelles is called cytoplasm. The space between inner membrane (IM) and outer membrane (OM) in gram-negative bacteria is the periplasm.

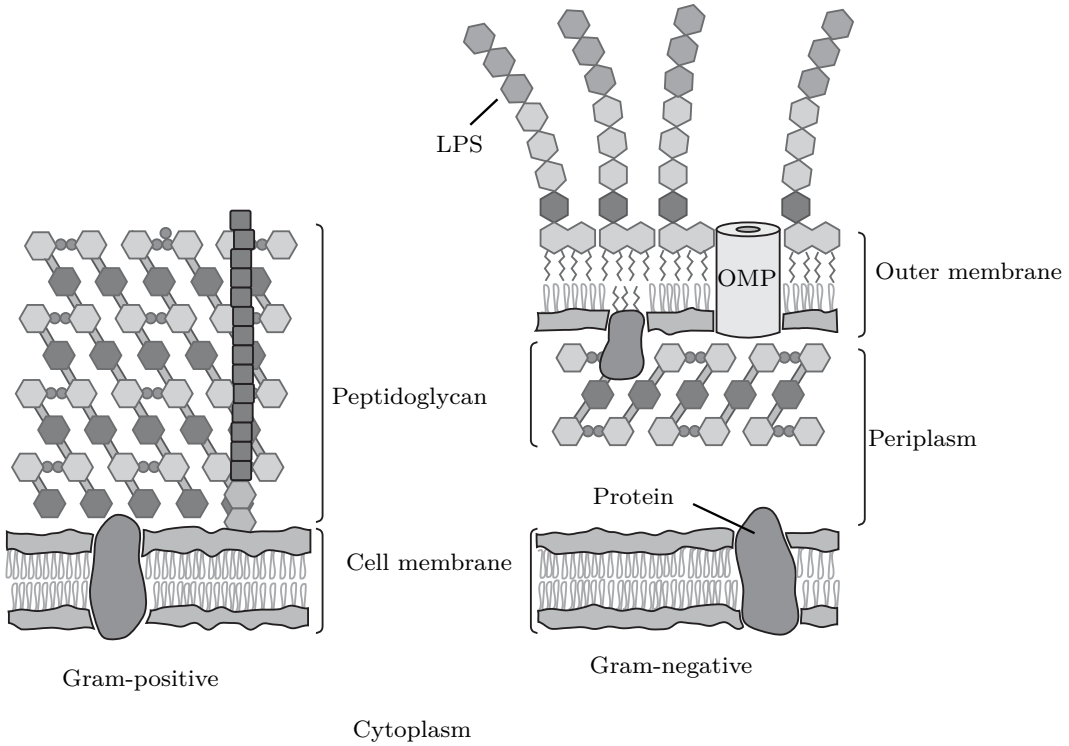


Figure 1.2. – Comparison of gram-positive and gram-negative cell envelopes. The picture is adapted from Silhavy et al. [14].

The outer membrane This organelle of the envelope is a unique feature of gram-negative bacteria that gram-positive bacteria lack. Like other membranes the OM is a lipid bilayer, but asymmetric with regards to its composition: the outer leaflet consists mainly of lipopolysaccharide (LPS), while the inner leaflet consists of phospholipids. LPS helps the barrier function of the OM by limiting the diffusion of hydrophobic molecules [15].
The OM is interspersed with beta-barrel proteins called the outer membrane porins (OMP), which facilitate the transport of small molecules across the OM.

The peptidoglycan cell wall Peptidoglycan is a mesh-like polymer consisting of alternating N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc) segments. The MurNAc units are connected by short peptides of alternating L- and D-amino

acids [16].
Since this structure is so rigid it determines the shape of the bacterial cell and functions like an exoskeleton.

The periplasm This is an aqueous compartment between the IM and the OM which is densely packed with proteins. Periplasmic binding proteins which aid the transport of sugars and amino acids are also located in this compartment.

The inner membrane The IM is a phospholipied bilayer with numerous embedded proteins. Transmembrane proteins carry selcted compounds from the periplasm to the cytoplasm and vice versa. Being also impermeable by electrolytes is a feature of the IM, which leads to the formation of electrochemical gradients. The gradient is then used by some transporting mechanisms as driving force for the translocation of compounds.

The cytoplasm The cytoplasm is all the material that is contained within a cell. It consists mostly of the cytosol, chromosomal DNA and ribosomes.

TRAP transporters are nested in the inner membrane of gram-negative bacteria, connecting the periplasmic and cytoplasmic space. They help transporting substrates into the cell where they can be of further use to the organism.

1.2. Crossing the membrane - crossing the barrier

1.2.1. Different sources of energy

The lipid bilayer is profoundly impermeable to most water soluble molecules. The mechanisms through which the passage becomes possible can be broadly divided in two categories: passive transport and active transport. Both mechanisms differ in the energetic costs required for those actions.
Passive transport requires no energy and the transported substrate is moved along the concentration gradient, leading finally to an equilibrium state of the transported substance on both sides of the membrane.
Active transport, however, requires an energy source such as ATP, light or an electrochemical gradient. By this mechanisms substances can also be transported against a concentration gradient. Here the substance can accumulate on the other side of the membrane and no equilibrium is reached.

Figure 1.3 summarizes all mechanisms that are deployed in bacterial membrane transport. The following paragraphs explain further the branches of the tree diagram in figure 1.3. The displayed classification is according to Cooper et al. [17], Oh et al. [18] and Lodish et al. [19].
Mechanisms of bulk transport via endocytosis, exocytosis and intracellular trafficking are not included in this summary, as they are not within the scope of this thesis.

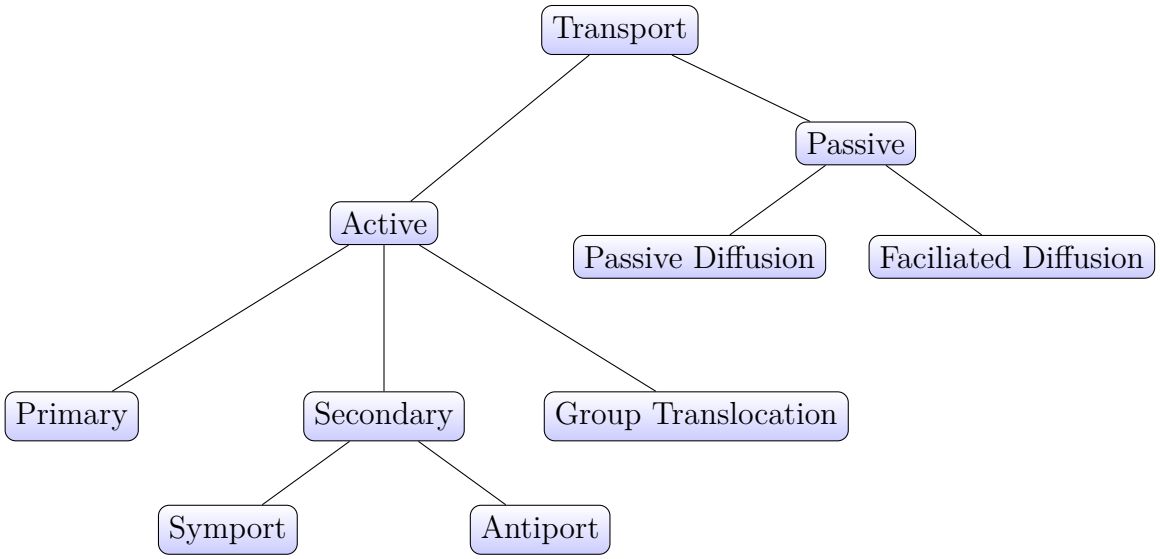


Figure 1.3. – Mechanisms by which small molecules and substrates are transported across the bacterial membrane. The act of transport can be mainly divided into being active or passive.

Passive Diffusion

Passive diffusion is directly executed by the membrane and does not need the help of imbedded proteins. The solute directly dissolves into the lipid bilayer and then crosses it by simple diffusion. Substances which pass this way are for example gases such as CO_2 and O_2, but also small uncharged molecules like ethanol or steroid hormones.

Facilitated Diffusion

Facilitated Diffusion is also known as carrier-mediated diffusion. It requires a membrane embedded protein which helps to transport the substrate. Proteins which apply this mechanism are often referred to as channels, pores or carriers. (The reverse statement, however, is not generally true.)
The transport of glycerol into *E. coli* is an example of facilitated diffusion. It is mediated by the glycerol facilitator GlpF [20].

Primary Active Transport

In primary active transport a source of energy is directly used to relocate the substrate against its concentration gradient. The energy source can be chemical, electrical or solar. Examples are bacterio- and halorhodopsins, which pump H^+ and Cl^- via light [21][22]. Another example are members of the ABC superfamily, which use the hydrolysis of ATP to ADP and Pi [23].

Secondary Active Transport

Secondary active transport is also known as cotransport. It couples the flow of one solute (such as Na^+ or H^+) down its concentration gradient to the pumping of a second solute against its concentration gradient. The installation of the concentration gradient is usually at the expense of a primary active transport.
Cotransport can be further divided into symport and antiport, depending on the direction of translocated substrate in relation to the flow of the counter ion.
Examples of for this transporting mechanism are members of the MFS superfamily [24] or the TRAP transporters, dicussed in this thesis.

Group Translocation

This is a mechanism by which the solute is modified in the process of transport. For example, in the phosphoryl transfer-driven group translocators, extracellular sugar is transported and comodified into an intracellular sugar-phosphate. The transport of N-acetylmuramic acid in *E. coli* is an example of group translocation [25].

1.2.2. Overview on transporters - The Transporter classification Database

Approximately 30% of the proteome of every cell consists of membrane proteins. And roughly 10% of membrane proteins are transporters [26].
Various approaches have been made to classify this category of proteins, using mainly sequence-based, structure-based oder functional-based grouping [27]. The results of these assembly drawings are summarized in various accessible databases on the internet such as TransportDB [28] and UCSF-FDATransPortal [29]. The profound consequence is the existence of many classification systems with conflicting information. Proteins which have been grouped based on a sequence-similarity scheme do not necessarily have to share identical functions. Some databases only list a specific group of transporters

such as Solute carriers [30] or ABC-transporters [31]. Others only provide information of transporters from few or only one organism (e.g. Human Transporter Database [32]). In this thesis the classification of membrane transporters follows the grouping scheme of Saier et al [33]. He and his group have created the Transporter Classification Database (TCDB, http://www.tcdb.org) with the intention of covering all membrane transport systems without confining to certain organisms [34]. With new knowledge about transporters, the database is continously growing [35].
The TCDB is the only transport protein classification database adopted by the International Union of Biochemistry and Molecular Biology (IUBMB). The classification is also also included in several other databases such as UniProt and the PDB.
Transporters in the TC system are classified according to function and phylogeny. The classification scheme contains five levels. The first number (V=1-5) indicates the class of the protein. This is followed by a letter (w) that indicates the protein subclass (e.g, the energy source). In summary the first two levels describe the function of the transporter. The next two numbers indicate the transport protein family and subfamily (X,Y). These two classify according to phylogenic similarity. The last number (Z) defines the substrate.

$$\underbrace{\underbrace{\underbrace{\underbrace{\underbrace{V.}_{\text{Class}}}_{\text{Subclass}}\ w.}_{\text{Family}}\ X.}_{\text{Subfamily}}\ Y.}_{\text{Transporter}}\ Z$$

Usually these transporter classes include completely different, non-overlapping families. But few transporters can employ more than one mode of action or can use an energy-coupling mechanism different from that of other members of the family. Thus, a limited amount of overlapping data exists. Also, homologous transport proteins are given the same TC-number if they share the same substrate.

The TC classification scheme acknowledges of five distinct classes of transporters: Channels and Pores, Electrochemical Potential-driven Transporters, Primary Active Transporters, Group Translocators and Transmembrane Electron carriers.

Class 1: Channels and Pores

The structures of this category consist mainly of α-helical channels or β-stranded porins, which spann across the membrane. But also channel-forming toxins, nonribosomally synthesized channels and holins belong in this category. The transporting mechanism is faciliated diffusion. Examples of this category are voltage-gated sodium channels [36] and outer membrane porins [37] (e.g. OmpF, PhoE [38]).

Class 2: Electrochemical Potential-driven Transporters

The defining feature of this class is that the transporters use a carrier-mediated process to translocate the solute. Some members use facilitated diffusion, not coupled to the use of a primary energy source. Most of them use preexisting electrochemical gradients as source of energy, e.g., of H^+ and Na^+. These systems are usually stereospecific. Examples of this transporter class are the Major Facilitator Superfamily (MFS) [39] and outer membrane transporters TonB [40].

Class 3: Primary Active Transporters

The members of this category use a primary source of energy to transport a solute against its concentration gradient. This is possibly the most heterogeneous class. Elements of this group are ABC transporters (ATP hydrolysis driven) [41], NADH Dehydrogenase Family (Oxidoreduction-driven transporter) [42] and the HelioRhodopsin family (light absorption-driven transporter) [43].

Class 4: Group Translocators

The proteins assigned to this class are complex transporter systems which couple the translocation of the substrate to an exergonic reaction. Usually this process involves the derivatization of the substrate, such as sugar phosphorylation using phosphoenolpyruvate (PEP) mediated by the phosphotransferase system (PTS) [44].

Class 5: Transmembrane Electron carriers

The systems of this category catalyze the flow of electrons across the membrane. It follows a donor-acceptor principle: the flow is directed from donors localized to one side of the membrane to acceptors localized on the other.

The best studied classes are by far 1-3. Therefore examples and comparisons will be mainly drawn from those classes.

1.3. Sialic acid transport

Under physiological conditions, sialic acid is bound to glyconjugates in mucus-rich organs of the human body. The gastrointestinal tract and the respiratory system are examples of sialic acid rich human environments. To make the bound sialic accessible, pathogens have to hydrolase the bond between the sialic acid and the glycoconjugate rest via a neuraminidase. Once released, import into the cell becomes possible.
This section outlines the transport from the external environment across both membranes of a gram-negative pathogen. Special focus is put on important transporter families.

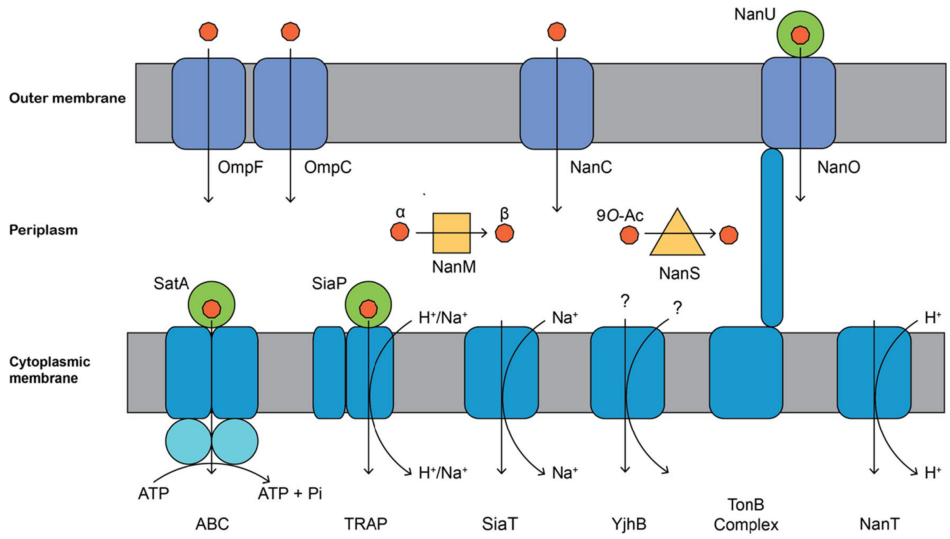


Figure 1.4. – Overview on sialic acid transport across the outer and inner membrane of gram-negative bacteria. Image taken from North et al. [45].

1.3.1. Crossing the outer membrane

The easiest ways to import sialic acid is by simple diffusion through porins with low specificity. The outer membrane channel proteins OmpF and OmpC from *E. coli* [46][47] are examples of this process. They are involved in the translocation of small, hydrophilic molecules up to a size of 600 Da.
Uptake can also occur with more specific channels: NanC from *E. coli* [48] is expected to be involved in the movement of negatively charged sialic acids.
The active transporter system NanOU [49][50] is higher in complexity as the above mentioned porines and channels. The NanOU system from *Bacteroides fragilis* is a TonB-dependent transport complex. NanU interacts with the outer membrane and consequently increases the uptake of sialic acid by NanO. This is especially helpful at low sialic acid concentrations.

After arriving in the periplasm, sialic acid is further processed: the sialate mutarotase NanM converts α-anomers of sialic acid into the β-form. It is assumed that the β-sialic acid anomer is the form in which cytoplasmic membrane transporters translocate sialic acid [51]. The mutarotase is expressed for example in *E. coli* [51] and in *V. cholerae* [52].
Further *E. coli* genes encodes for a protein (NanS), which functions as a sialate 9-*O*-acetylesterase [53]. NanS removes the acetyl-residues which were left from the scavenge.

1.3.2. Crossing the inner membrane

Four unique transporter families have been identified across pathogens, which translocate sialic acid from the periplasm into the cytoplasm of the cell. The transporters are members of the ABC family, MFS, SSS and TRAP.
Since all of these families are very important groups among transporters, they will be discussed in this chapter on the basis of archetypical examples which mediate the transport of sialic acid.

ABC transporters

ABC transporters hydrolyze ATP into ADP and P_i. The released energy from this chemical reaction is then coupled to the transport. ABC transporters are widespread in all domains of life and form one of the biggest protein families [54]. They are grouped into importers and exporters and the group of importers can be further subdivided into class I importer and class II importer.
They are characterized by two transmembrane domains (TMDs), which form the translocation pathway, and two nucleotide binding domains (NBDs), which are also called ATP binding cassette (ABC). Additionally to these four domains, ABC transporters possess a substrate binding protein (SBP), which captures the substrate and hands it over to the transporter. In gram-positive bacteria the SBP is anchored to the membrane whereas in gram-negative bacteria the SBP is freely soluble in the periplasm.
The SatABC system from *Streptococcus pneumoniae* is a good example of a sialic acid importer that belongs to the family of ABC transporters [55]. SatA is the periplasmic substrate binding domain and the integral membrane permease domains are SatB and SatC.

TRAP transporters

Sialic acids are also imported by a transporter family called TRAP transporters. Members of this family are the systems VcSiaPQM from *Vibrio cholerae* and HiSiaPQM from *Haemophilus influenzae*. Those are discussed in detail in section 1.5, since they are the main focus of this thesis.

Major facilitator superfamily

The major facilitator superfamily (MFS) is the largest family of secondary transporters [56]. The architecture of most members of this protein family consists of 12 transmembrane helices. Some have 14 transmembrane helices. Both termini are placed on the same (cytoplasmic) side of the membrane. All members of this family share a common fold called the MFS-fold. It displays a two-fold pseudosymmetry by which the fold is divided into two domains. Both domains exhibit six helices. Within each domain, the helices can be further subdivided into two bundles of three helices. Those two bundles are related to each other by an approximate 180° rotation.
Sequence analysis indicates that this pseudosymmetry is due to a gene duplication event.
The energy for the transport is gained with the help of a proton gradient.
A sialic acid symporters, which belong to the MFS is for example NanT from *E. coli* [57].

Sodium solute symporters

The family of Sodium solute symporters (SSS) belongs to the group of secondary transporters [58]. Sodium is co-transported (in symport) along with the substrate. The only two structures which could be solved for this transporter family until today is the sodium galactose transporter (SGLT) from the organism *V. parahaemolyticus* [59] and the sialic acid transporter SiaT from *Proteus mirabilis* [60]. Depending on the protein, the transporter is comprised of 13-15 helices. The core consists of two inverted repeats of 5 helices, also known as LeuT-like-fold.

1.4. The substrate N-acetylneuraminic acid Neu5Ac

In the previous chapter the different transporters for sialic acid transport were presented. This chapter focuses on the substrate of the VcSiaPQM and HiSiaPQM transporters. A short introduction of sialic acids in general is followed by an outline of how sialic acid is acquired and used by bacterial pathogens.

1.4.1. Introduction to sialic acids

Sialic acid (Sia) is a generic term for a family of α-keto sugars with a nine-carbon backbone [61]. It comprises over 50 naturally occuring molecules [62], the most common ones being N-acetylneuraminic acid (Neu5Ac), N-glycolylneuraminic acid (Neu5Gc) and N-acetyl-9-O-acetylneuraminic acid (Neu5,9Ac2). Figure 1.5 depicts the chemical structures of these three compounds.

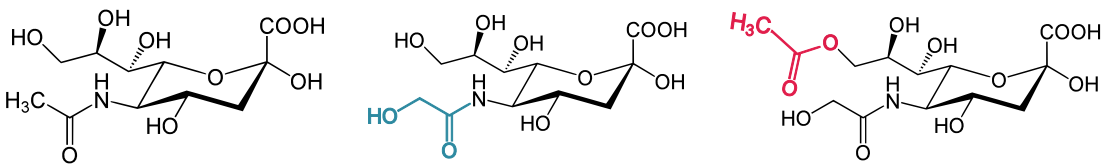


Figure 1.5. – The three most common sialic acids in nature. Neu5Ac also acts as a precursor in the biosynthesis of the other sialic acids. Figure adapted from Schauer [63].

With the exception of plants, they are found everywhere in nature. Since they appear in so many different species it is difficult to describe a general role of sialic acids [63][64][65]. In eucaryotes and procaryotes for example, Sias mainly occur as terminal components of cell surface glycoproteins and glycolipids. Here, they take part in in cell-to-cell communication and self-recognition [63].

In humans, sialic acid is most dominantly found in the brain. The central nervous system deploys sialic acid in mechanisms of neural cell adhesion [66]. Hence, the name "neuraminic acid" for some sialic acids. Further, they are present in mucin rich surfaces of the human body such as the intestinal system (especially the gut), the lungs and the vaginal tract.
These Sia rich surfaces are are vastly deployed by pathogens such as viruses [67] (e.g. MERS-CoV [68] and SV40 [69]), protozoa [70] (e.g. *Toxoplasma gondii* [71] and *Trypanosoma cruzi* [72]) and bacteria [73] (e.g *Vibrio cholerae* [74] and *Haemophilus*

influenzae [75]).

From the following section on I will use the terms Neu5Ac and sialic acid synonymously. Other Sias are not relevant for this thesis.

1.4.2. Sialic acid utilization by bacterial pathogens

Bacterial pathogens like *Vibrio cholerae* and *Haemophilus influenzae* use Neu5Ac to ensure their survival within the host. They use Neu5Ac either as source of food or molecular mimicry to mask themselves from the host [76][73][77]. Figure 1.6 gives an overview of the most important pathways used by prokaryotes.

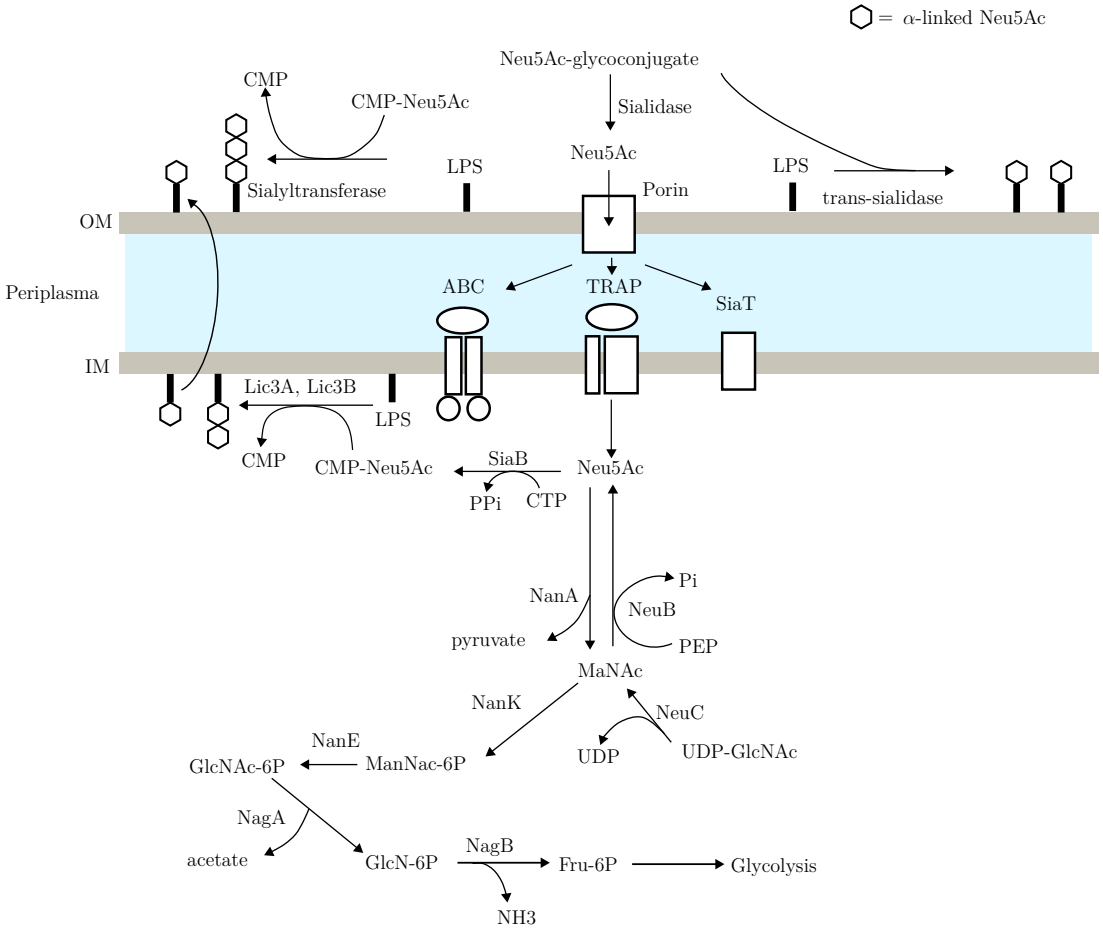


Figure 1.6. – Overview on important Neu5Ac catabolic pathways. Image modified from Severi et al. [73].

Sialic acid acquisition

Bacteria have several ways to pick up sialic acid from the "outside". They can either acquire Neu5Ac which was released from the host by other bacteria. Or, they use a sialidase (such as NanH, in the case of *Vibrio cholerae*) to remove terminal Neu5Ac from the glucoconjugate surface of the host.
Furthermore, a small fraction of free Neu5Ac exists from the host's own sialic acid recycling metabolism [78][79]. The sialic acid is then transported into the pathogen using the sialic acid transporters described in chapter 1.3.
If bacteria lack the genes for sialic acid uptake, they deploy another way of Neu5Ac acquisition. They produce this building block themselves via *de novo* biosynthesis. The metabolite UDP-N-acetylglucosamine UDP-GlcNAc is the initial precursor in this pathway. It is formerly produced in the hexosamine pathway. By the combined actions of UDP-GlcNAc 2-epimerase (NeuC) and Neu5Ac synthase (NeuB), Neu5Ac can be directly produced [80].

Sialic acid catabolism

After transport, Neu5Ac arrives in the cytoplasmic space, where it can be further catabolized. Through the cleavage of a pyruvate by N-acteylneuraminic acid lyase NanA, the sugar N-Acetylmannosamine ManNAc is generated. By enzymatic action of N-acytmannosamine kinase NanK and N-acetylmannosamine 6-phoshate epimerase NanE, ManNAc is transformed into N-acetylglucosamine-6-phosphate, GlcNAc-6-P. GlcNAc-6-P is further fragmented by N-acetylglucosamine-6-phosphate deacetylase NagA and glucosamine-6-phosphate deaminase NagB. This results in the sugar fructose-6-phosphate Fru-6-P. Fru-6-P can now be further recycled in the glycolitic pathway in the gain for energy.
It is still not clear how pathogens manage the balance between depletion and synthesis of Neu5Ac.

Incorporation of Neu5Ac on the cell surface

The acquired sialic acid is also used by several bacterial pathogens to decorate their cell surface, masking them from the immune system of the colonized host.
There are examples where sialylation of the LPS takes place outside of the cell and does not need to involve transport at all. This is performed either by a *trans*-sialidase (e.g. *Corynebacterium diphtheriae* [81]) or a sialyltranferase (e.g. *Neisseria gonorrhoeae* [82]).
Another way is to activate Neu5Ac in the cytoplasm by a by a CMP-Neu5Ac synthetase

(SiaB in *H. influenzae* [75], the synthethase of *V. cholerae* is still unnamed [83]). The activated CMP-Neu5Ac is then incorporated into the cytoplasmic LPS by sialyltransferases (Lic3A and Lic3B in the case of *H. influenzae* [84]). As a last step, the sialylated LPS is translocated to the surface of the cell by the regular LPS biosynthesis apparatus [85][86].

1.5. Tripartite ATP-independent periplasmic (TRAP) transporters

TRAP tranporters are a family of SBP-dependent secondary transporters (TC: 2.A.56), which are ubiquitous in bacteria and archaea but absent in eukaryotes. They belong to the large superfamily of ion transporters (IT superfamily) [87][88]. They transport a variety of different substrates like C4-dicarboxylates, α-Keto acids, taurine, lactat and sialic acid [89].
The first ever TRAP transporter was discovered in 1997 by Forward et al. [90]: it was the C_4-dicarboxylate transporter DctPQM from *R. capsulatus*.
DctPQM is a three-protein transporter. DctPQM possesses two transmembrane domains of unequal size, DctQ and DctM, and surprisingly also a SBP in the periplasm which was then called DctP. Up to that moment in time, ABC transporters were the only prokaryotic transporters known to utilise a SBP [91]. Hence, the name tripartite ATP-independent periplasmic transporters. (Since then, another family of secondary transporters has been found to deploy SBPs, the tripartite tricarboxylate transporters (TTTs). But this shall merely remain a sidenote.)
This TRAP transporter become the prototype of TRAP transporters. All TRAP transporters that display high sequence similarity with DctPQM are referred to as classical TRAP transporters, DctPQM-type transporters or TRAP-T transporters by the literature [92]. Figure 1.7 shows the current classification of TRAP transporters.

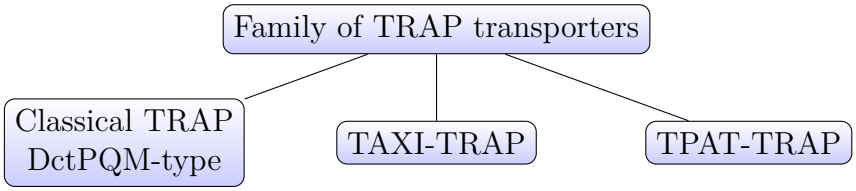


Figure 1.7. – Classification of the TRAP transporter family. They are divided into three subfamilies.

Another group of TRAP transporters exists which presents the SBP with very low sequence similarity to DctP. This SBP even forms a distinct group of its own called TRAP associated extracytoplasmic immunogenic (TAXI) proteins. The corresponding subgroup of transporters is called TAXI-TRAP.
The last existing group is the TPAT-TRAP group and was characterized by Deka et al. [93]. During their studies about the TRAP transporter in the organism *Treponema pallidum*, they found another essential protein belonging to the system, additionally to the P-, Q- and M-domains. This additional protein (T-domain) is soluble and

consists of 13 α-helices and one smaller helix. A few of these helices were homologous to a tetratricopeptide motif (TPR). This led to the name TPR-protein associated transporters (TPAT).
Figure 1.8 shows the organization of genes in the different TRAP transporter families. Genes for the Q (dark blue) and M domains (ligh blue) are part of all family members. The genes that encode for the SBPs directly flank these genes. Genes that code for homologs of DctP are indicated in red. Genes of the TAXI-SBP are shown in yellow and the gene which codes for the additional protein in TPAT system is colored in pink.

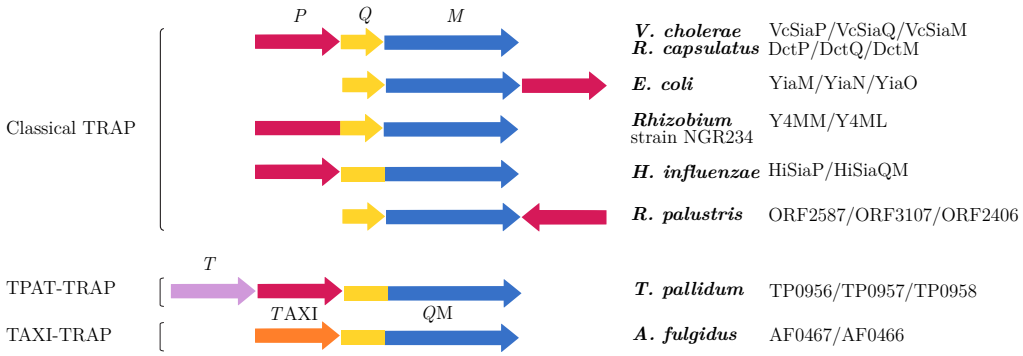


Figure 1.8. – Genetic organization of TRAP transporters. Picture modified from Kelly et al. [94]

Further, the image gives examples of organisms that are associated with these transporters. *Vibrio cholerae* and *Haemophilus influenzae* deploy a TRAP transporter that belongs to the classical group. *Treponema pallidum* uses a TPAT TRAP transporter and *Archaeoglobus fulgidus* has a TAXI TRAP transporter.

Structurally, TRAP transporters can be placed between ABC transporters and secondary transporters (see figure 1.9). The main transportation tunnel is formed by the larger M-domain which comprises 12 helices. The smaller of both transmembrane domains, the Q-domain, consists of 4 helices.
In some classical TRAPs Q- and M-domain are fused together by an additional helix as in the case of *H. influenzae*. Most of the TAXI transporters have fused QM as well. *Rhizobium* is, up to this day, the only organism in which P-domain and M-domain are fused together [94].

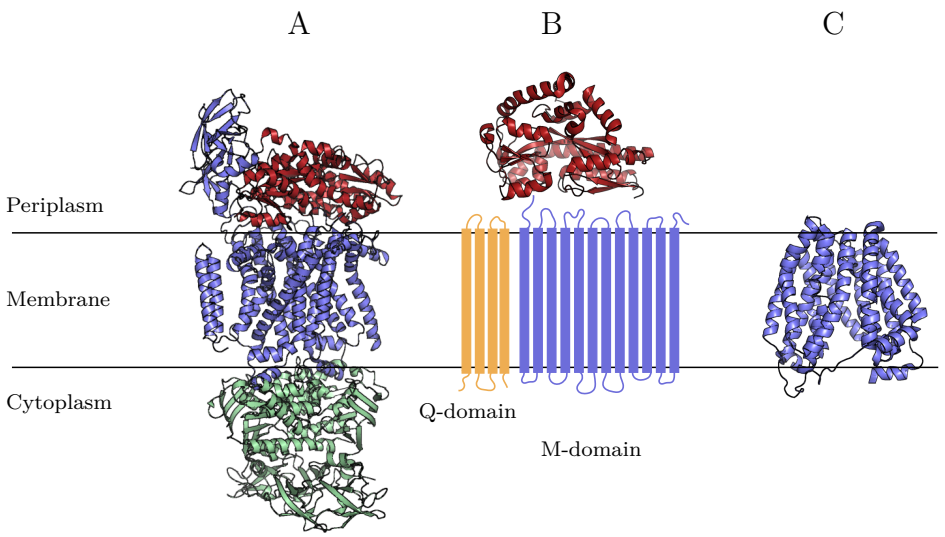


Figure 1.9. – Architectures of archetypal ABC (A), secondary (B) and TRAP transporters (C). ABC transporters (here represented by the maltose transporter $MalGFK_2$ [95]) consist of an SBP (red), two transmembrane domains (blue) and two nucleotide binding domains (green). Secondary transporters (represented by lactose permease [96]) are only composed of the transmembrane part. TRAP transporters are equipped with a transmembrane part (yellow Q-domain and blue M-domain) and an additional SBP.

The transporters VcSiaPQM and HiSiaPQM are attributed to classical TRAP transporters and therefore the other subgroups are not dealt with in this section.

1.5.1. Structures of the P-domain

High resolution X-ray structures of the *Vibrio cholerae* (VcSiaP), *Haemophilus influenzae* (HiSiaP) exist. Both SBPs have a sequence similarity of 50%. Further, there are homologous structures of *Fusobacterium nucleatum* (FnSiaP) and *Pasteurella multocida* (PmSiaP) available. With the exception of VcSiaP, all other P-domains have been investigated in the open as well as in the substrate-bound form. All four proteins share the same overall structure:
The P-domain consists of a two α/β-domains which are connected by the backbone of a long helix region. The binding cleft is located between these two α/β-domains. Binding of Neu5Ac bends the backbone region and leads to conformational changes with shifts of Cα positions up to 30 Å [97]. The closing mechanism resembles a venus flytrap. All SBPs of TRAP transporters belong to one cluster (cluster E) within the SBP classification system of Berntsson [91, 98].
Figure 1.10 A depicts the open P-domain (from *H. influenzae*). The hinge helix region

which bends upon substrate binding is depicted in orange and the substrate is shown in the middle of the protein in a purple color.

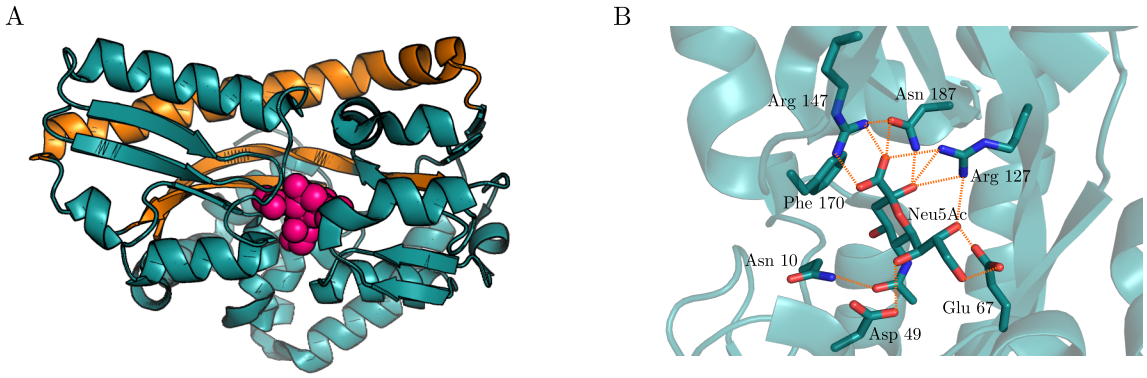


Figure 1.10. – P-domain of classical TRAP transporters. A shows the structure of substrate-bound HiSiaP. The helix hinge region is marked orange, the remaining domains of the protein are shown in teal and the substrate (Neu5Ac) in the center is depicted in purple. B depicts the interaction pattern of the residues of closed HiSiaP and the substrate Neu5Ac.

Setty et al. investigated the binding site of HiSiaP, FnSiaP and PmSiaP and found them to be highly conserved [99]. The hydrogen bonds and hydrophobic interactions, however, vary between the four SBPs. Out of these, the following have been identified to be the most important:
R145 interacts with the carboxylate group of Neu5Ac at the C1-position. This interaction is not only conserved within the comparison of the mentioned three TRAP SBPs, but conserved within all dicarboxylate-binding proteins [100].
R127 also forms a salt bridge with two oxygen-groups of the substrate: with O of the C1 as well as with O2.
Other interactions between closed HiSiaP and Neu5Ac are depicted in 1.10 B.

1.5.2. Structure and function of the QM domains

In VcSiaQM both subunits are expressed individually (4+12 helices). They form a tight 1:1 complex during protein purification, and separation of the two ruptures the process of transportation [101].
In HiSiaQM both subunits are fused together by an additional helix, resulting in a 17 helix transmembrane protein.
The M-domain is believed to form the translocation channel, since it is strongly conserved and member of the ion transporter superfamily [102].
The Q-domain, on the other hand, is known to be essential for transporter function,

but is weakly conserved. Therefore, its role is yet unclear. The following possible functions have been discussed [103][89]: mediator of the interaction between P-domain and M-domain, assembly factor to stabilize the M-domain and participant in the energy coupling of the transporter.
Up to this day no crystal structures of the transmembrane part of the TRAP transporter exists[1]. However, predictions of the membrane topology are available. Those predictions are based on a multiple sequencing alignment of seven TRAP transporters which translocate Neu5Ac [101]. The alignment includes four fused familiy members and three with seperate domains (see appendix A.2). The membrane components were identified by TMHMM [105] and the boundaries of the helices were identified by comparison with the multiple sequencing alignment [89]. Figure 1.11 shows the topology map for VcSiaQM.

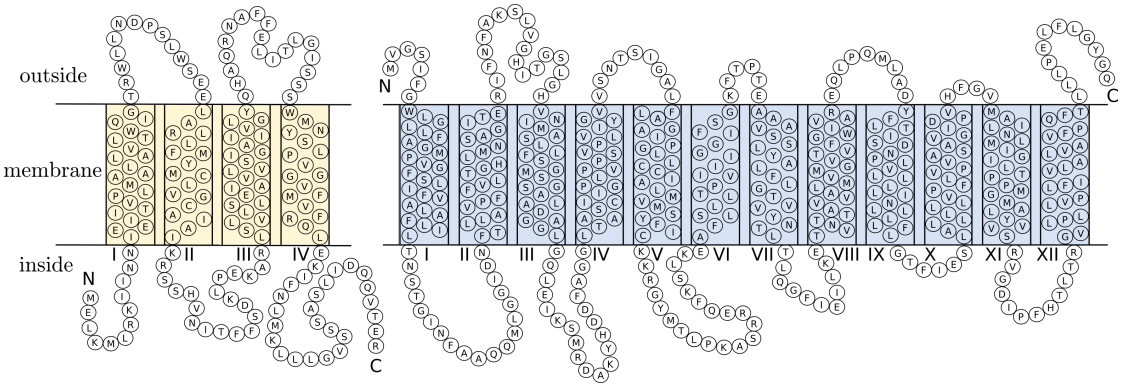


Figure 1.11. – Topology map of VcSiaQM. The transmembrane parts of the Q-domain are marked yellow and the transmembrane parts of the M-domain are indicated in blue.

Further, the structure of the TRAP transporter of YiaQM was calculated by Ovchinnikov et al. [106] as well as the docking of the P-domain. The structure differs from the topological predictions by having an inverted/twisted helix (at helix 11).

Studies by Mulligan et al. on HiSiaPQM have identified the transporter to be biologically unidirectional. The co-transport of at least two Na^+-ions are the driving force of this process [107].

1.5.3. Mechanism of transport

Based on the knowledge of ABC transporters and classical secondary transporters, a hypothetical mechanism of transport was proposed by Mulligan et al. [107] . Many

[1] A preprint of the first publication of this structure by Peter et al. does exist [104]

stages of the mechanism still do not have experimental support and are therefore an object of current research. The transporting cycle is summarized in figure 1.12.

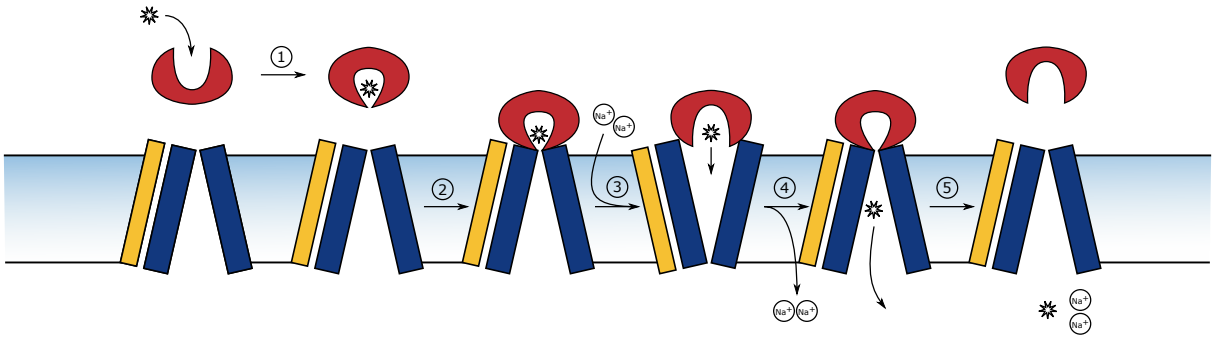


Figure 1.12. – Proposed mechanism of Neu5Ac transport for VcSiaPQM and HiSiaPQM. Figure adapted from Mulligan et al. [107].

As to be expected, the cycle of transport starts with the SBP binding to the substrate and undergoing conformational change (1). Now it resides in the closed and liganded state. Until now the resting state of the unliganded SBP ist unkown. This uncertainty leaves space for two possible speculations about the state of the SBP in solution:

1. The SBP rests in open conformation and the closing is strictly substrate induced.
2. The SBP is in equilibrium between the open and closed conformation before it scavenges the substrate.

Further, it is possible to have additional stable intermediates in solution, which have not been discovered yet by crystallography or another structural biology method.
After seizing the substrate, the closed and substrate bound SBP interacts with the transmembrane part SiaQM (2), which is thought to present itself in the periplasm-closed/cytoplasm-open conformation. This assumption is based on the comparison with two other secondary transporters LacY and GlpT, which are both resting in the outside-closed conformation [96][108]. Interaction (2) is very specific to the transporter: upon switching the SBPs of analogous transporters from different organisms, the transporter fails to function [107].
The transport is driven by binding of at least two Na^{2+} ions (3) which are co-transported into the cytoplasm (4). Either the binding of Na^{2+} or the the interaction with the SBP would induce the opening of SiaQM in an alternating access mechanism.

The transport of sialic acid via SiaPQM has been proven to be unidirectional. It is only reversible in the case of high Neu5Ac concentrations and unliganded SBP, which are highly unphysiological conditions.

1.6. Integrated structural biology – X-ray crystallography meets EPR

To understand a proteins function, the knowledge of its tertiary and quarternary structure is necessary. Only with a deeper understanding of the molecular mechanisms perfomend by proteins, manipulation of these become possible. And with manipulations, new technologies and drugs to help mankind are created [109][110].
For example, GCPR targeting drugs were identified through screening before structural information on the receptor was available. With structural information being accessible now, these drugs are improved [111][112].
The collecting of fitting complementary information in order to elucidate the structure-function relationship of biological systems is the main task of integrated structural biology. The figure 1.13 depicts the structure solving methods which were applied to structures published in the PBD (reference date is June 2019).

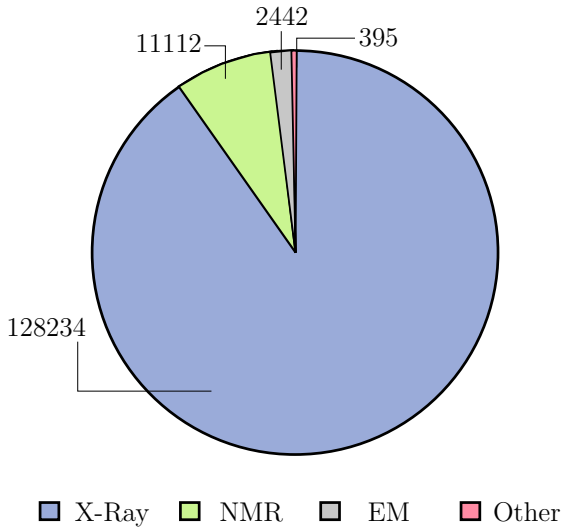


Figure 1.13. – Distribution of the different methods in the field of structural biology.

As can be seen from this statistic, the main methods for structure solving are X-ray, followed by NMR and then Cryo-EM. It is worthy to mention that with new developments in the field of cryo-EM, its structure-solving capacity has sped up incredibly and is becoming more and more a method of choice [113][114].
Already the combination of these main methods can be used to get a broader the picture of the studied protein systems [115]. Hybrid approaches which combined Xray with cryo-EM were famously able to solve the architecture of the 26S proteasome [116]. And combining NMR with Cryo-EM led to solving the bacterial type III secretion system [117].

These traditional techniques have been complemented by a growing number of experimental methods, which provide additional information to the protein system. These methods include small-angle X-ray scattering (SAXS), cross-linking, mass-spectrometry and EPR.

In this thesis I used mainly X-ray crystallography along side with EPR to investigate the TRAP transporter from *Vibrio cholerae*. While crystallography is well known within the structural biology community, EPR is still rather a niche technique. I shall provide an overview, which will be helpful in understanding my results. More thorough introductions to magnetic resonance in general and electron paramagnetic resonance in particular can be found in textbooks [118][119][120][121].

1.6.1. Site directed spin labeling with nitroxides

EPR requires the presence of a paramagnetic spin center to be applicable. Some proteins have naturally occuring spin centers in form of metal ions, or their cofactors contain a radical which can be used. Examples of these are Azurin (Cu^{2+}) and Cytochrome P450cam (Fe^{2+}, Fe^{3+}). Yet the vast majority of proteins is diamagnetic. For the investigation of these proteins, paramagnetic centers have to be introduced into the system. In common practice such paramagentic centers are bioconjugated to the protein postsynthetically, that is, after the protein has been expressed and purified.

Nitroxide labels

The most common spin centers used for this purpose are nitroxides because of their stable radical and their accessibility towards chemical manipulation. They were introduced by the group of McConnell in 1965. Nitroxide labels are composed of a heterocyclic ring (e.g pyrrolidine) which carries the nitroxyl radical $(N\text{-}O)^{\bullet}$ and a linker to connect the radical to the desired side-chain of the protein to be investigated.
The group of Berliner invented the cysteine specific label MTSSL (see figure 1.14) in 1982. Figure 1.14 shows MTSSL (X=H) and variations of this label. They are all small in size, comparable to a tryptophan side chain.

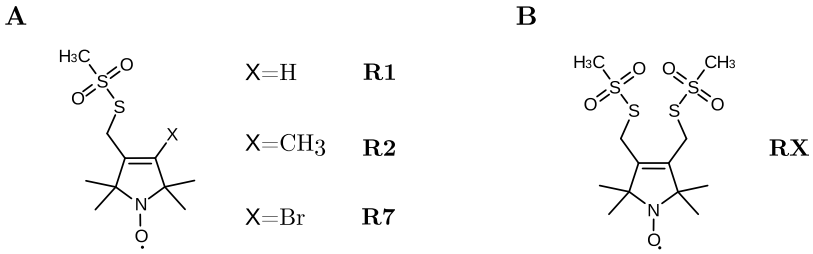


Figure 1.14. – A) The cysteine-specific label MTSSL (X=H) and variations of it. B) The label RX can even bind to two cysteine sites in close proximity.

The introduced methane sulfonyl group is a very reactive leaving group and forms stable disulfide bonds with the thiole group of cysteine. Further, the reaction can be reversed by incubating the labeled protein with mild reducing agents such as tris(2-carboxyethyl)phosphine (TCEP) or dithiothreitol (DDT).

Site-directed spin labeling of cysteines

In the 1980s, the group of Hubbell combined introducing point mutations into proteins (site-directed mutagenesis) with the bioconjugation of the cysteine specific MTSSL, thus inventing a general methodology for site-directed spin labeling (SDSL)[122]. The targeted amino acids are changed into cysteines and subsequently labeled with MTSSL. Here, one has to be careful not to disrupt the structure or the function of the protein. The side-chain which is now part of the protein and bears the nitroxide is called **R1** (see figure 1.14). Figure 1.15 summarizes this process.

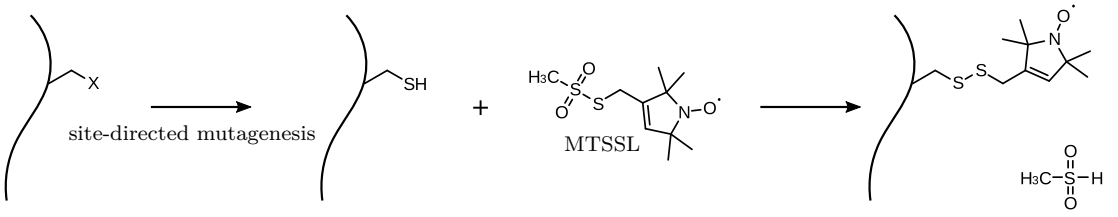


Figure 1.15. – Labeling of a protein backbone with MTSSL. First, the targeted amino acid is mutated into a cysteine. This can now react with MTSSL, forming an S-S bond under the release of the methane sulfonyl group.

MTSSL is disadvatageous when working under slightly reducing conditions. Then, the maleimide spin label (MSL) [123] or the iodoacetamide spin label (IASL) [124] can be used.

The presented method for introducing nitroxides into proteins has one major disadvantage: It fails when proteins contain cysteines which are essential to their function,

or when the protein simply contains a large number of cysteines. In this case, the naturally occuring cysteines cannot be removed. One way of overcoming this challenge is the introduction of unnatural amino acids and then labeling the protein again with a nitroxide [125]. The method of SDSL can also be applied to other biomolecules such as DNA and RNA [126].

Finding optimal positions for labeling

The answer to the question which spin labeling positions within the protein are optimal depends on the biological question one wants to address with EPR. Yet, there are some common requirements which need to be fulfilled for a successful experiment.
In general, spin labeling positions should be located on the surface of the protein. Those positions are easily accessible and it is less likely that the introduction of the label will interfere or disturb the structure of the protein.
Another principle is that the positions to be labeled lie in regions which are well structured. A less rigid region like a loop would add unnecessary flexibility to a possibly already flexible spin label. For these two demands, it is useful to have a structure or a hypothetical model.
For PELDOR experiments, the additional demand of having both labels in a certain distance range arises. Distances between 15-80 Å are reliably and precisely accessed. In fully deuterated samples, even ranges above 100 Å are possible [127].
The PyMol plugin MtsslWizard helps selecting labeling positions optimal for PELDOR measurements. The plugin contains more features (e.g. docking and trilateration), but since the scope of this thesis two nitroxides had to be introduced into the protein in order to investigate conformational changes, only the position optimization will be briefly discussed. More features of this tool can be found in the manual and in the publications of Hagelüken [128][129].
At first the two conformers (apo and holo, loaded as *.pdb from X-ray or NMR experiments) are investigated via a residue-by-residue comparison. The program calculates a distance map for each of the two conformers. Since it uses the distances between the Cβ atoms of the residues, it treats both of the loaded structures as "spin-labeled". To then obtain the changes between these two objects, the distance maps are subtracted from each other, forming a single map with the absolute difference distances.
Figure 1.16 shows an example of a difference distance map from the Cytochrome-Enzyme P450cam. This map now contains two very valuable pieces of information:
The red fields are the "conformational hotspots", those are the residues that should be considered for labeling. They are marked with HS1-HS3 in the image.
Furthermore, the darker regions provide insight into the amino acids which can be

considered rigid during the movement of the protein. These rigid regions are boxed in purple and marked with RB1-RB3.

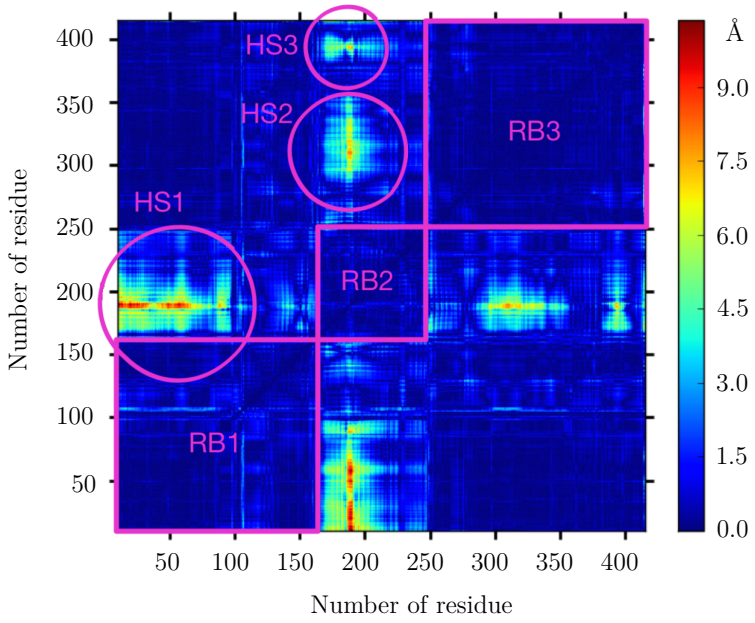


Figure 1.16. – Example of a difference distance map. The protein P450cam was used to generate this example. The conformational change of aminoacids is up to 9Å. The regions with most changes are marked as "hot spots" (HS1-HS3). The regions with least changes are the rigid body regions RB1-RB3. Image taken from [129].

The next step is to pick the right/fitting residues from the conformational hot spots and estimate whether they serve as a good position for labeling. A model of the spin label can now be attached to sites that are desirable for investigation. The algorithm of mtsslWizard can predict a distance distribution between two labels. First, it rotates the labels around the flexible bonds, leading to a bouquet of possible conformations of the label. Hereby, the algorithm discharges the labeling angles which would lead to a clashing conformation of the label with the surface of the protein. The remaining allowed conformations are then weighted evenly. The result is a cone of spin label conformations which represent the accessible volume of the spin label. Finally, a distance distribution between these two cones can be simulated and used to decide if the optimal distance range for PELDOR was achieved. Figure 1.17 depicts the principle. Other tools to plan the labeling and simulate distance distribution are MMM [130] and Pronox [131].

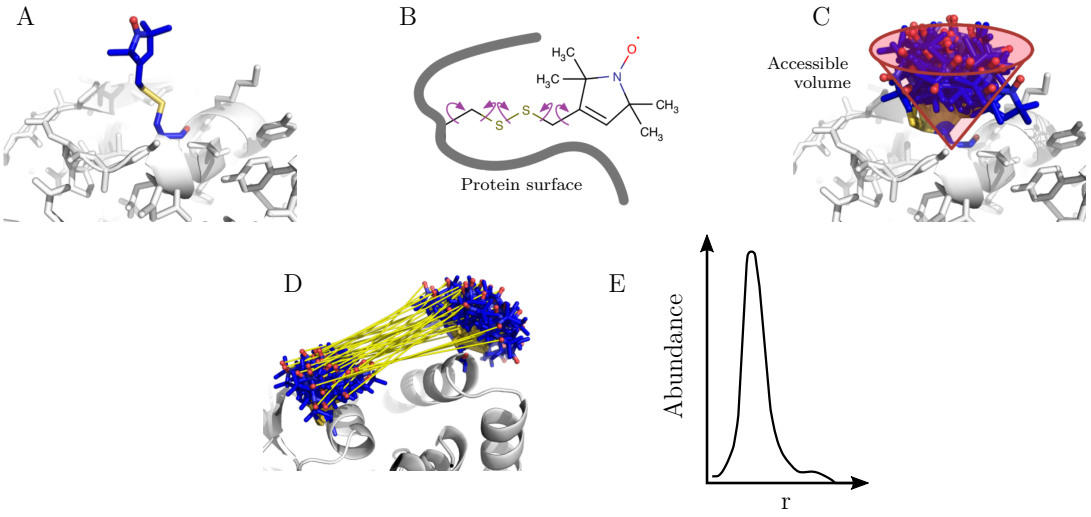


Figure 1.17. – The in silico spin-labeling algorithm used by the program MtsslWizard. A) The label MTSSL is connected to the protein. B) Different rotamers of the label are created. Rotamers which clash with the surface of the protein are eliminated. C) A collection of possible spin labels border the accessible volume of the label. D) + E) The distribution of the distances between all reasonable label conformations is calculated. Image taken from [129].

1.6.2. Continuous wave EPR of Nitroxide

In the practice of continuous wave EPR (cw-EPR), the spectrum is recorded under constant irradiation of microwaves at a fixed frequency while the magnetic field is swept over the region to be investigated. With this technique the nitroxide spin labels attached to the protein of interest can provide information on solvent accessibility, polarity, side chain mobility and can also reveal the distance between the spin labels [132].

In the nitroxide spin label, 95% of the spin population is distributed on the nitrogen and oxygen atoms. Therefore, the label can be described as a spin system with $S = \frac{1}{2}$ coupled to the nuclear spin $I = 1$ of the nitrogen.

The spectrum of a free nitroxide in water measured at room temperature consists of three isotropic lines of equal distance and intensity. They are caused by the hyperfine interaction of the electron with the nucleus ^{14}N.

The lines are centered at an average g value due to fast averaging of the anisotropies, this also leads to their equidistance.

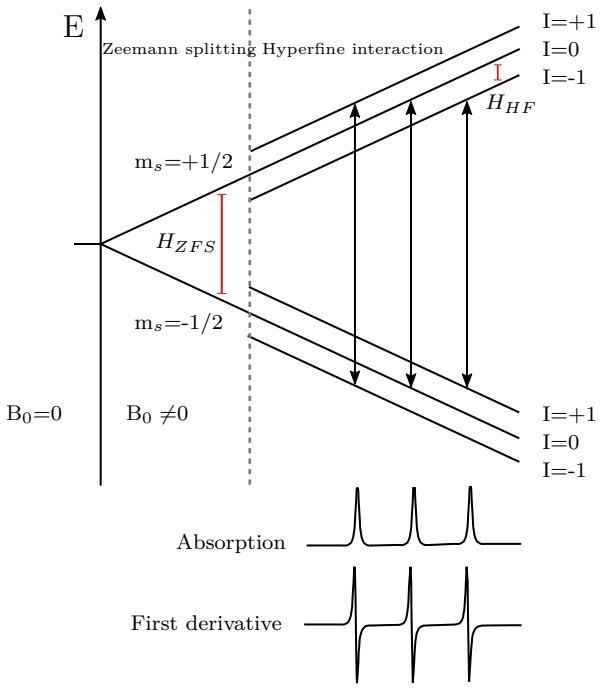


Figure 1.18. – cw-EPR spectrum of a nitroxide in a magnetic field B_0.

For technical reasons, most EPR machines record the first derivative of the absorption spectrum.

1.6.3. Dynamic regimes in cw-EPR

The shape of a cw spectrum is not only determined by static interactions (e.g., Zeemann-splitting, g-tensor, hyperfine couplings to spatially close nuclei), but also by dynamic processes which occur on the timescale of cw-EPR experiments. The most important dynamic process which contributes is molecular tumbling. It is characterized by the rotational correlation time τ_{corr}, which is the time after which molecules with initially identical orientation lose their alignment.

The timescale in cw-EPR experiments is limited by the spectral anisotropy $\Delta\omega$, which is the maximum difference between the line positions when the orientation of the paramagnet is varied. For nitroxides at X-band $\frac{\Delta\omega}{2\pi}$ is approximately 130 MHz.

Comparing $\Delta\omega$ to τ_{corr}^{-1}, four dynamic regimes become distinguishable: isotropic limit, fast motion, slow motion and rigid limit. The borders between these regimes are not sharp, but rather diffuse.

The rigid limit and the isotropic limit are the outerborder extremes of correlation time. In the *isotropic limit* the molecules are tumbling so fast that any aniosotropy is averaged out. Only isotropic properties such as the isotropic g-factor and the isotropic hyperfine coupling constants (a_{iso}) are measurable.

The opposite is true for the rigid limit. Due to the full immobilization of the molecules, the anisotropy of all interactions is fully revealed.
In the fast motion regime, the rotation of the molecule is so fast that the shape of the spectrum appears like in the isotropic limit, but with varying linewidths.
In the *slow motion* regime, the lines are substantially broadened and appear asymmetric due to the slow tumbling of the molecule.
The rotational correlation time of a spectrum can be obtained by numerical simulation of the spectra as explained in the following subsection.

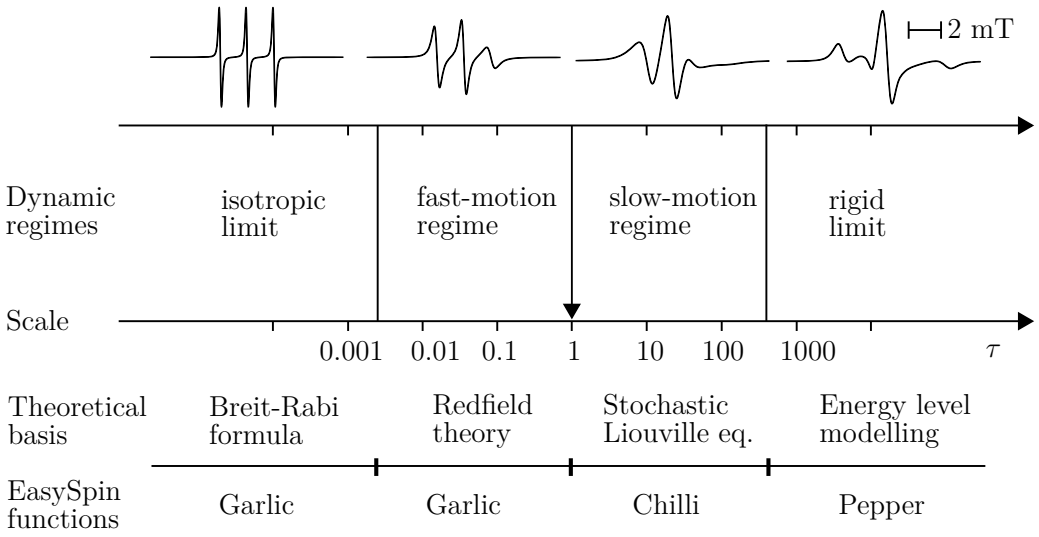


Figure 1.19. – Figure adapted from Stoll and Schweiger [120].

1.6.4. Simulation of cw-EPR spectra

Magnetic parameters are usually directly connected to the shape the cw-EPR spectrum. However, only in the simplest cases these parameters can be directly extracted from line positions, amplitudes and widths. Especially if many magnetic nuclei are present or interactions of similar strength start overlapping, the interpretation of the resulting shape becomes increasingly more difficult.
Numerical simulations of spectra are therefore a useful tool to understand the obtained data. Except for the study of spectral features in relation to the magnetic parameters and the extraction of parameters from a simulated spectrum, another advantage is offered by the simulation of spectra: it can predict whether a new experiment can give rise to new insights in the studied system.
The spectrum simulation is performend by the MATLAB based package *EasySpin* [133], which is able to simulate cw-nitroxide spectra in all of the above mentioned

dynamic regimes.
Figure 1.19 depicts the timescales of rotational dynamics and simulated spectra at X-band frequency. For each regime a different theory is used to describe the dynamics and therefore also a different algorithm is used.

Isotropic limit: $\frac{\tau_{corr}^{-1}}{\Delta\omega} > 1000$, $\tau_{corr} \approx$ 1ps

In the case of the nitroxide, one electron is coupled to one N^{14} nucleus. For this the isotropic spin Hamiltonian is analytically diagonalized by the function Garlic. The energy levels are described by the Breit-Rabi formulae as a function of the magnetic field. The respective resonance field is calculated numerically. Input parameters are g_{iso} and a_{iso} , the nuclear isotope type and the linewidth. Further, the sweep width and the microwave frequency have to be extracted from the experimental setup.

Fast-motion regime: $1000 > \frac{\tau_{corr}^{-1}}{\Delta\omega} > 1, \tau_{corr} \approx$ 0.1ns

In the fast-motion regime the lines do not shift, but experience broadening depending on m_I. The description is based on Redfield theory and the simulation is performed as in the case of the isotropic limit by the function Garlic.
The only difference is that each linewidth is now constructed individually. For this, the isotropic rotational correlation time τ_{corr} has to be passed to the program.

Slow-motion regime: $1 > \frac{\tau_{corr}^{-1}}{\Delta\omega} > 0.001, \tau_{corr} \approx$ 1ns

In this timeframe, the spectral lines are substantially broadened, asymmetric and distorted. The theory which forms the basis of the description of these spectra was developed by Schneider and Freed in the late 1980s and was established from stochastic Liouville equation (SLE) [134, 135].
The model to describe the molecular tumbling in this scenario is isotropic Brownian motion. Within this theory, τ_{corr} can be related to the rotational diffusion coefficient D_r by

$$\tau_{corr} = \frac{1}{6D_r} \tag{1.1}$$

and further, in the Stokes-Einstein hydrodynamic framework, it can be connected to the temperature T and solvent viscosity η via

$$\tau_{corr} = \frac{4\pi\eta r^3}{3kT} \tag{1.2}$$

Also τ_{corr} is linearly dependent on the molecular mass M [136] and can be estimated as :

$$\tau_{corr} = \frac{M}{2.55} \times 10^{-12} s \tag{1.3}$$

The function Chili is used for computing these spectra.

Rigid limit $\frac{\tau_{corr}^{-1}}{\Delta\omega} < 0.001, \tau_{corr} > 1\mu s$

In the rigid limit, paramagnetic molecules are fully immobilized and the spectrum reveals the full anisotropy of all interactions. Such spectra are obtained for powders, frozen solutions and glasses at low temperatures. These so-called powder spectra can be calculated with the function Pepper.

1.6.5. PELDOR

The PELDOR (pulsed electron electron double resonance) experiment makes use of pulses in two different microwave frequency ranges emitted into the frozen protein sample in order to obtain the distance between two spinlabels. Due to the orientation of the nitroxide-labels with respect to the magnetic field, different parts of the same nitroxide-spectrum can be excited with different microwave pulses.

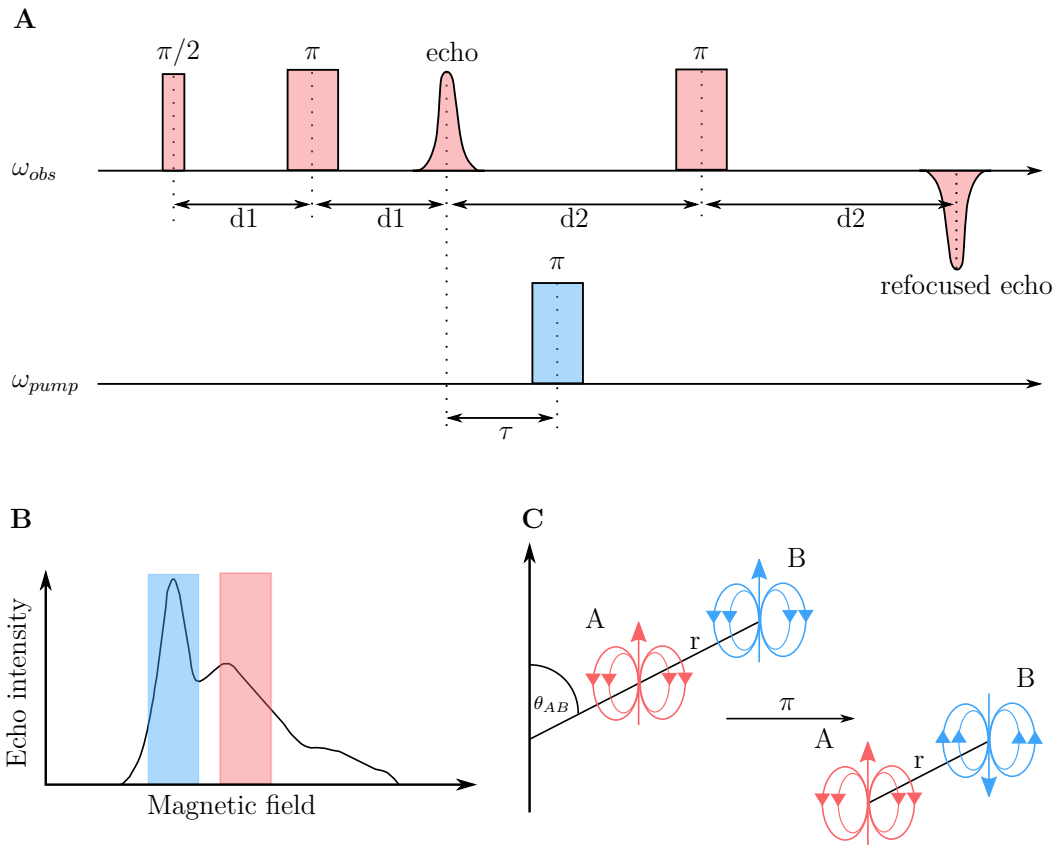


Figure 1.20. – A) 4-Pulse-sequence of PELDOR. B) Field sweep spectrum of Nitroxide. C) The effect of the pump-pulse.

Figure 1.20 A shows the four-pulse scheme, used in this thesis to measure spin-spin distances. The first two pulses $\pi/2$ and π create a so called primary echo of the spins excited with the frequency ω_{obs}. Due to dephasing of the spins, another π pulse at ω_{obs} has to be applied, refocussing the echo. The fraction of the excited nitroxide-spectrum is depicted in figure 1.20 B. A π pulse is applied at frequency ω_{pump} causing a spin-flip on the other nitroxide fraction (see figure 1.20 C). If both spin-fractions are coupled, this will cause a shift in the precession frequency of spin A.

$$\omega_{AB} = \frac{\mu_B^2 g_A g_B \mu_0}{4\pi\hbar} \frac{1}{r_{AB}^3} [1 - 3\cos^2(\theta)] \ , \tag{1.4}$$

where r_{AB} is the distance between spins A and B and θ_{AB} is the angle enclosed by their difference vector and the magnetic field.
As a result, the amplitude of the refocused echo signal changes in comparison to the echo signal in absence of this additional pump pulse. By varying the timing of the pump pulse, an oscillating signal is recorded, the so called time-trace, from which the distance distribution can be calculated.

PELDOR data analysis

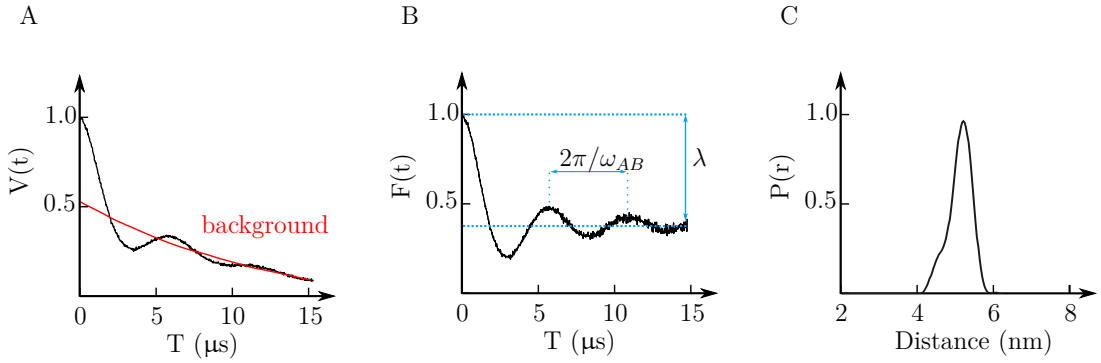


Figure 1.21. – The recorded PELDOR timetrace and data analysis with DeerAnalysis. A) The background (red line) is subtracted from the raw data. B) This is the resulting form factor F(t). It oscillates with the frequency $2\pi/\omega_{AB}$. λ is the modulation depth. C) The subsequent Tikhonov regularization gives the distance distribution of the spin labels.

The obtained signal $V(t)$ as depicted in figure 1.21 is a convolution of two parts: a so called form factor $F(t)$, which arises from the intramolecular interactions of the coupled spins, and an exponentially decaying background function $B(t)$ which is due to intermolecular interactions of the spins.

$$V(t) = F(t)B(t) \tag{1.5}$$

For a normalized signal (i.e. $V(0) = 1$), $B(t)$ is given by a stretched exponential

$$B(t) = \exp(-(kt)^{\frac{D}{3}}) \tag{1.6}$$

with k specifying the density of the spins in a space with the dimension D. Frozen glassy solutions of proteins are assigned a dimension of $D = 3$, whereas, e.g., reconstituted

membrane proteins can be viewed as trapped in a bilayer and therefore lose one degree of freedom ($D = 2$).
To deconvolute the form factor $F(t)$ from the background, the primary PELDOR data $V(t)$ has to be divided by the background $B(t)$. Considering one pair of spins AB with distance r_{AB} at angle θ_{AB} to the magnetic field, the form factor is a function of the dipolar frequency ω_{AB} and the modulation depth λ:

$$F(t) = 1 - \lambda[1 - \cos(\omega_{AB}t)] \tag{1.7}$$

The modulation depth quantifies the fractions of spins that are flipped after the pulse. In the sample, there are many well-separated spin pairs with random orientations. To obtain the final form factor, average equation 1.7 over all possible values of θ, leading to:

$$F(t) = 1 - \lambda + \lambda \int_0^{\pi/2} \cos[\omega_{AB}t]\sin(\theta)d\theta \tag{1.8}$$

In the next subchapter, the distance distrubution $P(r)$ will be retrieved from the form factor F(t).

1.6.6. Data analysis

A general reference for this section is [137]. The obtained PELDOR data can be now analyzed with the program DeerAnalysis, which is a MATLAB-based software. It calculates a distance distribution $P(r)$ from the diplolar evolution function $F(t)$.
The primary data $V(t)$ has to be deconvoluted from the background, thus leading to the form factor $F(t)$. Fitting the form factor $F(t)$ with simulated data $S(t)$ and subsequent Tikhonov regularization provides the interspin distance distribution $P(r)$. But calculating $P(r)$ from $F(t)$ is a mathematically ill-posed problem. Signal-to-noise ratio is crucial for the reliability of the obtained distribution, since small distortions in the data can have large effects on the distribution. To solve this problem the solution has to stabilized against noise-induced artefacts. This is done by requiring a certain degree of smoothness of $P(r)$.
From a given distance distribution $P(r)$ a simulated time domain signal $S(t)$ can be calculated by using a Kernel function $K(t, r)$ as

$$S(t) = \int_0^\infty K(t, r)P(r)dr \tag{1.9}$$

Assuming ideal pulses for the PELDOR experiment, $K(t,r)$ is given as:

$$K(t,r) = \int_0^{\pi/2} \cos[\omega_{AB} t] \sin(\theta) d\theta \tag{1.10}$$

This problem can be solved by using the method of Lagrange multipliers. The regularization parameter α is included in a cost function

$$G_\alpha(P) = \int_0^1 |S(t) - D(t)|^2 dt + \alpha \int_0^\infty \left| \frac{d^2}{dr^2} P(r) \right|^2 dr \tag{1.11}$$

The first part of the term calculates the error between simulated form factor S(t) and the experimental signal evolution D(t) as the mean square deviation.
The second part of the term calculates the roughness of P(r) as the square norm of the second derivative. α gives a compromise between these two summands.
Denote the minimizer of G_α by P_α, and its signal S_α. Introduce the α-dependent quantities

$$\rho(\alpha) = \int_0^1 |S_\alpha(t) - D(t)|^2 dt \tag{1.12}$$

and

$$\eta(\alpha) = \int_0^\infty \left| \frac{d^2}{dr^2} P_\alpha(r) \right|^2 dr \tag{1.13}$$

For data with good signal-to-noise ratio, the log-log plot of ρ against η is L-shaped. Then, the optimal regularization parameter α is obtained as the position of the kink.

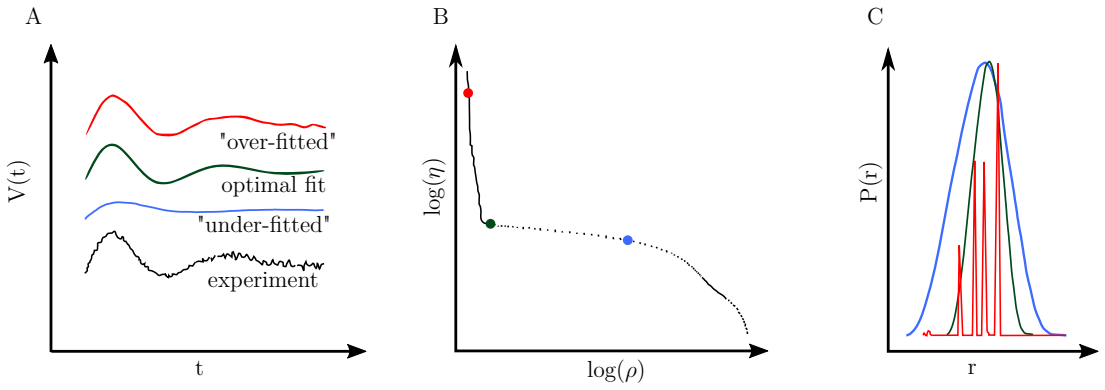


Figure 1.22. – Example of an L-curve calculated with DeerAnalysis. The green dot marks the optimal regularization parameter.

1.7. Aim of this thesis

TRAP transporters have been discovered in the late 1990s and since then only little has been understood regarding their mechanism of transport. This thesis aims to gain a better insight into the transporting mechanism of the substrate Neu5Ac by TRAP transporters, focusing mainly on the pathogen *Vibrio cholerae* and also *Haemophilus influenzae*. In both organisms the TRAP transporter is the only transporter of Neu5Ac into the cytoplasm of the cell. Also, in both organisms the virulence of pathogenity depends largely on the uptake of Neu5Ac from the environment. The understanding of the mechanism would open up a new target to disrupt the development of virulence of these pathogens. This aim is approached by different methods of integrated structural biology, such as X-ray crystallography, electron paramagnetic resonance spectroscopy and molecular biology techniques.

Investigation of the equilibrium At the beginning of the work for this thesis, crystal structures of VcSiaP in open conformation and of HiSiaP in open and closed conformation were available, as well as crystal structures from other TRAP P-domain homologs. It has not been studied before, whether the substrate induces the conformational change or if the P-domain switches constantly between open-, closed and possibly intermediate conformations.
To find this out, PELDOR in combination with SDSL shall be used to examine the conformational sampling of the P-domain

Investigation of the binding cleft It was already known that R145 acts like a filter for substrates by recognizing the carboxylic acid group of Neu5Ac.
Regarding the conformational switch mechanism, a synergetic interaction chain between the residues Arg127, Glu184 and His207 was proposed by Setty et al. [99] after detailed analysis of TRAP P-domain homologs. By mutating these allegedly important residues into residues of similar charge and length the mechanism of the binding cleft can also be studied by PELDOR.

Peptide scan and growth experiments During the course of the binding cleft study of VcSiaP, crystal structures of VcSiP were obtained by Martin Peter in the scope of his master thesis. One structure revealed a homo dimer of VcSiaP, connected by the His-tag cleaving site in the binding cleft. This finding inspired new methods to target the interaction between P and QM domains.

To investigate to what other possible QM structure related peptides the P-domain could bind, a peptide scan is combined with uptake assays and growth experiments.

Rx labeling of VcSiaP In order to relocate Neu5Ac into the QM-domains, an interaction between the P-domain and the QM-domains has to take place. However, this interaction has not been detected yet.
This challenging task is approached with labeling the P-domain with the more rigid Rx-label. The effect of spectral broadening upon increase molecular size due to complex formation seems feasible for this question.

2. Results

2.1. Studying the open-closed transition of VcSiaP with EPR techniques

The experiments presented in this subchapter were all aimed at investigating whether the open conformation of the SBP of *Vibrio choerae* coexists in solution with the closed conformation or not.
To study the SBP behavior in frozen solution with EPR spectroscopy, optimal labeling sites had to be identified. The process of selecting optimal labeling sites and the purification of the respective proteins is described. Furthermore, a model was built to aid the interpretation of the obtained data. The *cw*-EPR spectra and the recorded PELDOR data of the mutants is presented. The K_D of the binding cleft was determined with PELDOR spectroscopy and also with surface plasmon resonance (SPR). Finally, a crystal structure of a spin labeled mutant is described.

2.1.1. Finding optimal labeling sites for PELDOR spectroscopy

To study the open-closed transition of VcSiaP in frozen solution with PELDOR spectroscopy, spin labels had to be attached to the molecular surface of the protein. With mtsslWizard, a difference distance matrix (diffDM) was created, using the substrate-bound and substrate-free form of the homologous protein HiSiaP from *Haemophilus influenzae* (PDB-IDs: 3B50 [138], 2CEY [97]). Currently only a structure of open VcSiaP (PDB-ID: 4MAG [99]) exists while the closed, substrate-bound structure remains elusive. HiSiaP and VcSiaP have a sequence identity of 50%, which makes this substitution valid (see appendix A.2 for the multiple sequencing alignment of VcSiaP and HiSiaP).

Figure 2.1 shows the obtained diffDM. The violet regions indicate regions with no change between the Cβ-Cβ distances of the respective residues. The yellow "hot spots" on the other hand are regions where large distance changes between $12 - 18\,\text{Å}$ occur. From this matrix the residue pairs Q54/L173, Q110/L173, L173/S225 and Q54/Q110

were selected as labeling sites and are marked with white circles in figure 2.1. Residue pair Q54/Q110 functions as control, since it was not selected from the hot spot regions.

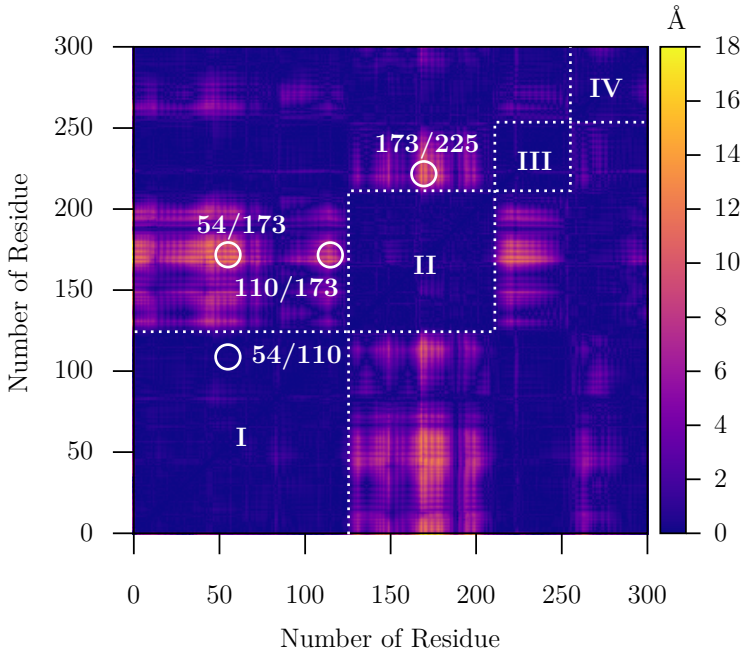


Figure 2.1. – The difference distance map of VcSiaP from mtsslWizard. It was calculated with open and closed HiSiaP as model for VcSiaP. The selected labeling sites are indicated with white circles. The dashed boxes border the rigid bodies of the protein.

The square boxes along the diagonal of the matrix can be understood as the subdomains of the protein which stay rigid during the conformational change. Four rigid domains (I-IV) were identified by this analysis and are depicted in figure 2.2. Region I is marked by residue section 1-127 (gray), region II by section 128-208 (green), region III consists of 209-260 (blue) and region IV of 261-307 (beige).

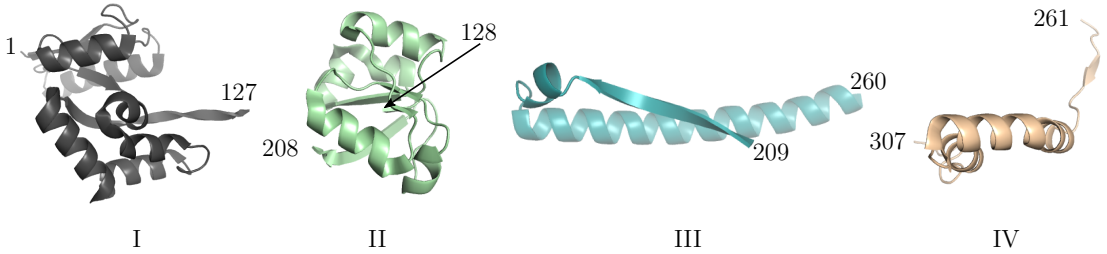


Figure 2.2. – The four rigid subdomains of VcSiaP. These subdomains were identified as rigid by the tool mtsslWizard. The range of amino acids the respective subdomains consists of is indicated.

These regions were then used for model building. The rigid domains of the open structure were superimposed on the closed and substrate bound structure of HiSiaP (PDB-ID: 3B50). The obtained model was then further used for predictions in this thesis (see chapter 2.1.5).

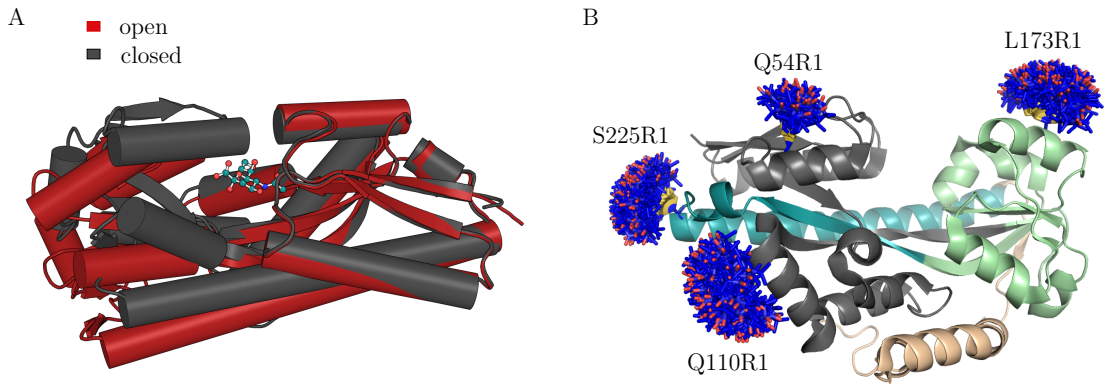


Figure 2.3. – Open and closed HiSiaP and VcSiaP with labels. Panel A is a superposition of open (red) and closed (gray) HiSiaP structures, which were used to create the difference distance matrix. Panel B shows VcSiaP with ensembles of MTSSL attached to all of the chosen labeling sites. The color coding of the domains is the same as used in figure 2.2.

Figure 2.3 A shows a superposition of the open and closed HiSiaP structures. The substrate Neu5Ac can be found in the center as ball and stick representation. In subfigure B the determined subdomains of VcSiaP are marked in their respective colors. Notice also that the selected labeling sites are located on three of four rigid subdomains.

2.1.2. Cloning, expression, purification and labeling

The VcSiaP genetic sequence was amplified from genomic *Vibrio Cholerae* DNA via PCR by Gregor Hagelüken and cloned into the pBADHisTev vector. The primers for amplification were designed to eliminate the N-terminal signal sequence of VcSiaP (MKTINKITIAILTLSAAASVNA), producing a construct with the length of 300 amino acids (see appendix section A.1).
Mutations were inserted into the sequence by quick change mutagenesis as described in section 4.2.1. The correct mutants were verified by sequencing.
The expression vectors of the selected mutants were then transformed into C43 cells as described in section 4.2.5.

Expression in LB media

The expression and purification of VcSiaP Q54R1/L173R1 was tested in 3 L LB media. At every purification step a protein sample was prepared as described in section 4.4.1 and an SDS-PAGE was performed. The gels from this protocol are depicted in figure 2.4.

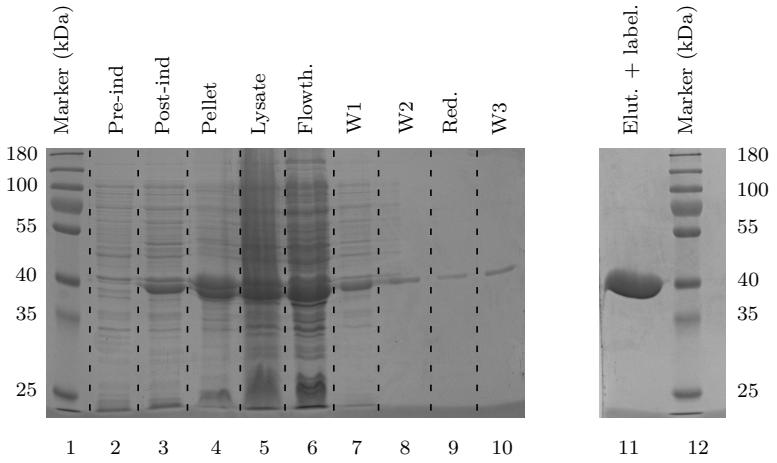


Figure 2.4. – SDS-PAGE of the expression and purification of VcSiaP Q54R1/L173R1 from 3 L LB media. Depicted are samples from protein expression (lane 2-4) as well as samples from the Ni^{2+} affinity chromatography. Lane 11 marks the eluted and MTSSL labeled protein.

After inducing with L-arabinose (500 mg/L) at an OD_{600} of 0.6, the protein is clearly expressed, which is proven by lane 3 of the gel depicted in figure 2.4. Here, a protein band arises at roughly 40 kDa. The construct of VcSiaP with Tev-cleavage site and His_6-tag at the N-terminus weights 37.7 kDa.

The cells were harvested and lysated. The cell debris (lane 4) was separated from the lysate (lane 5) via centrifugation and was disposed. The lysate was incubated with nickel-beads and transfered to a polystyrene column. The flowthrough is visible in lane 6. The beads were washed twice with 50 mL buffer (lane 7 and 8) and the cysteines were reduced with 1 mM TCEP (lane 9). The excess TCEP was removed with another washing step (lane 10) and the protein was labeled with sixfold excess MTSSL in the elution step. The labeling buffer also contained 500 mM imidazole (lane 11).
Finally the eluate was subjected to size exclusion chromatography on a Superdex 75 16/60 column. The chromatogram can be seen in figure 2.5.

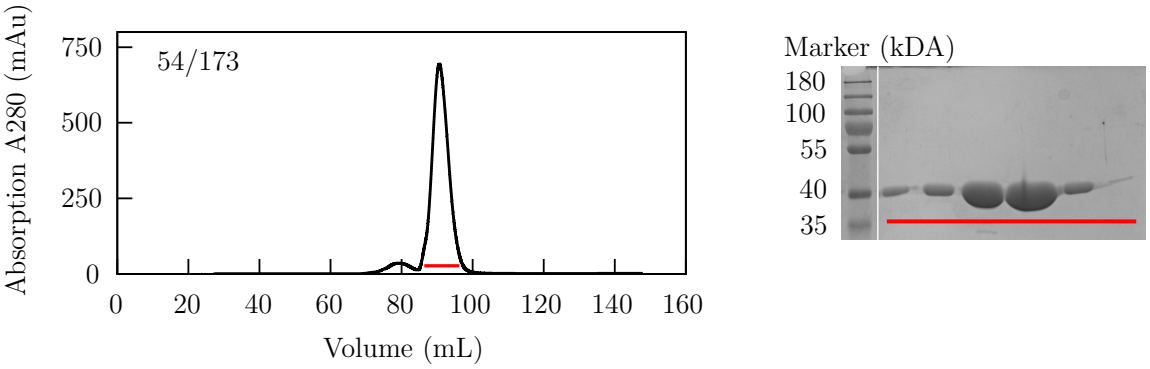


Figure 2.5. – Chromatogram of the size exlucion chromatography of VcSiaP Q54R1/L173R1 from 3 L LB media. The respective SDS PAGE is displayed on the right side of the chromatogram. The collected fractions are indicated in red.

To check whether the labeling was successful, a *cw*-X-band spectrum of the collected batch was recorded at room temperature (see section 1.6.2). Figure 2.6 shows the result of this experiment. The spectrum reveals three transition lines of unequal amplitude. In a *cw*-EPR spectrum with unbound nitroxide which can tumble freely in aqueous solution the transition lines are equidistant and of same intensity (see section 1.6.2). But, if bound to a macromolecule, the tumbling slows down and the anisotropy can no longer be averaged out, thus leading to a spectrum with broader and unequal lines, just as the one observed. The labeling was therefore successful. Furthermore, no additional and spiky lines emerged which adds to the conclusion that no free label contaminates the sample and the signal only stems from bound MTSSL.

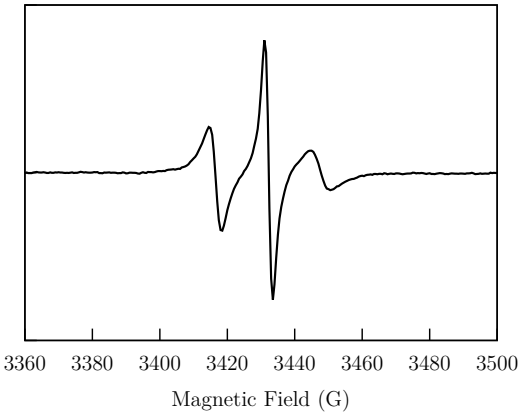


Figure 2.6. – *cw*-X-band spectrum of VcSiaP Q54R1/L173R1 which was expressed in LB medium. The measurement was performed at room temperature. The label is fully attached to the protein, no free label contaminates the sample.

A PELDOR timetrace of this batch was recorded and analyzed. It was recorded with a time window of 2.5 µs at 50 K. The sample was prepared as described in section 4.6.2. It revealed a surprisingly high amount of closed conformation, which arose from Neu5Ac in the LB medium. Figure 2.7 shows the result: the distance distribution peaks at 43 Å which can be attributed to the open conformation. But it also shows a peak at 27 Å which represents the closed conformation.

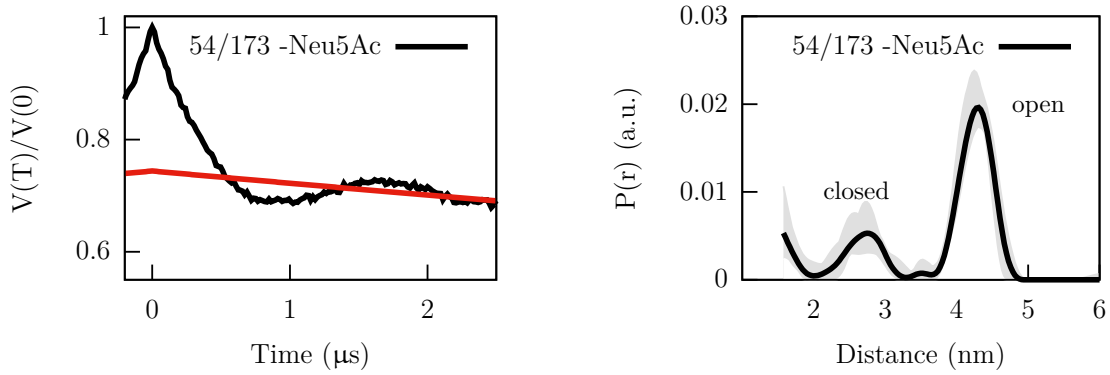


Figure 2.7. – PEDLOR timetraces and distance distributions of VcSiaP Q54R1/L173R1 from LB media protein expression. The background of the measurement is indicated by a red line in the timetrace plot. The error estimated by DeerAnalysis is marked in gray in the plot of the distance distribution.

To avoid contamination and co-purification with Neu5Ac from regular LB-media [139], an expression protocol in M9 media was developed and established (see section 4.4.3). The lag-phase of C43 cells in M9 media is longer than in regular LB media: the cells were therefore grown for 14-16 h prior to induction with L-arabinose (500 mg/L).

Expression in M9 minimal media

All four selected double-cysteine mutants were grown and expressed in M9 media. The purification was performed as described above. Figure 2.8 shows the final result of the purification.

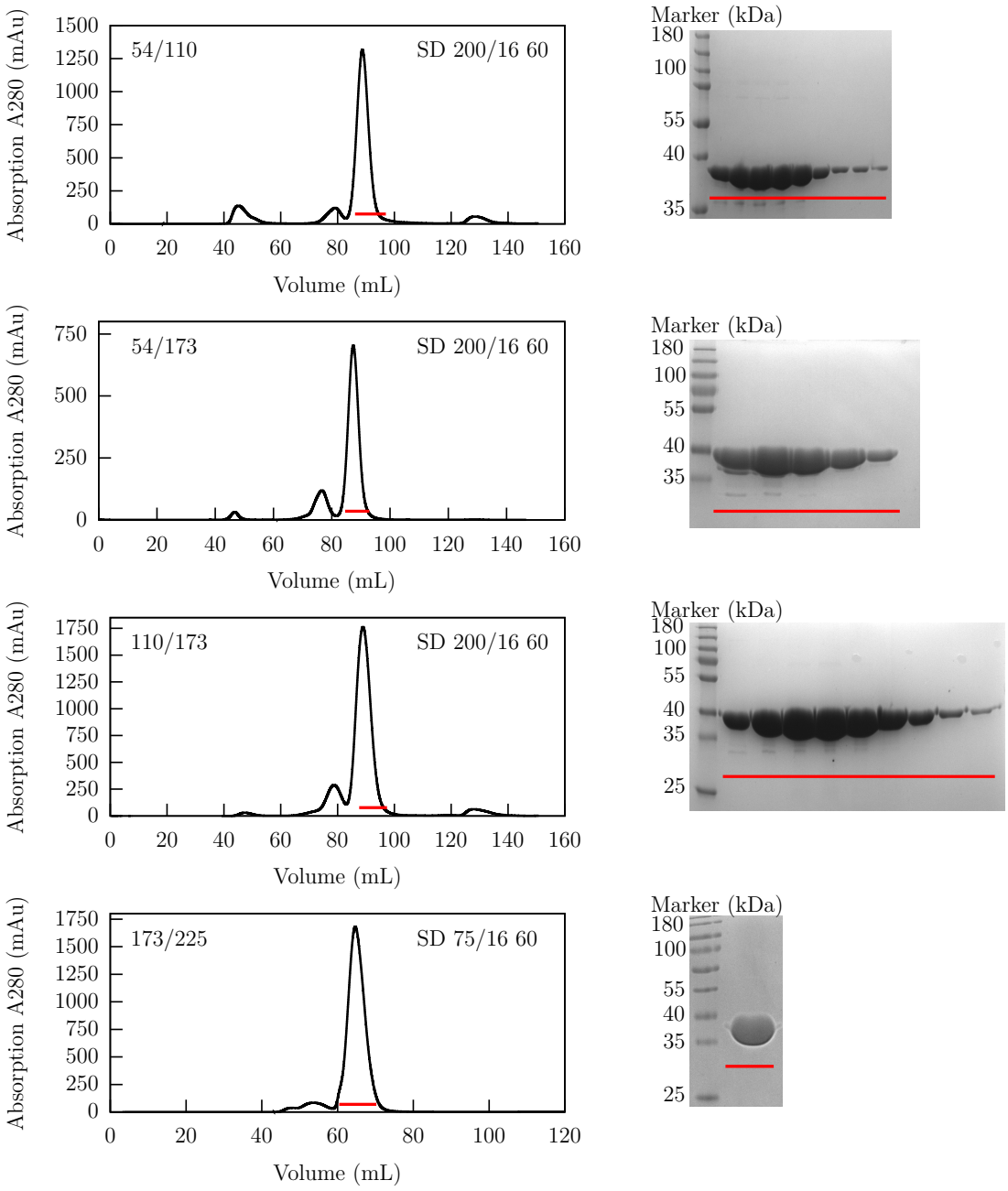


Figure 2.8. – Chromatograms of the size exclusion chromatography for the four mutants. The respective SDS PAGE is displayed on the right side of the chromatogram. The collected fractions are indicated in red.

The pooled fractions are marked in red. 6 L media yield roughly 50 mg of protein. When using a Superdex 200 16/60 column the protein peaks at around 90 mL. This is the case for the mutant pairs 54/110, 54/173 and 110/173. The size exclusion chromatography of mutant pair 173/225 was performed with a Superdex 75 16/60 column and therefore the protein fraction peaks at a different volume on to chromatogram. Here the peak arises at 77 mL.

The labeling efficiency was then determined with room temperature *cw*-EPR spectra at X-band frequency. An average labeling efficiency of $> 90\%$ was achieved. A detailed description of these results follows in chapter 2.1.4.

In a nutshell: the optimized M9 minimal media expression protocol delivered the desired protein constructs with high purity and very good labeling efficiency.
The question whether the label interferes with the protein structure still remains. In the following subchapter a crystal structure of VcSiaP Q54R1/L173R1 is presented which provides insight into the relation between the protein and the label.

2.1.3. Crystal structure of VcSiaP Q54R1/L173R1

To check if the introduced labels had altered the structure of the protein, crystallization plates were set up as described in chapter 4.5. JCSG+ crystallization screens with VcSiaP Q54R1/L173R1 yielded crystals in the folllowing condition: 0.2 M Potassium thiocyanate, 20% w/v PEG 3350. These crystals were picked and a dataset was collected at the Helmholtz Zentrum in Berlin with the electron storage ring Bessy II. The beamline Bessy BL 14.2. was used. The crystal diffracted up to a resolution of 2.99 Å. The data was processed and scaled in the space group $P2_1$. The structure was solved via molecular replacement using VcSiaP (PDB-ID:4mag) as a model. Positive difference electron density for the label was observed for only one of the two expected labeling sites. Here, models of the label could be placed into the electron density. Refinement was performed until R/R_{free} converged at 22.8/27.6. Crystallographic details on data collection and refinement are summarized in table 2.5.

Table 2.1. – Data collection and refinement statistics of VcSiaP Q54R1/L173R1.

Parameter	**Value**
Mutant	VcSiaP Q54R1/L173R1
Wavelength (Å)	0.89429
Resolution range (Å)	35.717 - 2.994 (3.029 - 2.994)
Space group	$P2_1$
Unit cell	134.3, 109.4, 156.8 / 90, 108.9, 90
Total reflections	840113 (41466)
Unique reflections	330905(41466)
Multiplicity	3.9 (3.7)
Completeness (%)	96.9 (90.0)
Mean I/sigma(I)	5.1 (0.6)
Wilson B-factor	66.5
R-merge	0.226 (2.438)
Reflections used in refinement	79710 (6978)
Reflections used for R-free	4090 (371)
R-work	0.2283 (0.3770)
R-free	0.2764 (0.4215)
RMS(bonds)	0.004

RMS(angles)	0.859
Ramachandran favored (%)	99.3
Ramachandran allowed (%)	99.9
Ramachandran outliers (%)	0.09
Rotamer outliers (%)	0.77
Clashscore	3.77

This structure contains 12 VcSiaP proteins (chain A−chain L) in the unit cell and reveals an incomplete labeling of the proteins. Instead of binding to the MTSSL label, the residue L173C formed a disulfide bond with another L173C residue of a neighboring VcSiaP protein. Thus, two VcSiaP proteins now form a bonded dimer. Figure 2.9 A shows such a dimer pair. The disulfide bond which connects both proteins is indicated with an arrow.

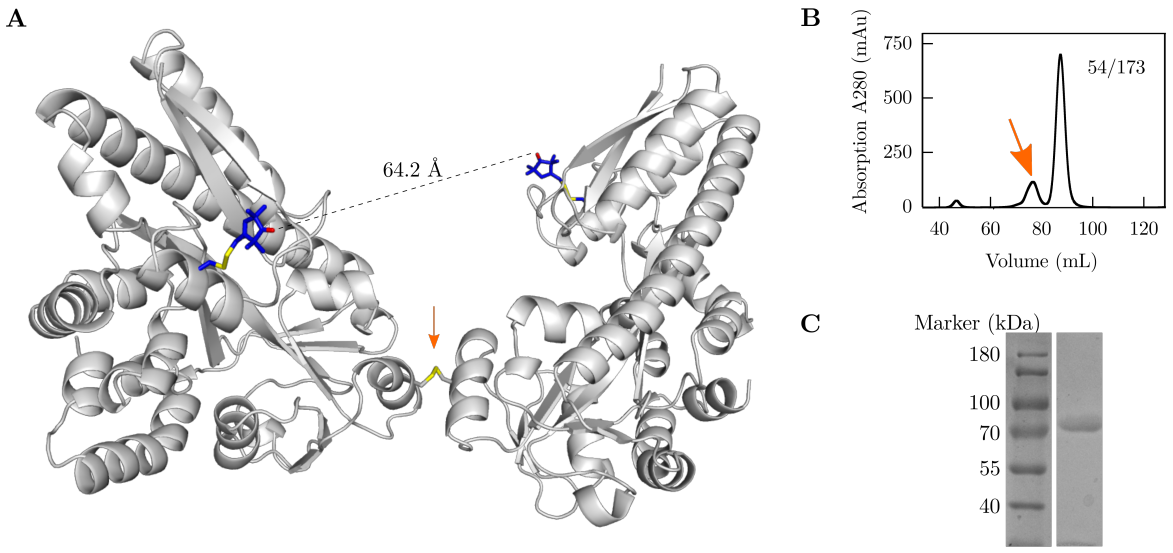


Figure 2.9. – Structure of connected VcSiaP dimer. Panel A shows two VcSiaP Q54R1 L173C proteins. Both are connected via a disulfide bond formed by the introduced cysteines at L173. The disulfide bond is marked by an orange arrow. In panel B the arrow indicates the peak of the crystallized VcSiaP dimer (at 76 mL) in the size exclusion chromatogram. The larger peak at 90 ml represents the monomer which has the spin labels attached at Q54 and L173.

This structure is an aberration from the double labeled monomers that should have crystallized. A look at figure 2.9 B and C hints at why this structure was obtained:

during the purification the cysteines have to be reduced with TCEP to prevent them from forming disulfide bonds with each other. Still, small fractions of the protein form dimers (compare figure 2.8). Panel B of the figure shows a dimer peak of the protein at 76 mL in the size exclusion chromatogram (SD200 16/60). It is marked with an orange arrow. Panel C of the same figure shows the respective SDS PAGE of the dimer peak (non-reducing gel, without β-Mercaptoethanol).
Figure 2.10 A visualizes the superposition of all 6 dimers: it shows the flexibility of the dilsulfide bond. Figure 2.10 B reveals the MTSSL label bound to the protein. Introduction of the label has not perturbed the overall structure of VcSiaP.

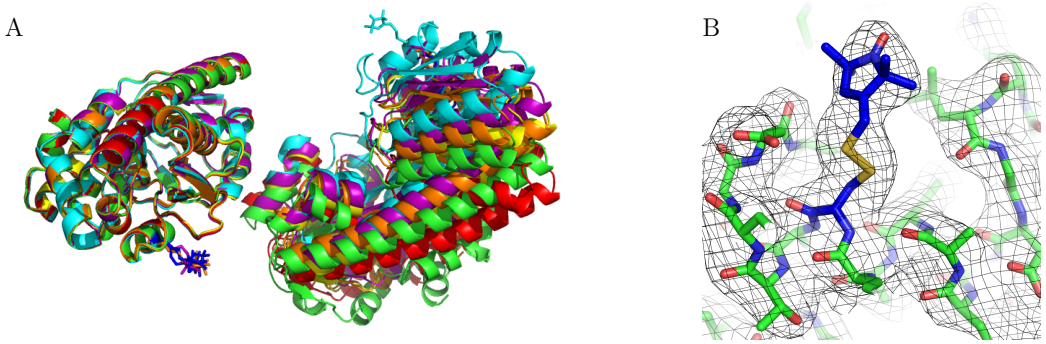


Figure 2.10. – VcSiaP dimers and label. Panel A shows the superposition of all VcSiaP dimers found in the crystal structure. Panel B shows how the electron density map fits perfectly around the MTSSL label.

Since there are 12 R1 spin labels available in this structure it is an excellent candidate for studying the dihedral angles. *In silico* spin labeling programs such as mtsslWizard help to plan and simulate experiments. To be able to improve the algorithms, experimental data has to be generated and compared with existing simulations.
Existing libraries such as RosettaEPR [140] need to be expanded by providing experimental data.
The R1 spin label contains five dihedral angles denoted by χ_1, χ_2, χ_3, χ_4 and χ_5 along the axis C_α-C_β-S_γ-S_δ-C_ϵ-C_3 (see appendix, section A.5). Furthermore, the definition of Tombolato et al [141] is used for the angle χ_5. Each of these dihedral angles has multiple rotameric states, which are also referred to as canonical angles. The description follows the nomenclature suggested by Lovell [142] for protein structures. For χ_1, χ_2 and χ_4 the stable rotamers are defined by the three angles 180°, +60° and −60° (here: 300°). For χ_3 and χ_5 the stable rotamers are at +90° and −90° (here: 270°).

Figure 2.10 displays polar plots of the dihedral angles (χ_1 – χ_5) of VcSiaP Q54R1/L173C (orange lines). The grey segments surround the canonical angles with an error intervall of ±30° [143]. The blue trajectories are the angular distribution of a rotamer library derived from MD simulations of the free MTSSL label performed by Polyhach [144].

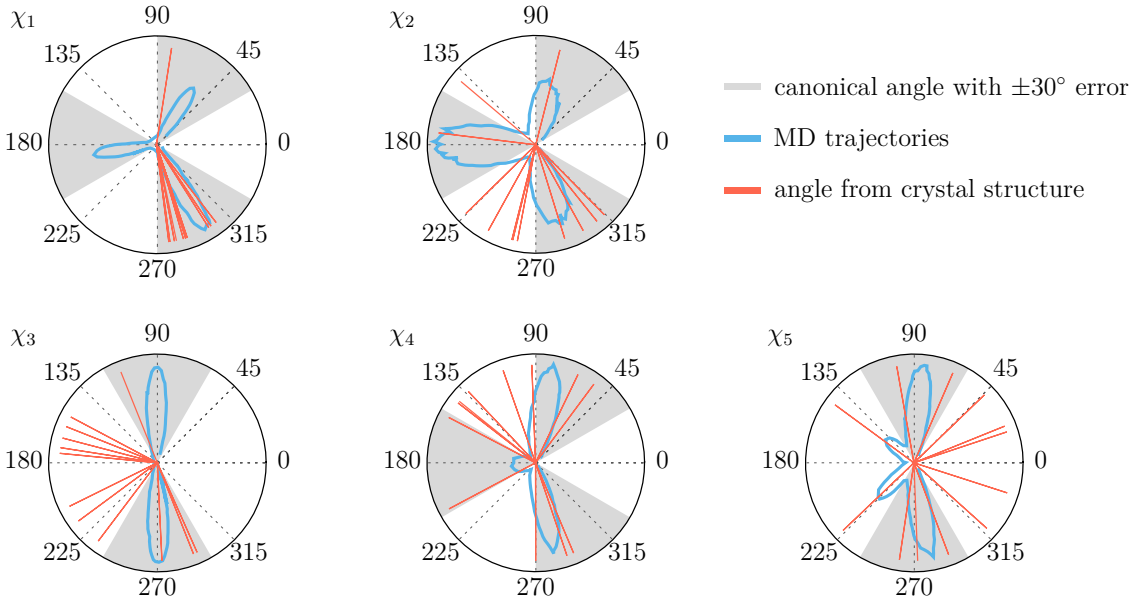


Figure 2.11. – Distribution of dihedral angles of the observed R1 side chains in VcSiaP Q54R1/L173C. The grey sections mark the canonical angles of the label with a ±30° confidence interval. The blue lines are from a computationally produced library for the free spin label.

For χ_1 the experimental angles fit perfectly into the range of the canonical angles and also have a great overlap with the MD trajectories. All rotamers except for one reside in the **m** conformer, which is the most common conformer for χ_1 [145]. Deviations occur for the other four angles.

I defined an outlier to be neither in the error range of the canonical angles nor covered by the trajectories of the MD simulations. Judging by this definition, there are no outliers for χ_2, but 5 angles which are outside of the canonical range. The same applies to χ_4. For the angles χ_3 and χ_5, the case is different and clear outliers can be observed. χ_3 counts 8 outliers while there are 5 for χ_5.

At this point it is helpful to look at the electron density of the label which is visualized in appendix section A.6. The electron density for the spin label on chains A – F is of very good quality, while chain L displays average quality and chains H and K reveal poor quality. However in all chains the position of the NO group could be identified

and subsequently the overall conformation of the label. With this kind of data, there is no reason to exclude these outliers [128].
Deviations from χ_3 have been observed in T4L K65R1 [146], TNFα T77R1 (PDB-ID:5UUI) [147], Azurin T30R1 crystal form II (PDB-ID:5I28) [148] and GB1 E15R1 (PDB-ID:5BMG) [149].

2.1.4. CW EPR data of the four mutants

To further investigate the transition from unbound to bound VcSiaP in solution *cw*-EPR spectra were recorded at room temperature using X-band frequency. The exact procedure is described in section 1.6.2.
Figure 2.12 summarizes the results. All spectra show the three transition lines, which are typical for the nitroxide label. The different heights of the peaks reveals that the label is bound to the protein and is therefore restricted in its mobility. No free label was observed.

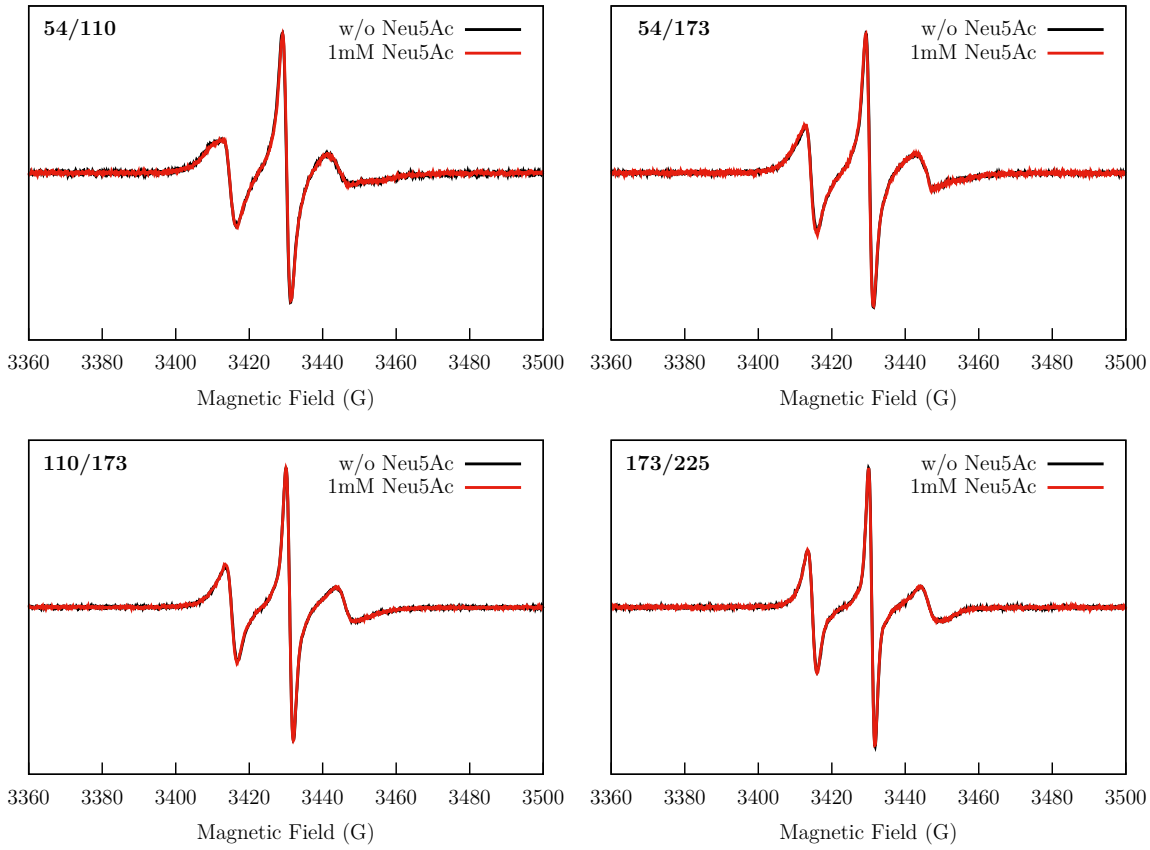


Figure 2.12. – *cw*-X-band spectra measured at room temperature of the four VcSiaP mutants. One measured without Neu5ac (black) and supplemented with 1 mM Neu5Ac (red). For each mutant both spectra are superimposed.

To determine the labeling efficiency of the four mutants, the double integration function of XENON software was used and the result compared with standard solutions of TEMPO. (A standard curve was recorded by Erik Schubert.) The obtained labeling efficiencies are outlined in table 2.2. An overall average labeling efficiency of 93.44% was achieved with the protein purification protocol established in section 2.1.2.

Table 2.2. – Labeling efficiencies of the four double cysteine mutants.

Mutant	54/110	54/173	110/173	173/225	Ø
Efficiency	97.36%	95.7%	108.7%	72%	93.44 %

As can be seen in figure 2.12 the spectra of the unbound (black) and Neu5Ac-bound (red) form of each mutant can be perfectly superimposed. This indicates that the mobility of the nitroxide label does not change when the substrate is attached to the protein.
This is contrary to the PELDOR measurements of the subsequent subchapter, where changes in the distance distribution are quite drastic upon substrate binding and are in support of rigid body movement.

2.1.5. PELDOR data of the selected VcSiaP double-mutants

PELDOR timetraces of the selected proteins were recorded as described in section 4.6.2. Figure 2.13 shows the background subtracted time traces and the distance distributions obtained with DeerAnalysis. Raw data can be found in appendix A.3. Each protein was measured with 1 mM Neu5Ac (black traces) and without Neu5Ac (red traces). Furthermore, the picture contains predicted distance distributions from mtsslWizard (blue shade). The orange enveloping curve is the error calculated with the evaluation feature of DeerAnalysis.
The time traces of the control, VcSiaP Q54R1/Q110R1, first show an initial decay and then submerge into a few clear oscillations before leveling out. The calculated distance distributions show a major peak at 27 Å followed by a smaller peak at 30 Å. The predicted distance distributions are in good agreement with the distributions from the experiment. As expected, both time traces are nearly identical. Rechecking figure 2.3 B explains why: both labels are located on the same rigid subdomain.
This is not the case for the other mutants which were selected from the hot spots of the diffDM.
The time trace of VcSiaP Q54R1/L173R1 without Neu5Ac (black) shows longer oscillations while the trace with Neu5Ac (red) flattens out very early. For this mutant the time traces differ the most compared to the other constructs. As a consequence distance distributions with main peaks at different distances are obtained: 43 Å for the apo protein and 27 Å for the protein with Neu5Ac. Also, here the predictions fit well to experimentally obtained distance distributions.

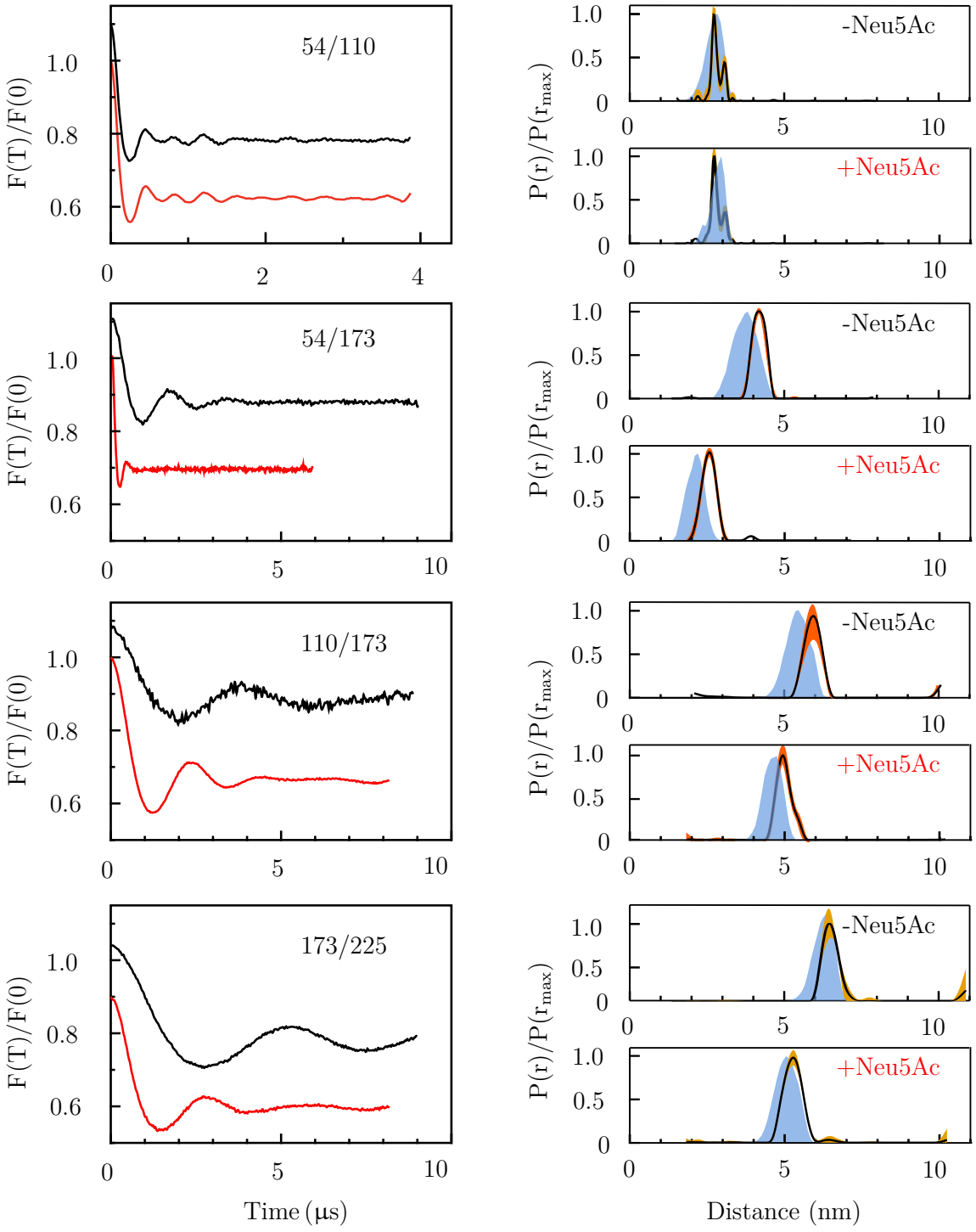


Figure 2.13. – Background corrected PELDOR time traces of the indicated VcSiaP double labeled mutants with the respective distance distributions.

Time traces of VcSiaP Q110R1/L173R1 and L173R1/S225R1 show longer oscillations than the other two constructs.

For VcSiaP Q110R1/L173R1 this leads to maxima of the distribution at 59 Å (-Neu5Ac) and 50 Å (+Neu5Ac). For VcSiaP L173R1/S225R1, the main peaks are

located at 64 Å (-Neu5Ac) and 53 Å (+Neu5Ac). Also for these two mutants the predictions fit well to the experimentally obtained distributions.
The clearest difference between substrate-bound and substrate-free form can be seen in VcSiaP Q54R1/L173R1. The difference between both forms is 16 Å. This mutant was therefore deemed fit for further investigation of the dynamics of VcSiaP.

2.1.6. PELDOR titration of VcSiaP with Neu5Ac

To quantitatively investigate the binding of Neu5Ac to the P-domain, the selected double spin-labeled mutant VcSiaP Q54R1/L173R1 was titrated with Neu5Ac ($0 - 600\,\mu M$) and subjected to PELDOR measurements.
This particular mutant was chosen for its long distance between the peaks of the distance distributions of the unbound and bound form of the protein. Additionally, both distributions have no overlap and therefore the analysis of the data will be facilitated.
Figure 2.14 shows the recorded and background-corrected Q-band PELDOR time traces. The step-wise transition from longer oscillations to fewer and shorter oscillations with increasing amount of substrate is clearly visible. The distance distributions were again calculated by DeerAnalysis and their total weight was normalized to unity. The distance distributions clearly show a shift between the two main peaks (27 Å and 43 Å), which were already determined in section 2.1.5. No additional peak arose in between.

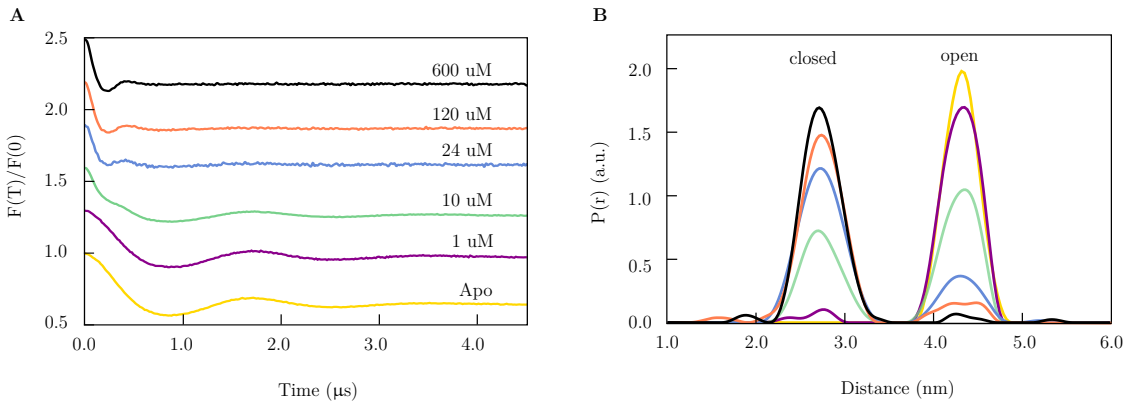


Figure 2.14. – Titration of VcSiaP Q54R1/L173R1 with Neu5Ac. Subfigure A shows the background corrected PELDOR timetraces and subfigure B displays the corresponding normalized distance distributions.

To be able to calculate the dissociation constant K_D, the obtained data has to be

translated into terms of concentration of bound ([$VcSiaP_{complex}$]) and unbound protein ([$VcSiaP_{free}$]). A binding isotherm can then be plotted with the obtained concentrations, followed by a Langmuir fit of equation 2.1 to calculate K_D [150].

$$[VcSiaP_{complex}] = [VcSiaP_{tot}] + [Neu5Ac_{tot}] + K_D - \frac{\sqrt{([VcSiaP_{tot}] + [Neu5Ac_{tot}] + K_D)^2 - 4 \cdot [VcSiaP_{tot}] \cdot [Neu5Ac_{tot}]}}{2} \quad (2.1)$$

Integration of distance distributions

Assuming that the area of the distribution with peak at 43Å is directly proportional to the unbound VcSiaP concentration [$VcSiaP_{free}$], the area of the distribution with peak at 27 Å is then directly proportional to the bound VcSiaP concentration [$VcSiaP_{complex}$]. Summing up both concentrations gives the total concentration of used protein [$VcSiaP_{tot}$] = 25 µM.
To calculate the binding curve (see figure 2.15) the two distance distributions of each sample were integrated and converted to the respective concentrations. The calculated bound-to-unbound ratio can be found in table 2.3. Since the shape of the distance distribution differs with the choice of the regularization parameter α, another method (linear combination fitting) was used to quantify the obtained data.

Fitting of timetraces as linear combinations of the open and closed form

This method focuses on analyzing the recorded timetraces instead of the distance distributions which are dependent on the regularization parameter α (see section 1.6.6) with regard to their the shape and width. Even though the program DeerAnalysis chooses the parameter α based on an L-curve criterion, there still remains a possible influence of the procedure on the calculated integrals.
All timetraces which stem from recordings with $1 - 120$ µM Neu5Ac were considered to be linear combinations of the boundary timetraces. The boundary timetraces are the ones which were recorded with 0 µM Neu5Ac (open conformation) and with 600 µM Neu5Ac (closed conformation).
To be able to calculate the combinations, the modulation depths of each time trace first had to be set to 100%. Then, the intermediate timetraces were fitted using equation 2.2.

$$f_{\text{intermediate}} = a \cdot f_{\text{open}} + (1 - a) \cdot f_{\text{closed}} \tag{2.2}$$

The obtained values for the mixing parameter a are listed in table 2.3. The error from the fitting procedure is stated in parenthesis.
Figure 2.15 shows the superposition of the measured timetraces (black) and the fitted timetraces (red). As can be seen, they align perfectly.

Calculation of the dissociation constant K_D

When comparing the bound-to-unbound ratios from both methods in table 2.3, one can see that the difference between the values is marginal. In fact the values obtained by the integration method are within the error range which arises from the fitting procedure of the linear combination method.

Table 2.3. – Bound-to-unbound ratios obtained from the integration and fitting method.

$[\mathbf{Neu5Ac}]_{\mathbf{tot}}$ (µM)	**VcSiaP open/closed (%)**	
	Fit (a in %)	**Integration**
0	100/0 (n.d.)	100/0
1	95/5 (± 2.0)	96/4
10	61/39 (± 0.4)	59/41
24	24/76 (± 0.8)	22/78
120	12/88 (± 0.8)	10/90
600	0/100 (n.d.)	3/97

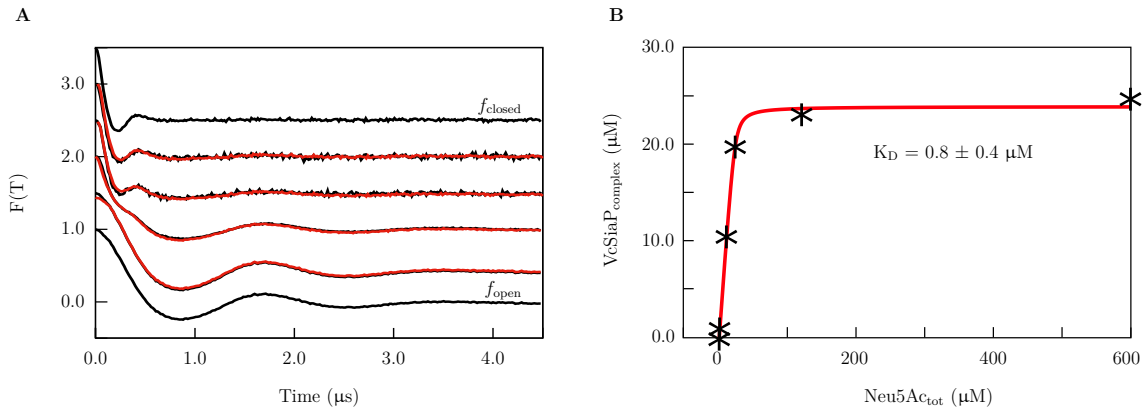


Figure 2.15. – Calculation of the K_D from PELDOR data. Panel A shows the two boundary timetraces which were used to calculate the mixing parameter a (f_{open} and f_{closed}). The fits (red) are superimposed with the measured timetraces (black). In Panel B the Langmuir fit of the PELDOR titration data is plotted. Average concentrations obtained from the integration method and the linear combination method were used.

Therefore, the bound-to-unbound ratios from table 2.3 were averaged, and then converted into the respective concentrations. Figure 2.15 shows the Langmuir isotherm one receives when plotting the deployed Neu5Ac concentration [$Neu5Ac_{tot}$] against the concentration of bound VcSiaP [$VcSiaP_{complex}$]. By fitting the data points to equation 2.1 a K_D of 0.8 ± 0.4 µM could be determined.
In this subsection the dissociation constant K_D of VcSiaP was acquired from PELDOR measurements. To test the validity of this approach, the K_D still has to be obtained by a standard method [151]. In the next section, the is K_D of VcSiaP is obatined from surface plasmon resonance (SPR) and compared to the current result.

2.1.7. Determination of the binding affinity of VcSiaP towards Neu5Ac

To determine the K_D of VcSiaP from SPR experiments, the protein had to be first biotinylated and then coupled to a streptavidin sensorchip surface. For this the mutant VcSiaP Q245C was chosen. In this construct the biotin is placed at the backbone of the protein, away from the binding cleft, to exclude possible interference with the substrate Neu5Ac.
The purification procedure starts with the same process as for the spin labeled double mutants in section 2.1.2: biotin was attached to the protein on a Ni^{2+} affinity column (just like MTSSL labeling) and then eluted with imidazole. Next, the protein was subjected to ion exchange chromatography, cleaved with TEV protease and finally purified via size exclusion chromatography. Subfigure 2.16 shows a SDS-PAGE of the final protein fractions from the Superdex 200 16/60 column. The red line indicates the collected fractions.

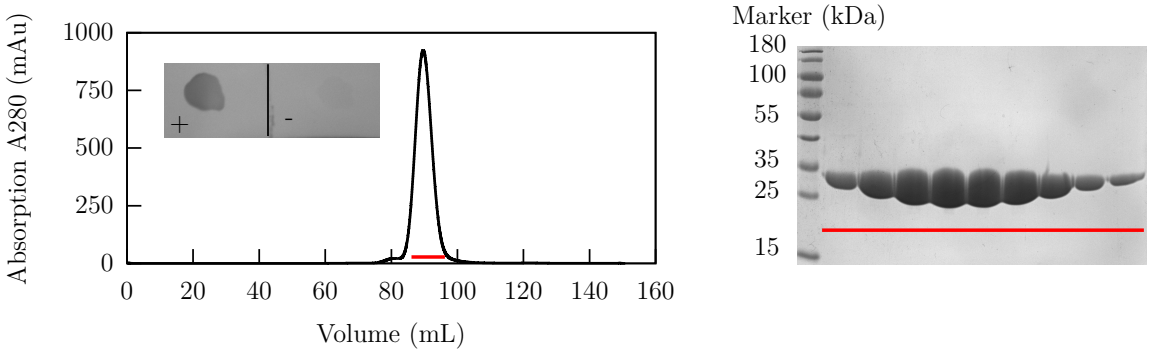


Figure 2.16. – Chromatogram of the size exclusion chromatography of VcSiaP Q245C. The respective SDS PAGE is displayed on the right side of the chromatogram. The collected fractions are indicated in red.

As can be seen in the image, the protein was obtained in high purity. To confirm the success of the biotinylation a dot blot with HRP-linked antibody S7075 (goat anti-biotin) was performed. The positive outcome can be seen in figure 2.16, where the dot blot is presented as inlet into the chromatogram.
A concentration series of Neu5Ac (5 nM−1 mM) was injected onto the sensorchip at continuous flow rate (30 µL/min). To remove any residual Neu5Ac, high salt injections were performed between the substrate injections. The salt was then washed off by buffer before the substrate injections continued. The exact protocol is described in section 4.4.5. Figure 2.17 shows the obtained sensogram. The binding curves were described by a Hill-fit which leads to a K_D of 0.473 ± 0.1 µM.

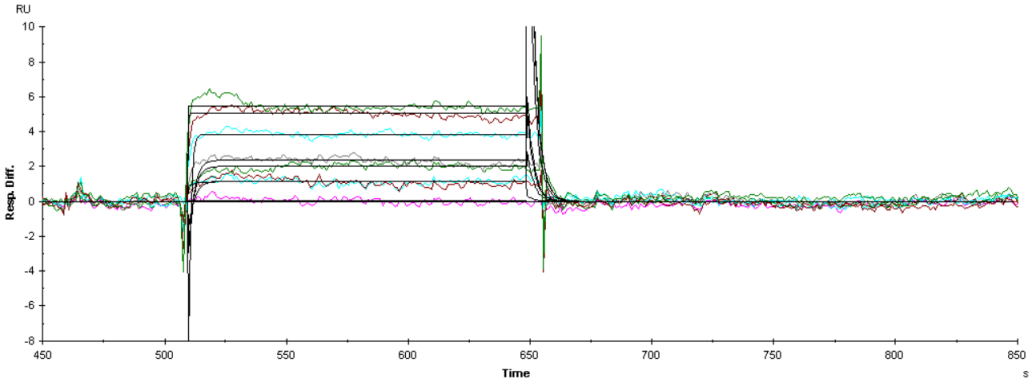


Figure 2.17. – SPR results for VcSiaP Q245Biotin. On top of the binding curves the Hill-fit is plotted.

It is noticeable that PELDOR described a slightly higher K_D (0.8 ± 0.4 µM) than SPR (0.473 ± 0.1 µM). The used concentration of VcSiaP in this is method should have been below the K_D to avoid substrate depletion. To obtain timetraces with good signal-to-noise ratio concentrations of 25 µM had to be used, which results in a very abrupt evolution of the binding isotherm. This impedes the fitting process and the accurate determination of K_D. Furthermore, the freezing process could have had an influence on the protein-substrate binding. Still, both values are reasonably close and within each others error ranges.

2.2. Investigation of the binding cleft

It still remains unclear how the bound substrate initiates the conformational change of VcSiaP, resulting in the closing of the protein. The three conserved amino acids R125, E184, H207 are proposed to play an important role in the transition of VcSiaP from open to closed state [99]. This subchapter focuses on those three amino acids which have been conserved through all known Neu5Ac TRAP transporters [152]. It aims at answering the question whether they are actually important in the switching mechanism and to which extent.
First, a structural analysis was performed to get first insights into the case. Then, they were systematically mutated, expressed and subjected to PELDOR measurements. At the end of this chapter I will give insight into X-ray structures that were obtained during this investigation.

2.2.1. Structural analysis

Figure 2.18 A shows the binding region of VcSiaP and the triad of the conserved amino acids R125, E184 and H207. Figure 2.18 B displays a superposition of open VcSiaP (blue) and the model of closed VcSiaP (orange). The model of closed VcSiaP was obtained by splitting unbound VcSiaP (PDB-ID:4mag) into the four rigid subdomains identified by the difference distance matrix in subchapter 2.1.1. Then, the regions were aligned with the respective regions of closed HiSiaP (PDB-ID:3B50).

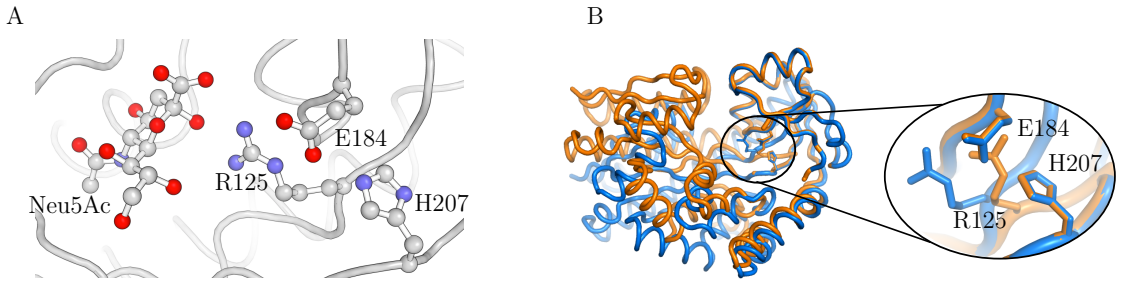


Figure 2.18. – The triad of conserved amino acids. Panel A shows the binding site of VcSiaP with the substrate Neu5Ac and the three amino acids R125, E184 and H207. Panel B shows a superposition of open VcSiaP and a model of the closed form of VcSiaP. While the residues E184 and H207 remain at their position, residue R125 undergoes a clearly visible shift.

Figure 2.18 B shows a close-up of the binding region. Here, it becomes clear that the amino acids E184 and H207 do not change their positions upon substrate binding, but R125 switches by 4 Å, making it the most important one of these three. This already

gives first insights into the significance of R125 with respect binding/closing. To further investigate the role of these amino acids, they were systematically mutated, expressed and subjected to PELDOR measurements.

2.2.2. Protein purification

Seven mutations were designed to test the role of the triad of amino acids in the opening and closing procedure of VcSiaP. Mutations of the chosen amino acid into alanine (R → A, E → A and H → A) are total knockouts of the function of the residue. More "conserved" options were tested also, which partially restored either the length or the charge of the respective residue (R → K, E → Q, E → D and H → Q).
These mutations were inserted into the VcSiaP Q54R1/L173R1 construct via QuickChange mutagenesis and transformed. This double spin labeled mutant was already chosen in other experiments for its clear switching from open to closed conformation.

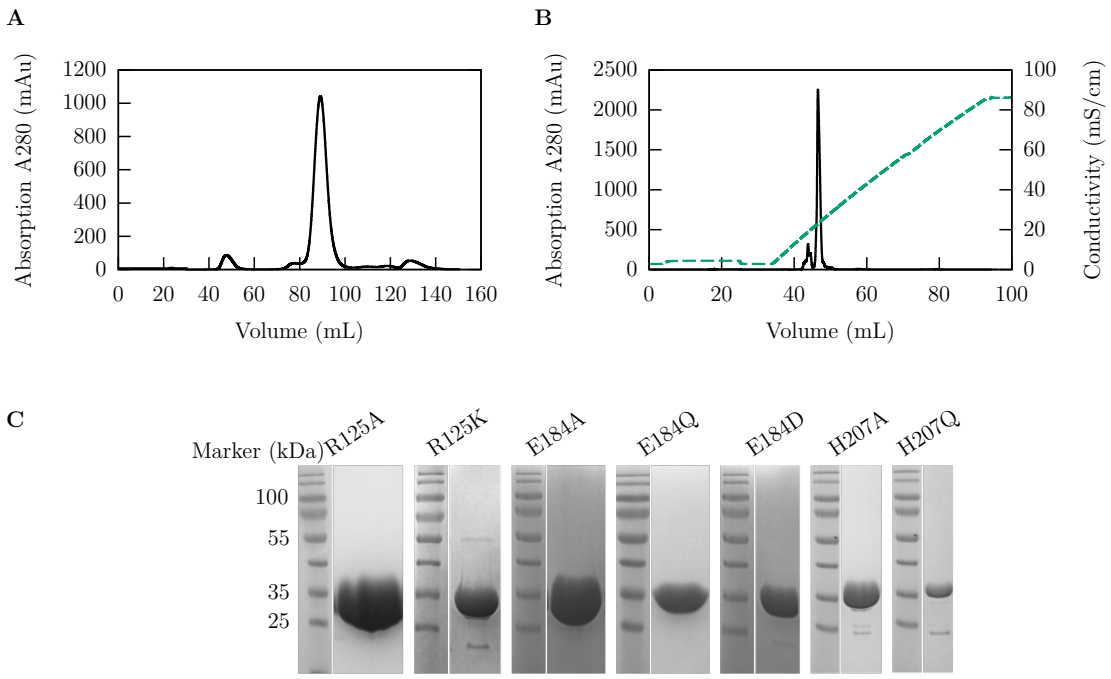


Figure 2.19. – Purification of the seven mutations which were introduced into the construct VcSiaP Q54R1/L173R1. Panel A shows the chromatogram of the size exlucion chromatography of the construct VcSiaP Q54R1/L173R1 R125A as example for all the other chromatograms. In panel B the ion exchange chromatogram is displayed. In panel C a montage of all SDS gels of the purified proteins is shown.

The purification starts with the same steps as for the spin labeled double mutants in section 2.1.2: the constructs were first washed on a Ni^{2+} affinity column and then labeled

with MTSSL in the elution step with imidazole. The elution was collected and subjected to size exclusion chromatography on a Superdex 200 16/60 column. Figure 2.19 shows the results of the purifications of all mutants. Panel A displays the chromatogram with the Superdex 200 16/60 column of the protein VcSiaP Q54R1/L173R1 R125A as example for all other chromatograms.
In the next step the protein was subjected to ion exchange chromatography with an ENrich Q 10/100 column. This was performed because in the expression process an *E. coli* protein emerged which could not be detached from the desired construct with regular size exclusion. The sequence analysis of the emerging protein can be found in appendix A.7.
To check if the labeling was performed correctly, *cw*-X-band spectra of the collected batches was recorded at room temperature (not shown). In the next section the results of the PELDOR measurements are presented.

2.2.3. PELDOR mutantional analysis

Figures 2.20 - 2.26 show the background subtracted time traces and the distance distributions obtained with DeerAnalysis. Raw data of the measurements can be found in appendix A.9. Each mutant was measured without Neu5Ac (indicated in figures as -Neu5Ac) and with 1 mM Neu5Ac (indicated in figures as +Neu5Ac). Further, the constructs VcSiaP Q54R1/L173R1 R125A and VcSiaP Q54R1/L173R1 R125A were also recorded with an excess amount of Neu5Ac (10 mM, indicated in figures as ++Neu5Ac). Each timetrace was recorded for 5 µs.
To quantify the influence of the mutations on the binding, the area under the distance distributions was integrated and the open/closed ratio determined. In subchapter 2.1.6 it was proven that the method of integrating the distance distributions and the method of fitting linear combinations of the timetraces are just as good, so I decided to calculate with the faster one. The ratio of open-to-closed conformation was determined for each sample and is stated in percent.
The results of this analysis are summarized by the end of this subchapter in figure 2.27 and table 2.4.

R125A

The background corrected timetraces for VcSiaP Q54R1/L173R1 R125A show two clear oscillations for all three measurements. For the sample without Neu5Ac (apo conformation) the distance distribution reaches a peak at 43 Å. Adding 1 mM of Neu5Ac (indicated as +NeuAc in figure 2.20) to the sample leads only to a hardly visible closed

conformation in the respective distance distribution. It is only at substrate concentrations of 10 mM (++NeuAc) that the closed conformation at 27 Å starts to become visible.

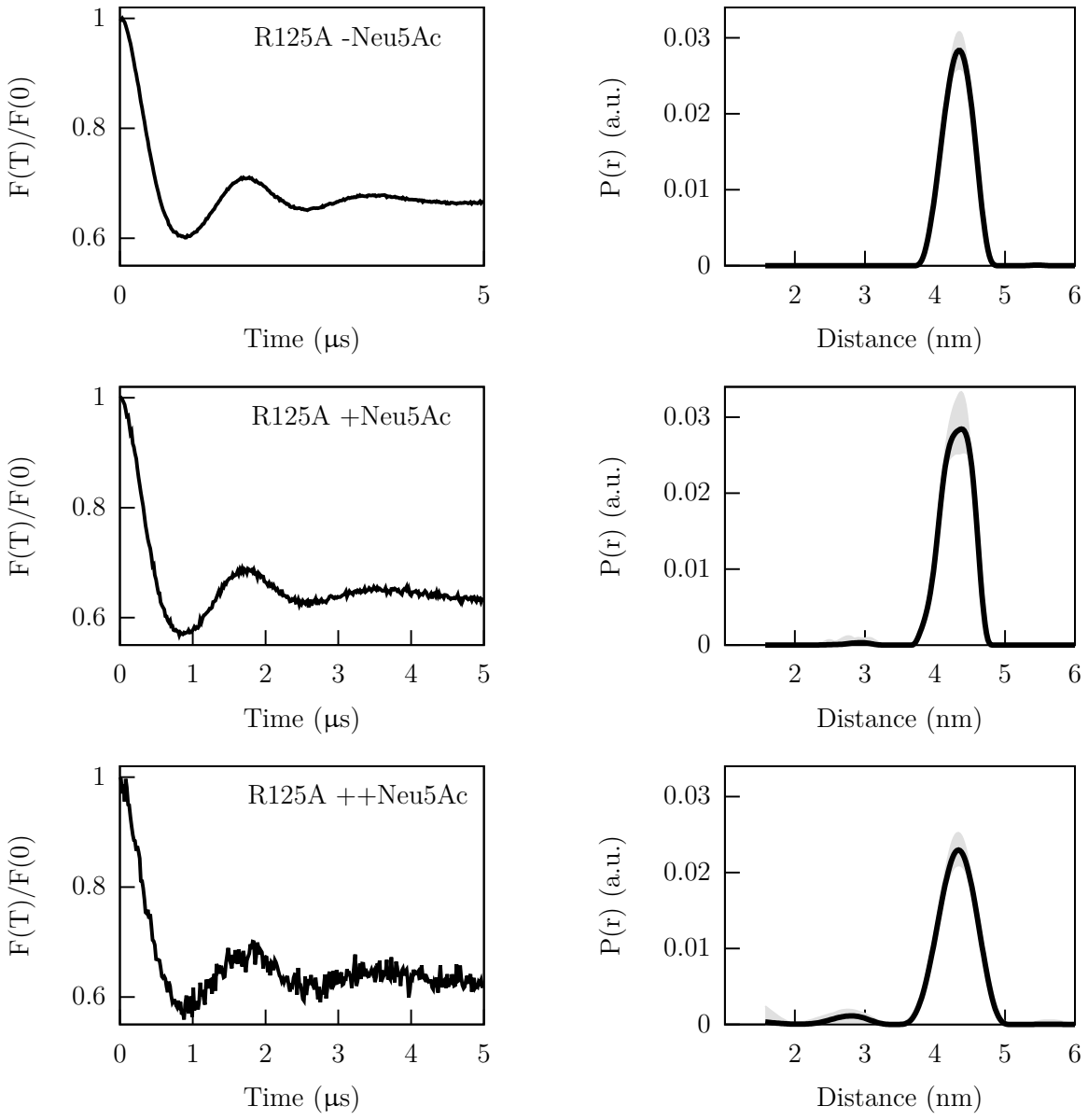


Figure 2.20. – PEDLOR timetraces and distance distributions of VcSiaP Q54R1/L173R1 R125A. The error estimated by DeerAnalysis is marked in gray.

R125K

The timetrace for VcSiaP Q54R1/L173R1 R125K without Neu5Ac shows again clearly two oscillations, which are slightly flattened with increasing amounts of Neu5Ac. Taking a look at the respective distance distributions the closed state emerges already at 1 mM Neu5Ac (open/closed ratio at 88/12). With 10 mM Neu5Ac the closed state reaches even 31% of the total amount of protein.

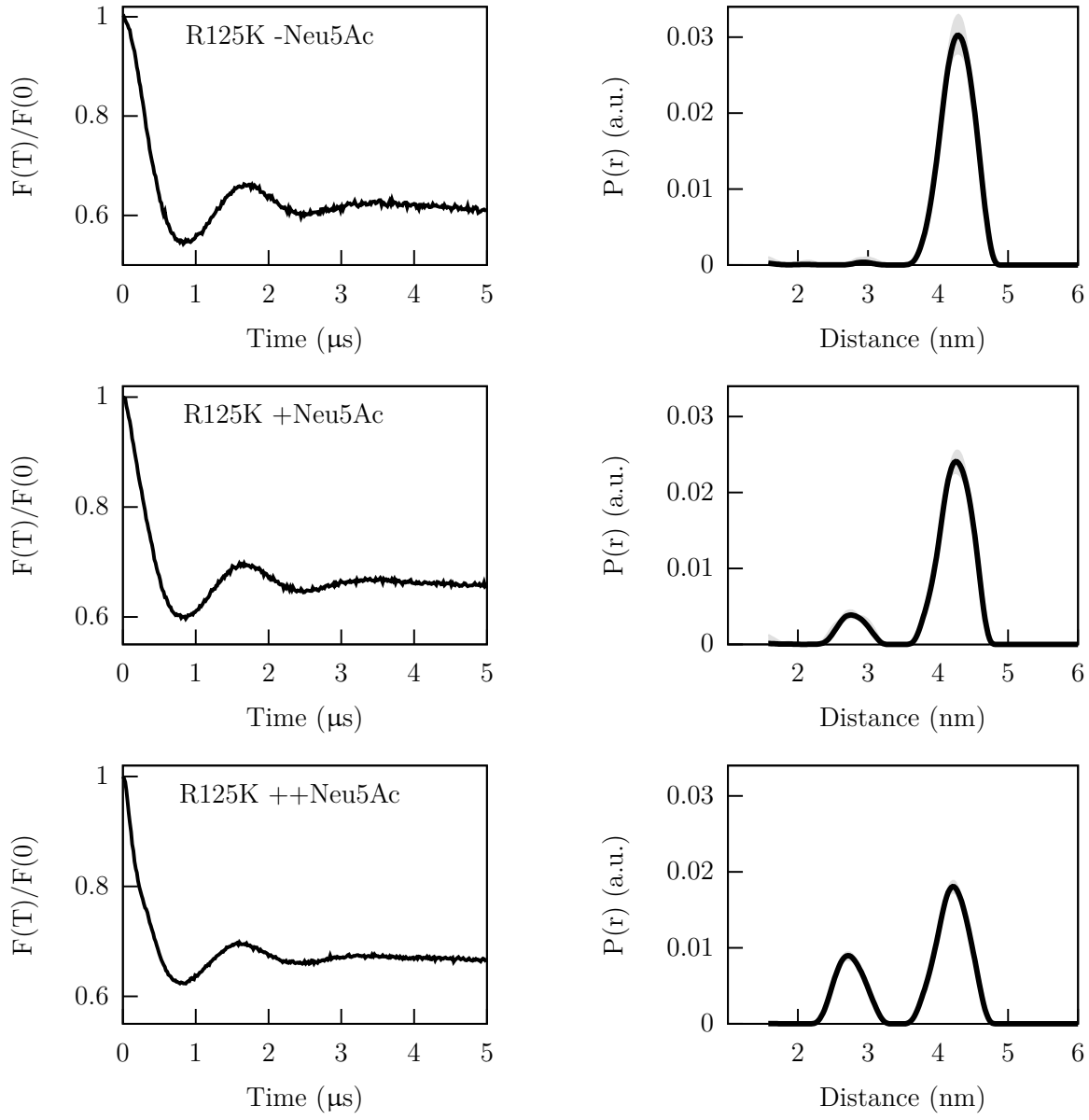


Figure 2.21. – PEDLOR timetraces and distance distributions of VcSiaP Q54R1/L173R1 R125K. The error estimated by DeerAnalysis is marked in gray.

E184A

The effect of the mutation is becoming more significant for the E184 residue. Whereas the timetrace of the sample without Neu5Ac has the same shape as the timetraces for the R125A and R125K mutants, the sample with Neu5Ac has a steeper fall at the beginning and clearly shows signs of two distances in the recording. The respective distance distributions confirm this initial assumption. The open/closed ration for the unbound state is 98/2. For the bound state it reaches upto 45/55.

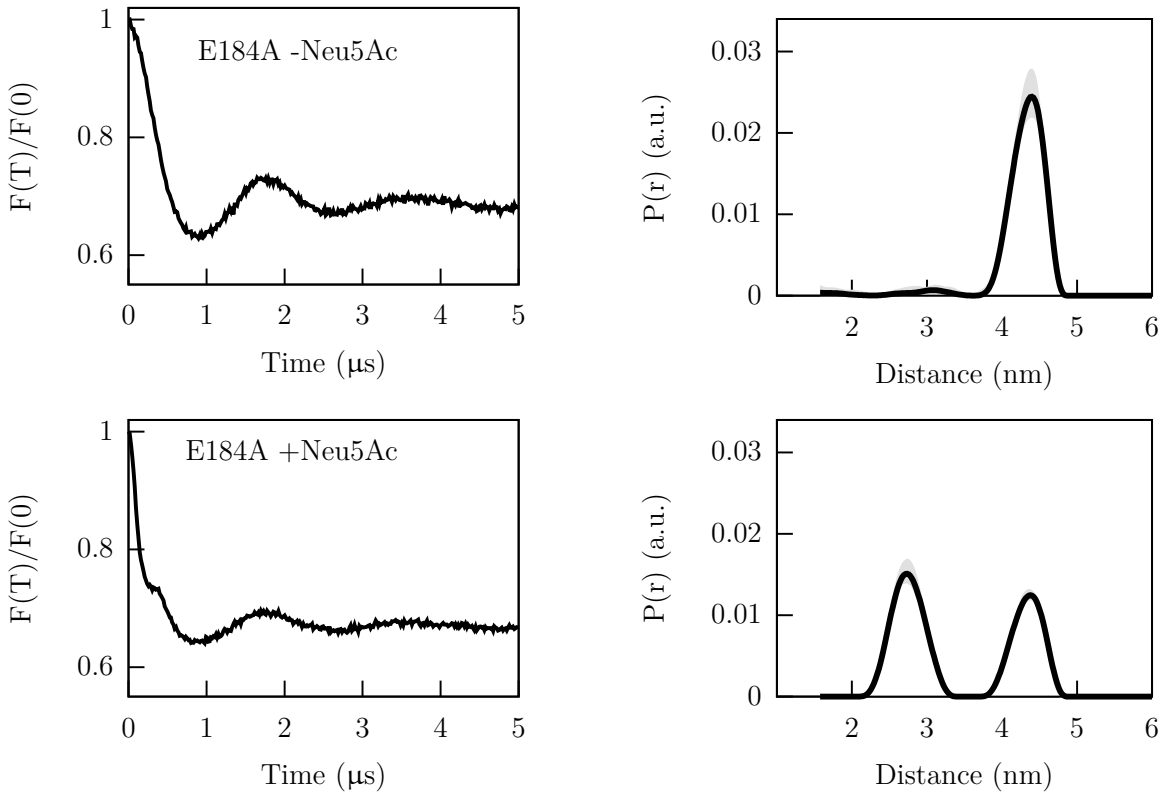


Figure 2.22. – PEDLOR timetraces and distance distributions of VcSiaP Q54R1/L173R1 E184A. The error estimated by DeerAnalysis is marked in gray.

E184D

This mutant has very little effect on the original switching mechanism of VcSiaP Q54R1/L173R1. The timetrace of the unbound state (-Neu5Ac) and the respective distance distribution reside mostly in the open state (open/closed ratio at 93/7). But with 1 mM Neu5Ac it shifts almost completely to the closed conformation (open/closed ratio at 9/91).

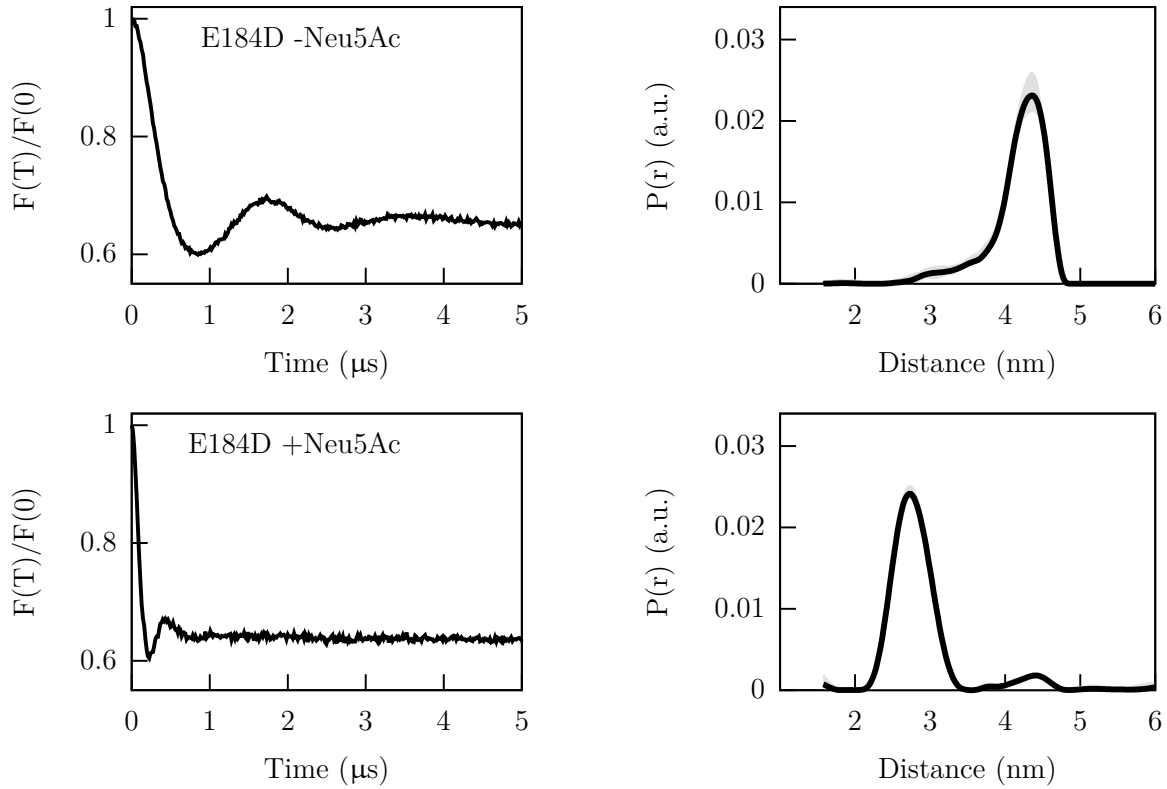


Figure 2.23. – PEDLOR timetraces and distance distributions of VcSiaP Q54R1/L173R1 E184D. The error estimated by DeerAnalysis is marked in gray.

E184Q

The course of the timetraces for the E184Q mutant resembles the recordings for E184A: Mostly, the open conformation is visible for the sample without Neu5Ac (open/closed ratio at 96/4). Then, however, an equilibrium of open and closed state is reached with 1 mM Neu5Ac (open/closed ratio at 39/61).

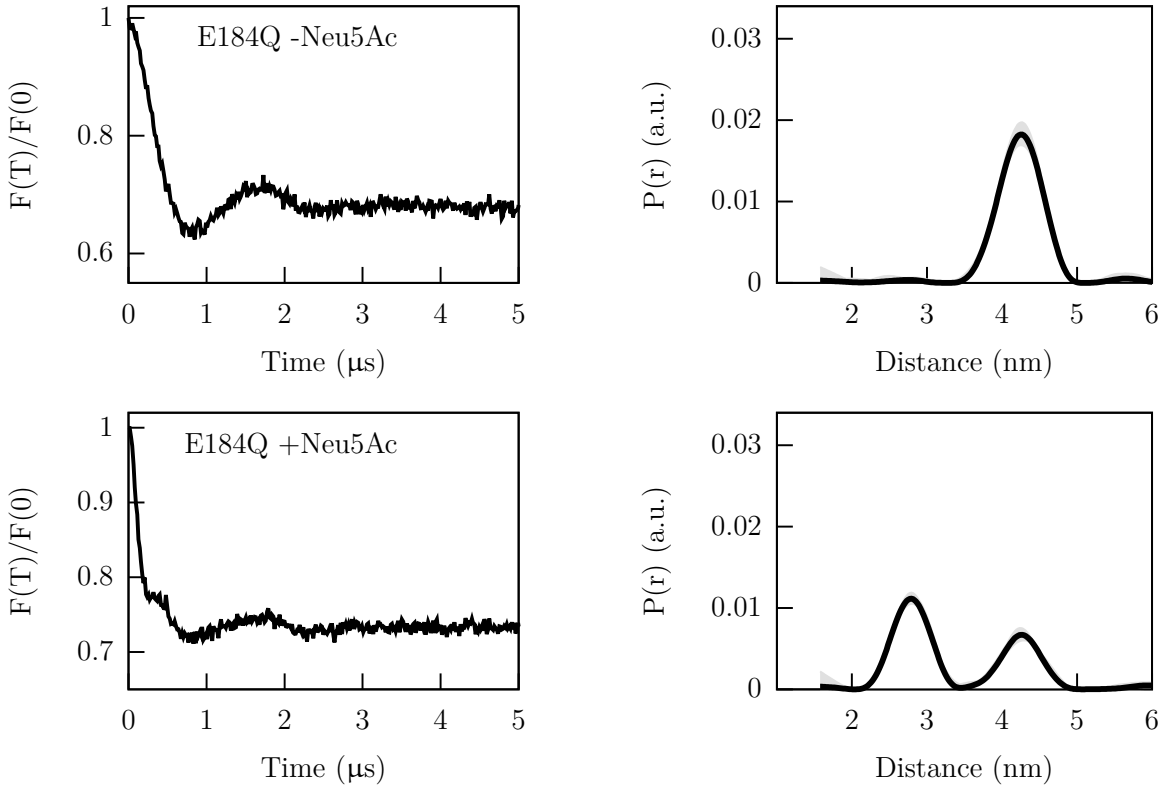


Figure 2.24. – PEDLOR timetraces and distance distributions of VcSiaP Q54R1/L173R1 E184Q. The error estimated by DeerAnalysis is marked in gray.

H207A

Just as for the mutant E184D the switching mechanism is hardly affected by this knockout mutant. The timetrace for the sample without Neu5Ac shows two clear oscillations, which is flattened out very steeply for the sample with Neu5Ac. The analysis of the two respective distance distributions reveals an open/closed ratio of 99/1 for the sample without substrate and a ratio of 7/93 for the sample with substrate.

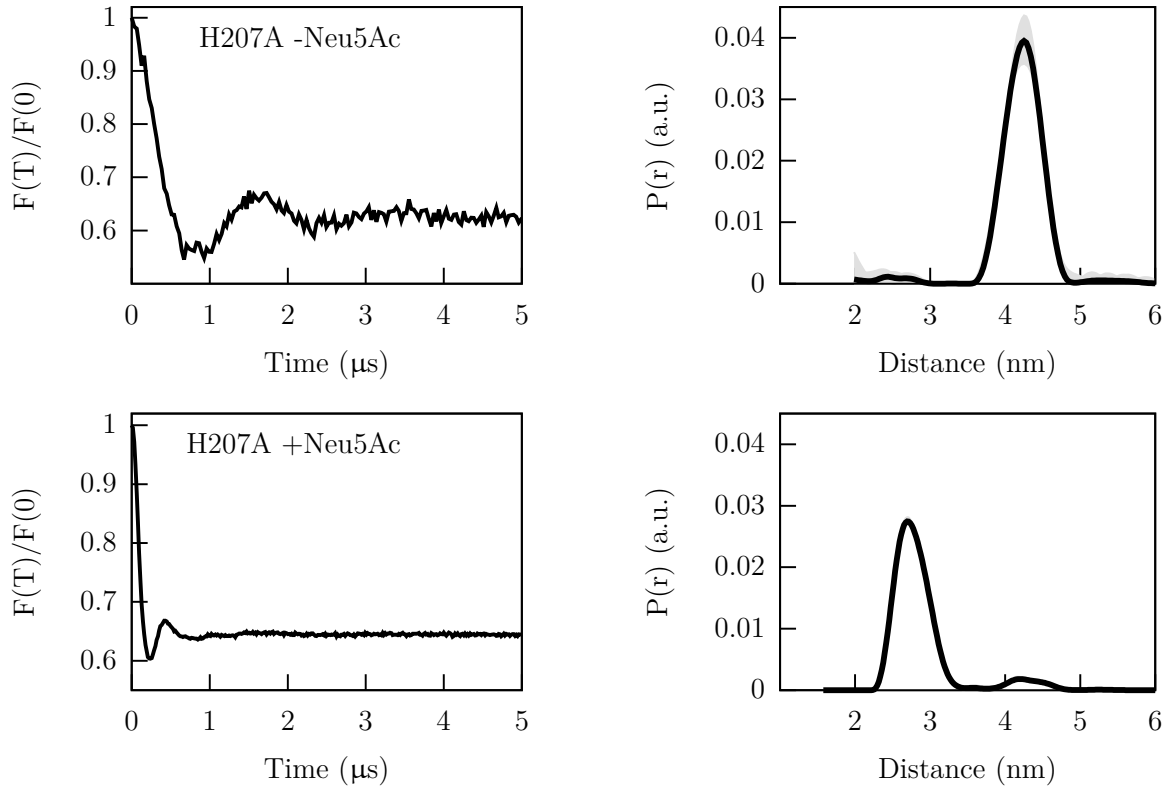


Figure 2.25. – PEDLOR timetraces and distance distributions of VcSiaP Q54R1/L173R1 H207A. The error estimated by DeerAnalysis is marked in gray.

H207Q

For this mutant the measurement reveals an open /closed ratio of 97/3 for the sample without Neu5Ac. The sample with 1 mM substrate leads to an open/closed conformation of 7/93.

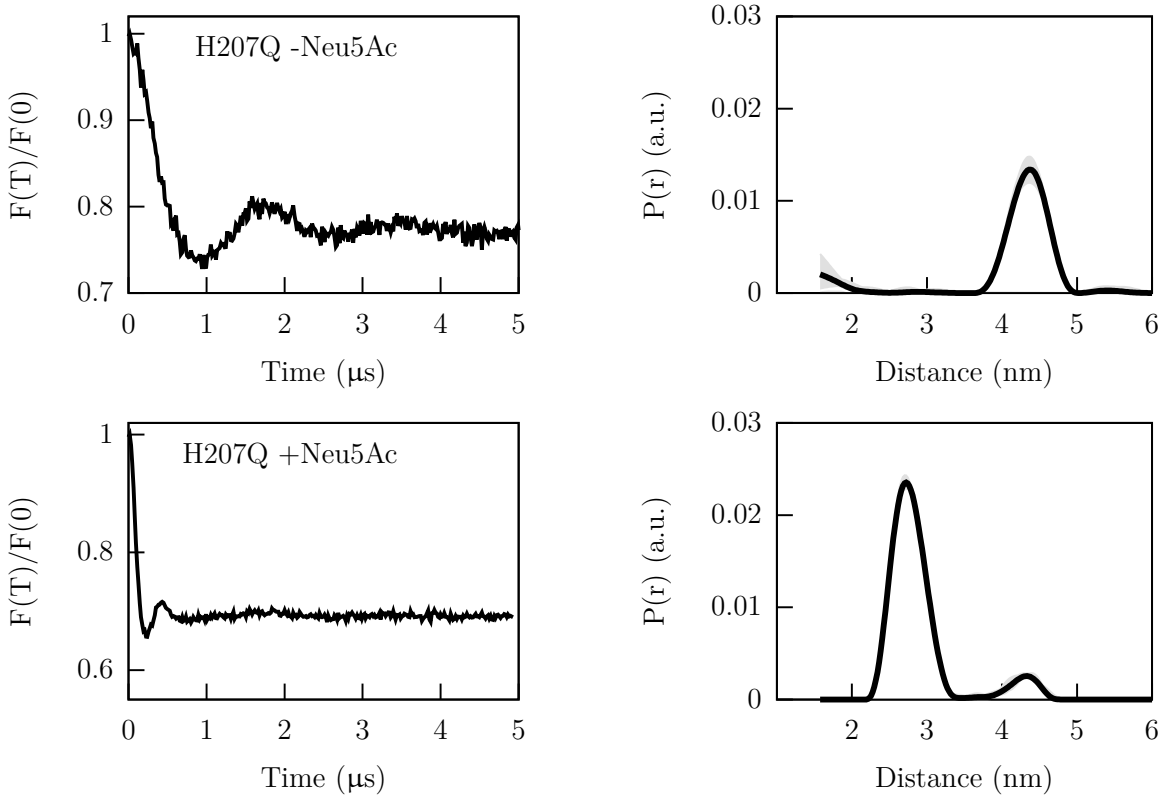


Figure 2.26. – PEDLOR timetraces and distance distributions of VcSiaP Q54R1/L173R1 H207Q. The error estimated by DeerAnalysis is marked in gray.

Summary

Figure 2.27 visualizes the results of this subchapter. Mutations at residue R125 had the greatest impact on the switching mechanism of VcSiaP Q54R1/L173R1. In particular R125A nearly eliminates the switching completely. Even high concentrations of Neu5Ac (10 mM) led only to a closed fraction of 4%. The switching is partially restored when alanine is substituted with lysine, but the contingent of closed conformation still only accounts for 31% with 10 mM Neu5Ac.
Mutating E184 into alanine leads to bigger fraction of closed conformation upon substrate binding: The open/closed conformation keep balance at a ratio of 45/55. But once the length of the original residue (A→Q) or the charge (A→D) is restored the impairment of the switch decreases.
Both mutations at residue H207 had very little effect on the switching. In both cases the open/closed ratio upon substrate binding lies at 7/93.

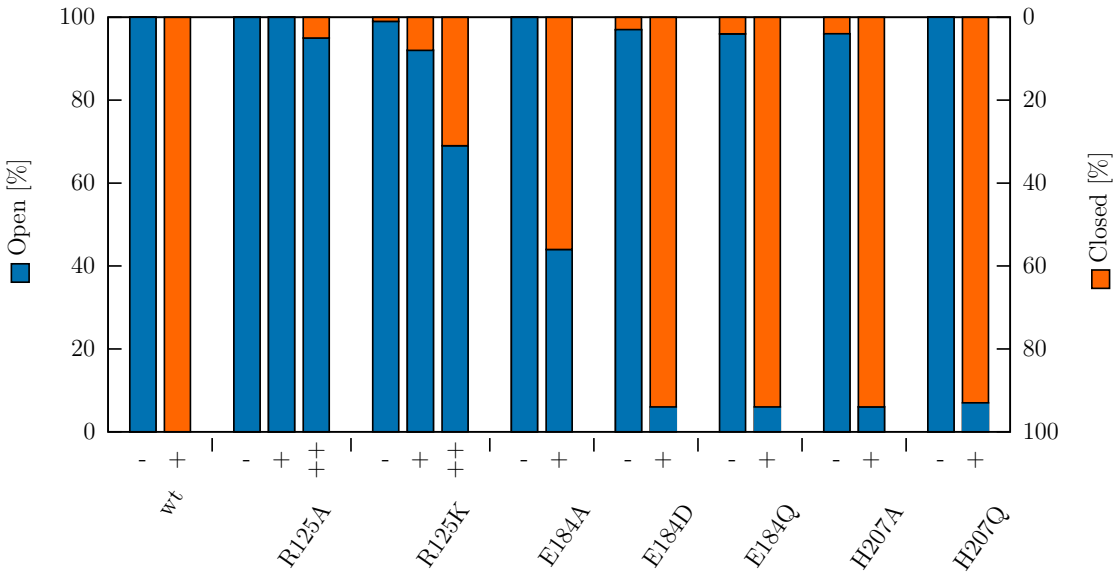


Figure 2.27. – Conformational changes of the seven investigated mutants upon binding of the substrate Neu5Ac. The blue bars represent the open configuration of the protein whereas the orange bars stand for the closed conformation. The amount of substrate used is indicated by - (0 mM), + (1 mM) and ++ (10 mM). The percentages of the conformation were determined by PELDOR spectroscopy.

Table 2.4 compiles the mutants, used substrate concentrations and the open/closed ratios of the proteins in percent.

Table 2.4. – Compilation of the results of the PELDOR mutational analysis of the three residues R125, E184 and H207.

Mutant	Neu5Ac (mM)	Unbound/Bound (%)
R125A	0	100/0
	1	99/1
	10	96/4
R125K	0	100/0
	1	88/12
	10	69/31
E184D	0	93/7
	1	9/91
E184Q	0	96/4
	1	39/61
E184A	0	98/2
	1	45/55
H207A	0	99/1
	1	7/93
H207Q	0	79/3
	1	7/93

2.2.4. Crystallization of VcSiaP Q54R1/L173R1 R125A

To ensure that the introduced mutations did not change the overall structure of the protein, efforts were made to crystallize the double spin labeled mutants of the binding cleft. First hits which looked promising were primarily obtained in the conditions A7 and D7 of the commercially available JCSG+ screen. These conditions were then optimized in a statistical manner, using the in-house software Xtals and the pipetting robot PrimaXtallo.

The condition 0.2 M lithium sulfate, 0.1 M Tris pH 8.5, 1.26 M ammonium sulfate gave birth to isometric crystals of around 0.1 mm and a fairly big one with a diameter of about 0.3-0.4 mm. These were harvested and cryo-protected with a mixture of 35% glycerol and reservoir solution.

Diffraction data sets were collected at the Helmholtz Zentrum in Berlin at the beamline Bessy BL 14.2. of the electron storage ring Bessy II. Two evaluable data sets were obtained during the measurement. Both diffracted up to a resolution of 2.1 Å.

The data sets were processed and scaled in the space groups $P2_122_1$ and $P6_3$, respec-

tively. In both cases the structure was solved via molecular replacement. In this process, VcSiaP (PDB-ID:4mag) was used as a search model. The summary of the crystallographic data and refinement statistics are listed in the tables 2.5 and 2.6.
In the following subsections I will shed light on the specialties of the obtained structures in more detail and explain what can be learned and derived from them.

VcSiaP Q54R1/L173R1 R125A does not get changed by R1 label

This structure exhibits two proteins in the unit cell (chain A and chain B). In difference to the "freak-like" structure from section 2.1.3, both chains are not fused together. Models of the label R1 could be placed on both labeling sites (Q54 and L173). Refinement was performed until R/R_{free} converged to 22.2/25.9.
The distance between both labels was measured to be 42.7 Å in chain A and 40.9 Å in chain B. This result is in agreement with the PELDOR measurements of VcSiaP R125A Q54R1 L173R1 in chapter 2.2.3.
Taking a closer look at the chains, one can observe that chain A fits perfectly onto the wildtype structure of VcSiaP (see figure 2.28), whereas chain B is trapped in a slightly more closed conformation. For better comparison, the structures where superimposed on residue regions 1-100 of the wildtype structure of VcSiaP, as visualized in figure 2.28.

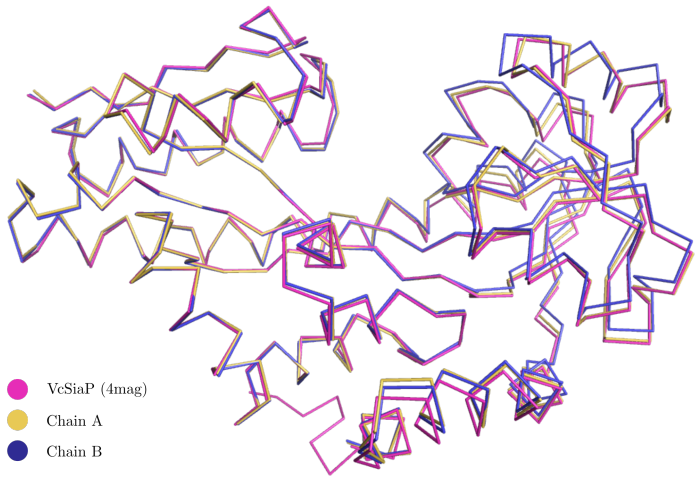


Figure 2.28. – Comparison of spin labeled VcSiaP R125A Q54R1/L173R1 with wildtype VcSiaP. The residues 1-100 of both chain A (yellow) and chain B (blue) are superimposed onto the open VcSiaP structure (magenta). Whereas chain A fits perfectly to the wildtype, chain B of the structure leans toward a slightly more closed conformation.

Figure 2.29 shows polar plots of the dihedral angles of the label. As mentioned earlier in section 2.1.3, this opportunity shall be taken to expand libraries of the dihedral angle of R1.

The angles χ_2 and χ_4 (purple) from the structure fit very neatly into the error range of the canonical angles (marked in grey). Deviations occur for the angles χ_1, χ_3 and χ_5.
A closer look at the electron density (see appendix A.10) reveals that these outliers mostly occur from the two labels A 173R1 and B 173R1. Here, the quality of the electron density is only mediocre. The other two labels, namely A 54 and B 54 are described by an electron density of excellent quality.
For B 54R1 all angles fit to the canonical angles of the free label. For the label A R1 one outlier occurs at $\chi_5 = 11°$. After all, the interaction of the bound spin-label with its protein environment can give rise to angles which differ from those of the free label. Furthermore, outliers for χ_5 have been observed in a crystal structures of azurin T30R1 (crystal forms I and II, PDB-IDs: 5I26, 5I28) [148].

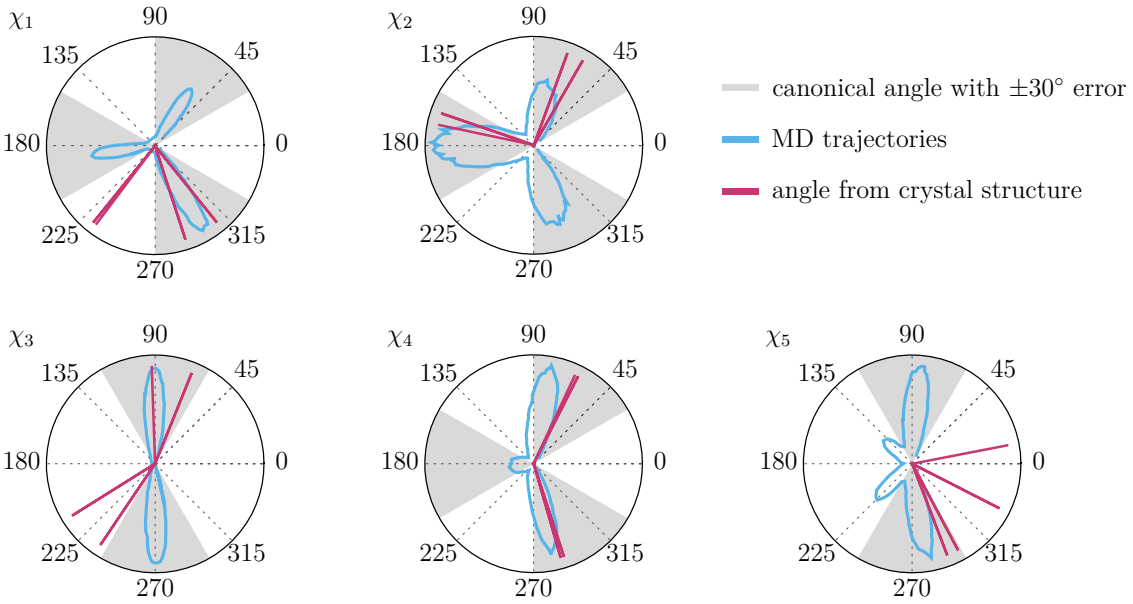


Figure 2.29. – Distribution of dihedral angles of the observed R1 side chains in VcSiaP Q54R1/L173C R125A. The grey sections mark the canonical angles of the label with a ±30° confidence interval. The blue lines are from a computationally produced library for the free spin label.

In summary, neither the inserted mutation R125A nor the attachment of the two R1 labels negatively interfered with the overall structure of VcSiaP.

Table 2.5. – Data collection and refinement statistics of VcSiaP Q54R1/L173R1 R125A.

Parameter	**Value**
Mutant	VcSiaP Q54R1/L173R1 R125A
PDB ID	5LTC
Wavelength	0.89429
Resolution range	45.29 - 2.101 (2.176 - 2.101)
Space group	$P2_122_1$
Unit cell	72.2697 78.1 116.24 90 90 90
Total reflections	73484 (6631)
Unique reflections	37987 (3606)
Multiplicity	1.9 (1.8)
Completeness (%)	97 (94)
Mean I/sigma(I)	7.25 (1.18)
Wilson B-factor	40.2
R-merge	0.043 (0.51)
R-meas	0.061 (0.73)
R-work	0.222 (0.357)
R-free	0.259 (0.359)
RMS(bonds)	0.005
RMS(angles)	0.78
Ramachandran favored (%)	98
Ramachandran allowed (%)	1.5
Ramachandran outliers (%)	0
Rotamer outliers (%)	0.78
Clashscore	7.60

VcSiaP Q54R1/L173R1 R125A binds to an artificial peptide

After solving the structure with molecular replacement, positive difference electron density was not only found at the two expected labeling sites Q54R1 and L173R1. The density map revealed furthermore strong difference electron density in the binding cleft of VcSiaP. The quality of the data was enough to presume it was a polypeptide bound to the substrate binding protein. But where did it come from? Certainly not from the crystallization conditions, since they lacked large molecules.
During refinement of the structure this question was answered: A part of the His_6 tag of the neighboring molecule (DYDIPTTENLYFGGAMG) had inserted itself into the binding pocket. Thus, two molecules form a yin and yang like dimer in the crystal.
Figure 2.30 displays the crystal structure of this dimer (panel A). Panel B shows the part of the His_6 tag, which sits in the binding cleft.
The N-terminal part of the His-tag (MSYYHHHHHH) is not visible in electron density and therefore assumed to be disordered. Refinement was performed until R/R_{free} converged to 0.19/0.23. Crystallographic details on data collection and refinement are summarized in table 2.6.
The most striking thing about this structure is that it is the first structure which shows/hints to the possibility of binding peptides by the SBP.

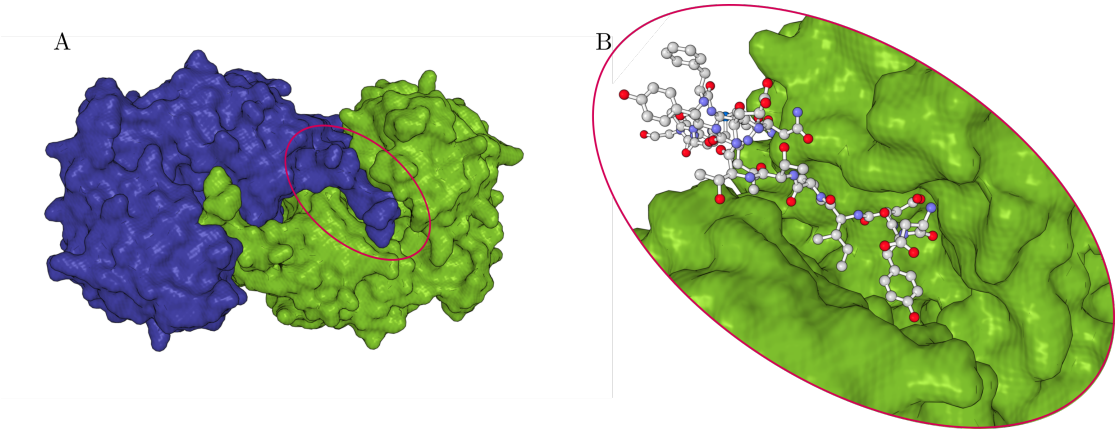


Figure 2.30. – Crystal structure of the VcSiaP interlocked dimer. A) Two molecules form a yin and yang like dimer in the crystal. B) Peptide sitting in the binding cleft.

In the following subsection the structure is compared to the open structure of VcSiaP to better classify and understand its conformation.

Table 2.6. – Data collection and refinement statistics of VcSiaP Q54R1/L173R1 R125A bound to an artificial peptide.

Parameter	**Value**
Mutant	VcSiaP Q54R1/L173R1 R125A
PDB ID	7A5C
Wavelength	0.89429
Resolution range	51.4-2.101 (2.176-2.101)
Space group	$P6_3$
Unit cell	118.70 118.70 113.75 90 90 120
Total reflections	1492314(134483)
Unique reflections	52966 (5254)
Multiplicity	10.6 (9.3)
Completeness (%)	100.0 (99.8)
Mean I/sigma(I)	67.6 (2.6)
Wilson B-factor	48.2
R-merge	0.094 (2.02)
R-work	0.1990 (0.3368)
R-free	0.2319 (0.3338)
RMS(bonds)	0.003
RMS(angles)	0.65
Ramachandran favored (%)	99
Ramachandran allowed (%)	1
Ramachandran outliers (%)	0
Rotamer outliers (%)	1.1
Clashscore	1.33

Peptide-bound structure is a trapped intermediate

To study the conformation of the peptide bound VcSiaP structure, a prediction of the distance distribution with mtsslWizard was performed. Figure 2.31 B shows the result

of this simulation: the distribution from the peptide-bound VcSiaP structure peaks at 3.4 nm (black curve).
When comparing this distribution to the measured PELDOR data on VcSiaP Q54R1 L173R1 one finds that the structure resides in an intermediate state between the open (purple envelope) and closed (mint envelope) form of the protein.
Figure 2.31 A shows the respective protein with the bound peptide (blue) and the predicted ensembles of the MTSSL label, which was used to predict the distance distribution.

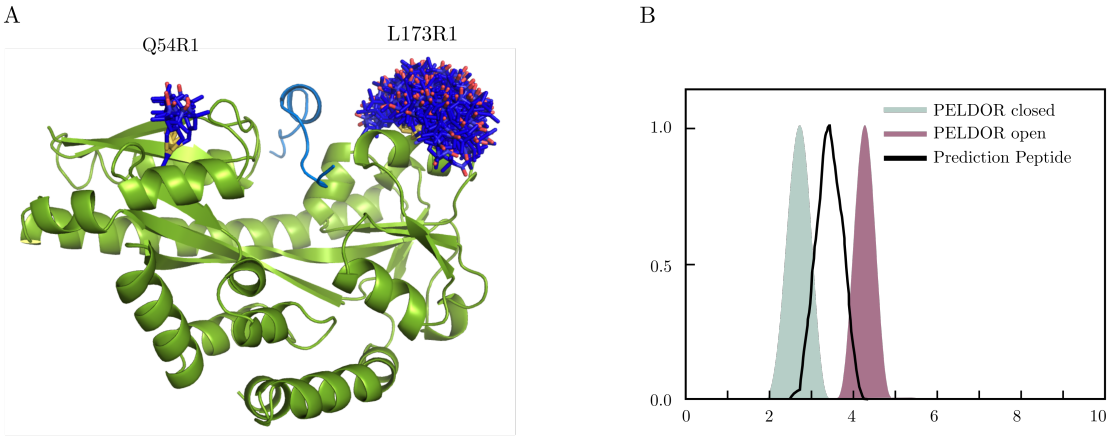


Figure 2.31. – Investigation of the conformation of the peptid-bound VcSiaP structure. A) The peptide bound strcuture of VcSiaP (green) with ensembles of the different label conformers at positions Q54 and L173. The peptide in the binding cleft is indicated in blue. B) The receptive predicted distance distribution (black), in comparison with measured PELDOR data (mint and purple).

Does binding to the peptide lead to the same movement as the one triggered by the original substrate Neu5Ac? To answer this question a difference distance matrix was calculated with mtsslWizard using the peptide-bound and open VcSiaP structure (4MAG). The top part of figure 2.32 shows the obtained matrix. The bottom part of the same figure depicts the difference distance matrix with open and closed HiSiaP, which was already presented in chapter 2.1.1.
Comparing both, it becomes clear that they have the same movement pattern. The same four rigid bodies could be identified as for the regular open to closed movement. The only differnce is that this happens on different length scales (upto 18 Å movement vs upto 5 Å movement).
This similarity in movement can be further visualized by aligning all the structures which were used in analyzing this movement. Figure 2.32 A shows the superpostion of VcSiaP + peptide (green), VcSiaP open (blue), HiSiaP + Neu5Ac (anthracite) and

HiSiaP open (grey).

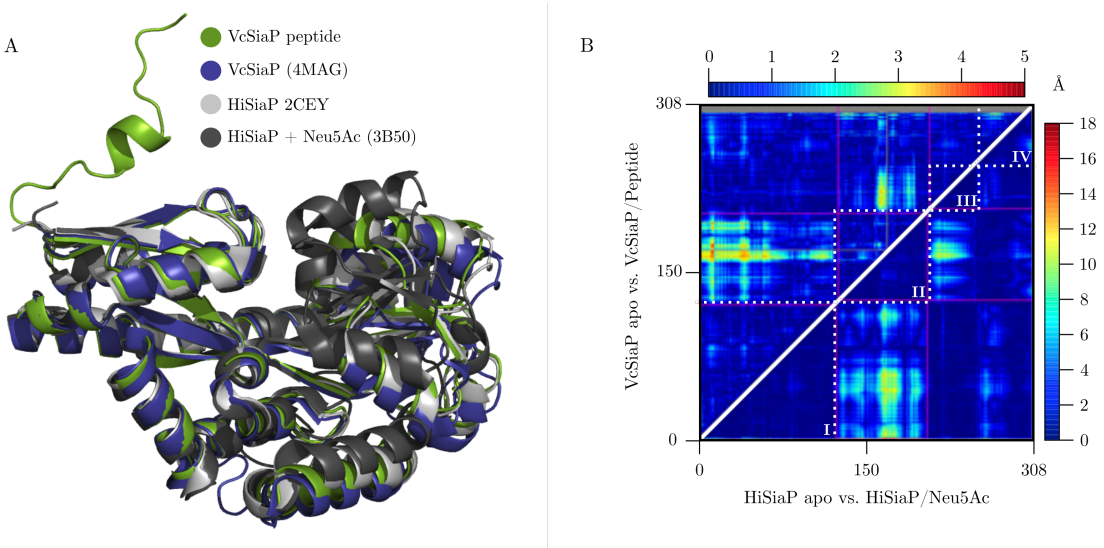


Figure 2.32. – Investigation of the semiclosed peptid-bound VcSiaP strcuture. A) Superposition of all protein structures used for this analysis. B) Difference distance map of VcSiaP from mtsslWizard. It was calculated with open and closed HiSiaP, as replacement for open and closed VcSiaP. The dashed boxes mark the rigid bodies of the protein.

A ligplot analysis reveals which interactions between the peptide and the residues of the binding cleft take place.

Interaction of the peptide with the binding pocket of the protein

The interaction of VcSiaP with the embedded peptide was studied with ligplot and is visualized in figure 2.33. The full length of the studied His_6 tag is also depicted in this picture.

The grey part of the sequence is disordered and therefore does not show up in the crystal structure. The red part sits directly in the binding cleft and the black part does not interact directly with important residues of the binding cleft.

The part of the peptide which sits in the cleft is involved in many interesting interactions.

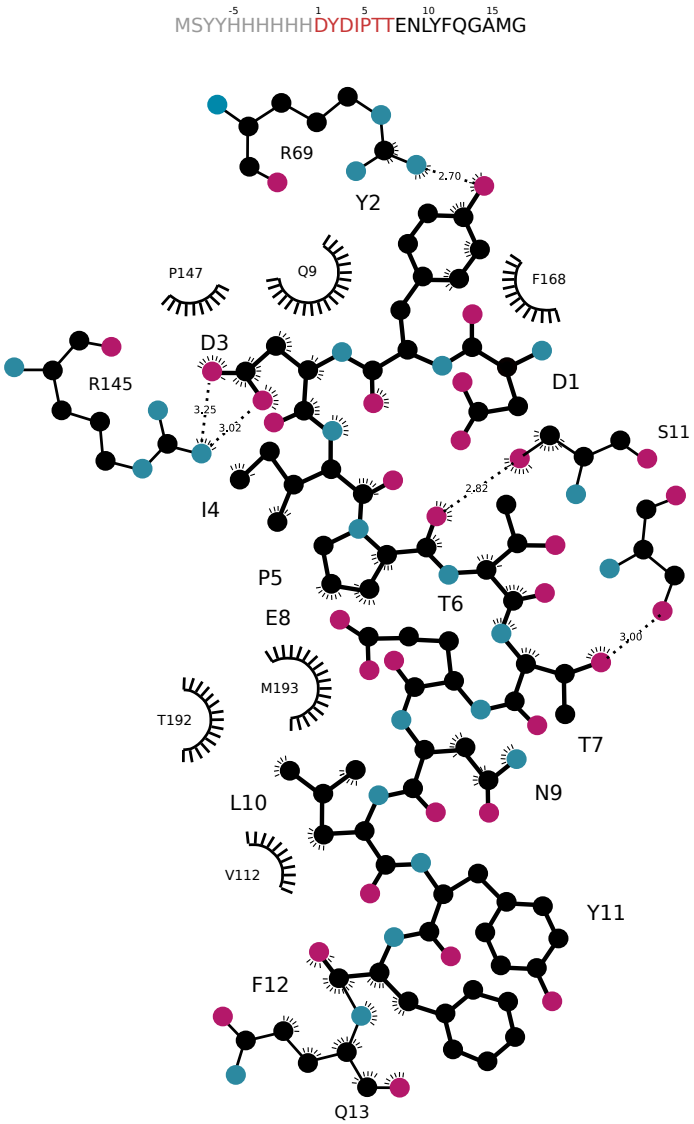


Figure 2.33. – Ligplot analysis of the substrate binding interactions between VcSiaP and the artifical peptide. Polar interactions are drawn as dashed lines, hydrophobic and Van-der-Waals interactions are represented by arcs. The distances are given in Å.

Most strikingly is that Asp3 (D3) forms a salt bridge with the crucial amino acid R145. R145 usually forms a salt bridge with the carboxylate group of TRAP transporter substrates. One must not forget that this is the crystal structure with the knockout mutant R125A. The mutant was modeled back into arginine in order to estimate this interaction better. This shows that the mutation does not interfere with this interactions. Arginine would further stabilize the interaction of D3 with R145.
There are several other interactions showing up: between Tyr2 (Y) of the peptide and and Arg69 of the P-domain, Thr7 (T) of the peptide and S42 of VcSiaP as well as multiple hydrophobic interactions.
The dimer seems stable in crystal structure. However, attempts to detect the dimer in size exclusion chromatography experiments failed. Neither the wildtype tagged VcSiaP, nor VcSiaP R125A with tag showed signs of dimerization.

The discovery of this peptide-bound structure was a milestone and turning point within this thesis. It is the first structure to show that peptides can be bound in the substrate binding cleft of VcSiaP. Furthermore, it raised a plethora of other questions and inspired more experiments, to be discussed in the following subchapters:
Can VcSiaP also bind to other peptides? Can it interact with the loops of the QM domains? Can a peptide-based inhibitor be found? How strong is the interaction between the peptide and VcSiaP?

2.3. Growth experiments

The discovery of the peptide-bound structure of VcSiaP inspired a new idea in the search for relevant interactions between the P-domain and the transmembrane parts of the transporter. The peptide sitting in the binding pocket of the protein is reminiscent of the so called "scoop-loop" of the maltose ABC transporter $MalEFGK_2$ [153]. This transporter consists of the maltose binding protein (MBP) MalE which is located in the periplasm and two membrane proteins MalF and MalG that form the pore for transport. MalK is in charge of hydrolysing ATP. Figure 2.34 A shows the interplay of MalE (magenta), MalF (yellow) and MalG (blue).
In the transporting cycle of $MalEFGK_2$ the periplasmatic loop P3 of MalG enters the substrate binding pocket of MalE. The P3 loop then supposedly "scoops" the substrate maltose out of the binding pocket. Figure 2.34 B presents how the P3 loop (yellow) fits perfectly in the binding pocket of MalE (magenta).

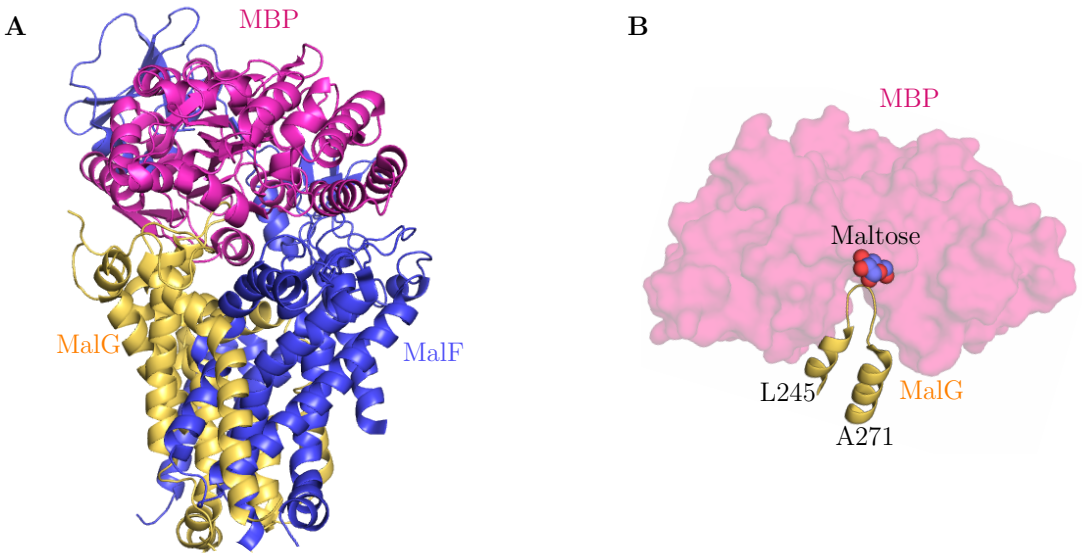


Figure 2.34. – The maltose ABC transporter $MalEFGK_2$. Panel **A** shows the MBP docked to the transmembrane part MalG and MalF of the ABC transporter (PDB-ID: 2R6G [95]). Panel **B** shows the scoop loop sitting in the MBP (PBD-ID: 1JW5 [153]).

This raised the following question for the transporter VcSiaPQM: is Neu5Ac possibly "scooped" from VcSiaP by a loop from VcSiaQ or VcSiaM? The structure in section 2.2 revealed a prominent interaction between Asp of the peptide and R145 of VcSiaP via a salt bridge. The same interaction takes place between R145 and the carboxylate group of the substrate upon binding.

To find possibly relevant amino acids of QM, which can also form a salt bridge with R145, alignments of homologous QM domains were checked for conserved glutamates and asparatates. To be able to interact with the binding cleft of the protein, these amino acids also have to be located in the periplasmic loops of the transporter. With the topology model from MULLIGAN [101], 7 amino acids were found to be conserved and periplasmic.
By replacing the original amino acid with an alanine (knockout) the selection was to be tested for its effect on the transporting process.
An *in vivo* transport assay consisting of a plasmid carrying the VcSiaPQM transporter (pES85) transformed into an *E. coli* ΔnanT strain (Sevy3) as "indicator strain" was used. Since the ΔnanT strain can not grow on Neu5Ac as sole carbon source its growth is a direct indicator for the functionality of the transporter. The selected mutations E47A, E48A, E47A/E48A, E119A, E237A, E293A and E420A were checked first on M9-agar plates and, additionally, their growth was monitored in M9-media supplemented with Neu5Ac or glucose (control) as sole carbon source.
Also, the peptide scan showed binding of the protein to loop 3, which was the only unmodified loop binding to VcSiaP. Therefore I included also the mutations F79E, A80E and F79E/A80E in the investigations. The peptide scan is explained in detail in the following subchapter 2.4.

2.3.1. First screening of mutants on M9 agar plates

Figure 2.35 shows the M9 plates on which the selected mutations were grown. Since pES85 is the plasmid with the unmodified full transporter, it functioned as positive control. Negative control was pwks30, which is the empty plasmid. The plates were prepared as described in chapter 4.3 and growth was checked after two days of incubation at 37 °C. Each plate was prepared three times to test reproducibility.
As expected the positive control grew on all plates and the negative control did not grow on any plate. Further growth could be observed with the mutants E48A, E47A/E48A, E74A, E119A, E293A, E420A and F79E. The mutants E47A, E237A, F79E, A80E did not grow on the plates.
Conspicuous is the growth of mutant E47A/E48A even though mutant E47A clearly does not show signs of growth. A possible reason for this could be that the plated cells started metabolizing available carbon sources such as IPTG or the polymerized agar.
Also noticeable is the unequal growth of the mutants: E420A and F79E show less surface coverage compared to the other segments. This observation lead to the interpretation that the cells could possibly differ in their growth rates due to the disrupted

transporter.
Regarding the "outlier" growth of E47A and the uneven growth of the cells, more controllable experiments needed to be performed for a distinct result. The growth of all mutants was checked in liquid M9 media supplemented with Neu5AC as sole carbon source in the following section.

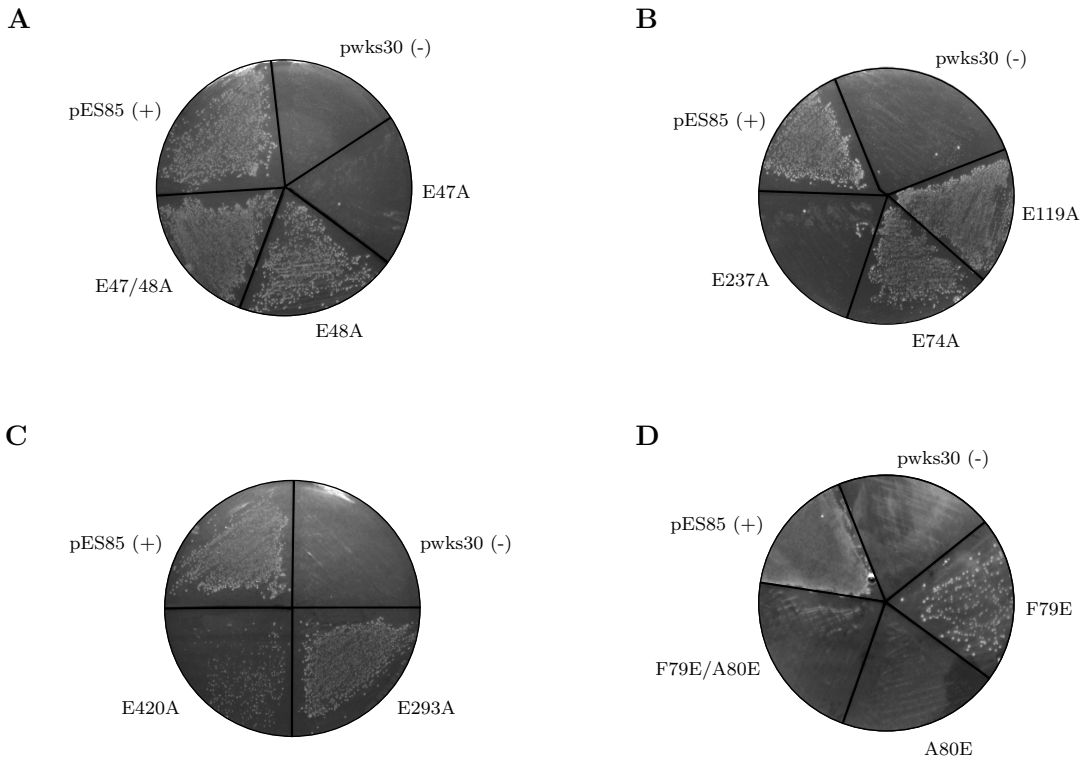


Figure 2.35. – Bacterial growth on M9 plates. The positive control pES85 (+) and negative control pwks30 (-) was applied on every plate. The white dots and areas indicate growth of the tested mutants.

2.3.2. Bacterial growth in liquid M9 media

The mutants to be investigated were prepared and grown in M9 media supplemented with Neu5Ac as sole carbon source as described in chapter 4.3. The cells were suspended in a 24-well plate as triplets and the OD(600) was measured every hour in order to monitor the growth. The growth was monitored for 10 hours starting from the moment IPTG was added to express the transporter. The OD(600) values of each triplet were averaged and the error was calculated as standard deviation.
The experiments which were individually performed on different days can be found in appendix figure A.23. The data of all individual growth experiments was merged and the result is depicted in figure 2.36.
In this series of experiments the mutants E47A, E47A/E48A, E237A, F79E, A80E and F79E/A80E did not grow. Notice that both mutants which were able to grow on M9 plates (E47A/E48A and F79E/A80E, see section 2.3.1) lost this feature in M9 media, which indicates that their growth on plates is an artifact of the plate growing method and handling. The mutants E48A, E119A, E74A E293A and E420A were all able to grow in M9 media. The results are summarized in table refcheck.

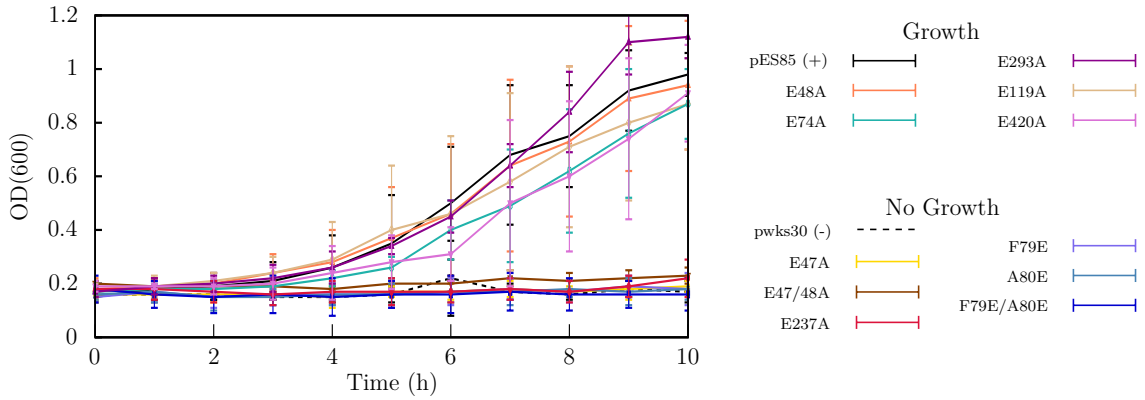


Figure 2.36. – Merged data of all M9 media bacterial growth subexperiments. The selected VcSiaQM mutations are listed in the figure. The media was supplemented with Neu5Ac as sole carbon source. The plasmid which functions as positive control (pES85) is indicated with a "+", the negative control (pwks30) is marked with a "-".

The uneven growth of the cells is reflected in the large errorbars. Basically, every measured point falls into the error range, making it impossible to draw conclusions about the growing rate of the "healthy" mutants.
To investigate whether the growing mutants can actually be distinguished by their

growing behavior, the experiment had to be repeated in a more controllable manner, which will be described in the following subsection.

Table 2.7. – Comparison of bacterial growth.

Mutant	Growth on M9 plate	Growth in M9 media
E47A	✗	✗
E48A	✓	✓
E47A/E48A	✓	✗
E74A	✓	✓
E119A	✓	✓
E237A	✓	✓
E293A	✗	✗
E420A	✓	✓
F79E	✗	✗
A80E	✗	✗
F79E/A80E	✓	✗

2.3.3. M9 media bacterial growth in baffled flasks

The mutants E48A, E119A, E74A E293A and E420A were grown in 25 mL M9 media using shikane flasks as described in chapter 4.3. The mutant E48A and the positive control pES85 were investigated on a separate day and were therefore monitored for 11 h. Mutants E74A, E119A, E293A, E420A were monitored for 14 h. Additionally after 24 h the OD(600) was measured for every mutant. The result is depicted in figure 2.37.
In this series of experiments each mutant was simultaneously grown in three shikane flasks. The obtained OD(600) was merged and the error calculated as standard deviation. Figure 2.37 summarizes the results.

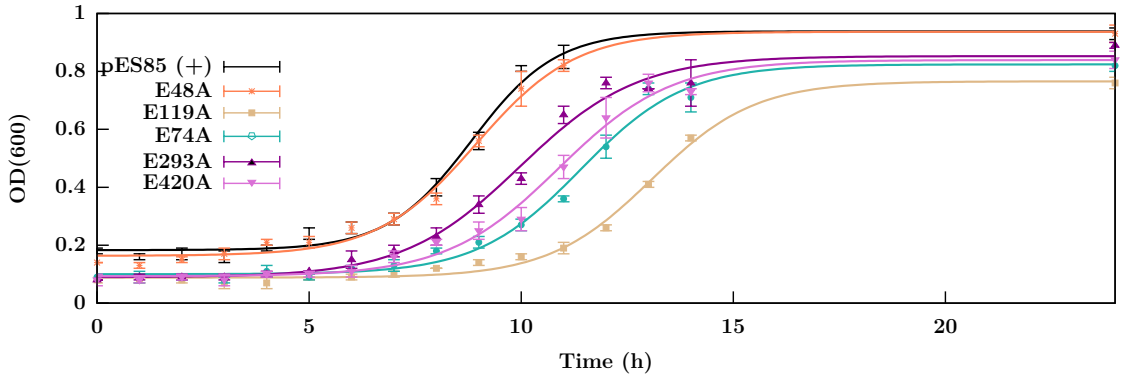


Figure 2.37. – Bacterial growth in liquid M9 media using shikane flasks. The selected VcSiaQM mutations are listed in the figure. The media was supplemented with Neu5Ac as sole carbon source.

The positive control pES85 and E48A had an OD(600) around 0.18 as starting point. After 24 h both reached an OD(600) of 0.9.
Mutants E74A, E119A, E293A, E420A started at an OD(600) around 0.1 and reached overall lower OD(600) after 24h compared to pES85 and E48A. This is due to the difference in starting points.
The data points were fitted with the Boltzmann sigmoidal fitting function [154] which is used to fit bacterial growth data [155, 156]

$$f(x) = a + \frac{b-a}{1+\exp(\frac{c-x}{d})} \tag{2.3}$$

The slope of the fit was used to calculate the maximal growth rate (OD(600)/h):

$$f'(c) = -\frac{a}{d \cdot 4} \tag{2.4}$$

It shall be noted that this performed fit had the objective of obtaining comparable

variables for this particular set of experiments. It is not supposed to be a model for the growth: modeling bacterial growth still remains a challenging task [157].
The error was calculated as Gaussian error propagation of the fitting function f(x)

$$\sqrt{\left(\frac{\Delta a}{4 \cdot d}\right)^2 + \left(\frac{a \cdot \Delta d}{4 \cdot d^2}\right)^2} \tag{2.5}$$

and the lag-phase was determined as axis intersection of the linear part of the fitted growth curve and the x-axis.
Table 2.8 summarizes the results of this analysis. The error is listed in parenthesis.

Table 2.8. – Growth parameters of bacterial growth in shikane flasks.

Mutant	**Growth Rate (OD(600)/h)**	**Lag-Phase (h)**
pes85	0.19 (12%)	6.84
E48A	0.17 (11%)	6.66
E74A	0.15 (12%)	8.99
E119A	0.14 (8%)	10.66
E293A	0.14 (12%)	7.24
E420	0.14 (12%)	8.22

The growth rates fall in a small range from $0.14 - 0.19\,\mathrm{OD(600)/h}$ with pES85 being the fastest to grow and E119A the slowest to grow mutant.
Boiling down all results of this subchapter one has to come to the conclusion that among the mutants which grow in liquid M9 media there are only small differences in growing rates.
But the biological significance of the mutations which did not show growth is still unclear: further experiments need to be performed and the folding of the transporter needs to be confirmed.

2.4. Peptide Scan

Chapter 2.2.4 presented the crystal strcuture of VcSiaP bound to the His_6-tag of a neighboring SBP. The discovery of the first TRAP SBP binding to a peptide discloses new and unworn research paths: Are there other peptides that act as potential binders? And among these binders, is there a possibility of increasing the interaction in order to inhibit the SBP?
But before submerging oneself into these questions, it needs to be checked if this new interaction is not simply the observation of a crystal packing artifact. This chapter therefore focuses on studying whether VcSiaP can also bind this "tag-peptide" and other peptides in solution. For this a cellulose membrane was used, onto which several different peptides are anchored. This so called "peptide-scan" was used in a western-blot style to identify interactions between the anchored peptides and a fluorescent-labeled binder-molecule.
Figure 2.38 shows a cartoon scheme of a peptide scan and illustrates the process of the experiment. VcSiaP was marked with the label Dylight800 and functioned as the binder molecule. The scans were incubated with P-domain for 1 h. Then, the excess P-domain was washed off with buffer. After the incubation period and the washing-step, the fluorescent signal displayed the hits and thereupon possible binders.

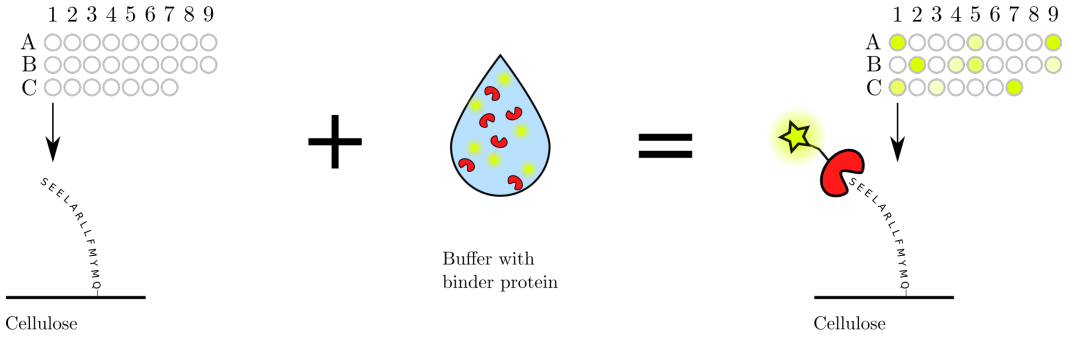


Figure 2.38. – Schematic representation of the peptide scan experiment. Selected peptides are spot-synthesized on a membrane in form of a grid. This membrane is incubated with buffer containing the fluorescence labeled binder protein and the excess of protein is then washed off. The remaining membrane contains peptides onto which binding of the protein was possible and which can be read out by an imaging system.

The next subchapter describes the purification the labeled binder-molecules and is followed by the presentation of the scan results. An explanation of which peptides were selected is given in the subchapter of the respective results.

2.4.1. Purification of VcSiaP Dylight800 and HiSiaP Dylight800

The constructs VcSiaP Q245C and HiSiaP K254C were chosen to be labeled with a fluorescent dye. In both mutants the introduced cysteine is placed at the backside of the protein, ensuring the labeling can not interfere with the binding site. The label Dylight800 contains a maleimide group which makes it sulfhydryl-reactive and therefore it can bind to cysteine.
Just like other VcSiaP constructs discussed previously in this chapter, VcSiaP Q245C and HiSiaP K254C are constructed into the pBadHisTev vector.
The purification starts with similar steps as the purification of the the spin labeled double mutants in section 2.1.2: The protein was bound to a Ni^{2+} affinity column, washed twice with 50 mL buffer and the cysteines were reduced with 1 mM TCEP. Differing from the protocol of the the spin labeled double mutants, the protein was now eluted with imidazole and 1 mM TCEP, ensuring reduced cysteines. Furthermore, the protein was subjected to ion exchange and size exclusion chromatography as described in subchapter 4.4.4 and the His-tag was cleaved with TEV protease.
For labeling, the TCEP was removed through size exclusion chromatography on a SD 75 10/300. Figure 2.39 A shows the chromatogram of 1 mg HiSiaP K254C prior to labeling. Panel B shows the collected protein fractions of VcSiaP Q245C and HiSiaP K254C. As can be seen, both proteins were obtained in high purity.

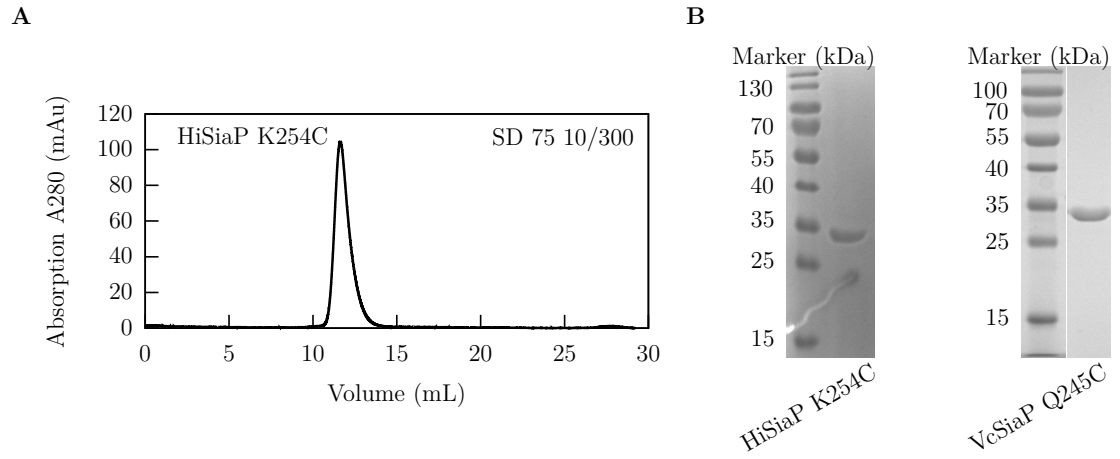


Figure 2.39. – Purification of HiSiaP K254C and VcSiaP Q245C. The proteins were subjected to size exclusion chromatography on a SD 75 10/300. Panel A shows the chromatogram of HiSiaP K254C. Panel B shows the SDS PAGE gels of both proteins. Directly after elution the fractions were collected and labeled with Dylight800.

2.4.2. On inhibitor loops

The scan analyzed in this subchapter was designed on the basis of the tag-peptide found in the VcSiaP binding cleft. The peptide integrates well into the substrate cleft, but the interaction still has to be extremely weak, since the tagged SBP was only observed as a monomer in size exclusion chromatography. Modifying the peptide seemed to be a reasonable way to increase the interaction. First the peptide sequence `DYDIPTTENLYFQ` was changed gradually into the peptide sequence of the loop which connects helix 1 and helix 2 of VcSiaQ (`SEELARLLFMYM`). This loop has two adjoining glutamates which are known to be conserved among TRAP transporters and therefore seem important for the transporting process. Also, the interaction between D3 of the tag-peptide and R145 of VcSiaP binding cleft indicated that conserved glutamates and/or asparatates in the QM-loops could possibly form a salt-bridge with R145 as well. This gradual change from one sequence into the other is depicted in red in figure 2.40, sequences A1-B1.

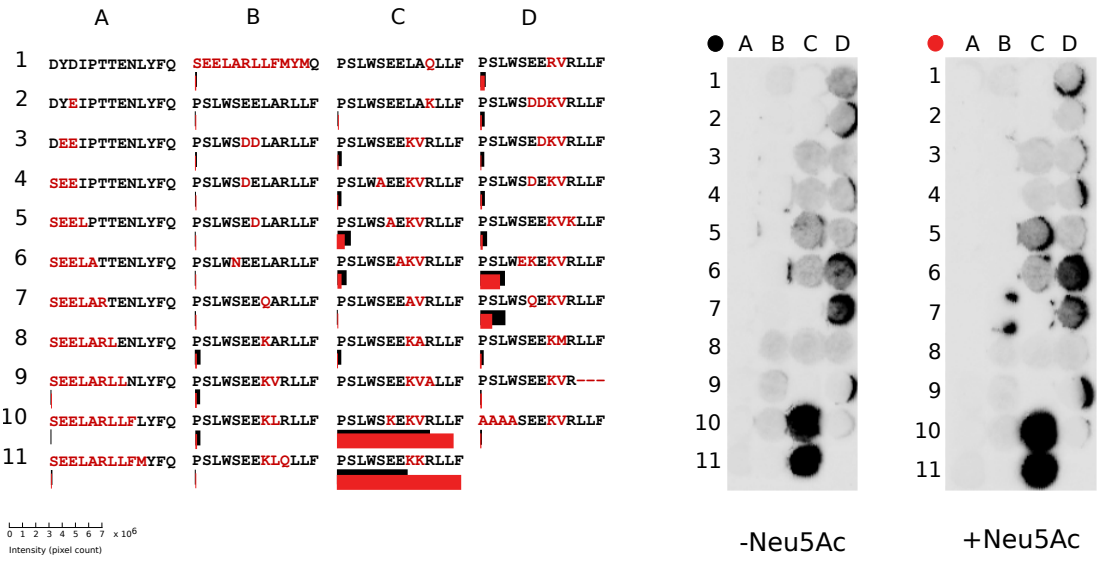


Figure 2.40. – Binding of VcSiaP to peptides anchored on cellulose membranes. The spots of the left side of the picture are the peptides whose sequences appear on the right side. The experiment was once performed without Neu5Ac (black bars) and once as a competition experiment with Neu5Ac (red bars).

Furthermore, different positions in the sequence `PSLWSEELARLLF` were mutated, in order to provide a better binding partner, which could eventually even become an inhibitor. Those mutation selections were based on the attempt to "pull" the peptide deeper into the SBP binding pocket through ionic interactions. These mutations are depicted in

red on positions B2-D10 of figure 2.40.
The scan was incubated with fluorescent labeled VcSiaP for 1 h. Then the excess SBP was washed off with buffer as described in section 4.7. A picture of the scan was recorded with an Odysee imaging system. The hits on the scan were analyzed with the software Image Studio Lite (5.2.5) and the intensities were scaled to the background to become comparable. The hit with the highest intensity was measured to be 7×10^6 pixel counts. This hit therefore became the maximum of the intensity scale which is indicated in all further peptide scan pictures of this thesis. The length of the bars in figure 2.40 correlate with the measured intensities of the scan and function as an additional measure of the relative binding strength, alongside with the original image of the scan. A table which provides an overview of the scaled data can be found in appendix D.
No fluorescence signal was detected on the position of the tag-peptide (A1). This result is not too surprising, since the interaction has to be very weak. Also, gradually changing this peptide sequence into the Q-loop hardly increased the affinity towards VcSiaP (A2-B1, black bars). The affinity improved when a positive charge in form of a lysine is introduced next to the double glutamates (B2 compared to B8, B9). Also eliminating the charge of the first glutamate, either by forming it into a glutamine, or by turning it into a lysine paired with a displaced glutamate, results in higher affinity towards the SBP (B9 compared to D7, D6). When taking a look at the crystal structure, this observation can be rationalized: the additional lysine can potentially interact with residue E16, which sits deep in the binding cleft. Figure 2.41 underlines this trail of thought (magenta 1).

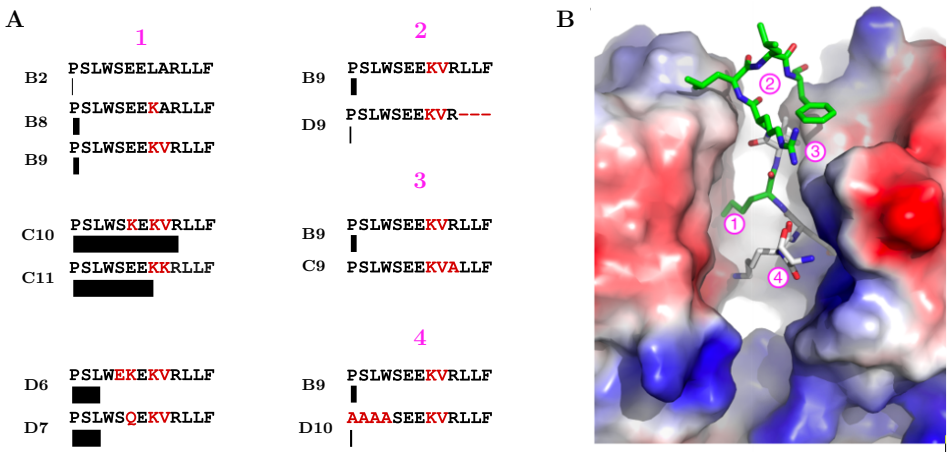


Figure 2.41. – Strongest hits of the inhibitor loops scan. A) Summary of the strongest binders from the peptide scan. B) Peptide B10 (the strongest binder, modelled onto the structure of the tag-peptide. The surface of the SBP is colored according to its electrostatic potential (white: neutral, red: negative, blue; positive).

Panel B shows the electrostatic surface of the binding cleft. The strongest binder of scan (B10) is modelled into the cleft, replacing the original tag-peptide. The lysine (magenta 1) comes in close proximity to the positive charged (red) surface. Adding a second positive charge potentially pulls the peptide further into the binding pocket, giving rise to a possible interaction with E66 of the binding cleft (figure 2.41, magenta 1). Also, the hydrophobic C-terminal sequence LLF contributes to the binding (B9 compared to D9). This sequence can interact with a hydrophobic surface on the outer corner of the binding cleft (magenta 2). The arginine at position 10 of the peptide also contributes to the interaction (B9 compared to D9, magenta 3), as well as the N-terminal sequence PSLW (B9 compared to D10, magenta 4). These two findings however can not be explained by the model in panel B. This hints to the possibility of the peptides not just being in a linear chain, but actually forming structures or different conformations depending on the sequence.

As a following step, a competition experiment with the native substrate Neu5Ac was performed on an identical scan. Here, 1mM Neu5Ac was incubated with fluorescent VcSiaP for 1h. The measured intensities of this experiment are indicated by the red bars in figure 2.40. Under this condition, most low affinity hits nearly vanished (B8-B10, C3-C4, C8, D1-D5, D8). Exceptions are the two strongest binders (C10 and C11). Those notably reveal a slight increase in affinity towards VcSiaP. It could be that these two peptides actually "favor" the closed form of the SBP.

In a final step, another identical membrane was also incubated with fluorescent HiSiaP as binding molecule. The result is depicted in appendix figure A.28. HiSiaP binds

to the same peptides, thus giving rise to quite a similar pattern, but with strongly decreased affinity. Since a HiSiaP structure with a peptide is nonexistent, explaining these findings will not be possible.

2.4.3. On VcSiaQM transporter loops

This experiment aimed at finding possible VcSiaQM loops, which the SBP VcSiaP could interact with.

Therefore, relevant periplasmatic loops of VcSiaQM were identified by multiple sequencing alignment (see appendix A.2). If the sequence was highly conserved, then the region of the sequence was located within QM-domain topology maps of Mulligan. Each spot on these two scans consists either of one loop of the transmembrane proteins, if it was short enough or of a fragment of the loop. The fragments belonging to the same loop overlap by at least one amino acid. The amino acid length varies between 8 and 10 amino acids.

Interestingly, all glutamates and aspartic acids in the loops are highly conserved, so knockout mutations of these peptides with an alanin replacing the glutamate or aspartic acid were ordered on the peptide scan as well.

Figure 2.42 visualizes which loops were selected for the scan. The chosen loops and their location on the scan are indicated in green. The red spots show which highly conserved amino acids where changed into alanine. The numbers of the periplasmatic loops are displayed in magenta.

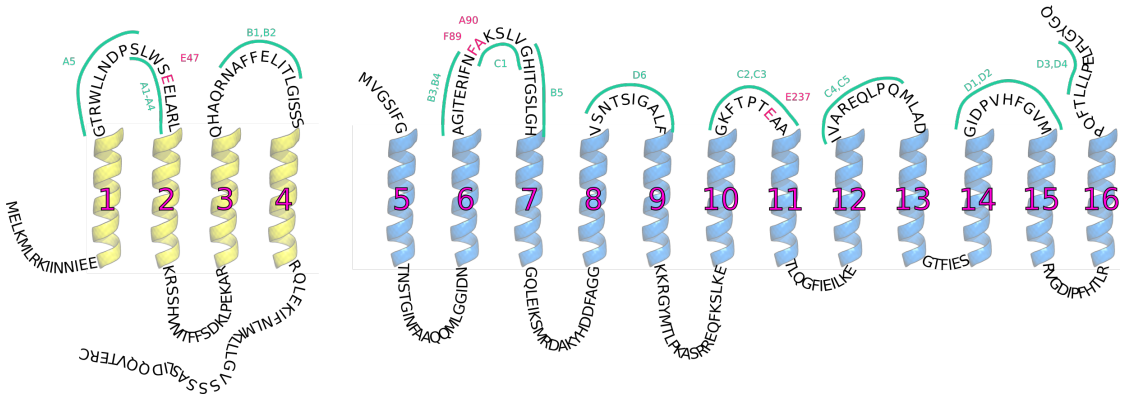


Figure 2.42. – Schematic VcSiaQM topology map as proposed by Mulligan et al. [101]. The four helices which compose the Q-domain (1-4) are shaded in yellow, whereas the blue ones compose the M-domain (5-16). The chosen loops and their respective position on the scan are marked in green. The pink AA indicate the mutations which did not growth in the M9 media in chapter 2.3.

Fluorescent VcSiaP was incubated with a cellulose membrane onto which these chosen loops of VcSiaQM were spot-synthesized. Figure 2.43 summarizes the result.

Again, the experiment was once performed without Neu5Ac (black bars) and then on an identical scan with 1 mM Neu5Ac (red bars). The inserted alanin knockouts are indicated in red in the sequences.

Only one of the peptides which correspond to the native periplasmatic loops show an interaction that can be detected by this experiment. The only original loop sequence which shows significant binding is on spot C1. It is attributed to the loop in the M-domain, which connects helix 6 and helix 7. Mutating the two conserved hydrophobic residues in this loop to glutamate (F97E and A80E) resulted in no growth in the M9 liquid media, ergo disrupting the transporter. This suggests an important role of this loop during the transporting process. Which one still needs to be disclosed by more experiments.
Strikingly, all alanin-mutated sequences bind with higher affinity than the original loop sequences. The peptide on spot A4 dominates the scan as strongest binder. In this peptide sequence the two highly conserved glutamates have been replaced by alanines (E47/48A). The mutations E47 and E47/48A also show no growth in chapter in the M9 media in chapter 2.3.
The peptide on spot B4 followes as second strongest binder. But in this case the mutation E74 grows in M9.

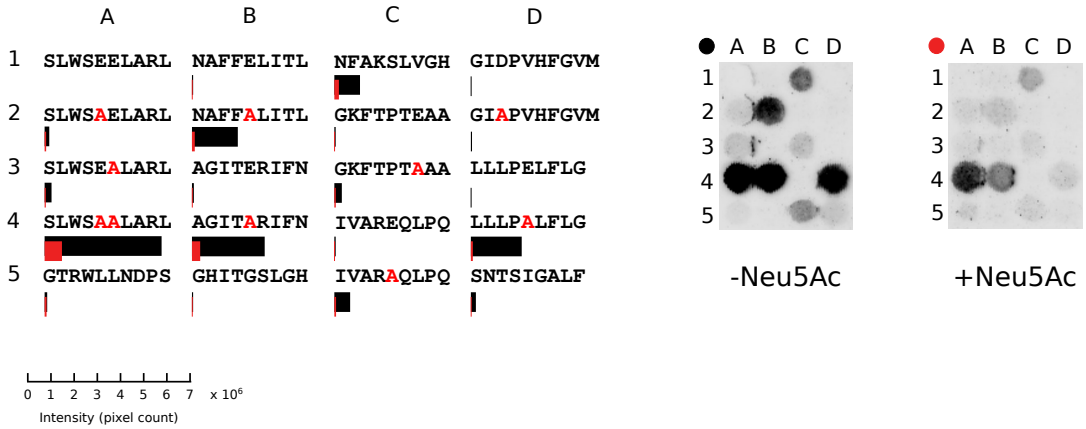


Figure 2.43. – Binding of VcSiaP to VcSiaQM-peptides anchored on cellulose membranes. The spots of the left side of the picture are the peptides which sequences appear on the right side. The experiment was once performed without Neu5Ac (black bars) and once as a competition experiment with Neu5Ac (red bars).

The competition experiment with 1 mM Neu5Ac showed reduced interaction between VcSiaP and the scan.
The experiment was also repeated with fluorescent HiSiaP as binding molecule. The result is depicted in appendix figure A.27. Interestingly, HiSiaP binds to the same peptides, giving rise to a similar pattern, but with strongly decreased affinity. This is expected, since HiSiaP discriminates between VcSiaQM and HiSiaQM.

2.5. PELDOR on selected peptides

The discovery of the peptide-bound structure of VcSiaP lead to the search for more binding partners or potential inhibitors. One of these follow-up experiments is the peptide scan, whose results are presented in section 2.4. The peptide scan is a cellulose membrane on which several peptides are directly spot-sythesized. This membrane was incubated with fluorescent labeled VcSiaP and the hits were identified by an imaging system.

The experiments presented in this chapter built on the obtained results of the peptide scan. The most promising hits of the scan were identified and ordered as lyophilized peptides. The peptides were dissolved in water and 250 µM peptide was added to 25 µM double spin labeled VcSiaP Q54R1/L173R1. PELDOR time traces were recorded to investigate whether a closing of the protein would be detectable or not.

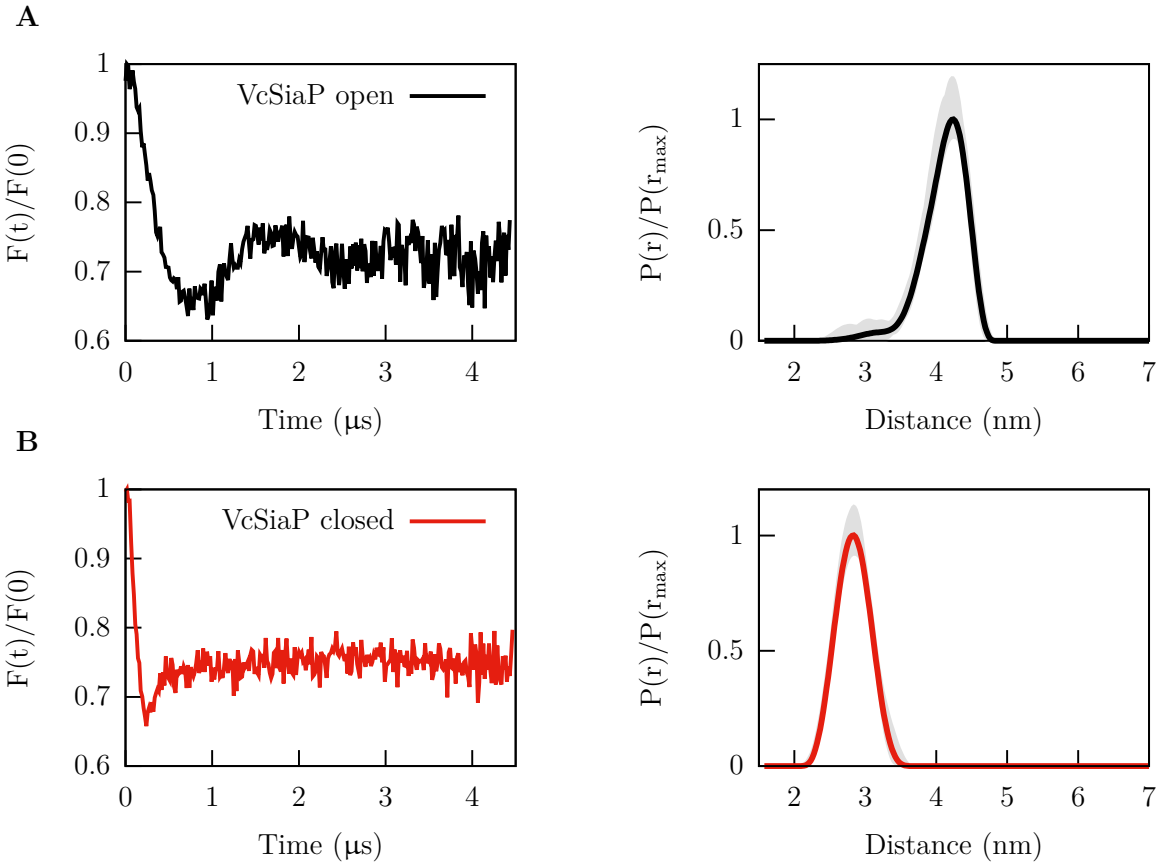


Figure 2.44. – PEDLOR timetraces and distance distributions of VcSiaP Q54R1/L173R1 without ethgly. A) shows the open conformation where as B) was measured with 1 mM Neu5Ac and presents itself therefore in closed state. Both samples were measured at 50 K. The error of the normalized distance distribution $P(r)/P(r_{max})$ was estimated by DeerAnalysis and is marked in gray.

To exclude a potential disturbing effect of the glass forming additive ethylene glycol (ethgly), it was excluded from the measurments. Glass forming additives are usually used in PELDOR samples to prevent protein aggregation and ice crystal formation during the freezing process and therefore to increase signal to noise ratio in the measurement [158].
Figure 2.44 shows the background-corrected timetraces of the protein with their respective distance distribution. The raw-data can be found in appendix A.12.
At the first glance it is clear that ethylene glycol is missing from the measurement when compared to the data in chapter 2.2. The timetraces were recorded for 4.5 µs. Previous data were recorded for 9 µs for the open conformation and 5.5 µs for the closed conformation. Furthermore, the timetraces contain more noise and therefore had to be measured longer (1463 scans for the open and 2473 scans for the closed conformation). The timetrace of the open VcSiaP (black) reveals two oscillations, whereas the timetrace of the closed VcSiaP (red) flattens out very fast.
Confirming the results in chapter 2.2, the distance distribution of the open conformation (black) peaks at 43 Å and the closed conformation (red) peaks at 27 Å.
Figure 2.45 depicts the timetrace and distance distribution of VcSiaP Q54R1/L173R1 and the peptide DYDIPTTENLYFQ, which was found in the binding cleft of the VcSiaP Q54R1/L173R1 R125A crystal structure in section 2.2.4. The time trace discerns two oscillations. When comparing it to the time trace in figure 2.44 (black), their shapes are hardly distinguishable. Also, the respective distance distribution reveals the open VcSiaP conformer.

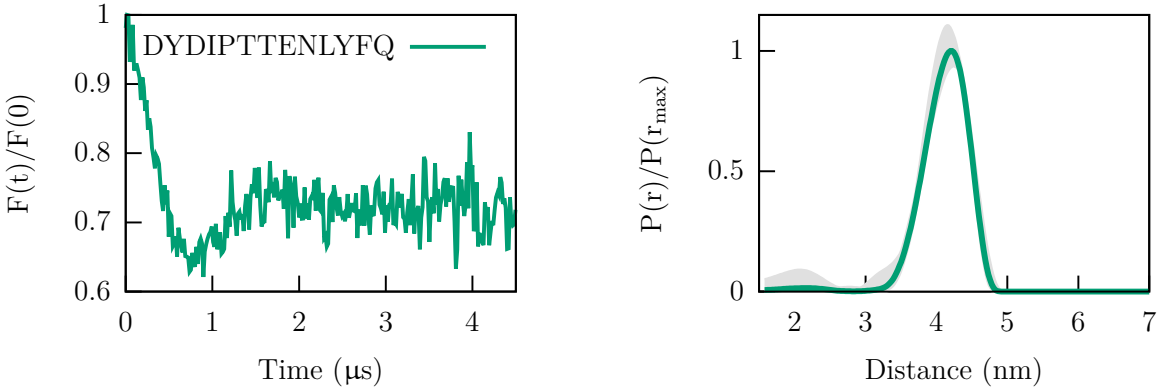


Figure 2.45. – PEDLOR timetraces and distance distributions of 25 µM VcSiaP Q54R1/L173R1 and 250 µM peptide DYDIPTTENLYFQ. The sample was measured at 50 K without ethgly. The error of the normalized distance distribution $P(r)/P(r_{max})$ was estimated by DeerAnalysis and is marked in gray.

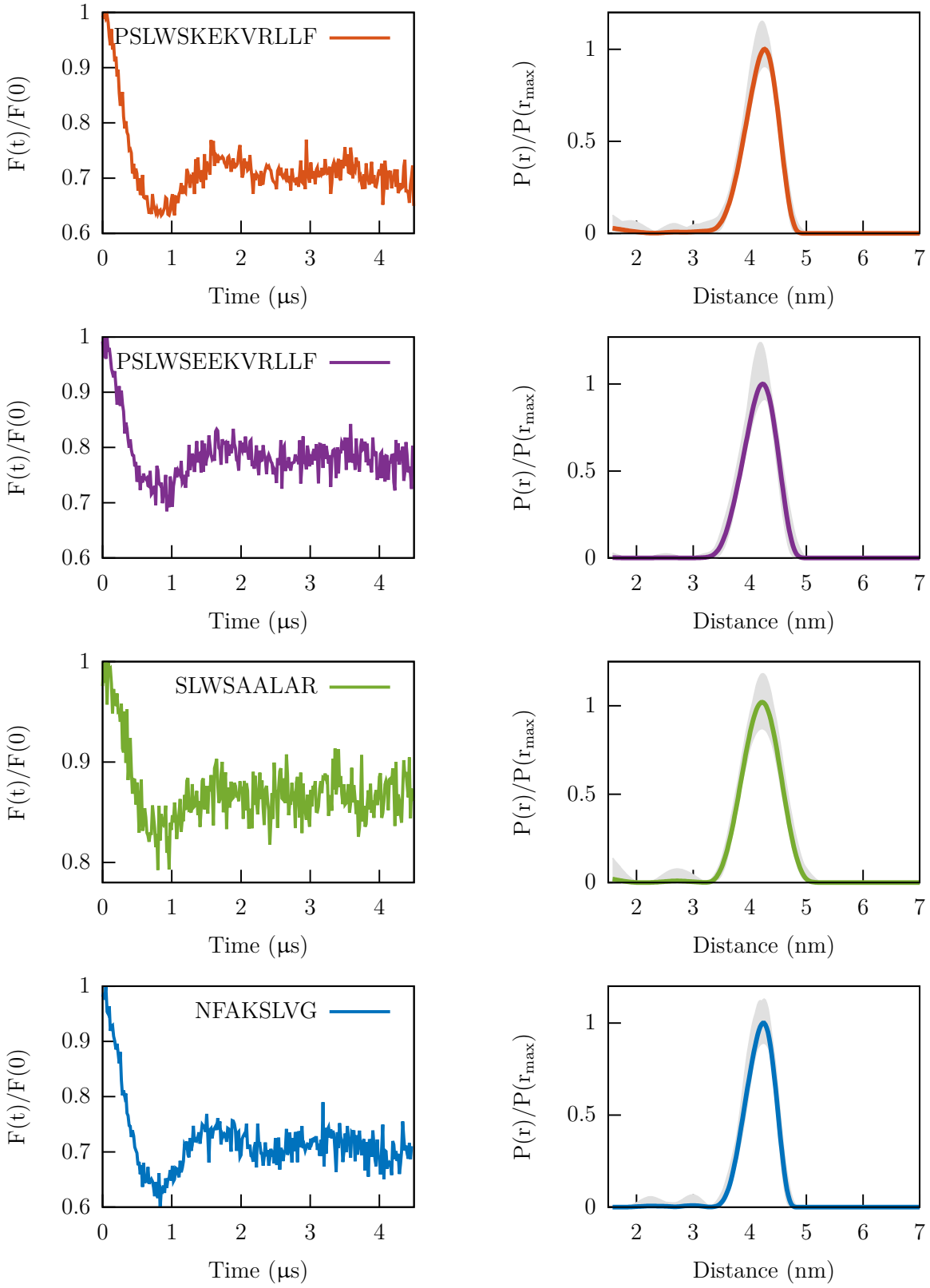


Figure 2.46. – Timetraces and distance distributions of 25 µM VcSiaP Q54R1/L173R1 and 250 µM of the selected peptides. All samples were measured at 50 K without ethgly. The error of the normalized distance distribution $P(r)/P(r_{max})$ was estimated by DeerAnalysis and is marked in gray.

Figure 2.46 shows the time traces and distance distributions of VcSiaP Q54R1/L173R1 with the peptides identified from the peptide scan.
The peptide SLWSAALAR binds stronger to the peptide scan than the original sequence SLWSEELAR and was therfore of strong interest. Also the peptide PSLWSKEKVRLLF is a stronger binding variation of the sequence PSLWSEEKVRLLF, which displays no binding in the scan. However it was included in the experiment at hand, for comparative purposes.
The peptide NFAKSLVG was included in this series of measurements for the opposite reason: it is the only periplasmatic loop sequence which binds without an introduced mutation.
When comparing the timetraces of these four measurements one finds that they all resemble the open conformation shapewise. The timetraces with PSLWSEEKVRLLF (purple) and SLWSAALAR (lightgreen) also have more narrow modulation depths. Yet, the result of these measurements is the same: no closed conformation could be detected with the selected peptides as substrate.

2.6. Investigation of VcSiaP and VcSiaQM interaction

During the transporting process an interaction between the transmembrane part VcSiaQM and the substrate binding domain VcSiaP has to take place. This is derived from the fact that the substrate has to be parted from the SBP and transferred to the QM-domains, where the actual transport into the cytoplasm takes place. However, this interaction has not yet been proven by an experiment.
This chapter tries to capture the interaction of VcSiaP with VcSiaQM by experiment. The method of choice for this undertaking is *cw*-EPR and the well-known effect of spectral broadening due to an increase in molecular size.
The construct VcSiaP Q245C was picked to be labeled with MTSSL. The single label in this case is placed on the backside of the protein which prevents it from interfering with any possible interactions taking place near the binding cleft. To estimate if any effect would be actually expected upon titrating this single spin labeled mutant with VcSiaQM, a simulation of the spectrum was calculated with the EasySpin package.

2.6.1. Simulation of the spectra

To investigate whether an effect would be visible in the *cw*-EPR spectrum due to the increased mass upon PQM-complex formation, *cw*-EPR spectra were simulated.
First the g-values and an initial rotational correlation time τ were obtained by fitting the experimentally obtained spectrum of VcSiaP Q245R1 with the "chili" routine of EasySpin. Then, the rotational correlation time τ was adjusted to the weight of the transporter complex VcSiaPQM according to equation 1.3 and a simulation was performed with "chili" and the fitted g-values.
The resulting *cw*-EPR spectra are summarized in figure 2.47. The black spectrum (VcSiaP) is the fitted spectrum and the red spectrum (VcSiaPQM) is the spectrum with adjusted τ and table 2.9 lists the parameters used for the simulation. The figure shows clearly a broadening of the spectrum with formation of the complex. Yet, one needs to keep in mind that the red line is 100% complex, which is unlikely to form and stay stable in solution. Nevertheless, the simulation seems to indicate that the experiment is feasible.

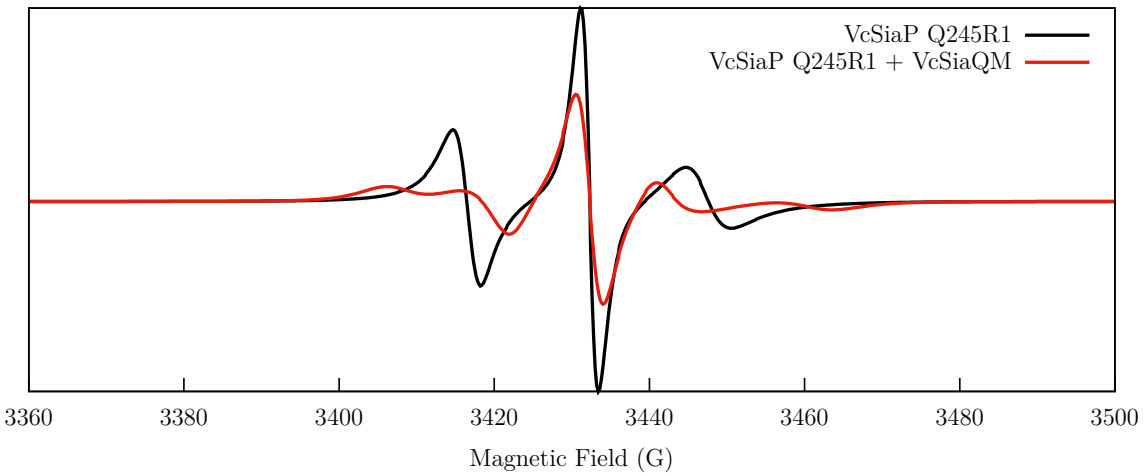


Figure 2.47. – Simulation of *cw*-EPR spectra of VcSiaP Q245R1. The calculation was carried out using the "chili" routine of the simulation package EasySpin. The parameters used for calculation are listed in table 2.9. The simulation of the individual P-domain (black) shows the three transition lines of unequal height, that are typical for the R1-label bound to a protein in solution. The red spectrum belongs to calculation of the heavier VcSiaPQM-complex and shows clear signs of broadening.

Table 2.9. – Parameters used for simulating *cw*-EPR spectra.

Name of protein	**Parameters**
VcSiaP Q245R1	g= (2.00572, 2.00599, 2.00323) | τ= -8.61381 | LWPP= (0.0433,0) | A= (22.5451 18.1733 94.912)
VcSiaP Q245R1 + QM	g= (2.00572, 2.00599, 2.00323) | τ= -8.139554288 | LWPP= (0.12904 0) | A= (22.5451 18.1733 94.912)

2.6.2. Purification of VcSiaP Q245/R1

The purification started following the same protocol as used on the spin labeled double mutants in section 2.1.2: the constructs were first washed on a Ni^{2+} affinity column and then labeled with MTSSL in the elution step with imidazole. The elution was collected and subjected to ion exchange chromatography with an ENrich Q 10/100 column. This was followed by a size exclusion chromatography on a Superdex 200 16/60 column. Figure 2.48 shows the chromatogram and an SDS PAGE of the pure

protein.

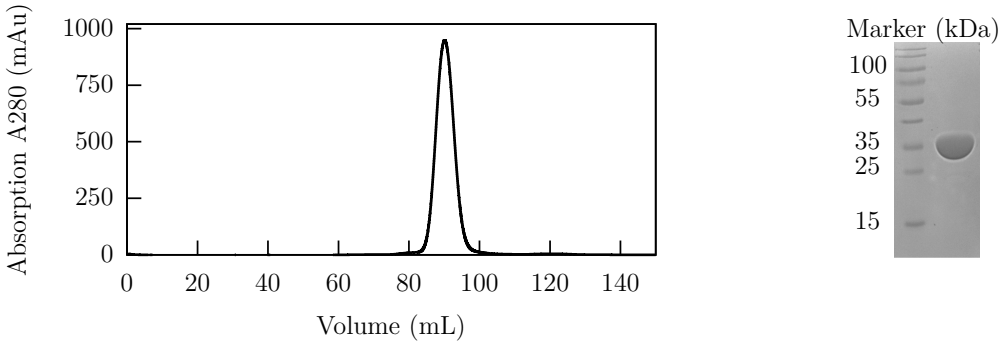


Figure 2.48. – Purification of the construct VcSiaP Q245R1. The chromatogram of the size exclusion chromatography on a SD 200 16/60 and the respective SDS PAGE gel is depicted.

Further, for the measurements, the protein VcSiaQM was needed and provided to me by my co-worker Martin Peter.

2.6.3. CW-measurements of VcSiaP/R1 with VcSiaQM

The backbone mutant VcSiaP Q245R1 (25 µM) was titrated with tenfold excess of VcSiaQM (250 µM) and *cw*-EPR spectra were recorded at room temperature.
Figure 2.49 A shows the superposition of the recorded spectrum without VcSiaQM (black) and with VcSiaQM (red). Both spectra can be almost perfectly aligned, no significant change was observed.
Testing the hypothesis that interaction would only make sense in case of bound VcSiaP, the mutant was supplemented with VcSiaQM an 1 mM of Neu5Ac. Figure 2.49 B depicts this recorded spectrum (red) and compares it again to the spectrum of only VcSiaP Q245R1 (black). Again, in this case no change in the shape of the spectrum is detectable.
A possible explanation would be that the tumbling movement of the label is still too fast to detect a formed complex of VcSiaPQM. A label that attaches to the protein with more rigidity could possibly provide the sensitivity needed to detect the complex in this experimental setting. Therefore the measurements were repeated with the bipedal spin label Rx. This label is more restricted in its conformation and dynamics than its monopedal counterpart R1.
Ideal positions that could be labeled with the Rx label were identified via molecular modeling to be i-i+4 position of an α-helix. In the case VcSiaP Q245C, the second cysteine was therefore successfully placed at position D241C. Figure 2.50 shows a model of VcSiaP D241-Q245Rx. The label was modeled to the protein using Coot.

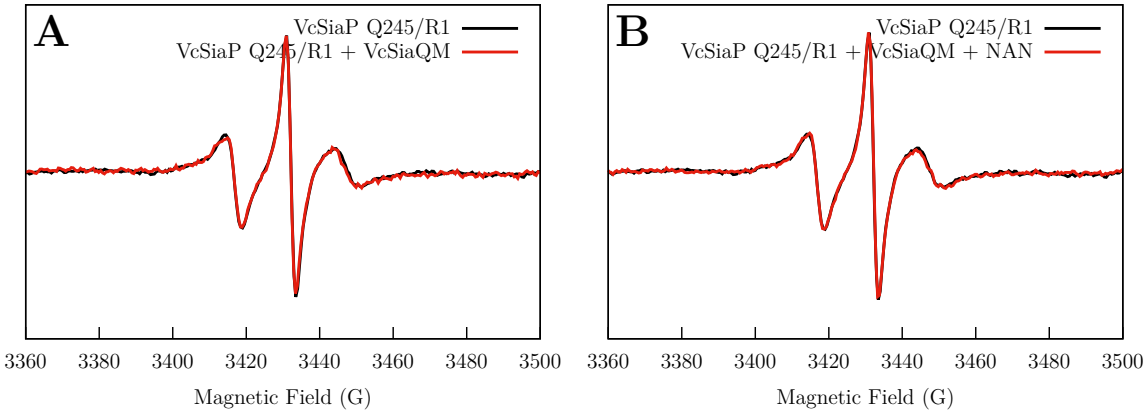


Figure 2.49. – *cw*-X-band spectra of VcSiaP Q245R1 measured at room temperature. A) The spectra of VcSiaP Q245R1 without any supplements (black) and with VcSiaQM (red) are superimposed. B) The spectrum of the labeled protein (black) is superimposed with the spectrum of VcSiaP Q245R1 + VcSiaQM + Neu5Ac (red). No changes can be observed in both panels.

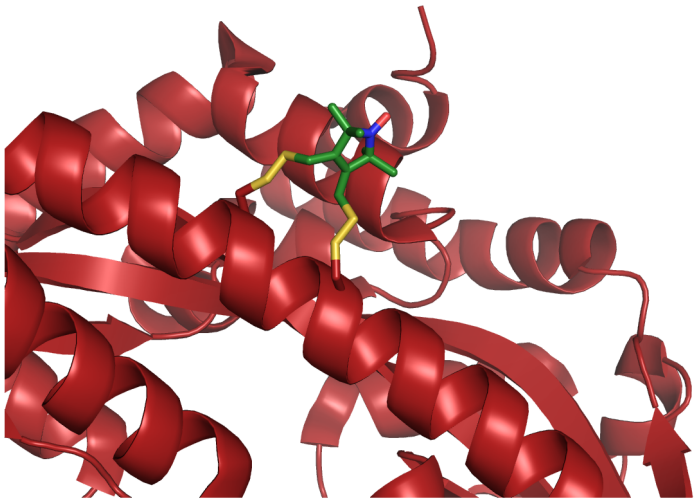


Figure 2.50. – Model of VcSiaP D241-Q245Rx. The picture illustrates, that the label has to span one turn of helix for a successful binding.

2.6.4. Purification of VcSiaP D241-Q245Rx and VcSiaQM

VcSiaP D241-Q245Rx

In a first attempt to label this protein, the regular protocol used for labeling VcSiaP Q54R1 L173R1 was followed. In this protocol the cysteines of the protein are labeled while bound to Ni^{2+} beads on a polystyrene column with excess label during the elution step. Repeating this for VcSiaP D241-Q245Rx, leads to a protein which gives a strong free label signal in *cw*-EPR measurements. Figure 2.51 shows the cw spectrum from this labeling attempt. The spectrum of VcSiaP D241-Q245Rx (red) is superimposed with the spectrum of VcSiaP Q245R1 (black). The shape of the spectrum of the Rx labeled protein is an overlap of different signals: it is slightly broadened due the use of the more rigid Rx label, but the two spikes that are marked with asterisks are clearly due to label behaving as if it was free in solution.

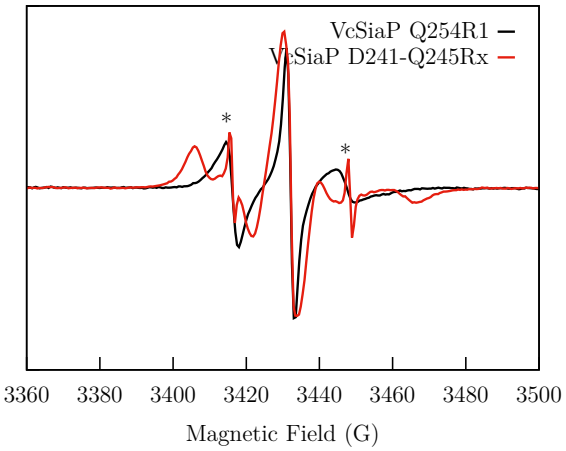


Figure 2.51. – *cw*-X-band spectra of VcSiaP D241-Q245Rx from first labeling attempt. The recorded signal of VcSiaP D241-Q245Rx (red) is superimposed on the spectrum obtained from VcSiaP Q245R1 (black). The spikes which arise from the free label are marked with asterisks.

This made the whole batch unusable for the planned experiment. Therefore the labeling procedure had to be optimized for this protein. Instead of directly labeling the protein on column, it was washed, eluted with TCEP and imidazole and then subjected to ion exchange chromatography on an ENrich Q 10/100 column. After determining the protein concentration, the protein was rebound to the Ni-resin and the TCEP was removed with washing buffer. For on-column labeling, 0.5 equivalent of label was dissolved in 5 ml of buffer and slowly added to the protein-bound resin. This was left for incubation for 1 h. This step was repeated and the protein was again incubated for 1 h with the label.

After labeling, the protein was eluted and the His-Tag cleaved with TEV protease in an overnight dialysis. On the next day the protein was subjected to size exclusion chromatography and dimerized protein was separated from the monomeric fraction.
Figure 2.52 shows the chromatogram of the protein and its respective SDS PAGE (non-reducing, without β-Mercaptoethanole). This batch contains a small dimer peak that is indicated with a blue line. The fractions collected and used for recordings is indicated by a red line.

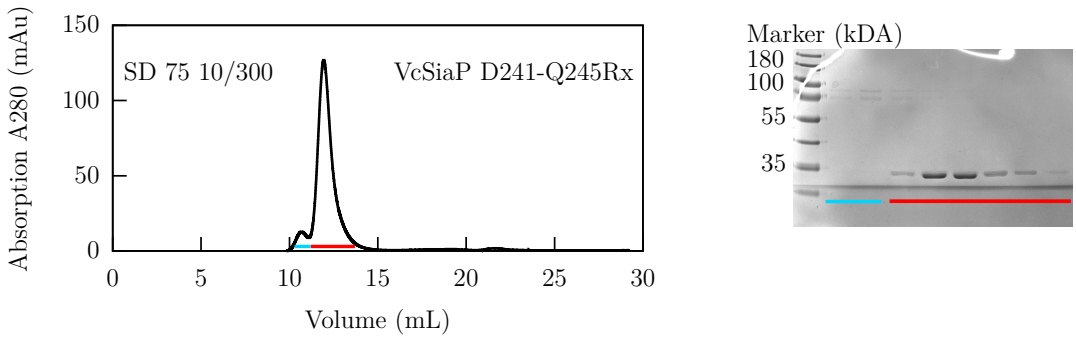


Figure 2.52. – Purification of VcSiaP D241-Q245Rx. In the final purification step the proteins was subjected to size exclusion chromatography on a SD 75 10/300. The respective SDS PAGE (non-reducing gel) of the collected fractions shows a dimerized portion of the protein (marked with a blue line.) The monomeric portion is marked with a red line.

This protocol yielded properly labeled protein, which was usable for the experiment. In section 2.6.5 the *cw*-spectrum will be shown and discussed.

VcSiaQM

The VcSiaQM construct was cloned into the pBadHisTev vector and carries a N-terminal His_{14}-tag which is only attached to the Q-domain. Since both form a strong complex in solution, this set-up leads to a co-purification of the M-domain.
The purification of a membrane protein differs from the purification of soluble proteins. Since the targeted protein is incorporated in the membrane of the bacterial cell (in this case the expression system was MC1061), the working buffer has to imitate this environment for the protein. Usually this is done by extracting the protein from the membrane with a detergent and upholding a certain detergent concentration in the buffer:
After harvesting and pelleting the cells, the protein was subtracted from the pellet over night using 1.5% DDM in the regular QM buffer. The membrane debris was then separated from the solubilized protein by centrifugation. The protein solution was then

bound to Ni^{2+} beads on a polystyrene column and further purified with washing steps. The elution was performed with imidazole. Finally the protein complex was subjected to size exclusion chromatography on a SD 200 10/300 column. Figure 2.53 shows the result of this run and the respective SDS PAGE of the collected fractions.
The line with the Q-domain emerges slightly above the 15 kDA marker. The size of the tagged Q-domain amounts to 19.5 kDa. The line with the M-domain is visible slightly above the 35 kDA marker. This domain appeared at a position that corresponds to a mass smaller than the actual mass of the protein 45.5 kDa. This is a well known effect with integral membrane proteins [159], due to incomplete unfolding by SDS.

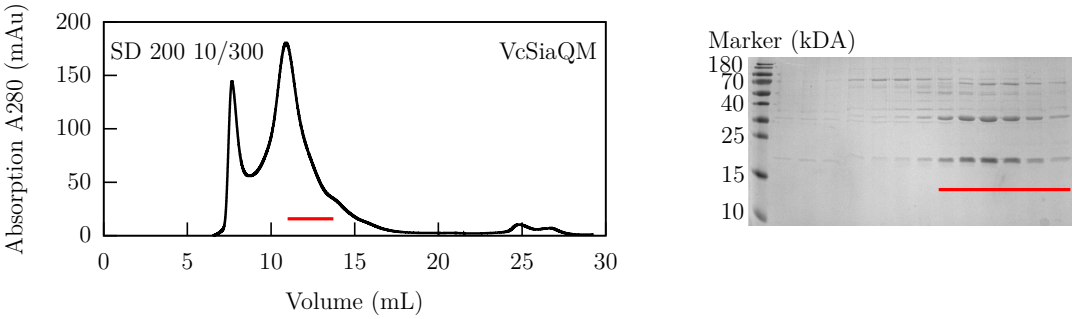


Figure 2.53. – Chromatogram of the size exclusion chromatography of VcSiaQM. The used column was a SD 200 10/300. The collected fractions are marked with a red line. The respective SDS PAGE shows the Q-domain appear at 20 kDa and the M-domain emerge at 37 kDa. In solution both domain form a tight complex.

2.6.5. CW EPR spectra of VcSiaP D241-Q245Rx

Figure 2.54 shows the recorded *cw*-EPR spectra of VcSiaP D241-Q245Rx. The measurements were performed at room temperature in VcSiaQM buffer (50 mM KH_2PO_4 at pH=7.8, 200 mM NaCl and 20% glycerol) on an EMXmicro spectrometer equipped with an ER4119HS resonator. The single recordings were not superimposed, but stacked for easier comparison.
At first, a spectrum of 25 µM protein without any additional substrates was recorded (purple). The shape of this spectrum is clearly broadened: Little is left of the three transition lines that are typical for the protein-bound R1 label. The most prominent transition line (center peak) has its inflexion point at 3509 G.
The spectrum of 25 µM VcSiaP D241-Q245Rx with 1 mM Neu5Ac (green) does not reveal any binding activity of the P-domain, the shape of the spectrum does not differ from the shape of the spectrum of the sole Rx labeled protein.
The spectrum of 25 µM P-domain with 25 µM QM-domain (blue) shows spikes of free

label at 3482 G and 3517 G (marked with asterisks). These also appear in the spectrum of 25 µM VcSiaPQM with 1 mM Neu5Ac (red). Other than that no significant spectral broadening could be observed between the spectra without QM-domain and with QM-domain.

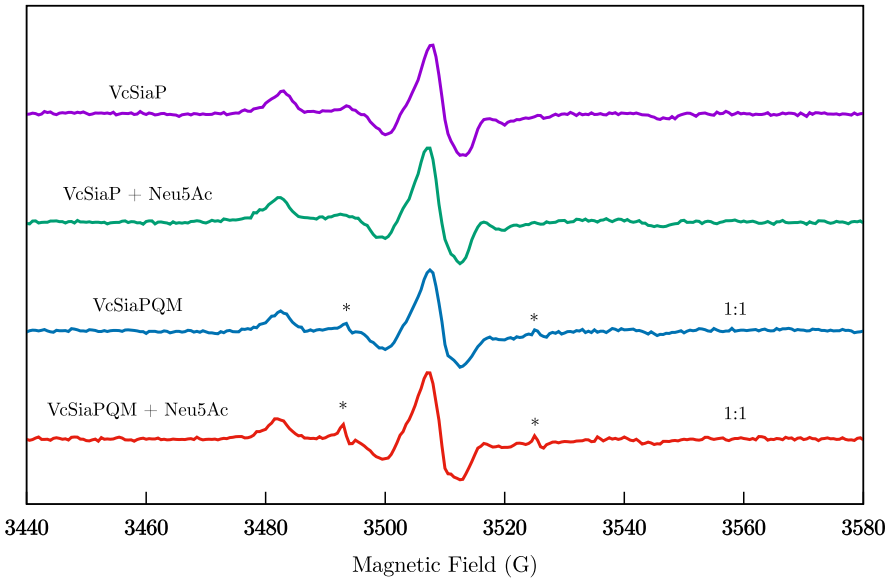


Figure 2.54. – *cw*-EPR spectra of VcSiaP D241-Q245Rx supplemented with VcSiaQM. The Rx-labeled P-domain (25 µM) was measured without any supplements (purple), with 1 mM Neu5Ac (green) and with 25 µM VcSiaQM (blue). The red spectrum is combination of 25 µM VcSiaP D241-Q245Rx, 25 µM VcSiaQM and 1 mM Neu5Ac. Spikes of free label occur at 3482 G and at 3517 G. Those are marked with an asterisk.

The idea that maybe more QM-domain needs to be added in order to force complex formation presents itself. The limiting factor in this experiment is the amount of VcSiaQM. The homolog HiSiaQM expresses as fused protein and is easier to purify and concentrate, therefore the same set of experiments was repeated with HiSiaP V250-K254Rx and HiSiaQM. The results are depicted in the following section.

2.6.6. Purification of HiSiaP V250-K254Rx

HiSiaP V250-K254Rx was purified following to the new improved protocol mentioned in section 2.6.4.
Figure 2.55 depicts the size exclusion chromatogram and the respective SDS PAGE of the collected fractions. The blue line indicates the improperly labeled and therefore dimerized portion of this purified batch. The red line indicates the collected monomer.

In the next subsection the cw-EPR spectra of HiSiaP V250-K254Rx will be presented and discussed. The protein HiSiaQM was purified by Martin Peter.

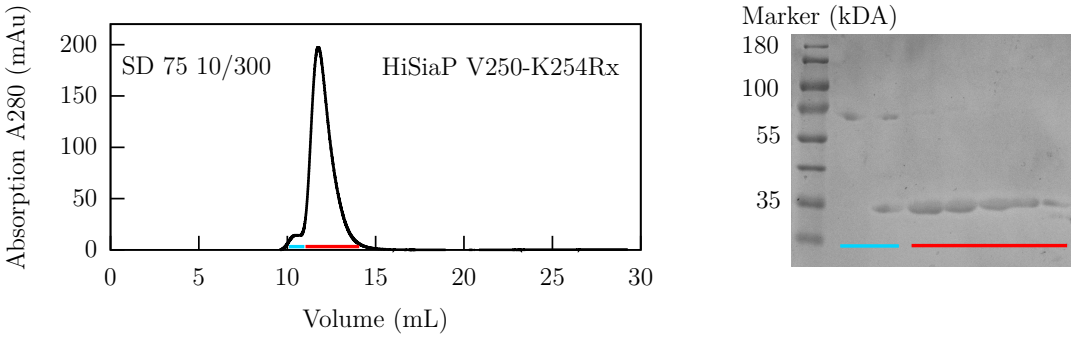


Figure 2.55. – Purification of HiSiaP V250-K254Rx. The proteins was subjected to size exclusion chromatography on a SD 75 10/300. The respective SDS PAGE (non-reducing gel) of the collected fractions shows a dimerized portion of the protein (marked with a blue line.) The monomeric portion is marked with a red line.

2.6.7. CW EPR spectra of HiSiaP V250-K254Rx

Figure 2.56 shows the recorded *cw*-EPR spectra of 25 µM HiSiaP V250-K254Rx.

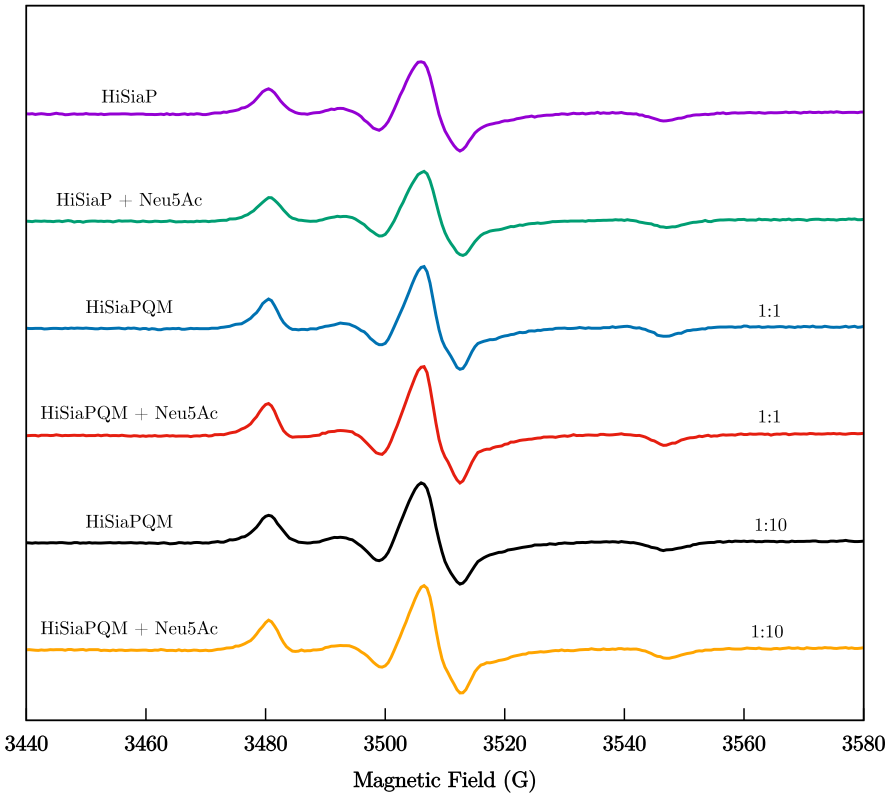


Figure 2.56. – *cw*-EPR spectra of HiSiaP V250-K254Rx supplemented with HiSiaQM. The Rx-labeled P-domain (25 µM) was measured without any supplements (purple), with 1 mM Neu5Ac (green) and with 25 µM HiSiaQM (blue). The red spectrum is combination of 25 µM HiSiaP D241-Q245Rx, 25 µM HiSiaQM and 1 mM Neu5Ac.

The measurements were performed at room temperature in QM buffer (50 mM KH_2PO_4 at pH=7.8, 200 mM NaCl and 20% glycerol) on an EMXmicro spectrometer equipped with an ER4119HS resonator.

Just as expected from the preceding experiment, there is no detectable difference in the spectrum upon substrate binding (compare purple line to green line). Also, when the QM-domain is added in a 1:1 ratio to the P-domain no difference can be seen (compare purple and blue lines). Even at an excess of 10-fold QM-domain, still no signs of complex formation, i.e., spectral broadening can be observed (compare purple line to black line and green line to yellow line).

3. Discussion

The work of this thesis is dedicated to investigating the transport mechanism of the TRAP transporter VcSiaPQM from *Vibrio cholerae* and is divided into three main research topics:
One main field is elucidating the function and dynamics of the substrate binding protein VcSiaP. Another one is searching and optimizing possible peptide binding partners of VcSiaP. These optimized petide binders are also a starting point in the search of future inhibitors.
And the last task tackles the question: how does the interaction between P and QM domains take place?

3.1. Dynamic studies on VcSiaP

The whole transporter consists of three parts. Two are embedded in the membrane (VcSiaQ and VcSiaM). The third part is an SBP (VcSiaP) located in the periplasmic space of the gram-negative cell.
VcSiaP scavenges the substrate of the TRAP transporter, which is the sialic acid Neu5Ac. Subsequently after the capture, VcSiaP delivers Neu5Ac to the transmembrane parts, which further translocate it to the cytoplasmic space.
At the beginning of the investigations, crystal structures of P-domains of related transporters were available in the open (substrate-free) and also in the closed (substrate-bound) forms [97, 100, 138]. Further, the open structure of VcSiaP [99] was also known. But little was known about the dynamics of VcSiaP [160].
The state in which VcSiaP resides prior to Neu5Ac binding was completely unclear. Does VcSiaP exists in an equilibrium of open and closed conformation in the absence of substrate? Another possible dynamic was to remain in the open form and then transform into the closed form once substrate was available: like a switch, but induced by substrate.
A look at the well studied ABC transporters reveals that a conclusion about the state of SBPs in solution is not possible by comparison. This question has to be answered for each transporter individually through experiment. The SBP of the maltose ABC

transporter for example fluctuates between the open and closed state in the absence of the substrate [161]. On the other hand the glutamine ABC transporter GlnBP remains open when no substrate is present [162]. And in the case of the leucine transporter LeuT, no clear state can be confined, but merely a shift of the equilibrium occurs [163]. To tackle this question PELDOR (pulsed electron electron double resonance), also known as DEER (double electron electron resonance), became the structural biology tool of choice.

3.1.1. VcSiaP functions like a switch

To make the changes of conformation "visible" for the PELDOR technique, spin labels were introduced into the protein. The labeling sites were identified, based on a difference distance matrix, using the open and closed crystal structure of the homolog protein HiSiaP. The identified sites were labeled with MTSSL and subjected to PELDOR measurements. From this data distance distributions were subtracted with the program Deeranalysis. Four double-spin labeled constructs were investigated: 54/173, 110/173 and 173/225 with 54/110 serving as a control.
Further, four rigid subdomains of VcSiaP could be identified via the difference distance matrix. Those were superimposed onto the closed HiSiaP structure, to obtain a working model of closed VcSiaP. This working model and the open VcSiaP structure were used to predict distance distributions for the chosen labeling sites.
The analyzed data clearly shows only one distribution in the absence of the substrate Neu5Ac. In the presence of excess Neu5Ac, this distribution shifts to another single distribution peaking at a smaller distance. Also, for the pair 54/173 these two distributions are clearly separated and do not have overlapping regions.
The predicted distance distributions fit the measured data nicely and support the described findings. The crystal structure of doubly spin labeled open VcSiaP mutant 54/173 R125A also displays a distance which perfectly fits into the high probability part of the respective distribution.
However, since the samples were flash-frozen and PELDOR can only detect the steady state of the protein, one can not rule out transient states completely, which might have snapped to one of the states and therefore could not be detected.
Taking all presented arguments into account, the SBP can be interpreted as fully open without fluctuating to the closed conformation, within the measurement scope of PELDOR. Here, VcSiaP functions like a "switch". The protein is open at all times, unless binding of the substrate forces it to undergo conformational change to the closed form. One can argue that evolution formed VcSiaP in such a manner

to close only in contact with substrate in order to avoid unnecessary closing of the transmembrane parts SiaQM and therefore conserve energy[89].

MD (molecular dynamics) and RAD (restrained average dynamics) simulations performed by Marinelli and Fiorin [164] on VcSiaP deliver strong computational arguments which backup the finding that VcSiaP only undergoes change when binding the substrate:
In their study, semi-closed states of VcSiaP snapped into an open-state conformation after a few nanoseconds runtime. Their work also suggest that the fully-closed conformation is not accessible in absence of the substrate.
Previously performed simulations by the same group on the SBP TeaA of the ecotin TRAP transporter TeaABC also predicted these results [165]. For this SBP, the complete (substrate-free) closure is energetically unfavorable, since the closed state is approximately 3 kcal/mol costlier than the open conformation. According to their calculations, it is therefore less than 1% probable without the substrate.

These findings further support the idea that VcSiaP will only return to the open conformation after having interacted with VcSiaQM and therefore after having passed on the substrate [89].

3.1.2. Transition from open-form to closed-form

To observe how exactly the transition of open-form VcSiaP to closed-form VcSiaP takes place, the mutant 54/173 was titrated with the substrate and again subjected to PELDOR measurements. Again, the distance distributions were calculated from the measured time traces.
The open-conformation peak in the distribution gradually decreases, while the peak of the closed conformation rises with higher substrate concentrations. No stable intermediate states appear. The measurements reflect that, in solution with varying concentrations of Neu5Ac, an equilibrium exists only between the open and closed states.
A method to study the movement of the protein in a bigger picture, which also captures thermodynamically unstable states, would be to rapid freeze quench the protein with its substrate [166, 167]. The downside, however, is an experimentally very challenging set-up. And reasoning from initial freeze quench trials on the protein VcSiaP (data unpublished), the rigid-body movement is far below 12 ms.
Furthermore, the dissociation constant K_D was determined by linear combination

fitting of the PELDOR timetraces of the fully opened and fully closed conformations. The ratio of both is a function of substrate concentration. The obtained K_D value is close to the value obtained by surface plasmon resonance, in the scope of this thesis. Both values again are in alignment with values obtained from experiments like isothermal titration calorimetry (ITC) and tryptophan fluorescence quenching [99, 101, 107].
Integration of the distance distributions and recalculation of substrate concentrations lead to the same results as linear combination fitting of the PELDOR timetraces. The marginal concentration differences are within the error range of the fitting procedure.

It remains to judge the method in general for obtaining the dissociation constant. One has to consider that the PELDOR samples were flash-frozen. So, a K_D that was obtained via PELDOR can differ from a K_D that was accessible from a measurement performed at room temperature. Furthermore, the PELDOR samples are supplemented with 50% ethylene glycol which possibly has an effect on measurement, the extent of which still needs to be quantified.

3.1.3. Investigation of the closing-mechanism

The SBP VcSiaP has a well defined binding pocket [100], in which the substrate is snugged in tightly upon binding. The two domains around the binding pocket are connected with a hinge region, which bends when closing takes place.
The role of "selectivity filter" R145 for substrate binding has been already well investigated [168] but it is not clear yet what triggers the closing mechanism. The interplay of the totally conserved amino acids R125, E184 and H207 was predicted to be necessary for the conformational change between the open and closed SBP [99]:
Comparisons between the open ligand-free and closed ligand-bound structures of HiSiaP show that E184 forms a tight salt bridge with R125 before binding Neu5Ac. Upon ligand binding, this interaction gets weakend and E184 forms a stable salt bridge with H207. Müller et al. implicated that H207 plays a crucial role in protein folding and could be the amino acid which triggers the movement [97].
Various mutations were subsequently introduced into this amino acid triad and PELDOR was used to check upon their ability to undergo transition from one state to the other.

R125 was mutated into alanine (R127A) and into lysine (R127K). Only a small

fraction of the mutant R125A was able to transform into the closed state in the presence of substrate (1%).
The measurement was repeated with higher concentration of Neu5Ac and still the fraction of closed conformation was very small (4%). Without substrate R125A remained in the open conformation, as did all the other introduced mutations.
For the mutant R125K a larger fraction (15%) of closed state was detected with PELDOR. Excess of Neu5Ac lead to an increase of the closed state in this mutant (33%).
Previous work on these mutants show decreased binding affinity [99, 169], which was also observed in the PELDOR experiments: R127A >3mM, and R127K around 1mM. When taking a look at the crystal structure of HiSiaP one can see why this is the case. The δ-guanido group of R125 forms a salt bridge with the C1 carboxyl and the C2 carboxyl group of the substrate Neu5Ac. Mutation to alanine completely destroys this possibility and it can only be partially restored by reintroducing positive charge via lysine.

The residue E184 plays a different role. Mutation E184A results in quite a lot of closed conformation in presence of substrate (55%). With E184Q the fraction of the closed protein is a little bit higher (61%) and with E184D the protein was nearly fully closed (91%). In conclusion this data reveals that E184 plays an important, but not an essential, part when it comes to the function of VcSiaP. As Setty et al. already observed, E184 seems to function as ionic stabilizer for the essential residue R125. Also, a decrease in binding affinity was also observed for E184Q [99]: 1.78 µM.

Interestingly, the residue H207 plays a subordinate role when it comes to binding of Neu5Ac. The H207A mutant was also nearly fully closed (93%) in presence of the substrate. The same can be said about the H207D mutant.
For H207A, a decrease of binding affinity was measured as well, but it is not as drastic as for the other mutants [99]: 0.662 µM.

In an earlier study by Johnston et al., this particular triad was investigated in *H. influenzae* with respect to the sialylation of the LPS of the cell [138]. Different degrees of sialylation resulted as an aftermath of the introduced mutations. R125A resulted in no sialylation of the LPS. With the mutations R125K and E184Q partial sialylation was possible. And mutation H207A resulted in full sialylation of the LPS. These results harmonize with the PELDOR studies of this thesis.
The fact that all mutants were able to adopt the closed conformation is a very strong

argument against the claim that the interplay of the amino acid triad triggers the closing mechanism of VcSiaP. What can be observed from experiments is that these amino acids have a role when it comes to the binding of the substrate Neu5Ac. And their importance decreases with increasing distance to the essential amino acid R125. R125 is a critical amino acid for high affinity substrate binding, but not essential for the closing mechanism. E184 seems to function as ionic stabilizer for the essential residue R125. The expected role of H207 as mechanical trigger of the movement could be disproven. It only further stabilized the binding of Neu5Ac.
So, inserting mutations in the triad does interrupt the interplay of salt bridge formation, but the effect of this is a decrease in the binding affinity.
This implies that another mechanism has to explain the closing mechanism of VcSiaP and other TRAP SBPs.

3.1.4. A peptide can interact with selectivity filter R145

During the investigation of the R125-E184-H207-triad, efforts were made to crystallize the investigated mutations. During this period the mutation VcSiaP R125A Q54C L173 formed a crystal and the structure was solved at 2.1 Å via molecular replacement with the native VcSiaP.
Surprisingly, this structure revealed also positive electron density in the binding cleft, which then turned out to be part of the His_6-tag of the neighboring VcSiaP protein. Thus forcing two VcSiaP proteins into dimerization within the crystal structure.
There are also other TRAP SBPs which have been found to dimerize [170][171], but the dimerization of this structure is mediated through an "artificially introduced peptide". This "peptide" shows interactions with many closeby residues of VcSiaP which which support the binding : R69 of the SBP interacts with Y2 of the peptide, S42 of the SBP interacts with T7 of the peptide and various hydrophobic interactions. But the most striking interaction is the saltbridge it forms with residue R145.
Residue R145 alongside with R125 is one of the most important residues when it comes to binding of the native substrate Neu5Ac. It is involved in substrate binding of nearly all known TRAP SBPs so far [168].

3.1.5. A peptide can trigger partial closing

To quantify the effect of this "peptide binding" on the overall structure of VcSiaP, a difference distance matrix was calculated. The pattern of this diffDm map was then compared to the diffDM pattern that is produced by the native substrate. This analysis revealed very clearly that the binding of the substrate had triggered partial closing of

VcSiaP.
The comparison clearly shows the same rigid body movements that are involved with binding of the substrate. Therefore the exact same closing mechanism that is applied for the substrate also takes place for the peptide. Further the crystal structure captures an intermediate state, a snap picture of the closing mechanism which no experiment yet tracked.
One can argue that for inducing the closing mechanism, first an interaction of the substrate with the selectivity filter R145 has to take place, and then through a pattern of varying salt bridges on both sides of the two domains, the fly trap is pulled together. The domain closes mechanically, when the substrate interacts with enough residues on both sides of the two domains.
Why is this a reasonable hypothesis? There had been a search for triggering residues for a while now, and no trigger residue has been found yet. Also the fact that other substrates than Neu5Ac can be bound by the P-domain hints towards this direction.

3.2. The search for possible peptide binders and inhibitors

TRAP transporters of pathogens like *Vibrio cholerae* and *Haemophilus influenzae* are the perfect target for highly needed antibiotic development: the TRAP transporters of these pathogens are absent in humans. During drug development the specificity can be optimized for this transporter without interfering with other transporters which are present in the human body.
This motivation let to manually designing and searching for possible inhibitors on the peptide scan. The partially closed crystal structure of VcSiaP binding to the artificial peptide was the perfect starting point of this endeavor.

3.2.1. VcSiaP and can bind to peptides in solution

The partially closed crystal structure of VcSiaP with the peptide in the binding cleft shows that the binding cleft can also hold on to artificially designed peptides. To rule completely out that this is a crystal packing artifact, several peptides were spot synthesized on a membrane and incubated with VcSiaP onto which a florescence label was attached. Several different hits were obtained which prove that VcSiaP can bind peptides in solution.

3.2.2. Peptides could be optimized for binding

The sequence of the His_6-tag was gradually changed on the membrane into the loops sequence which connects helix 1 and helix 2 of VcSiaQ. This was considered a good starting point for the search of binders since the loop has two strongly conserved glutamates and this architecture strongly resembles the "scoop loop" of the ABC-transporter $MalFGK_2$:
In the ABC-transporter $MalFGK_2$ the opening of the SBP takes place by loops on the surface of the TMs which crack-open the SBP and scoop the substrate out of the binding pocket.
The fact that the crystal structure of VcSiaP binds a peptide and is locked into a transient state suggests the possibility that loops of VcSiaQM could also act as a scoop loop. Further, the interaction of D3 of the peptide with R145 of VcSiaP gives rise to the possibility that the loops could interfere with this selectivity filter to release the substrate.
Furthermore, the loop connecting helix 1 and helix 2 was manually optimized. This resulted in stronger hits on the peptide scan. Some of the introduced changes could be explained with the peptide-bound crystal structure. But a more in-depth explanation requires a crystal structure with the peptide binders.

3.2.3. The interaction of VcSiaP with the peptides is very weak

To get a rough impression of the binding strength, a competitive experiment with 1mM Neu5Ac was performed on identical membranes. Most peptides showed lower affinity under these experimental conditions, which indicated that the peptides definitely bind weaker than the substrate Neu5Ac. Attempts to capture the interaction with PELDOR or to measure the K_D with MST for the strongest identified binders were unsuccessful. However, Niels Schneberger was able to estimate the K_D of the strongest binder within the range of 30 µM [172]. This is a K_D which is very low for drug target optimization [173].

3.3. Interaction between VcSiaP and VcSiaQM

In this work, it has been shown that without binding to the substrate, VcSiaP resides in an open state. It has been proposed that this open substrate-unbound state is not recognized by the transmembrane domains QM. The closed SBP on the other hand is supposed to have a conformation which allows to form a complex with the transmembrane components. Upon complex formation the substrate is then passed

on to the transporter. The interaction with the TM lowers the binding affinity of the substrate to the SBP, which is how the substrate is released. As explained in the previous section, the loops of VcSiaQM could act as a scoop loop to aid substrate release. The remaining question is: Which of the loops is involved?
In the crystal structure with the bound His_6-tag peptide the selectivity filter residue R145 interacted with D3 of the peptide. This pivotal interaction lead to the search for conserved glutamates (E) and asparatates (D) in the periplasmic loops of the QM-domains. Several residues were identified and further tested in an transport assay on the overall impact on substrate transport.

3.3.1. Are the glutamates in the QM-loops relevant for transport?

For the transport assay, an *E. coli* ΔnanT strain was used, which was no longer able to grow solely on Neu5Ac. This property was reintroduced into the strain with a plasmid containing VcSiaPQM. The individual mutants VcSiaQ E47A, E48A, E47A/E48A, E119A and VcSiaM E74A, E237A, E293A and E420A were introduced into the plasmid and then transformed into the *E. coli* ΔnanT strain. Growth on Neu5Ac as a sole carbon source indicates functionality of the transporter, whereas lack of growth shows the disruption. Then the mutants were observed for growth in M9-media supplemented with Neu5Ac.
Of the Q-domain mutants, E47A showed no growth, whereas the E48A mutant grew normally. Consistent with this result, the E47A/E48A double mutant did also not grow. The other mutations of the Q-domain E48A and E119A grew normally.
Of the M-domain only E237A showed no growth, the other mutations E74A, E293A, E420A grew normally on Neu5Ac.

3.3.2. VcSiaP binds to periplasmic loops on the peptide scan

Another peptide scan was designed onto which the periplasmic loops were synthesized. Additionally, peptides were added to the scan in which the important glutamates of the loops were replaced by alanine. Here a very interesting "binding-pattern" was found: all of the loop peptides with the introduced mutant bind with stronger affinity than the original loops with the residue glutamate. The only exception to this is the transmembrane loop connecting helix 6 and helix 7 of the transporter.
Two conserved hydrophobic residues were identified within this loops, mutated (F79E and A80E) and also checked in the transport assay. Both mutants showed no growth on Neu5Ac in the assay, which indicates their importance for the transporter.

Repeating the experiment with fluorescent HiSiaP instead of VcSiaP showed a dractic decrease in binding, but also gave rise to a different binding pattern. This hints that the interaction between VcSiaP and VcSiaQM is very specific. Exchanging the SBP of a system with a homologous SBP is not going to trigger the interaction with the transmembrane domains. This is in agreement with experiments from Mulligan et al. where VcSiaP failed to deliver Neu5Ac in a protoliposome assay.

3.3.3. Do the mutants grow with different speed?

All mutations were also checked for different growth rates in M9 media, but only two categories emerged here: no growth or growth at the same rate within the errors of the experiment.

3.3.4. On the hunt for the interaction between P- and QM-domain

An interaction between the P- and QM-domains of some form has to take place. How else is the substrate parted from the SBP and transfered to the QM-domains. The experimental evidence however has not yet been delivered. This question was approached with *cw*-EPR: upon complex-formation, broadening of the spectra is expected.
The mutant VcSiaP Q245C was labeled with MTSSL and subjected to *cw*-EPR measurements with excess VcSiaQM and with excess VcSiaQM and Neu5Ac. No signs of binding or interaction were visible in the spectra.
Transferring the principle of the experiment to the more rigid system VcSiaP D241-Q245Rx did not lead to a visualization of the interaction in the spectrum.
Thus a different experiment is needed to prove the interaction between these domains.

3.4. Outlook

TRAP transporters are an interesting target for the development of antibiotics. They are widespread among various procaryotic pathogens but are absent in the human body.
This thesis proved that the P-domain of the transporter can bind to peptides, which opens up a path for the search of peptide inhibitors. The peptide screens have shown that with only 43 tested peptides an increase in affinity could be accomplished already. Screening approaches such as Bac2Hybrid assays might be efficient ways to further optimize the inhibition of the P-domain.

4. Material and Methods

4.1. Consumables and material

4.1.1. Chemicals

In table 4.1 the commonly used chemicals for this thesis and their suppliers are listed. The table is divided into general chemicals, various kinds of label, supplements for bacterial growth and inducers, material for DNA work and antibodies. The chemicals were stored according to the manufacturer's guidelines.
Some chemicals were directly available through the main chemical supply of the University of Bonn (Zentrale Chemikalienversorgung Endenich (ZVE)). All biochemical experiments were performed in laboratories which comply S1 safety regulations. Waste disposal was handled according to these regulations.

Table 4.1. – List of chemicals and their suppliers.

Chemical	**Source**
General chemicals	
Acetic acid	VWR Chemicals
Acrylamide (Rotiphorese Gel 30)	Roth
Agar	AppliChem
Agarose	Roth
Albumin Fraktion V (BSA)	Roth
Ammonium persulfate (APS)	Roth
β-DDM	Roth
β-Mercaptoethanol (BME)	Roth
Bromophenol blue	Roth
$CaCl_2$	Roth
Coomassie Brillant Blue R250	Roth
D_2O	Deutero

Ethylenediaminetetraacetic acid (EDTA)	Roth
Ethanol	ZVE
Glycine	Roth
Glycerol	Roth
Imidazole	Roth
KH_2PO_4	Roth
KCl	Roth
Li_2SO_4	Roth
Milkpowder	Roth
$MgSO_4$	Roth
Methanol	ZVE
NH_4Cl	Roth
$NiSO_4$	Fisher Scientific
Na_2HPO_4	Roth
NaOH	Roth
NaCl	Roth
Sodium cacodylate	Roth
Sodium dodecyl sulfate (SDS) solution (20%)	AppliChem
Trisodium citrate	Roth
Tween 20	Roth
TMB-Blotting Solution (1-Step Ultra)	ThermoFisher Scientific
TCEP	Roth
TEMED	Roth
Tryptone	Roth
Tris	Roth
TES PUFFERAN	Roth
Triton X-100	Roth
Xylene cyanol	Roth
Yeast extract	Roth
Label for spectroscopy	
(1-Oxyl-2,2,5,5-tetramethyl-δ3-pyrroline-3-methyl) Methanethiosulfonate (MTSSL, R1)	Toronto Research Chemicals

3,4-Bis-(methanethiosulfonylmethyl)-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-1-yloxy Radical (Rx)	Toronto Research Chemicals
Fluorophore labels	
DyLight 800 Maleimide	ThermoFisher Scientific
Fluorescein-5-Maleimide	ThermoFisher Scientific
NHS-Fluorescein (5/6-carboxyfluorescein succinimidyl ester)	ThermoFisher Scientific
Biotin label	
EZ-Link Maleimide-PEG2-Biotin	ThermoFisher Scientific
Antibiotics	
Ampicillin sodium salt	Roth
Kanamycinsulfate	Roth
Chloramphenicol	Roth
Inducers	
Isopropyl-β-D-thiogalactopyranosid (IPTG)	Carbolution
L-(+)-Arabinose	Roth
Supplement for bacterial growth	
N-acetylneuraminic acid (Neu5Ac)	Carbosynth Limited
D-(+)-Glucose	Roth
DNA work	
Ethidium bromide solution (1%)	Applichem
GeneRuler DNA Ladder	ThermoFisher Scientific
6× loading buffer	
Protein work	
PageRuler Plus Prestained Protein Ladder	ThermoFisher Scientific
Antibodies	
Mouse α 6x-His Epitope Tag	Pierce
Goat anti-Mouse IgG (H+L) Cross-Adsorbed Secondary Antibody	Pierce
Goat anti-biotin, HRP-linked Antibody S7075	NEB

4.1.2. Peptides

Tabel 4.2 summarizes all peptides which were used to study the interaction between VcSiaP and VcSiaQM. The peptides were dissolved in water by shaking, heating and sonification.

Table 4.2. – List of used peptides.

Name	Peptidesequence	Source
P1	SLWSEELARL	peptides&elephants
P2	SLWSAALARL	peptides&elephants
P3	PSLWSKEKVRLLF	peptides&elephants
P4	NFAKSLVGH	peptides&elephants
P5	PSLWSEEKVRLLF	peptides&elephants
P6	DYDIPTTENLYFQ	peptides&elephants

4.1.3. Instruments and columns

Table 4.3 summarizes the hardware used to perform experiments and gather data. Table 4.4 lists the columns used with the chromatographs for protein purification.

Instruments

Table 4.3. – List of used instruments.

Instrument	Tradename	Manufacturer
	EPR instruments	
EPR spectrometer	Elexys	Bruker Corporation
EPR spectrometer	EMXnano	Bruker Corporation
EPR spectrometer	EMXmicro	Bruker Corporation
Q-band resonator	ER 5106QT-2	Bruker Corporation
X-band resonator	ER4119HS	Bruker Corporation

Helium cryostat	CF935	Oxford instruments
Temperature control	ITC 502	Oxford instruments
TWT amplifier	150 W	Applied Systems Engineering
Biolaboratory instruments		
Cell disruptor	Cell disruptor Benchtop	Constant Systems Limited
Waterdesalinator	Milli-Q Direct	Merck Millipore
+4 °C fridge	-	Gram BioLine
-20 °C freezer	-	Liebherr
-80 °C freezer	-	New Brunswick Scientific
Blotting cell	Trans-Blot Semi-Dry Transfer Cell	Bio Rad
Gel elctrophoresis cell	Mini-PROTEAN Tetra Cell	Bio Rad
Chromatograph	Äkta pure	GE Healthcare
Chromatograph	Äkta avant	GE Healthcare
Incubator #1	I26	New Brunswick Scientific
Incubator #2	Ecotron	Infors HT
Ultracentrifuge	Optima XPN	Beckman Coulter
Thermocycler	Mastercycler nexus	Eppendorf
Centrifuge #1	5810R	Eppendorf
Centrifuge #2	5424R	Eppendorf
Centrifuge #3	Avanti J-26 XP	Beckman Coulter
Table top centrifuge	Mini Star	VWR
Rotor #1	JA-25.50	Beckman Coulter
Rotor #2	JLA-8.100	Beckman Coulter
Rotor #3	50.2 Ti	Beckman Coulter
Balance #1	PCB	Kern
Balance #2	AES	Kern
Block heater	Thermomixer comfor	Eppendorf
Spectrophotometer	NanoDrop 2000	ThermoFisher Scientific
Autoclave	Laboklav ECO	SHP Steriltechnik

SPR Analyzer	Biacore 3000	GE Healthcare
MST instrument	Monolith NT.115	NanoTemper Technologies
Ultrasonic bath	USC100T	VWR
Odyssey Imaging	-	LI-COR Biotechnology
Gel Imaging System	-	Peqlab
pH meter	pHenomenal	VWR
Shaker	Vortex Genie 2	Scientific Industries
Microwave	-	Clatronic
Bunsen burner	Bochem	Fisher Scientific

Columns

The listed columns were used on the chromatographs during protein purification. Additional to that a 10 mL on-bench column equipped with a frit was used which was filled with Ni-NTA resin. Also a PD-10 column was used to remove excess dyes after labeling of protein. All columns were stored in 20% ethanol.

Table 4.4. – List of used columns.

Columns for chromatography systems	**Manufacturer**
ENrich Q 10/100	Bio-Rad
HiLoad 16/600 Superdex 75 prep grade	GE Healthcare
HiLoad 16/600 Superdex 200 prep grade	GE Healthcare
Superdex 200 10/300 GL	GE Healthcare
Other columns	**Manufacturer**
PD-10	GE Healthcare
Polystyrene column with Ni-NTA	-

4.1.4. Kits and other material

In table 4.1.4 and table 4.6 additional and disposable materials are listed.

Kits

Table 4.5. – List of used kits.

Kit	Manufacturer
Molecular biology	
QIAquick R PCR Purification Kit	Qiagen
GeneJET Plasmid Miniprep Kit	ThermoFisher Scientific
Crystallography	
JCSG	Molecular Dimensions
SPR	
Biotin CAPture Kit	GE Healthcare

Other materials

Table 4.6. – List of additional instruments and materials.

Material/Instrument	Manufacturer
Crystallography	
96-well plate	Jena Bioscience
Sealing foil	Roth
EPR	
X-band tube	Wilmad
Q-band tube	Wilmad
10 µL cappilary	Hirschmann
Superglue	Uhu
Bacterial growth	
Schikane flask 1 L	Thomson
Schikane flask 50 mL	Thomson
Schikane flask 25 mL	Thomson
Petri dishes	Sarstedt

MST	
Capillaries (Monolith NT.115)	NanoTemper
Protein purification	
Ni-NTA resin	Nalgene
On bench column	Thomon
PVFD membranes	Roth
Blotting papers	Roth
Concentrator Vivaspin 2/10.000 mwcoPES	Sartorius
Concentrator Vivaspin 6/10.000 mwcoPES	Sartorius
Concentrator Vivaspin 20/10.000 mwcoPES	Sartorius
Reaction tubes (1.5 mL)	Eppendorf
Reaction tubes (15 mL, 50 mL)	Sarstedt
Filtropur BT50 filter	Sarstedt

4.1.5. Buffers

All listed buffers, which were used on chromatographs or for SPR experiments were sterile filtered with a Filtropur BT50, 0.2 µm prior to usage.

Table 4.7. – List of used buffers and their field of application.

Name	Composition	Application
Transferbuffer (Towbin)	25 mM Tris (pH= 8.3 - 8.5) , 192 mM Glycin, 0.05% SDS	Westernblot
TBST (Roche)	50 mM Tris, 150 mM NaCl, 0.1% Tween 20	Westernblot
Blocking Solution #1	5% (w/v) Milkpowder, low fat in TBST	Westernblot
Blocking Solution #2	1% (w/v) BSA in other buffer A	Peptide Scan
4 × SDS-loading dye (reducing)	62.5 mM Tris, 10% Glycerol, 20% SDS (10%), 10 % 2-Mercaptoethanol, 1 tip Bromophenolblue, pH 6.8	SDS-PAGE

8 × SDS-loading dye (non-reducing)	125 mM Tris, 20% Glycerol, 40% SDS (10%), 1 tip Bromophenolblue, pH= 6.8	SDS-PAGE
Separating gel buffer	1.5 M Tris pH= 8.8, 10% SDS, 12% Acrylamide or 15% Acrylamide	SDS-PAGE
Stacking gel buffer	1.5 M Tris pH= 8.8, 10% SDS, 5% Acrylamide	SDS-PAGE
SDS-PAGE Running buffer	25 mM Tris, 192 mM Glycine, 0.1% SDS, pH= 8.3	SDS-PAGE
Coomassie Stain	45% Methanol, 10% Acetic acid, 2.5 g/L Coomassie Brilliant Blue R-250	SDS-PAGE
6 × Loading buffer	10 mM Tris, 60 mM EDTA, 60% Glycerole, 0.03% Xylene cyanol, 0.03% Bromophenol blue	Agarose gel
TAE buffer	40 mM Tris, 20 mM Acetic acid, 1 mM EDTA	Agarose gel
PBS buffer	137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4 pH= 7.4	Protein labeling
PELDOR buffer	100 mM TES, 100 mM NaCl,in D2O,pH= 7.5	EPR
Buffer A	50 mM Tris, 50 mM NaCl pH= 8.0	Protein purification
Reduction buffer	Buffer A with 1 mM TCEP	Protein purification
Elution buffer	Buffer A with 250 mM Imidazole	Protein purification
Ion Exchange buffer A	50 mM Tris, pH= 8.0	Protein purification
Ion Exchange buffer B	50 mM Tris, 1 M NaCl, pH= 8.0	Protein purification
Reducing Ion Exchange buffer A	Ion Exchange buffer A with 1 mM TCEP	Protein purification
Reducing Ion Exchange buffer B	Ion Exchange buffer B with 1 mM TCEP	Protein purification

MST buffer	Buffer A with 0.05% Triton X-100	
QM buffer	50 mM KH_2PO_4 pH=7.8, 200 mM NaCl, 20% Glycerol	Protein purification
Extraction buffer	QM buffer with 1.5% DDM	Protein purification
QM Protein buffer	QM buffer with 0.03% DDM	Protein purification
QM washing buffer	QM buffer with 0.03% DDM and 25 mM Imidazole	Protein purification
QM elution buffer	QM buffer with 0.03% DDM and 250 mM Imidazole	Protein purification

4.1.6. Enzymes and buffers

Table 4.8. – List of used enzymes and their buffers.

Enzyme/buffer	**Concentration**	**Source**
Pfu Polymerase	2.4 U/µL	Thermo Fisher Scientific
Pfu buffer	10×	Thermo Fisher Scientific
DpnI restriction enzyme	10 U/µL	Thermo Fisher Scientific
Tango buffer	10×	Thermo Fisher Scientific
TEV protease	1 mg/mL	Schiemann lab

4.1.7. Oligonucleotides

In table 4.9 all of the primers used to introduce mutations into the P-domains via quick-change mutation are listed. The triplet in which the mutation was inserted into is underlined. The overlapping sequences are highlighted in grey.

Table 4.9. – List of DNA oligonucleotides used for quick change mutation. The overlapping sequences are marked in grey. The mutated triplets are underlined.

Mutation		Primersequence 5'→3'
VcSiaP Q54C	for	CTT CAG TGC TTG ACG CTG GGA GAT CTC GAT ATA AC
	rev	CGT CAA GCA CTG AAG CAT GGC ACG ATC ACC
VcSiaP L173C	for	GTT TAT TGC GCG CGT CAG ACC AAT GCC GTA GAT G
	rev	CAG CGC GCA ATA AAC TTC AGA GAA TGA CAT CGG GG
VcSiaP Q110C	for	T GAA ATG CTC TGT AAG TTC A AC TGG CGT GCT TTG GAC AC
	rev	T GAA CTT ACA GAG CAT TTC A TC ACG AAC ACC TTG ACC AAA G
VcSiaP S225C	for	G ATC ATT TCT GAA TCT ACT TGG CAG AAG CTT TGT GAT ACG GAT
	rev	CAC GGC TTT CTG AAT GAT GTC TTT ATC CGT ATC ACA AAG CTT
VcSiaP Q245C	for	CAT ACA TGC ACC GTT AAA ACC CAA GAG GAA GAA TTG
	rev	AAC GGT GCA TGT ATG AGC ATC TCC CAC TTT CTG
VcSiaP R125A	for	GGT ACC GCT GAA ACC AC T TCA AAC CGT CCC CTC AAT TCG
	rev	GT GGT TTC AGC GGT ACC GTT ATA CCA AGT GTC CAA AGC AC
VcSiaP R125K	for	GGT ACC AAA GAA ACC AC T TCA AAC CGT CCC CTC AAT TCG
	rev	G TGG TTT CTT TGG TAC C GT TAT ACC AAG TGT CCA AAG CAC

VcSiaP E184A	for	GGG CAA GCA AAC CCG TTA CCA ACA ATT AAA ACG ATG AAG TTC TAT
	rev	CGG GTT TGC TTG CCC ATC TAC GGC ATT GGT CTG CAG
VcSiaP E184Q	for	GGG CAA CAA AAC CCG TTA CCA ACA ATT AAA ACG ATG AAG TTC TAT
	rev	CGG GTT TTG TTG CCC ATC TAC GGC ATT GGT CTG CAG
VcSiaP E184D	for	GGG CAA GAC AAC CCG TTA CCA ACA ATT AAA ACG ATG AAG TTC TAT
	rev	CGG GTT GTC TTG CCC ATC TAC GGC ATT GGT CTG CAG
VcSiaP H207A	for	ATG ACA CAT GCT ATC GTT AAC GAT CAA ATG GTG ATC ATT TCT
	rev	AAC GAT AGC ATG TGT CAT GGC TAA GTT CTT TTG CAC TTC
VcSiaP H207Q	for	ATG ACA CAT CAA ATC GTT AAC GAT CAA ATG GTG ATC ATT TCT
	rev	AAC GAT TTG ATG TGT CAT GGC TAA GTT CTT TTG CAC TTC
VcSiaP H207N	for	ATG ACA CAT AAT ATC GTT AAC GAT CAA ATG GTG ATC ATT TCT GAA TCT
	rev	GTT AAC GAT ATT ATG TGT CAT GGC TAA GTT CTT TTG CAC TTC ATA
VcSiaQ E47A	for	TG TGG AGT GCA GAG CT T GCC CGA TTA CTC TTC ATG TAT ATG
	rev	AGC TCT GCA CTC CAC A AA GAA GGA TCA TTG AGA AGC CAAC
VcSiaQ E48A	for	TG TGG AGT GAA GCG CT T GCC CGA TTA CTC TTC ATG TAT ATG

	rev	AGC GCT TCA CTC CAC A AA GAA GGA TCA TTG AGA AGC CAA C
VcSiaQ E47/48A	for	G CTT CTC AAT GAT CCT TCT TTG T GG AGT GCA GCA CTT G
	rev	CAT ATA CAT GAA GAG TAA TCG GG C AAG TGC TGC ACT CC
VcSiaQ E119A	for	TTC TTT GCA TTG ATC ACA TTA GGG ATA TCG AGT AGC TGG ATG
	rev	TAA TGT GAT CAA TGC AAA GAA TGC ATT ACG CTG GGC ATG TTG ATA
VcSiaM E74A	for	GGT ATT ACA GCA AGA ATC TTC AAC TTC GCG AAG TCT CTC GTT G
	rev	GAA GAT TCT TGC TGT AAT ACC GGC ACT ATT CAT TAG ATG GCC AG
VcSiaM E237A	for	CCA ACT GCA GCG GCT GCT GTT TCA TCT CTC TAC GCA CTA TTT
	rev	AGC CGC TGC AGT TGG CGT AAA TTT ACC AGA GAA GAT CCC AC
VcSiaM E293A	for	GCT CGC GCA CAG TTA CCA CAG ATG TTG GCT GAT TAT TTC TTA
	rev	TAA CTG TGC GCG AGC CAC AAT CCA CCC AAA AAC AGT C
HiSiaP V250C/K254C	for	TTC TGT GAT GGA GAG TGT GAT TTA GTC ACA TTC TTT GAA AAA CAA GGC
	rev	ACA CTC TCC ATC ACA GAA TAA TTT AGT GTG ATA TTT TGC TGC ATT TTC GG
HiSiaP K254C	for	CA AAA TAT CAC ACT AAA TTA TTC GTA GAT GGA GAG TGT GAT TTA GTC
	rev	CAC GCC TTG TTT TTC AAA GAA TGT GAC TAA ATC ACA CTC TCC

4.1.8. Plasmids

Table 4.10 lists all plasmids used in the course of this thesis. The plasmid pBADHis-TEV was a gift from Liu. The plasmid pWKS30 was a gift from E. Severi.

Table 4.10. – List of plasmids and their application.

Name	Application	Source
pBADHisTEV	Protein expression	Naismith lab, St. Andrews
pWKS30	Bacterial growth experiments: negative control	E. Severi from Thomas lab, York
pES85	Bacterial growth experiments	E. Severi from Thomas lab, York

4.1.9. Bacterial strains

For plasmid preparation PCR products were transformed and grown in D5hα cells. The expression of periplasmic protein was performed in the C43 cell line, which was specially designed for membrane proteins. The expression of the QM-domains was performed in the cell line MC1061. All monitored bacterial growth experiments (see chapter 4.3) were carried out in Sevy3 cells, which was a gift from the Thomas group at the University of York. In Sevy3 the sialic acid transporter ΔnanT is knocked out and all sialcis acid catabolic genes are constitutivly activated through knocked-out repressor genes (ΔnanT, ΔnanR, ΔnagC).

Table 4.11. – Bacterial strains for mutations, protein expression and growth experiments.

Name	Genotype	Source
D5hα	F– endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20 ϕ80dlacZδM15 δ(lacZYA-argF)U169, hsdR17(rK–mK+), λ–	Invitrogen
C43	F –ompT hsdSB (rB- mB-) gal dcm (DE3)	Lucigen
Sevy3	F- LAM- rrnB3 DElacZ4787 hsdR514 DE(araBAD)567 DE(rhaBAD)568 rph-1 ΔnanT ΔnanR ΔnagC	E. Severi from Thomas lab, York
BL21(DE3)	B F– ompT gal dcm lon hsdSB(rB–mB–) λ(DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5]) [malB+]K-12(λS)	New England Biolabs
MC1061	araD139 Del(araA-leu)7697 Del(lac)X74 galK16 galE15(GalS) lambda- e14- mcrA0 relA1 rpsL150(strR) spoT1 mcrB1 hsdR2	Thomas lab, York

4.1.10. Media and supplement for bacterial culture

The media listed in table 4.12 was sterilized by autoclaving prior to use. Depending on the purpose various additives are added to the M9 media protocol. M9 media used for protein expression was further supplemented with 50 mL/L glycerol. M9 media for the bacterial growth experiments was supplemented with either 1 mg/mL Neu5Ac or 1 mg/mL D-(+)-glucose (as positive control).

Table 4.12. – List of media and agar for bacterial growth and washing of cells.

Name	Composition	Application
Luria-Bertani (LB) medium	10 g/L Typtone, 5 g/L Yeast extract, 10 g/L NaCl, pH=7.0 with NaOH	Growth of *E. coli*
LB agar	Composition of LB media with additional 15 g/L Agar	Growth of *E. coli*
M9 medium #1	13.4 g/L Na_2HPO_4, 6 g/L KH_2PO_4, 1 g/L NaCl, 2 g/L NH_4Cl, 100 µL $CaCl_2$ (1 M), 2 mL $MgSO_4$ (1 M) pH=7.4 with NaOH	Growth of *E. coli*
M9 medium #2	Composition of M9 media with additional 50 mL glycerol	Washing of *E. coli*, growth of *E. coli*
M9 agar	Composition of M9 media with additional 15 g/L Agar and Neu5Ac (1 mg/mL) or D-(+)-glucose (1 mg/mL)	Growth of *E. coli*

Table 4.13. – Antibiotics and supplements.

Name	Stock concentration	Application
$MgSO_4$	1 M	M9 media supplement
$CaCl_2$	1 M	M9 media supplement
Ampicillin	100 mg/mL	*E. coli* selection
Kanamycin	50 mg/mL	*E. coli* selection
Chloramphenicol	34 mg/mL	*E. coli* selection
IPTG	1 M	Protein expression
L-(+)-Arabinose	500 mg/L or 50 mg/L	Protein expression

4.1.11. Bioinformatic tools and other software

Each instrument software and the respective analysis software used in this thesis are listed in table 4.14.
DNA and protein sequences for primer design and determination of protein parameters were obtained from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database.
Oligonucleotides for QuikChange mutation were designed with Geneious. Also, analysis of sequencing results was performed with this program.
To calculate physico-chemical properties of proteins, such as molecular weight and extinction coefficient, the ProtParam tool of the ExPASy server was used.
Protein concentrations were calculated from the absorption of the sample at 280 nm measured with a NanoDrop spectrophotometer, using the respective extinction coefficient according to the Lambert-Beer law.

Table 4.14. – List of used computer software, plug-ins and databases.

Software	Source	Application
	EPR-Spectroscopy	
MATLAB R2012b	MathWorks	EPR analysis
DEERAnalysis 2016	EPR Group, ETH Zurich	EPR analysis
DEERAnalysis 2018	EPR Group, ETH Zurich	EPR analysis
Xenon 1.1b.108	Bruker	*cw*-spectra recording and analysis
Xenon nano 1.2a.2	Bruker	*cw*-spectra recording and analysis
Xepr 2.6b.151	Bruker	PELDOR recording
mtsslWizard	Gregor Hagelüken, University of Bonn	EPR analysis
	Databases and online tools	
KEGG	Kanehisa lab, Kyoto University	Genome database
T-Coffee	Notredame lab, CRG Barcelona	Multiple sequence alignments

ProtParam	Swiss Institute of Bioinformatics, Lausanne	Prediction of protein properties
	MST	
MO.Affinity Analysis	NanoTemper Technologies	Analysis of MST data
	SPR	
BIAevaluation	GE Healthcare	Analysis of SPR data
	General software tools	
Gnuplot	Open Source	Fitting and plotting of data
Texmaker	Pascal Brachet, Joël Amblard	Thesis writing
Inkscape	Open Source	Graphic drawing
	FPLC software	
Unicorn 7	GE Healthcare	Operation of Äkta Avant
Unicorn 6	GE Healthcare	Operation of Äkta Pure
	Molecular biology	
Geneious	Biomatters	Primer design, analysis of sequencing results
Image Studio Lite 5.0	LI-COR Biosciences	Analysis of the peptide scan
	Crystallography	
COOT	Paul Emsley, MRC lab, Cambridge	Structure refinement
CCP4	Rutherford lab, Harwell Campus, Oxfordshire	X-ray data analysis and structure solving
Ligplot [174]		Ligand-protein interaction analysis
Phaser	Read lab, Cambridge	Structure refinement
Pymol	Open Source, maintained by Schrödinger	Structure visualization

4.2. Molecular biology

4.2.1. QuikChange Mutagenesis

Site specific point mutations were introduced in the constructs via QuikChange mutagenesis as described by Liu [175]. The pipetting scheme for the mutagenesis is summarized in table 4.15. The primers with the desired mutation and the corresponding temperatures are displayed in table 4.9 and table 4.17. Table 4.16 summarizes the protocol for the polymerase chain reaction (PCR).

Table 4.15. – Pipetting scheme for QuikChange mutagenesis.

Component	**Control-sample [µL]**	**Sample [µL]**
DNA template	-	0.5
Pfu buffer	5	5
dNTPs (10 mM)	1	1
Forward Oligo(100 µM)	0.5	0.5
Reverse Oligo(100 µM)	0.5	0.5
Pfu Polymerase	0.5	0.5
H_2O	41.5	42
$\sum$	50	50

Table 4.16. – PCR protocol for QuickChange mutagenesis.

No. of Cycles	**Temperature [°C]**	**Time [s]**	**Process**
1	95	300	Initial denaturation
3	95	60	Denaturation
	$T_{m,no}$-5	60	Annealing
	72	900	Elongation
15	95	60	Denaturation
	$T_{m,comb} - 5$	60	Annealing
	72	900	Elongation
2	95	60	Denaturation
	43	60	Annealing
	72	900	Elongation

In table 4.16 $T_{m,no}$ is the non-overlapping primer region and $T_{m,comb}$ is the overlapping primer region.

Table 4.17. – Temperatures for QuickChange mutagenesis

Inserted mutation	$\mathbf{T_{m,no}}$	$\mathbf{T_{m,comb}}$
VcSiaP Q54C	49.7	41.9
VcSiaP L173C	51.5	41.9
VcSiaP Q110C	52.4	45.6
VcSiaP S225C	54.4	43.5
VcSiaP Q245C	48.2	39.2
VcSiaP R125A	54.4	44.7
VcSiaP R125K	53.5	44.6
VcSiaP E184A	53.5	49.5
VcSiaP E184Q	53.5	49.5
VcSiaP E184D	53.9	47.4
VcSiaP H207A	51.4	43.5

VcSiaP H207Q	53.2	44.6
VcSiaP H207N	50.5	38.9
VcSiaQ E47A	49.2	40.9
VcSiaQ E48A	49.2	40.9
VcSiaQ E47/48A	46.7	39.7
VcSiaQ E119A	49.4	39.6
VcSiaM E74A	50.0	43.5
VcSiaM E237A	51.4	42.4
VcSiaM E293A	49.1	39.7
HiSiaP V250C K254C	51.5	40.8
HiSiaP K254C	51.6	45.8

4.2.2. Agarose Gel Electrophoresis

Obtained PCR products were checked via agarose gel electrophoresis with 1% agarose gels. The gels were prepared by adding agarose to TAE buffer and heating in the microwave until the mixture was fully melted. The solution was then supplemented with 2 µL ethidiumbromide solution and placed in the gel tray for cooling.
The samples were prepared by adding 1 µL of 6× Loading buffer to 5 µL of PCR product. The gel was run for 30 min at 100 V and then visualized by the gel imaging system.

4.2.3. Restriction Enzyme Digestion

To remove the non-mutated DNA template from the PCR product, it had to be incubated with DpnI. 45 µL of PCR product was supplemented with 5 µL TANGO reaction buffer and 0.5 µL of DpnI enzyme and then incubated for 2 h at 37°C.
Then the enzymes and primers were removed using the PCR purification kit according to manufacturer's instructions. Now the PCR product was ready to be transformed into competent cells.

4.2.4. Preparation of competent cells

The cells used in this study (see table 4.11) have been grown and made competent in the laboratory of Prof. Schiemann.
For preparation of competent cells 5 mL of LB-medium was inoculated with 1 colony of the plated cell from the respective agar-plate and grown over night at 37°C while being gently shaken. 1 mL of this bacterial preculture was then used to inoculate 100 mL LB-media. This culture was incubated for about 2 h at 37°C until an OD_{600} of 0.4-0.6 was reached. This culture was then chilled on ice for 30 min and transferred to two prechilled 50 mL falcon tubes. The bacterial suspension was centrifuged for 5 min at 4°C and 2800 rpm. After discarding the supernatant, each pellet was resuspended in prechilled 12.5 mL $CaCl_2$ (100 mM) and 12.5 mL MgSO4 (40 mM) and again chilled on ice for 30 min.
The suspension was centrifuged again at 2800 rpm for 5 min at 4°C and the remaining pellets resuspended in 2.5 mL $CaCl_2$ (100 mM) and 2.5 mL MgSO4 (40 mM). Prechilled, sterile glycerol was added to 10% of total volume and the cells were aliquoted into 50 µL aliquots. The aliquots were flash-frozen with liquid nitrogen and stored at -80°C.
To check the cells for contamination they were streaked onto LB-agar plates without antibiotic, with kanamycin, ampicillin and with chloramphenicol and grown over night at 37°C.

4.2.5. Transformation into competent cells and isolation of plasmid DNA

For transformation, 50 µL of competent cells were incubated on ice with 1-5 µL (1-100 ng) PCR product for 20 min. The cells were heat-shocked for 45 s at 42°C and then immediately put on ice for another 5 min. After adding 1 mL of LB-media the cells were incubated for 1 h at 37°C. Finally 100 µL cell suspension were plated on LB plates supplemented with selection antibiotic and incubated over night at 37°C.
For small-scale plasmid preparation, 10 mL LB-medium supplemented with respective antibiotic were inoculated with a selected single colony and grown over night at 37°C. Plasmid DNA was isolated using the GeneJET Plasmid Miniprep Kit according to manufacturer's instructions. To verify the mutation, plasmids were send to Microsynth SEQLAB (Göttingen) for sequencing.

4.2.6. Glycerol stocks

To conserve the right plasmid-bacterial colony combination for protein expression, glycerol stocks were prepared.
First an overnight culture was prepared with 25 ml LB media supplemented with 25 µL ampicillin and inoculated with a single colony picked from the previously prepared agar plates. This mixture was incubated over night at 37°C under constant shaking at 180 rpm.
Then, 1120 µL of the preculture were mixed with 280 µL of glycerol (100%, sterile), flash frozen and stored at -80°C.

4.3. Bacterial growth experiments

4.3.1. M9 plate growth

All experiments performed in this section were carried out as sterile as possible. The media and pipette tips were sterilized by autoclaving. The flasks and well-plates were sterilized by boiling water in the microwave. The agar was sterilized by heating in the microwave. All instruments were flamed with a Bunsen burner prior to usage and transfer of sterile liquids was performed next to the Bunsen burner flame.

Preparation of M9 plates

250 mL of M9 media were filled in a 500 mL bottle and supplemented with 5 µL $CaCl_2$ (1 M) and 1 mL $MgSO_4$ (1 M), which leads to precipitation. The bottle was shaken vigorously to evenly distribute the precipitant. After adding agar the mixture was boiled in the microwave until everything was fully dissolved. After letting the hot M9 agar cool for some time, 30 mL was transferred to a 50 mL falkon tube and left for cooling a little further.
Meanwhile 30 mg of sugar (either D-(+)-glucose or Neu5Ac) was dissolved in 1 mL of M9 media, supplemented with ampicillin (30µL for 30 mL of M9 agar) and put on ice, until the agar was cold enough.
The sugar-ampicillin mix was added to the agar and mixed gently by inverting the tube for a few times.
15 mL of agar were poured into an empty and sterile agar plate, yielding two plates. After solidification of the agar 15 µL of IPTG is added to 70 mL of M9 media and plated on the surface of one plate until it is fully soaked up. This has to be repeated for the second plate.

Preparation of cells and distribution on plates

Glycerol stocks of the mutants were prepared as described in section 4.2.6. The constructs consisted of mutations in the VcSiaPQM genes on the pES85 plasmid. They all were transformed into Sevy3 cells.
The tip of a sterile pipette was dipped in the glycerol stock and the transferred amount of cells were dissolved in an Eppendorf tube with 1 mL of M9 media. This step was important because it diluted the contamination with glycerol. The tube was then shaken (Vortexer) to properly distribute the cells. A streaking loop was flamed with the Bunsen burner, dipped in the cell dilution and the transferred cells were distributed on a segment of the M9 plate.
The finished plate was placed into an incubator at 37° C and left to incubate for two days. After the incubation time of two days, a picture was taken with the gel imaging system.

4.3.2. M9 media growth

25 mL LB preculture was supplemented with 25 µL ampicillin and inoculated with cells from the glycerol stock of the investigated mutant. The preculture was grown at 37° C for 8 h at 180 rpm. The cells were centrifuged at 4000 rpm for 10 min and then resuspended in M9 media. This washing step was repeated for two more times.
Then a M9 preculture was prepared. 25 mL M9 media was supplemented with 200 µL glycerol (0.4% v/v) and inoculated with 250 µL of the washed cells. The second preculture was grown at 37° C for 10 h at 180 rpm. The cells were washed three times and the main culture was prepared.
The main culture consisted of M9 media supplemented with ampicillin (100 µg/mL), IPTG (1 mM) and either glucose (positive control, 1 mg/mL) or Neu5Ac (1 mg/mL) and cells that were diluted to an OD(600) = 0.1 − 0.2.
Cell growth was monitored in a 24 well plate. Three of the four wells in a row were filled with 1 ml culture of one mutant supplemented with Neu5Ac and the last well was filled with culture supplemented with glucose. The well plate was covered to avoid evaporation. Growth was monitored by measuring the OD(600) every hour for 10 h. The OD(600) of the triplets was averaged and plotted against the elapsed time t.
The growth experiments were repeated in 25 mL schikane flasks and the growth was monitored for 24 h. Again, triplets of each mutant were monitored and the OD(600) averaged.

4.4. Proteomics

4.4.1. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis

To monitor the purification progress and the purity of the proteins, analysis by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed. Depending on the size of the protein, 12% or 15% polyacrylamide gels were prepared. All buffers mentioned in this chapter are listed in table 4.7.

Preparation of gels

To prepare gels for electrophoresis, 5 mL of separating gel buffer (either 12% or 15%) was supplemented with 50 µL APS (10%) and 2 µL TEMED, then transferred into the gel casket and topped with 70% ethanol to ensure a clear border. After polymerization of the running gel, 1.5 mL of stacking gel buffer was supplemented with 10 µL APS (10%) and 1 µL TEMED, poured on top and interspersed with a comb.
After polymerization, the comb was removed, the finished gel was placed in the running casket which was then filled with SDS-PAGE running buffer. The protein sample was then loaded into the pockets.

Sample preparation

To be able to properly analyze the protein samples, they had to be denatured and coated uniformly with a negative charge. This ensured that their migration in the gel matrix was proportional to their actual size.
Each loaded sample was prepared by mixing 30 µL of protein solution with 10 µL 4× SDS-loading dye and then heated for 5 min at 95 ° C.
Native gels were prepared in the same manner, but the SDS-loading dye was free of 2-β-mercaptoethanol.
Samples for membrane proteins were mixed with the regular SDS-loading dye, but not heated.

Electrophoresis

After loading the protein samples in to the pockets of the stacking matrix, the electrophoresis was performed for 50-60 min at 175 V. In order to estimate the size of proteins, a protein ladder was also loaded onto the gel and run under the same conditions. The obtained gels were stained with coomassie stain for 10 min and then washed with hot water to remove excess stain. A picture of the gel was then taken with an imaging system.

4.4.2. Western Blot

Prior to performing a western blot, a SDS-PAGE gel of the proteins of interest is needed. Additionally it should also contain a protein with the required tag (in this thesis TEV protease with His_6-tag) to indicate the correct handling of this subexperiment.
A PVFD membrane in the size of the SDS-Page gel was activated for a few seconds with methanol then washed for 5 min in distilled water and equilibrated for 15 min in transfer buffer. The SDS-PAGE gel was also equilibrated in transfer buffer for 15 min. Then four pieces of soaking paper (also in the size of the protein gel and the PVFD membrane) are soaked into transfer buffer and assembled in a semi-dry blotting cell as shown in figure 4.1.

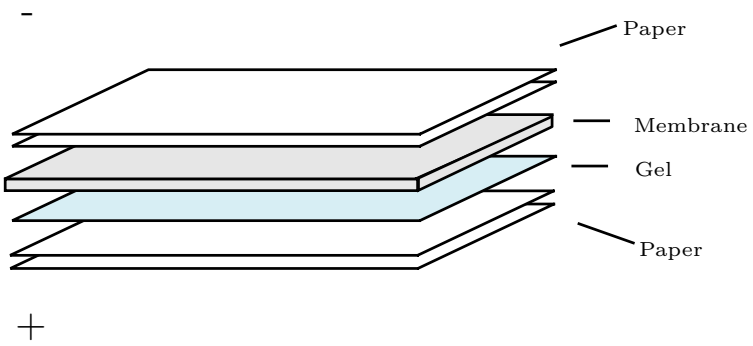


Figure 4.1. – Arrangement of protein gel, blotting paper and PVFD membrane in the blotting chamber. The flow of current is indicated.

The proteins were then blotted onto the membrane in a semi-dry transfer cell at a constant electric current of 300 mA and a maximum voltage of 25 V for 50 min.
To saturate unspecific proteins which might interfere with the antibodies, the membrane was incubated for 1 h with blocking solution #1 at room temperature. Then, the membrane was washed three times for 10 min each with TBST buffer.
The membrane was incubated over night at 4° C with 5 mL of TBST supplemented with the primary antibody (mouse anti His-tag, working dilution 1:1000). On the next day, the membrane was washed three times with 10 mL TBST and incubated again with 10 mL of TBST supplemented with the second antibody (goat anti-mouse IG, working dilution 1:5000) for 1 h at room temperature. After washing off the second antibody with 10 ml for three times again, the membrane was developed with 5 ml TMB blotting solution and the reaction was stopped with water after 30 min.

4.4.3. Protein expression

Regular expression of VcSiaP in LB media is accompanied with a contamination of the substrate Neu5Ac. This contamination leads to a significant amount of closed protein, when the sample is subjected to PELDOR experiments (see appendix section 2.1.2). To bypass this issue a protocol to express VcSiaP in M9 media was established and optimized.

VcSiaP wildtype and modifications

VcSiaP and modifications of the protein are pBADHisTEV-constructs, which were transformed into C43 expression cells.
First a preculture has to be prepared out of 50 mL LB-media supplemented with 50 µL ampicillin (100 mg/mL) and inoculated with one pipette tip of the glycerol stock of the desired construct. The preculture was grown for 12-14 h at 37° C and then centrifuged for 15 min at 4000 rpm. The supernatant was discharged and the obtained pellet was resuspended in 50 mL M9-media. This washing cycle of centrifugation and resuspension was repeated for three times to fully remove remains of the LB-media.
1 L of main culture is prepared out of 1 L M9-media supplemented with 1 mL ampicillin (100 mg/mL), 2 mL $MgSO_4$ (1 M) and 100 µL $CaCl_2$ (1 M). This is then inoculated with 5 mL washed preculture and grown at 37°C shaking at 180 rpm. After 14-16 h the culture had an OD(600) of $0.6-1$, which is optimal for induction of protein expression. The expression is induced by adding 500 mg L-(+)-arabinose. The main culture was incubated with shaking for additional 3 h until the cells were harvested.
The cells were harvested by centrifugation of the suspension for 20 min at 4000 rpm. They were flash-frozen with liquid nitrogen and stored at -80 °C.
Usually 6 L were prepared and yielded 50 mg of protein.

HiSiaP K254C

This protein was a pBADHisTEV-construct in BL21(DE3) expression cells. It was expressed following the VcSiaP protocol.

VcSiaQM

VcSiaQM is expressed entirely in LB media. A preculture was prepared out of 50 mL LB-media supplemented with 50 µL ampicilin (100 mg/mL). It was inoculated with one pipette tip of the glycerol stock of VcSiaQM. This preculture was also grown at 37° C for 12-14 h and 180 rpm.

For the main culture 1 L of LB media was supplemented with 1 mL of ampicilin (100 mg/mL) and inoculated with 10 mL of preculture. The main culture was grown at 37° C until an OD(600) of 1 was reached and then expression was induced with 50 mg of L-(+)-arabinose. The protein was expressed for 2 h at 37° C and at 180 rpm.
The cells were harvested by centrifugation for 20 min at 4000 rpm. The cell pellet was flash-frozen with liquid nitrogen and stored at -80 °C.

4.4.4. Protein purification and labeling

In this thesis various modifications of VcSiaP were purified. Modified HiSiaP and unmodified VcSiaQM were also needed for experiments.
The soluble, periplasmic proteins were all lysed in the same manner and bound to Ni-NTA resin for further purification. All mentioned buffers are listed in table 4.7.

Cell lysis of the P-domains

The cell pellet was weighted and resuspended in buffer A ($m_{(\text{pellet})}\, g^{-1} \times$ 5 mL). The cells were then lysed with a celldisruptor twice at 30 kPsi and the remaining lysate was centrifuged at 20,000 rpm for 20 min at 4°C. The cell debris was discharged.
The supernatant was mixed with Ni-beads, which were preequilibrated with buffer A. To ensure binding of the his-tagged protein to the beads, the mixture was incubated under gentle rotation at room temperature for 1 h.

VcSiaP Q54C/L173C, VcSiaP Q110C/L173C, VcSiaP L173C/S225C, VcSiaP Q54C/Q110C

The mixture was transferred to a polystyrene column and the Ni-beads were washed twice with 50 mL buffer A. To ensure that all the cysteines remain unbound and free for labeling, the beads were washed with 50 mL reducing buffer, which contains 1 mM TCEP. To remove the excess TCEP another washing step with 50 mL buffer A was included. The protein was labeled and eluted simultaneously with 15 mL elution buffer which contained imidazole and MTSSL. The amount of MTSSL was added as six times excess of an estimated amount of the protein (10 mg/L). MTSSL was stored in a stock concentration of 100 mM in DMSO at -80° C.
The elution was collected and concentrated to a volume of 5 mL at 4000 rpm. The concentrated sample was then loaded onto a pre-equilibrated Superdex 200 16/60 or a Superdex 75 column and eluted with buffer A.
The elution peak was collected and the protein visualized via SDS-PAGE. The labeling was checked via *cw*-EPR. The protein was concentrated to roughly 20-30 mg/mL

distributed into 50 µL Eppendorf tubes, flash-frozen with liquid nitrogen and stored at -80°C.

VcSiaP Q54R1/L173R1 {X} | {R125A, R125K, E184A, E184D, E184Q, H2017A, H207Q}

The seven VcSiaP mutants for the binding cleft experiments were purified as described in the protocol for VcSiaP Q54C/L173C, VcSiaP Q110C/L173C, VcSiaP L173C/S225C and VcSiaP Q54C/Q110C.
Additionally to this protocol ion exchange with an ENrich Q 10/100 post size exclusion chromatography was performed as an optimization to remove remains of a co-expressed *E. coli* protein (see results chapter 2.1.2).
The protein was collected, the volume determined and then diluted with 50 mM Tris (pH = 8) to a NaCl concentration of roughly 20 mM. Without this dilution step, no binding to the ENrich Q 10/100 is possible. The protein was eluted from the column with a salt gradient, created by ion exchange buffer A and ion exchange buffer B.

VcSiaP Q245R1

The purification of this protein follows the same protocol as for VcSiaP Q54C/L173C, except for an additional ion exchange step, which was now introduced before the size exclusion chromatography:
After elution from the Ni-column with elution buffer B, the protein was diluted to salt concentration of roughly 20 mM NaCl and loaded onto an ENrich Q 10/100 column. The protein was eluted from the column with a salt gradient, created by ion exchange buffer A and ion exchange buffer B. This elution was collected and concentrated to a volume of 5 mL at 4000 rpm. The concentrated sample was then loaded onto a pre-equilibrated Superdex 200 16/60 or a Superdex 75 column and eluted with buffer A.

VcSiaP Q245Biotin

After batch-binding to the Ni-NTA resin for 1 h, the mixture was transferred to a polystyrene column and the Ni-beads were washed twice with 50 mL buffer A. The protein was then washed with 50 mL reduction buffer and the excess TCEP was washed off with 50 mL PBS buffer. 50 mL of PBS buffer were supplemented with Maleimide-PEG2-Biotin (see figure 4.2) as six times excess of an estimated amount of the protein (10 mg/L) and incubated with the protein-bound resin for 1 h at room temperature while being mildly shaken.

The resin was transferred back to the column and washed with 50 mL buffer A to remove excess biotin-label. The protein was eluted with 15 mL elution buffer.
The labeled protein was then subjected to ion exchange and size exclusion chromatography as described in the protocol for VcSiaP Q245R1.
The labeling was checked via dot blot: a small amount of anti-biotin HRP-linked antibody was applied on a PVFD membrane and left to dry. Then the labeled protein was applied to it. The development was visible after a few minutes and documented with a gel imaging system.

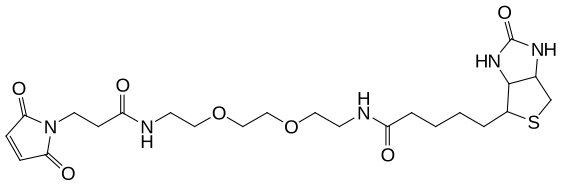
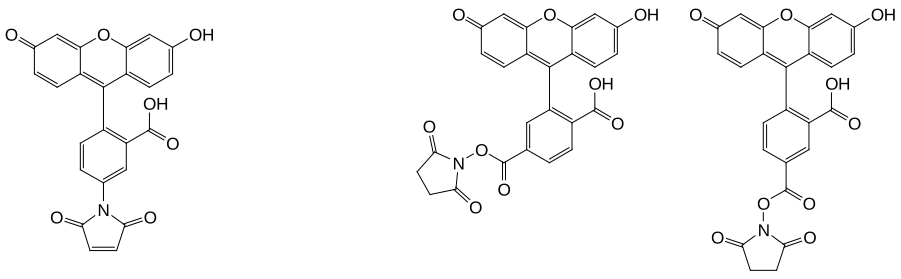


Figure 4.2. – Maleimide-PEG2-Biotin used for surface plasmon resonance experiments.

VcSiaP Q245Fluorescein / wildtype VcSiaP NHS-Fluorescein

For Microscale Thermophoresis measurements (see chapter 4.4.6) the protein VcSiaP had to be labeled with fluorescein. Two different labels were tested with this method: Fluorescein-5-Maleimide which binds to cysteines, and NHS-Fluorescein which targets primary amino groups (see picture 4.3).



Fluorescein-5-Maleimide NHS-Fluorescein

Figure 4.3. – Fluorescein dyes for labeling. NHS-Fluorescein can be ordered as a mixed isomer of 5- and 6-carboxyfluorescein succinimidyl ester. Fluorescein-5-Maleimide is available as a single isomer.

The cells were lysed and bound to the Ni NTA-resin as described in section 4.4.4. After batch-binding to the Ni-NTA resin for 1 h, the mixture was transferred to a polystyrene column and the Ni-beads were washed twice with 50 mL buffer A.
In the case of VcSiaP Q245C this was followed by a wash with 50 mL reduction buffer and a direct elution with 15 mL elution buffer which was supplemented with 1 mM

TCEP. Up to the point of labeling all of the buffers for VcSiaP Q245C had to contain 1 mM TCEP.
For wild type VcSiaP the reduction step was redundant and the protein was directly eluted with 15 mL elution buffer. The protein was subjected to ion exchange chromatography with ion exchange buffer A and ion exchange buffer B. This step was then followed by a size exclusion chromatography on the SD 200 16/60 column and buffer A. The protein was collected and the His-tag cleaved with TEV protease (1:10) for 4 h. At this point the protocol varies, depending on the label:

- Fluorescein-5-Maleimide : Another size exclusion was performed to remove TCEP from the buffer. The protein was directly collected after detection and the label was added in fourfold excess to the protein. Labeling was performed over night at 4° C. A final size exclusion was run to remove excess of label.

- NHS-Fluorescein: Another size exclusion was performed to exchange the buffer from buffer A to PBS buffer. This step is crucial, otherwise the label will bind to the Tris in the buffer instead. The protein was collected and NHS-fluorescein added (1:1). Labeling was performed over night at 4° C. On the next day, a final size exclusion chromatography was run with buffer A to remove excess of label and to exchange the protein again into its original buffer.

VcSiaP Q245Dylight800 / HiSiaP K254Dylight800

After batch-binding to the Ni-NTA resin for 1 h, the mixture was transferred to a polystyrene column and the Ni-beads were washed twice with 50 mL buffer A. This was followed by a reduction washing step with 50 mL reduction buffer and an elution with 15 mL elution buffer containing 1 mM TCEP.
The protein was loaded onto an ENrich Q 10/100 column and eluted with ion exchange buffer A and ion exchange buffer B. The protein fraction was collected and loaded onto a Superdex 200 16/60 column. The size exclusion column was run with buffer A supplemented with 1 mM TCEP. The eluted protein fraction was collected and the His-tag cleaved with TEV protease (1:10) for 4 h. The cleaved tag and the remaining protease were removed by the Ni-column and 1 mg of the protein was loaded onto a Superdex 200 10/300 column and eluted with buffer A to remove TCEP.
The peak of the protein was directly collected and the concentration of the protein was determined. Then the protein was supplemented with four fold excess of Dylight800 label. The labeling took place over night at 4° C. On the following day, the excess label was removed with an on bench PD-10 column.

VcSiaP {D241/Q245}Rx / HiSiaP {V250/K254}Rx

In a first attempt to purify and label these proteins, the same protocol as for VcSiaP Q54C/L173C, VcSiaP Q110C/L173C, VcSiaP L173C/S225C and VcSiaP Q54C/Q110C was followed using the Rx-label instead of MTSSL. This lead to dissatisfying labeling, therefore the protocol was optimized.

After batch-binding to the Ni-NTA resin, the mixture was transferred to a polystyrene column and the Ni-beads were washed twice with 50 mL buffer A. This was followed by a reduction washing step with 50 mL reduction buffer and an elution with 15 mL elution buffer containing 1 mM TCEP.
The protein was subjected to ion exchange on an ENrich Q 10/100 column and eluted with ion exchange buffer A and ion exchange buffer B. The protein was collected and the amount of protein determined. Then, the protein was incubated again with Ni-NTA resin for 1 h for the labeling procedure. To remove TCEP the Ni-column was washed with 50 mL buffer A.
0.5 equivalent label were dissolved in 5 mL buffer A and incubated with the Ni-bound protein for 1 h. Another 0.5 equivalent label was added to the protein and again incubated for 1 h.
The resin was washed with 100 mL buffer A to remove unbound label and the protein was eluted with elution buffer. TEV protease was added and the protein was subjected to dialysis over night at 4° to 2 L of buffer A.
On the next day, tag and TEV protease were removed with Ni-NTA and the protein subjected to size exclusion chromatography with a Superdex 200 16/60 column.

VcSiaQM

The VcSiaQM construct is created as N-terminally His_14-taged protein in the vector pBadHisTev. Only the Q-domain is tagged, but the M-domain co-purifies and together they form a strong complex.
The VcSiaQM cell pellet is weighted and then resuspended in five times (w/v) QM buffer. The cells were lysed three times using a celldisruptor at 30 kPsi. The lysate is centrifuged for 1 h at 300,000×g with a ultracentrifuge and the supernatant is discarded. The remaining pellet is resuspended in 50 mL extraction buffer and was left for incubation over night on a rotating wheel at 4°C. This extraction was then centrifuged for 1 h at 300,000×g to separate the membrane debris from the solubalized protein.
The extraction was mixed with Ni-NTA resin pre-equilibrated with QM protein buffer.

This mixture was incubated under gentle rotation at 4°C for 2 h.
The resin was then transferred into a polystyrene column and washed twice with 50 mL QM washing buffer and finally eluted with 15 mL QM elution buffer. The eluate is concentrated at 4000 rpm to a volume of 500 µL and then loaded onto a pre-equilibrated Superdex 200 10/300 column. The size exclusion was run with QM protein buffer.
The elution peak was collected and the protein visualized via SDS-PAGE. The protein was concentrated, aliquoted and flash-frozen with liquid nitrogen.

4.4.5. Surface Plasmon Resonance Spectroscopy

Experiments to determine the K_D between VcSiaP and Neu5Ac were performend with a Biacore 3000 (GE Healtcare) on a chip from the Biotin CAPture Kit (GE Healthcare) at room temperature. Prior to use the chip was activated according to manufacturer's protocol.
At first, 50 µL of biotinylated VcSiaP (see chapter 4.4.4) at a concentration of 1 µM was streptavidin-captured on the chip at a flow rate of 10 µL/min. For the binding analysis Neu5Ac injections were performed at a flow rate of 30 µL/min at the concentration range of 5 nm to 1 mM. The continuous flow buffer used contained 50 mM Tris at a pH of 7.8 and 50 mM NaCl. To remove any residual bound Neu5Ac, regeneration injections with 500 mM NaCl were performed.
The obtained binding curves were analyzed with the BIAevaluation software. The curves were referenced against a protein free channel on the chip. The K_D was calculated via the Hill-fits available in the software.

4.4.6. Microscale Thermophoresis

Theory

The migration of molecules in a temperature gradient is referred to as thermophoresis and was discovered by Ludwig and Soret in the 19th century [176]. Microscale thermophoresis (MST) is a technique, which makes use of this physical phenomenon in order to detect binding events and to quantify the biomolecular interaction [177].
During the experiment a molecule with fluorecent label is filled into a glass capillary and heated with an infrared laser. Upon heating the migration of the molecules can be detected via fluorescence measurement.
In the steady state, thermophoresis is balanced by mass diffusion, which can be de-

scribed by the following spatial concentration distribution:

$$\frac{c_1}{c_0} = \exp\left(-S_T \Delta T\right) \tag{4.1}$$

where c_0 ist the concentration before heating, c_1 is the hot equilibrium concentration, S_T is the Soret coefficient and ΔT is the temperature difference. The Soret coefficient can be expressed by

$$S_T = \frac{A}{kT}\left(-s_{\text{hyd}} + \frac{\beta\sigma_{\text{eff}}^2}{4\varepsilon\varepsilon_0 T}\lambda_{DH}\right) \tag{4.2}$$

with A the particle surface area, k the Boltzmann constant, T is the temperature, s_{hyd} the hydration entropy, β the factor resulting from temperature derivatives, σ_{eff} the effective surface charge density, ε the dielectric constant, ε_0 is the vacuum permittivity and λ_{DH} is the Debye-Hueckel length.
The migration of molecules in a temperature gradient depends on physical properties like the size, the charge and the hydration shell. When binding a ligand at least one of these properties changes.

The labeled protein is supplemented with decreasing amount of ligand, which are distributed in up to 16 capillaries. The fluorescence is monitored before, after and during infrared radiation.

$$F_{\text{norm}} = (1 - x) \cdot F_{\text{norm}}(\text{unbound}) + x \cdot F_{\text{norm}}(\text{bound}) \tag{4.3}$$

here x describes the fraction of labeled molecules bound to their targets.

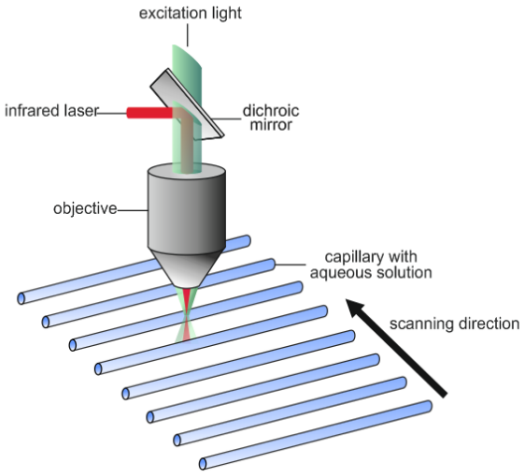


Figure 4.4. – Set-up of the microscale thermophoresis experiment. The figure was adapted from the Monolith NT.115 user manual.

Figure 4.4 shows the setup of the experiment. At first the initial fluorescence is detected. Then, a focused infrared laser (IR) is used to create a local and steep temperature gradient within the glass capillaries, which contain the protein samples. This way the samples are typically heated by 1-6 K. This increase of temperature leads to an abrupt change in fluorescence intensity due to the thermophoretic movement of the labeled protein (temperature jump, T-jump).
The IR-laser shares the same optical path as the detection unit and ensures the coupled detection of temperature increase and fluorescence signal alteration.
The intensity reaches a plateu, in which the thermodiffusion is counterbalanced by mass diffusion. After heating, the fluorescence signal recovers due to cooling of the sample and the backdiffusion of particles.

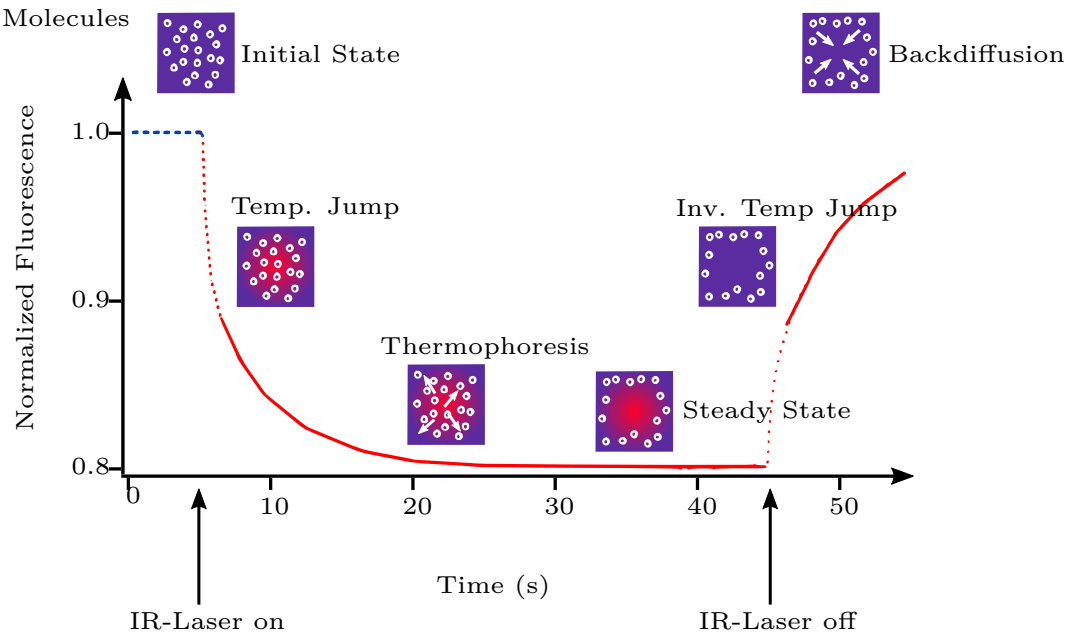


Figure 4.5. – Measured timetrace of the fluorescence signal during a MST experiment. The fluorescence of a sample inside a glass capillary is measured. The normalized fluorescence in the heated spot is plotted against time. The figure was adapted from [178].

The timetrace of the fluorescence signal can be divided into six states. In the *initial state* all molecules undergo unidirected brownian motion. As a result the fluorescence signal is constant. After irradiation with the IR laser the so called *temperature jump* takes place: the fluorescence signal starts to decrease exponentially. This dropping of signal intensity is due to changes in the local surrounding of the fluorophore, which are induced by heating. This effect sets in after approximately 50 ms. Then, on the timescale of diffusion (approx. 1s) the effect of *thermophoresis* kicks in. The signal reaches a *steady state*, when thermophoresis and regular diffusion average out. After reaching the steady state, the IR laser is turned off the sample starts to cool. This causes an *inverted temperature jump* followed by *back-diffusion* of the molecules.

Experiment

Experiments were performed on a Monolith NT.115 instrument at 25 °C. Fluorescein-labeled VcSiaP (see chapter 4.4.4) was used at a concentration of 100 nM. The ligand Neu5Ac was titrated in 1:2 dilutions ranging from 2000 µM to 0,06 µM, using MST buffer. Each dilution-protein mixture was filled into capillaries and placed in the slots on the sample tray. Thermophoresis was induced in each capillary by a laser using 5 s/ 30 s/ 5 s as off/on/off times. The instrument parameters were set to 40% LED power

and 10% MST power. The obtained curves were analyzed by MO.Affinity Analysis using the signal from thermophoresis and the T-jump.

Table 4.18. – Settings for MST

Measurement Settings	
MST-power	20% or 40%
Excitation-Power	40%
Excitation type	Blue
Thermostat. Setpoint	25°
Analysis Settings	
Evaluation strategy	Thermophoresis with T-Jump
Cold Region Start	-1
Cold Region End	0
Hot Region Start	20.67
Hot Region End	21.67

4.5. X-ray Crystallography

4.5.1. Crystallization

Initial screening

To identify suitable crystallization conditions the initial screen JCSG was used. 0.5 µL of protein sample at a concentration of roughly 15 mg/mL were mixed manually with 0.5 µL reservoir solution of the various conditions. The drops were pipetted on a 96-well 2-drop MRC crystallization plate as sitting-drops over 50 µL reservoir solution. The plate was sealed with foil. The screens were stored in a vibration-free environment at room temperature and were checked regularly for growth by light microscopy.

Optimization of initial conditions

Crystallization parameters of the initial hits (i.e. salt concentration, precipitant concentration and pH) were statistically varied by a program written by Gregor Hagelüken. The droplet sizes and ratios were used as during the screening of initial conditions.

Identification of protein crystals

Crystals in promising conditions have to be checked whether they are protein crystals or merely salt crystals from the reservoir solution. One way to check for protein crystals is to dye them with methylene blue, which incorporates into the solvent channels of the crystal. For this purpose, methylene blue was diluted with reservoir solution in the ratio 1:5 and 0.5 µL of this mixture was added to the drop containing the crystals. The drop was checked after several hours for changes in the crystal colors.
Another way to check for protein crystals is to test if they diffract when irradiated with x-rays. Crystals in promising conditions were picked and a diffraction pattern was recorded using a X8-KappaApex II diffractometer. This initial diffraction pattern additionally gives insight about the quality of the crystal.

Harvest of crystals

Crystals were picked with cryogenic loops and dipped into a cryoprotectant solution consisting of 35% glycerol and 65% of the respective reservoir solution. The loops which now contain the crystals were flash-frozen with liquid nitrogen and stored in liquid nitrogen until measurement.

4.5.2. X-ray data collection and processing

X-ray data were collected at the Helmholtz Zentrum in Berlin with the electron storage ring Bessy II [179]. The used beamline MX 14.3 is operated at a fixed energy of 13.87 keV and equipped with a MarMosaic 224 CCD-detector.

The obtained data were processed using the CCP4i program package [180]. Indexing and integration was performend with the subprogram iMOSFLM [181]. The structure factor amplitudes were calculated with the program CTRUNCATE [180].
The structure was solved by molecular replacement with the subprogram PHASER [182] from the Phenix program [183] package. The VcSiaP open structure served as model (PDB:4mag). The structure was optimized and validated by the webtool MOLPROBITY [184].

4.6. Electron paramagnetic resonance spectroscopy

4.6.1. Continuous wave EPR

Sample preparation

All protein samples were measured in their original post-size-exclusion chromatography buffer (buffer A for R1-labeled and QM buffer for Rx labeled proteins). The samples were diluted to the desired concentration (usually 25-100 µM) and then soaked into a 10 µL capillary. To prevent the sample from leaking, the bottom of the capillary was sealed with superglue. The sealed capillary was then placed in a quartz glass X-band tube.

Instrumental setup

The *cw*-EPR experiments were performend either on an EMXmicro EPR-spectrometer or an EMXnano EPR-spectrometer. The EMXmicro was equipped with an ER 4119HS resonator, the EMXnano had a built-in resonator. The *cw*-EPR spectra were recorded at room temperature.

Measurement

Spectra of the protein samples were recorded on the EMXmicro, using a microwave power of 17.44 mW, a video amplifier gain of 60 dB, a modulation amplitude of 1 G, a time constant of 20.48 ms, a conversion time of 51.58 ms, and a resolution of 10 points/G.
The parameters for the EMXnano EPR-spectrometer were: a microwave power of 2.51 mW, a video amplifier gain of 30 dB, a modulation amplitude of 1 G, a time constant of 0.01 ms, a conversion time of 30.00 ms and a resolution of 8 points/G.

4.6.2. Pulsed EPR

Sample preparation

The total volume of every sample was 80 µL consisting of 25 µM protein in PELDOR buffer and 50% deuterated ethylene glycole as glass forming additive. Depending on the experiment an additional substrate was added (either 1 mM Neu5Ac, 10 mM Neu5Ac or 250 µM peptide).

The sample was transferred into a quarz glass tube and left to settle down and incubate on ice for 30 min.

Instrumental setup

The pulsed EPR measurements were performed on an ELEXSYS E580 spectrometer equipped with an ER5106QT-II resonator at Q-band frequencies. Pulses were amplified with a 150 W traveling wave tube (TWT) pulse amplifier. The temperature was set and controlled with a CF935 liquid helium flow cryostat in combination with a Mercury iTC503S Temperature controller.
All pulsed experiments were recorded at 50 K.

Echo detected field swept EPR spectrum

The Hahn-Echo sequence $\pi/2$ - τ - π - τ - Echo [185] with pulse lengths of $\pi/2 = 12$ ns and $\pi = 24$ ns and $\tau = 400$ ns was applied. The occurring echo was integrated and the magnetic field was swept. The recorded spectrum shows the integrated echo intensity as function of the magnetic field. 1 scan was recorded with a shot repetition time of 999.60 µs and 50 shots per point. The sweeping range of the magnetic field was 11900-12100 G. The microwave frequency was 33.7 GHz.
For further pulsed EPR experiments the magnetic field was set to the value of the highest signal amplitude in the field swept spectrum.

Electron Spin Echo Envelope Modulation

The interpulse-delay time τ of the Hahn-Echo sequence was extended by an increment dx, resulting in the sequence $\pi/2$ - $(\tau + dx)$ - π - $(\tau + dx)$ - Echo. The used parameters were $\pi/2 = 12$ ns, $\pi = 24$ ns, initial $\tau = 200$ ns and $dx = 8$ ns. 4 scans were recorded with a shot repetition time of 1020 µs. τ was incremented for 3096 points and a two-step phase cycle was applied.
The recorded trace shows the spin echo decay (acquired by point detection of the maximal amplitude of the Hahn echo) with increasing inter-pulse delay time. The delay time at which the echo intensity approached 0 was determined to be the length of the time traces for the PELDOR experiments.

Transient nutation

For this experiment the Hahn-Echo sequence was extended by an additional π pulse, leading to the sequence: π - T - $\pi/2$ - τ - π - τ - Echo. The time T of this additional π

pulse is varied and the effect on the maximum amplitude of the echo is detected. The time point at the first minima in the recorded spectrum is the optimal pump pulse length.
A two-step phase cycling was applied.

PELDOR

The four-pulse PELDOR sequence $\pi/2$ - τ - π - τ - Echo1 - τ - π - τ - Echo2 was applied [186]. The pump pulse π was irradiated between Echo1 and π. The length of the pump pulse was determined by a transient nutation experiment. The frequency of the pump pulse was set to the maximum of the field sweep spectrum and the observer pulses were applied at a frequency offset of -80 MHz relative to the pump frequency. An 8-step modulation averaging was applied with an increment time of 16 ns and a two-step phase cycle was used.

Data Analysis

All pulsed EPR data was analyzed with the DeerAnalysis MATLAB package. The background was substracted from the time traces and the distance distribution calculated via Tikhonov regularization. Optimal regularization parameter was chosen according to the L-curve criterion.

4.7. Peptide Scan

Three peptide scans were manually designed as described in section 2.4 and ordered from peptides&elefants. They consisted of a cellulose membrane on which the ordered peptides were directly spot synthesized. Figure 4.6 shows a simple depiction of the three ordered scans. The size of the picture is accurate.

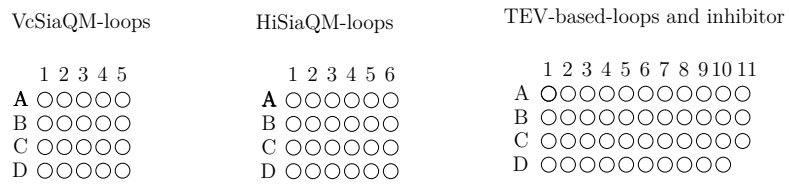


Figure 4.6. – Schematic depiction of the used peptide scan.

The scans were incubated with Dylight800 labeled P-domain (regular experiment) or incubated with labeled P-domain and Neu5Ac (competitive experiment).

Regular experiment

Prior to usage, the scan had to be soaked with water and washed three times for 15 s with methanol to activate the surface of the anchored peptides. The scan is then washed with 10 mL MiliQ water for 10 min. To avoid unspecific binding of the protein to the membrane, the scan is blocked with Blocking Solution #2 (1% BSA in 15 mL buffer A) for 1 h.
The scan was washed for three times with 10 mL Buffer A for 10 min to fully remove the blocking solution.
The scan was incubated for 1 h with 10 mL buffer A supplemented with 1 µL of Dylight800 labeled protein (either VcSiaP or HiSiaP, 1 mg/mL). The excess protein was then washed off the scan with 10 mL Buffer A for 10 min for three times.
Pictures of the scan were recorded with an Odyssey Imaging System at an excitation wavelength of 800 nm.

Competitive experiment

The assay/scan was also executed as competitive experiment with Neu5Ac: 10 mL of buffer A were supplemented with 1 µL of Dylight800 labeled protein (either VcSiaP or HiSiaP, 1 mg/mL) and 1 mM Neu5Ac. This mixture was first equilibrated for 15 min and then incubated with the scan for 1 h. Further handling took place as described with the regular scan.

Data Analysis

The data was analyzed with the software Image Studio Lite. During this procedure, part of the empty membrane was defined as background and the fluorescence-intensities of the peptides were determined.

5. Bibliography

[1] "List of Blueprint priority diseases". In: https://www.who.int/blueprint/priority-diseases/en/ (accessed 31.10.2019).

[2] "Pandemic preparedness". In: https://www.euro.who.int/en/health-topics/communicable-diseases/influenza/pandemic-influenza/pandemic-preparedness (accessed 31.10.2019).

[3] Huremović D. "Brief History of Pandemics (Pandemics Throughout History)". In: *Psychiatry of Pandemics: A Mental Health Response to Infection Outbreak* (2019), pp. 7–35.

[4] F. R. Chowdhury et al. "Pandemics, pathogenicity and changing molecular epidemiology of cholera in the era of global warming". In: *Annals of clinical microbiology and antimicrobials* 16.10 (2017).

[5] D. Hu et al. "Origins of the current seventh cholera pandemic". In: *Proceedings of the National Academy of Sciences of the United States of America* 113.48 (2016), pp. 7730–7739.

[6] "Cholera - the forgotten pandemic". In: https://www.who.int/cholera/the-forgotten-pandemic/en/ (accessed 31.10.2019).

[7] M. Ali et al. "Updated Global Burden of Cholera in Endemic Countries". In: *PLoS neglected tropical diseases* 9.6 (2015), pp. 1–13.

[8] N. J. Beeching, C. A. Hart, and B. I. Duerden. "Tropical and exotic infections". In: *Journal of medical microbiology* 49.1 (2000), pp. 5–27.

[9] R. I. Glass et al. "Cholera in Africa: lessons on transmission and control for Latin America". In: *The Lancet* 338.8770 (1991), pp. 791–795.

[10] F. Federspiel and M. Ali. "The cholera outbreak in Yemen: lessons learned and way forward". In: *BMC Public Health* 18.1338 (2018), pp. 1–8.

[11] S. Almagro-Moreno, K. Pruss, and R. K. Taylor. "Intestinal Colonization Dynamics of Vibrio cholerae". In: *PLoS pathogens* 11.5 (2015), pp. 1–11.

[12] W. S. Jermyn and E. F. Boyd. “Characterization of a novel Vibrio pathogenicity island (VPI-2) encoding neuraminidase (nanH) among toxigenic Vibrio cholerae isolates”. In: *Microbiology* 148.11 (2002), pp. 3681–3693.

[13] S. Almagro-Moreno and E. F. Boyd. “Insights into the evolution of sialic acid catabolism among bacteria”. In: *BMC evolutionary biology* 9.118 (2009), pp. 1–16.

[14] T. J. Silhavy, D. Kahne, and S. Walker. “The Bacterial Cell Envelope”. In: *Cold Spring Harbor perspectives in biology* 2.5 (2010), pp. 1–16.

[15] X. Wang and P. J. Quinn. “Lipopolysaccharide: Biosynthetic pathway and structure modification.” In: *Progress in Lipid Research* 49.2 (2010), pp. 97–107.

[16] K. H. Schleifer and O. Kandler. “Peptidoglycan types of bacterial cell walls and their taxonomic implications.” In: *Bacteriological review* 36.4 (1972), pp. 407–477.

[17] G. M. Cooper. "The Cell, a molecular approach". Sinauer Associates Inc., 2000.

[18] D. Oh, H. Han, and G. L. Amidon. "Drug Transport and Targeting - Intestinal Transport". Pharmaceutical Biotechnology, 1999.

[19] H. Lodish et al. "Molecular Cell Biology". W. H. Freeman, 2016.

[20] C. Maurel et al. “Functional characterization of the Escherichia coli glycerol facilitator, GlpF, in Xenopus oocytes.” In: *The Journal of biological chemistry* 269.16 (1994), pp. 11869–11872.

[21] J. K. Lanyi and B. Schobert. “Mechanism of proton transport in bacteriorhodopsin from crystallographic structures of the K, L, M1, M2, and M2’ intermediates of the photocycle.” In: *Journal of molecular biology* 328.2 (2003), pp. 439–450.

[22] M. Kolbe et al. “Structure of the light-driven chloride pump halorhodopsin at 1.8 A resolution”. In: *Science* 288 Kolbe, M., Besir, H., Essen, L. O., & Oesterhelt, D. (2000). Structure of the light-driven chloride pump halorhodopsin at 1.8 A resolution. Science (New York, N.Y.),5470 (2000), pp. 1390–1396.

[23] S. Wilkens. “Structure and mechanism of ABC transporters”. In: *F1000prime reports* 7.14 (2015), pp. 1–9.

[24] N. Yan. “Structural Biology of the Major Facilitator Superfamily Transporters”. In: *Annual review of biophysics* 44 (2015), pp. 257–283.

[25] U. Dahl et al. “Identification of a phosphotransferase system of Escherichia coli required for growth on N-acetylmuramic acid”. In: *Journal of bacteriology* 186.8 (2004), pp. 2385–2392.

[26] E. Padan. "Bacterial Membrane Transport:Secondary Transport Proteins." In: Encyclopedia of Life Sciences. John Wiley & Sons, 2019.

[27] M. Viereck et al. “Transporter taxonomy - a comparison of different transport protein classification schemes”. In: *Drug discovery today. Technologies* 12 (2014), pp. e37–e46.

[28] Q. Ren, K. Chen, and I. T. Paulsen. “TransportDB: a comprehensive database resource for cytoplasmic membrane transport systems and outer membrane channels”. In: *Nucleic acids research* 35 (2007), pp. D274–D279.

[29] K. M. Morrissey et al. “The UCSF-FDA TransPortal: A Public Drug Transporter Database”. In: *Clinical pharmacology and therapeutics* 92.5 (2012), pp. 545–546.

[30] M. A. Hediger et al. “The ABCs of membrane transporters in health and disease (SLC series): introduction”. In: *Molecular aspects of medicine* 34.2-3 (2013), pp. 95–107.

[31] G. Fichant, M. J. Basse, and Y. Quentin. “ABCdb: an online resource for ABC transporter repertories from sequenced archaeal and bacterial genomes”. In: *FEMS microbiology letters* 256.2 (2006), pp. 333–339.

[32] A. Y. Ye et al. “Human Transporter Database: Comprehensive Knowledge and Discovery Tools in the Human Transporter Genes”. In: *PloS one* 9.2 (2014), e88883.

[33] M. H. Saier. “A functional-phylogenetic classification system for transmembrane solute transporters”. In: *Microbiology and molecular biology reviews* 64.2 (2000), pp. 354–411.

[34] M. H. Saier. “Molecular Phylogeny as a Basis for the Classification of Transport Proteins from Bacteria, Archaea and Eukarya”. In: *Advances in microbial physiology* 40 (1998), pp. 81–136.

[35] M. H. Saier et al. “The Transporter Classification Database (TCDB): recent advances”. In: *Nucleic acids research* 44.D1 (2016), pp. D372–D379.

[36] M. J. Lenaeus et al. “Structures of closed and open states of a voltage-gated sodium channel”. In: *Proceedings of the National Academy of Sciences of the United States of America* 114.15 (2017), pp. E3051–E3060.

[37] J. Vergalli et al. “Porins and small-molecule translocation across the outer membrane of Gram-negative bacteria.” In: *Nature reviews microbiology* 18.3 (2020), pp. 164–176.

[38] S. W. Cowan et al. “Crystal structures explain functional properties of two E. coli porins”. In: *Nature* 358.6389 (1992), pp. 727–733.

[39] V.S. Reddy et al. “The major facilitator superfamily (MFS) revisited”. In: *The FEBS journal* 279.11 (2012), pp. 2022–2035.

[40] N. Noinaj et al. “TonB-dependent transporters: regulation, structure, and function”. In: *Annual review of microbiology* 64 (2010), pp. 43–60.

[41] D. C. Rees, E. Johnson, and O. Lewinson. “ABC transporters: the power to change”. In: *Nature reviews. Molecular cell biology* 10.3 (2009), pp. 218–227.

[42] J. M. Berrisford, R. Baradaran, and L. A. Sazanov. “Structure of bacterial respiratory complex I”. In: *Biochimica et biophysica acta* 1857.7 (2016), pp. 892–901.

[43] K. Kovalev et al. “High-resolution structural insights into the heliorhodopsin family”. In: *Proceedings of the National Academy of Sciences of the United States of America* 117.8 (2020), pp. 4131–4141.

[44] E. Gabor et al. “The phosphoenolpyruvate-dependent glucose-phosphotransferase system from Escherichia coli K-12 as the center of a network regulating carbohydrate flux in the cell”. In: *European journal of cell biology* 90.9 (2011), pp. 711–720.

[45] R. A. North et al. “Just a spoonful of sugar...: import of sialic acid across bacterial cell membranes”. In: *Biophysical reviews* 10.2 (2018), pp. 219–227.

[46] S. W. Cowan et al. “The structure of OmpF porin in a tetragonal crystal form”. In: *Structure* 3.10 (1995), pp. 1041–1050.

[47] A. Baslé et al. “Crystal structure of osmoporin OmpC from E. coli at 2.0 A”. In: *Journal of molecular biology* 362.5 (2006), pp. 933–942.

[48] C. Wirth et al. “NanC crystal structure, a model for outer-membrane channels of the acidic sugar-specific KdgM porin family”. In: *Journal of molecular biology* 394.4 (2009), pp. 718–731.

[49] C. Phansopa et al. “Structural and functional characterization of NanU, a novel high-affinity sialic acid-inducible binding protein of oral and gut-dwelling Bacteroidetes species”. In: *The Biochemical journal* 458.3 (2014), pp. 499–511.

[50] S. Roy, C. W. Douglas, and G. P. Stafford. “A novel sialic acid utilization and uptake system in the periodontal pathogen Tannerella forsythia”. In: *Journal of bacteriology* 192.9 (2010), pp. 2285–2293.

[51] E. Severi et al. “Sialic acid mutarotation is catalyzed by the Escherichia coli beta-propeller protein YjhT.” In: *The Journal of biological chemistry* 283.8 (2008), pp. 4841–4849.

[52] S. Almagro-Moreno and E. F. Boyd. “Sialic acid catabolism confers a competitive advantage to pathogenic vibrio cholerae in the mouse intestine”. In: *Infection and immunity* 77.9 (2009), pp. 3807–3816.

[53] S. M. Steenbergen, J. L. Jirik, and E. R. Vimr. “YjhS (NanS) is required for Escherichia coli to grow on 9-O-acetylated N-acetylneuraminic acid”. In: *Journal of bacteriology* 191.22 (2009), pp. 7134–7139.

[54] C. Thomas and R. Tampé. “Structural and Mechanistic Principles of ABC Transporters”. In: *Annual review of biochemistry* 89 (2020), pp. 605–636.

[55] C. Marion et al. “Sialic acid transport contributes to pneumococcal colonization”. In: *Infection and immunity* 79.3 (2011), pp. 1262–1269.

[56] D. Drew et al. “Structures and General Transport Mechanisms by the Major Facilitator Superfamily (MFS)”. In: *Chemical reviews* 121.9 (2021), pp. 5289–5335.

[57] J. Martinez, S. Steenbergen, and E. Vimr. “Derived structure of the putative sialic acid transporter from Escherichia coli predicts a novel sugar permease domain”. In: *Journal of bacteriology* 177.20 (1995), pp. 6005–6010.

[58] T. Henriquez et al. “Prokaryotic Solute/Sodium Symporters: Versatile Functions and Mechanisms of a Transporter Family”. In: *International journal of molecular sciences* 22.4 (2021), p. 1880.

[59] A. Watanabe et al. “The mechanism of sodium and substrate release from the binding pocket of vSGLT”. In: *Nature* 468.7326 (2010), pp. 988–991.

[60] W. Y. Wahlgren et al. “Substrate-bound outward-open structure of a Na+-coupled sialic acid symporter reveals a new Na+ site”. In: *Nature communications* 9.1, 1753 (2018).

[61] F. G. Blix, A. Gottschalk, and E. Klenk. “Proposed nomenclature in the field of neuraminic and sialic acids”. In: *Nature* 179.4569, 1088 (1957).

[62] T. Angata and A. Varki. “Chemical diversity in the sialic acids and related alpha-keto acids: an evolutionary perspective”. In: *Chemical reviews* 102.2 (2002), pp. 439–469.

[63] R. Schauer. “Sialic acids as regulators of molecular and cellular interactions”. In: *Current opinion in structural biology* 19 (2009), pp. 1–8.

[64] S. Kelm and R. Schauer. “Sialic acids in molecular and cellular interactions”. In: *International review of cytology* 175 (1997), pp. 137–240.

[65] A. Varki. “Biological roles of oligosaccharides: all of the theories are correct”. In: *Glycobiology* 3.2 (1993), pp. 97–130.

[66] E. Gascon, L. Vutskits, and J. Z. Kiss. “Polysialic acid-neural cell adhesion molecule in brain plasticity: from synapses to integration of new neurons”. In: *Brain research reviews* 56.1 (2007), pp. 101–118.

[67] U. Neu, J. Bauer, and T. Stehle. “Viruses and sialic acids: rules of engagement”. In: *Current opinion in structural biology* 21.5 (2011), pp. 610–618.

[68] W. Li et al. “Identification of sialic acid-binding function for the Middle East respiratory syndrome coronavirus spike glycoprotein”. In: *Proceedings of the National Academy of Sciences of the United States of America* 114.40 (2017), pp. E8508–E8517.

[69] U. Neu et al. “Structural basis of GM1 ganglioside recognition by simian virus 40”. In: *Proceedings of the National Academy of Sciences of the United States of America* 105.13 (2008), pp. 5219–5224.

[70] A. Guha-Niyogi, D. R. Sullivan, and S. J. Turco. “Glycoconjugate structures of parasitic protozoa”. In: *Glycobiology* 11.4 (2001), 45R–59R.

[71] V. G. Monteiro, C. P. Soares, and W. de Souza. “Host cell surface sialic acid residues are involved on the process of penetration of Toxoplasma gondii into mammalian cells”. In: *FEMS microbiology letters* 164.2 (1998), pp. 323–327.

[72] A. B. Lantos et al. “Sialic Acid Glycobiology Unveils Trypanosoma cruzi Trypomastigote Membrane Physiology”. In: *PLoS pathogens* 12.4 (2016), e1005559.

[73] E. Severi, D. W. Hood, and G. H. Thomas. “Sialic acid utilization by bacterial pathogens”. In: *Microbiology* 153.9 (2007), pp. 2817–2822.

[74] N. D. McDonald et al. “Host-Derived Sialic Acids Are an Important Nutrient Source Required for Optimal Bacterial Fitness In Vivo”. In: *mBio* McDonald, N. D., Lubin, J. B., Chowdhury, N., & Boyd, E. F. (2016). 7.2 (2016), e02237–15.

[75] E. Vimr, C. Lichtensteiger, and S. Steenbergen. “Sialic acid metabolism’s dual function in Haemophilus influenzae”. In: *Molecular microbiology* 36.5 (2000), pp. 1113–1123.

[76] E. Vimr and C. Lichtensteiger. “To sialylate, or not to sialylate: that is the question”. In: *Trends in microbiology* 10.6 (2002), pp. 254–257.

[77] B. L. Haines-Menges et al. “Host Sialic Acids: A Delicacy for the Pathogen with Discerning Taste”. In: *Microbiology spectrum* 3.4 (2015).

[78] B. K. Sohanpal et al. “Integrated regulatory responses of fimB to N-acetylneuraminic (sialic) acid and GlcNAc in Escherichia coli K-12”. In: *Proc Natl Acad Sci U S A* Sohanpal, B. K., El-Labany, S., Lahooti, M., Plumbridge, J. A., & Blomfield, I. C. (2004). Integrated regulatory responses of fimB to N-acetylneuraminic (sialic) acid and GlcNAc in Escherichia coli K-12. Proceedings of the National Academy of Sciences of the United States of America, 101(46),46 (2004), pp. 16322–16327.

[79] B.K. Sohanpal et al. “Multiple co-regulatory elements and IHF are necessary for the control of fimB expression in response to sialic acid and N-acetylglucosamine in Escherichia coli K-12”. In: *Mol Microbiol.* (2007).

[80] Y. Li and X. Chen. “Sialic acid metabolism and sialyltransferases: natural functions and applications”. In: *Applied microbiology and biotechnology* 94.4 (2012), pp. 887–905.

[81] A. L. Mattos-Guaraldi, L. C. Formiga, and A. F. Andrade. “Trans-sialidase activity for sialic acid incorporation on Corynebacterium diphtheriae”. In: *FEMS microbiology letters* 168.2 (1998), pp. 167–172.

[82] S. Gulati et al. “Utilizing CMP-Sialic Acid Analogs to Unravel Neisseria gonorrhoeae Lipooligosaccharide-Mediated Complement Resistance and Design Novel Therapeutics”. In: *PLoS pathogens* 11.12 (2015), e1005290.

[83] S. Bose et al. “Structural and functional characterization of CMP-N-acetylneuraminate synthetase from Vibrio cholerae”. In: *Acta crystallographica. Section D, Structural biology* 75.Pt 6 (2019), pp. 564–577.

[84] K. L. Fox et al. “Identification of a bifunctional lipopolysaccharide sialyltransferase in Haemophilus influenzae: incorporation of disialic acid”. In: *The Journal of biological chemistry* 28.52 (2006), pp. 40024–40032.

[85] M. P. Bos, V. Robert, and J. Tommassen. “Biogenesis of the Gram-Negative Bacterial Outer Membrane”. In: *Annual Review of Microbiology* (2007).

[86] B. W. Simpson et al. “Lipopolysaccharide transport to the cell surface: biosynthesis and extraction from the inner membrane.” In: *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* (2015).

[87] S. Prakash et al. “The ion transporter superfamily”. In: *Biochimica et biophysica acta* 1618.1 (2003), pp. 79–92.

[88] J. S. Chen et al. “Phylogenetic characterization of transport protein superfamilies: superiority of SuperfamilyTree programs over those based on multiple alignments”. In: *Journal of molecular microbiology and biotechnology* 21.3-4 (2011), pp. 83–96.

[89] C. Mulligan, M. Fischer, and Thomas G. H. “Tripartite ATP-independent periplasmic (TRAP) transporters in bacteria and archaea”. In: *FEMS microbiology reviews* 35.1 (2010), pp. 68–86.

[90] J. A. Forward et al. “TRAP transporters: a new family of periplasmic solute transport systems encoded by the dctPQM genes of Rhodobacter capsulatus and by homologs in diverse gram-negative bacteria”. In: *Journal of bacteriology* 179.17 (1997), pp. 5482–5493.

[91] G. H. Scheepers, J. A. Lycklama a Nijeholt, and B. Poolman. “An updated structural classification of substrate-binding proteins”. In: *FEBS Letters* 590.23 (2016), pp. 4393–4401.

[92] L. T. Rosa et al. “Tripartite ATP-Independent Periplasmic (TRAP) Transporters and Tripartite Tricarboxylate Transporters (TTT): From Uptake to Pathogenicity”. In: *Frontiers in cellular and infection microbiology* 8.33 (2018).

[93] Deka R. K. et al. “Structural, bioinformatic, and in vivo analyses of two Treponema pallidum lipoproteins reveal a unique TRAP transporter”. In: *Journal of molecular biology* 416.5 (2012), pp. 678–696.

[94] D. J. Kelly and G. H. Thomas. “The tripartite ATP-independent periplasmic (TRAP) transporters of bacteria and archaea”. In: *FEMS microbiology reviews* 25.4 (2001), pp. 405–424.

[95] M. L. Oldham et al. “Crystal structure of a catalytic intermediate of the maltose transporter”. In: *Nature* 450.7169 (2007), pp. 515–521.

[96] J. Abramson et al. “Structure and mechanism of the lactose permease of Escherichia coli”. In: *Science* 301.5633 (2003), pp. 610–615.

[97] A. Müller et al. “Conservation of structure and mechanism in primary and secondary transporters exemplified by SiaP, a sialic acid binding virulence factor from Haemophilus influenzae”. In: *The Journal of biological chemistry* 281.31 (2006), pp. 22212–22222.

[98] R. P. Berntsson et al. “A structural classification of substrate-binding proteins”. In: *FEBS letters* 584.12 (2010), pp. 2606–2617.

[99] T. Gangi Setty et al. “Bacterial periplasmic sialic acid-binding proteins exhibit a conserved binding site”. In: *Acta crystallographica. Section D, Biological crystallography* 70.7 (2014), pp. 1801–1811.

[100] M. Fischer et al. “Caught in a TRAP: substrate-binding proteins in secondary transport”. In: *Trends in microbiology* 18.10 (2010), pp. 471–478.

[101] C. Mulligan et al. “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”. In: *The Journal of biological chemistry* 287.5 (2012), pp. 3598–3608.

[102] R. Rabus et al. “TRAP transporters: an ancient family of extracytoplasmic solute-receptor-dependent secondary active transporters”. In: *Microbiology* 145.12 (1999), pp. 3431–3445.

[103] N. R. Wyborn et al. “Topological analysis of DctQ, the small integral membrane protein of the C4-dicarboxylate TRAP transporter of Rhodobacter capsulatus”. In: *FEMS microbiology letters* 194.1 (2006), pp. 13–17.

[104] M. F. Peter et al. “The structure of HiSiaQM defines the architecture of tripartite ATP-independent periplasmic (TRAP) transporters”. In: *bioRxiv* doi: 10.1101/2021.12.03.471092 (2021).

[105] A. Krogh et al. “Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes”. In: *Journal of molecular biology* 305.3 (2001), pp. 567–580.

[106] S. Ovchinnikov, H. Kamisetty, and D. Baker. “Robust and accurate prediction of residue-residue interactions across protein interfaces using evolutionary information”. In: *eLife* 3 (2014), e02030.

[107] C. Mulligan et al. “The substrate-binding protein imposes directionality on an electrochemical sodium gradient-driven TRAP transporter”. In: *Proceedings of the National Academy of Sciences of the United States of America* (2009).

[108] Y. Huang et al. “Structure and mechanism of the glycerol-3-phosphate transporter from Escherichia coli”. In: *Science* 301.5633 (2003), pp. 616–620.

[109] B. Leader, Q. J. Baca, and D. E. Golan. “Protein therapeutics: a summary and pharmacological classification”. In: *Nature reviews. Drug discovery* 7.1 (2008), pp. 21–39.

[110] L. Hu, J. P. Fawkett, and J. Gu. “Protein target discovery of drug and its reactive intermediate metabolite by using proteomic strategy”. In: *Acta Pharmaceutica Sinica B* 2.2 (2012), pp. 126–136.

[111] D. C. Swinney and J. Anthony. “How were new medicines discovered?” In: *Nature reviews. Drug discovery* 10.7 (2011), pp. 507–519.

[112] A. Manglik et al. “Structure-based discovery of opioid analgesics with reduced side effects”. In: *Nature* 537.7619 (2016), pp. 185–190.

[113] Callaway E. “Revolutionary cryo-EM is taking over structural biology”. In: *Nature* 578.7794 (2020), p. 201.

[114] E. Callaway. “The revolution will not be crystallized: a new method sweeps through structural biology”. In: *Nature* 525.7568 (2015), pp. 172–174.

[115] M. G. Rossmann et al. “Combining X-ray crystallography and electron microscopy”. In: *Structure* 13.3 (2005), pp. 355–362.

[116] G. C. Lander et al. “Complete subunit architecture of the proteasome regulatory particle”. In: *Nature* 482.7384 (2012), pp. 186–191.

[117] A. Loquet et al. “Atomic model of the type III secretion system needle”. In: *Nature* 486.7402 (2012), pp. 276–279.

[118] D. Goldfarb and S. Stoll. "EPR Spectroscopy: Fundamentals and Methods". Wiley, 2018.

[119] A. Schweiger and G. Jeschke. "Principles of Pulse Electron Paramagnetic Resonance". Oxford University Press, 2001.

[120] M. A. Hemminga and L. Berliner. "ESR Spectroscopy in Membrane Biophysics". Springer, 2007.

[121] M. H. Levitt. "Spin Dynamics: Basics of Nuclear Magnetic Resonance". Wiley, 2008.

[122] C. Altenbach et al. “Structural studies on transmembrane proteins. 2. Spin labeling of bacteriorhodopsin mutants at unique cysteines”. In: *Biochemistry* 28.19 (1989), pp. 7806–7812.

[123] L. J. Berliner et al. “A novel reversible thiol-specific spin label: papain active site labeling and inhibition”. In: *Analytical biochemistry* 119.2 (1982), pp. 450–455.

[124] K. I. Furukawa and Y. Tonomura. “The conventional and saturation transfer electron paramagnetic resonance of spin-labeled myosin subfragment-1 in the presence of F-actin and nucleotides”. In: *Journal of biochemistry* 92.4 (1982), pp. 1219–1225.

[125] M. R. Fleissner et al. “Site-directed spin labeling of a genetically encoded unnatural amino acid”. In: *Proceedings of the National Academy of Sciences of the United States of America* 106.51 (2009), pp. 21637–21642.

[126] G. W. Reginsson and O. Schiemann. “Spin Labeling of DNA and RNA”. In: *Encyclopedia of Biophysics* (2013).

[127] R. Ward et al. “EPR distance measurements in deuterated proteins”. In: *Journal of magnetic resonance* 207.1 (2010), pp. 164–167.

[128] G. Hagelueken et al. “MtsslWizard: In Silico Spin-Labeling and Generation of Distance Distributions in PyMOL”. In: *Applied magnetic resonance* 42.3 (2012), pp. 377–391.

[129] G. Hagelueken, D. Abdullin, and O. Schiemann. “mtsslSuite: Probing Biomolecular Conformation by Spin-Labeling Studies.” In: *Methods in enzymology* 563 (2015), pp. 595–622.

[130] G. Jeschke. “MMM: A toolbox for integrative structure modeling”. In: *Tools for Protein Science* 27.1 (2018), pp. 76–85.

[131] M. M. Hatmal et al. “Computer modeling of nitroxide spin labels on proteins”. In: *Biopolymers* 97.1 (2012), pp. 35–44.

[132] H. J. Steinhoff. “Inter- and intra-molecular distances determined by EPR spectroscopy and site-directed spin labeling reveal protein-protein and protein-oligonucleotide interaction”. In: *Biological chemistry* 385.10 (2004), pp. 913–920.

[133] S. Stoll and A. Schweiger. “EasySpin, a comprehensive software package for spectral simulation and analysis in EPR”. In: *Journal of magnetic resonance* 178.1 (2006), pp. 42–55.

[134] D. J. Schneider and J. H. Freed, eds. "Calculating Slow Motional Magnetic Resonance Spectra" In: Spin Labeling. Biological Magnetic Resonance, vol 8. Springer, 1989.

[135] D. J. Schneider and J. H. Freed, eds. "Spin Relaxation and Motional Dynamics" In: Advances in Chemical Physics: Lasers, Molecules, and Methods, Volume 73. John Wiley & Sons, Inc., 2007.

[136] N. R. Zaccai, I. N. Serdyuk, and J. Zaccai. *Methods in Molecular Biophysics*. Cambridge University Press, 2017.

[137] G. S. Worswick et al. “Deep neural network processing of DEER data”. In: *Science advances* 4.8 (2018), eaat5218.

[138] J. W. Johnston et al. “Characterization of the N-Acetyl-5-neuraminic Acid-binding Site of the Extracytoplasmic Solute Receptor (SiaP) of Nontypeable Haemophilus influenzae Strain 2019”. In: *The Journal of biological chemistry* 283.2 (2007), pp. 855–865.

[139] E. Severi et al. “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”. In: *Molecular microbiology* 58.4 (2005), pp. 1173–1185.

[140] N. S. Alexander et al. “RosettaEPR: rotamer library for spin label structure and dynamics”. In: *PloS one* 8.9 (2013), e72851.

[141] F. Tombolato, A. Ferrarini, and J. H. Freed. “Dynamics of the nitroxide side chain in spin-labeled proteins”. In: *The journal of physical chemistry, B* 110.51 (2006), pp. 26248–26259.

[142] S. C. Lovell et al. “The penultimate rotamer library”. In: *Proteins* 40.3 (2000), pp. 389–408.

[143] Z. Guo et al. “Structural determinants of nitroxide motion in spin-labeled proteins: solvent-exposed sites in helix B of T4 lysozyme”. In: *Protein science* 17.2 (2008), pp. 228–239.

[144] Y. Polyhach, E. Bordignon, and G. Jeschke. “Rotamer libraries of spin labelled cysteines for protein studies”. In: *Physical chemistry chemical physics* 13.6 (2011), pp. 2356–2366.

[145] D. Marsch. "Spin-Label Electron Paramagnetic Resonance Spectroscopy". CRC Press, 2020.

[146] R. Langen et al. “Crystal structures of spin labeled T4 lysozyme mutants: implications for the interpretation of EPR spectra in terms of structure”. In: *Biochemistry* 39.29 (2000), pp. 8396–8405.

[147] B. Carrington et al. “Natural Conformational Sampling of Human TNFalpha Visualized by Double Electron-Electron Resonance”. In: *Biophysical journal* 113.2 (2017), pp. 371–380.

[148] D. Abdullin, G. Hagelueken, and O. Schiemann. “Determination of nitroxide spin label conformations via PELDOR and X-ray crystallography”. In: *Physical chemistry chemical physics* 18.15 (2016), pp. 10428–10437.

[149] T. F. Cunningham et al. “Rotameric preferences of a protein spin label at edge-strand beta-sheet sites”. In: *Protein science* 25.5 (2016), pp. 1049–1060.

[150] E. C. Hulme and M. A. Trevethick. “Ligand binding assays at equilibrium: validation and interpretation”. In: *British journal of pharmacology* 161.6 (2010), pp. 1219–1237.

[151] W. Ma, L. Yang, and L. He. “Overview of the detection methods for equilibrium dissociation constant KD of drug-receptor interaction”. In: *Journal of pharmaceutical analysis* 8.3 (2018), pp. 147–152.

[152] J. W. Johnston et al. “Characterization of the N-acetyl-5-neuraminic acid-binding site of the extracytoplasmic solute receptor (SiaP) of nontypeable Haemophilus influenzae strain 2019”. In: *The Journal of biological chemistry* 283.2 (2008), pp. 855–865.

[153] X. Duan and F. A. Quiocho. “Structural evidence for a dominant role of nonpolar interactions in the binding of a transport/chemosensory receptor to its highly polar ligands”. In: *Biochemistry* 41.3 (2002), pp. 706–712.

[154] J. L. Nadeau. "Introduction to Experimental Biophysics: Biological Methods for Physical Scientists". CRC Press, 2017.

[155] Sevcik C. “Caveat on the Boltzmann distribution function use in biology”. In: *Progress in biophysics and molecular biology* 127 (2017), pp. 33–42.

[156] J Resendiz-Munoz et al. “Mathematical model of Boltzmanns sigmoidal equation applicableto the set-up of the RF-magnetron co-sputtering in thin filmsdeposition of BaxSr1-xTiO3”. In: *Bulletin of materials science* 40.5 (2017), pp. 1043–1047.

[157] M. Peleg and M. G. Corradini. “Microbial growth curves: what the models tell us and what they cannot”. In: *Critical reviews in food science and nutrition* 51.10 (2011), pp. 917–945.

[158] I. de Vera et al. “Pulsed EPR distance measurements in soluble proteins by site-directed spin labeling (SDSL)”. In: *Current protocols in protein science* 74 (2013), pp. 17.17.1–17.17.29.

[159] A. Ward et al. “Expression of prokaryotic membrane transport proteins in Escherichia coli”. In: *Biochemical Society transactions* 27.6 (1999), pp. 893–899.

[160] J. Lecher et al. “Journal of molecular biology”. In: *The crystal structure of UehA in complex with ectoine-A comparison with other TRAP-T binding proteins* 389.1 (2009), pp. 58–73.

[161] C. Tang, C. D. Schwieters, and G. M. Clore. “Open-to-closed transition in apo maltose-binding protein observed by paramagnetic NMR”. In: *Nature* 449.7165 (2007), pp. 1078–1082.

[162] G. A. Bermejo et al. “Ligand-free open-closed transitions of periplasmic binding proteins: the case of glutamine-binding protein”. In: *Biochemistry* 49.9 (2010), pp. 1893–1902.

[163] D. P. Claxton et al. “Ion/substrate-dependent conformational dynamics of a bacterial homolog of neurotransmitter:sodium symporters”. In: *Nature structural & molecular biology* 17.7 (2010), pp. 822–829.

[164] F. Marinelli and G. Fiorin. “Structural Characterization of Biomolecules through Atomistic Simulations Guided by DEER Measurements”. In: *Structure* 27.2 (2019), pp. 359–370.

[165] F. Marinelli et al. “Evidence for an allosteric mechanism of substrate release from membrane-transporter accessory binding proteins”. In: *Proceedings of the National Academy of Sciences of the United States of America* 108.49 (2011), pp. E1285–E1292.

[166] R. Pievo et al. “A rapid freeze-quench setup for multi-frequency EPR spectroscopy of enzymatic reactions”. In: *Chemphyschem* 14.18 (2013), pp. 4094–4101.

[167] T. Hett et al. “Spatiotemporal Resolution of Conformational Changes in Biomolecules by Combining Pulsed Electron-Electron Double Resonance Spectroscopy with Microsecond Freeze-Hyperquenching”. In: *Journal of the American Chemical Society* 143.18 (2021), pp. 6981–6989.

[168] M. Fischer et al. “Tripartite ATP-independent Periplasmic (TRAP) Transporters Use an Arginine-mediated Selectivity Filter for High Affinity Substrate Binding”. In: *The Journal of biological chemistry* 290.45 (2015), pp. 27113–27123.

[169] A. P. Hopkins. “Molecular and biochemical characterisation of SiaP as a sialic acid binding protein component of a TRAP transporter for sialic acid”. PhD thesis. University of York, 2010.

[170] S. Gonin et al. “Crystal structures of an Extracytoplasmic Solute Receptor from a TRAP transporter in its open and closed forms reveal a helix-swapped dimer requiring a cation for alpha-keto acid binding”. In: *BMC structural biology* 7.11 (2007).

[171] N. Akiyama, K. Takeda, and K. Miki. “Crystal structure of a periplasmic substrate-binding protein in complex with calcium lactate”. In: *Journal of molecular biology* 392.3 (2009), pp. 559–565.

[172] M. F. Peter et al. “Triggering Closure of a Sialic Acid TRAP Transporter Substrate Binding Protein through Binding of Natural or Artificial Substrates”. In: *Journal of molecular biology* 433.3 (2021), p. 166756.

[173] S. B. Shuker et al. “Discovering high-affinity ligands for proteins: SAR by NMR”. In: *Science* 274.5292 (1996), pp. 1531–1534.

[174] R. A. Laskowski and M. B. Swindells. “LigPlot+: multiple ligand-protein interaction diagrams for drug discovery”. In: *Journal of chemical information and modeling* 51.10 (2011), pp. 2778–2786.

[175] H Liu and J. H. Naismith. “An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol”. In: *BMC biotechnology* 8.91 (2008).

[176] C. Ludwig. “Diffusion zwischen ungleich erwärmten Orten gleich zusammengesetzter Lösungen”. In: *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften: Mathematisch-Naturwissenschaftliche Classe* 20.539 (1856).

[177] M. Asmari et al. “Thermophoresis for characterizing biomolecular interaction”. In: *Methods* 146 (2018), pp. 107–119.

[178] M. Jerabek-Willemsen et al. “Molecular Interaction Studies Using Microscale Thermophoresis”. In: *ASSAY and Drug Development Technologies* (2011).

[179] U. Mueller et al. “Facilities for macromolecular crystallography at the Helmholtz-Zentrum Berlin”. In: *Journal of synchrotron radiation* 19.3 (2012), pp. 442–449.

[180] M. D. Winn et al. “Overview of the CCP4 suite and current developments”. In: *Acta crystallographica. Section D, Biological crystallography* 67.4 (2011), pp. 235–242.

[181] T. G. Battye et al. “iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM”. In: *Acta crystallographica. Section D, Biological crystallography* 67.4 (2011), pp. 271–281.

[182] A. J. McCoy et al. “Phaser crystallographic software”. In: *Journal of applied crystallography* 40.4 (2007), pp. 658–674.

[183] P. D. Adams et al. “PHENIX: a comprehensive Python-based system for macromolecular structure solution”. In: *Acta crystallographica. Section D, Biological crystallography* 66.2 (2010), pp. 213–221.

[184] V. B. Chen et al. “MolProbity: all-atom structure validation for macromolecular crystallography”. In: *Acta crystallographica. Section D, Biological crystallography* 66.1 (2010), pp. 12–21.

[185] E. L. Hahn. “Spin Echoes”. In: *Physical Review* 80.4 (1950), pp. 580–594.

[186] M. Pannier et al. “Dead-time free measurement of dipole-dipole interactions between electron spins”. In: *Journal of magnetic resonance* 213.2 (2011), pp. 331–340.

[187] T. N. Petersen et al. “SignalP 4.0: discriminating signal peptides from transmembrane regions”. In: *Nature methods* 8.10 (2011), pp. 785–786.

[188] J. J. Almagro Armenteros et al. “SignalP 5.0 improves signal peptide predictions using deep neural networks”. In: *Nature biotechnology* 37.4 (2019), pp. 420–423.

[189] A. Waterhouse et al. “SWISS-MODEL: homology modelling of protein structures and complexes”. In: *Nucleic acids research* 46.W1 (2018), W296–W303.

A. Appendix of data

A.1. SignalP

Figure A.1 summarizes the analysis of the signaling sequence prediction performed with the SignalP server [187][188]. The initial 22 amino acids of the protein (MKTINKITI-AILTLSAAASVNA) were identified to be the targeting peptide for the pathway which delivers the transporter to the periplasmic space. The identified cleavage site (VNA|A) also satisfies the -3,-1 rule that states residues at positions -1 (in this case A) and -3 (in this case V) relative to the cleavage site must be small and neutral to ensure proper cleavage. The rest of the sequence belongs to the SBP itself.

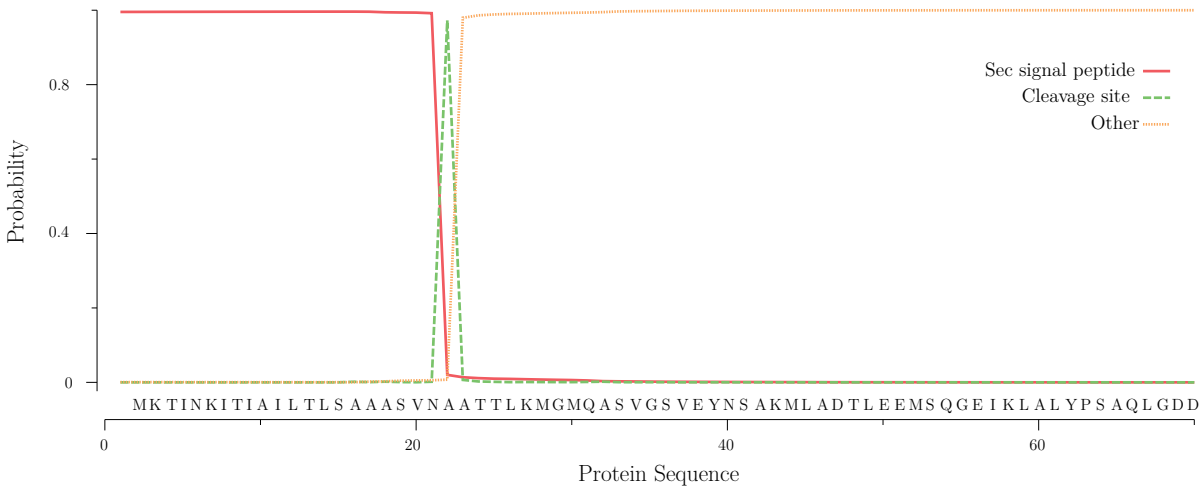


Figure A.1. – Identification of the signaling peptide in VcSiaP. The red line indicates the probability of the sequence being part of the signaling peptide. The green line indicates the cleavage site and the yellow line marks the rest of the sequence as functional SBP.

A.2. Multiple Sequence Alignment

Figure A.2 depicts a multiple sequence alignment of VcSiaP and HiSiaP. Figure A.3 shows a multiple sequence alignment of QM-domains from seven closely related organisms which possess a TRAP transporter.
The listed organisms are *Vibrio cholerae* (VcSiaQM), *Vibrio vulnificus* (VvSiaQM), *Photobacterium profundum* (PbpraSiaQM), *Pasteurella multocida* (PmSiaQM), *Haemophilus somnus* (HsSiaQM), *Haemophilus influenzae* (HiSiaQM) and *Fusobacterium nucleatum* (FnSiaQM).

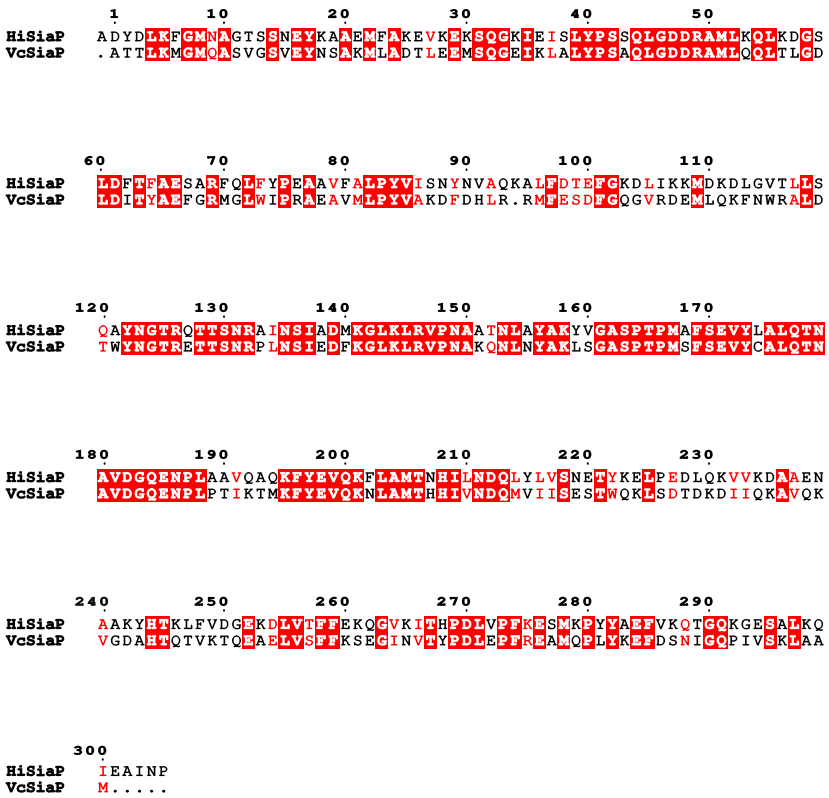


Figure A.2. – Multiple sequence alignment of VcSiaP and HiSiaP. The sequence identity is 50%.

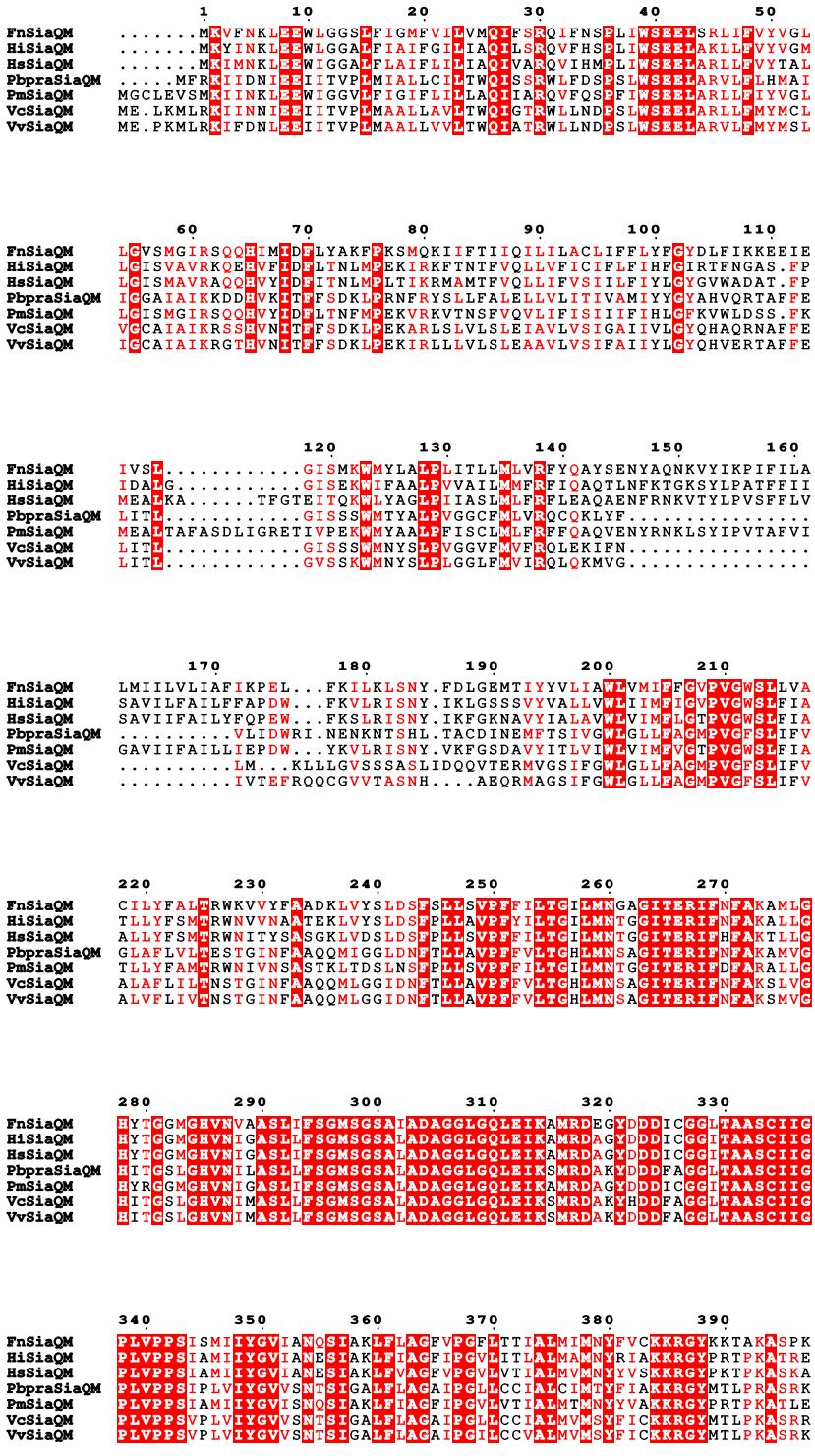


Figure A.3. is spread on two pages. Caption is given on the next page.

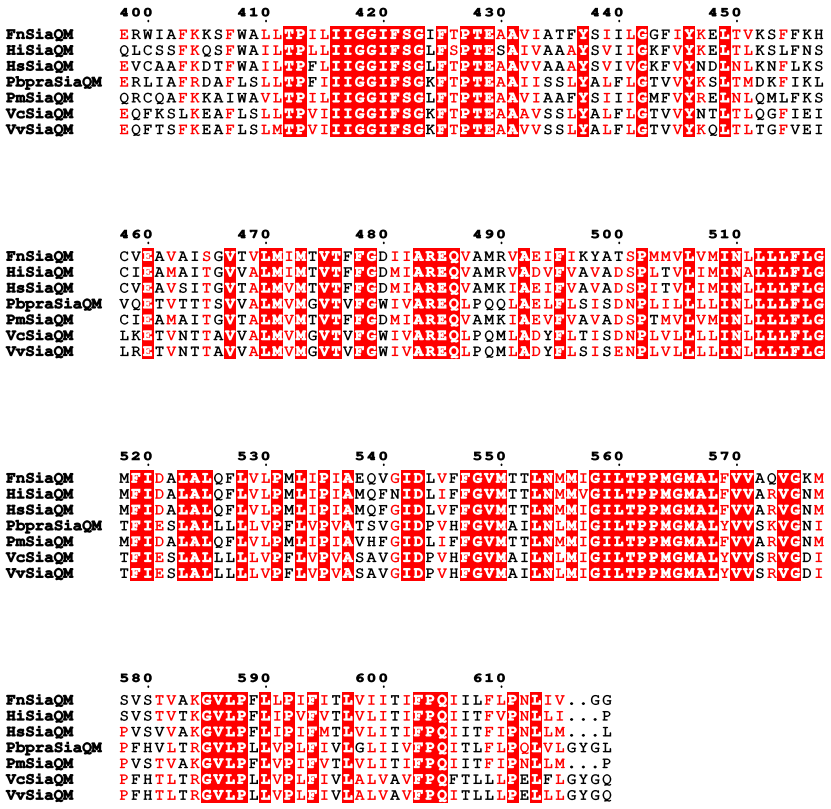


Figure A.3. – Multiple sequence alignment of seven QM-domains.

A.3. PELDOR raw data of the four spin-labeled double-cysteine mutants

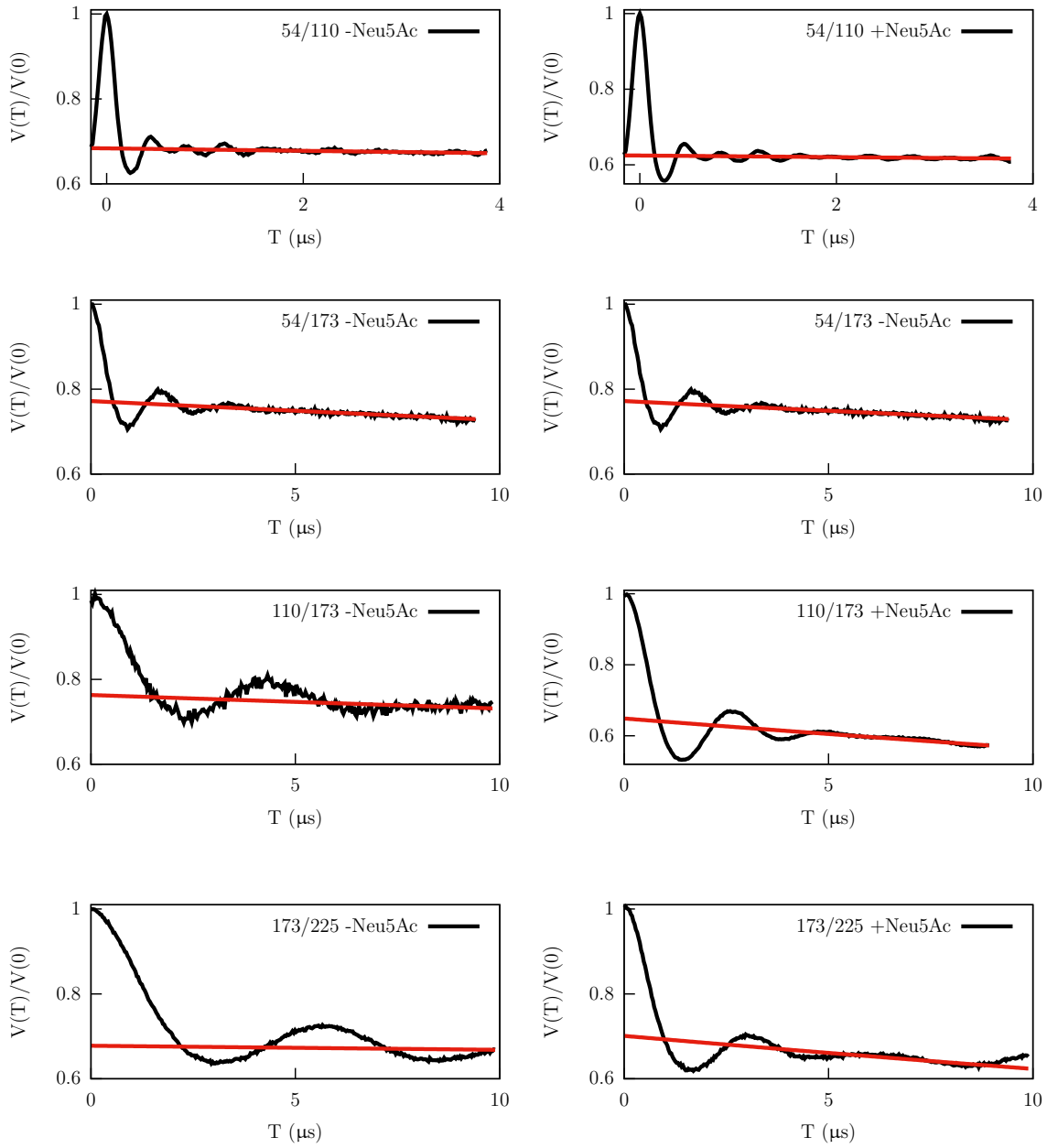


Figure A.4. – PELDOR raw data of the four VcSiaP spin-labeled double-cysteine mutants. The timetraces are not background corrected. The respective background which was then subtracted from the recording is indicated by the red line. Each mutant was recorded once without the substrate Neu5Ac and once with 1 mM Neu5Ac.

A.4. PELDOR raw data of the binding assay

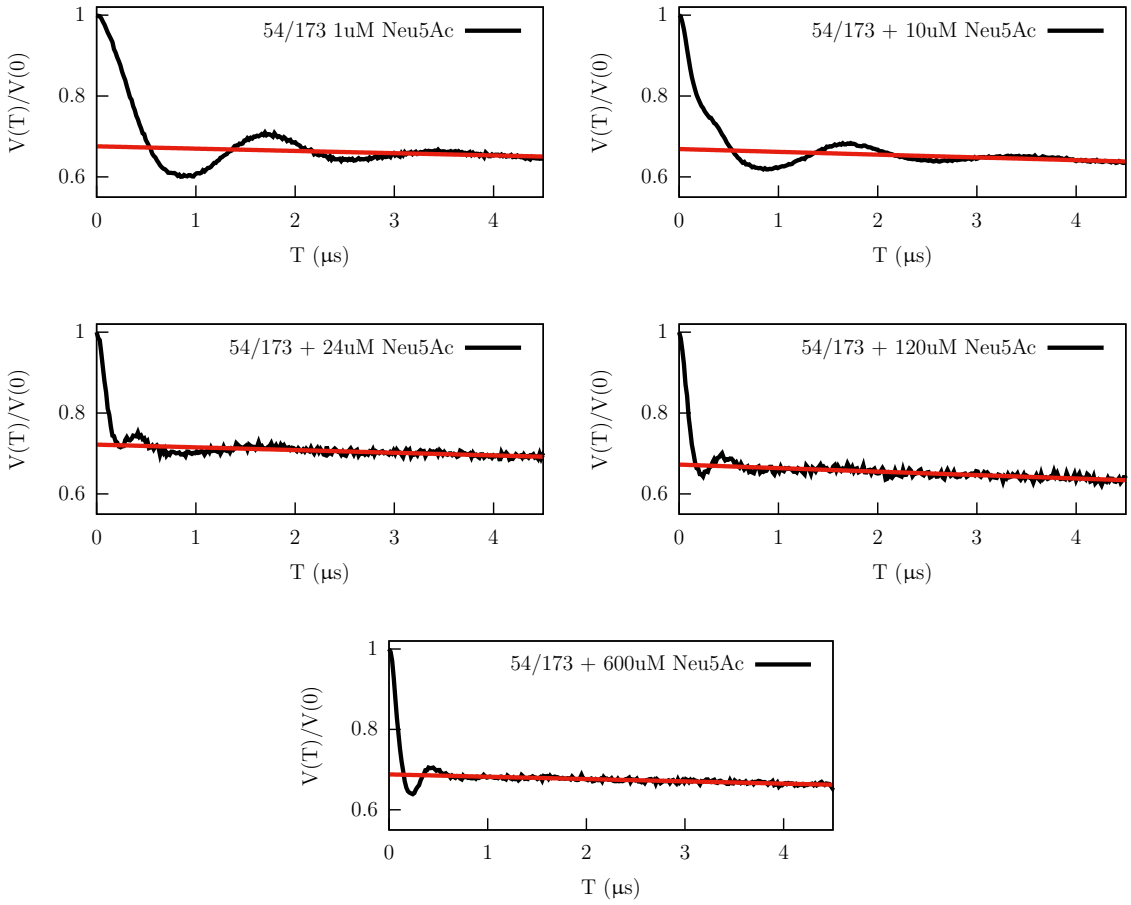


Figure A.5. – PELDOR raw data of the binding assay. The binding assay was recorded with the VcSiaP Q54R1/L173R1 mutant and Neu5Ac concentrations ranging from 1 µM − 600 µM. The timetraces are not background corrected. The respective background which was then subtracted from the recording is indicated the red line.

A.5. Dihedral angles of the R1 label

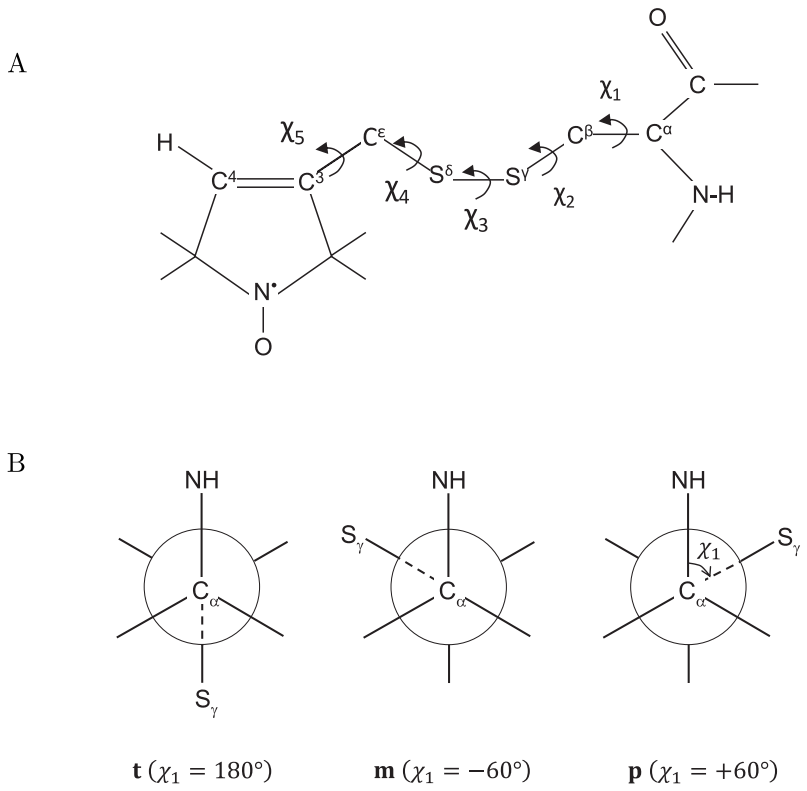
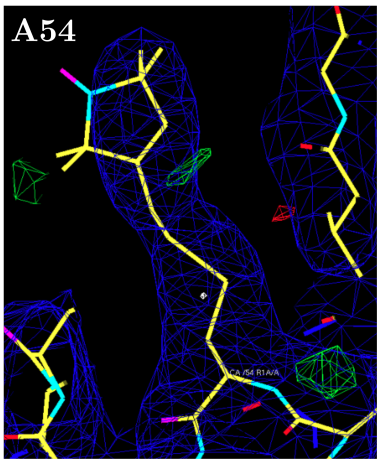


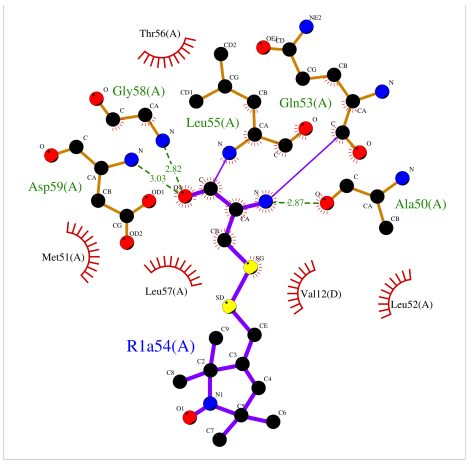
Figure A.6. – Dihedral angles of MTSSL. Panel A shows the dihedral angles $\chi_1 - \chi_5$ of the spin label MTSSL. Panel B shows the preferred rotamers of χ_1.

Table A.1. – Dihedral angles of R1 side-chains. The greek and numerical subscripts represent side chain atoms of the label and the nitroxide ring (see figure A.6).

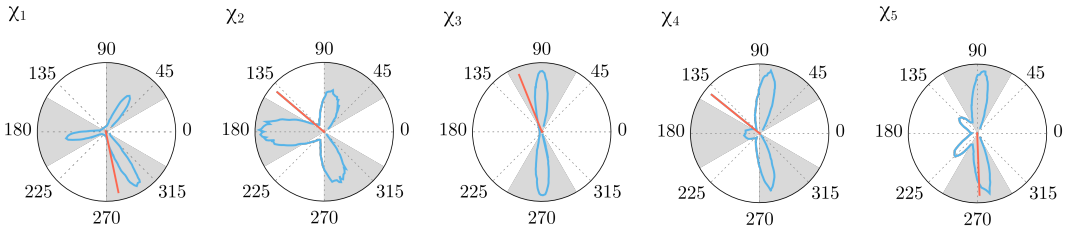
Dihedral angle	Bond
χ_1	NH-C_α-C_β-S_γ
χ_2	C_α-C_β-S_γ-S_δ
χ_3	C_β-S_γ-S_δ-C_ϵ
χ_4	S_γ-S_δ-C_ϵ-C_3
χ_5	S_δ-C_ϵ-C_3-C_4

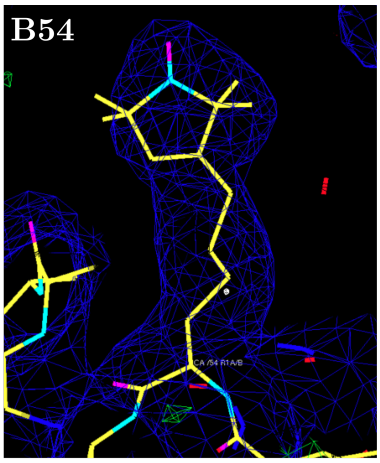
A.6. Dihedral angles of the R1 label in the VcSiaP Q54R1/L173C dimer

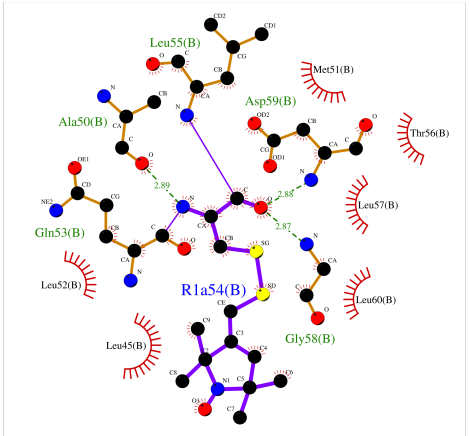




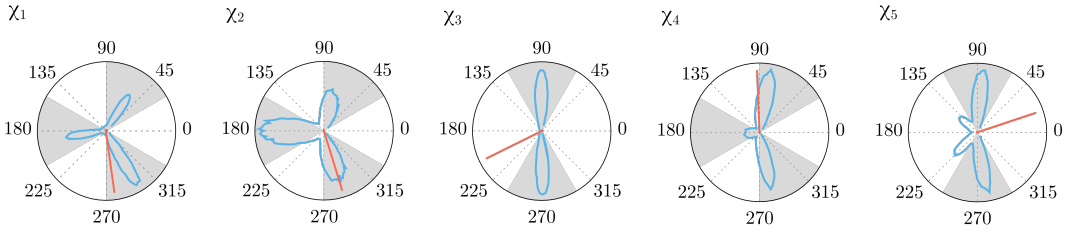
$\chi_{1\text{-}5}$ =281°, 139°, 111°, 141°, 272°

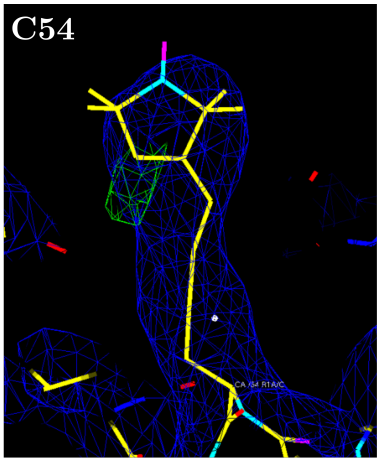


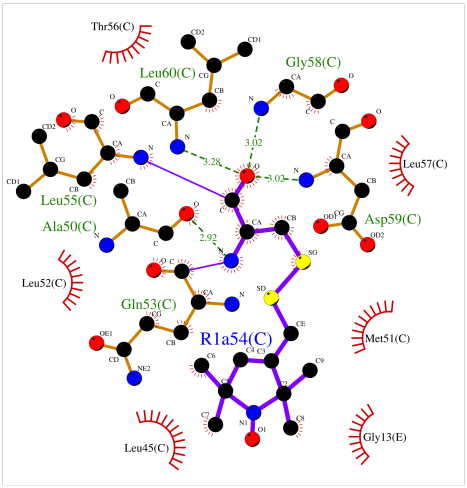




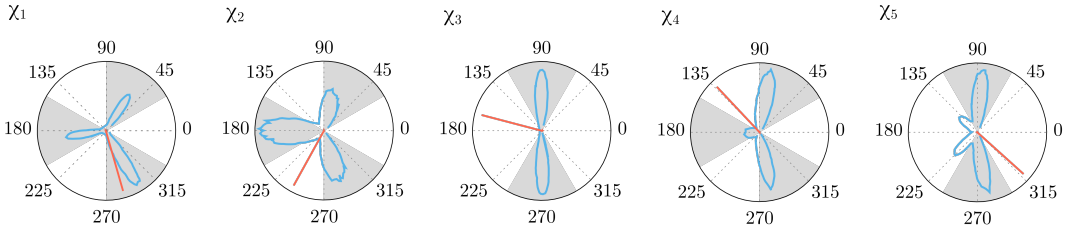
$\chi_{1\text{-}5}$ =280°, 287°, 206°, 92°, 18°

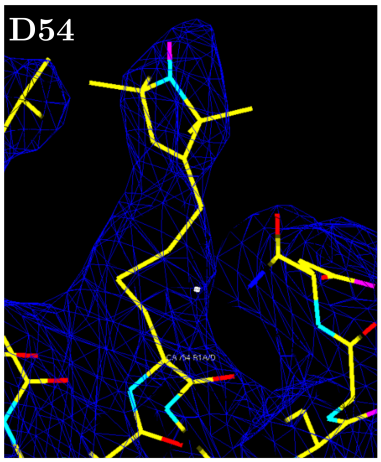


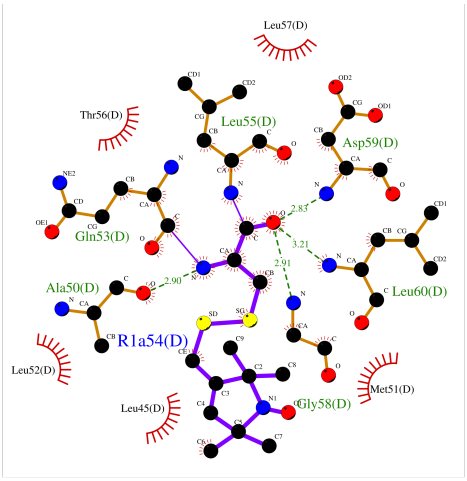




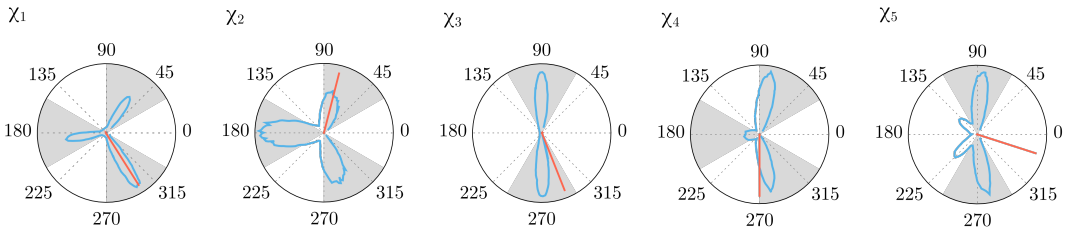
χ_{1-5} =288°, 256°, 165°, 133°, 318°

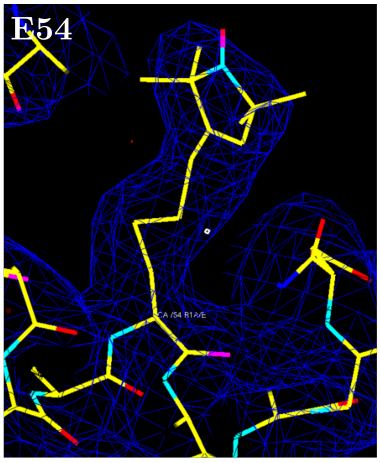


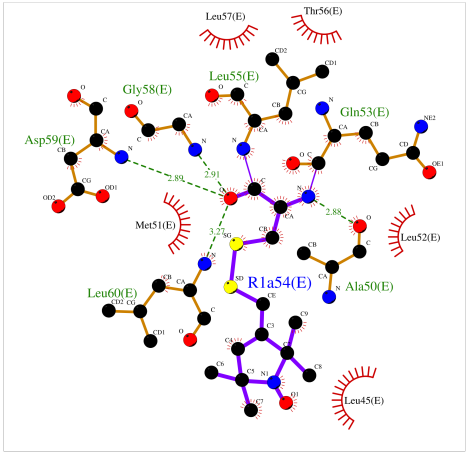




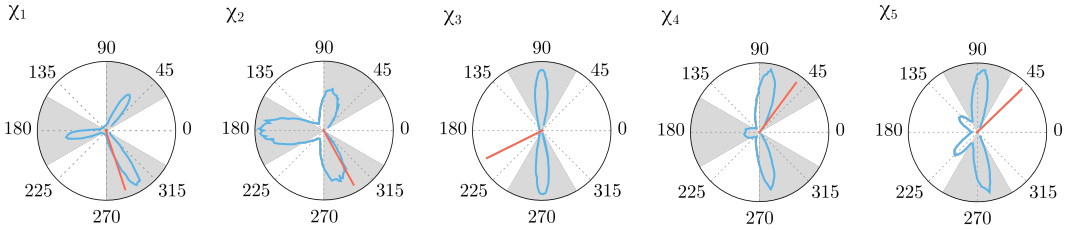
χ_{1-5} =303°, 75°, 291°, 270°, 342°

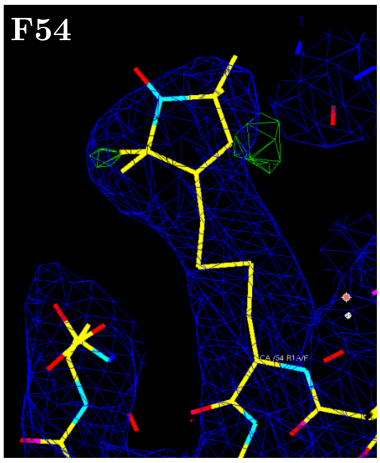


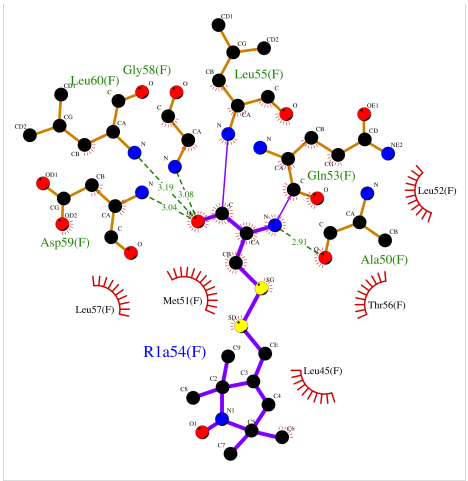




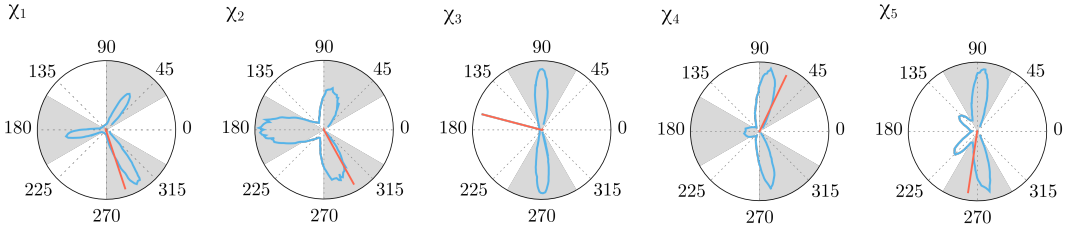
$\chi_{1\text{-}5}$ =287°, 308°, 216°, 53°, 44°

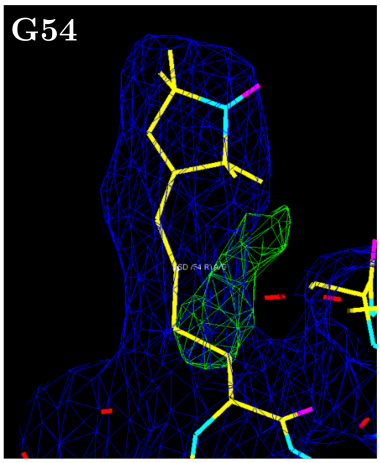


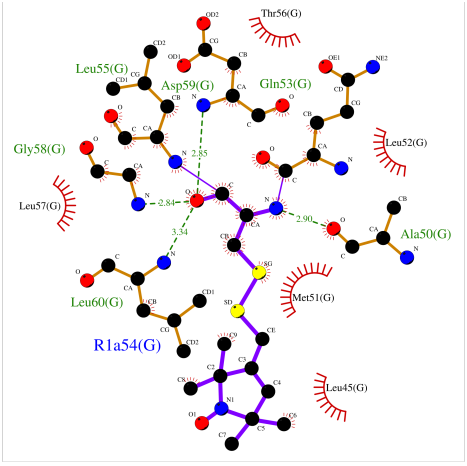




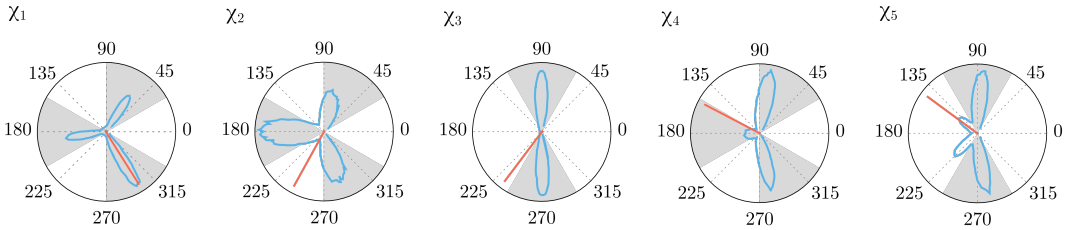
$\chi_{1\text{-}5}$ =277°, 299°, 171°, 64°, 261°

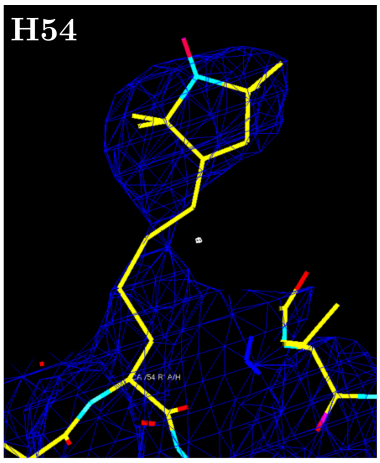


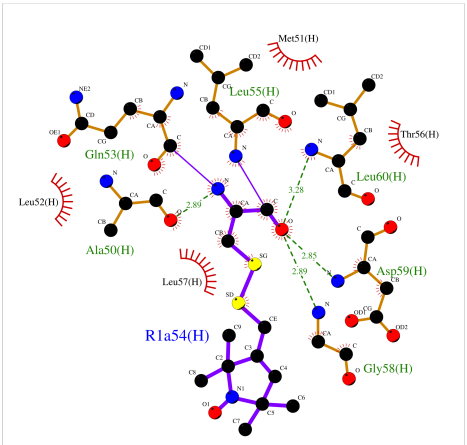




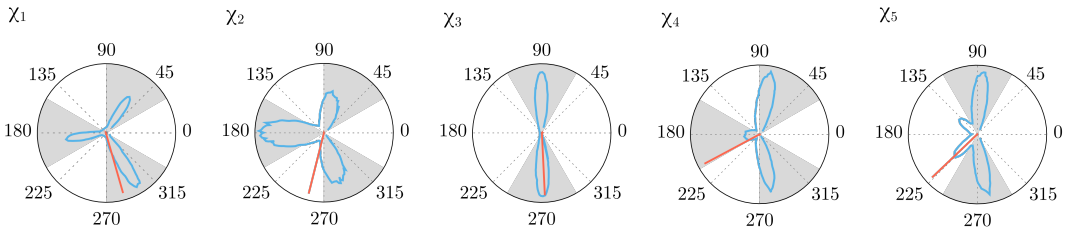
$\chi_{1\text{-}5}$ =307°, 255°, 232°, 152°, 143°

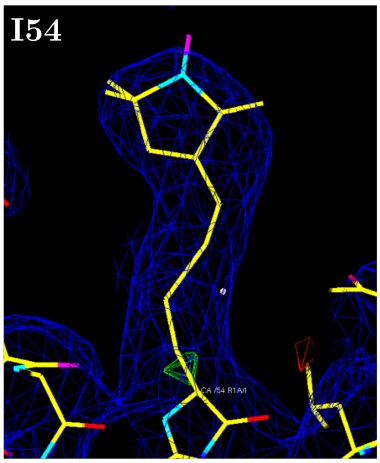


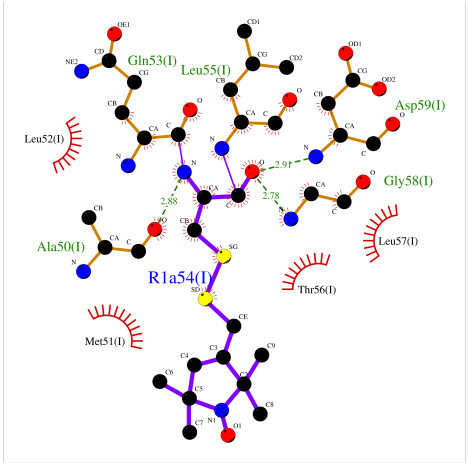




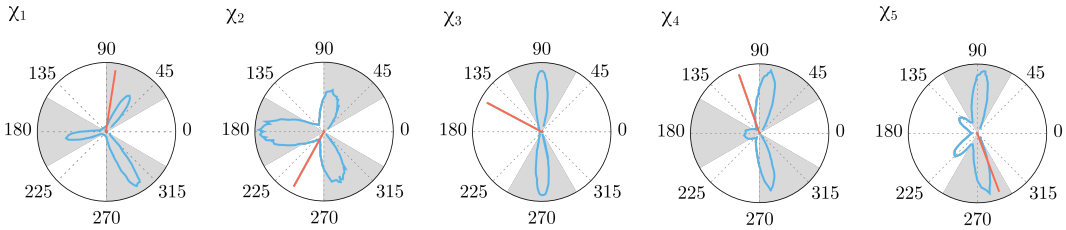
$\chi_{1\text{-}5}$ =286°, 259°, 272°, 208°, 223°

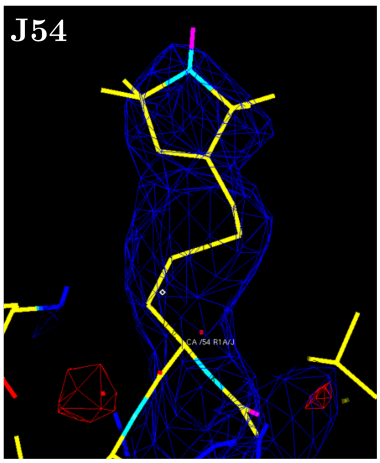


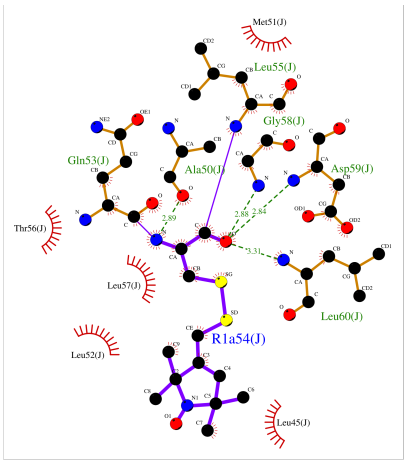




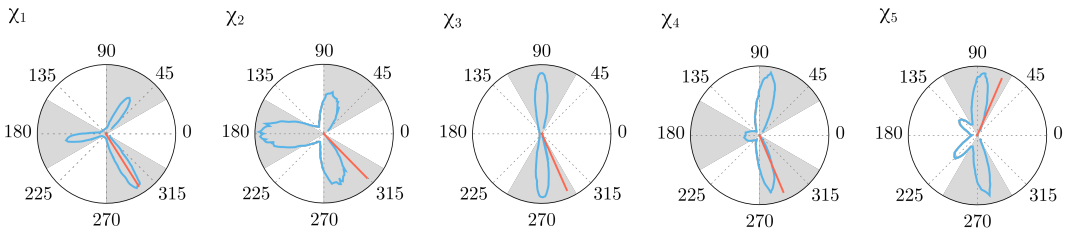
$\chi_{1\text{-}5}$ =81°, 241°, 152°, 109°, 290°







$\chi_{1\text{-}5}$ =302°, 314°, 294°, 292°, 66°



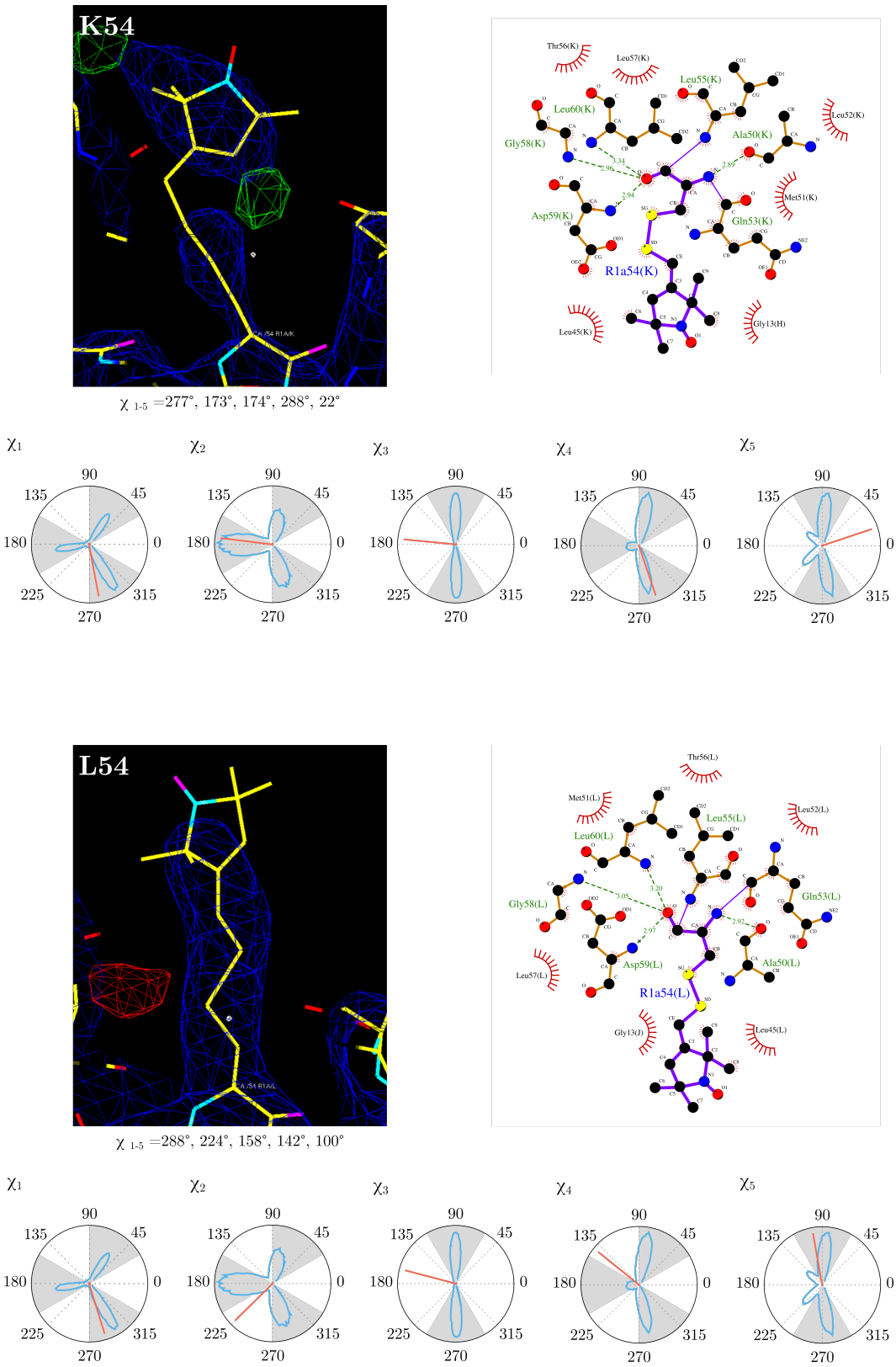


Figure A.7. – Detailed visualization of the dihedral angles of the R1 label in the VcSiaP Q54R1/L173C dimer. For each label (A 54−L 54) the electron density (blue mesh), the Ligplot and the polar plots of the angles are shown.

A.7. Protein mass spectrometry of a copurified E. coli protein

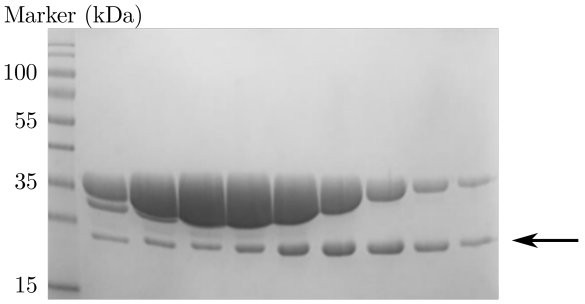


Figure A.8. – SDS PAGE of a size exclusion chromatogram of VcSiaP R125K.

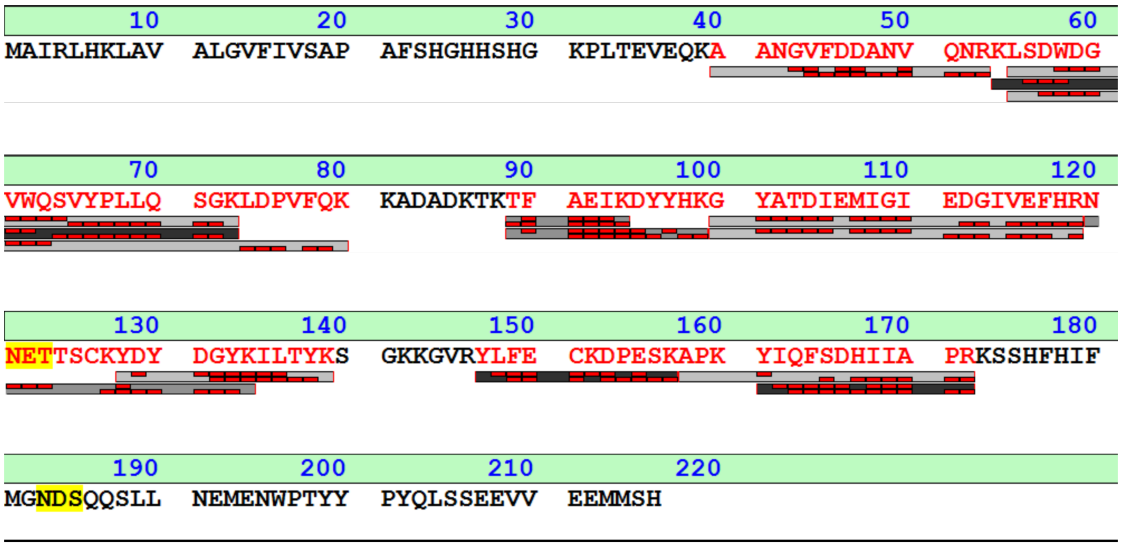


Figure A.9. – Sequencing analysis of an emerging *E.coli* proteins during expression and purification. Alignment with zinc/cadmium-binding protein results in 54.60% sequence coverage.

A.8. CW of the binding cleft

R125A and R125K

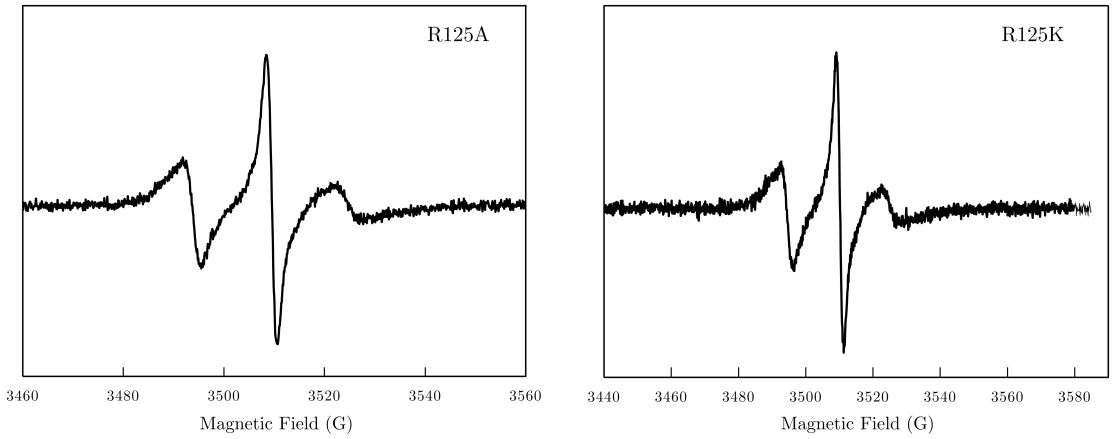


Figure A.10. – *cw*-X-band spectra measured at room temperature of VcSiaP Q54R1/L173R1 R125A and VcSiaP Q54R1/L173R1 R125K.

H207A and H207Q

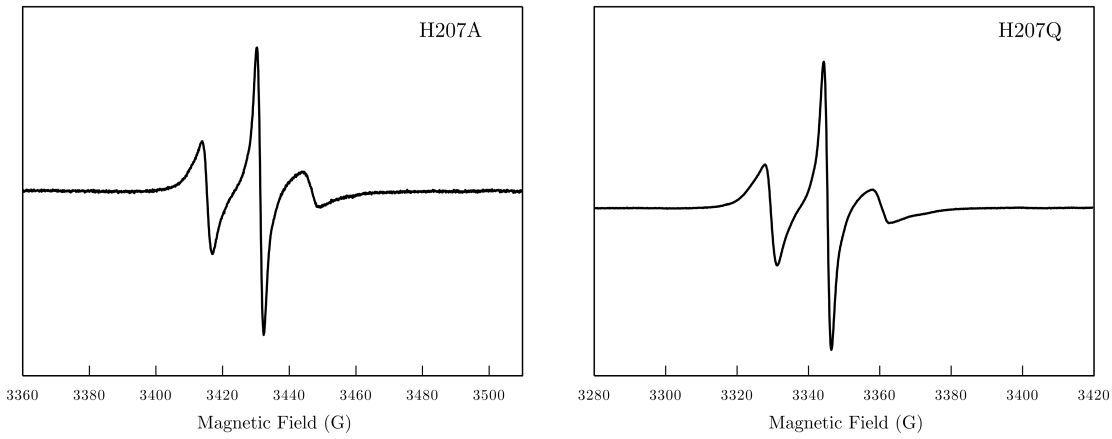


Figure A.11. – *cw*-X-band spectra measured at room temperature of VcSiaP Q54R1/L173R1 H207A and VcSiaP Q54R1/L173R1 H207Q.

E184A, E184Q and E184D

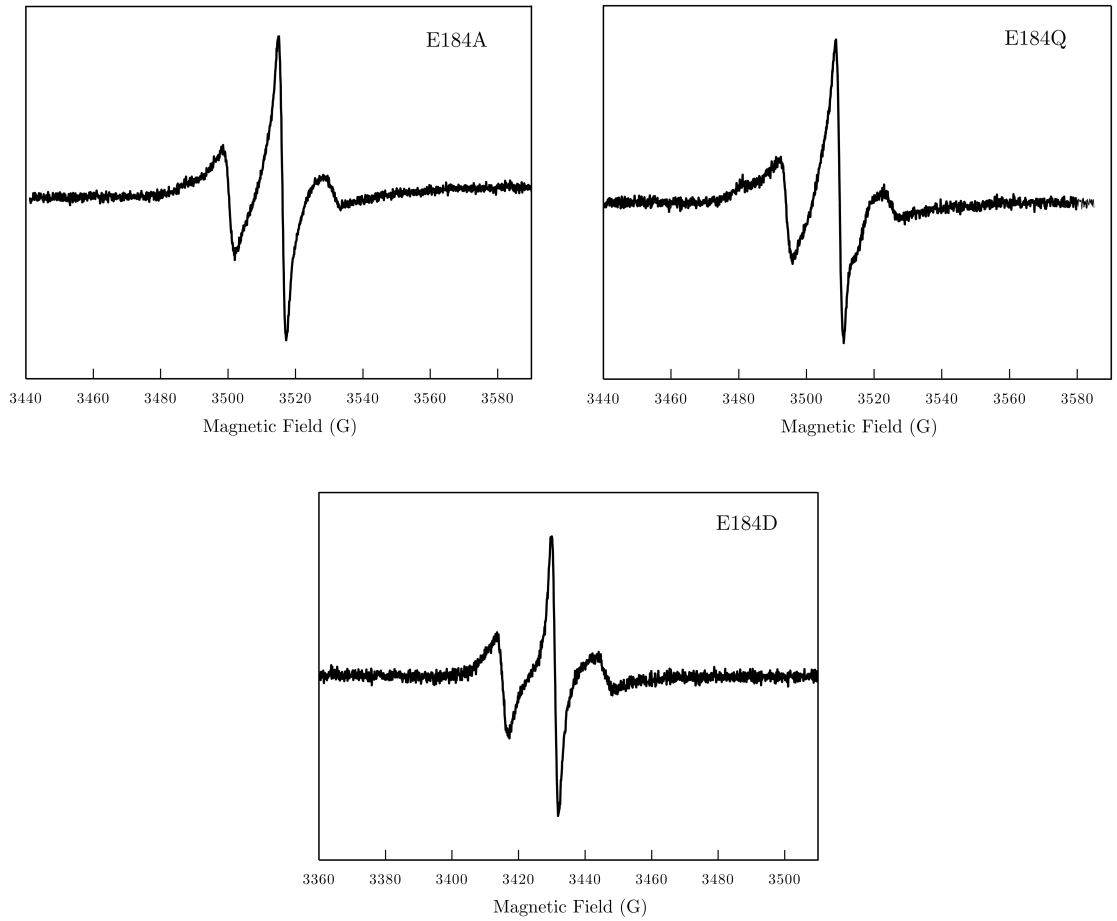


Figure A.12. – *cw*-X-band spectra measured at room temperature of VcSiaP Q54R1/L173R1 E184A, VcSiaP Q54R1/L173R1 E184Q and VcSiaP Q54R1/L173R1 E184D.

A.9. PELDOR of the binding cleft

Figures A.13−A.19 show the raw PELDOR data of the binding cleft mutational analysis. The timetraces are not background corrected. The respective background which was then subtracted from the recording is indicated by red lines in every plot.

VcSiaP Q54R1/L173R1 R125A

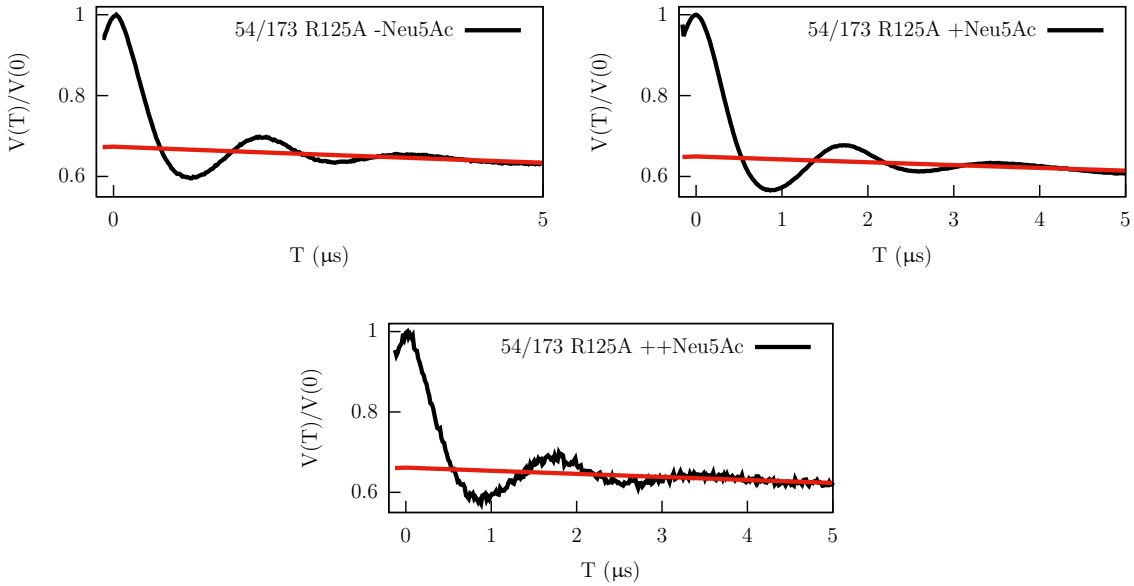


Figure A.13. – PELDOR raw data of the VcSiaP Q54R1/L173R1 R125A mutant.

VcSiaP Q54R1/L173R1 R125K

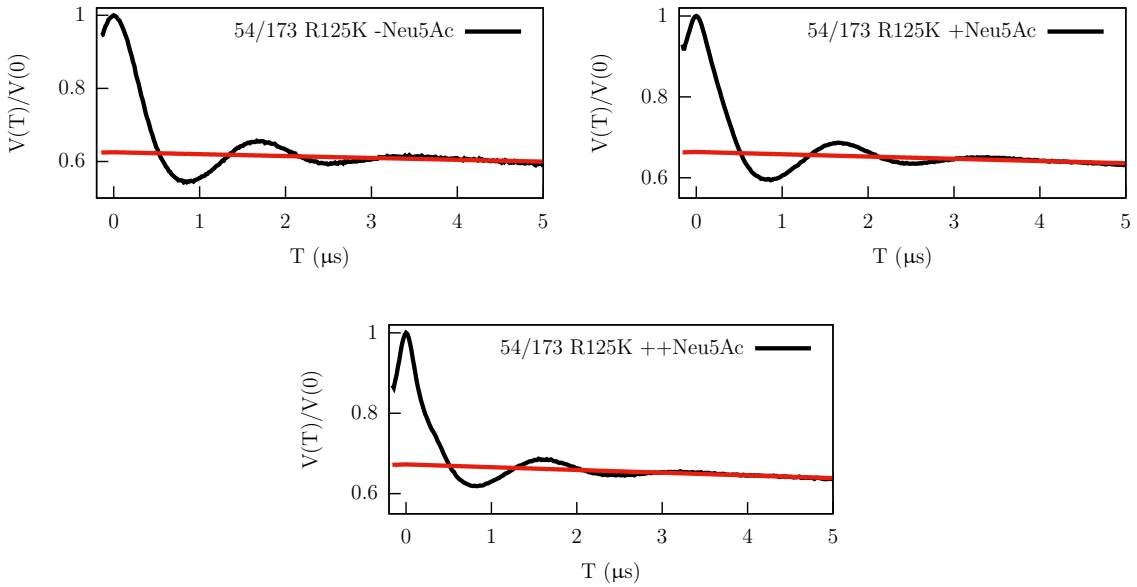


Figure A.14. – PELDOR raw data of the VcSiaP Q54R1/L173R1 R125K mutant.

VcSiaP Q54R1/L173R1 E184A

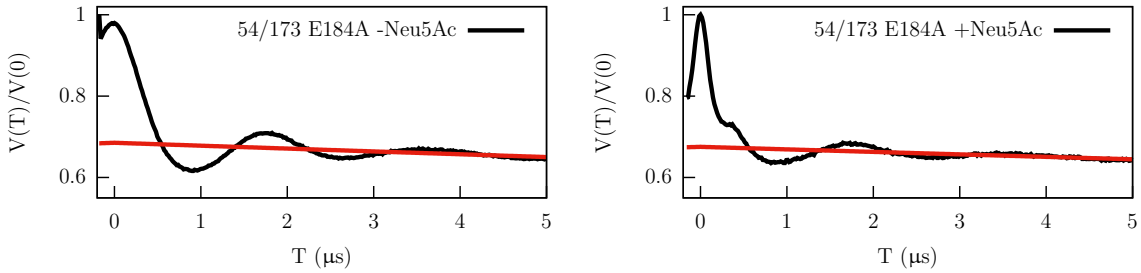


Figure A.15. – PELDOR raw data of the VcSiaP Q54R1/L173R1 E184A mutant.

VcSiaP Q54R1/L173R1 E184D

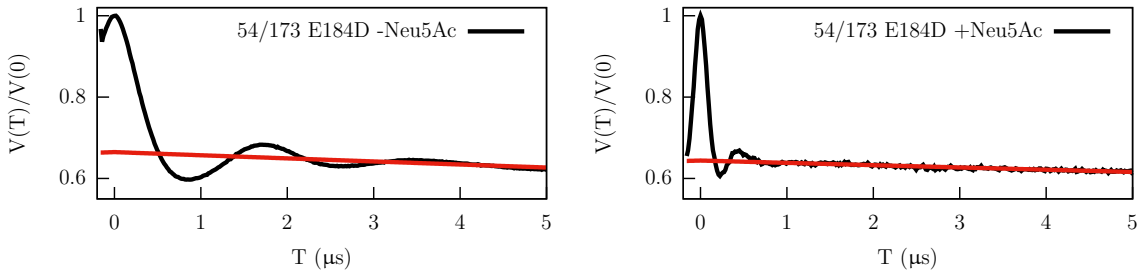


Figure A.16. – PELDOR raw data of the VcSiaP Q54R1/L173R1 E184D mutant.

VcSiaP Q54R1/L173R1 E184Q

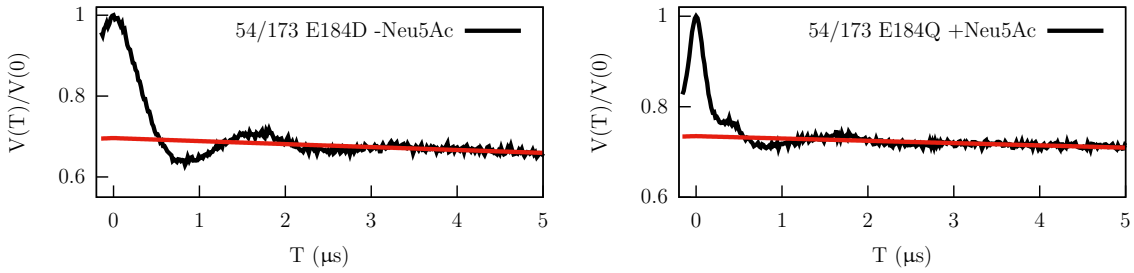


Figure A.17. – PELDOR raw data of the VcSiaP Q54R1/L173R1 E184Q mutant.

VcSiaP Q54R1/L173R1 H207A

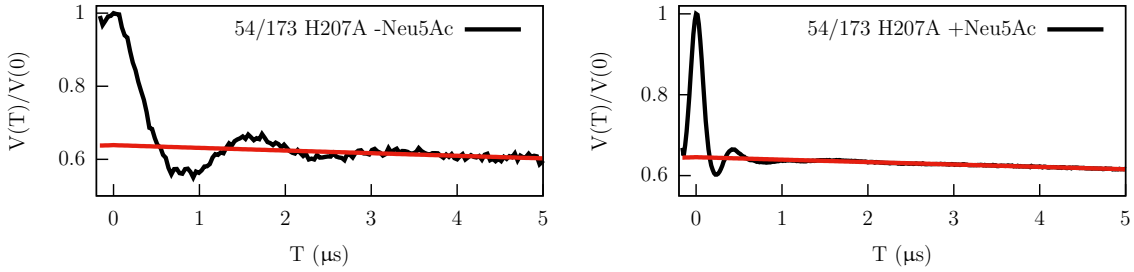


Figure A.18. – PELDOR raw data of the VcSiaP Q54R1/L173R1 H207A mutant.

VcSiaP Q54R1/L173R1 H207Q

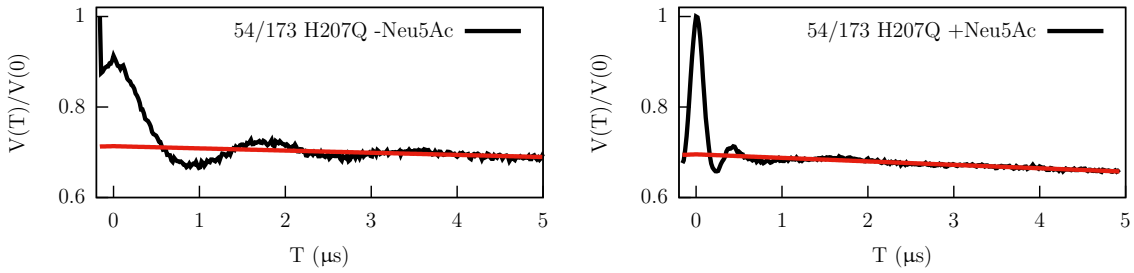
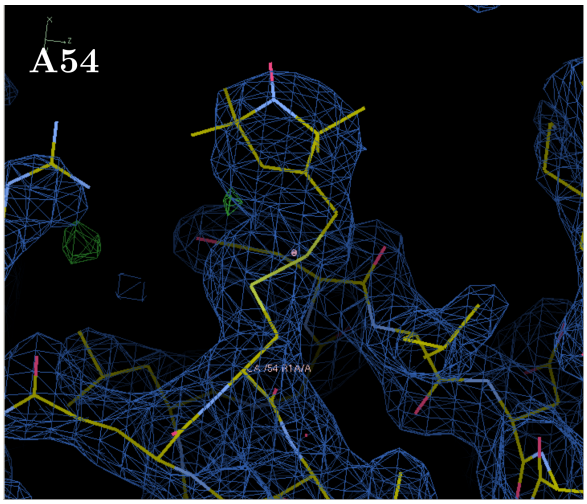


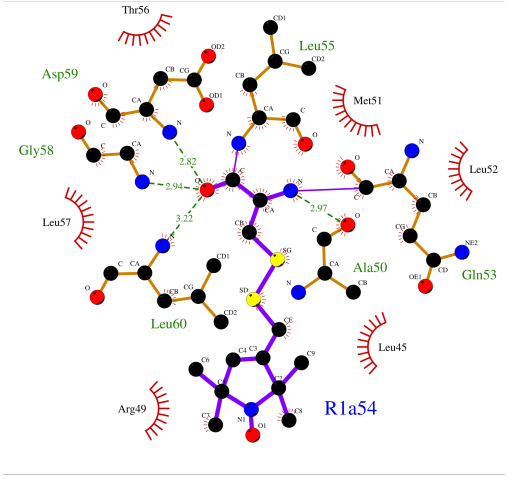
Figure A.19. – PELDOR raw data of the VcSiaP Q54R1/L173R1 H207Q mutant.

A.10. Dihedral angles of the R1 label in VcSiaP Q54R1/L173R1 R125

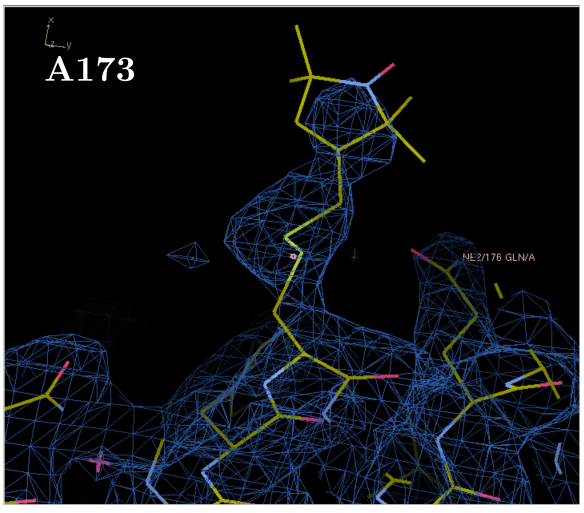
The electron density (blue mesh), the Ligplot and the polar plots of the angles are shown.

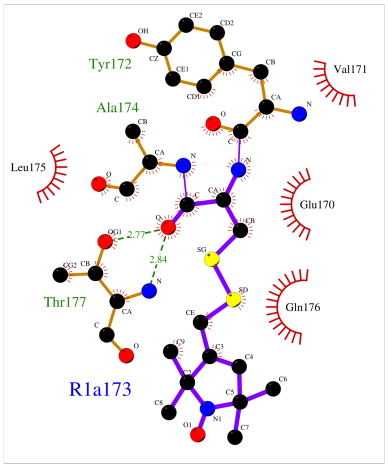
Chain A





$\chi_{1\text{-}5}$ = 288°, 168°, 92°, 65°, 11°

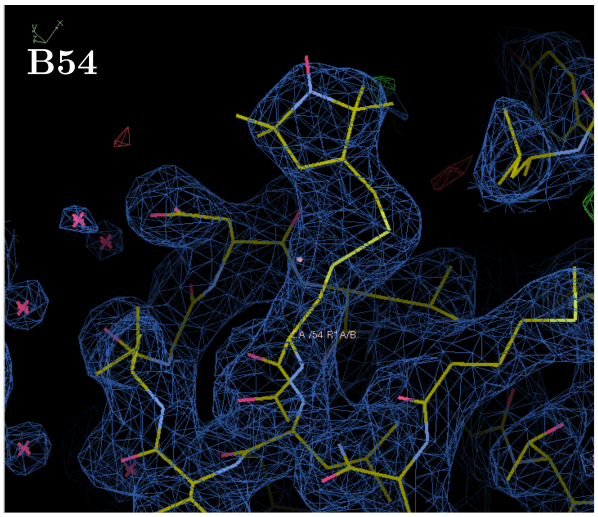


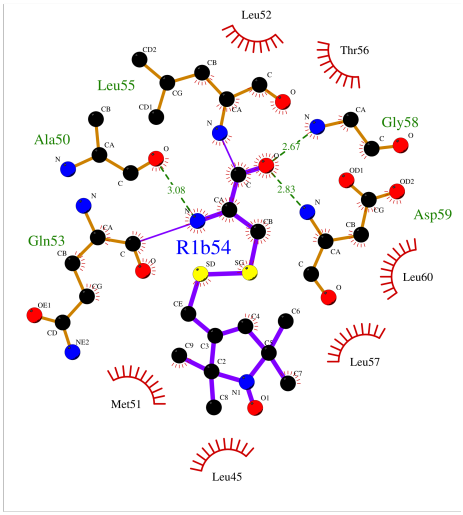


$\chi_{1\text{-}5}$ = 231°, 70°, 212°, 288°, 333°

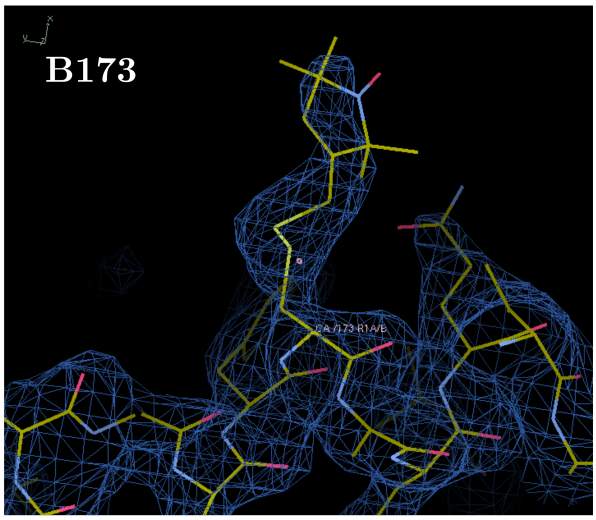
Figure A.20. – Detailed visualization of the dihedral angles of the R1 label of chain A in the crystal structure of VcSiaP Q54R1/L173R1.

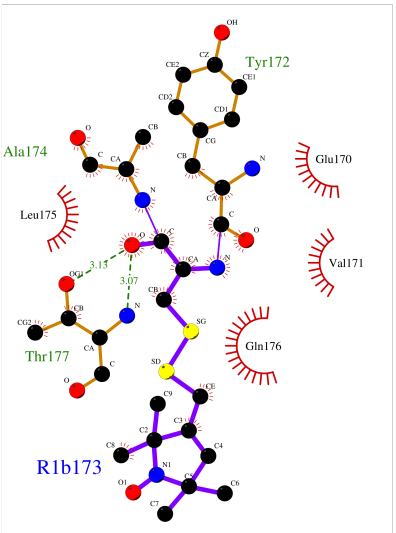
Chain B





$\chi_{1\text{-}5}$ = 309°, 161°, 68°, 63°, 291°



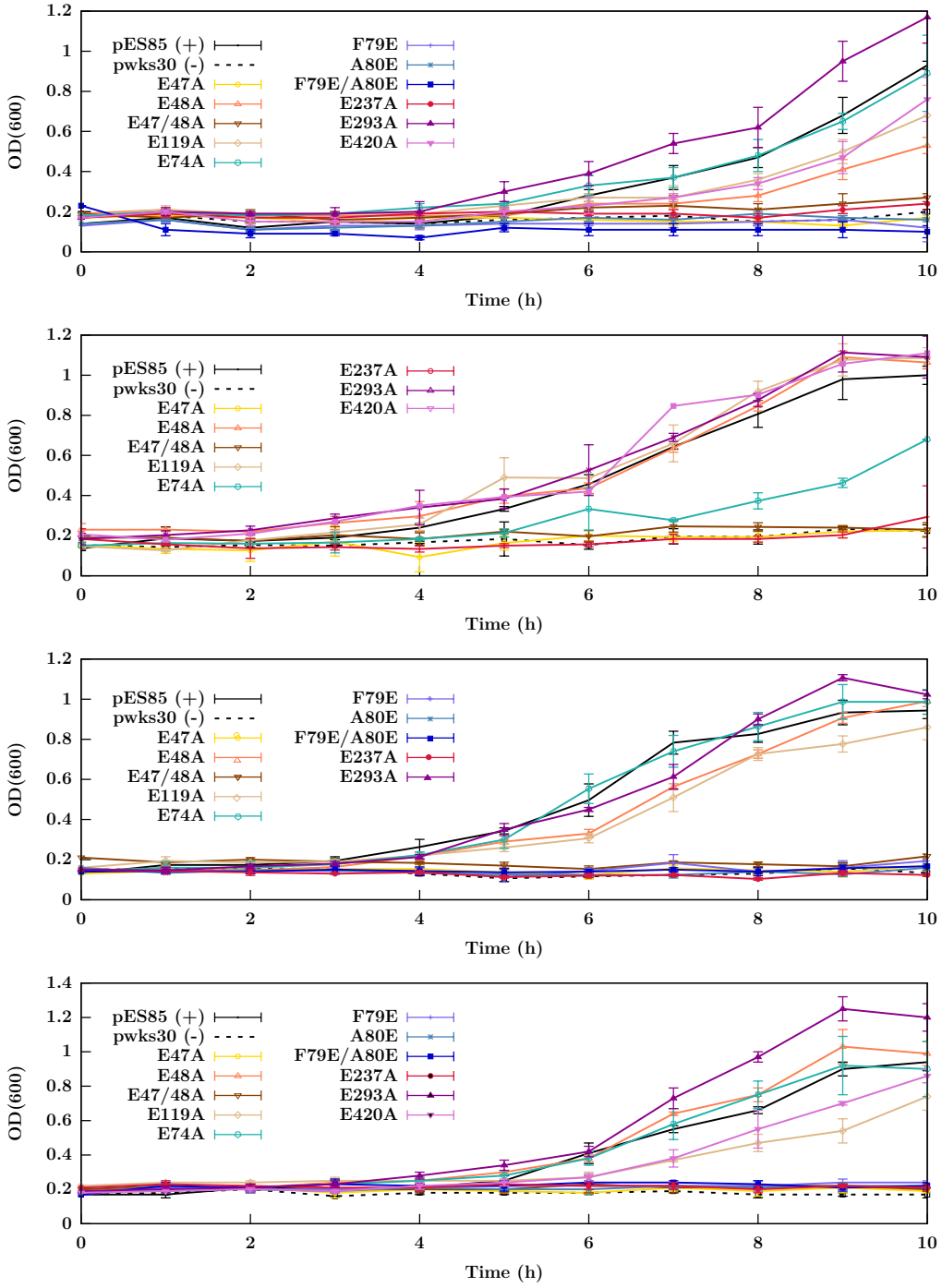


$\chi_{1\text{-}5}$ = 233°, 60°, 236°, 286°, 298°

Figure A.21. – Detailed visualization of the dihedral angles of the R1 label of chain B in the crystal structure of VcSiaP Q54R1/L173R1.

A.11. Growth

Figure A.23 lists all single growth experiments which were merged into figure 2.36 in subchapter 2.3.2.



A.23 is spread on two pages. Caption is given on the next page.

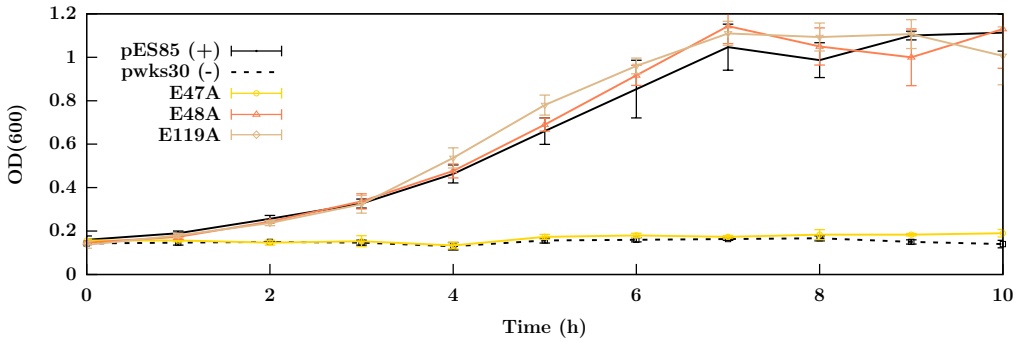


Figure A.23. – Growth experiments of selected VcSiaQM mutations in liquid M9 media. The media was supplemented with Neu5Ac as sole carbon source. The plasmid which functions as positive control (pES85) is indicated with a "+", the negative control (pwks30) is marked with a "-".

A.12. PELDOR of the peptides

Figures 1-2 show the raw PELDOR data of the double spin labeled VcSiaP with selected peptides. The timetraces are not background corrected. The respective background which was then subtracted from the recording is indicated by the red line in every plot.

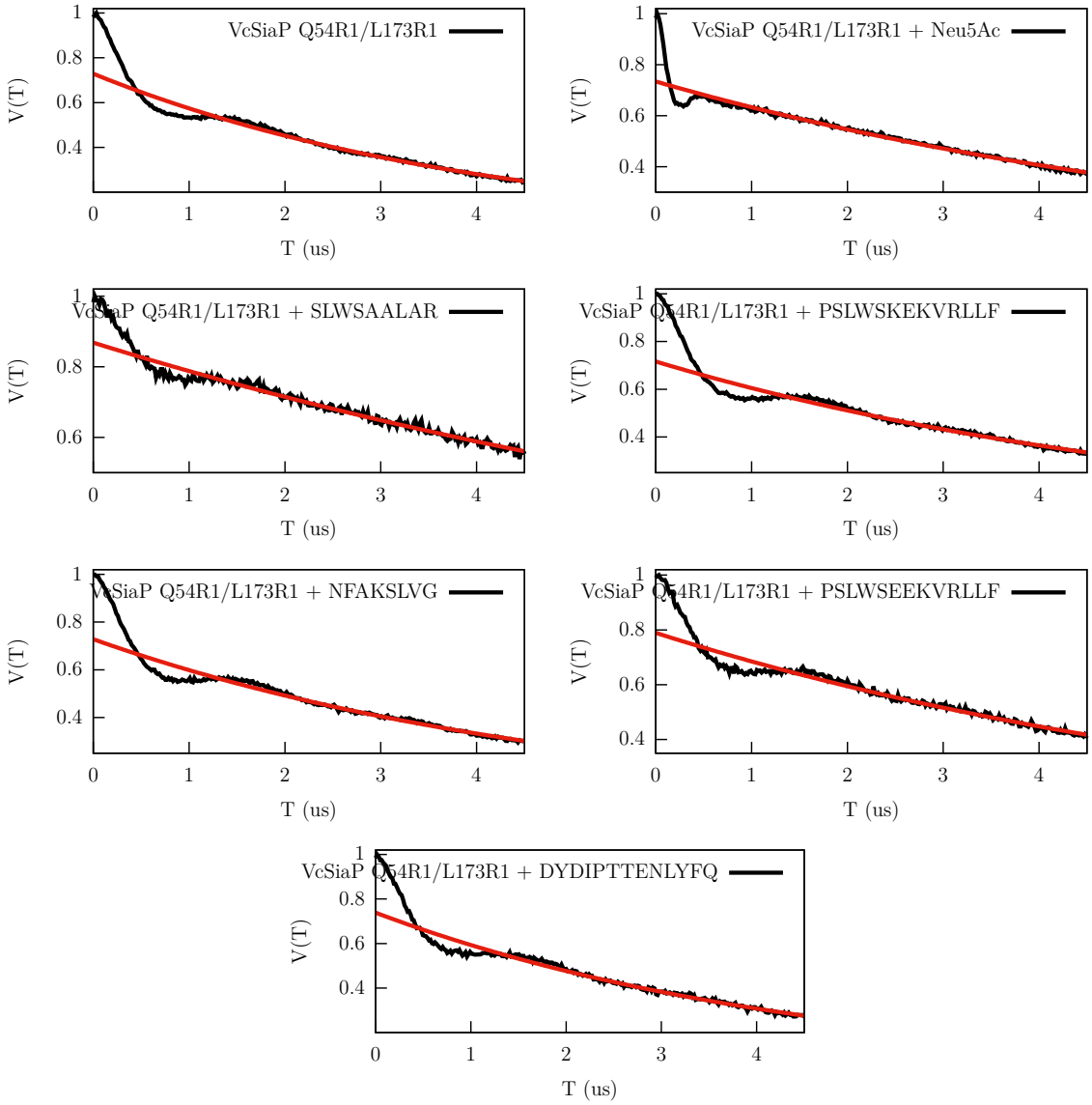


Figure A.24. – PELDOR data of the peptite experiments.

A.13. Homology model of VcSiaPQM

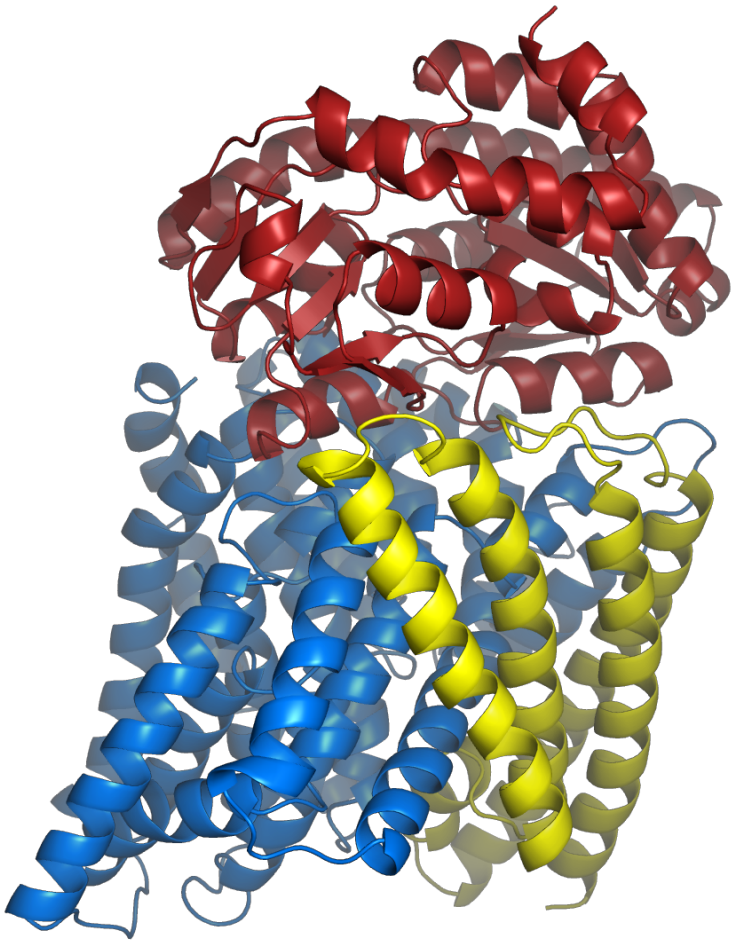


Figure A.25. – Homology model of VcSiaPQM obtained by SWISS-MODEL using [189]

A.14. Peptide Scan

Additional to the results presented in section 2.4, experiments were performed with fluorescent HiSiaP K254Dylight800 on membranes with the "inhibitor loops", on periplasmic VcSiaQM transporter loops and on periplasmic HiSiaQM transporter loops.

A.14.1. Binding of HiSiaP to inhibitor loops

The analogous result to this experiment with VcSiaP can be found in figure 2.40. HiSiaP binds to the same peptides as VcSiaP, but with different affinities.
The strongest binders for VcSiaP (C10 and C11) show decreased intensities for HiSiaP binding. But also the strongest binders for HiSiaP (C5, C6, D6, D7) have lower affinity towards VcSiaP.
Since there is no structure for HiSiaQM available, interpreting these results is difficult.

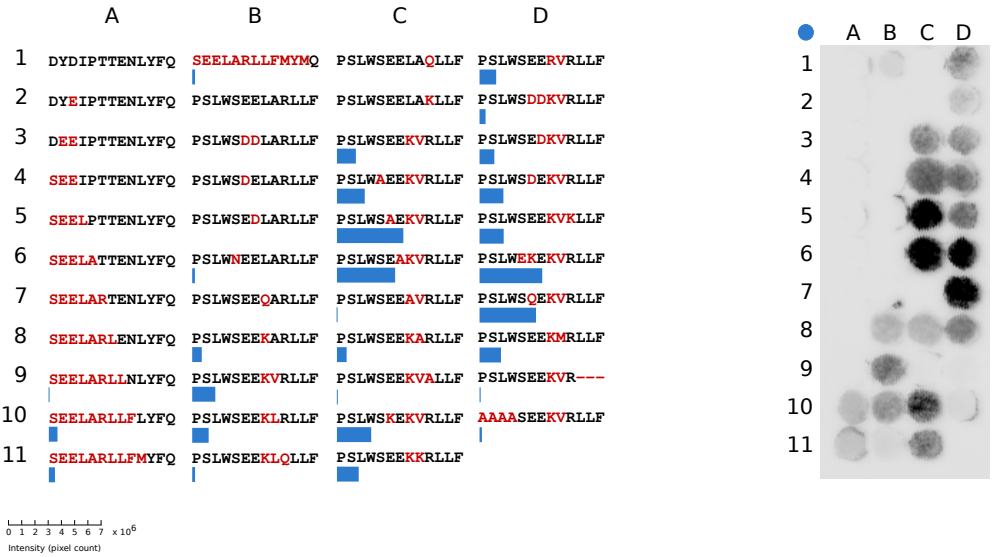


Figure A.26. – Binding of HiSiaP to peptides anchored on a cellulose membrane. The spots of the right side of the picture are the peptides whose sequences appear on the left side. The intensities of the binding are shown as blue bars.

A.15. Binding of HiSiaP to VcSiaQM transporter loops

The analogous result of this experiment with VcSiaP can be found in figure 2.43. HiSiaP bind to the same peptides that VcSiaP binds to, put with strongly decreased affinity. This is in good agreement with the fact that the P-domain discriminates between HiSiaQM and VcSiaQM. HiSiaP can not trigger the transporting mechanism of the VcSiaQM transporter and vice versa.

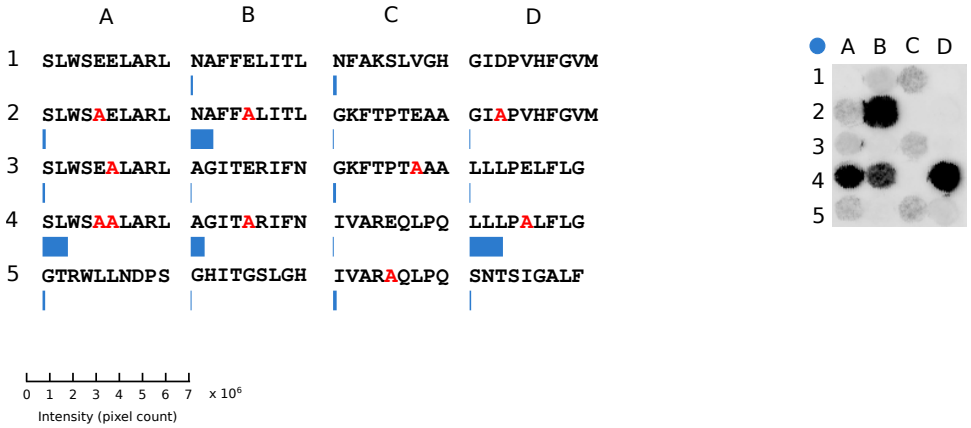


Figure A.27. – Binding of HiSiaP to periplasmic VcSiaQM transporter loops anchored on a cellulose membrane. The spots of the right side of the picture are the peptides whose sequences appear on the left side. The intensities of the binding are shown as blue bars.

A.15.1. Binding of HiSiaP to HiSiaQM transporter loops

This experiment was once performed without Neu5Ac (black bars) and then on an identical scan with 1 mM Neu5Ac (red bars). The result is depicted in figure A.28. The inserted alanin knockouts are indicated in red in the sequences.
All alanin-mutated sequences bind with higher affinity than the original loop sequences. This was also the case for the membrane where the binding of VcSiaP onto VcSiaQM loops was investigated (see figure 2.43). The strongest binders are B2 and C2.
All affinities decrease in the competition experiment with the substrate Neu5Ac.

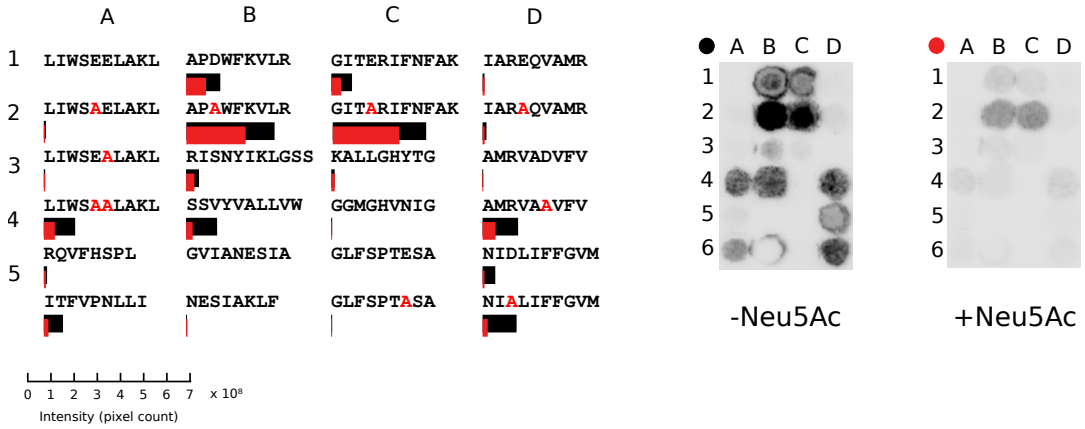


Figure A.28. – Binding of HiSiaP to periplasmic HiSiaQM transporter loops anchored on a cellulose membrane. The spots of the right side of the picture are the peptides whose sequences appear on the left side. The experiment was once performed without Neu5Ac (black bars) and once as a competition experiment with Neu5Ac (red bars).

A.16. Readout of the peptide scan

The tables A.2-A.4 summarize the fluorescence-intensities (in pixel count) of all peptide scan experiments as obtained by the software Image Studio Lite after background subtraction.

A.16.1. VcSiaP and HiSiaP binding to VcSiaQM loops

Table A.2 lists the intensities which are depicted in figure 2.43 and figure A.27.

Table A.2. – Intensities of VcSiaP and HiSiaP binding to periplasmic VcSiaQM transporter loops.

Position	VcSiaP-Neu5AC	VcSiaP+Neu5AC	HiSiaP-Neu5AC
A1	0	0	0
A2	48580	50772	121365
A3	109426	37280	90886
A4	2497875	711680	1073932
A5	41652	70799	100578
B1	10248	15395	83405
B2	780926	102387	953796
B3	31761	15938	22089
B4	1363166	343914	572322
B5	18258	13500	17359
C1	400857	186287	141908
C2	16276	4131	2768
C3	90836	49247	110034
C4	10568	0	3277
C5	178426	57279	139132
D1	0	0	0
D2	7600	0	2504
D3	0	0	10546
D4	623948	68624	1407608
D5	37898	12191	52932

A.16.2. VcSiaP and HiSiaP binding to inhibitor loops

Table A.3 lists the intensities which are depicted in figure 2.40 and figure A.26.

Table A.3. – Intensities of VcSiaP and HiSiaP binding to inhibitor loops.

Position	VcSiaP-Neu5AC	VcSiaP+Neu5AC	HiSiaP-Neu5AC
A1	0	0	0
A2	0	0	0
A3	0	0	0
A4	0	0	0
A5	0	0	0
A6	0	0	0
A7	0	0	0
A8	0	0	0
A9	8326	6535	7634
A10	18166	8754	467697
A11	39972	23864	341315
B1	45350	38854	143669
B2	9465	10013	0
B3	35188	7375	0
B4	8061	6355	0
B5	10088	8394	0
B6	11778	6235	116
B7	13319	42331	0
B8	161879	56662	520303
B9	147757	57231	1281071
B10	143528	69913	908155
B11	14059	9533	146150
C1	3064	2818	0
C2	9978	9353	0
C3	187360	108407	1045983
C4	166632	105139	1542095
C5	435002	390939	3701663
C6	321270	281302	3242143
C7	13356	10313	173
C8	154115	58311	526767

C9	7557	7734	28413
C10	2982143	6994783	1903967
C11	2466792	6806569	1218727
D1	266144	249433	917165
D2	217630	109456	337080
D3	172668	104420	827035
D4	175638	140306	1320623
D5	247205	174363	1335967
D6	757726	997884	3496527
D7	926149	704769	3146783
D8	145694	95965	1184319
D9	43442	138117	47799
D10	52432	38704	117228

A.16.3. HiSiaP binding to HiSiaQM loops

Table A.4 lists the intensities which are depicted in figure A.28.

Table A.4. – Intensities of HiSiaP binding to periplasmic HiSiaQM transporter loops.

Position	**HiSiaP-Neu5AC**	**HiSiaP+Neu5AC**
A1	0	0
A2	86753	49477
A3	51502	45178
A4	1477715	506966
A5	140813	73377
A6	898730	207221
B1	1600903	933486
B2	4176079	2797238
B3	606706	384417
B4	1458391	300688
B5	0	0
B6	29377	52246
C1	969758	455344

C2	4423418	3155906
C3	145683	149413
C4	21112	20690
C5	0	0
C6	22987	10143
D1	75996	84892
D2	171678	152941
D3	1563	8440
D4	1674389	606032
D5	590163	81450
D6	1606060	230042

A.17. Microscale Thermophoresis

Based on the peptide scan, interesting potential binding partners were identified. Those selected peptides were subjected to PELDOR measurements as explained in section 2.5. Since no binding to the protein VcSiaP was detectable with PELDOR, a new method needed to be tested. Microscale thermophoresis seemed a reasonable method to check for protein substrate interactions. The protein VcSiaP was labeled with fluoresceine and mixed with its original substrate Neu5Ac to test the method.

A.17.1. Purification of VcSiaP with Fluorescein-maleimid and NHS-fluorescein

For the MST measurements VcSiaP Q245C was labeled with fluorescein-5-maleimide and NHS-fluorescein as described in chapter 4.4.4. Figure A.29 shows the chromatograms and the SDS PAGEs from the purification.

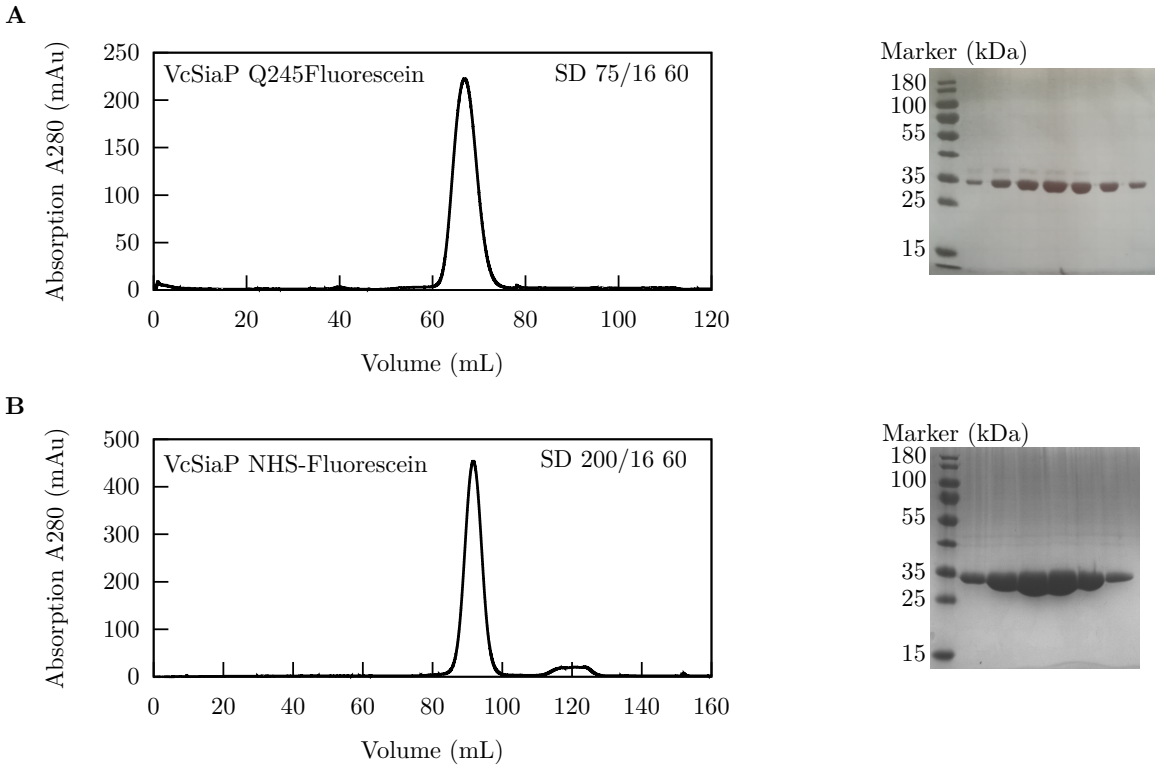


Figure A.29. – Chromatograms of the size exclusion chromatography of fluorescein labeled VcSiaP. The respective SDS PAGE is displayed on the right side of the chromatogram.

A.17.2. MST of VcSiaP Q245Fluorescein with Neu5Ac

Figure A.30 depicts the results of the MST measurements performed with 100 nm VcSiaP Q245Fluorescein and Neu5Ac (range 2.000 nM to 0.1 nM) as described in section 4.4.6.

The experiment was once performed with 20% MST-Power (subfigure A) and with 40% MST-Power (subfigure B) and the change in fluorescence was recorded and analyzed with MO.Affinity Analysis.

The change in fluorescence intensities was not enough to calculate binding affinities with the analysis program.

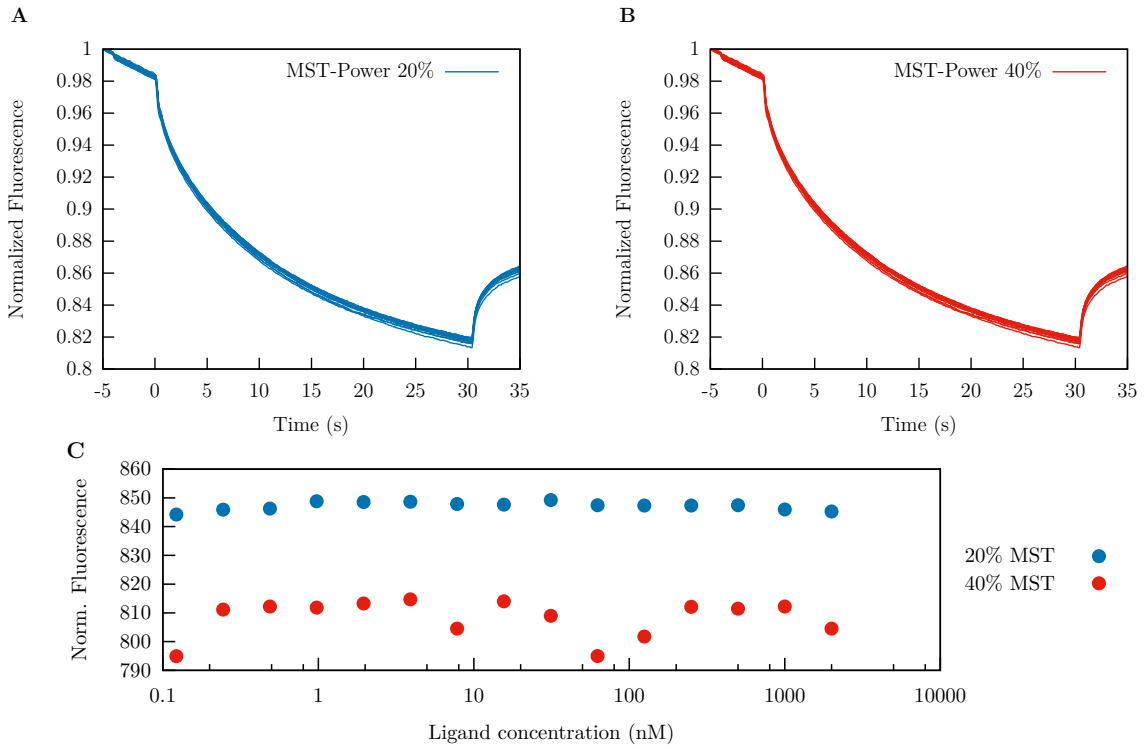


Figure A.30. – Microscale Thermophoresis of VcSiaP Q245Fluorescein with Neu5Ac. A) shows timetraces with 20% MST-power (blue) and B) shows timetraces with 40% MST-power. A dissociation constant could not be calculated from this data set.

A.17.3. MST of VcSiaP NHS-Fluorescein with Neu5Ac

Figure A.31 depicts the results of the MST measurements performed with 100 nm VcSiaP Q245Fluorescein and Neu5Ac (range 10.000 nM to 0.3 nM) as described in section 4.4.6.

The experiment was once performed with 20% MST-Power (subfigure A) and with 40% MST-Power (subfigure B) and the change in fluorescence was recorded and analyzed with MO.Affinity Analysis. Subfigure C depicts the curves which were automatically fitted by MO.Affinity Analysis and the respective constants K_D.

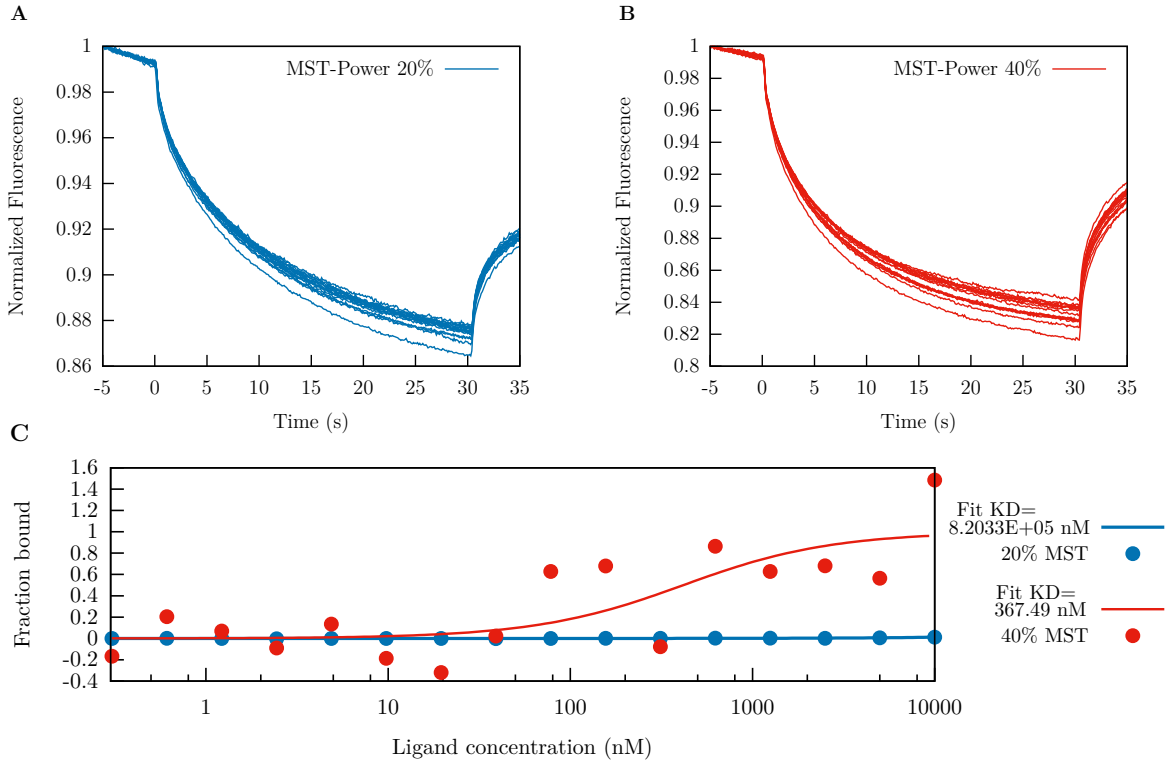


Figure A.31. – Microscale Thermophoresis of VcSiaP, labeled with NHS-Fluorescein. Neu5Ac as used as the substrate. A) shows timetraces with 20% MST-power (blue) and B) shows timetraces with 40% MST-power. In subfigure C) the ligand concentration was plotted against the bound fraction of protein and the respective K_Ds were determined by MO.Affinity Analysis.

Table A.5 shows the statistics of the fit. The set of measurements with 20% MST-power (blue) shows no plausible binding at all and the set with 40% MST-power (red) is too noisy. Therefore, this experiment was not usable to detect binding with the peptides.

Table A.5. – Analysis results of the MST of VcSiaP NHS-Fluorescein with Neu5Ac.

MST Power	Protein conc.	K_D	K_D confidence	Std. Error of Reg.
20%	100 nm	8.2033×10^5	$\pm 3.0095 \times 10^7$	1.3093
40%	100 nm	367.49	±446.1	3.9369

List of Figures

List of Tables