Structural and Functional Characterization of TRAP Transporters from Bacterial Pathogens

Dissertation

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

vorgelegt von

Martin Friedrich Peter

aus
Bad Neuenahr-Ahrweiler

Bonn, 2021

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

1. Gutachter: PD Dr. Gregor Hagelüken

2. Gutachter: Prof. Dr. Ulrich Kubitscheck

Tag der Promotion: 20.8.2021

Erscheinungsjahr: 2021

Für Mama und Papa

In ewiger Dankbarkeit

Table of Contents

Abstract

The class of tripartite ATP-independent periplasmic (TRAP) transporter proteins are widespread in the kingdoms of prokaryotes and archaea, where they are responsible for translocation of substrates across the cell membrane. The three components are a periplasmic substrate-binding protein (SBP), also referred as P-domain, and two integral membrane proteins, a smaller Q-domain and a larger M-domain. The two most extensively studied TRAP transporters HiSiaPQM and VcSiaPQM from the pathogens *Haemophilus influenzae* and *Vibrio cholerae* are known as the sole pathway for the uptake of the sugar sialic acid into these bacteria. For both organisms, sialic acid is of key importance for their virulence and helps the bacteria to circumvent the immune response of the host. Since neither of the two bacteria can produce sialic acid by itself, the uptake mechanism is essential. Therefore, TRAP transporters are potential targets for antibiotics, especially because they are absent from eukaryotic cells. However, the entire structure and function of TRAP transporter proteins are still unknown, which hampers the efficient development of compounds that inhibit the uptake of sialic acid.

In this work, it was attempted to investigate structural and functional aspects of TRAP transporters with an integrative structural biology approach. In the beginning, two integrative methods, namely PELDOR (pulsed electron electron double resonance) and smFRET (single molecule Förster resonance energy transfer) spectroscopy, were compared using three protein model systems. The cross-validation revealed an overall good agreement between these two methods and demonstrated the distinct experimental strengths and advantages of each method. In a next step, both methods were employed to address functional aspects of the uptake mechanism from TRAP transporters. In addition to other findings, smFRET confirmed the substrate-induced conformational change of TRAP SBPs upon substrate binding. Furthermore, the substrate bound crystal structure of VcSiaP was successfully solved for the first time, which subsequently allowed a comprehensive analysis of the protein-substrate interaction. A complex water network between protein and substrate was described and binding studies as well as additional crystal structures containing selected residue mutations revealed a high impact of the water network onto the substrate binding process.

For the membrane proteins of TRAP transporters, an efficient expression and purification protocol was developed by using various membrane mimetic systems. The proteins were reconstituted into different kinds of nanodiscs which for the first time allowed their handling in aqueous solution in a native-like lipid environment. Next, the membrane proteins were prepared and selectively labelled for functional studies by PELDOR spectroscopy. However, due to the yet unknown structures of the transmembrane Q- and M-domains, the PELDOR spectroscopy yielded only limited insight into the structure-function correlation of the studied transporter protein. Thus, the subsequent experiments were focused on the structural elucidation of the TRAP transporter membrane proteins using a two-pronged methodical approach, consisting of X-ray crystallography and cryo-electron microscopy. To cope with the difficulties, which still currently exist for the structure determination of membrane proteins, TRAP transporter specific and high affinity nanobodies were successfully generated. Even though the protein structure could still not be solved on an atomic scale, this work for the first time visualized TRAP transporter membrane domains at low resolutions with cryo-EM. Major necessary steps have been accomplished through this work and pave the way for a full structural description of TRAP transporters in subsequent studies.

In addition, *in vivo* studies were performed to investigate the effect of nanobody binding on the function of the TRAP transporter HiSiaPQM. The presence of several different types of nanobodies showed an inhibiting effect on uptake of sialic acid into the cells. This sets the path for a new class of TRAP transporter inhibitors which represent a promising starting point for the development of therapeutic strategies against the pathogenic bacteria *Haemophilus influenzae.*

Acknowledgments

First of all, I would like to thank PD Dr. Gregor Hagelüken for his great supervision during my PhD. He inspired my fascination for the subject of this work and introduced me helpfully into the used techniques. Thank you for all our exciting discussions, the development of new ideas and the high level of readiness to help, in scientific challenges and beyond.

Additionally, my thanks go to Prof. Olav Schiemann and Prof. Matthias Geyer. They both allowed me to not only use their laboratory but also welcome me cordially as part of their group. Thank you for all of the suggestions and honest opinions during scientific discussions.

I would like to thank Prof. Ulrich Kubitscheck, Prof. Arne Lützen and Prof. Christoph Wilhelm for their participation in my PhD committee. In addition, I would like to acknowledge our cooperative labs, particularly Christian Gebhardt and Prof. Thorben Cordes for cooperation on the work with smFRET spectroscopy, Veronika Heinz and Prof. Christine Ziegler for our EM studies and Prof. Gavin H. Thomas for sharing his expertise on TRAP transporters with us.

Further, I would like to thank the *Konrad-Adenauer Stiftung* for admission of a PhD fellowship for the last years. Besides a financial support which facilitated my work, I am thankful for the possibilities to participate in many workshops and seminars which represented a great alternation from the lab-work and showed me new, interesting aspects on many different topics.

During my PhD-thesis, I had the chance to work in two laboratories in which I felt both welcome and get to know many inspiring people and found new friends. My gratitude goes to all members from the groups of Prof. Olav Schiemann and Prof. Matthias Geyer, for their supporting help in the lab-work and also for the variety of the daily routine. Above all, I would like to thank my both master students Peer Deeping and Niels Schneberger, which I had the pleasure to supervise, for their supportive and motivated work.

As mentioned before, I had and made many friends during the last years who gratefully helped and supported me in many different ways and everyone in their own style. Above all, my thanks go to my whole family, the families of my uncles Thomas Dahm and Markus Dahm, but especially to my brother Christoph, my twin Michael and my sister Caroline. We had to challenge sorrowful times which showed me the importance of a caring family and I am infinitely grateful for your support that help all these years and through these difficult times. I am glad, thankful and proud to have you as my brothers and sister. My last gratitude goes to Maria, for being a lovely partner, to have an open ear for all my problems and nonsense and the greatest backing I can ever imagine.

List of abbreviations

ABC	ATP-binding cassette
AC	Affinity chromatography
AIDS	Acquired immune deficiency syndrome
APS	Ammonium peroxydisulfate
ATP	Adenosine triphosphate
AU	Absorption unit
AV	Accessible volume
CDR	Complementarity-determining regions
CMP	Cytidine monophosphate
COVID	Coronavirus disease
CV	Column volume
DDM	n-dodecyl-β-D-maltoside
DEER	Double electron electron resonance
DIBMA	Diisobutyl maleic acid
DM	n-decyl-β-D-maltoside
DMPC	1,2-dimyristoyl-sn-glycero-3-phosphocholine
DNA	Deoxyribonucleic acid
DOPC	1,2-dioleoyl-sn-glycero-3-phosphocholine
DP	Differential heat power
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethlyenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
EM	Electron microscopy
EPR	Electron paramagnetic resonance
ES	Energy-stoichiometry
ESEEM	Electron spin echo envelope modulation
FPLC	Fast protein liquid chromatography
FPS	FRET-restrained positioning and screening
FRET	Förster resonance energy transfer
GST	Glutathione *S*-transferase
HDX	Hydrogen-deuterium exchange
Hib	*Haemophilus influenzae* type b
H. influenzae	*Haemophilus influenzae*
HIV	Human immunodeficiency virus

HMM	Hidden Markov Model
HPLC	High pressure liquid chromatography
IM	Inner membrane
IPTC	Institute for Physical and Theoretical Chemistry
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
ISB	Institute of Structural Biology
ITC	Isothermal titration calorimetry
K_D	Dissociation constant
LB	Lysogeny broth
LCMS	Liquid chromatography mass spectroscopy
LCP	Lipidic cubic phase
LPS	Lipopolysaccharide
M	Marker
MD	Molecular Dynamics
MFS	Major facilitator superfamily
MSP	Membrane scaffold protein
MTSSL	Methanethiosulfonate spin label
MW	Molecular weight
MWCO	Molecular weight cut-off
NAN	*N*-acetylneuraminic acid
NBD	Nucleotide binding domain
Nb	Nanobody
NCS	Nucleobase cation symport
ND	Nanodiscs
Neu5Ac	*N*-acetylneuraminic acid
NHS	*N*-Hydroxysuccinimide
NMR	Nuclear magnetic resonance
NSS	Neurotransmitter solute symporter
NTA	Nitrilotriacetic acid
dNTP	Deoxynucleotide triphosphate
OD	Optical density
OG	n-octyl-β-D-glucoside
OM	Outer membrane
OMP	Outer membrane protein
ORF	Open reading frame
PAGE	Polyacrylamid gelelectrophoresis
PBST	Phosphate-buffered saline (Tween)
PCR	Polymerase chain reaction
PDB	Protein data bank
PDB-ID	Protein data bank-identifier
PEG	Polyethylene glycol
PELDOR	Pulsed electron electron double resonance
PVDF	Polyvinylidene fluoride

rcf	Relative centrifugal force
RMSD	Root-mean-square deviation
RNA	Ribonucleic acid
RT	Room temperature
RU	Resonance unit
SAXS	Small angle X-ray scattering
SBP	Substrate-binding protein
SD	Superdex
SDS	Sodium dodecyl sulfate
SDSL	Site-directed spin labelling
SEC	Size-exclusion chromatography
SI	Supporting Information
SMA	Styrene maleic acid
SNR	Signal-to-noise
SPR	Surface plasmon resonance
SSS	Sodium solute symporter
TAE	Tris-acetate-EDTA
TB	Terrific Broth
TBST	Tris-buffered saline with Tween
TCDB	Transporter classification database
TCEP	Tris-(2-carboxyethyl)-phosphine
TEMED	Tetramethylethylenediamine
TEV	Tobacco etch virus
TMB	Tetramethylbenzidine
TMD	Transmembrane domain
TMH	Transmembrane helix
TMR	Tetramethylrhodamin-5-maleimide
TRAP	Tripartite ATP-independent periplasmic
TTT	Tripartite tricarboxylate transporter
UAA	Unnatural amino acid
UC	Ultracentrifugation
UV	Ultraviolet
V. cholerae	*Vibrio cholerae*
VPI	*Vibrio* pathogenicity island
WHO	World health organization
wt	Wildtype

Chapter 1

Introduction

1.1 Bacterial pathogens

In the course of writing this thesis in 2020, the world is challenged by the COVID-19 pandemic and is facing with unprecedented challenges whose end and consequences are not foreseeable. The global spread of the virus within several weeks showed how fast such an infectious disease can grow to a worldwide threat and change everyday life of almost everyone. This includes not only infected people, deceased patients and their relatives, but in the form of restrictions and containments also a change in public and social life and an accompanying economic slowdown.

Even if the current pandemic is caused by a virus, several further kinds of pathogens can cause human infection diseases, such as fungi, parasites or bacteria. Considering all pathogens together, the World Health Organization (WHO) assigns infection diseases to one of the major causes for death, all over the world. In 2004, the world health report attributed more than 25% of all deaths to infection diseases, which represented more than 14 million deaths worldwide (World Health Organization, 2004). Even if death owing to infections have declined during the last 15 years, 5 of the top 10 death reasons in low-incoming countries are still infection diseases, such as diarrhoeal, malaria or tuberculosis (World Health Organization, 2020d). In high-income countries, the causes of death are significantly different but, as shown by the COVID-19 pandemic, also these countries are far from immune against the impact of infection diseases (World Health Organization, 2020d).

The investigations in this work are related to bacterial pathogens. Nowadays, bacterial infections are mostly treated with an antibiotic drug. Since their first description in the late 18th and early 19th century, including the famous work of Alexander Fleming on penicillin, a large range and diversity of antibiotics have become available (Nicolaou and Rigol, 2018). The development of these drugs can be seen as one of the most important progresses in medicine because they lower the mortality of a bacterial infection immensely. This is especially apparent in history, where some of the most fatal pandemics are ascribed to pathogenic bacteria. For example, the plague outbreak in Europe in the middle ages, also known as the black death, was caused by the bacterium *Yersinia pestis* and killed more than 200 million people (Haensch et al., 2010). Today, this bacterium, especially its mechanism of pathogenicity and the reaction of the human immune system is well known and various antibiotics are developed that help to prevent a spread and can be used for the treatment of infected patients (DeLeo and Hinnebusch, 2005; Perry and Fetherston, 1997).

Even if important goals in the fight against pathogenic bacteria have been accomplished today, current developments and observations in combination with the experiences from the pre-antibiotic era hint at a possible bad scenario for the future. The introduction and widespread use of antibiotics led to an adaption of several bacteria in developing a resistance against the antibiotics and can no longer be treated with such a substance (Ventola, 2015). Corresponding to the WHO, the increasing antibiotic resistance is one of the biggest threats, not only to the health systems but also for general developments, as shown by the economics consequences of the current COVID-19 pandemic (World Health Organization, 2020a). A few years ago in 2015, the organization started the global antimicrobial resistance surveillance system (GLASS), to monitor and describe the current developments of antibiotic resistances (World Health Organization, 2020b). On the basis of this report, an already high level of antibiotic resistance can be found all over the world.

An important and famous example of antibiotic resistant bacteria are methicillin-resistant *Staphylococcus aureus,* also known as MRSA. The first case of these resistant bacteria could be observed only two years after the introduction of methicillin as antibiotic drug in 1960 (Ventola, 2015). Since then, the infections with MRSA have greatly increased, as for example indicated by the doubled number of infections in the US between 1999 and 2005 (Klein et al., 2007). Overall, the spread and growth of MRSA infections lead to a high stress on health systems with higher treatment costs, more serious progression of the disease and higher risk of deadly infections (Klein et al., 2013, Lambert et al., 2011).

Another infamous pathogenic bacterium is *Mycobacterium tuberculosis,* responsible for the respiratory disease tuberculosis (pulmonary tuberculosis). Due to the annually, latest global tuberculosis report in 2020, more than 10 million people got sick of tuberculosis and 1.4 million of these illnesses ended fatally in 2019 (World Health Organization, 2020c). As a result, the WHO report classifies this infectious disease as the most dangerous one (prior to the COVID-19 pandemic arises). A medical treatment as well as full recovery from a tuberculosis disease is possible, but especially low-income countries cannot medicate many patients and are therefore affected with a higher number of serious cases and higher death rates (World Health Organization, 2020c). Another, very worrying fact for *M. tuberculosis* is an increase in its antibiotic resistance: In 2007, only for 7% of all tuberculosis cases a rifampicin resistance was obtained, while in 2019, already 61% of the patients showed such a resistance (World Health Organization, 2020c).

The development of growing antibiotic resistance, the potential danger that results from pathogenic bacteria without a usable drug and the nowadays fast outbreak of a pandemic in the globalized world, underscore the importance of continuous research on pathogenic bacteria and useful drugs. To avoid drastic developments towards a dangerous post-antibiotic era, the WHO analysed the resistances of several pathogenic bacteria in 2017 and published a list of twelve pathogens, for which new antibiotics are urgently needed (World Health Organization, 2017a, World Health Organization, 2017c).

1.1.1 Virulence of *Haemophilus influenzae*

One representative of the above mentioned WHO list and one of the two pathogenic bacteria that are related to this work is *Haemophilus influenzae,* a Gram-negative pathogen. This bacterium was first described at the end of 18th century and was mistakenly classified as the causative agent for the viral influenza (Turk, 1984). From this pathogen, bacterial strains in the encapsulated and nonencapsulated forms are known and the latter ones are further divided up into six serotypes a-f, depending on the composition of the capsule-forming polysaccharides (Pittman, 1931, World Health Organization, 2007). The *H. influenzae* type b

(Hib) is the best known and also the most virulent serotype and responsible for around 95% of all *H. influenzae* diseases (World Health Organization, 2007). An infection with *H. influenzae* can cause many serious diseases, for example pneumonia, pericarditis, septic arthritis, meningitis or epiglottitis (World Health Organization, 2007). Especially the two latter diseases are known to be caused by a *H. influenzae* type b infection and if left untreated, usually end fatally (Barbour, 1996).

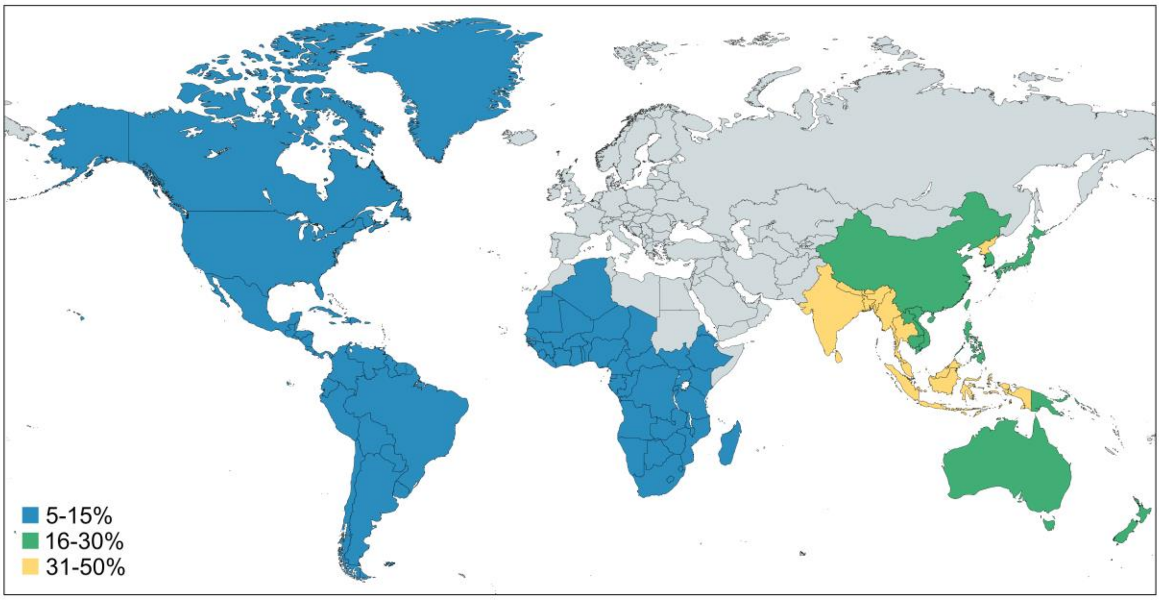


Figure 1-1: Worldwide spread of ampicillin resistance of *H. influenzae*. The ampicillin resistance of the pathogen bacterium *H. influenzae* in different regions of the world are shown in percentage. The figure was adapted from (World Health Organization, 2017a) (map created with: www.mapchart.net).

Even though people of every age can be infected, children are particularly at risk and over 90% of all infections are observed in children under 5 years (World Health Organization, 2013). From year 2000 to 2008, the number of WHO member countries using antibiotics against Hib infections have more than doubled from 62 to 136 but still in 2008, 203000 deaths of children under 5 years are reported in relation with a Hib infection (371000 in the year 2000) (World Health Organization, 2013). Even in the case of a successful treatment, affected children are faced with serious and permanent impairments such as blindness or learning disabilities (Edmond et al., 2010, World Health Organization, 2007). Unfortunately, *H. influenzae* is one of the pathogens with an increasing occurrence of antibiotic resistance. The weekly epidemiological record of the WHO in 2013 about Hib describes a detected resistance of this pathogen to the standardly used antibiotic ampicillin of around 20%-30%, spread in many regions of the world (Figure 1-1) (World Health Organization, 2013, World Health Organization 2017a). The reported high fatality and infection rates despite vaccines and drugs, as well as the serious diseases that are caused by an infection in combination with the increasing antibiotic resistance are responsible for the WHO's classification of *H. influenzae* type b as pathogen, for which new antibiotics are urgently needed.

Considering the pathogenicity of Hib and its impact on and response of the human immune system after infection, an analysis of the cellular surface of the pathogen is helpful. Logically, the outer surface of the pathogen is the first point for an interaction between the bacterium and the host cells immune system. In addition to the mentioned capsule, the outer membrane of the encapsulated *H. influenzae* strains includes lipopolysaccharides (LPS) (Hood et al., 2004). In a general view, the LPSs of *H. influenzae* consist of an oligosaccharide chain that is linked to a membrane-anchored lipid A moiety in the outer membrane of the bacterium (Hood et al., 2004). A more detailed introduction about the composition and synthesis of the LPSs, as well as the impact of the LPS composition to the virulence of *H. influenzae*, is given in Chapter 1.2.

In general, bacterial LPSs are strong stimulators of the innate immune system (Alexander and Rietschel, 2001, Rosadini and Kagan, 2017). However, in *H. influenzae* the LPSs in combination with the polysaccharide capsule are assigned to be mainly responsible for the resistance of the pathogen against the humane immune system and nonencapsulated mutants showed a reduced virulence (Kimura and Hansen, 1986, Moxon et al., 1984, Zwahlen et al., 1983). Today, it is known that these components inhibit processes in the complement system, especially the "alternative pathway" of C3b, and by this, avoid initial destruction of the pathogen through phagocytosis or bacteriolysis and the formation of an inflammatory response (Hetherington et al., 1993, Severi et al., 2007, Sukupolvi-Petty et al., 2006, World Health Organization, 2007). Thus, several classes of anticapsular antibodies, IgG, IgM and IgA are mainly involved in the host immune reaction, while subclasses of the IgG antibody are assigned to the key role to defend from the bacterial infection (Jennings, 1983, Ramadas et al., 1986). The detection and opsonization of the *H. influenzae* capsule fragments by antibodies promote bactericidal activity and phagocytosis, that result in a destruction of the bacteria (Granoff, 2001, World Health Organization, 2007).

The described bacterial pathogenicity and the host immune response give an explanation for the high numbers of *H. influenzae* infections with serious diseases and fatal outcome in children. The polysaccharide capsule directly stimulates the antibody producing B cells of the adaptive immune system without helper T cells (T independent antigen) and this kind of antigens are poorly immunogenic in the first years of life (World Health Organization, 2007). Thus, the immune response through generation and production of specific antibodies is decreased and the distribution of bacteria are less effectively detected by the defence mechanism.

1.1.2 Virulence of *Vibrio cholerae*

A second bacterium that relates to this work is *Vibrio cholerae*, the causative agent of the infamous cholera. Hints of the bacterial infection through description of typical symptoms go back to the 5th century BC and in the last 200 years since 1817, seven worldwide cholera pandemics have been recorded (Harris et al., 2012). The origin of the seventh pandemic is located in the year 1961 in the city Makassar (Sulawesi, Indonesia), distributed over large parts of the world and is still ongoing nowadays (Hu et al., 2016). During the last years, mostly local outbreaks that are assigned to this pandemic are observed, as for example in Zimbabwe, Dominican Republic or Haiti (Islam et al., 2011, Pan American Healt Organization, 2018). In Haiti, the endemic outbreak is correlated with the strong and destructive earthquake in 2010 and in the 7 years after the natural disaster, more than 800000 infections with *V. cholerae* were reported with more than 9 000 fatal cases (Pan American Healt Organization, 2018). For 2016, the official record of the WHO quantifies 132121 reported cases of infections of *V. cholerae* with 2 420 deaths worldwide (World Health Organization, 2017b). According to estimates, the real infection number is significant higher with around 2.86 million cases and 95 000 deaths (Ali et al., 2015). Furthermore, an estimated 1.3 billion people are affected and under risk of an *V. cholerae* infection, underlying the importance for continuous research on this pathogen and development of cheap, fast and effective drugs, under consideration of a grow of antibiotic resistance (Ali et al., 2015).

Until today, more than 200 serogroups of *V. cholerae* are known and classified, depending on the composition of the LPSs on the bacterial surface (Chatterjee and Chaudhuri, 2003, Shimanda et al., 1994). Only two variants of these serogroups, O1 and O139 (named after the O antigen of the LPS), are pathogenic and can cause the disease cholera, while the other ones are none pathogenic strains for humans (Harris et al., 2012). The virulence is on the one hand traced back to the cholera toxin, which causes the diarrheal symptoms of a disease and on the other hand a toxin-corregulated pilus (TCP), which allows the colonization of the pathogen in the human intestinal tract (Li et al., 2003, Tacket et al., 1998). The cholera toxin disrupts the adenylate cyclase system of the host cells, which results in a secretion of chloride ions, followed by water and further electrolytes into the intestinal tract, leading to strong diarrhoea (Cassel and Pfeuffer, 1978, Gill, 1976, Harris et al., 2012). The genes of the toxin and TCP are located in the genome of a bacteriophage derived CTXϕ DNA fragment and a *Vibrio* pathogenicity island (VPI), also proposed as filamentous phage (VPIϕ) (Karaolis et al., 1999, Li et al., 2003).

An infection with the pathogen *V. cholerae* is mostly caused by a direct contact with contaminated water or food. The contamination of water is often correlated with warm temperatures and poor sanitary facilities, whereby both factors increase the number of *V. cholerae* bacteria (Chowdhury et al., 2017). After infection with the pathogen, the innate and adaptive immune systems are involved in a host cell response. For the innate immune system, an upregulation of genes that are correlated to the innate defence mechanism are observed in patients (Flach et al., 2007). Following this, a complex response of the adaptive immune system is formed and antibodies from the IgG, IgM and IgA classes are observed in correlation to an infection (Flach et al., 2007).

One dangerous consequence of a serious *V. cholerae* infection is a massive water loss with up to 1 L per hour, which, if untreated, results in a strong dehydration and 70% fatality after only a few hours (Harris et al., 2012). For a treatment of a cholera disease, the patients are in general initially supported with an oral or intravenous rehydration therapy of liquid, nutrients and electrolytes to replace the large loss of water (Harris et al., 2012, Morris, 2003). Supportively, if possible, a patient can be treated with antibiotics to decrease the time and impact of the diarrhoea (Morris, 2003). Unfortunately, also for this pathogenic bacterium, the large scale application of antibiotics during the last decades led to an increased resistance, which was firstly observed 40 years ago in 1979-1980 in Bangladesh in 21% of investigated isolates (Glass et al., 1980). Since then, several other local studies on characterization of *V. cholerae* detected a multiple-drug resistance of this pathogen, indicating a problem that is spread in many affected regions (Dalsgaard et al., 2001, Dromigny et al., 2002). Importantly, not only developing countries are affected, also in the United Stated in the 1990s, an increase of antibiotic resistance from 8% in the first five years up to 31% in the last five years of the decade were observed (Steinberg et al., 2001).

Cholera is a good example of a bacterial disease that on the one hand can be treated effectively with comparably easy interventions but on the other hand is still present all over the world with partially large local outbreaks and a high number of deaths. Especially the further development of the antibiotic resistance of *V. cholerae* poses a great danger for serious diseases and can strongly increase the number of fatal infections. Therefore, in analogy to the previously presented *H. influenzae*, new strategies for antibiotic compounds to these pathogens are of high interest.

1.2 Sialic acid plays an important role in bacterial pathogenicity

One important component for the two pathogens *H. influenzae* and *V. cholerae* is the sugar sialic acid. The term sialic acid describes a group of α-keto acids with a nine-carbon sugar backbone and from this basic structure, more than 50 modifications are known with a widespread presence in eukaryotes and prokaryotes (Angata and Varki, 2002). In many publications and also in this work, the term sialic acid is not used for the whole group of molecules, but for the main representant *N*-acetylneuraminic acid (Neu5Ac, NAN, 2-keto-3-deoxy-5-acetamido-D-*glycero*-D-*galacto*-nonulosonic acid). The chemical structure of Neu5Ac is shown in Figure 1-2.

Sialic acid and its derivates are involved in many important metabolic processes within in prokaryotes and in eukaryotes. In eukaryotic cells, including humans, the sialic acid is often positioned on the surface of the cells and is thereby responsible for a cell-cell interaction, interaction with small molecules such as ligands or antibodies or an overall stabilization of the membrane (Sillanaukee et al., 1999, Vimr et al., 2004). In prokaryotes, sialic acid is often observed in pathogens and regarded as an important factor for bacterial virulence (Vimr et al., 2004). Figure 1-2 illustrates a general overview of the sialic acid pathways from pathogenic bacteria *H. influenzae* and *V. cholerae*. Here, the use of sialic acid can be divided into two pathways: The coverage and modification of the bacterial surface and the use of the sugar as source for generation of energy (Figure 1-2) (Severi et al., 2007).

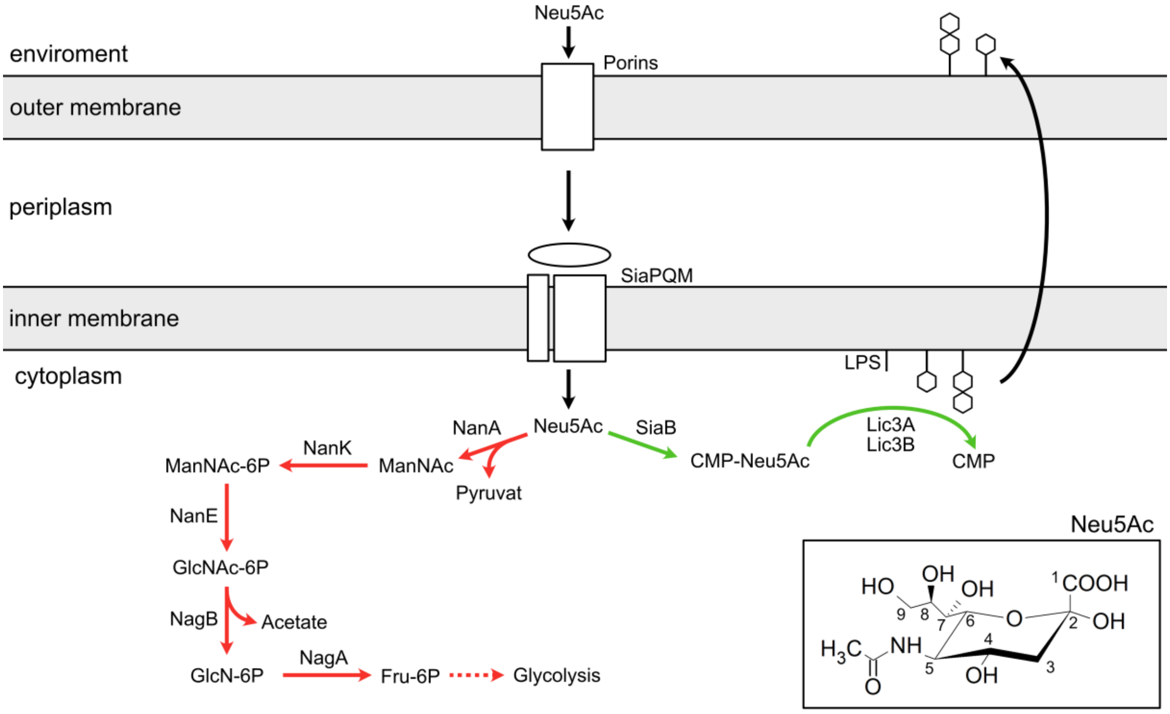


Figure 1-2: Sialic acid metabolism in pathogenic bacteria. Schematic of the known metabolism pathways of sialic acid in the two pathogens *H. influenzae* and *V. cholerae*, adapted from (Severi et al., 2007). The nutrient pathway is indicated in red, the LPS sialylation pathway in green.

1.2.1 Sialic acid metabolism in *H. influenzae*

In the pathogen *H. influenzae* type b, both of these sialic acid pathways are present, while the coordination between them is not known (Vimr et al., 2000). The first description of the sialic acid pathway for the usage as nutrient was for *Escherichia coli* K-12 bacteria (Plumbridge and Vimr, 1999, Vimr and Troy, 1985). During the analysis of the pathway in *H. influenzae*, a high similarity between both bacteria were observed (Vimr et al., 2000). The central key point for this nutrient pathway is the *nan* operon (for Hib more specific the *nan-nag* operon), which encodes at least the three proteins NanA (aldolase), NanE (epimerase) and NanK (kinase) and in Hib additionally the two proteins NagA (deaminase) and NagB (deacetylase) (Figure 1-3) (Vimr et al., 2000, Vimr et al., 2004). As illustrated in Figure 1-2, NanA initially catalyses the conversion from sialic acid to pyruvate and ManNAc and the latter product is processed by phosphorylation and epimerization through the enzymes NanK and NanE. In the final step, GlcN-6-P is converted by NagA and NagB with detachment of acetate and ammonia to receive Fru-6-P, which can be used as an educt for generation of energy in the glycolysis.

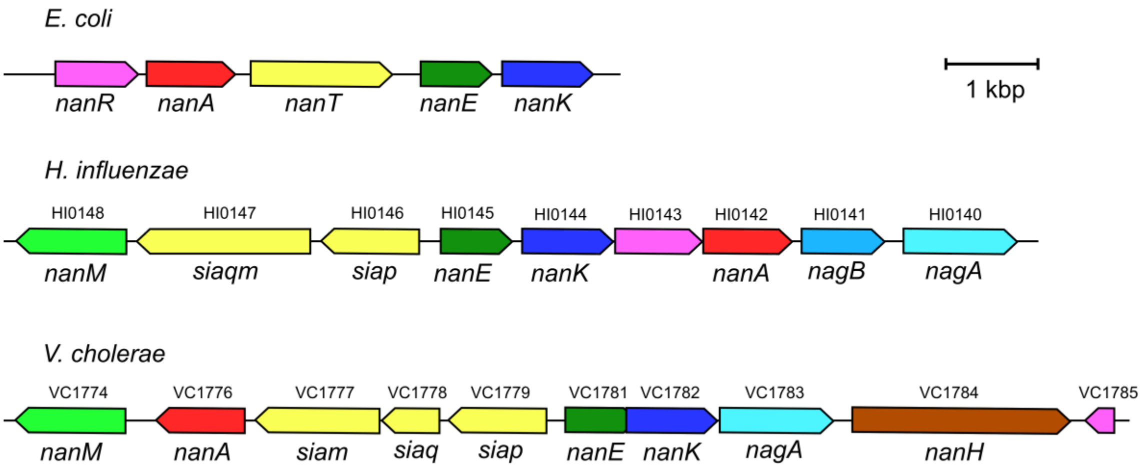


Figure 1-3: Schematic from *nan-nag* gene regions in selected bacteria. The *nan* region from *E. coli* (top) or the linked *nan-nag* gene regions from *H. influenzae* (middle) and *V. cholerae* (bottom) are shown. The *nan-nag* region of *V. cholerae* is located on the VPI-2 (Almagro-Moreno and Boyd, 2009). The indicated genes are on the same size-scale. Similar genes are illustrated by colour as follows: transcriptional regulators (magenta), aldolases (light red), transporter proteins (yellow), epimerases (dark and bright green), kinases (dark blue), deaminases (blue), deacetylases (bright blue) and sialidase (brown) (figure was modified after (Vimr et al., 2004) and (Severi et al., 2005) and gene information were received from UniProtKB (The UniProt Consortium, 2019) (access date: 08.01.2021)).

The second sialic acid pathway in *H. influenzae*, the modification and decoration of the bacterial cellular surface with sialic acid, is presumably very important for their virulence, as the LPSs were predicted to be involved in the resistance of Hib against the hosts immune system (Kimura and Hansen, 1986, Moxon et al., 1984, Zwahlen et al., 1983). The first evidence for the incorporation of sialic acid into the surface-fixed LPSs were described in the early 1990s during the analysis of the LPS composition of *H. influenzae* (Mandrell et al., 1992, Phillips et al., 1993). As described above, sialic acid is often presented on the cellular surface of eukaryotes. The isolation of sialic acid from the LPS of this pathogen provided the first clue, that Hib uses the sialic acid to imitate the hosts cell surface as a kind of invisibility cloak to prevent the detection by the hosts immune system.

Importantly, the synthesis of the LPS fragments that are ultimately fixed on the outer surface is performed in the cytoplasm (Sukupolvi-Petty et al., 2006). The LPS pathway for sialic acid starts with the attachment of CMP to Neu5Ac by the synthetase SiaB, which was determined as an essential step for the sialylation of LPSs (Figure 1-2) (Hood et al., 1999). For the attachment of sialic acid to a LPS basic structure, four sialyltransferases were described in several strains of *H. influenzae*, Lic3A, Lic3B, LsgB and SiaA (Fox et al., 2006, Hood et al., 2001, Jones et al., 2002). Today, the two closely related enzymes Lic3A and Lic3B are considered as the main pathways for sialylation of LPSs in *H. influenzae*, whereas Lic3A is responsible for a monosialylation and Lic3B can attach mono- and disialic acid to the LPSs (Fox et al., 2006).

At about the same time of the discovery and description of the LPS sialylation mechanism, the exact composition and structure of *H. influenzae* LPSs, especially regarding the presence of sialic acid, were analysed. The first LPSs were described for two different strains, the common pathogenic strain type b (RM153) (Masoud et al., 1997, Schweda et al., 2000) and strain type d (RM118) (Risberg et al., 1999). The initially observed compositions are shown in Figure 1-4 A. The inner-core or basic structure of both LPSs consists of a trisaccharide of heptose, attached to a lipid A moiety that fixes the LPS in the outer membrane of the bacterium. Each of the three heptoses serves as an anchor point for the binding of further hexose residues, that are elongated in different ways for the two mentioned *H. influenzae* strains. In the first description, for none of the LPSs from the two strains, the attachment of sialic acid as nine-carbon sugar was observed (Masoud et al., 1997, Schweda et al., 2000, Risberg et al., 1999).

In further studies on *H. influenzae* type d (RM118), the bacteria were cultured in the presence of supplemented sialic acid in the environment and the LPS analysis of these bacteria revealed a significant alteration of the composition and containment of sialic acid (Figure 1-4 B) (Cox et al., 2002, Hood et al., 2001). Following these results, the LPSs from the pathogenic strain type b were also investigated under colonization with sialic acid supplemented environment and also for this strain, an alternation of the LPSs was observed (Figure

1-4 B) (Hood et al., 2004). Under these conditions, both strains indicated not only single sialic acid sugars in their LPS, but furthermore the attachment of a tetrasaccharide, consisting of three hexose and one sialic acid residue at the terminal position. In a comparable study, the pathogenic *H. influenzae* strain type b showed a higher tendency of the attachment of the tetrasaccharide motif to the LPS than the *H. influenzae* strain type d (Hood et al., 2004).

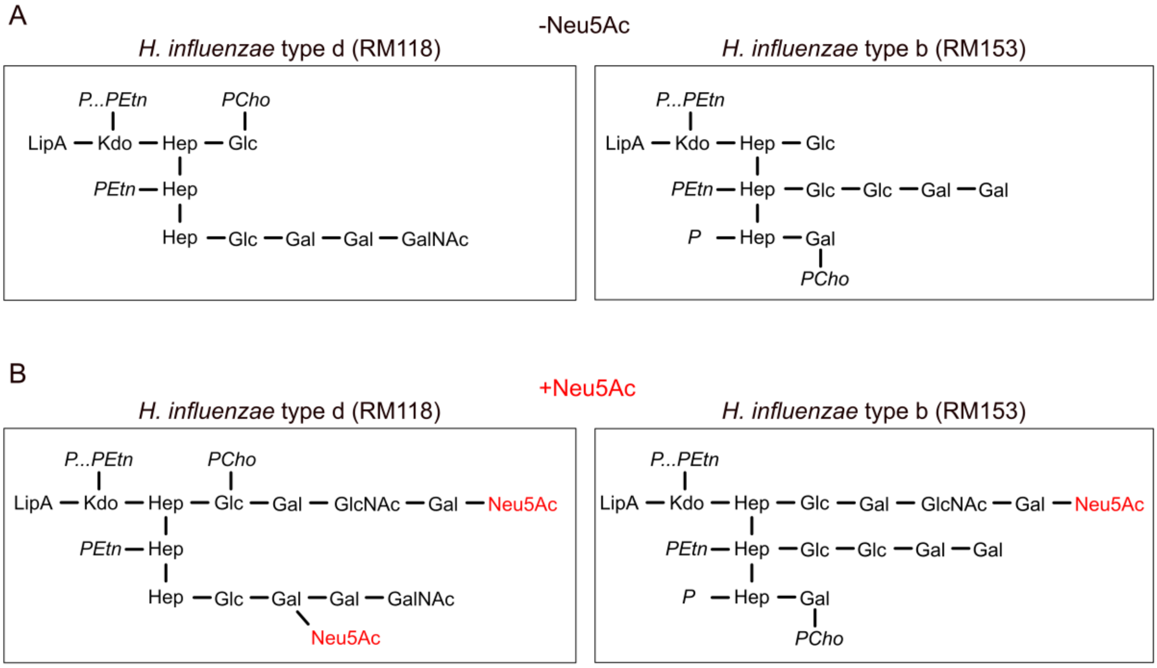


Figure 1-4: Composition of LPSs from *H. influenzae*. A) Schematic of the LPS composition from *H. influenzae* type d (left, strain RM118) and type b (right, strain RM153) after colonization without sialic acid in the environment. **B)** Same as A) but after colonization with sialic acid in the environment. The figure is modified after (Hood et al., 2004) (abbreviations: LipA, lipid A; Kdo, 2-keto-3-deoxyoctulosonic acid; Hep, L-*glycerol*-D-*manno*-heptose; Glc, D-Glucose; Gal, D-galactose; GalNAc, *N*-acetylgalactosamine; GlcNAc, *N*-acetylglucosamine; *PEtn*, phospoethanolamine; *P*, phosphate; *PCho*, phosphocholine).

As mentioned before, the LPSs from *H. influenzae* were predicted to be important for the virulence and involved in the resistance of Hib against the immune response of the host (Kimura and Hansen, 1986, Moxon et al., 1984, Zwahlen et al., 1983). The influence of the presence of sialic acid on the composition and structure of the LPS suggests a direct correlation between the sialic acid and the virulence of the pathogenic bacterium. This hypothesis was experimentally confirmed by several studies with knock-out variations of different enzymes that are involved in the LPS sialylation pathway (Figure 1-2). For example, the knock-out of the CMP-Neu5Ac SiaB synthetase gene and an accompanying disrupted attachment of sialic acid into the LPSs revealed a decreased resistance of the bacteria against human serum (Fox et al., 2006, Hood et al., 1999). Another study investigated an infection of chinchillas with *H. influenzae* to cause an otitis media and for bacteria with knock-out SiaB, as well as Lic3A sialyltransferase, a strongly weakened course of disease was obtained (Bouchet et al., 2003). These studies, as well as the described LPS sialylation, are all in agreement with the hypothesis that pathogens such as *H. influenzae* use sialic acid to imitate the cell surface of the host and by this, hide from an immune response.

1.2.1.1 Source of sialic acid

One question that is important regarding the impact of sialic acid on the virulence of *H. influenzae* is the origin of sialic acid. For bacteria in general, sialic acid can originate from a *de novo* synthesis, such as in *E. coli* with the sialic acid synthase NeuB from ManNAc and pyruvate (Vimr et al., 1989, Vimr et al., 2004). Additionally, bacteria were observed to be able to uptake sialic acid from their environment and transport it into their cytoplasm for further metabolism. This was for the first time described for *E. coli* with the sialic acid transporter NanT (Vimr and Troy, 1985). For the pathogen *H. influenzae,* the way of a *de novo* synthesis is unlikely as the bacteria lack a gene for the NeuB synthase and only sialic acid free LPSs could be isolated if no exogenous sialic acid source was available (Hood et al., 2001, Hood et al., 2004, Vimr et al., 2004). However, many years after description of the NanT sialic acid transporter from *E. coli,* a transporter for this substrate in *H. influenzae* was described by several groups, suggesting that this pathogen gains the sialic acid only from an exogenous source (Allen et al., 2005, Severi et al., 2005). The transporter was assigned to the class of TRAP transporters (tripartite ATP-independent periplasmic transporter), a bacterial transporter system that is investigated intensively in this work and is described in detail in Chapter 1.4. In addition to these transporters, many further sialic acid transporters in other pathogenic bacteria are known nowadays, for example in *Streptococcus pneumoniae* or *Staphylococcus aureus* (North et al., 2018).

The exogenous source for the sialic acid is the direct environment of the pathogen, represented in case of an infection by human cells and serum. The sugar is commonly presented in human serum with around 1.5 to 2.2 mM, but only small amounts are free (around 2 µM) and the main part (around 99.9%) is in a bound state, e.g. in oligosaccharides (Sillanaukee et al., 1999). Due to this, some bacteria, such as *V. cholerae*, feature a sialidase enzyme (neuraminidase) that removes sialic acid from the oligosaccharides and allows the subsequent import into the bacterium (Ada and French, 1959, Corfield, 1992). However, *H. influenzae* bacteria have no sialidase genes in their genome but the previously mentioned study with infection of chinchillas strongly suggested the host-derived sialic acid in their LPS (Bouchet et al., 2003). Therefore, *H. influenzae* is possibly using the minimal amounts of natively, free sialic acid. Alternatively, the *H. influenzae* is involved in an interbacterial competition on the sialic acid, where other bacteria release sialidase to also receive sialic acid, such as *Streptococcus pneumoniae* (Shakhnovich et al., 2002).

Due to the impact to the virulence of *H. influenzae*, the pathways that are involved in LPS sialylation, the uptake of sialic acid and attachment to the LPSs, are interesting and potential targets for attempts to reduce the pathogens virulence. Mutational inactivation of involved proteins already showed a reduced virulence of the pathogen (Bouchet et al., 2003, Fox et al., 2006). However, the LPS synthesis and attachment of sialic acid are performed in the cytoplasm and a potential inhibitor, which blocks the function of a protein during LPS synthesis, depends on its specific and efficient transporter over the cell membrane. Even if methods and ways for this purpose are available, the membrane transport for drug targets is still often a challenging task (Yang and Hinner, 2015). To avoid this, the focus of this work is not on the cytoplasmic LPS sialylation but on the specific transport of sialic acid into *H. influenzae* and the involved transporter proteins. Until today, the function of this system is still not well understood. The known characteristics and a comparison with other transporter systems is described in detail in Chapter 1.3 and Chapter 1.4.

1.2.2 Sialic acid metabolism in *V. cholerae*

Similar to *H. influenzae,* the pathogen *V. cholerae* also uses a TRAP transporter for sialic acid uptake (Severi et al., 2005, Vimr et al., 2004). Due to the high similarity of the predicted structures and function of the TRAP transporter from both organisms, this work includes investigations on both transporters, while the focus is mainly on the transporter from *H. influenzae* (Mulligan et al., 2011).

In contrast to *H. influenzae*, *V. cholerae* were described to use the sialic acid only as source for energy and carbon (Almagro-Moreno and Boyd, 2009). The incorporation of sialic acid into surface exposed LPSs as for *H. influenzae* were not observed during studies of the chemical composition of LPSs from *V. cholerae* (Chatterjee and Chaudhuri, 2003). Furthermore, a genomic analysis revealed no genes that encode relevant enzymes for the LPS sialylation, such as SiaB or NeuA CMP-Neu5Ac synthetases (Almagro-Moreno and Boyd, 2009). In contrast, genes that are responsible for the nutrient pathway as indicated in Figure 1-2, such as NanA (Neu5Ac aldolase), NanK (ManNAc kinase) and NanE (ManNAc-6P epimerase) were detected (Figure 1-3) (Jermyn and Boyd, 2002). Even if the use of sialic acid as nutrient is not directly correlated to the virulence of the pathogen, NanA, NanK and NanE knock-out mutant strains were identified with a significant disadvantage in colonization-competition assays compared to the wildtype strain (Almagro-Moreno and Boyd, 2009). This study suggests that the pathogenicity and virulence of *V. cholerae* also correlates with sialic acid metabolism, but until today there is no direct evidence about the mechanism.

Analogous to *H. influenzae*, *V. cholera* bacteria cannot use a *de novo* synthesis route as source of sialic acid, a gene for the essential NeuB sialic acid synthetase or related enzymes was not detected in a genomic analysis (Almagro-Moreno and Boyd, 2009). Consequently, this pathogen also relies on an exogenous source of sialic acid and a transport of this molecule across the cell membrane into the cytoplasm. As mentioned above, the transport in this bacterium is also carried out by a specific translocation with a TRAP transporter system that is studied and investigated in detail in this thesis (Severi et al., 2005). As the TRAP transporters from both organisms have the same target substrate, they are quite similar and the known details about the function are comparable (Chapter 1.3) (Mulligan et al., 2011, Severi et al., 2005).

One important feature in the sialic acid metabolism of *V. cholerae* that is not present in *H. influenzae,* is the sialidase NanH (neuraminidase), which is used for cleavage of sialic acid from oligosaccharides from other cells and subsequent uptake into the pathogen (Figure 1-3)

(Ada and French, 1959, Vimr et al., 1988). Importantly, in *V. cholerae* the neuraminidase is not only responsible for the metabolism of sialic acid from oligosaccharides but is also involved in the binding and uptake of the cholera toxin of targeted cell and in this way, influence the virulence of the pathogen (Galen et al., 1992). One particularity of *V. cholerae* with respect to sialic acid metabolism is that the corresponding genes for metabolism (NanA, NanK and NanE), the TRAP transporter proteins and the sialidase (NanH) are located on the genetic fragment *Vibrio* pathogenicity island-2 (VPI-2) (Figure 1-3) (Jermyn and Boyd, 2002). This fragment was described to exclusively exist in pathogenic strains of *V. cholerae* and therefore again highlights a presumably high importance of sialic acid metabolism for the pathogenicity and virulence of this bacterium (Almagro-Moreno and Boyd, 2009).

1.3 Translocation of sialic acid across the cellular membrane

Every living cell is surrounded by a membrane that builds up an outer shell and by this, separates the interior of the cell from the environment. A biological membrane is mainly composed of amphipathic lipids such as phospholipids and glycolipids (Lodish et al., 2000). The composition of a cell membrane defines which molecules can cross the membrane by diffusion and for which molecules the membrane is impermeable and another way for translocation is needed (Cooper, 2000). Some bacterial organisms are not only surrounded by just one membrane (Gram-positive), but also from a second bilayer (Gram-negative) (Figure 1-2). The pathogens *H. influenzae* and *V. cholerae*, as well as the mentioned *E. coli* bacterium, belong to the group of the Gram-negative bacteria and therefore feature an outer and an inner membrane with the periplasmic volume between them (Vimr et al., 2004).

Overall, the membrane and its components are not only a barrier between the cytoplasm of a cell and the environment but also involved in many essential processes, for example cell-cell interactions, anchor points for the cytoskeleton or even as virulence factors for pathogens, as presented above (Lodish et al., 2000). Furthermore, every cell membrane contains a large variety of membrane proteins, which constitute between 25-75% of the total membrane mass (Cooper, 2000). One important task of several membrane proteins is the uptake and extrusion of molecules to cross the impermeable membrane or to transport against an energy gradient. The overall importance of membrane transporters becomes clear with the presumption that around 20-30% of the genomes open reading frames (ORFs) of eukaryotes and prokaryotes are encoding for membrane proteins (Wallin and Heijne, 1998). The essential impact is also underlined with the estimation that around 50-60% of FDA-approved (food and drug administration of the US) pharmaceutical drugs have a membrane protein as their target (Overington et al., 2006, Cellular gatekeepers, 2016).

Just as the molecules which are translocated over the membrane have a large diversity, the class of membrane transporter proteins are vary in structure and function (Saier, 2000). The following chapters introduce into structural and functional mechanisms of bacterial membrane transporters and finally, the responsible TRAP transporters for sialic acid import into *H. influenzae* and *V. cholerae* are presented.

1.3.1 Overview of bacterial active membrane transporters

Until today, membrane proteins from the two major groups of primary active and electrochemical potential-driven (secondary active) transporters are assigned to be responsible for sialic acid uptake in bacteria (North et al., 2018). Until today, only the structure for one sialic acid transporter from the secondary active group is known, while the other transporters, including those from *H. influenzae*, *V. cholerae* or *E. coli*, remain unknown (Wahlgren et al., 2018). To get an idea about the general process and mechanistic of membrane transporters in bacteria, the groups of primary and secondary active transporters are firstly described in the following chapters. The presented information about the two major groups might be helpful for the interpretation and comparison of the (sub-)group of sialic acid TRAP transporters from *H. influenzae* and *V. cholerae*.

1.3.1.1 Structure and function of primary active transporters

One class of transporters that is responsible for sialic acid transport in bacteria are the ATP-binding cassette transporters (ABC transporters), the most prominent representatives of the primary active transporter group which represent one of the largest superfamily of proteins (Davidson et al., 2008, Saurin et al., 1999, North et al., 2018). In general, this group can be divided into the two main functional classes of importers and exporters with common structure and function (Licht and Schneider, 2011). Until today, several bacterial sialic acid ABC importers were identified, e.g. for *Haemophilus ducreyi* and *Streptococcus pneumoniae*, but their structures are still unknown (Marion et al., 2011, Post et al., 2005). Therefore, an overview about the general class of bacterial ABC importers is presented in the following paragraph, regarding their structure and function.

As illustrated in Figure 1-5, the structure of bacterial ABC importers consists of five functional domains, two transmembrane domains (TMD, blue), two nucleotide binding domains (NBD, green) and one optional substrate binding protein (SBP, red) (Licht and Schneider, 2011). ABC importers are mainly assigned to type I and type II classes, for example MalE-FGK$_2$ and BtuCD-F in Figure 1-5, with variations in the number of transmembrane helices (TMH) (Korkhov et al., 2012, Oldham et al., 2007, Thomas and Tampé, 2020). Nonetheless in both types, the TMHs form the general translocation pathway for the substrates. In general, the TMDs have a low degree of conservation between different ABC transporters and are therefore presumed as an important determinant for substrate specificity (Licht and Schneider, 2011). In contrast to the TMDs, the NBDs of ABC transporters are more conserved and certain functional motifs, for example the Walker A and Walker B motifs, can be found in canonical ABC transporter (Figure 1-5, green) (Schneider and Hunke, 1998, Thomas and Tampé, 2020). These motifs correlate with the general function of the NBDs, to generate the necessary energy for substrate transport. In a nutshell, the conserved motifs of the NBDs bind ATP (adenosine triphosphate) from the cytoplasm and thereupon induce a short dimerization of the NBDs, which is again weakened and separated after hydrolysis of ATP (Thomas and Tampé, 2020). This process is accompanied by changes of the conformation of the TMDs.

Additional to the four essential components, some ABC transporters have a substrate-binding protein (SBP) as further functional domain (Figure 1-5, red) (Thomas and Tampé, 2020). This protein often occurs in prokaryotic ABC importers and only a few exceptions of importers without an SBP are known (Licht and Schneider, 2011, Rodionov et al., 2006, Rodionov et al., 2009). Because SBPs are the subject for a methodical comparison in result Chapter 2, they are introduced with more detail. In Gram-negative bacteria, the SBPs often exist as individual proteins that freely roam in the periplasmic space, as the MalE SBP from the *E. coli* MalE-FGK$_2$ transporter in Figure 1-5 (Oldham et al., 2007). Alternatively, for Gram-positive bacteria and archaea without a periplasm, SBPs are known that are attached to the cell membrane in the near of the transmembrane proteins via a lipid motif or a transmembrane segment in their structure (Albers et al., 1999, Sutcliffe and Russell, 1995). Further, SBPs can be also directly attached to the N- or C-terminus of the transmembrane domains (van der Heide and Poolman, 2002).

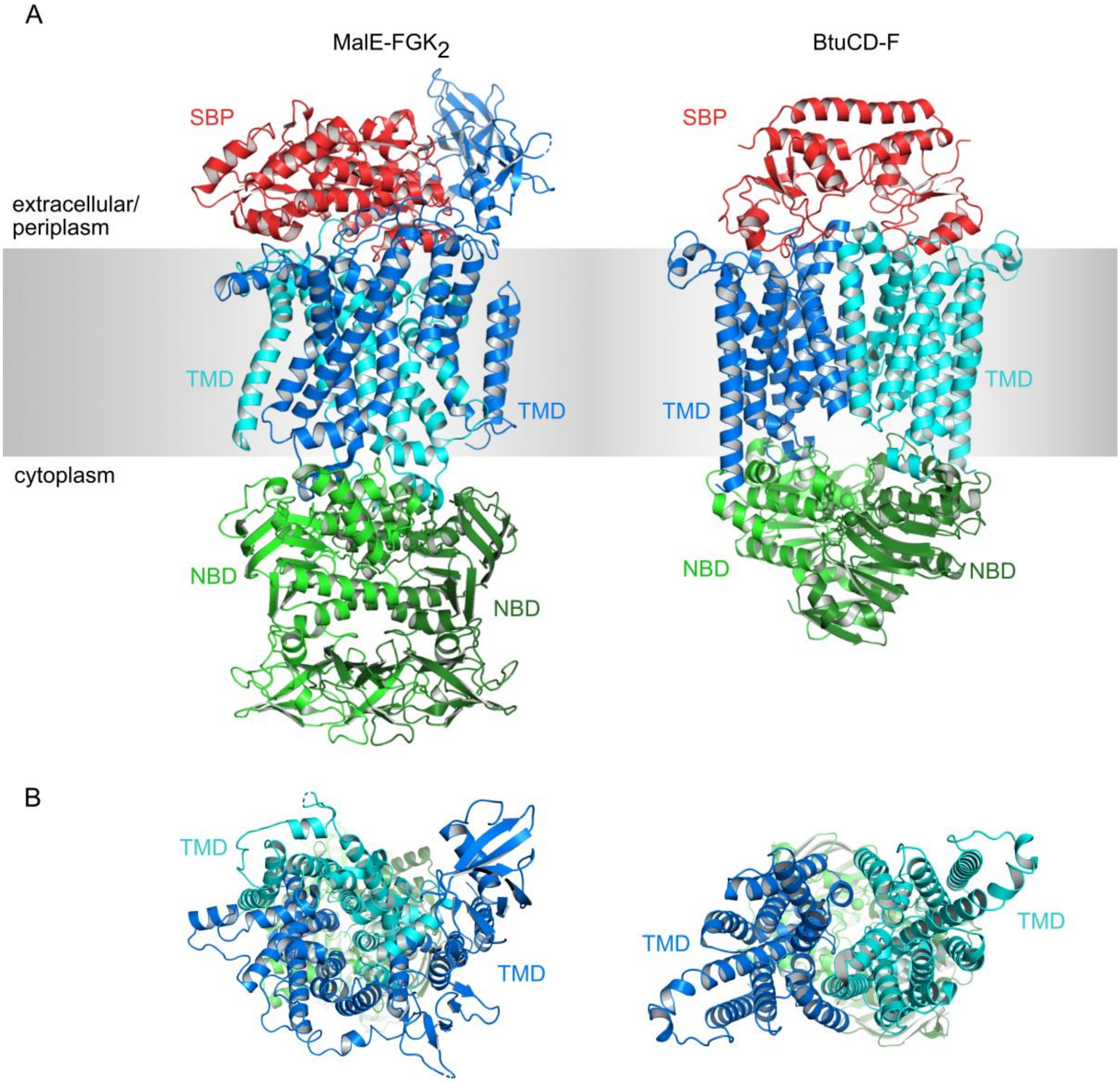


Figure 1-5: Structures of bacterial ABC transporters. A) Bacterial primary ABC importers type I MalE-FGK$_2$ (left, PDB-ID: 2R6G (Oldham et al., 2007)) and type II BtuCD-F (right, PDB-ID: 4FI3 (Korkhov et al., 2012)). The SBPs are indicated in red, the TMDs in blue and the NBDs in green. The membrane region is represented in grey. **B)** Views from the extracellular space through the transporter into the cytoplasm. Same proteins and colours as in A), for a better overview the SBPs are not shown.

The overall structure of SBPs is conserved with two domains which are connected by a hinge region (Berntsson et al., 2010). This creates a binding pocket in which the SBPs can bind the target substrate with typically affinities of around 0.01-10 µM (Berntsson et al., 2010). As implied by their name, the main task of SBPs is the binding of the substrate of the transporter and a following transfer of the substrate to the transmembrane domains through formation of an SBP-TMD complex (Hollenstein et al., 2007, Hvorup et al., 2007, Oldham et al., 2007). Due to the reason that the bacteria with such ABC transporters do often occur in substrate-limited environment, the SBPs are presumed to be responsible for a higher substrate accumulation and by this, for a higher substrate translocation rate (Licht and Schneider, 2011). This is also supported by the fact that the SBPs are often present in a higher concentration than the membrane transporter, which concentration in the membrane is limited (Shilton, 2008). In addition to these aspects, binding of SBPs to the membrane domains are postulated to have an important role in different steps of the transport process, for example an activating effect on ATP hydrolysis or stabilization of transition states (Davidson et al., 1992, Shilton, 2008).

For the overall ABC transporter function, some parts of the translocation process are already described in detail and transport cycles of type I and type II transporter classes have been postulated (Locher, 2016, Thomas and Tampé, 2020). In a nutshell, the SBP scavenges the substrate from the environment or periplasm and forms a complex with the TMDs. A crystal structure of a MalE-FGK$_2$ complex revealed a loop of a TMD bound in the substrate binding cleft of the SBP, suggesting a “scoop loop” mechanism in which the loop helps to dislocate the substrate from the SBP and deliver it to the translocation pathway in the TMDs (Oldham et al., 2007). Afterwards, through hydrolysis of ATP in the NBDs, the TMDs presumably change their conformations from at least an outward-faced open to an inward-faced open conformation (Thomas and Tampé, 2020). This conformational change allows the substrate to enter the cytoplasm and the ABC transporter can readapt the initial states.

1.3.1.2 Structure and function of secondary active transporters

The electrochemical potential-driven (secondary active) transporters form the second group of membrane transporters that are responsible for sialic acid transport in bacteria (North et al., 2018). As for the ABC transporter, the proteins of this group can be further divided into subclasses. Today, the common structural fold of the transporters is used for classification and by this, the two large main groups of major facilitator superfamily (MFS) fold and sodium-coupled leucine transporter (LeuT) fold are specified (Boudker and Verdon, 2010). Importantly, while the proteins in the MFS fold are also mostly functional assigned as one family, the sequence of LeuT fold proteins are classified to different functional families, e. g. neurotransmitter solute symporter (NSS), sodium solute symporter (SSS) and nucleobase cation symport 1 (NCS1) (Forrest et al., 2011). Additional to these fold types, some other structures of secondary transporters are known that are not assigned to these two main folds and summarized in small classes or as single structures, for example the Na^+/H^+ antiporter NhaA or the glutamate transporter Glt_{Ph} (Boudker and Verdon, 2010, Hunte et al., 2005, Yernool et al., 2004). Sometimes, the MFS and LeuT folds are classified regarding their function as proton and sodium transporter classes, respectively. However, this is not strictly valid as sodium MFS fold and proton LeuT fold transporters are also known (Ethayathulla et al., 2014, Shaffer et al., 2009). For sialic acid import into bacteria, representatives from both fold-types are known. Exemplary, the already mentioned NanT transporter from *E. coli* is assigned to the group of MFS fold and the only structurally known sialic acid transporter SiaT from *Proteus mirabilis* to the LeuT fold (sodium solute symporter (SSS) family) (North et al., 2018, Wahlgren et al., 2018). The sialic acid transporters from *H. influenzae* and *V. cholerae* could not be assigned to one of the main fold types and known structural and functional information recommend an own class or subclass of these transporters (Chapter 1.4).

In general, the known structures of secondary transporters differ from the presented ABC transporters and at a first glance, seem to be less complex (Figure 1-6). Both, the MFS and LeuT fold, consist of a transmembrane part without extra-membrane domains. Analogous to the ABC transporters, the TMHs of the secondary transporters form the translocation pathway for the substrate through the membrane (Figure 1-6). The transporters from the MFS fold group are built up from 12 TMHs that can be divided into an N- and C-domain (Figure 1-6, blue and cyan) (Shi, 2013). Based on several available structures, such as the lactose permease LacY (Figure 1-6), the MFS fold transporters are proposed as monomeric functional unit and no dimeric of trimeric MFS fold transporters have been described to date (Abramson et al., 2003, Shi, 2013). The transporters of the LeuT fold consist of two motifs with 5 TMHs, resulting in 10 TMHs that build the structure of the transporter (Figure 1-6) (Shi, 2013). In contrast to the MFS fold, transporters from the LeuT fold group are described in a monomeric, dimeric and trimeric state (Gao et al., 2009, Ressl et al., 2009, Yamashita et al., 2005). Further, also secondary transporters that are not assigned to these two major folds, for example the Na^+/H^+ antiporter NhaA, can be structurally described with two motifs (Figure 1-6) (Hunte et al., 2005). Despite structural differences, the protein structure from the most secondary transporters is based on structural motifs with 10-14 TMHs (Boudker and Verdon, 2010).

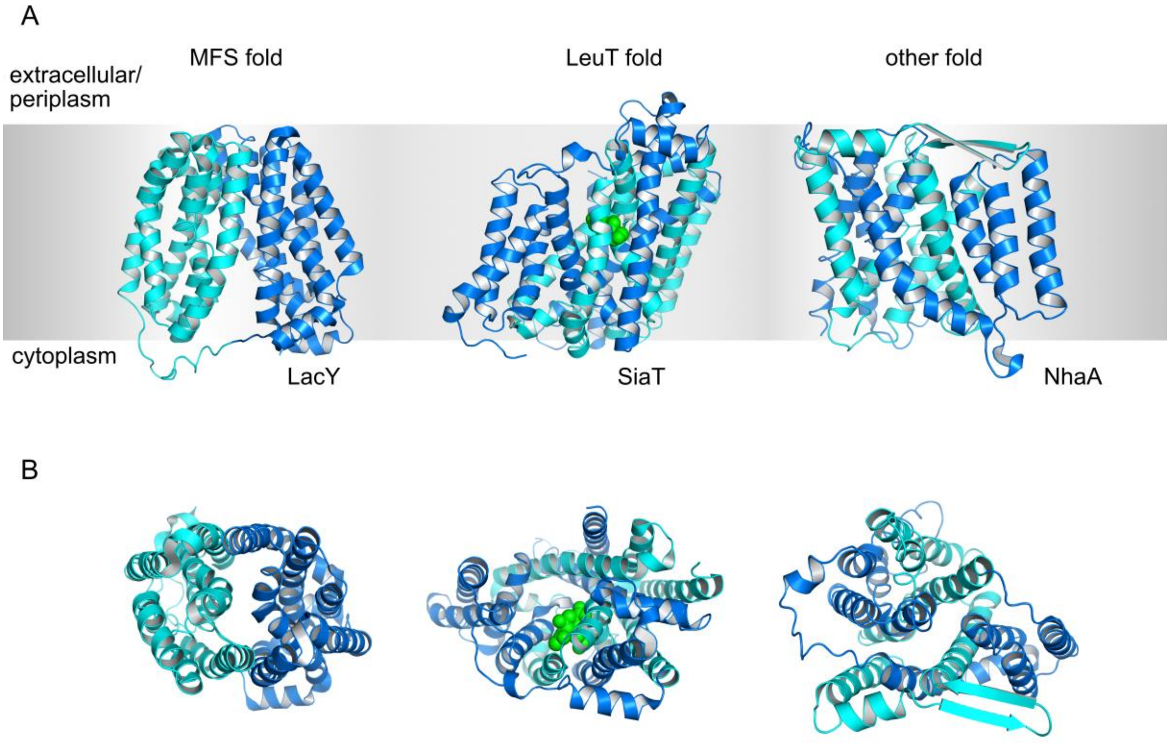


Figure 1-6: Structures of secondary transporters. A) Bacterial secondary transporters LacY from the MFS fold class (left, PDB-ID: 1PV6, (Abramson et al., 2003)), SiaT from the LeuT fold class with bound sialic acid in green (middle, PDB-ID: 5NVA (Wahlgren et al., 2018)) and NhaA (right, PDB-ID: 1ZCD (Hunte et al., 2005)). The TMDs are indicated in blue colours and the membrane area is represented in grey. **B)** Views from the extracellular space through the transporter into the cytoplasm. Same proteins and colours as in A).

The translocation mechanism of a substrate with secondary transporters through a cell membrane is probably accomplished with changes between different conformational states of the TMDs (Shi, 2013). Secondary transporters were postulated several years ago with changes between an outward-facing open, an outward-facing occluded, an inward-facing occluded and an inward-facing open conformation, also known as alternating-access mechanism (Krishnamurthy et al., 2009, Reyes et al., 2009, Weyand et al., 2008). Subsequent from these postulations, all states were successfully observed in structural investigations of secondary transporters and used to confirm and further refine the transport mechanism (Boudker and Verdon, 2010, Forrest et al., 2011, Shi, 2013).

Another important point concerning the transport mechanism is the way of gaining the energy that is needed for substrate translocation. In general, secondary transporters use an electrochemical potential difference that exists across the membrane and do not directly depend on the use of ATP. For this, the secondary transporters use a coupling mechanism whereby not only the substrate is translocated through the membrane but also one or more ions, in most cases a proton or sodium-ion (Forrest et al., 2011). In spite of different ions, the functional mechanism of proton or sodium-ion coupling transporters are postulated to be quite similar (Shi, 2013). Noticeable, also secondary transporters are known which depend on the additional presence of a chloride anion or on the presence of potassium cations (Forrest et al., 2007, Zomot et al., 2007). On the basis of coupling mechanisms, the secondary transporters are separated into symporters, in which the substrate and ion cross the membrane in the same direction, and antiporters, in which both cross the membrane in opposing direction (Forrest et al., 2011).

Even if many details about the coupling mechanism are still unknown today, several crystal structures of sodium-coupling transporters revealed bound cations between the transmembrane helices, for example the sialic acid SiaT transporter, and through this, allowed the identification of conserved sodium binding sites (Krishnamurthy et al., 2009, Wahlgren et al., 2018, Watanabe et al., 2010, Weyand et al., 2008). An integrative study on the coupling mechanism of a sodium-dependent transporter revealed a conformational change of a TMH upon release of the ion, which opens the translocation pathway in direction of the cytoplasm and allows the substrate to enter the cell (Watanabe et al., 2010). The analysis of proton coupling transporters are more difficult because the protons are not visible in crystal structures. However, amino acids that should be involved in proton binding were described in mutational studies on lactose permease symporter LacY (Sahin-Tóth et al., 2000).

The sialic acid transporter in *H. influenzae* and *V. cholerae* are functionally assigned to the group of secondary transporters. They are not mentioned until this point, because they have features and motifs that are unique and not observed for any of the described transporter classes. Further, there are currently no structures available for TRAP transporters and the knowledge about their function is rather small. In the following chapter, the transporters are introduced in detail and compared with the presented structural and functional knowledge from primary ABC and classical secondary active transporters.

1.4 Sialic acid TRAP transporter in *H. influenzae* and *V. cholerae*

During the characterization of a transport system for C4-dicarboxylates malate and succinate for the bacterium *Rhodobacter capsulatus* in the early 1990s, the group of David J. Kelly (University of Sheffield) described the existence and need of an SBP for uptake of this substrate (Shaw and Kelly, 1991). Further studies on this SBP, called DctP, confirmed a high binding affinity to the substrates and binding characteristics that were previously observed for SBPs from ABC transporters (Quiocho, 1990, Walmsley et al., 1992). However, a sequence analysis indicated only two corresponding genes which encode for integral membrane proteins (DctQ and DctM) and no characteristic genes for NBDs of ABC transporter were observed (Forward et al., 1997). Functional studies with transporter class specific inhibitors showed a high sensitivity of this tripartite DctPQM system to secondary transporter inhibitors (Forward et al., 1997). This observation suggested a membrane potential-dependent, ATP-independent transport mechanism. Further, all three genes were identified as essential for the substrate transport and homologous genes were identified for other Gram-negative bacteria (Forward et al., 1997). This description of a secondary transporter system, which functionally depended on the presence of an SBP, led to the formation of the new class of tripartite ATP-independent periplasmic (TRAP) transporters (Forward et al., 1997).

After the initial description and classification, several following genome sequence analyses confirmed the existence of this transporter class in many bacteria and archaea and showed that the transporters are commonly used by these organisms (Kelly and Thomas, 2001, Mulligan et al., 2007, Rabus et al., 1999). Simultaneously, the results revealed the fact that these SBP-dependent secondary transporters are absent from eukaryotes, similar to the SBP-dependent ABC importers (Davidson et al., 2008, Kelly and Thomas, 2001). After the description of TRAP transporters, a second group of SBP-dependent secondary transporters, the tripartite tricarboxylate transporter (TTT), was described and due to no sequence similarity to TRAP transporters, it was defined as an individual class (Winnen et al., 2003). The functional units of these TTTs are similarly postulated to be one SBP and two TMDs but the current knowledge about this group is even more sparse than for the TRAP transporters (Winnen et al., 2003).

Considering only the class of TRAP transporter, the gene analyses also revealed significant variance in the SBP, which lead to a further distinction and formation of subclasses. In addition to classical TRAP transporters, which have a high sequence identity with the initially described DctP from *R. capsulatus*, a second group of TAXI proteins (TRAP-associated extracytoplasmic immunogenic) with low identity to DctP was defined (Kelly and Thomas, 2001, Rabus et al., 1999). A decade after definition of these two groups, an additional type of TRAP transporters was identified during the analysis of a TRAP-related transporter system from the pathogenic bacterium *Treponema pallidum* (Deka et al., 2012). Additional to the classical SBP and TMDs, a further correlated gene was observed and the corresponding protein was identified with similarities to a tetratricopeptide repeat (TPR) motif, which is also responsible for the new name TPAT transporter (TPR-protein-associated TRAP) (Deka et al., 2012). Functional studies on TPAT transporters revealed a certain interaction between the TPR and SBP and suggest a mechanism, in which especially hydrophobic substrates are bound and translocated through the membrane (Brautigam et al., 2012). However, until today both identified TRAP transporter subclasses, the TAXI TRAP and TPAT transporters, are less studied than the DctP TRAP transporters and also the present work and following paragraphs are focused on the classical TRAP transporters.

1.4.1 Substrates of TRAP transporters

The widespread utilization of TRAP transporters in bacteria and archaea suggests that the systems are not only responsible for the initially described transport of C4-dicarboxylates or derivates, but also for many other molecules. Indeed, until today, several molecules that are translocated by TRAP transporters have been identified (Mulligan et al., 2011, Rosa et al., 2018). Examples for substrates are ectoine and 5-hydroxyectoine, both translocated from the bacterial TeaABC and UehABC TRAP transporters and are used by the bacteria as compatible solute against osmotic stress or as nitrogen and carbon source (Grammann et al., 2002, Kuhlmann et al., 2008, Lecher et al., 2009). Further identified example substrates of TRAP transporters are monocarboxylic 2-oxo-acids (α-Keto acids) such as pyruvate (RRC01191 (Thomas et al., 2006)), the amino acid glutamate (GtrABC (Quintero et al., 2001)) and recently discovered five different sugar acids (dctPQM$_{Am}$ (Meinert et al., 2017, Thomas, 2017)). Furthermore, a study analysed in a high throughput approach several hundred TRAP transporter SBPs and identified more than 40 new SBP ligands, immensely increasing the number of identified TRAP transporter substrates (Vetting et al., 2015).

Among these substrates, the best studied TRAP transporter systems are specific for the substrate sialic acid (*N*-acetylneuraminic acid). During investigations of genes for sialic acid metabolism from *H. influenzae* and *V. cholerae*, the observation of correlated TRAP transporter encoding genes suggested a first connection between this substrate and TRAP transporters (Figure 1-3) (Kolker et al., 2004, Vimr et al., 2004). Shortly after this hypothesis, further studies proved the uptake of sialic acid in *H. influenzae* by a SiaPQM TRAP transporter (Allen et al., 2005, Severi et al., 2005). As mentioned before, *H. influenzae* has no genes for a *de novo* synthesis of Neu5Ac and investigations of SiaPQM in *H. influenzae* (named HiSiaPQM) identified the TRAP transporter as the sole route for sialic acid uptake in this pathogen (Severi et al., 2005). During investigations of sialic acid metabolism from *V. cholerae*, also in this pathogen a SiaPQM transporter (named VcSiaPQM) could be identified as the only uptake pathway of sialic acid in these bacteria (Almagro-Moreno and Boyd, 2009, Chowdhury et al., 2012, Mulligan et al., 2009). The genes for this transporter are localized in the VIP-2, similar to the other genes that are involved in the sialic acid metabolism and by this, support the suggested importance of the sialic acid for the virulence of this pathogen (Figure 1-3) (Almagro-Moreno and Boyd, 2009, Jermyn and Boyd, 2002). Even though the sialic acid transporters HiSiaPQM and VcSiaPQM are the best described TRAP transporters, the following chapters about structural and functional aspects reveal a large number of unknown details.

1.4.2 Structural and functional aspects of TRAP transporters

1.4.2.1 Substrate binding proteins of TRAP transporters

Until today, only TRAP SBPs from Gram-negative bacteria have been structurally described (Müller et al., 2006). Similar to several ABC transporters, they are postulated to freely move in the periplasm (Shaw and Kelly, 1991). For Gram-positive bacteria, genes that correspond to TRAP transporter proteins are also detected and a similar behaviour as for anchored ABC transporter SBPs via a fusion protein or lipid anchor is presumed (Chapter 1.3.1.1) (Albers et al., 1999, Fischer et al., 2010, Sutcliffe and Russell, 1995). Also, TRAP transporters are known in which the SBP is either fused to the Q-domain, as in a TRAP transporter from *Rhizobium* (sp. strain NGR234, YAMM transporter), or to the M-domain, as predicted for a C4-dicarboxylate related TRAP transporter in *Acidaminococcus intestini* (strain RyC-MR95) (Rabus et al., 1999).

A breakthrough in the mechanistic understanding of TRAP transporter SBPs was the first crystal structure of the SBP HiSiaP from HiSiaPQM (Müller et al., 2006). In this study, the protein was observed in two states, a substrate-free and substrate-bound state. The structure revealed an overall fold that is similar to SBPs from ABC transporters, consisting of two domains that are connected by a hinge region and forming a cleft between them (Figure 1-7, A) (Berntsson et al., 2010). The comparison of apo and substrate-bound structure of HiSiaP showed that one molecule sialic acid binds in this cleft and triggers a large conformational change (Figure 1-7, A) (Müller et al., 2006). Until today, several other protein structures of TRAP transporter SBPs have been described, including sialic acid SBPs from several other organisms such as *Vibrio cholerae* (VcSiaP)*, Pasteurella multocida* (PmSiaP) and *Fusobacterium nucleatum (*FnSiaP) (Gangi Setty et al., 2014, Rosa et al., 2018). In contrast to HiSiaP, these three SBPs are either described in the substrate-free or substrate-bound conformation. Overall, the structures of the known TRAP transporter SBPs are quite similar and assigned to the Cluster E of the SBP classification, which includes nowadays more than 70 individual structures (including SBPs from the TTT class) (Berntsson et al., 2010, Scheepers et al., 2016).

While the presented sialic acid SBPs are all observed and presumed to be functionally active in a monomeric state, three SBP crystal structures from different TRAP transporters were observed with a tendency to form dimeric or tetrameric assemblies (Akiyama et al., 2009, Cuneo et al., 2008, Gonin et al., 2007). Some of these studies proposed a possible functional impact of this oligomerization that was also observed in solution, even if the consequences for the transport process remained elusive.

On the basis of several described TRAP SBP structures, the binding interactions of the SBP with its substrate can be analysed on an atomic scale. In general, the overall location of the substrates is conserved. Due to the variability of substrates, the interacting amino acids change for different SBPs (Rosa et al., 2018). One exception is an arginine side chain at position 147 in HiSiaP, which is highly conserved for TRAP SBPs (96.8% in 6142 sequences (Rosa et al., 2018)). This residue has a crucial role in substrate binding and by this, is presumed to act as a selectivity filter for substrates with a carboxylate group (Figure 1-7, B and C) (Fischer et al., 2015, Johnston et al., 2008). The substrate-bound structures of the sialic acid SBPs HiSiaP, FnSiaP and PmSiaP show a high similarity of the substrate interacting amino acids (Gangi Setty et al., 2014). The characterization of substrate binding revealed high affinities and similar binding characteristics of these three SBPs with binding strengths around 16-120 nM (Gangi Setty et al., 2014, Johnston et al., 2008, Mulligan et al., 2009, Severi et al., 2005). For sialic acid TRAP transporter SBP from *V. cholerae*, VcSiaP, only a substrate-unbound protein structure is known and the binding strength was determined in a previous study to 306 nM (Gangi Setty et al., 2014).

Additional to the native ligand, an interaction study of HiSiaP with derivates of sialic acid sugar group revealed a high specificity of the protein and only for the similar Neu5Gc (*N*-glycolylneuraminic acid), a comparable binding affinity was observed (Severi et al., 2005). The mutational characterization of amino acids which were postulated to be important for substrate binding gives nowadays a detailed characterization about the protein-substrate interaction in sialic acid SBPs (Gangi Setty et al., 2014, Glaenzer et al., 2017, Johnston et al., 2008). Recently, it was shown that also the highly ordered water network in the binding cleft has a huge impact on the substrate affinity, which is therefore postulated to contribute to the affinity and specificity of the protein (Darby et al., 2019). Based on the work of our group, a crystal structure and binding studies showed that VcSiaP is also able to bind short peptides and by this, occupies the substrate binding cleft (Peter et al., 2021). Even if binding studies resulted in a low affinity for the peptide of around 30 µM, this structure was the first evidence of an artificial bound substrate to a sialic acid SBP and allowed to postulate further details about the closing mechanism.

Furthermore, the dynamic behaviour of the closing mechanism upon substrate binding in VcSiaP was studied previously (Glaenzer et al., 2017). Detection of structural changes with distance measurements in frozen solution confirmed the existence of only two conformations, the open state in the absence of substrate and the closed state upon substrate addition (Glaenzer et al., 2017). No intermediate states were observed in this study, which supported the “Venus flytrap” mechanism for sialic acid TRAP SBPs, as postulated for the first TRAP transporter SBP DctP (Mao et al., 1982, Walmsley et al., 1992, Glaenzer et al., 2017).

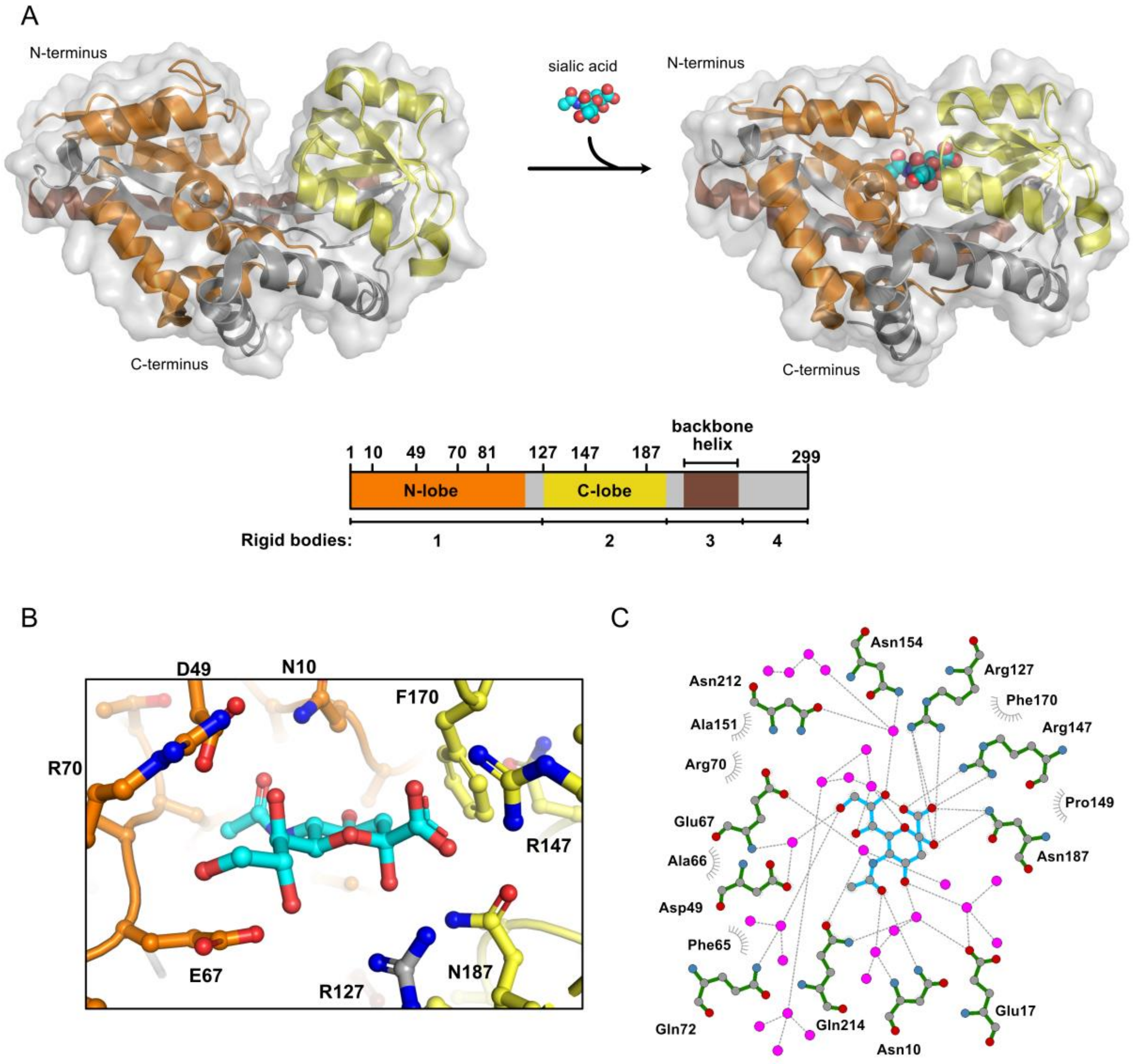


Figure 1-7: Structure of sialic acid TRAP transporter SBP. A) Protein structures of HiSiaP in cartoon representation with transparent surface in a substrate free (left, PDB-ID: 2CEY (Müller et al., 2006)) and substrate bound (right, PDB-ID: 3B50 (Johnston et al., 2008)) conformational state (N-lobe in orange, C-lobe in yellow, backbone helix in brown and remaining motifs in grey). The substrate sialic acid is illustrated as a ball-and-stick model (coloured cyan) and an outline of the primary structure of HiSiaP with structural motifs, rigid bodies and positions of selected amino acids are given. **B)** Representation of the binding cleft from substrate bound HiSiaP with sialic acid and interacting amino acids, same colours as in A). **C)** Analysis of the substrate-protein interaction (LigPlot (Laskowski and Swindells, 2011)), the amino acids are shown in green, the substrate in cyan and the water molecules in magenta.

1.4.2.2 Transmembrane domains of TRAP transporters

The transmembrane domains of TRAP transporters are much less well understood than the SBPs, which is caused by the generally higher requirements of membrane proteins in terms of the experimental handling in functional and structural studies (Chapter 4.1). Until today, no experimentally determined protein structures of TRAP transporter transmembrane domains are known. However, theoretically models, based on the topology or evolutionary coupling predictions, are available (Mulligan et al., 2012, Ovchinnikov et al., 2014).

During the first description of TRAP transporters, two transmembrane domains, a smaller (DctQ) and a larger one (DctM), were identified (Forward et al., 1997). Additionally, the amino acid sequence of the two transmembrane domains were analysed regarding their hydrophobic character and the domains were postulated to consist of 4 and 12 transmembrane helices for DctQ and DctM, respectively (Forward et al., 1997, Kelly and Thomas, 2001). This is a common feature of TRAP transporters and the sialic acid transporters from *H. influenzae* and *V. cholerae* also have this canonical architecture (Figure 1-8) (Mulligan et al., 2012). A comprehensive gene analysis of TRAP transporters from several organisms revealed that the two membrane domains are sometimes fused into one peptide chain, as for example in HiSiaPQM (Figure 1-8 A) (Kelly and Thomas, 2001). However, even though the proteins are not covalently fused and expressed as single domains, as for VcSiaPQM (Figure 1-8 B), a stable complex between the two domains was observed (Mulligan et al., 2012).

Due to the lack of high-resolution structures, the knowledge about the function and mechanistic relevance of the two TMDs in the transport process is sparse. Based on the prediction of 12 TMHs, the larger M-domain is presumed to be comparable to the MFS group of secondary transporters and due to this, postulated to be responsible for the translocation pathway of the substrate through the membrane (Forward et al., 1997, Shi, 2013). In contrast, the function of the smaller Q-domain is totally unknown. Gene analysis of TRAP transporters revealed a lower sequence identity of this domain to other homologues as for the M-domain (Rabus et al., 1999). Nevertheless, functional studies of the DctPQM TRAP transporter showed that the presence of the Q-domain is essential for substrate transport and that the TRAP transporter without this domain was not functional (Forward et al., 1997). One hypothetical task is that the Q-domain in the membrane serves as an anchor and recognition site for the corresponding SBP, which might explain the combined and unique presence of the Q- and SBP domains in TRAP transporters (Kelly and Thomas, 2001, Rabus et al., 1999). The observation of TRAP transporters with covalently fused PQ-domains, as well as the described fusion of QM-domains, support the idea that the Q-domain mediates the interaction between the SBP and M-domain (Rabus et al., 1999). Another possibility that is not directly linked to the transport process is a regulatory or chaperone-like function of the Q-domain (Kelly and Thomas, 2001, Rabus et al., 1999). Finally, the Q-domain might coordinate the coupling of an ion gradient to the substrate translocation, as similar domains were never observed in SBP-dependent ABC transporters (Kelly and Thomas, 2001).

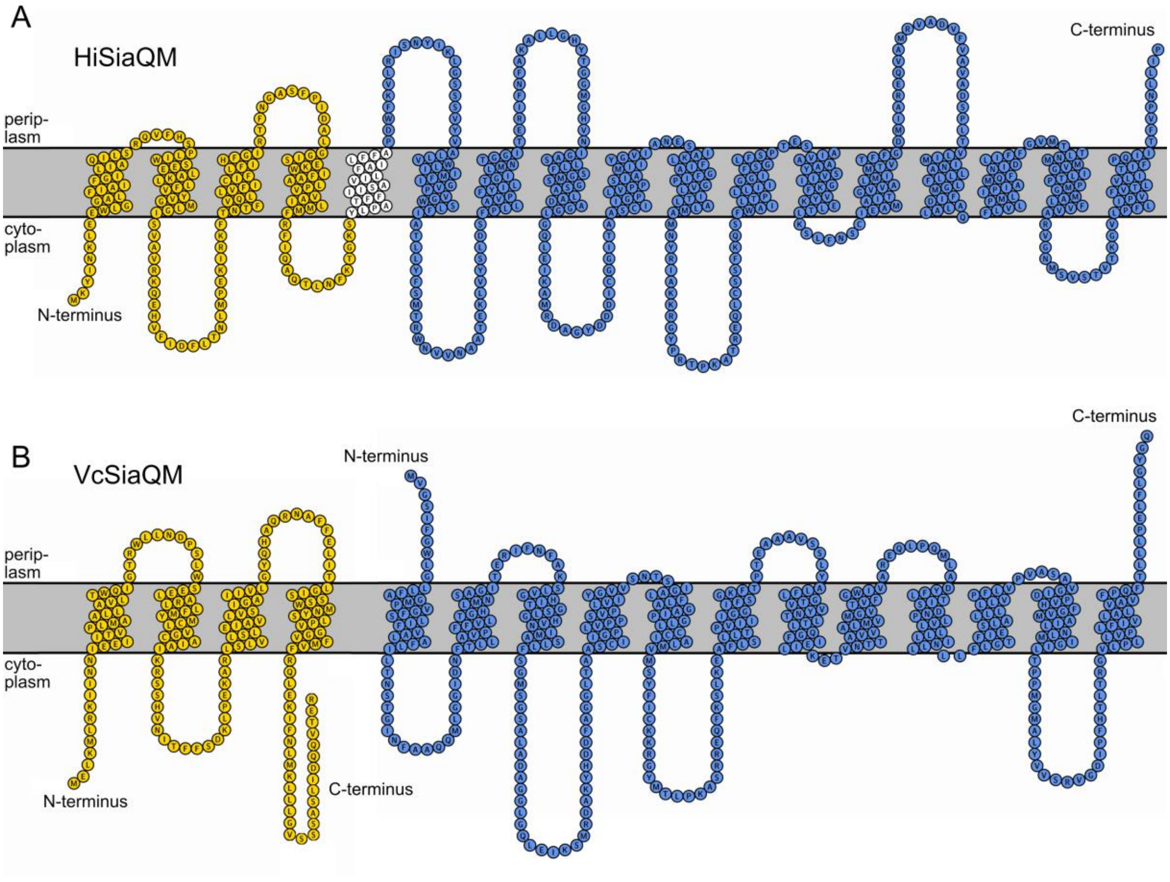


Figure 1-8: Topology models of TRAP transporter TMDs. A) Prediction of the TMHs from the sialic acid TRAP transporter membrane domains from *H. influenzae.* The Q-domain is illustrated in yellow, the M-domain in blue and connecting TMH in white. The TMH prediction was taken from the corresponding UniProtKB entry P44543 (access 17.11.2020) (The UniProt Consortium, 2019) and the illustration was prepared with Protter (Omasits et al., 2014), modified after (Mulligan et al., 2012). **B)** Same as A) but for sialic acid TRAP transporter membrane domains form *V. cholerae,* UniProtKB entry Q9KR65 and Q9KR66 (access 17.11.2020) (The UniProt Consortium, 2019).

1.4.3 Proposed transport mechanism of TRAP transporters

Regarding the transport pathway for sialic acid uptake in *H. influenzae* and *V. cholerae,* the first barrier is not the inner membrane (IM) with the specific TRAP transporter, but the crossing of the outer membrane (OM) into the periplasm (Figure 1-2). In general, small molecules such as sialic acid can enter the periplasm by diffusion through specific or unspecific porins or with an active transport system (North et al., 2018). For *E. coli*, a specific sialic acid outer membrane porin, NanC, was described and identified as essential for sialic acid uptake into the bacterium in the absent of unspecific porins (Condemine et al., 2005, Wirth et al., 2009). However, for sialic acid uptake in *H. influenzae* and *V. cholerae,* no NanC homologs or other sialic acid specific porins were described and the translocation of sialic acid over the OM are therefore presumably occurs with unspecific porins, such as the OmpP2 porin from *H. influenzae* (Coulton et al., 1992, Severi et al., 2005).

As mentioned before, the TRAP transporters HiSiaPQM and VcSiaPQM from *H. influenzae* and *V. cholerae* are the only routes for uptake of the sialic acid across the IM into the cytoplasm (Figure 1-2) (Chowdhury et al., 2012, Severi et al., 2005). Following the postulated transport cycle for TRAP transporters from Mulligan et al. (Mulligan et al., 2011), the substrate in the periplasm is initially recognized by a freely moving SBP, which tightly binds the substrate with high affinity and specificity and thereby adopts the closed conformation (Figure 1-9) (Gangi Setty et al., 2014, Glaenzer et al., 2017, Johnston et al., 2008, Müller et al., 2006). This state is presumably recognized with high specificity by the membrane domains, that rest in an outward-faced closed conformation as expected for classical secondary transporters (Boudker and Verdon, 2010). The presumed high specificity of the interaction between the SBP and the membrane domains is based on the experimental observation that the exchange of HiSiaP to the similar VcSiaP in a functional assay was not able to transport the substrate, probably due to no interaction between VcSiaP and HiSiaQM (Mulligan et al., 2009).

After formation of the tripartite PQM complex, which was never experimentally observed until today, the SBP presumably again adopts an open-like state and by this, releases the substrate to the TMD, similar as described for SBP dependent ABC transporter (Figure 1-9) (Thomas and Tampé, 2020). One possible mechanism for this process is a scoop loop mechanism, previously observed for MalE-FKG$_2$, in which a loop of the TMD interacts with the substrate binding cleft of the SBP and by this, supports the detachment of the substrate and enables the entering to the translocation pathway in the TMDs (Oldham et al., 2007). Either during formation of the tripartite complex, the substrate releasing of the SBP or shortly after these events, the TMDs adopt an outward-faced open state and incorporate the substrate from the SBP. After this step, the TMDs presumably undergo a conformational shift to an inward-faced open state (Figure 1-9). The conformational changes of the TMDs rely on an electrochemical potential of sodium ions over the membrane and a co-translocation of most likely two sodium ions during substrate transport (Mulligan et al., 2009). After the conformational change, the sialic acid is released from the TMDs into the cytoplasm and can be used for further metabolism (Chapter 1.2). Finally, the transporter domains adopt the starting conformations for a new transport cycle. This means that the tripartite complex dissociates, the SBP adopts the open conformation and the TMDs change back to the outward-faced closed state.

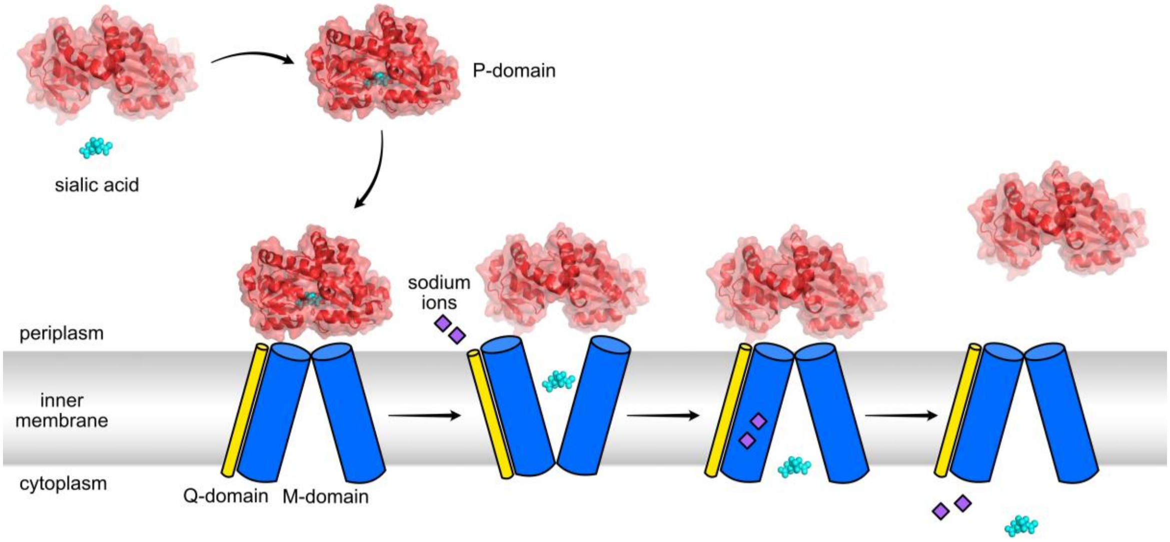


Figure 1-9: Hypothetical transport mechanism of TRAP transporters. The illustrated transport mechanism is based on current knowledge about TRAP transporters in combination with the function of classical secondary transporters, adapted and modified from (Mulligan et al., 2011). The SBP is indicated in red in the two described states, substrate unbound (PDB-ID: 2CEY (Müller et al., 2006)) and substrate bound (PDB-ID: 3B50 (Johnston et al., 2008)). The TMDs are located in the inner membrane (grey) and the Q- (yellow) and M-domain (blue) are schematically indicated, as well as the substrate sialic acid as ball-and-stick model in cyan and the co-translocated sodium ions as purple squares.

Overall, even if the transport process of TRAP and classical secondary transporters are presumably quite similar, there are important differences. TRAP transporters catalyse only a unidirectional translocation and are not bidirectional, as observed for secondary transporters (Mulligan et al., 2011). This fact is likely due to the essential presence of the SBP for transport. Another important fact about TRAP transporters reveals the inspection of the energy balance. If only indeed two sodium ions are needed for substrate translocation, the overall energy balance for TRAP transporter is lower than for ATP-consuming ABC transporters (Mulligan et al., 2009). At the same time, due to the SBP, the TRAP transporters possibly have the similar advantages as SBP-dependent ABC transporters, in which the SBPs for example enable an efficient substrate accumulation or a stabilization of transition states of the TMDs (Chapter 1.3.1.1).

1.5 Research Aim

The aim of this thesis is to gain knowledge about the class of the sialic acid TRAP transporters from the pathogens *H. influenzae* and *V. cholerae*. As presented above, the sialic acid transport by the TRAP transporters HiSiaPQM and VcSiaPQM are the only uptake routes for these pathogens and strongly correlate with their virulence, as proven for *H. influenzae* (Allen et al., 2005, Severi et al., 2005). Through combination of several orthogonal bioanalytical methods, this work should achieve new insights into structural and functional parts, as well as essential interactions in the TRAP transporter system.

Among other methods, two integrative methods that should be used in this work for the characterization of the TRAP transporters are smFRET (single molecule Förster resonance energy transfer) and PELDOR (pulsed electron electron double resonance) spectroscopy. Due to a comparable outcome of these distance measuring methods in integrative structural biology experiments, a comprehensive comparison should reveal the advantages and strengths of each method. Further, the comparison should show in which way the results of both methods can be combined to overcome complementary limitations.

Because the structure of the SBPs from the two sialic acid TRAP transporters HiSiaPQM and VcSiaPQM are only partially described, the focus on the SBPs in the present work is set on the missing crystal structure of VcSiaP in a substrate-bound conformation. The new substrate-bound protein structure will allow a comparison of the protein-substrate interactions between sialic acid SBPs. Differences that are observed should be investigated with mutational analyses to further characterize the protein-substrate interaction. Additional to these studies, smFRET measurements should be used to analyse the closing mechanism of sialic acid SBPs in solution.

A further aim of this work was to gain structural and functional insights into the TMDs of TRAP transporters. In the beginning, working procedures for expression, extraction and purification of the membrane proteins HiSiaQM and VcSiaQM must be developed and established. While doing this, the reconstruction of TRAP transporters in different membrane mimetic systems should be tested, as membrane proteins are natively embedded in and interacting with a lipidic environment. Afterwards, the TRAP transporters should be analysed in structural and functional studies. Prior to functional PELDOR spectroscopic studies, a procedure towards specifically labelled constructs of the membrane proteins should be developed. Then, PELDOR spectroscopy distance measurements should be used to detect conformational states in the membrane proteins and analyse interactions with various additives, such as the substrate sialic acid, the corresponding SBP or the membrane mimetic system. Ideally, the results should experimentally validate current theoretical models of the proteins.

For further studies, TRAP transporter nanobodies should be generated through immunization of a camelid, as these specific binders are helpful tools for many bioanalytical studies, especially for membrane proteins. For the case of strong binding nanobodies, an influence of the binding on the function of the transporter system can be analysed in functional studies. Further, they can be combined with structural studies. In this work, the two major methods for structure elucidation of membrane proteins, X-ray crystallography and cryo-EM (electron microscopy) should be used to analyse the TRAP transporter membrane domains. This two-pronged approach should enhance the chance of the successful description of the membrane proteins structure. New functional aspects and especially the description of the first structure of the membrane protein of TRAP transporters can be possibly used to develop specific inhibiting strategies against the uptake of sialic acid and by this, open avenues towards reduction of the virulence of the two pathogens *H. influenzae* and *V. cholerae*.

Chapter 2

Comparison of methods for distance measurements in proteins

Preamble:

The work in this chapter was accomplished in cooperation with Christian Gebhardt and Thorben Cordes (University of Munich). They supplied the purified MalE and SBD2 cysteine mutants and performed smFRET measurements and analysis.

The content of this chapter is published in: Peter, M. F., Gebhardt, C.; Mächtel, R.; Glaenzer, J.; Thomas, G. H.; Cordes, T.; Hagelueken, G.; Cross-validation of distance measurements in proteins by PELDOR/DEER and single-molecule FRET. *BioRxiv,* **2020**, https://doi.org/10.1101/2020.11.23.394080.

2.1 PELDOR and smFRET spectroscopy in integrative structural biology

As presented in Chapter 1.3, nearly all proteins and especially membrane proteins are dynamic molecules and rely on conformational changes to fulfil their biological task, such as the transport of a substrate across a membrane. While often difficult, the dynamical characteristics can be studied using structural methods as presented in Chapter 7.1. But, in two of the three major methods for structure determination, X-ray crystallography and cryo-EM, the proteins are not in their native state (crystallized or frozen) and the resulting structures are snapshots of the dynamic behaviour.

Overall, the determination of a dynamic process with X-ray crystallography and cryo-EM can result in a time-consuming and expensive approach and therefore, several alternative methods have been developed to gain such insights, for example hydrogen-exchange mass spectrometry (HDX-MS, (Wales and Engen, 2006)), small-angle X-ray scattering (SAXS (Tuukkanen et al., 2017)) pulsed electron electron double resonance spectroscopy (PELDOR, also known as double electron electron resonance (DEER) (Jeschke, 2018c)) and Förster resonance energy transfer (FRET, (Kapanidis et al., 2004)). These methods are often described as orthogonal and integrative methods, meaning that the results can be combined with available protein structures. Thereby, the methods are not only helpful for interpretation of dynamical processes in the protein but also useful to cross-validate protein structures. Even if the initial sample preparation often needs time-consuming optimization of the protein sample, a suitable set-up can be used with a higher throughput of functional studies than with the structure determining methods. However, it should be noted that while these methods are often helpful and rich on functional information, the determination or availability of a protein structure is most of the time a prerequisite and without a structure, the results from all orthogonal methods are often difficult to interpret.

The present work is focused on the two integrative methods PELDOR and smFRET spectroscopy (Figure 2-1). Both techniques are commonly used for functional studies of proteins and both methods are suitable to measure distances between spectroscopically active centres on a nanometer scale. If these centres are located at certain positions, a variation of experimental conditions that triggers or stabilizes different conformational states, for example addition of a substrate, can be detected as a change in the measured distance. The combination of several PELDOR or smFRET measurements at different conditions can lead to a detailed insight into the dynamic and function of a protein.

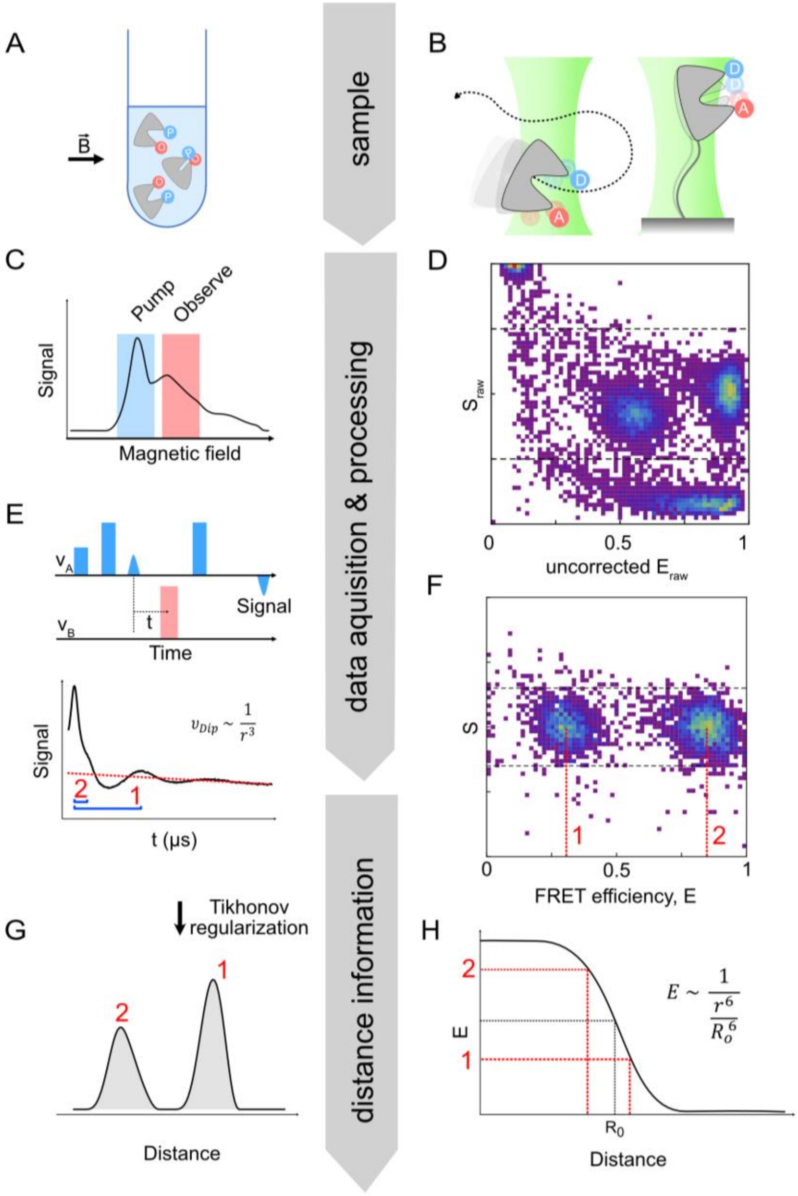


Figure 2-1: Distance measurements of proteins with PELDOR and smFRET spectroscopy. Sample preparation: A) The PELDOR experiment is performed in frozen solution in a glass tube inside a magnetic field B. O and P labels represent observer and pump spins, which are often chemically the same type of label. **B)** The smFRET experiment is performed in liquid solution on free diffusion (left) or surface-coupled proteins (right). D and A represent the donor and acceptor fluorophores of the FRET pair. **Data acquisition & processing: C & E)** Field-sweep spectrum of a double spin labelled protein sample. By using two microwave frequencies ν_A and ν_B, subpopulations of the spin centres serve as pump or observer spin. Application of a PELDOR 4-pulse sequence (top) leads to a signal that is integrated by using different times of the pulse sequence T and result in the PELDOR time trace (bottom). The oscillation frequency in the time trace contains the distance information of the sample. **D)** Typical raw ES histogram of a solution-based smFRET experiment

with low and high FRET efficiency populations. The populations in the corner and at the bottom represent donor- and acceptor-only labelling protein, respectively. **F)** Removing of the single-type labelled populations and applying correction procedures lead to the device-independent FRET efficiency distribution. **Distance information: G)** The PELDOR time trace in E) can be converted after background correction (red dashed line in E) via Tikhonov regularization to a PELDOR distance distribution. **H)** The theoretical relation between the measured FRET efficiency E and the Förster radius of the dye pair are used to calculate the experimental distance between the fluorophores. The figure is modified from (Peter et al., 2020).

Even if the result of PELDOR and smFRET experiments is quite similar at first glance, the comparison of the experimental parameters and conditions reveals some important differences. The standard smFRET experiment has a time-resolution in the sub-millisecond range and can be performed at low concentrations (single molecules) in liquid solution at room temperature and at physiological conditions (Kapanidis et al., 2004, Oikawa et al., 2018, Okamoto et al., 2020). In contrast, the standard PELDOR experiment on biomolecules is performed at 50 K in a frozen solution and for this, the samples are snap-frozen in liquid nitrogen (Glaenzer et al., 2017, Jeschke and Polyhach, 2007, Peter et al., 2019). Standardly 80 µL of 10-30 µM labelled protein are used for PELDOR spectroscopy, while a recent study revealed a significantly lower concentration of around 100 nM (Ackermann et al., 2020). For time-resolved experiments, the sample can be prepared with rapid freeze-quench (RFQ), where the samples are frozen after a defined time (Aitha et al., 2015, Schmidt et al., 2020). Additionally, PELDOR distance measurements or related EPR pulse experiments can be performed in cells but this application is nowadays not a standard experiment, especially in functionally interesting cell lines (Fleck et al., 2020, Igarashi et al., 2010, Jassoy et al., 2017).

A schematic of the general procedure for PELDOR and smFRET experiments is shown in Figure 2-1. The PELDOR experiment is the most commonly used EPR (electron paramagnetic resonance) pulse sequence for the determination of distances in biomolecules and allows to determine distances by measuring the dipolar coupling between two or more unpaired electrons (spin centres) (Jeschke and Polyhach, 2007). For other applications with other kinds of spin centres, for example metal ions, other pulsed EPR experiments, such as DQC (double-quantum coherence EPR), SIFTER (single frequency technique for refocusing dipolar couplings) or RIDME (relaxation induced dipolar modulation enhancement) have been developed (Abdullin and Schiemann, 2020, Schiemann and Prisner, 2007). The distance measurements with smFRET rely on the detection of the non-radiative energy transfer between two dipolar-coupled fluorophores (Roy et al., 2008). The measurement requires two kinds of fluorophores (donor and acceptor) with distinct, spectral properties, which allow the non-radiative energy transfer from the donor to the acceptor fluorophore (Roy et al., 2008). The detected energy transfer efficiency depends on overlapping of donor emission and acceptor absorption spectral areas, their relative orientation and the distance between the fluorophores, which can be calculated if the other parameters are known or experimentally determined (Lakowicz, 2006).

The present work focuses on an experimental comparison of the nowadays standardly used PELDOR and smFRET experiments for functional studies of proteins. It should be mentioned that many further developments and special variations for both methods exist that can be helpful for specific applications and questions. For a full description of the theoretical background for each method or developments, many detailed theoretical reviews and textbooks are available (Dimura et al., 2016, Goldfarb and Stoll, 2018, Jeschke, 2012, Lakowicz, 2006, Roy et al., 2008, Schiemann and Prisner, 2007).

2.1.1 Protein labelling for PELDOR and smFRET spectroscopy

Because the most native proteins contain neither a spin centre nor fluorophores that could be used for distance measurements, the artificial introduction of labels into proteins is necessary for their investigation with PELDOR or smFRET spectroscopy. A notable exception are proteins that natively contain paramagnetic metal ions, which can be directly detected with EPR spectroscopy as spin centre (Abdullin et al., 2015). Nevertheless, also for the majority of these proteins, at least one other spin centre for distance measurements must be introduced. One further exception is represented by dimeric or oligomeric proteins, in which repeating single spin centres in each unit can be used for PELDOR distance measurements (Ward et al., 2014). The introduction of fluorophores or spin labels is often accomplished with site-directed labelling of cysteines with maleimide or thiosulfonate ester functionalized labels (Hubbell et al., 1996, Klare and Steinhoff, 2009, Toseland, 2013) (Figure 2-2). As label positions, native cysteines at suitable locations can be used or cysteines can be artificially introduced into the protein at desired positions. If labelling of native cysteines, introduction of artificial cysteines or the removing of native cysteines to avoid random labelling are not feasible, alternative methods are available. For example, unnatural amino acids can be incorporated into the protein and either already contain a spectroscopically active centre or can be selectively labelled after incorporation (Fleissner et al., 2009, Lee et al., 2016b, Roy et al., 2008, Schmidt et al., 2014). Further, specific binders such as nanobodies or ligands can be selectively labelled and then added to the target protein to infiltrate the label (Bordenave et al., 2016, Galazzo et al., 2020, Yin et al., 2018).

For smFRET, several fluorophore labels that can be used as FRET pair with compatible spectral properties are known and commercially available, also with different linkers (Figure 2-2) (Gust et al., 2014, Roy et al., 2008). The distance measurements with PELDOR spectroscopy on proteins are most of the time performed with the standard nitroxide spin label methanethiosulfonate (MTSSL, attached to a cysteine referred as R1) (Berliner et al., 1982). This label is small and minimally invasive during attachment to a protein, as the overall size is comparable with the size of an arginine side chain (Figure 2-2). In contrast, the spectroscopical properties of MTSSL are also responsible for several limitations of the PELDOR experiment, for example that measurements have to be done at 50 K or an incompatibility with the reducing conditions that occur in living cells (Azarkh et al., 2011, Jeschke, 2012). Therefore, promising alternative spin labels such as trityl-, Cu- or Gd-labels are developed but have not yet been applied to a broad range of proteins (Cunningham et al., 2015, Fleck et al., 2020, Potapov et al., 2010).

Independently from the method or kind of label, the requirements for suitable labelling positions in proteins are similar. Firstly, the positions should have a high solvent accessibility to avoid steric hindrance during the labelling reaction. As mentioned before, the mutated or labelled position should have no impact to the function of the protein and the functional activity of the labelled proteins should be validated with additional experiments, for example functional assays. Further, the two label positions should have a distance of around 2.0 to 8.0 nm for both methods (Jeschke, 2012, Roy et al., 2008). Larger distances are also feasible, for example farFRET with fluorophore labelled samples or up to 16 nm with fully deuterated spin labelled proteins in PELDOR experiments (Krainer et al., 2015, Schmidt et al., 2016b). Further, the change of distance that occurs due to different states of the proteins should be as large as possible to precisely detect this change. For this purpose, software programs exist that help in the identification of suitable labelling positions (Beckers et al., 2015, Hagelueken et al., 2015, Jeschke, 2018b, Kalinin et al., 2012, Muschielok et al., 2008). Importantly, to reliably calculate the expected distances and distance changes, the structure of the protein or homologous proteins should be known.

After determination of distances in proteins with PELDOR or smFRET experiments, the results can be compared with the protein structure. To consider label specific characteristics such as linker length and mobility of the label, software tools are available that help to compare the experimental measured distance with the protein structure, such as MMM or mtsslWizard for PELDOR or FPS for smFRET (Hagelueken et al., 2015, Jeschke, 2018b, Kalinin et al., 2012).

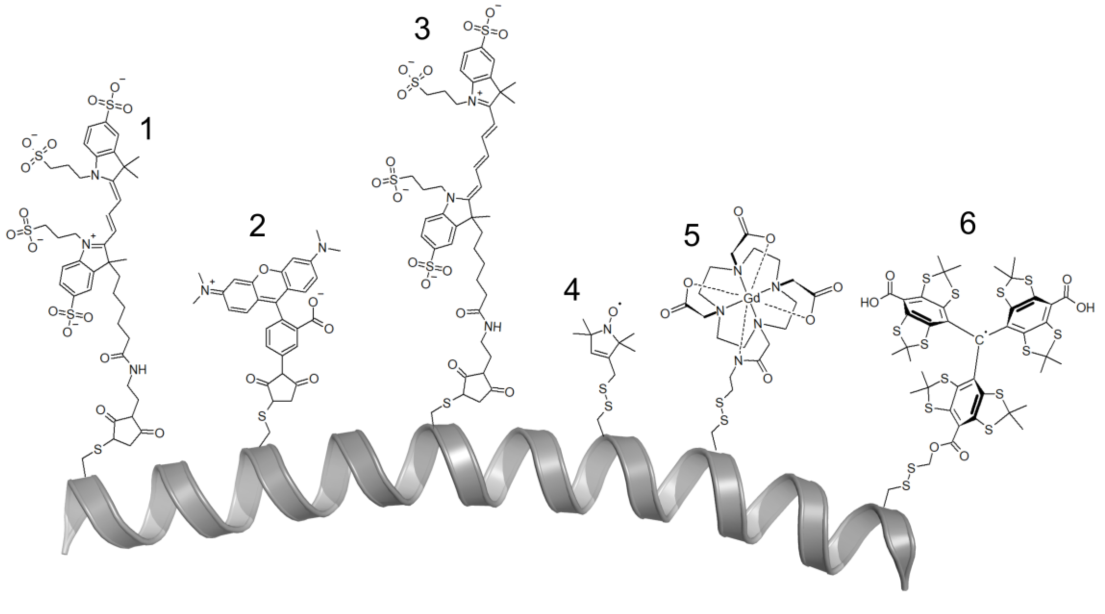


Figure 2-2: Chemical structures of exemplary fluorescence or spin labels. Maleimide-thiol adducts of the fluorophores **1)** AlexaFluor 555 (Gebhardt et al., 2020), **2)** TMR (tetramethylrhodamin-5-maleimide) (Sauer et al., 2010) and **3)** AlexaFluor 647 (Gebhardt et al., 2020). Contrary, the **4)** nitroxide MTSSL- (Berliner et al., 1982), **5)** DOTA-Gd- (Yagi et al., 2011) and **6)** trityl-spin labels (Reginsson et al., 2012) are attached to a cysteine residue via a disulfide bond. The peptide chain is represented as a grey α helix. Importantly, other labels and variable functionalized-linkers are available and partially mentioned in the main text. The figure was taken from (Peter et al., 2020).

PELDOR and smFRET spectroscopy are commonly used to measure distances in proteins on a comparable size scale but a direct experimental comparison of both methods is still missing today. In the following chapters, three proteins from the group of transporter SBPs were selected as model systems for such an experimental comparison (HiSiaP, MalE and SBD2). As described above, the proteins have the advantage that large conformational changes can be triggered easily by addition of a substrate. The same cysteine mutants of each SBP were labelled with fluorophores and spin labels and measured with smFRET or PELDOR spectroscopy with the nowadays standard procedure for protein distance measurements of both methods. The results are then compared and analysed to give an idea about the strengths and advantages, as well as possible limitations of each method.

2.2 Comparison 1: Sialic acid binding protein HiSiaP

2.2.1 Choice of label positions and *in silico* predictions

The SBP of sialic acid TRAP transporter from *H. influenzae* was used as first model system. The positions that are well suited for labelling and measurable distance changes upon a conformational change of the protein were identified with a difference distance map from the software program mtsslWizard (Hagelueken et al., 2015) (Figure 2-3 A). The positions are also partially based on previous distance measurements in our group with the homologous protein VcSiaP, in which these positions were identified with the same software tool (Glaenzer et al., 2017). The map was constructed with the two known conformations of the protein, the open, substrate unbound (PDB-ID: 2CEY (Müller et al., 2006)) and the closed, substrate bound (PDB-ID: 3B50 (Johnston et al., 2008)). Large structural changes between these conformations are visible as dark regions in the map and were used to select pairs of labelling positions with unambiguous distance changes between the conformations, indicated by coloured circles in the difference distance map (58/134, 55/175, 175/228 and 112/228, Figure 2-3 A).

After identification of labelling positions, the protein structures were firstly labelled and investigated *in silico*. This procedure should verify the accessibility of the chosen labelling positions by inspection of the protein structure and confirm the detectability of the conformational change of the introduced labels. Additionally, the *in silico* distances for each double mutant were used for a later comparison of the experimental results to the protein structures. The *in silico* labels were calculated with standard programs for both methods, mtsslWizard (Hagelueken et al., 2015) for the used nitroxide spin label MTSSL and FPS (Kalinin et al., 2012) for the used fluorophores AlexaFluor 555 and AlexaFluor 647 (geometric parameters for the fluorophores are summarized in Table 9-8). The calculated accessible volumes (AV) of both kind of labels at selected positions for both conformations of HiSiaP are shown in Figure 2-3 B. For a better overview, only the AV of fluorophore AlexaFluor 647 is presented, the second AlexaFluor 555 has a very similar AV, due to similar geometric parameters (Table 9-8). As expected from the chemical structure of MTSSL and the fluorophores (Figure 2-2), the calculated AV for the fluorophore labels are much larger than for the spin labels (Figure 2-3 B). Despite this volume difference, the *in silico* distances that were calculated with the AVs are quite similar for both types of labels, as illustrated next to the experimental data in Figure 2-5 F. A clear change of the *in silico* distances between the apo (without substrate) and holo (with substrate) conformations of HiSiaP was observed for all mutants, confirming the useful labelling positions.

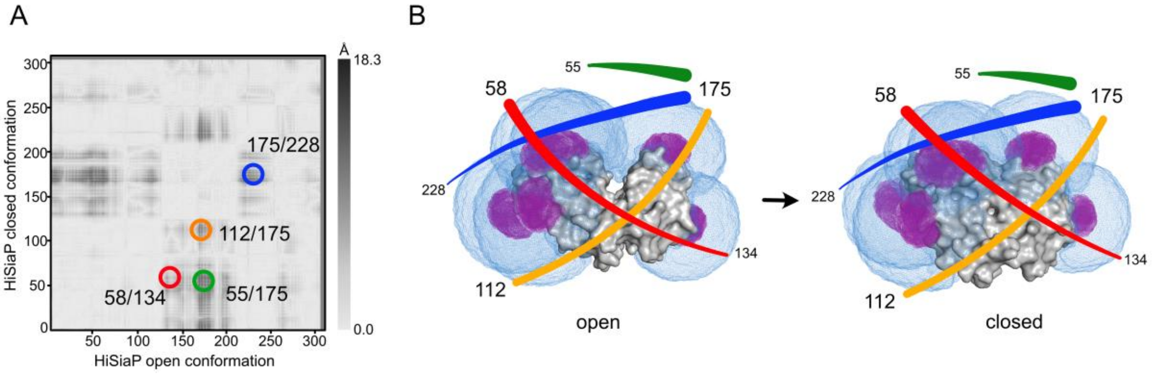


Figure 2-3: Label positions and *in silico* labelling of HiSiaP. A) Difference distance map of open (PDB-ID: 2CEY (Müller et al., 2006)) and closed (PDB-ID: 3B50 (Johnston et al., 2008)) conformation of HiSiaP for identification of regions with high conformational changes, calculated with mtsslWizard (Hagelueken et al., 2015). The positions that were used for construction of double mutants are indicated with coloured circles. **B)** HiSiaP (grey) in the open, substrate free (left) and closed, substrate bound (right) conformation from A). The AVs of spin label MTSSL (magenta) at labelling positions from A) were calculated with the mtsslWizard (Hagelueken et al., 2015). AVs of FRET label AlexaFluor 647-maleimide (blue) at the same labelling positions were calculated with FPS (Kalinin et al., 2012). The double mutants that were used for distance measurements are indicated with coloured curves, corresponding to A). The figure was modified from (Peter et al., 2020).

2.2.2 Labelling of HiSiaP double mutants

Following the identification and classification of suitable labelling positions as described above, cysteines were introduced at the desired positions to construct the four different double cysteine mutants. The mutants were expressed in M9-minimal medium to avoid unintended contaminations of the protein with their native substrate through standard LB medium (personal communication with Janin Glänzer, University of Bonn). By using the standard purification procedure for sialic acid TRAP transporter SBPs from Chapter 9.1.4.1, all mutants were produced with good yields and high purity, as illustrated by the final SEC runs and SDS-PAGEs of all mutants in Figure SI-1.

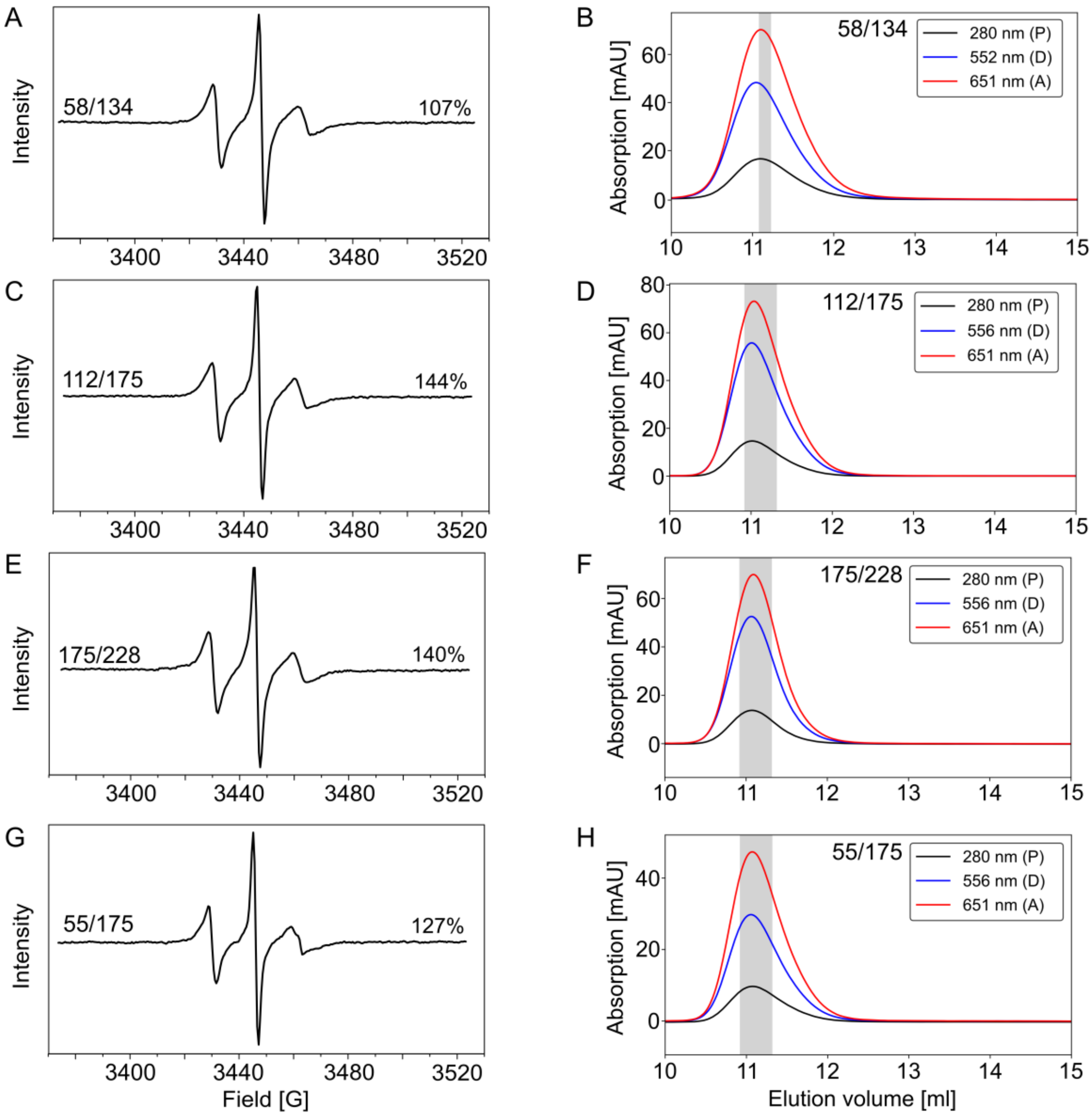


Figure 2-4: Labelling verification of HiSiaP double cysteine mutants. A) X-band RT cw-EPR spectrum of nitroxide spin labelled HiSiaP double cysteine mutant 58R1/134R1 at 25 µM protein concentration (350 scans). The labelling efficiency was determined via integration of the signal. **B)** Absorption chromatogram of fluorophore labelled HiSiaP double cysteine mutant 58C/134C (AlexaFluor 555 and AlexaFluor 647) after elution of SEC column. The absorption was detected at three different wavelengths, for the protein (black) and specific for the donor (blue) and acceptor (red) absorptions. The fraction that was used for smFRET experiments is indicated with a grey bar. Labelling and SEC were performed by C. Gebhardt. **C, E, G)** Same as A) but for other HiSiaP double mutants, as mentioned in the figure. **D, F, H)** Same as B) but for other HiSiaP double mutants, as mentioned in the figure. The figure was modified from (Peter et al., 2020).

For PELDOR experiments, all purified mutants were labelled with the nitroxide spin label MTSSL, which is currently the most commonly used EPR spin label for PELDOR distance measurements on proteins. The spin labelling was performed as described in Chapter 9.1.6.1. To confirm successful labelling, a X-band cw-EPR spectrum at room temperature (RT) was recorded for each of the HiSiaP double mutants (Figure 2-4 A, C, E and G). All spectra showed an anisotropic triplet signal that is typical for a nitroxide spin label bound to a large biomolecule (NejatyJahromy and Schubert, 2014). Additionally, the cw-EPR spectra revealed the successful separation of excess spin label from the labelling procedure, which would otherwise cause a strong overlaying second, isotropic triplet signal. The labelling efficiencies for all mutants were determined by quantification of the signal with the spectrometer software and are given next to each spectrum (Figure 2-4). The high labelling efficiencies of more than 100% are presumably caused by an imprecise protein concentration determination via UV/Vis absorption, due to a low absorption coefficient of HiSiaP through the absence of native tryptophan residues (Table 9-6).

Further, the cysteine mutants were stochastically labelled with a mixture of the commonly used FRET pair fluorophores AlexaFluor 555 and AlexaFluor 647. The excess label was separated with a SEC run and the successful labelling was confirmed simultaneously through detection of the fluorophore specific absorption of the eluted protein (around 555 nm for donor, 651 nm for acceptor) (Figure 2-4 B, D, F and H). The chromatograms for all mutants indicated absorption of both fluorophore wavelengths at the elution volume of the protein, which was itself detected at 280 nm absorption. The labelling efficiencies for all mutants were determined to be higher than 90% and a good donor to acceptor ratio of around 50:50 was observed. Overall, the labelling procedures of the four HiSiaP mutants for both distance determining methods were successfully accomplished.

2.2.3 Distance measurements with PELDOR and smFRET on HiSiaP

Each double spin labelled HiSiaP mutant was investigated with PELDOR spectroscopy in the absence or presence of the substrate sialic acid as described in Chapter 9.2.3. The raw PELDOR time traces and the processing to obtain the PELDOR distance distributions are shown in Figure 2-5 A-C for one double mutant and for the other mutants in Figure SI-2. All recorded PELDOR time traces had a high quality with good modulation depth of around

30-40%, a clear visible oscillation of the signal and a high signal to noise ratio, similar to the already published homologous protein VcSiaP (Glaenzer et al., 2017). The distance distributions were calculated after background correction with DeerAnalysis (Jeschke et al., 2006) and revealed a single, well-defined and sharp peak for all measurements (Figure 2-5 C and Figure SI-2). Due to the high quality of the detected PELDOR time trace signal, the error margin of the distance distributions was quite low.

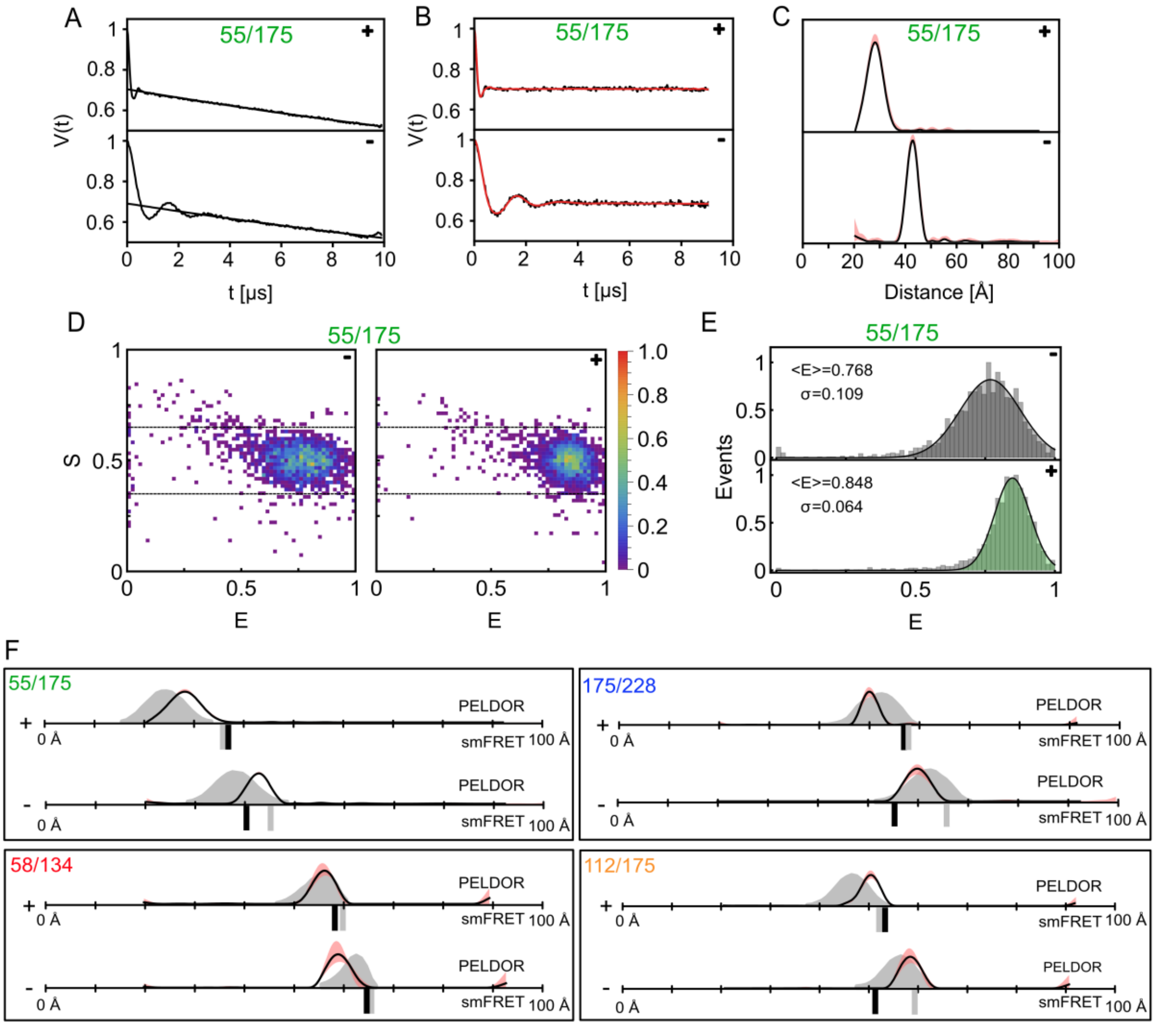


Figure 2-5: Distance measurements of HiSiaP with PELDOR and smFRET. A) Raw PELDOR time traces for spin labelled double mutant HiSiaP 55/175 at apo (-) and holo (+ = 1 mM substrate) conditions. The backgrounds that were generated with DeerAnalysis (Jeschke et al., 2006) are shown as black line. **B)** Background-corrected time traces from A) with fit (red) for distance calculation. **C)** Distance distributions from B) with validation of the distribution (red), both calculated with DeerAnalysis. **D)** ES-2D-histograms of fluorophore labelled (AlexaFluor 555 and AlexaFluor 647) double mutant HiSiaP 55/175 for apo (-) and holo (+) conditions. **E)** 1D-E-histograms from corresponding experiment in D) for apo (grey) and holo (green). The histograms were fitted with a gaussian distribution, the mean energy values and standard deviations are mentioned in the figure. **F)** Summary of PELDOR and smFRET distance measurements for several HiSiaP double

mutants for apo (-) and holo (+) conditions. The PELDOR distance distributions are shown above (simulations as grey areas (Hagelueken et al., 2015), experiments as black curves, validation in red) and the FRET distances below the distance-axis (simulation as grey bars (Kalinin et al., 2012), experiment as black bars). The raw data, as well as the processing steps towards the distances are shown in A-E) exemplary for one mutant and for all other mutant in Figure SI-2 and Figure SI-3. The smFRET experiments were performed and analysed by C. Gebhardt, the figure was modified from (Peter et al., 2020).

For each of the four double mutants, a shift of the distance peak towards a smaller distance was observed in the substrate supplemented PELDOR measurement (Figure 2-5 F). Additionally, the experimental distance distributions were in good agreement with the *in silico* determined distances from the crystal structure, considering the error of such predictions of ± 3 Å (Hagelueken et al., 2015) (Figure 2-5 F). Consequently, the experimental PELDOR distance distributions for each mutant represent the open and closed state of HiSiaP from the corresponding crystal structure and can be used to study the conformational change upon substrate binding. Importantly, no intermediate states were observed in these experiments, as well as no substrate-free closed or substrate-supplemented open conformations by using a high excess of sialic acid. These observations were in agreement with other results that are presented in Chapter 3 and findings from a previous study on the homologous protein VcSiaP (Glaenzer et al., 2017).

The smFRET measurements were performed with freely diffusing proteins and the raw data are shown in Figure 2-5 D for mutant 55/175 and for the other mutants in Figure SI-3. All data revealed good quality ES-histograms with clearly defined smFRET populations. However, some measurements revealed rather broad distributions that could result from an additional conformational flexibility of the protein or fluorophores. Another reason might be an unwanted effect, caused from the choice of labels and labelling positions. The populations were used to calculate the experimental smFRET distances, the results are illustrated in combination with the *in silico* predictions in Figure 2-5 F. Considering the different labels with varying linker lengths and sizes, an overall good correlation between the PELDOR and smFRET determined distances was observed. The smFRET measurements in the holo conditions after addition of the substrate were all in good agreement with the expectation from the simulation (Figure 2-5 F). However, all apo measurements, except mutant 58/134, indicated significant differences between the experimental and predicted distances, partially larger than the known experimental error of around ± 5 Å (Hellenkamp et al., 2018). For mutant 58/134, the experimental and predicted distances were in good agreement and a conformational change towards a closed protein state upon substrate addition was represented by this mutant. For the other three mutants, the measured apo distances were 5-10 Å lower than the predicted ones from the crystal structure and due to this, the conformational change from the open to a closed state could not be concluded with certainty (Figure 2-5 F). For the two mutants 112/175 and 175/228, the measured distances from the apo experiments were even smaller than for the holo experiments. The reasons for these deviations of the smFRET experiments from their predictions and the difference to the PELDOR experiments that fit nicely to the crystal structures are examined in detail in the following chapters.

2.2.4 Investigations of unexpected distances in smFRET experiments

Since previous studies strongly indicated a substrate-bound closed state for HiSiaP under holo conditions, a totally new structural feature for these three smFRET mutants is unlikely to be responsible for the discrepancies (Fischer et al., 2015, Johnston et al., 2008, Müller et al., 2006). However, from a structural perspective, mutational modifications or attachments of labels to a protein can indeed trigger unexpected functional behaviours. To investigate if the unexpected distances for the apo measurements are caused by the labelled mutants itself, for example a stabilization of the closed conformation or an uncontrolled fluctuation between both states, a burst-variance analysis was performed (Figure 2-6). As for all smFRET measurements, this experiment was performed by Christian Gebhardt from the group of Thorben Cordes (University of Munich). The result indicated that in the absence of Neu5Ac, the three fluorophore labelled mutants 58/134, 175/228 and 55/175 all exist only in a single conformation and no switching between two conformations was detected. Importantly, the experiments were limited to a millisecond timescale and motions faster than 500 µs cannot be detected with this method.

Since none of the results pointed towards a disturbed structure or function, an unintended interaction of the fluorophores with the proteins surface was investigated. Such interactions can strongly immobilize the label at a certain position and the measured experimental distance would then not correlate to the *in silico* predictions, because the AV calculation is based on a freely moving label (Kalinin et al., 2012). Both used fluorophore labels have sulfonic acid groups which could in principle easily interact with positively charged areas on the surface of the protein (Figure 2-2). The interaction will depend on the type of label and the labelling position, which allow the label to reach certain protein regions. Because smFRET uses two different kind of dyes, it is possible that only one dye is immobilized at one position, while the other dye at the same position cannot interact with the protein and is not immobilized. The stochastic labelling procedure will therefore causes two smFRET populations or a broadening of the smFRET population, as observed in the histograms for all three HiSiaP apo measurements with deviations from the simulation (Figure 2-5 D and E, Figure SI-3).

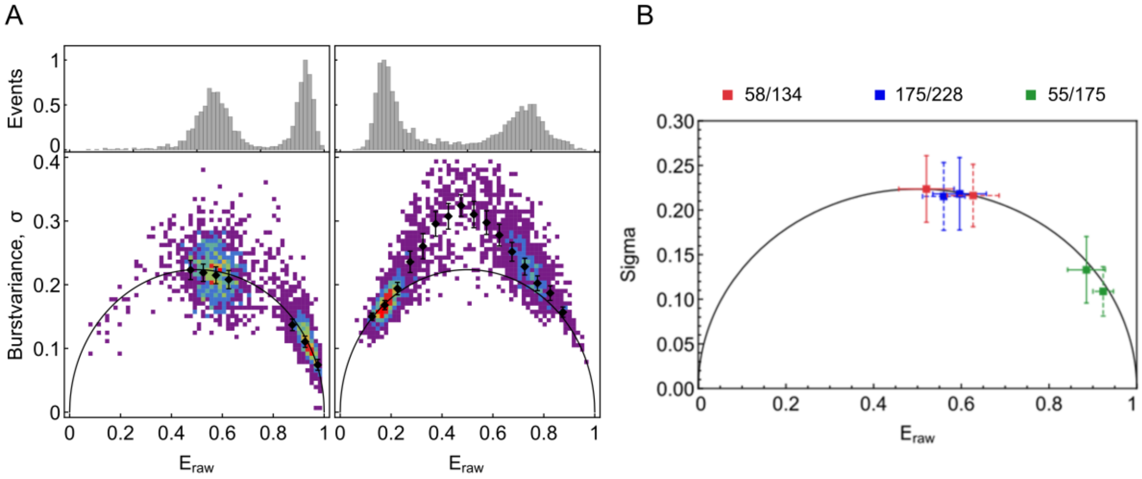


Figure 2-6: Burst-variance analysis of HiSiaP data. A) Example burst-variance analysis of mutants 55/175 and 175/228 in their holo states (left) and control experiment with fluctuating DNA-hairpin (right) from (Tsukanov et al., 2013). Standard errors from the recorded data are shown in black. **B)** Population mean and standard deviation of the burst variance analysis for three HiSiaP mutants in apo (solid) and holo conditions (dashed) are shown. The analysis was performed by C. Gebhardt, the figure was taken from (Peter et al., 2020).

To prove that immobilized fluorophores are the reason for the unintended distances in the apo measurements, the labels and labelling positions were investigated with determination of the fluorescence anisotropy and lifetime. For an unambiguous result, single cysteine mutants with only one labelling position were chosen. Because all mutants with deviations contain the label site at residue 175, this position was postulated to be responsible or involved in the unintended distances and was chosen as one single cysteine mutant. In contrast, the double mutant HiSiaP 58/134 was the only one for which both conditions, apo and holo, correlated well with the corresponding simulations and PELDOR results. The single cysteine mutant HiSiaP 58 was therefore chosen and should represent a control measurement with low or even no immobilization of the spin label. The two single cysteine mutants 58C and 175C were constructed, expressed and purified and were received in high yields with good purity, similar to the double mutants before (Figure SI-4). In contrast to the smFRET measurements above, the single mutants were labelled separately with AlexaFluor 555 and AlexaFluor 647. The anisotropy decay was subsequently analysed for these mutants with both fluorophores in the apo and holo conditions (Figure 2-7 A and B). For the label position 58, no change of the anisotropy between apo and holo condition was observed for both fluorophores. However, the fluorophores at position 175 showed a slower decay of the anisotropy for the apo condition compared to the holo measurements (Figure 2-7 A and B). Similar observations were detected for lifetime measurements of both mutants (Figure 2-7 C and D). Position 58 revealed an equal result for apo and holo conditions, while the lifetime for labels at position 175, especially for fluorophore AlexaFluor 555, was increased at apo condition compared to the holo condition.

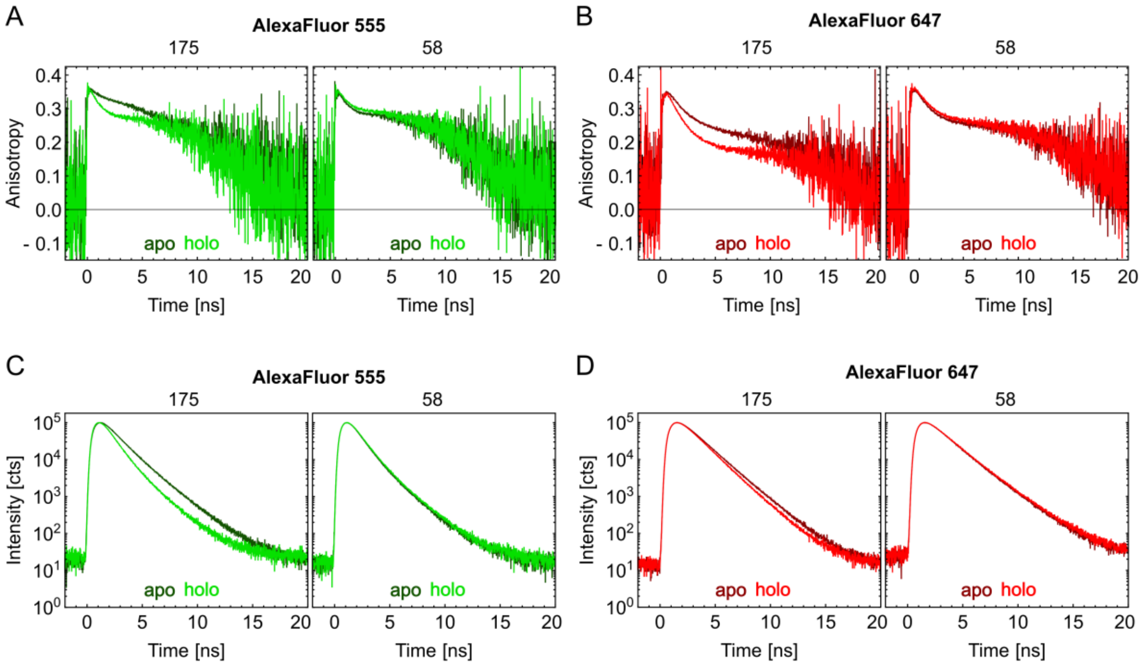


Figure 2-7: Fluorescence anisotropy and lifetime measurements on HiSiaP single labelled mutants. A) Anisotropy decay curves of HiSiaP single labelled mutants 175 (left) and 58 (right) labelled with AlexaFluor 555 for apo (dark green) and holo conditions (bright green). **B)** Same as A) but with fluorophore AlexaFluor 647 for apo (dark red) and holo conditions (bright red). **C)** Lifetime decay curves for same mutants, labelled with AlexaFluor 555 and same conditions and mutants as in A). **D)** Same as C) but with fluorophore AlexaFluor 647. The experiments were performed by C. Gebhardt, the figure was modified from (Peter et al., 2020).

The different lifetime increases for the two fluorophores at position 175 explained the observed broadening of the FRET populations in the double mutant histograms. Due to a stochastic labelling, half of the proteins were labelled with acceptor at position 175 and donor at the other positions, while the second half was labelled the other way around (by just considering the proteins with two different kind of labels). The changed lifetime of donor AlexaFluor 555 at position 175 will therefore cause another FRET state compared to proteins in which AlexaFluor 647 is attached to this position. Provided that the other labelling positions have no impact onto the lifetime that will exactly counteract the effect on position 175, two distinct FRET states are detected which cause a broadening of the FRET populations. Further, a change in lifetime, orientation and/or fluorophore disposition have an experimental influence on the Förster radius and an impact on the FRET-efficiency, which are both used to determine the fluorophore distance (Chapter 9.2.4). Taken together, the results suggested that the unintended FRET distances in the apo measurements of three HiSiaP double mutants with a not detectable closing mechanism and high experiment-simulation discrepancy were caused by a fluorophore-protein interaction of the AlexaFluor dyes at position 175.

2.2.5 Use of an alternative FRET pair

To avoid such disrupting protein-fluorophore interactions and by this, further prove the unintended interaction between the AlexaFluor dyes and HiSiaP protein, alternative label positions or other FRET pairs can be used. Alternative labelling positions require the production of a new mutant, whereas another FRET pair can directly be used for labelling of available double mutants. Therefore, the FRET pair with TMR and Cy5 dye was tested, which have a completely different molecular structure and are not negatively charged as the AlexaFluor dyes (Figure 2-2) (Gust et al., 2014, Sauer et al., 2010). To confirm a changed behaviour of these alternative dyes, the fluorophores were attached to the single cysteine mutant HiSiaP 175 and the anisotropy decay and lifetime were analysed (Figure 2-8). Indeed, contrary to the AlexaFluor dyes at this position, the TMR and Cy5 fluorophores indicated no significant change of the anisotropy decay between apo and holo conditions and thereby revealed an almost ideal behaviour with high rotational freedom (Figure 2-8 A).

Further, the detected lifetimes for TMR are also unchanged for both conditions, while the Cy5 dye revealed a small increase for the apo sample (Figure 2-8 B).

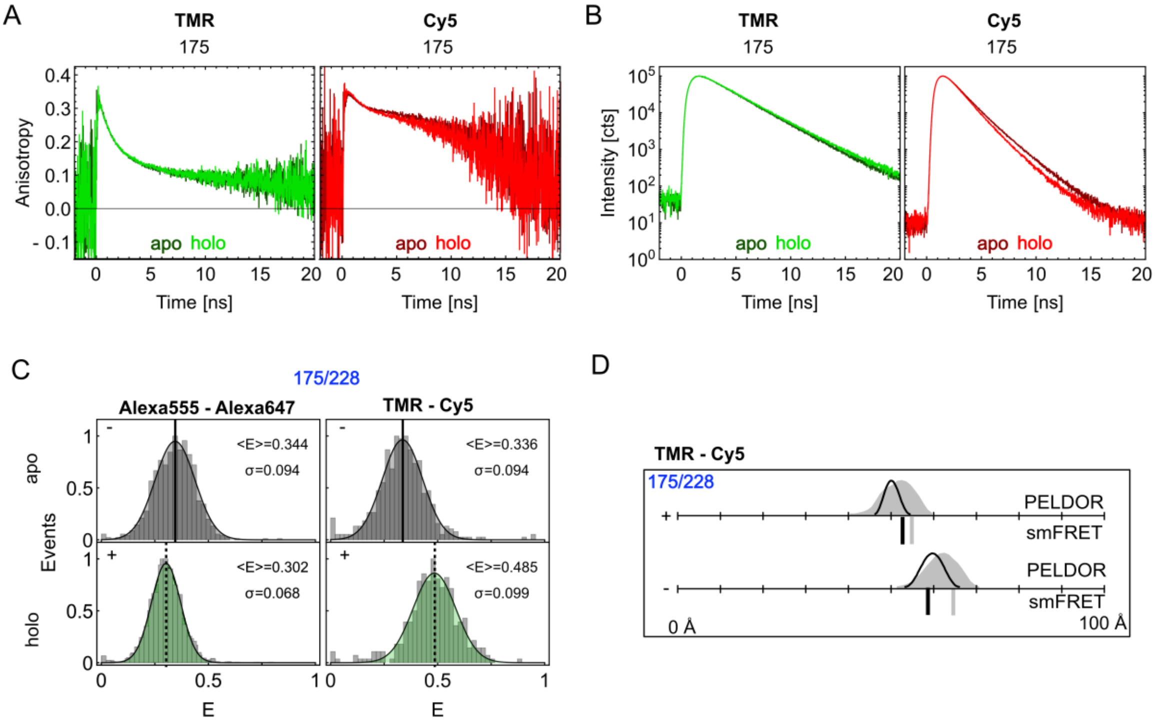


Figure 2-8: Alternative FRET pair characterization for HiSiaP mutants. A) Anisotropy decay curves of HiSiaP single mutant 175 labelled with TMR (donor, left) and Cy5 (acceptor, right) for apo (dark colours) and holo conditions (bright colours). **B)** Lifetime decay curves for same mutant, same fluorophores and conditions as in A). **C)** 1D-E histograms of HiSiaP double mutant 175/228 labelled with FRET pairs AlexaFluor 555 – AlexaFluor 647 (left) and TMR-Cy5 (right). For both label pairs, the apo (grey, above) and holo (green, below) measurements are illustrated and the mean FRET efficiencies and standard deviations of the gaussian fits are mentioned. **D)** Comparison of distances for HiSiaP 175/228 from smFRET experiments with FRET pair TMR-Cy5 to simulations and PELDOR spectroscopy results, similar to Figure 2-5 F. The experiments were performed by C. Gebhardt, the figure was modified from (Peter et al., 2020).

After this initial characterization, the smFRET distance measurements on HiSiaP double mutant 175/228 were repeated using the alternative FRET pair. In contrast to the AlexaFluor dyes, the 1D-E-histograms for the alternative FRET pair revealed a clear shift towards a higher FRET energy upon addition of the substrate (Figure 2-8 C). The resulting FRET distances are illustrated in Figure 2-8 D in combination with the simulations and the PELDOR spectroscopy distances. The smFRET distances with the TMR-Cy5 dyes for the apo and holo conditions were in agreement with the predicted distances, considering the known experimental error of around ± 5 Å (Hellenkamp et al., 2018). The conformational change from an open to a closed conformation was represented through the measurement with this FRET pair, contrary for the comparable measurements with AlexaFluor dyes. In summary, after optimization of the smFRET conditions and presumable prevention of fluorophore-protein interactions, an overall good agreement between the crystal structure and the experimental PELDOR and smFRET distances were observed for HiSiaP.

2.3 Comparison 2: Maltose binding protein MalE

2.3.1 Choice of label positions and *in silico* labelling

The SBP MalE from ABC transporter MalE-FKG$_2$ was used as further system for the comparison of PELDOR and smFRET. Previously, this protein was analysed with smFRET to investigate the conformational change and substrate binding mechanism of this protein (de Boer et al., 2019b). For the present work, four double cysteine mutants of MalE were used and distances between the label positions and the conformational changes upon substrate binding were analysed (Figure 2-9 A and B).

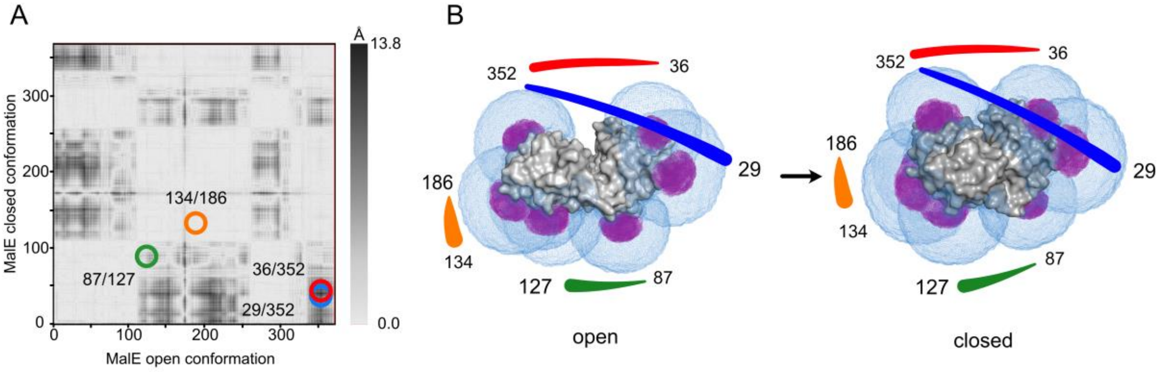


Figure 2-9: Label positions and *in silico* labelling of MalE. A) Difference distance map of open (PDB-ID: 1OMP (Sharff et al., 1992)) and closed (PDB-ID: 1ANF (Quiocho et al., 1997)) conformation of MalE to identify regions with high conformational changes, calculated with mtsslWizard (Hagelueken et al., 2015). The positions that were used for construction of double mutants are indicated. **B)** Representation of MalE (grey) in the open, substrate free (left) and closed, substrate bound (right) conformation that were used in A). The AVs of spin label MTSSL at seven different labelling positions were calculated with the mtsslWizard (magenta) (Hagelueken et al., 2015). AVs of FRET label AlexaFluor 647-maleimide were calculated with FPS (Kalinin et al., 2012) for the identical positions (blue). The double mutants that were used for distance measurements are indicated with coloured curves, corresponding to A). The figure was modified from (Peter et al., 2020).

The cysteines of the two mutants 29/352 and 36/352 are located at the entrance site of the binding cleft. Hence, a decrease of the distance between the two corresponding labels in the closed conformation of the protein was predicted (Figure 2-9). In contrast, the cysteines of double mutant 87/127 are located on the back of the binding cleft. This should lead to an increased distance during the transition from the open to the closed state (Figure 2-9). For the fourth mutant, the two cysteines are located in one rigid body of the protein without a conformational change upon substrate binding. This mutant was used as a negative control (Figure 2-9). All label positions were again initially analysed via *in silico* labelling as described for HiSiaP with both types of labels, the spin label MTSSL and the fluorophore AlexaFluor 647. The AV calculations are shown in Figure 2-9 B on the surface of the protein and indicated a similar result as for HiSiaP with large, homogenous AV predictions for the fluorophore and smaller AVs for the spin label. Additionally, the simulated distances of the double mutants indicated the changes upon substrate addition that were expected from the calculated distance difference map (Figure 2-11 F).

2.3.2 Labelling of MalE double mutants

The MalE double cysteine mutants were provided by the group of Thorben Cordes. For preparation of the mutants for smFRET and PELDOR spectroscopy, the purified proteins were labelled with the same labels and labelling procedures as described for HiSiaP (Chapter 9.1.6).

The successful spin labelling of the mutants with MTSSL was confirmed with cw-EPR spectroscopy. For each sample, an anisotropic triplet signal was obtained, caused by a low rotation correlation time through successful binding of the label to a large molecule (Figure 2-10). Low intensity features of an isotropic triplet signal were detected in the sample with mutant 87/127 (Figure 2-10 A). This observation indicated a second species of spin label with another rotation correlation time, which is presumably represented by unsuccessfully separated, unbound label. This result might also explain the high labelling efficiency of 115% for this mutant. Due to the same labelling procedures for all mutants and the only small amount of free label, the sample was used for PELDOR measurements without further optimization. The other three mutants revealed high labelling efficiencies of around 79-94% (Figure 2-10 C, E, G). For smFRET experiments, the fluorophore labelled proteins were separated from excess dye and detected with the specific wavelength with a SEC run. The elution profiles of all four mutants indicated the successful binding of the fluorophores and sufficient efficiencies of more than 70% (Figure 2-10).

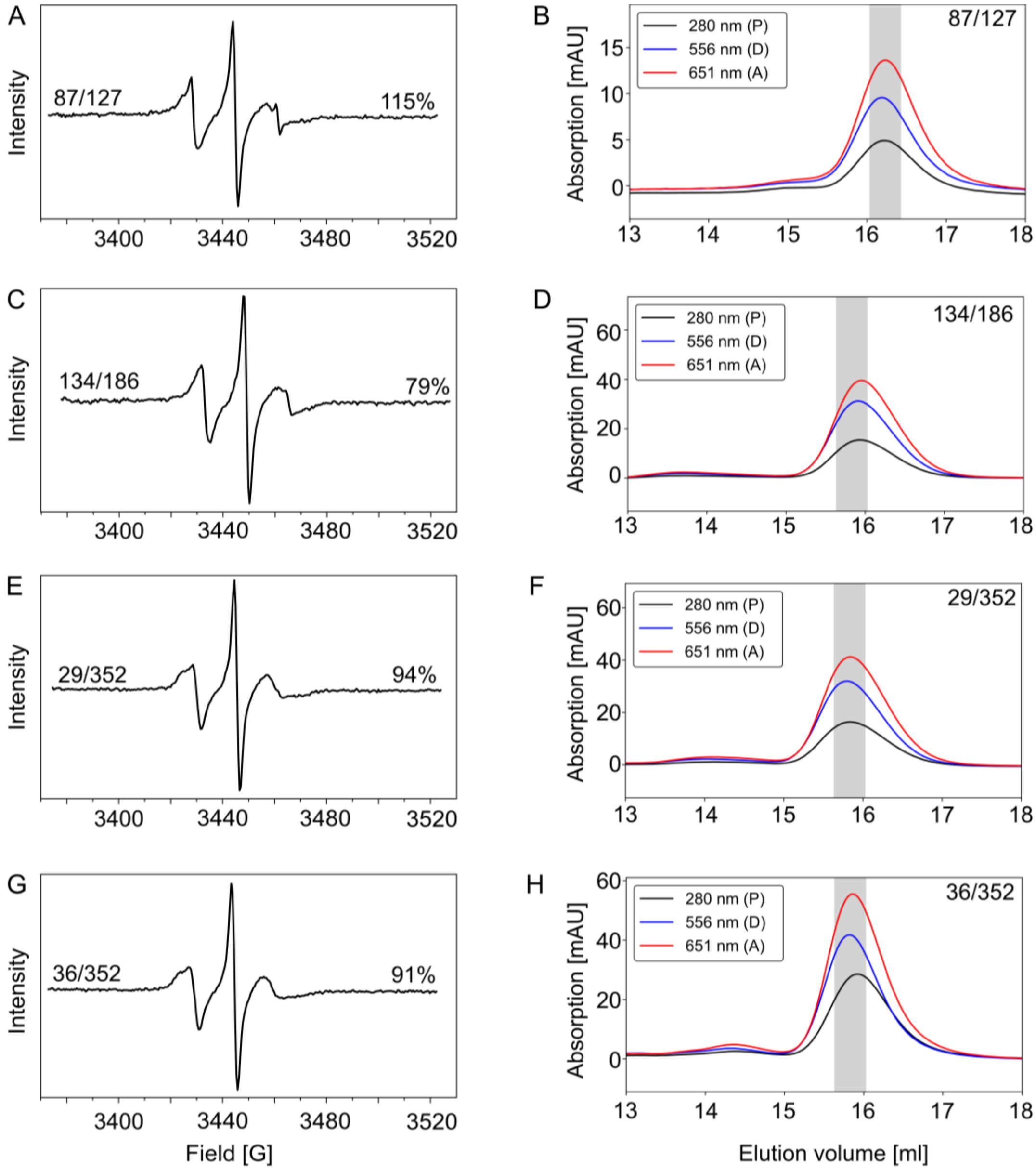


Figure 2-10: Labelling verification of MalE double cysteine mutants. A) X-band RT cw-EPR spectrum of nitroxide spin labelled MalE double cysteine mutant 87R1/127R1 at 25 µM protein concentration (350 scans). The labelling efficiency that was determined via integration of the signal is mentioned. **B)** Absorption chromatogram of fluorophore labelled HiSiaP double cysteine mutant 87C/127C (AlexaFluor 555 as donor and AlexaFluor 647 as acceptor) after elution of SEC column. The absorption was detected at three different wavelengths, for the protein (black) and specific for the donor (blue) and acceptor (red) absorptions. The fraction that was used for smFRET experiments is indicated with a grey bar. Labelling and SEC were performed by C. Gebhardt. **C, E, G)** Same as A) but for other MalE double mutants. **D, F, H)** Same as B) but for other MalE double mutants. The figure was modified from (Peter et al., 2020).

2.3.3 Distance measurements with PELDOR and smFRET on MalE

The distance measurements of the labelled MalE proteins were performed with the standard set-ups, analogous to the HiSiaP comparison. For the three MalE mutants 29/352, 36/352 and 134/186, PELDOR time traces with a good modulation depth of around 20-30% and a visible oscillation were obtained (Figure 2-11 A and Figure SI-5). Without an obvious reason, the quality for the time trace of the fourth mutant 87/127 was noticeably lower than for the other mutants (Figure SI-5). The raw data and processing steps for calculation of the distance distribution are illustrated in Figure 2-11 A-C for mutant 36/352 and in Figure SI-5 for the other mutants. The smFRET experiments for fluorophore labelled MalE mutants revealed good quality populations in the 2D-ES-histograms and a single, gaussian-shaped distribution for the FRET energy in the 1D-E-histogram (Figure 2-11 D and E and Figure SI-6). The experimental distances obtained from both methods are shown in Figure 2-11 F for the apo and holo condition of each mutant and in combination with the simulated data.

For the apo conditions, all experimental PELDOR and smFRET distances closely matched the simulated predictions, except double mutant 29/352 (Figure 2-11 F). While for the PELDOR experiment the measured distance was still in agreement with the prediction, the experimental smFRET distance was around 7 Å smaller than the prediction and by this, not in the range of the experimental error of FRET with around ±5 Å (Hellenkamp et al., 2018). Since both methods showed a tendency of shorter experimental distances, the MalE protein structure at position 29 was inspected, as position 352 was also present in mutant 36/352 which closely fits to the prediction (PDB-ID: 1OMP (Sharff et al., 1992)). The residue is located at a terminal position of a helix at the proteins surface and the analysis of the crystal packing showed a direct contact of this region with the C-terminal helix of a neighbouring protein. This observation suggested a possible local rearrangement of the protein chain in the crystal structure, which possibly explains the deviation of the distance measurements with both methods in a solution state of the protein. To experimentally confirm this hypothesis, further double mutants with new labelling positions next to position 29 might be helpful.

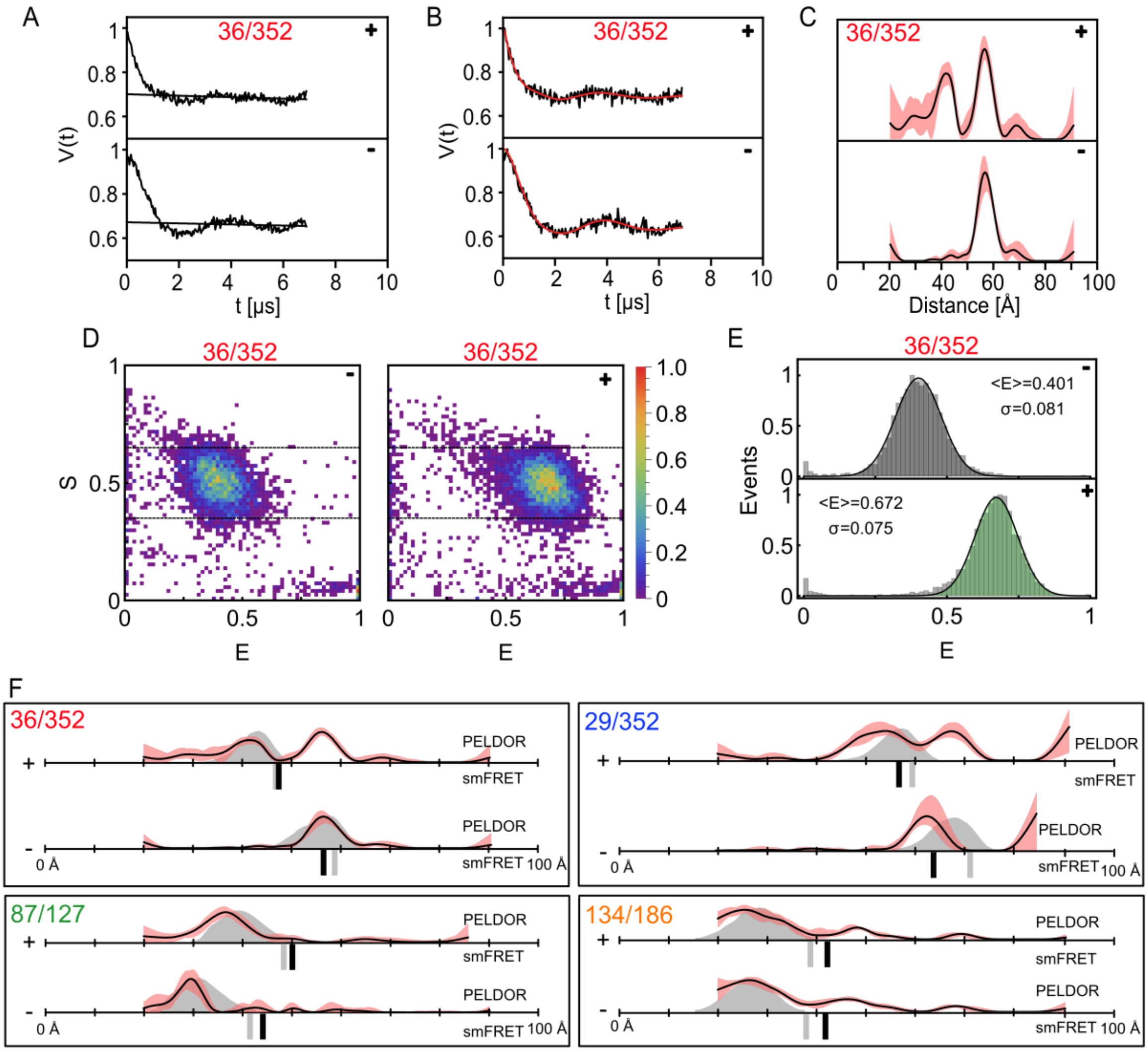


Figure 2-11: Distance measurements of MalE with PELDOR and smFRET. A) Raw PELDOR time trace for spin labelled double mutant MalE 36/352 at apo (-) and holo (+ = 1 mM substrate) conditions. The backgrounds that were generated with DeerAnalysis (Jeschke et al., 2006) are shown as black line. **B)** Background-corrected time traces from A) with fit (red) for distance calculation. **C)** Distance distributions from B) with validation (red), both calculated with DeerAnalysis. **D)** ES-2D-histograms of fluorophore labelled (AlexaFluor 555 and AlexaFluor 647) double mutant MalE 36/352 for apo (-) and holo (+) conditions. **E)** 1D-E-histograms from corresponding experiment in D) for apo (grey) and holo (green). The histograms were fitted with a gaussian distribution, the mean energy values and standard deviations are mentioned in the figure. **F)** Summary of PELDOR and smFRET distance measurements for several MalE double mutants for apo (-) and holo (+) conditions. The PELDOR distance distributions are shown above (simulations as grey areas (Hagelueken et al., 2015), experiments as black curves, validation in red) and the FRET distances below the distance-axis (simulation as grey bars (Kalinin et al., 2012), experiment with black bars). The raw data, as well as the processing steps towards the distances are shown in A-E) exemplary for one mutant and for all other mutant in Figure SI-5 and Figure SI-6. The smFRET experiments were performed and analysed by C. Gebhardt, the figure was modified from (Peter et al., 2020).

The addition of the substrate maltose in high excess showed no effect for the control mutant 134/186, as the PELDOR and smFRET distances were nearly unchanged (Figure 2-11 F).

Further, the mutant 87/127 on the back of the protein indicated a conformational change towards the closed state by an expected increase of the distance (Figure 2-11 F). The experimental distances of these two mutants for both methods were in good agreement with the predicted distances. The other two double mutants 29/352 and 36/352 after addition of the substrate revealed smFRET distances that correlated well with the simulated data while the PELDOR data showed unexpected distance distributions (Figure 2-11 F). For both mutants, a clear double peak was observed in the distance distribution. The double peak was already manifested in a slight multifrequency oscillation of the original PELDOR time trace (Figure 2-11 A and Figure SI-6).

The simulated distances allowed the assignment of the two experimental peaks to the open and closed protein conformation and thereby indicated a mixture of the protein states in the sample, despite a large excess of substrate with around 66 times higher molar concentration. The deviation for mutant 29/352 that was described for the apo measurements was still detectable for the holo measurements, albeit the difference was smaller. Contrary to the two double mutants 29/352 and 36/352, the mixture of the two conformational states was not represented in the distance distribution of mutant 87/127. For this mutant, only a broader distance distribution at holo conditions was observed compared to the apo experiments. Due to the lower quality of the PELDOR time traces for this mutant, it is possible that the mixture of conformational states was just observed as single, broader peak and not as two distinctly separated distance peaks.

2.3.4 Investigations of unexpected distances in PELDOR experiments

One important difference between standard PELDOR and smFRET spectroscopy measurements is the experimental temperature which can influence a binding process, as previously observed for the OppA ABC transporter SBP (Norcross et al., 2019). In PELDOR spectroscopy, the samples are snap-frozen after preparation and incubation with the substrate and measured for spectrophysical reasons at 50 K, whilst smFRET is performed in liquid solution at room temperature. To analyse if the observed partial closure of the protein is derived from a changed binding affinity between substrate and protein, the PELDOR experiments for the two mutants 29/352 and 36/352 were repeated with a 10 times higher amount of substrate (10 mM instead of 1mM maltose). The PELDOR experiments revealed similar distance distributions for both substrate concentrations and still revealed a mixture of both protein conformations (Figure 2-12 A). Although the closed state for mutant 36/352 was minimally increased, the mixture of both states still dominated. Additional to these measurements, a further analysis of the direct temperature influence onto the measurement is hard to achieve. PELDOR experiments in the used set-up with nitroxide spin labels cannot be performed at higher temperatures and contrary, smFRET measurements in frozen solution are also not standard procedures. Alternative spin labels such as trityls give the option for pulsed-EPR experiments at room temperature but are currently under development and not used in standard set-ups (Fleck et al., 2020).

Another experimental condition that differs between both methods is the addition of high amount of cryo-protectant to the PELDOR sample, prior to snap-freezing. Standardly the spin labelled proteins were measured in 50% *d*-ethylene-glycol (deuterated) to receive an amorphic, glass-like state of the frozen solution. Additionally, the deuterated cryo-protectants lead to an increase of the T_2 relaxation times and thereby allow the determination of longer PELDOR distances (Georgieva et al., 2012, Jeschke, 2012, Schmidt et al., 2020). To find out if the addition of cryo-protectant influenced the substrate binding or closure transition in the presented experiments, the PELDOR experiments of the double mutants 29/352 and 36/352 were repeated with differing amounts and another kind of cryo-protectant.

First, a sample of mutant 29/352 was prepared with a lower amount of 25% *d*-ethylene-glycol (Figure 2-12 B). As visible from the distance distribution, the mixture of the two states was unchanged, the single peaks for each distance are just less resolved due to the lower quality of the PELDOR time trace (Figure SI-7). In a second measurement, the alternative cryo-protectant *d*-glycerol (25%) was used and the distance distribution of this measurement indeed revealed a varied mixture of the two conformational states (Figure 2-12 B). The amount of the closed state was clearly increased compared to the measurements with the other cryo-protectant, but still a representative distance signal for the open conformation was detectable (estimated to be around 30%). As consequence, the PELDOR measurements for this mutant were repeated without any addition of a cryo-protectant. The intensity of the echo signal for this measurement was very low and the time scale of the PELDOR time trace had therefore been shortened, whereby the long distances are more error-prone than for the other measurements (Figure SI-7) (Jeschke, 2012, Glaenzer et al., 2018). However, an oscillation of the signal was detected on a short time scale and the calculated distance distribution for this measurement indicated only a monomodal distance peak. The same effect was obtained for the second mutant 36/352, the removal of cryo-protectant led to a change of the PELDOR distribution to a single peak (Figure 2-12 B and Figure SI-7). Noticeably, the experimental holo PELDOR distances for both mutants without cryo-protectant were smaller than the simulated distances, especially mutant 29/352 indicated a deviation between the experimental and predicted distance.

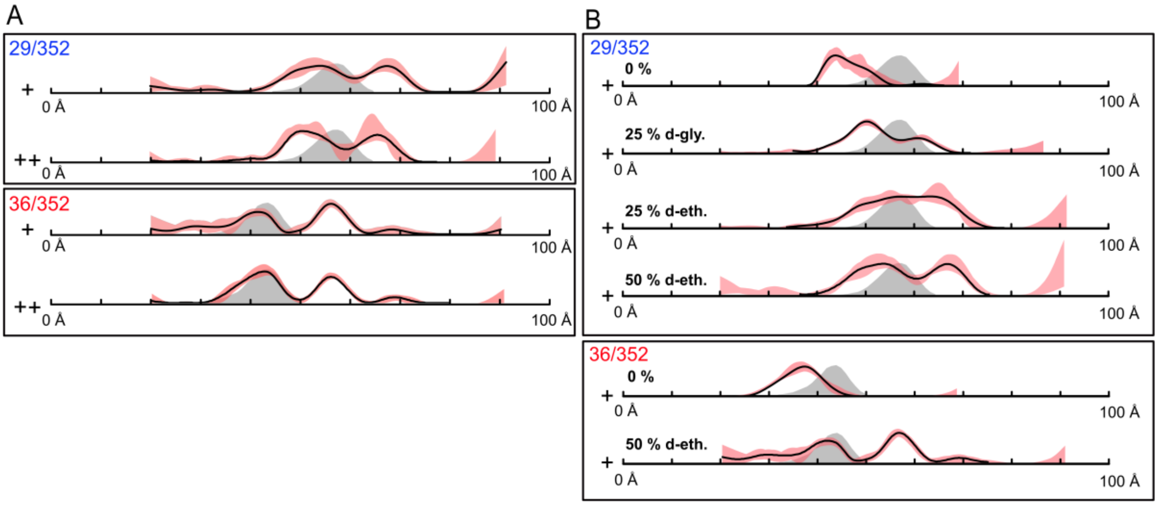


Figure 2-12: Variations in sample preparation of MalE for PELDOR. A) PELDOR distance distributions for the two spin labelled MalE mutants 29/352 and 36/352 with standard substrate concentration (+) and a 10 times higher amount (++) (simulations as grey areas (Hagelueken et al., 2015), experiments as black curves, validation in red). Data were analysed with DeerAnalysis (Jeschke et al., 2006), the raw PELDOR data are shown in Figure SI-5. **B)** PELDOR distance measurements for same mutants as in A) with variations in the amount and kind of supplemented cryo-protectant as mentioned in the figure. The raw data, as well as the processing steps are shown in Figure SI-7. The figure was modified from (Peter et al., 2020).

The experiments above suggested a strong dependence of the conformational change from MalE on the use of cryo-protectant. Through reducing or variation of the protectant, the PELDOR distance distributions revealed different results. Hence, as the smFRET distances were in good agreement with the expectations and simulations, this method is more suitable for functional studies of MalE by using standard conditions. Nevertheless, a further optimization or variation of the PELDOR experimental conditions can also presumably lead to a full functional description. Therefore, general binding and interaction studies between MalE and the cryo-protectants will be useful and cryo-protectants without interactions should be used.

2.4 Comparison 3: Glutamate/Glutamine binding protein SBD2

2.4.1 Choice of label positions and *in silico* labelling

As third model system, the substrate binding domain 2 (SBD2) of the ABC transporter GlnPQ was used, previously studied with smFRET spectroscopy (Gouridis et al., 2015). The regions with high conformational changes between the open and closed conformation were confirmed by generating a difference distance matrix with the mtsslWizard software (Hagelueken et al., 2015) (Figure 2-13 A). Two double mutants were used, which are located on the opposite side of the protein and revealed a high change of the distance between the cysteine positions upon substrate binding (319/392 and 369/451). The labels, as well as the distances between the labels were again determined *in silico*. The AVs for the spin label MTSSL and fluorophore AlexaFluor 647 are shown in Figure 2-13 B on the surface of the two conformational states, indicating similar AVs for both types of labels as before. The *in silico* distance predicitions of the two double mutants revealed the expected behaviour of distance changes upon conformational change of the protein (Figure 2-15 F).

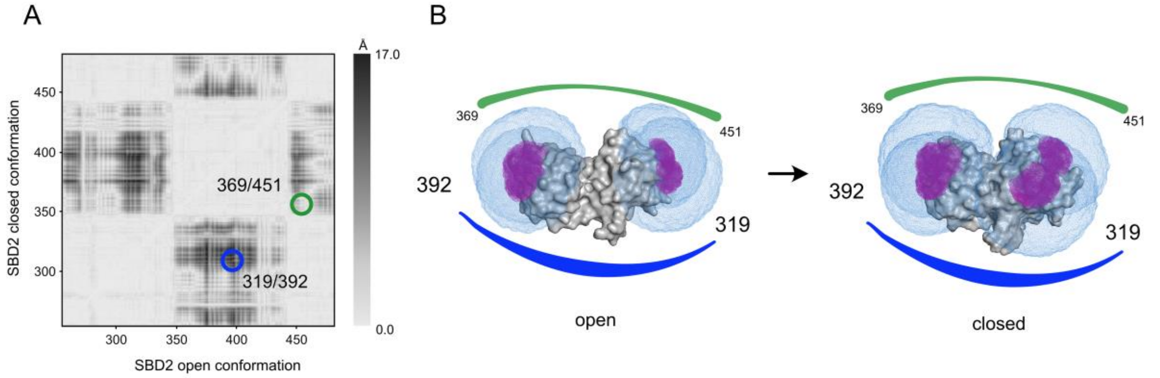


Figure 2-13: Label positions and *in silico* labelling of SBD2. A) Difference distance map of open (PDB-ID: 4KR5 (Fulyani et al., 2013)) and closed (PDB-ID: 4KQP (Fulyani et al., 2013)) conformation of SBD2 for identification of regions with structural changes, calculated with mtsslWizard (Hagelueken et al., 2015). The positions that were used for construction of double mutants are indicated. **B)** Representation of SBD2 (grey) in the open, substrate free (left) and closed, substrate bound (right) conformation that were used in A). The AVs of spin label MTSSL at four different labelling positions are calculated with mtsslWizard (magenta) (Hagelueken et al., 2015). The AVs of FRET label AlexaFluor 647-maleimide were calculated with FPS (blue) (Kalinin et al., 2012) for the identical positions. The double mutants that were used for distance measurements are indicated with coloured curves, corresponding to A). The figure was modified from (Peter et al., 2020).

2.4.2 Labelling of SBD2 double mutants

Following the procedure from the other two model systems, the two SBD2 mutants were constructed, expressed and purified by the laboratory of Thorben Cordes (University of Munich). The mutants were subsequently labelled with either the spin label MTSSL or a mixture of the two fluorophores AlexaFluor 555 and AlexaFluor 647 by using the standard labelling procedures as described in Chapter 9.1.6. The spin labelling was again confirmed by recording of cw-EPR spectra. For both mutants, an anisotropic triplet signal was measured, indicating the successful attachment of the label to the protein SBD2 (Figure 2-14 A and C). Interestingly, the cw-EPR spectrum for mutant 319/392 revealed at least two spin label species, one with a lower and one with a higher mobility. This second species might be caused by an unsuccessful separation of the excess label that is still present unbound in the sample or an attached label in a very flexible region of the protein. The high quantity of determined labelling efficiency of around 111% suggested an unsuccessful separation. However, as before for one MalE mutant, the sample was used for the distance measurement to ensure a high level of comparability for all experiments. The spin labelling efficiency of the other mutant was determined to a value of 91%.

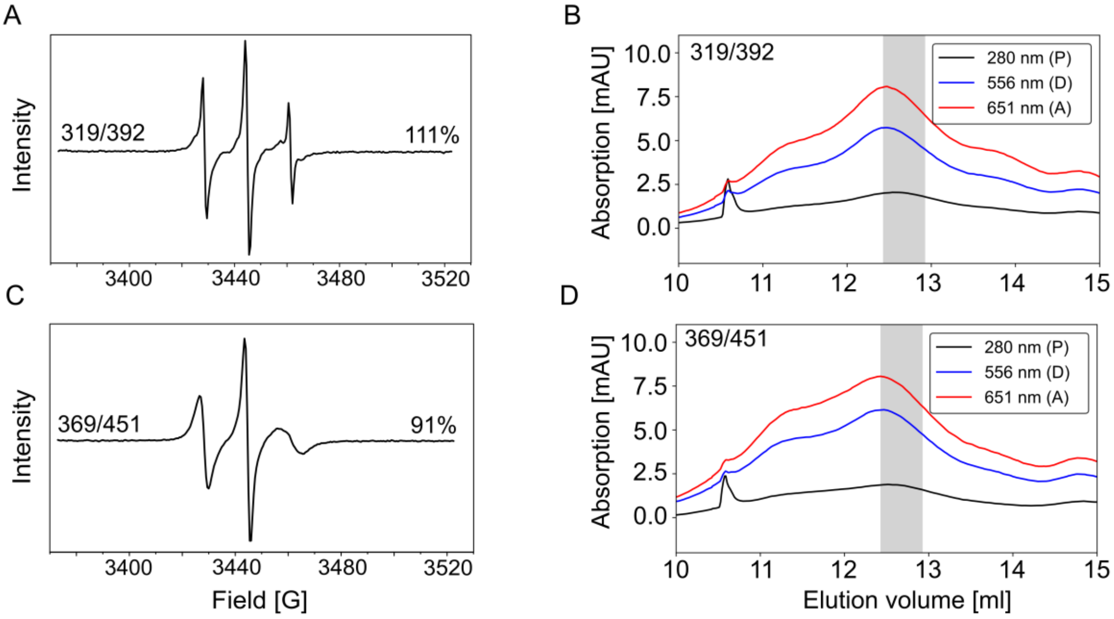


Figure 2-14: Labelling verification of SBD2 double cysteine mutants. A) X-band RT cw-EPR spectrum of nitroxide spin labelled SBD2 double cysteine mutant 319R1/392R1 at 25 µM protein concentration (350 scans). The labelling efficiency that was determined via integration of the signal is mentioned in the figure. **B)** Absorption chromatogram of fluorophore labelled HiSiaP double cysteine mutant 319C/392C (AlexaFluor

555 as donor and AlexaFluor 647 as acceptor) after elution of SEC column. The absorption was detected at three different wavelengths, for the protein (black) and specific for the donor (blue) and acceptor (red) absorptions. The fraction that was used for smFRET experiments is indicated with a grey bar. Labelling and SEC were performed by C. Gebhardt. **C)** Same as A) but for other SBD2 double mutants. **D)** Same as B) but for other SBD2 double mutants. The figure was modified from (Peter et al., 2020).

The fluorophore labelling of SBD2 for smFRET experiments was performed and confirmed with the specific detection of the elution from a SEC by the group of Thorben Cordes. Even if the overall elution profile showed a low intensity for the protein-wavelength at 280 nm, the two fluorophore absorptions were successfully identified (Figure 2-14 B and D). Again, the labelling efficiency of both mutants was higher than 70% and the labelled mutants were used for the smFRET measurements.

2.4.3 Distance measurements with PELDOR and smFRET on SBD2

The smFRET experiments on the labelled SBD2 mutants resulted in clear populations and single distributions of FRET energy in a 1D-E-histogram (Figure 2-15 D, E and Figure SI-9). The calculated smFRET distances indicated an overall good agreement with the simulated distances for both conditions (Figure 2-15 F). For the holo measurements, the experimental distances closely matched to the simulated ones, whereby in the apo conditions the predicted distances for both mutants were slightly larger. Nevertheless, the discrepancy is still in agreement with the experimental error for the smFRET experiments of around ± 5 Å and therefore, the smFRET experiments for this protein successfully represented the expected conformational change (Hellenkamp et al., 2018).

The standard PELDOR experiments for both SBD2 mutants revealed time traces with good modulation depth, good signal-to-noise ratio and a clear visible oscillation (Figure 2-15 A and Figure SI-8). In contrast to the smFRET experiments, the calculated PELDOR spectroscopy distance distributions showed a different behaviour than expected (Figure 2-15 F). The apo measurements for both mutants resulted in a mixture of two distinct peaks, similar to the MalE holo measurements of mutants 29/352 and 36/352 (Figure 2-11 F). The comparison with the simulated PELDOR distances assigned the two distance peaks to the open and closed protein states and suggested thereby a mixture of the two conformations in the PELDOR sample without substrate. Through addition of substrate, the ratio of the two peaks was clearly shifted to the shorter distance, albeit a small amount of the large distance was still observable (Figure 2-15 F). Such closed states in the absence of the substrate were described before for other SBPs but are not known for SBD2 and are not in agreement with the smFRET data (de Boer et al., 2019a, Tang et al., 2007).

To investigate the reason for the deviations, the experimental conditions were varied for the SBD2 protein, similar as for the MalE model system. First, the holo measurement for mutant 369/451 was repeated with a higher excess of substrate to analyse if the remaining open state protein can be triggered to adopt the closed state (Figure 2-15 F). Both PELDOR distance distributions, with standard and 10 times higher substrate concentration, showed the same small amount of the open conformation, indicating no obvious change of the substrate binding affinity due to the freezing of the sample. In the next step, the PELDOR samples were analysed regarding contaminations. For this, the PELDOR experiment was repeated with a significantly lower protein concentration of a factor 10 to see if a possible equilibrium between open and closed states can be changed. However, the PELDOR distance distribution with low concentration of mutant 369/451 was unchanged compared to the standard concentration experiment (Figure SI-8). Additionally, the protein sample after spin labelling procedure was analysed with LCMS (liquid-chromatography mass spectrometry) to check for the presence of the substrate in the apo sample. For this purpose, the protein was aggregated by heat, pelleted by centrifugation and the supernatant investigated with LCMS. The obtained mass spectrum revealed no signals in the supernatant of the solution that could be assigned to the detected spectrum of a glutamine positive control run (Figure SI-10).

Because the previous model system MalE indicated a strong influence of the cryo-protectant on the conformational states, the apo measurement of SBD2 mutant 369/451 was also performed without cryo-protectant. However, this distance distribution still featured the similar mixture of the two conformational states than the standard measurement (Figure 2-11 F). Small changes towards the open conformation were observable in the cryo-protectant free distribution, but this could also be the result of a lower quality of the time trace without cryo-protectant, as visible from the validation of the distribution (Figure SI-7). In summary, the reason for the mixture of two distances in the PELDOR distances distributions could not be described. Possible further reasons that might cause such an effect will be discussed in Chapter 2.6.

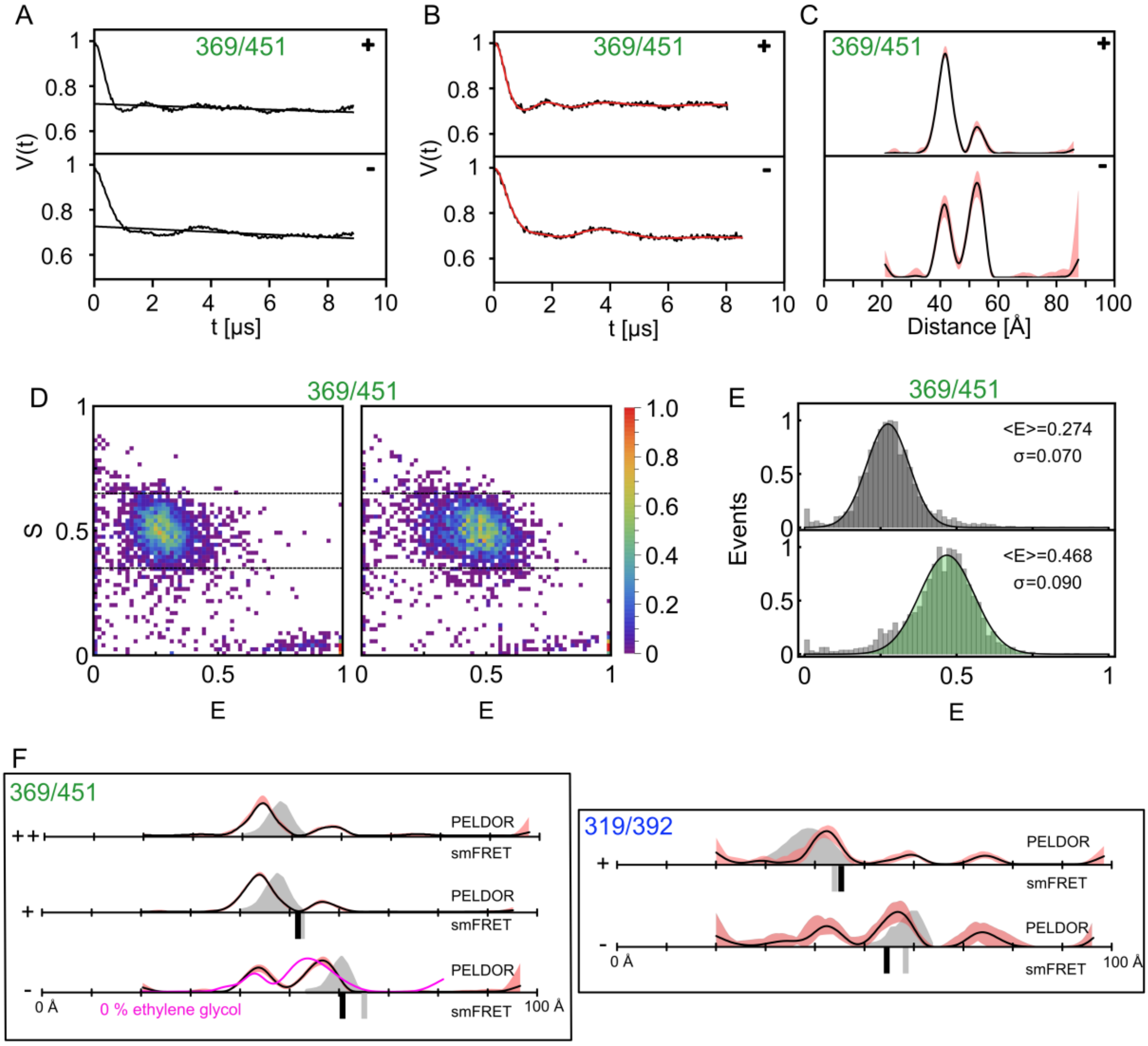


Figure 2-15: Distance measurements of SBD2 with PELDOR and smFRET. A) Raw PELDOR time trace for spin labelled double mutant SBD2 369/451 at apo (-) and holo (+ = 0.1 mM substrate) conditions. The backgrounds that were generated with DeerAnalysis (Jeschke et al., 2006) are shown as black line. **B)** Background-corrected time traces from A) with fit (red) for distance calculation. **C)** Distance distributions from B) with validation of the distribution (red), both determined with DeerAnalysis. **D)** ES-2D-histograms of fluorophore labelled (AlexaFluor 555 and AlexaFluor 647) double mutant SBD2 369/451 for apo (-) and holo (+) conditions. **E)** 1D-E-histograms from corresponding experiment in D) for apo (grey) and holo (green). The histograms were fitted with a gaussian distribution, the mean energy values and standard deviations are mentioned in the Figure. **F)** Summary of PELDOR and smFRET distance measurements for two SBD2 double mutants at different conditions (++ = 1 mM substrate). The PELDOR distance distributions are shown above (simulations as grey areas (Hagelueken et al., 2015), experiments as black curves, validation in red) and the FRET distances below the distance-axis (simulation as grey bars (Kalinin et al., 2012), experiment with black bars). The raw data, as well as the processing steps towards the distances are shown in A-E) exemplary for one mutant and for all other mutant in Figure SI-8 and Figure SI-9. The PELDOR apo measurement for mutant 369/451 was repeated without cryo-protectant (magenta, raw data in Figure SI-7). The smFRET experiments were performed and analysed by C. Gebhardt, the figure was modified from (Peter et al., 2020).

2.5 Influence of label parameters on the accuracy of distance predictions

As seen from the HiSiaP model system in comparison 1, a protein-label interaction can cause large uncertainties in the determined distance and can lead to a deviation between experimental and predicted distances. The consideration of such interactions into the calculation of the label conformers are still challenging and under development (Chapter 2.6) (Spicher et al., 2020). The current standard methods are based on the prediction of AVs in which sterically hindered conformations are deleted from the label rotamer ensembles and no protein-label interactions are considered (Hagelueken et al., 2015, Jeschke, 2018b, Kalinin et al., 2012). Due to the discrepancies in the smFRET results for HiSiaP, the impact of neglected protein-label interactions, as well as the linker length of a label, on the predicted distributions are analysed in the following chapter.

For this, a simulation was set-up with two spherical regions that are 50 Å apart from each other and whose sizes represent the linker length of the label and is varied in two simulations approaches for a 10 Å linker and a 20 Å linker (Figure 2-16 A, simulation was performed by Gregor Hagelüken, University of Bonn). Each of the two regions for one simulation contained 1000 individual atoms, representing the conformations of a label in an ensemble. By default, the atoms were randomly distributed in the spherical regions to imitate a completely free label. To include label-protein interaction in the simulation, a defined percentage of atoms (10, 50 and 100%) from the randomly distributed atoms were fixed at a certain position. These atoms represented an immobilized label that strongly interacts with the surface of the protein and is not evenly distributed in the AV. After definition of the parameters, the geometric average distance of all approaches with varying linker length and different percentage of fixed atoms were determined and compared to the distance of 1000 randomly distributed atoms. This algorithm was processed 1000 times to generate a statistical distribution of interaction sites.

The distributions for two linker lengths and three percentages of immobilized labels are illustrated in Figure 2-16 B. For the case of low immobilization (10% fixed atoms, magenta), the distributions of the predicted error remained very sharp, even after the doubling of the linker length to 20 Å. The immobilization of 50% of the atoms led to a significant broadening of the distributions. For the simulation with a shorter linker, the error distribution was still in an acceptable region of around ± 10 Å. In contrast, the error of the longer linker was broader with around ± 20 Å and will therefore presumably cause great differences between the experimentally determined and simulated distances. The effect was even more dominant in the case of a fully immobilized spin label (100%, cyan). In this case, the use of both linker lengths caused a large error from theoretical distances of a free moving label and will generate a significant difference of the experimental distance to the predicted ones.

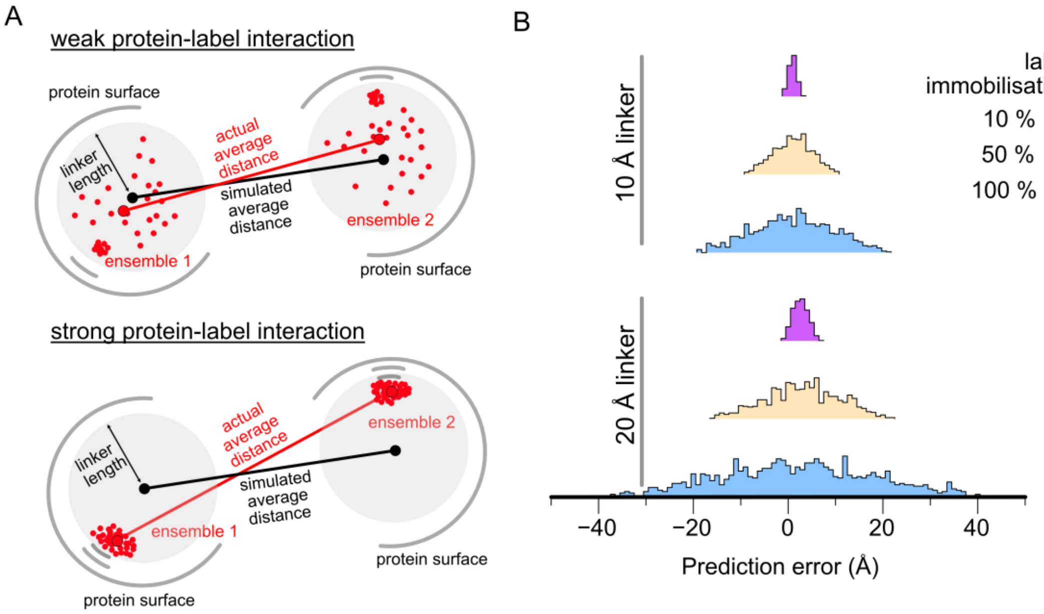


Figure 2-16: Influence of the linker length and immobilization of labels on the correlation between experimental and predicted distances. A) Schematic of the simulation approach. Two spheres (50 Å apart from each other) were generated and filled with 1000 atoms (red balls, representing the labels). The size of the spheres (representing the linker length) as well as the amounts of atoms that are fixed at one place (representing immobilized labels) are varied for the simulations. The averaged distance of all atoms (experimental distance, red) is compared to the average distance for randomly distributed atoms (simulated distance, black), representing the AV calculation (simulation was performed by Gregor Hagelüken, University of Bonn). **B)** Histograms of the error between the experimental and simulated distance from A). The simulations were performed with a short (10 Å) and long linker (20 Å) as well as each linker with three different percentages of immobilized labels (10%, 50% and 100%). For a distribution of the predicted errors, each simulation was run 1000 times and the results are combined in the histograms. The figure was taken from (Peter et al., 2020).

The crystal structure of the matrix metalloproteinase-12 (MMP-12) in complex with a fluorophore Cy5.5 labelled inhibitor is a suitable example for a completely immobilized label (PDB-ID: 5L79 (Bordenave et al., 2016)). The Cy5.5 and the AlexaFluor dyes from the present work have a similar structure with negatively charged sulfonic acid groups (Figure 2-2 and (Bordenave et al., 2016)). In the crystal structure, the label is fixed in a certain position on the surface of the protein and one sulfonic acid group is bound in a small surface pocket between positively charged lysine 233 and arginine 249 (Figure 2-17 A). As consequence, the immobilized label is resolved and could be built into the electron density after diffraction experiment, contrary to a potential free moving label. To compare the immobilized label with simulated label ensembles, the AV of the label was calculated (Figure 2-17 B). The comparison of the AV with the immobilized label from the crystal structure indicated that such a label position in smFRET distance measurements presumably results in significant deviations of the experimental and simulated distances, as observed for the AlexaFluor pair in HiSiaP.

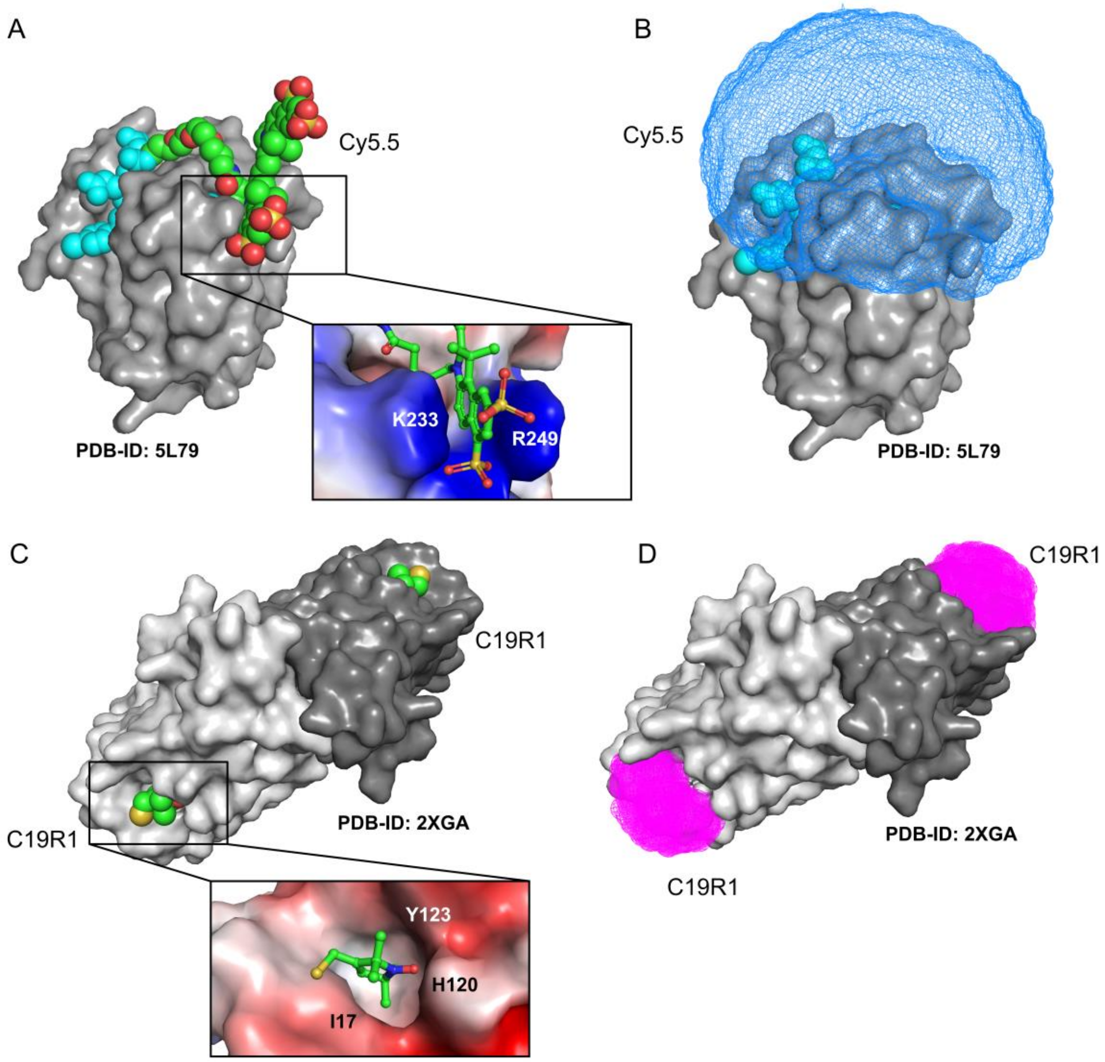


Figure 2-17: Crystal structures with immobilized labels. A) Surface representation (grey) of MMP-12 crystal structure in complex with Cy5.5 labelled (green) inhibitor (cyan) (PDB-ID: 5L79 (Bordenave et al., 2016)). The interaction-area between label and protein is highlighted with the surface charge of the protein, calculated

with PyMOL (Schrödinger LLC) (positive regions in blue, negative regions in red). **B)** Same protein and representation as A) but with AV of the fluorophore (blue) at the labelling position, calculated with FPS (Kalinin et al., 2012). **C)** Surface representation (grey) of dimeric chaperone Spa15 crystal structure with spin label MTSSL at position 19 (green) (PDB-ID: 2XGA (Lillington et al., 2011)). The interaction-area between label and protein is highlighted with the surface charge of the protein, calculated with PyMOL (Schrödinger LLC) (negative regions in red). **D)** Same protein and representation as in C) but with AVs of the spin labels (magenta) at same labelling positions, calculated with mtsslWizard (Hagelueken et al., 2015).

Importantly, also examples for PELDOR spin labels are known that interact with the protein and are immobilized in one or a few conformations. The crystal structure of single spin labelled, dimeric Spa15 chaperone from *Shigella flexneri* indicated a fixed conformation of the nitroxide MTSSL in a surface pocket of the protein (Figure 2-17 C) (Lillington et al., 2011). Similar to the fluorophores Cy5.5 from the example above, the AV prediction would result in a large difference between experimental and simulated distances for such PELDOR experiments (Figure 2-17 D).

2.6 Discussion and conclusion

2.6.1 Overall comparison of the results from PELDOR and smFRET experiments

In the last chapters, distance measurements with the two methods PELDOR and smFRET on three SBPs (HiSiaP, MalE and SBD2) were presented and analysed. In summary, the experimental quality of the majority of measurements were quite good and allowed for the calculation and interpretation of the label distances on the proteins with structural interpretations. Despite some deviations, the resulting PELDOR and smFRET distances were also in a good consistency with each other and with the *in silico* predictions from the crystal structures. The experimental and simulated distances from both methods differed by 5 Å for the majority of experiments, while the overall difference distribution is spread in a range of ± 10 Å.

However, the investigation of the difference between the experimental and simulated data revealed significant deviations for some measurements. The simulations and examples from Chapter 2.5 showed that even moderate protein-label interactions can cause such significant deviations, as the *in silico* predictions are not consider such a behaviour. Additional to protein-label interactions, some measurements showed further special features, for example biomodal PELDOR distance distributions, that were not observed by the other method and contradicted the expectations from the crystal structure. In the following paragraphs, the different observations from the three model systems, as well as possible variations and modifications that could prevent such deviations, are discussed.

2.6.2 Preventing protein-label interactions

The first comparison on the HiSiaP protein is a good example to show the influence of protein-label interactions on distance measurements. The commonly used AlexaFluor FRET pair resulted in unexpected distances in the smFRET experiments. As shown in the simulation from Chapter 2.5, the strong interaction of a label, either a spin or a fluorescence label, with the protein will cause an error-prone measurement and thereby to a misleading interpretation for the function of the protein. One point that is not significant for the proteins in this comparison but for experiments with membrane proteins is an additional possible interaction of the label with the membrane mimetic system that would result in similar deviations (Hughes et al., 2014).

Based on the experimental experiences, there are several possible ways to avoid such protein-label interactions. One solution can be the use of shorter linkers, whereby a possible deviation will have less impact on the error as shown by the simulation in Chapter 2.5. Regarding the used spin label MTSSL, the linker length, as well as the overall label size, is already relatively small and has a similar size compared to an arginine residue. One alternative version of the commonly used MTSSL is a double-functionalized nitroxide spin label (Rx), which is connected to two nearby located cysteines in a protein (Fleissner et al., 2011). The label is more rigid, can reach fewer protein surface regions and the probability of a protein-label interaction is reduced. However, the requirements to the label position are higher than for a mono-functionalized label and already for the standard MTSSL, often only a small number of labelling positions exists that are suitable for functional studies (Chapter 2.1). Furthermore, for distance measurements with the bi-functionalized Rx label, four mutations are necessary, if no native cysteines can be labelled or other spin centres can be used. Additionally, the connection of two protein regions with one label can cause steric problems and can more easily disturb the structure of the protein. For smFRET fluorophores, free rotation of the labels is a requirement for predictions of the AVs and to exclude orientational effects on the FRET efficiencies (Kalinin et al., 2012). The shortening of the label linker with a probably more rigid label will therefore cause further challenges for the simulation, as well as for the experimental distance calculation.

The second way to avoid deviations between simulation and experiment is the prevention of specific protein-label interactions, especially interactions that cause a 100% immobilized label as presented for the matrix metalloproteinases MMP-12 (Figure 2-17). For proteins, the charge distribution on the surface is often complex and negatively and positively charged regions, as well as uncharged regions, are almost randomly distributed over the surface. Therefore, if immobilized labels are detected or supposed, the use of labels with other functional groups and variably charged regions are a possible alternative. For example, in the smFRET experiments with HiSiaP, the change to another FRET pair resulted in the expected distances that were not obtained with the first used dyes. However, not every kind of label is available with alternative charges or other functional groups. Also, a synthetical modification of charged groups to construct new labels often goes hand in hand with a changed water solubility and hydrophobicity. This effect can cause a strong variation of the labelling procedure and can result in lower labelling efficiencies (Fleck et al., 2021, Qu et al., 2019).

A further alternative to prevent disruptive protein-label interactions in a certain labelled protein is the use of other labelling positions in a different region of the protein. For these new positions, the requirements are similar as for the initially used labelling positions: large distance change between label positions upon conformational change of the protein, measurable distances between the labels, no functional impact and high accessibility for the label (Chapter 2.1). As mentioned before, many proteins do not have a lot of positions which fulfil such requirements, and the number of suitable positions are often rather small. However, the smFRET results of HiSiaP showed that a change of the labelling positions, as for the double mutant 58/134 compared to the other three mutants, can significantly change the observation and result in the expected behaviour. Indeed, it is not to be ruled out that the labels at new positions also interact with the surface of the protein. However, the distance measurements with several double mutants can be used to identify mutants with immobilized labels or to mutually confirm the measured distances and by this, strengthens the reliability of the results.

Another way to prevent a large difference between experiment and simulation is the inclusion of potential protein-label interactions into the simulation, which can increase the prediction accuracy of the label conformations. Consequently, the simulated data is in a higher agreement with the experimental result and misleading interpretations or deviations are less frequent. Further, such simulations can be used prior to the measurements to identify regions with a potential high immobilization of the label. However, such a theoretical calculation by MD simulations (molecular dynamics) is still time-consuming and a comparison of the standard AV approaches with such MD simulations showed no significant increase on the prediction accuracy (Alexander et al., 2013, Hagelueken et al., 2012, Jeschke, 2013). Nevertheless, new algorithms and programs that try to solve these bottlenecks and give a higher accuracy for the predicted distances are under constant development (Spicher et al., 2020, Spicher and Grimme, 2020).

2.6.3 Influence of cryo-protectants on distance measurements

The second model system MalE indicated an unexpected behaviour in the standard PELDOR experiments, which could be presumably ascribed to the supplementation of cryo-protectant. As mentioned above, the addition of cryo-protectant is necessary to avoid protein aggregation and increase the T_2 relaxation time to receive higher quality time traces and longer detectable distances (Georgieva et al., 2012, Jeschke, 2012, Schmidt et al., 2020). While in the most PELDOR studies no effect from the cryo-protectant on the distance measurement is mentioned, examples of proteins are known that can interact with a small cryo-protectant molecule (Georgieva et al., 2012, Schmidt et al., 2020, Vagenende et al., 2009). Particularly the chosen class of SBPs are prone to a changed structural or functional behaviour, as they have a native cavity in which the binding of a small molecule can trigger a large conformational change.

The first step in analysing a possible impact of a cryo-protectant on the conformational state of a protein is the use of a different type from a large arsenal of such compounds (Vera and Stura, 2014). As seen for MalE, the distance distribution changed significantly by replacing the *d*-ethylene glycol with *d*-glycerol and supplied a clue for an initially difficult to interpret distance distribution. Despite the simple exchange of cryo-protectants, PELDOR sample preparation with rapid-freeze quenched (RFQ) allows the use of a lower concentration or even a complete absence of a cryo-protectant (Georgieva et al., 2012, Schmidt et al., 2020). However, the production of RFQ samples requires further equipment and complicates the standard sample preparation of snap-freezing. Another conceivable alternative is a room temperature PELDOR measurement in which no cryo-protectant is needed. As mentioned before, room temperature PELDOR experiments are currently not standard experiments, due to spectroscopical reasons (Jeschke, 2012). Alternatively, new spin labels are under current development that may enable room temperature PELDOR experiments, such as trityl radicals (Fleck et al., 2020, Reginsson et al., 2012). Nevertheless, a comprehensive functional investigation of a protein with nowadays standard procedure for PELDOR spectroscopy should involve control measurements with varied or without cryo-protectants to exclude an impact of these additives on the measured distances. For the proteins in the presented comparison, the two model systems MalE and SBD2 were analysed with alternative or without cryo-protectant, while for HiSiaP, the homologous protein from *V. cholerae*, VcSiaP, was previously studied without cryo-protectant and no functional change was observed (Peter et al., 2020).

2.6.4 Deviations of PELDOR and smFRET distances with unknown reasons

The third SBP model system SBD2 resulted in an unexpected PELDOR distance distribution, for which no explanation could be experimentally found. Hypothetically, again the freezing conditions of the PELDOR measurements could be the reason, which can have an influence on the conformational state of the protein. Due to the freezing procedure, it is possible that a conformational state is generated and stabilized that is not visible in the room temperature smFRET measurements. Furthermore, it could be possible, that also the smFRET measurements do not represent the correct equilibrium between the two conformational states. Nevertheless, to find out which method represents the real behaviour of the protein in solution, more parameters need to be analysed, for example other labelling positions, alternative labels for PELDOR and smFRET or the mentioned RFQ sample preparation for PELDOR measurements. Additionally, a further cross-validation with another orthogonal method or structural analysis of the labelled proteins can help to analyse the conformational states that were observed in the PELDOR and smFRET experiments.

2.6.5 Conclusion

For the first time, an experimental approach was performed in this work, which compared the commonly used integrative structural biology methods PELDOR and smFRET spectroscopy in high detail. The three model systems from this cross-validation comparison showed strengths and advantages, but also some limitations of the two methods. The unexpected results, caused by various reasons such as characteristics of the labels or addition of supplements, give a good impression about the analysis and reliability of resulting distances. Consequently, the comparison showed that it is important to keep in mind that also these methods have restrictions, since they are often used to circumvent limitations of X-ray crystallography or cryo-EM. The comparison also indicated that a cross-validation of the results, especially if they are inconsistent with expectations from the protein models, is important and essential to avoid misleading interpretation.

Chapter 3

Sialic acid binding mechanism of TRAP transporter SBP VcSiaP

Preamble:

Preliminary work on the SBPs of TRAP transporters, especially the expression and purification routine, was done in cooperation with Janin Glänzer (University of Bonn). The first sialic acid bound structure of VcSiaP was solved with the help of Niels Schneberger during his Master thesis. The smFRET experiments were performed in cooperation with Christian Gebhardt, Marijn de Boer and Thorben Cordes (University of Munich and University of Groningen).

Parts of this chapter were published in: Peter, M. F.; Gebhardt, C.; Glaenzer, J.; Schneberger, N.; de Boer, M.; Thomas, G. H.; Cordes. T.; Hagelueken, G.; Triggering closure of a sialic acid TRAP transporter substrate binding protein through binding of natural and artificial substrates, *Journal of Molecular Biology,* **2021**, *433* (3), 166756.

3.1 Structural elucidation of sialic acid bound VcSiaP

Previous studies on the function of the SBP VcSiaP from the sialic acid TRAP transporter from *V. cholerae* revealed a factor 10 decreased binding affinity to the substrate compared to the three related SBPs, HiSiaP, FnSiaP and PmSiaP from homologous transporters (Gangi Setty et al., 2014). Further, even though the sequence identity between VcSiaP and HiSiaP with 49% similar amino acids is quite high (67.3% similarity, calculated with EMBOSS Needle Pairwise Alignment, (Madeira et al., 2019)), VcSiaP seems to not interact with the membrane domains HiSiaQM from the *H. influenzae* sialic acid transporter (Mulligan et al., 2009). In an attempt to find an explanation of these differences as well as thoroughly describe VcSiaP, this chapter is focused on structural studies that describe the missing substrate bound conformation, analyse the substrate binding interactions and compare it to related structures.

For the crystallization experiments, the protein VcSiaP was expressed and purified as described in Chapter 9.1.4.1. The expression was performed in M9-minimal medium to prevent contaminations with sialic acid during purification, as previous studies revealed a proportion of around 20% substrate bound VcSiaP after expression in standard LB medium (personal communication with Janin Glänzer, University of Bonn). In contrast, the M9-medium expression showed no contamination and the substrate could thus be selectively added to the protein in appropriate amounts before further usage or analysis. A typical result of purification from VcSiaP for crystallization is shown in Figure 3-1. The SEC after Ni^{2+}-AC indicated an intense, single peak at around 70 mL elution volume which corresponds to a protein mass of around 30 kDa (Figure 3-1 A). The SDS-gel of the SEC fractions showed a high excess of the target VcSiaP with an intense protein band at around 35 kDa ($VcSiaP_{tagged}$: 37 040 Da) (Figure 3-1 B). Additionally, two other protein bands were visible at smaller weights, from which the lowest one was previously identified as a zinc/cadmium-binding protein (Master thesis (Peter, 2016)). After digestion of the N-terminal affinity His_6-tag with TEV protease and an inverse Ni^{2+}-AC, the VcSiaP protein was obtained as a pure solution with minimal amounts of impurities (Figure 3-1 C, flowthrough fraction). The SDS-gel further indicated a successful cleavage of the tag by a shift of the VcSiaP band towards smaller masses ($VcSiaP_{untagged}$: 33 755 Da). The digestion of the affinity tag should avoid an interaction of amino acids from the tag with the binding cleft of the protein, as observed in another VcSiaP structure (PDB-ID: 7A5C (Peter et al., 2021)).

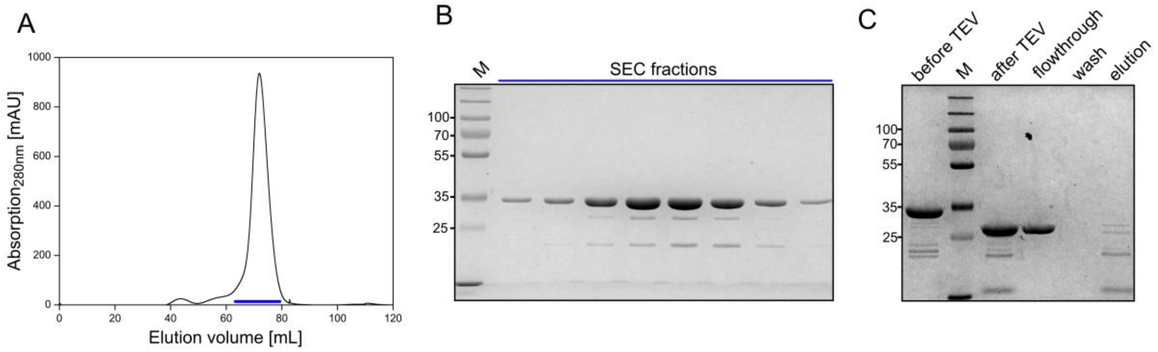


Figure 3-1: Purification of VcSiaP for crystallization. A) SEC run on a HiLoad SD 75 16/600 column with detection at 280 nm absorption (buffer: 50 mM Tris (pH 8), 50 mM NaCl). **B)** SDS-gel with samples from the SEC eluted fractions in A). **C)** SDS-gel with samples from TEV-protease digestion of VcSiaP and separation of digested products with inverse Ni^{2+}-AC. The flowthrough fraction was concentrated to around 30 mg/mL and used for crystallization.

3.1.1 Protein structure of VcSiaP with sialic acid

The purified and cleaved VcSiaP protein was supplemented with sialic acid and crystallization trials were prepared in sitting drop 96-well plates with commercially available screens (Chapter 9.2.7). A single large crystal was detected after 55-135 days at 20 °C in the *Morpheus* screen condition F4 (Figure 3-2, ingredients see figure description). The crystal had a size of around 600 x 70 x 70 µm and a rhombic prism-shaped morphology. The crystal was collected, frozen with cryoprotectant (35% glycerol) and diffraction images were recorded at the SLS (swiss light source) at the Paul Scherrer Institute (Switzerland). The diffraction images revealed a high resolution of more than 2.0 Å and well-defined diffraction spots. The diffraction images were processed as described in Chapter 9.2.7, the phases were obtained by molecular replacement using the homologous substrate bound structure of HiSiaP (PDB-ID: 3B50 (Johnston et al., 2008)) and the structure was solved at a resolution of 1.68 Å.

The electron density after molecular replacement unambiguously revealed the substrate sialic acid in the binding pocket of the protein (Figure 3-3 A). The structure was refined as described in Chapter 9.2.7, until good quality indicators such as R-factors and geometric parameters were obtained. The final refinement parameters, as well as an overview of data collection and crystallographic parameters, are listed in Table 3-1. The refined structure of the protein is illustrated in Figure 3-3. As expected, the protein has the overall structure of classical SBPs (Berntsson et al., 2010). The protein consists of two lobes, an N-terminal (orange) and C-terminal (yellow) lobe, connected by β-sheets and a long backbone-helix (brown) (Figure 3-3 B and C). Sialic acid was embedded between these two lobes and allowed for the observation of specific substrate-amino acid interactions, which have already been proposed by binding studies and comparison with other sialic acid SBPs (Gangi Setty et al., 2014, Johnston et al., 2008, Müller et al., 2006). For example, direct interactions between the substrate and two arginine residues at positions 125 and 145, as well as with asparagine 185 were observed (Figure 3-3 A). Especially the interaction between the carboxy group of the substrate and the guanidino group of arginine 145 appears very strong, as this interaction was already proposed as a substrate selection filter (Fischer et al., 2015). A further classification of the binding interactions and comparison with other proteins is presented in Chapter 3.2.

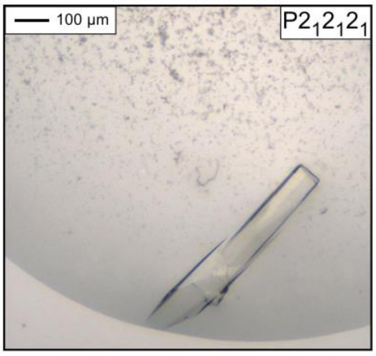


Figure 3-2: Crystal of VcSiaP with sialic acid. Protein crystal in sitting drop well of condition *Morpheus* F4 by using VcSiaP supplemented with sialic acid (condition from Molecular Dimension (UK): 0.12 M monosaccharide mix (0.2 M D-glucose, 0.2 M D-mannose, 0.2 M D-galactose, 0.2 M L-fucose, 0.2 M D-xylose, 0.2 M N-acetyl-D-glucosamine), 0.1 M buffer solution (pH 6.5, 1 M, imidazole, MES monohydrate) and 37.5% v/v precipitant mix (25% v/v MPD, 25% PEG 1000, 25% w/v PEG 3350)). The crystal grew between 55-135 days after preparation.

In the next step, the newly described substrate bound structure of VcSiaP was compared to the substrate free structure (PDB-ID: 4MAG (Gangi Setty et al., 2014)). For this purpose, the N-lobes of the proteins were superposed (Figure 3-3 D). As expected from various experiments and the related protein HiSiaP, a clear conformational change from an open state without substrate, to a closed state with bound substrate was observed (Glaenzer et al., 2017, Müller et al., 2006). An analysis of the conformational change by calculation of a difference distance map with the program mtsslWizard (Hagelueken et al., 2015) revealed the movement of four rigid bodies and is in agreement with the map that was previously generated with a model VcSiaP structure (Glaenzer et al., 2017) (Figure 3-3 E). Further, the VcSiaP difference distance matrix was compared with a similar map for HiSiaP (Figure 3-3 E). Overall, the conformational changes of both structures upon substrate binding are very similar and the same rigid body movements were identified for these two TRAP transporter SBPs. Additional to these equal rigid bodies, the size of the conformational changes is also on a similar scale of around 13.2 - 13.5 Å.

Table 3-1: Data collection, cell parameters and refinement statistics of VcSiaP with sialic acid. Values in the parentheses represent the shell of highest solution.

	VcSiaP (sialic acid)
Space group	$P2_12_12_1$
Unit cell (Å)	48.4, 106.7, 134.5
Wavelength (Å)	1.000009
Total reflections	572385 (40060)
Unique reflections	79908 (7732)
Completeness (%)	99.6 (97.8)
Multiplicity	7.2 (5.2)
Resolution (Å)	44.05-1.68 (1.74-1.68)
I/sigma (I)	9.74 (1.3)
$CC_{1/2}$	0.998 (0.779)
Wilson B-factor (Å^2)	23.74
B-factors (Å^2) (protein/solvent/ligand)	30.9/47.3/18.6
R_{merge}	0.089 (0.92)
R_{work}	0.213 (0.386)
R_{free}	0.247 (0.371)
RMSD bonds (Å)	0.019
RMSD angles (°)	1.8
Ramachandran favoured/allowed/forbidden	98.0/2.0/0
Unfavourable rotamers (%)	1.69
Molprobity Score (Chen et al., 2010)	1.66
PDB-ID	7A5Q

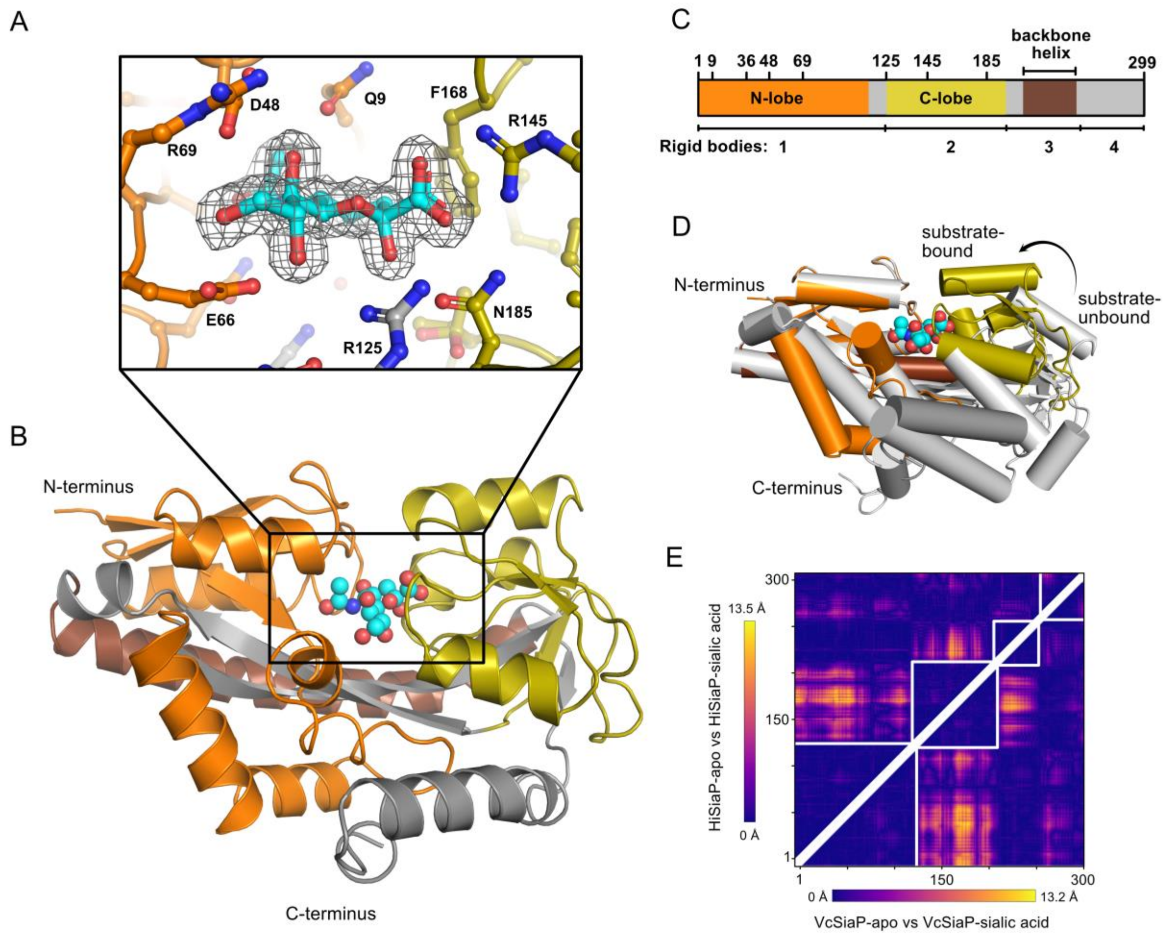


Figure 3-3: Structure of sialic acid bound VcSiaP. A) Substrate-binding site of the refined VcSiaP structure with experimentally obtained electron density ($2mF_0\text{-}DF_c$, sigma = 1.0) and ball-and-stick model of sialic acid (cyan). The amino acids in the direct environment of the substrate are illustrated and the carbon atoms are colour coded as in B-C), depending on the structural motifs. **B)** Overall structure of VcSiaP with bound substrate sialic acid as cyan spheres, the N-lobe in orange, C-lobe in yellow, backbone helix in brown and remaining motifs in grey. **C)** Outline of the primary sequence of VcSiaP with structural motifs, rigid bodies for conformational changes and positions of selected, substrate-binding involved amino acids. **D)** Comparison of the sialic acid-bound conformation of VcSiaP (coloured cartoon) with the substrate-free conformation (white cartoon). For comparison, the N-lobe domains (residues 1-113) were aligned with an RMSD of 0.483. **E)** Difference distance matrix calculated with mtsslWizard (Hagelueken et al., 2015) between the substrate-bound and substrate-free VcSiaP conformations (bottom right) (PDB-ID: 4MAG (Gangi Setty et al., 2014)) and the apo (PDB-ID: 2CEY (Müller et al., 2006)) and holo structure (PDB-ID: 3B50 (Johnston et al., 2008)) of HiSiaP (top left). The rigid bodies are highlighted as white squares.

3.1.2 Binding characterization and comparison with other SBPs

Four SBPs from sialic acid TRAP transporter from *H. influenzae*, *V. cholerae*, *P. multocida* and *F. nucleatum* were already investigated intensively regarding their binding behaviour towards sialic acid. In one study with all four SBPs, an overall 10 times lower binding affinity of VcSiaP compared to the other three SBPs was observed (VcSiaP = 306 nM, FnSiaP

= 45.5 nM, PmSiaP = 19.7, HiSiaP = 16.6 nM) (Gangi Setty et al., 2014). Since only the substrate bound protein structures from the three stronger binding SBPs were available at the time, the reason for this observation remained unknown. The first description of sialic acid bound VcSiaP in the present work allowed a structural comparison between all four SBPs, as well as an investigation of possible reasons for the lower binding affinity of VcSiaP. The initial inspection of VcSiaP in complex with sialic acid revealed no alteration of amino acids that are obviously important for the direct substrate interaction (Figure 3-3 A and Figure 3-5 A). Because the two proteins VcSiaP and HiSiaP were available and already purified several times before (Figure 3-1 and Figure SI-11), the binding of sialic acid to both proteins was reinvestigated to confirm the different binding strengths. Therefore, both proteins were analysed with isothermal titration calorimetry (ITC), which allowed the determination of the binding affinities as well as thermodynamic parameters such as reaction enthalpy and entropy (Chapter 9.2.5). Figure 3-4 showed examples for detected ITC thermograms and binding isotherms for the titration of sialic acid to VcSiaP and HiSiaP. For every protein, the ITC experiment was repeated four times to exclude faulty measurements and minimize the experimental error (Table 3-2).

Overall, both proteins revealed high-quality differential power (DP) peaks for each injection of sialic acid to the protein solution. The recorded data showed a high signal-to-noise ratio (SNR) and a clear binding event, noticeable through the decrease and levelling out of the detected DP. As described in Chapter 9.2.5, the reaction thermodynamics and binding affinities were determined through fitting of a “one set of sites” binding isotherm to the integrated DP peaks with the manufacture’s software. As expected from the high SNR, a good sigmoidal fit was obtained for both proteins (Figure 3-4 B and D). The averaged results from repeated experiments are summarized in Table 3-2. The dissociation constant K_D and the reaction enthalpy ΔH were obtained from the slope of the binding isotherm and the integrated DP intensity, respectively. Overall, slight differences of the determined binding parameters with the literature data were observed, especially for the dissociation constants. (Gangi Setty et al., 2014). In the literature, the dissociation constant of VcSiaP was described by a factor 18 weaker than for HiSiaP, while the presented ITC experiments revealed a factor of 3.6. Even if the tendency for a weaker binding of VcSiaP to sialic acid is consistent in both studies, reasons for the deviation are not obviously identifiable. Possible reasons which might cause such deviations are discussed in Chapter 3.4.

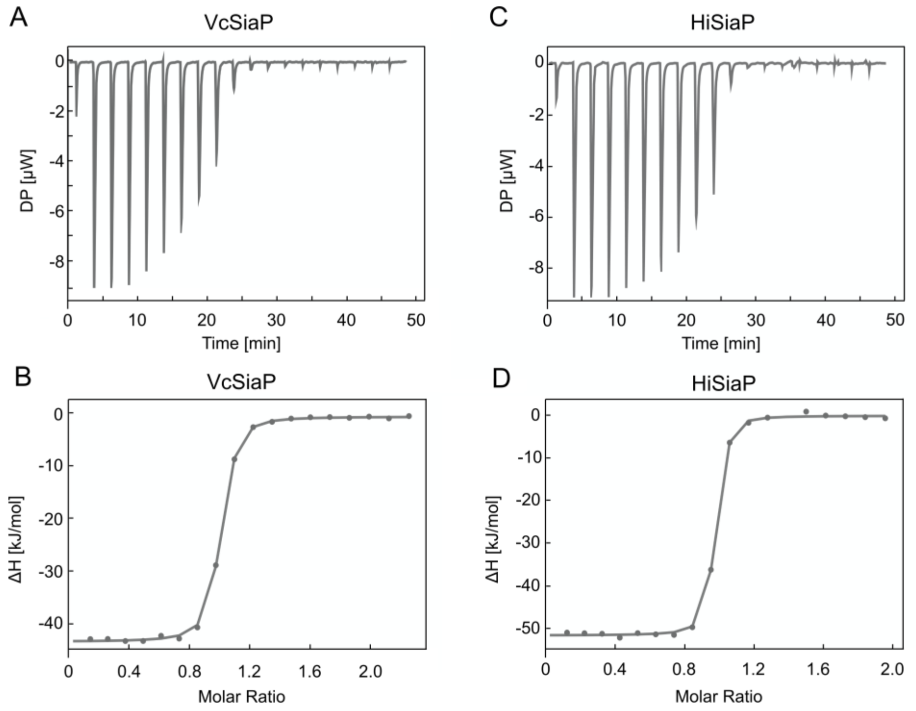


Figure 3-4: ITC experiments with TRAP SBPs. A) ITC thermogram for titration of sialic acid (1.2 mM, 2 µL injections) to VcSiaP (100 µM, wildtype, no tag) in a calorimetry measuring cell at 25 °C. **B)** Automatically calculated binding isotherm from the injection peaks for determination of thermodynamic parameters. The binding affinity from the isotherm is given in the figure. **C-D)** Same as A-B) but for HiSiaP protein, same concentrations.

Table 3-2: Thermodynamic and binding parameters for VcSiaP and HiSiaP. The listed values and errors are averaged from four experiments at 25 °C. * Data are taken from (Gangi Setty et al., 2014) with converted unit from calory to joule.

	VcSiaP$_{exp.}$	**VcSiaP$_{lit.}$***	**HiSiaP$_{exp.}$**	**HiSiaP$_{lit.}$***
N (sites)	0.9 ± 0.002	/	0.8 ± 0.002	/
K_D (nM)	203 ± 17.5	306	56 ± 7.1	16.6
ΔH (kJ/mol)	-42.8 ± 0.3	-29.9	-62.5 ± 0.4	-73.2
ΔG (kJ/mol)	-38.2	-32	-41.6	-44.5
-TΔS (kJ/mol)	4.5	-2.1	20.9	28.7

Nevertheless, to elucidate reasons for the weaker sialic acid binding affinity for VcSiaP compared to the other SBPs, a more detailed structural analysis was performed. For this, the interactions between all known sialic acid TRAP transporter SBPs were analysed by inspection of the structures with PyMOL (Schrödinger LLC) and calculation of the interactions with LigPlot (Laskowski and Swindells, 2011). As described above, the amino acids, which were found to be essential for substrate binding and that have a direct interaction with the substrate, are conserved for all SBPs (Figure 3-5 A). These include the proposed substrate selective filter, arginine 145, as well as the arginine 125 and asparagine 185 (numbering corresponding to VcSiaP) (Fischer et al., 2015, Johnston et al., 2008, Müller et al., 2006).

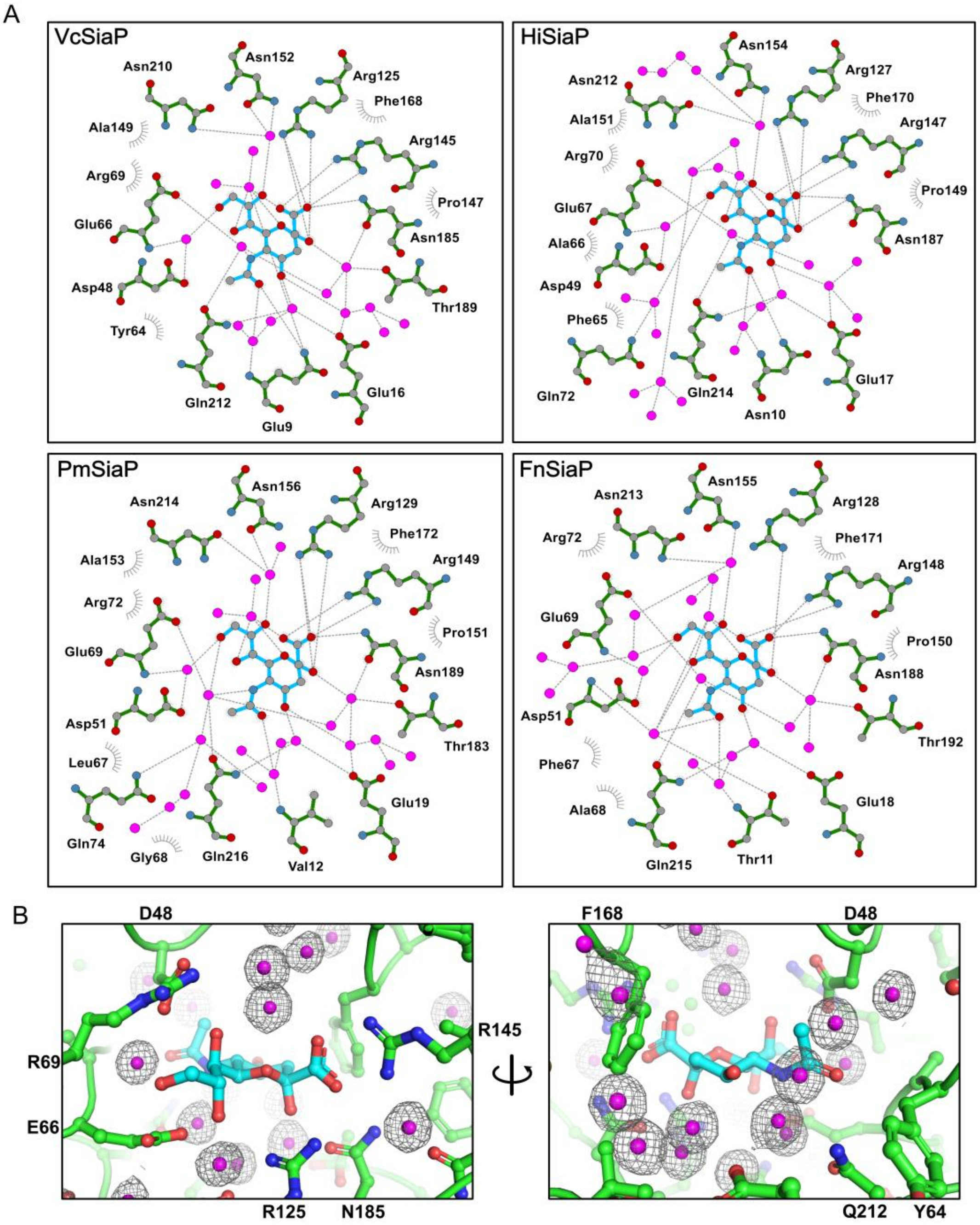


Figure 3-5: Interaction between TRAP SBPs and sialic acid. A) Analysis of interactions between sialic acid (cyan) and SBPs from sialic acid TRAP transporters (HiSiaP PDB-ID: 3B50 (Johnston et al., 2008), PmSiaP PDB-ID: 4MMP (Gangi Setty et al., 2014), FnSiaP PDB-ID: 4MNP (Gangi Setty et al., 2014)), calculated with LigPlot (Laskowski and Swindells, 2011). The amino acids are coloured in green and water molecules in magenta. **B)** Ball-and-stick model of the substrate binding site from the new VcSiaP substrate-bound structure. The protein is coloured in green and the substrate in cyan. The water molecules are displayed as magenta spheres with the experimentally obtained electron density ($2mF_0$-DF_c, sigma level = 1.0).

Further inspection of the VcSiaP substrate binding site revealed a complex water network with highly located water molecules at fixed positions, visible by a defined electron density (Figure 3-5 A and B). A similar water network was recently described as fundamental for a high substrate affinity in HiSiaP (Darby et al., 2019). This water network can be described as a “second-shell” of substrate binding, as amino acids that presumably interact with the water network have an important impact on the substrate affinity (Darby et al., 2019). Considering the water networks, three different regions of interacting amino acids were identified by comparison of all four TRAP SBPs (Figure 3-5 A). The impact of these second-shell variations on the water network in the binding cleft and from this, on the substrate affinity, was investigated in a mutational binding study with combination of ITC experiments and protein crystal structures in the following chapter.

3.2 Dissecting the substrate-protein interaction at molecular level

Since the SBPs HiSiaP and VcSiaP were successfully purified in our laboratory several times before, the comparison of the water network in sialic acid SBPs was focused on these proteins. Furthermore, HiSiaP is the most studied TRAP SBP and had been investigated in several studies before (Darby et al., 2019, Fischer et al., 2010, Johnston et al., 2008, Müller et al., 2006). In the following chapters, the regions that differ in the second shell of TRAP SBPs are described and mutations that were designed for functional investigations are explained. For each region, the substrate binding is analysed for the constructed mutants with a combined approach of ITC experiments for determination of thermodynamic parameters and determined crystal structures for a structural description of the water network.

3.2.1 Structural and functional characterization of region I

The first region of second-shell variations between the sialic acid SBPs is located at the terminal hydroxy-group of the 1,2,3-trihydroxylprop-1-yl chain of sialic acid, illustrated in Figure 3-6 A. In HiSiaP, and also in PmSiaP and FnSiaP, a water molecule (W1-1, Figure 3-6 A) is positioned in the direct environment of the terminal hydroxy group, stabilized by hydrogen bonds to the substrate and glutamine 72, as well as to surrounding water molecules W1-2 and W1-3. The substrate bound VcSiaP structure is the only structure that revealed the absence of W1-1. The analogous position is occupied by a methionine side chain (M81) which has no polar interaction with the substrate (Figure 3-6 A). At the corresponding position, all other three SBPs have a residue with a smaller side chain that do not disturb or influence the water network around water 1-1, for example a small alanine (A82) in HiSiaP. To investigate the correlation between the presence or absence of this water network on the substrate binding, the methionine 81 in VcSiaP was mutated to an alanine (M81A) to create the space for the water network. Based on this mutant, the glutamine 72 from HiSiaP was mutated in the corresponding position for glycine 71 in VcSiaP (M81A G71Q), which potentially stabilizes the water 1-1. Further, by inspection of the local environment, a glutamic acid was identified at position 78 in VcSiaP, which is not present in HiSiaP and might disturb the inserted glutamine in the mutation G71Q. Therefore, a third mutant was created with mutation of this glutamic acid to an alanine, as in HiSiaP (M81A G71Q E78A). All three VcSiaP mutants were constructed, expressed in M9-minimal medium and purified with the standard protocol, including TEV cleavage of the affinity tag (Chapter 9.1.4.1). The final SEC runs and SDS-gels are shown in Figure SI-12 and revealed a high purity of the protein samples for all mutants.

3.2.1.1 ITC binding studies with VcSiaP mutants of region I

First, the binding behaviour of the VcSiaP second-shell mutants was investigated with ITC regarding changes in the binding affinity or thermodynamics, such as enthalpy or entropy. Similar to the wildtype, every mutant was analysed in at least three separated runs to reduce the influence of experimental fluctuations or outliers. One exemplary measurement for each mutant is illustrated in Figure 3-6. All measurements revealed a good quality with clear DP peaks for each substrate injection and a significant change of the integral of these peaks upon substrate addition. Only the double mutant M81A G71Q indicated a lower signal intensity of the DP peaks, resulting in a lower SNR. The thermodynamic parameters for all three mutants are averaged and summarized in Table 3-3. Overall, as expected from the structural proximity of the mutated residues to the location of bound sialic acid in the binding cleft, each mutant showed a changed binding behaviour compared to the VcSiaP wildtype protein.

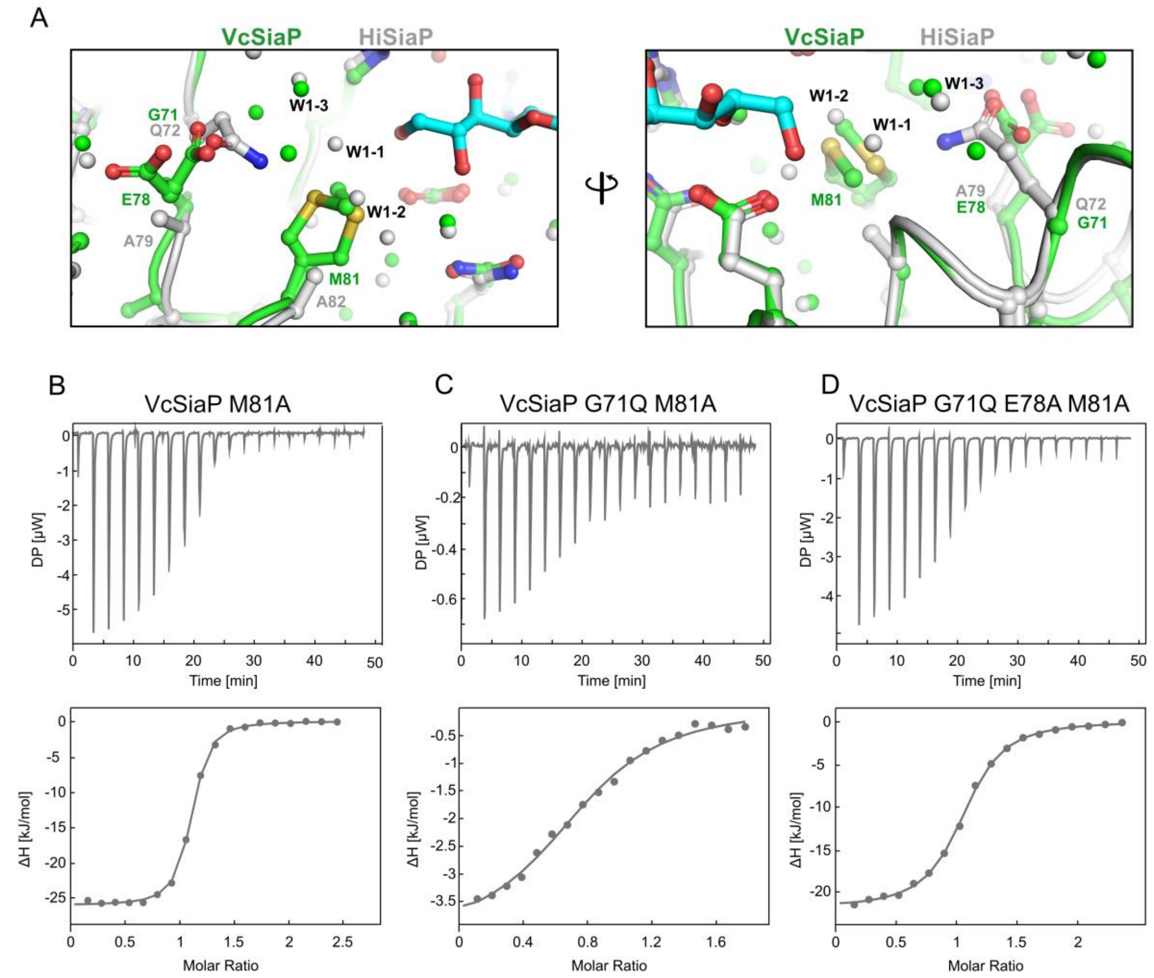


Figure 3-6: Comparison and ITC experiments of mutants from second shell region I. A) Binding site at region I around water molecule W1-1 and VcSiaP methionine 81 residue with alignment of protein structures from VcSiaP (green) and HiSiaP (grey, PDB-ID: 3B50 (Johnston et al., 2008)). The water molecules are coloured in green for VcSiaP and grey for HiSiaP. The right and left illustrations are rotated views of the binding region. **B)** Top: ITC experiment with measured differential heat power for titration of sialic acid (1.2 mM) to VcSiaP M81A mutant (100 µM) at 25 °C (19 injections with 2 µL, except the first with 0.4 µL). Bottom: Automatically calculated binding isotherm from the injection peaks for determination of thermodynamic parameters and binding affinities. **C-D)** Same as B) but for other mutants.

The VcSiaP mutant M81A indicated a decrease of the binding affinity by a factor of 2.5 to 515 nM and a change of the reaction enthalpy from ΔH -42.8 kJ/mol to -26.9 kJ/mol, compared to the wildtype. Using the definition of Gibbs energy ($\Delta G = -RT \ln(K_D) = \Delta H - T\Delta S$, Chapter 9.2.5), ΔG is slightly changed for this mutant, based on the change in the K_D. Since ΔH has a strongly reaction-unfavourable shift for this mutant, the calculation of -TΔS leads to a significant change from 4.5 kJ/mol to -9 kJ/mol (Table 3-3). As the change of entropy in a biomolecular reaction is mostly caused by an alternate behaviour from the solvent rearrangements and the conformational freedom of the protein, the mutant M81A indicated a modified substrate binding process compared to the wildtype (Du et al., 2016). A more pronounced, similar effect was obtained for the double mutant M81A G71Q. The binding constant for this mutant was further decreased from 203 nM to 9400 nM and the enthalpy was strongly changed to a value close to zero (from -42.8 kJ/mol to -3.6 kJ/mol). The lower affinity resulted in an increase of the reaction enthalpy ΔG and in combination with ΔH in a strong decrease of the entropy from 4.5 kJ/mol to -25.5 kJ/mol. Compared to the first single mutant M81A, the weakening of substrate binding is even more pronounced in this double mutant, the introduction of glutamine at position 71 showed no stabilizing effect as assumed in the comparison to HiSiaP.

Table 3-3: Thermodynamic and binding parameters for VcSiaP mutants in region I. The listed values and errors are averaged from at least three experiments at 25 °C.

	VcSiaP M81A	**VcSiaP M81A G71Q**	**VcSiaP M81A G71Q E78A**
N (sites)	1.03 ± 0.004	0.87 ± 0.018	1.04 ± 0.008
K_D (nM)	515 ± 45	9400 ± 1700	3000 ± 300
ΔH (kJ/mol)	-26.9 ± 0.2	-3.6 ± 0.2	-22.3 ± 0.4
ΔG (kJ/mol)	-36	-29.1	-31.7
-TΔS (kJ/mol)	-9	-25.5	-9.5

The rationale behind the third mutation in this region, the deletion of glutamic acid 78 to an alanine, was to remove steric constrains for the introduced glutamine 71. The binding affinity of the triple mutant was determined to 3000 nM and therefore lies between the single and double mutant values (Table 3-3). Within the margin of error, the reaction enthalpy was in the similar range as the single mutant (-22.3 kJ/mol to -26.9 kJ/mol). These parameters led to a decrease of the enthalpy ΔG and a nearly constant ΔS value for the triple mutant, compared to the single mutant. Consequently, the triple mutant also revealed a changed binding behaviour compared to the wildtype protein and the change of these three residues were not beneficial for substrate binding. As postulated during the analysis of the binding region and design of the mutants, the results from the triple mutant suggest that the even weaker binding in the double mutant is presumably partially caused by a steric hindrance between glutamine 71 and glutamic acid 78, which is not present in the triple mutant.

3.2.1.2 Structural characterization of VcSiaP mutants from region I

To further interpret the thermodynamic parameters from the ITC experiments, the three second-shell mutants were co-crystallized with sialic acid. The substrate was added in a 10 times molar excess directly before preparation of the crystal plate. As before for the wildtype protein, several crystallization screens were used for an initial screening of crystallization conditions. For each of the mutant, crystals were observed in at least one condition and diffraction data sets could be collected with resolutions higher than 2 Å for at least one crystal of each mutant at PETRA III at the Helmholtz Research Centre DESY (Hamburg, Germany) or SLS (swiss light source) at the Paul Scherrer Institute (Switzerland).

The crystals with the highest resolution diffraction for each mutant that were used for processing and structure refinement are shown in Figure 3-7. The crystals for VcSiaP M81A and VcSiaP M81A G71Q E78A had a rhombic prism-shape, comparable to the wildtype protein crystal in Figure 3-2. The form of the crystal for mutant VcSiaP M81A G71Q seemed to be similar but was hard to determine more precisely, due to overlaying needles. The ingredients of all crystallization conditions are presented in the description of Figure 3-7.

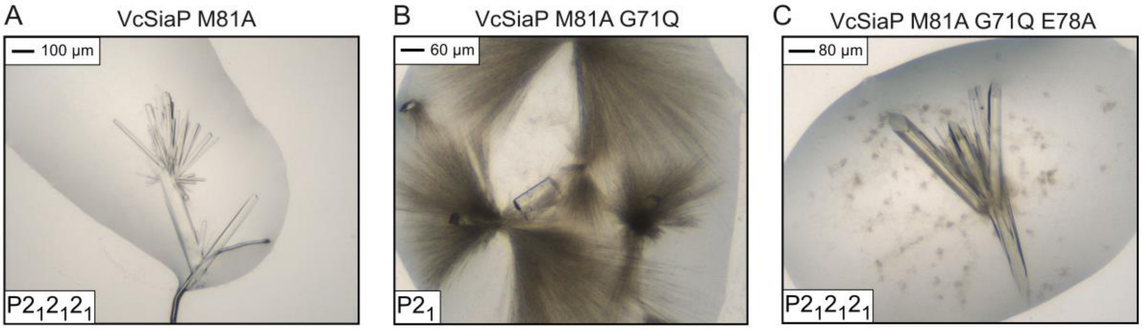


Figure 3-7: Protein crystals of VcSiaP second-shell mutants from region I. A) Crystal of VcSiaP M81A with sialic acid in condition *Morpheus* E12, grown between 26-39 days after preparation (condition from Molecular Dimension (UK): 0.12 M ethylene glycols (0.3 M diethylene glycol, 0.3 M triethylene glycol, 0.3 M tetraethylene glycol, 0.3 M pentaethylene glycol), 0. 1 M buffer solution (pH 8.5, 1 M, Tris (base), BICINE), 37.5% v/v precipitant mix (25% v/v MPD, 25% PEG 1000, 25% w/v PEG 3350)). **B)** Crystal of VcSiaP M81A G71Q with sialic acid in condition *Morpheus* A5, grown directly after preparation (condition from Molecular Dimension (UK): 0.06 M divalents (0.3 M magnesium chloride hexahydrate, 0.3 M calcium chloride dihydrate), 0.1 M buffer solution (pH 7.5, 1 M, sodium HEPES, MOPS (acid)), 30% v/v precipitant mix (40% v/v PEG 500 MME, 20% w/v PEG 20000)). **C)** Crystal of VcSiaP M81A G71Q E78A with sialic acid in condition *PACT* G4, grown between 20-26 days after preparation (conditions from Molecular Dimension (UK): 0.2 M Potassium thiocyanate, 0.1 M Bis-Tris propane (pH 7.5), 20% w/v PEG 3350).

All diffraction data were processed as for the wildtype protein before and the phases were again obtained by molecular replacement, using the previously solved VcSiaP substrate bound structure from Chapter 3.1.1 as search model. The electron density clearly revealed the substrate sialic acid in the binding pocket of each chain in the asymmetric unit from every crystallized mutant. The refinement was carried out for each mutant as described for the wildtype crystal in Chapter 3.1.1. An overview of data collection, crystallographic and refinement parameters is listed in Table 3-4. Importantly, for all mutants, a high resolution was obtained that allowed the localization of water molecules and investigations of amino acid interactions.

Table 3-4: Data collection, cell parameters and refinement statistics of VcSiaP M81A, M81A G71Q and M81A G71Q E78A with sialic acid. Values in the parentheses represent the shell of highest solution.

	M81A	**M81A G71Q**	**M81A G71Q E78A**
Space group	$P2_12_12_1$	$P2_1$	$P2_12_12_1$
Unit cell (Å)	48.5, 107.1, 135.0	55.1, 132.5, 85.8	76.6, 103.9, 160.1
Wavelength (Å)	0.97630	0.97630	0.97630
Total reflections	271038 (8292)	1314298 (156677)	2767960 (327711)
Unique reflections	62375 (3337)	204009 (29124)	209435 (29079)
Completeness (%)	94.6 (87.1)	97.7 (86.3)	96.4 (83.5)
Multiplicity	4.3 (2.5)	6.4 (5.4)	13.2 (11.3)
Resolution (Å)	49.78-1.80 (1.84-1.80)	85.76-1.47 (1.56-1.47)	63.43-1.47 (1.56-1.47)
I/sigma (I)	10.6 (1.0)	9.18 (0.92)	22.60 (2.43)
CC$_{1/2}$	0.999 (0.491)	0.997 (0.434)	1.000 (0.790)
Wilson B-factor (Å^2)	19.76	20.25	15.76
B-factors (Å^2) (protein/solvent/ligand)	28.2 / 41.3 /18.8	19.8 / 24.0 /15.7	18.7 /32.0/ 11.5
R$_{meas}$	0.074 (1.294)	0.113 (1.416)	0.008 (0.967)
Twinning operator	/	h, -k, -l	/
R$_{work}$	20.24 (0.329)	0.189 (0.353)	0.158 (0.256)
R$_{free}$	23.69 (0.364)	0.212 (0.340)	0.184 (0.2600)
RMSD bonds (Å)	0.006	0.002	0.013
RMSD angles (°)	0.95	0.88	1.36
Ramachandran favoured/allowed/forbidden	98.5 / 1.5 / 0	98.7 / 1.3 / 0	98.6 / 1.4/ 0

Unfavourable rotamers (%)	0.19	0.10	0.09
Molprobity Score (Chen et al., 2010)	1.46	1.35	1.21

A comparison of the refined and solved crystal structures of the second-shell mutants from region I is shown in Figure SI-13. Clearly, all mutants adopted the same overall protein structure of the closed, substrate bound VcSiaP conformation that was observed for the wildtype protein before (Chapter 3.1.1). The only small variation was observed for mutant M81A G71Q E78A, in which a C-terminal helix was bent a little bit towards the protein. One possible reason for this observation is a crystal packing effect, where a neighbouring protein chain causes steric stress towards the helix. Indeed, the inspection of symmetry related chains next to this helix revealed a direct contact with the neighboured chain that can presumably trigger the alternative helix position (Figure SI-14).

The crystal structure of VcSiaP mutant M81A with focus on region I is shown in Figure 3-8 A for chain A and a comparison of both protein chains in the asymmetric unit in Figure 3-8 B. The electron density unambiguously indicated the mutated residue at the expected position as well as several defined water molecules next to the substrate. For a comparison of the water network in the mutant, the crystal structure was superposed to the VcSiaP and HiSiaP wildtype structures (Figure 3-8 C and D). The comparison with the VcSiaP wildtype structure revealed that the change of the methionine to an alanine led to formation of a water network next to the substrate, which was occupied in the wildtype structure through the large methionine side chain (W1, W2 and W3, Figure 3-8 C). Several further water molecules that were already present in the VcSiaP wildtype structure were also detected for this mutant and seemed to be nearly unchanged and localized at the same positions (W4, W5 and W6, Figure 3-8 C). Also, further amino acids in this region, such as the glutamic acid 78, were located at similar position within this mutant with slightly different conformations. The comparison of the mutated structure with HiSiaP in Figure 3-8 D indicated an equal position of the alanine residue in both structures, as well as a similar water network. Importantly, the positions of the water molecules next to the substrate differ between the mutated VcSiaP structure and HiSiaP wildtype. Additionally, the two asymmetric chains of the mutant structure in Figure 3-8 B revealed that the positions of the water molecules do not perfectly fit to each other. These small local shifts may occur due to a different stabilization of the water network by amino acid residues that vary from VcSiaP to HiSiaP. By considering the 2.5 times lower binding affinity for this mutant, the shifts in the two mutant chains indicated that the water network has rather low interactions and no stabilizing effect on the substrate, as expected for this mutant. Otherwise, the water molecules should be strongly fixed at the same positions in the two chains, as seen in further mutant structures in the following paragraphs. However, the influence of the mutation on the substrate binding is rather low, although the mutation is in the direct environment of the bound substrate.

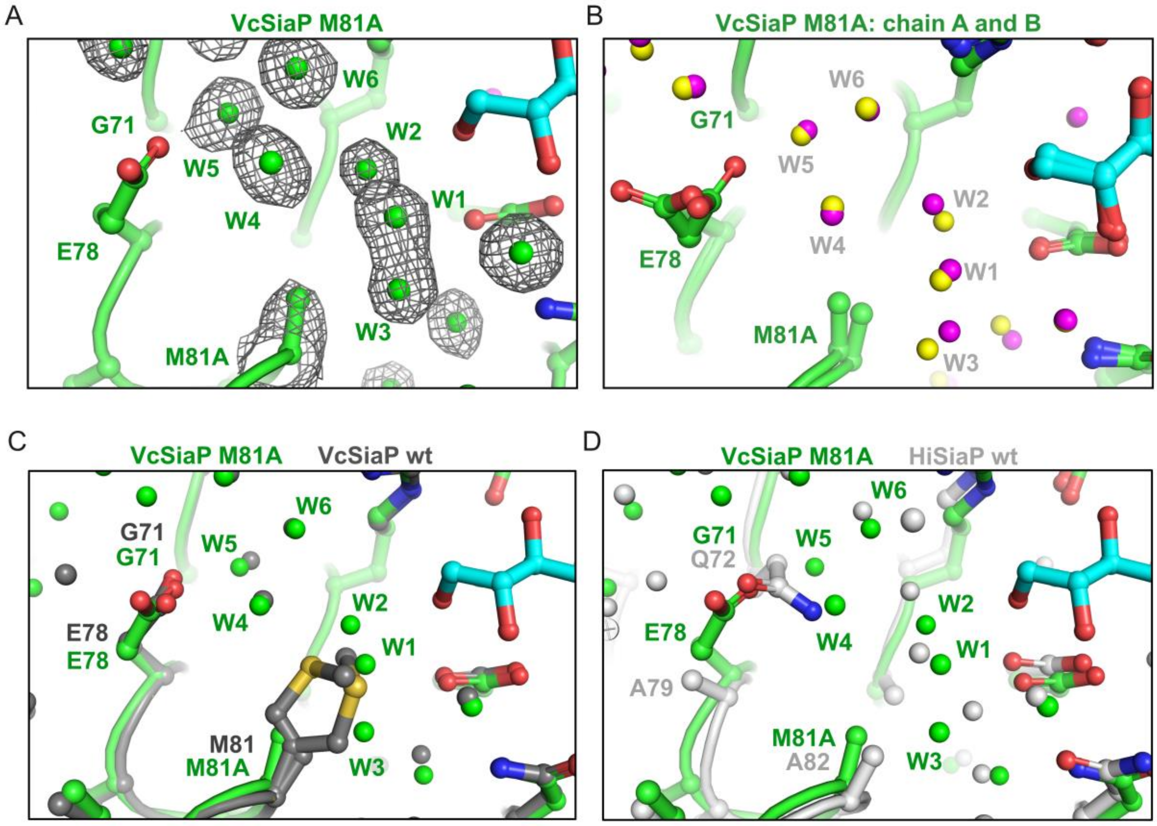


Figure 3-8: Structural details of VcSiaP M81A. A) Mutant VcSiaP M81A (green) with detailed view of the mutated region. The electron density of the mutated amino acid and water molecules in the surrounding are illustrated as grey meshes ($2mF_0$-DF_c, sigma level = 1.0). The waters are shown as green spheres and the substrate as cyan ball-and-stick model. **B)** Alignment of the protein chain A and B in the asymmetric unit from VcSiaP M81A. The water molecules of the two chains are highlighted in different colours. **C)** Detailed structural view of the mutated region from protein VcSiaP M81A (green) and aligned structure of VcSiaP wildtype, substrate bound protein (dark grey) from the previous chapter. **D)** Same as C) but with alignment of HiSiaP wildtype, substrate bound structure (grey, PDB-ID: 3B50 (Johnston et al., 2008)).

For the second mutant VcSiaP M81A G71Q, the crystal structure with focus on region I for chain A is shown in Figure 3-9 A and a comparison of all four protein chains from the asymmetric unit in Figure 3-9 B. Also, for this crystal structure, the electron density revealed the mutations at expected position and a water network in the substrate binding cleft. The alignment of this mutant structure with both wildtypes structures is shown in Figure 3-9 C and D, similar as for the first mutant. The comparison of the four chains and the alignment of chain A to the wildtype VcSiaP structure revealed several changes that presumably occur due to the introduction of the large and polar glutamine at position 71 (Figure 3-9 B and C). All residues of glutamic acid 78 and three of four glutamine 151 residues in the four chains were positioned in another conformation compared to the wildtype structure (Figure 3-9 B). Interestingly, the dominant alternative conformation of glutamine 151 in the mutant structure was comparable with threonine 153 in HiSiaP, which also faces towards the water network (Figure 3-9 D). The introduced glutamine 71 showed several different conformations in the four protein chains, suggesting no strong conserved interactions for this residue. A similar observation was made for the water network next to the substrate, which was not present in the VcSiaP wildtype structure (W1, W2 and W3, Figure 3-9 C). As for the first single mutant, the comparison of all four protein chains from the double mutant revealed significant changes of the water molecule positions in this binding region (Figure 3-9 B). These two coherent observations, the flexible positions for glutamine 71 and the partially disordered water network, are presumably the reason for the low binding affinity for this mutant of around 9400 nM. The introduction of the glutamine 71 disrupted the substrate binding and showed no strong interaction with the water network next to the substrate.

A high steric stress for the introduction of G71Q was postulated and to reduce such a disrupting effect on the substrate binding, the residue E78 was mutated to a smaller side chain in the triple mutant M81A G71Q E78A. The observation of alternative conformations for E78 upon mutation G71Q in all four chains of the double mutant supports such a postulated steric hindrance (Figure 3-9 B). As for the other two mutants, the electron density unambiguously indicated the mutated three residues and defined water molecules next to the substrate (Figure 3-10 A). The comparison of all four protein chains for this mutant is shown in Figure 3-10 B and with the two SBP wildtype structures in Figure 3-10 C and D. Importantly, the conformations of the mutated residues M81A, G71Q and E78A, as well as the previous mentioned Q151, are all conserved in the four different protein chains (Figure 3-10

B). This observation suggests a more defined interaction of these residues in this mutant, compared with the other mutants with higher conformational flexibilities. However, even if the conformations are conserved, the comparison of the triple mutant with the HiSiaP structure revealed an alternative conformation of the introduced glutamine 71 (Figure 3-10 D). Furthermore, the alternative conformation for glutamine 151 that was observed in the double mutant was also detected in the triple mutant, even though a potential steric stress from the double mutant between residues G71Q, E78 and Q151 might be reduced in the triple mutant (Figure 3-10 C and D). Therefore, a potential additional stabilizing interaction of the conformation for Q151 exists in the mutant structure, which is not present in the VcSiaP wildtype.

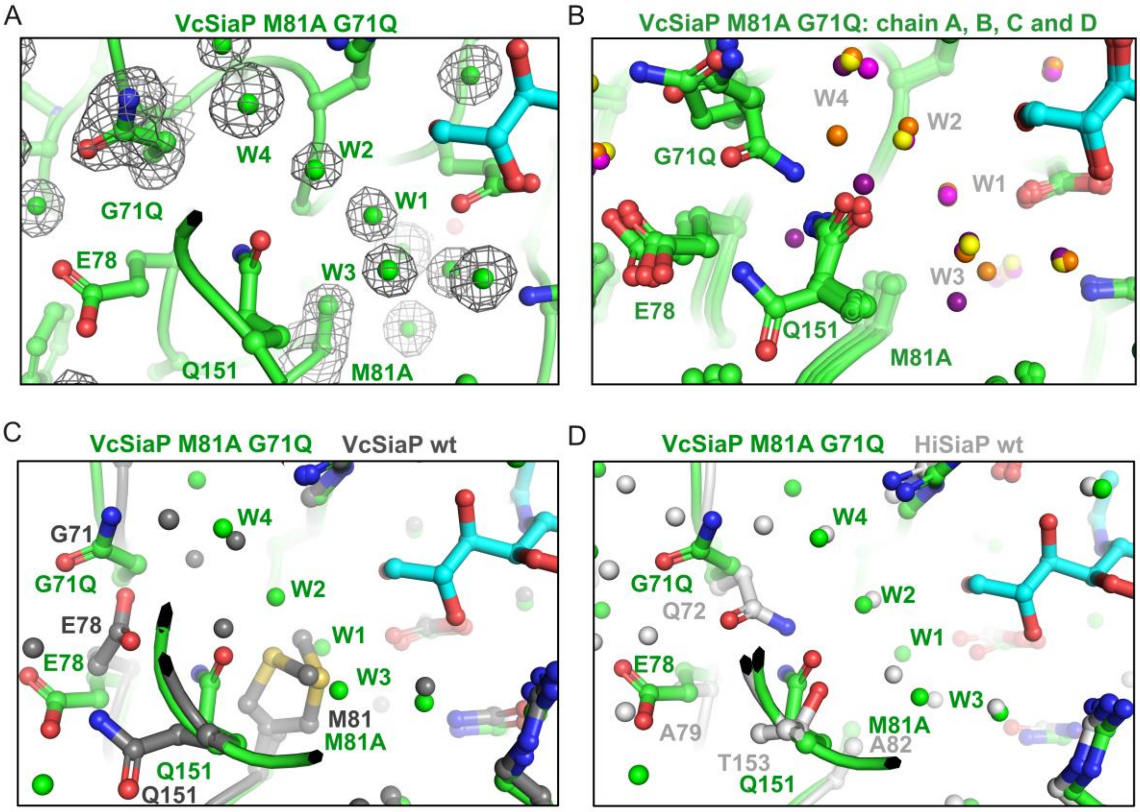


Figure 3-9: Structural details of VcSiaP M81A G71Q. A) Mutant VcSiaP M81A G71Q (green) with detailed view of the mutated region. The electron density of the mutated amino acid and water molecules in the surrounding are illustrated as grey meshes (2mF$_0$-DF$_c$, sigma level = 1.0). The water is shown as green spheres and the substrate as cyan ball-and-stick model. **B)** Alignment of the four protein chains in the asymmetric unit from VcSiaP M81A G71Q. The water molecules of both chains are highlighted in different colours. **C)** Detailed structural view of the mutated region from protein VcSiaP M81A G71Q (green) and aligned structure of VcSiaP wildtype, substrate bound protein (dark grey) from the previous chapter. **D)** Same as C) but with alignment of HiSiaP wildtype, substrate bound structure (grey, PDB-ID: 3B50 (Johnston et al., 2008)).

In contrast to the first two mutants, the comparison of the four protein chains from the triple mutant revealed also highly conserved positions for the water molecules (Figure 3-10 B). However, the positions of the water molecules W1, W2 and W3 are still slightly different in the mutant structure compared to the HiSiaP structure (Figure 3-10 D). This might be caused by further changed interactions between the protein and water network, such as the glutamine 151 residue. Nevertheless, the overall conservation suggests a defined stabilization of the water network through the introduced amino acids. This high degree of conservation in the substrate binding region, including the residue conformations and the water network, are presumably the reason for the three times higher binding affinity for this triple mutant compared to the double mutant. However, the binding affinity was still much weaker compared to the VcSiaP and HiSiaP wildtype proteins.

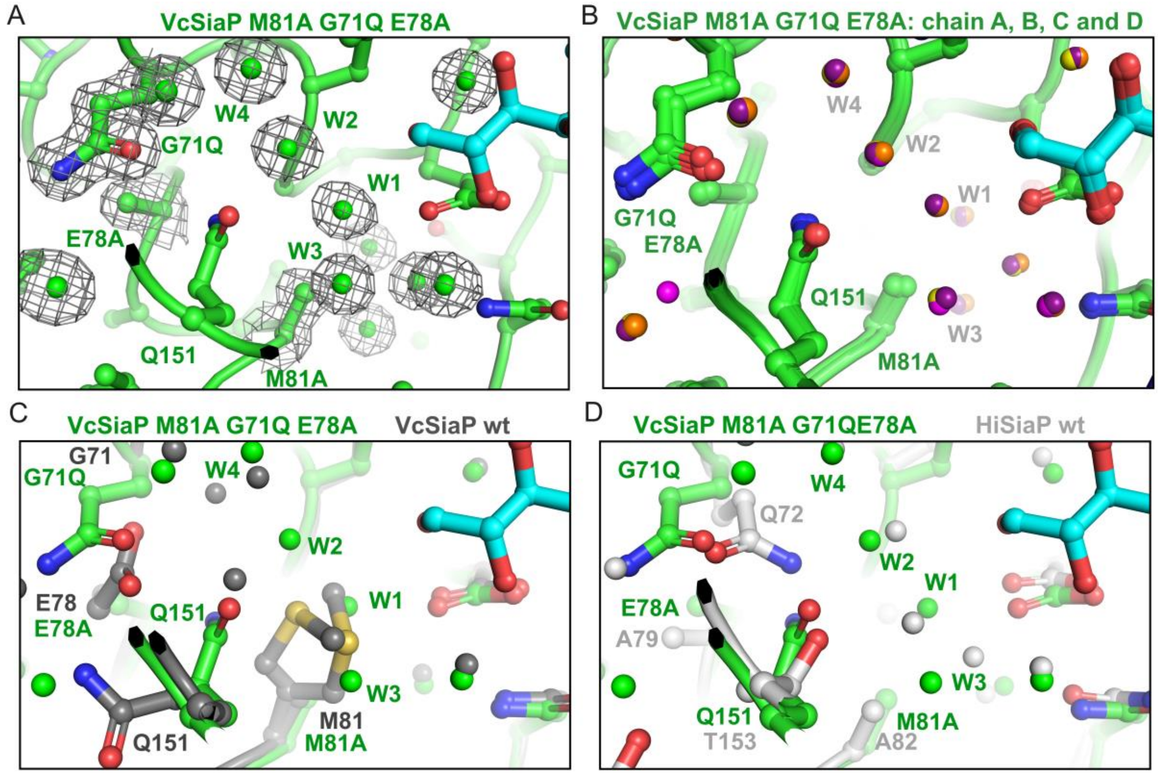


Figure 3-10: Structural details of VcSiaP M81A G71Q E78A. A) Mutant VcSiaP M81A G71Q E78A (green) with detailed view of the mutated region. The electron density of the mutated amino acids and water molecules in the surrounding are illustrated as grey meshes ($2mF_0$-DF_c, sigma level = 1.0). The water is shown as green spheres and the substrate as cyan ball-and-stick model. **B)** Alignment of the four protein chains in the asymmetric unit from VcSiaP M81A G71Q E78A. The water molecules of both chains are highlighted in different colours. **C)** Detailed structural view of the mutated region from protein VcSiaP M81A G71Q E78A (green) and aligned structure of VcSiaP wildtype, substrate bound structure (dark grey) from the previous chapter. **D)** Same as C) but with alignment of HiSiaP wildtype, substrate bound protein (grey, PDB-ID: 3B50 (Johnston et al., 2008)).

For all three proteins with mutations in the identified second-shell binding region I, a new water network between the protein and substrate was clearly observed in the electron density through the removal of methionine 81. However, the binding affinity of the substrate was not increased in one of the three mutants compared to the wildtype proteins. On the other hand, the changes of the amino acid residues in this binding region showed a strong influence on the water network. Overall, the thermodynamic parameters and the crystal structures of the three mutants indicated highly complex interactions between the residues, the water molecules and the substrate, that make it difficult to predict modifications with an increased substrate affinity.

3.2.2 Structural and functional characterization of region II

The water molecule W2-1 (Figure 3-11 A), which interacts with the substrate in a complex way, was identified as a further region with variations in the second-shell interactions. Unlike the first described region, the water network with this central molecule W2-1 is observed in both structures for HiSiaP and VcSiaP. However, the amino acids that are presumably interact with this water vary for all four known structures, a glutamine (Q9) for VcSiaP, an asparagine (N10) for HiSiaP, a valine (V12) for PmSiaP and a threonine (T11) for FnSiaP. In VcSiaP, glutamine 9 is close enough to stabilize the water W2-1 (3.1 Å), while the asparagine 10 residue from HiSiaP contains the same functional group but is too far away for a strong hydrogen bond (3.8 Å) (Figure 3-11 A). In both proteins, W1-1 is also involved in interactions with further water molecules (W2-2 and W2-3). The amino acids which interact with this water network are almost completely identical, apart from serine 44 in HiSiaP and the corresponding alanine 43 in VcSiaP (Figure 3-11 A). To experimentally analyse the impact of the shorter residue, the glutamine 9 in VcSiaP was mutated to an asparagine (Q9N). Further, the alanine at position 43 was mutated to a serine to investigate a possible further interaction of the water network in construct with the mutated glutamine 9 (Q9N A43S). The purification and measurement of this mutant was performed with the help of Seungkyeong Nam during her bachelor thesis. The two mutants were purified in the same way as mutants for region I, the final SEC runs and SDS-gels are shown in Figure SI-12 and revealed again a high purity of the protein samples.

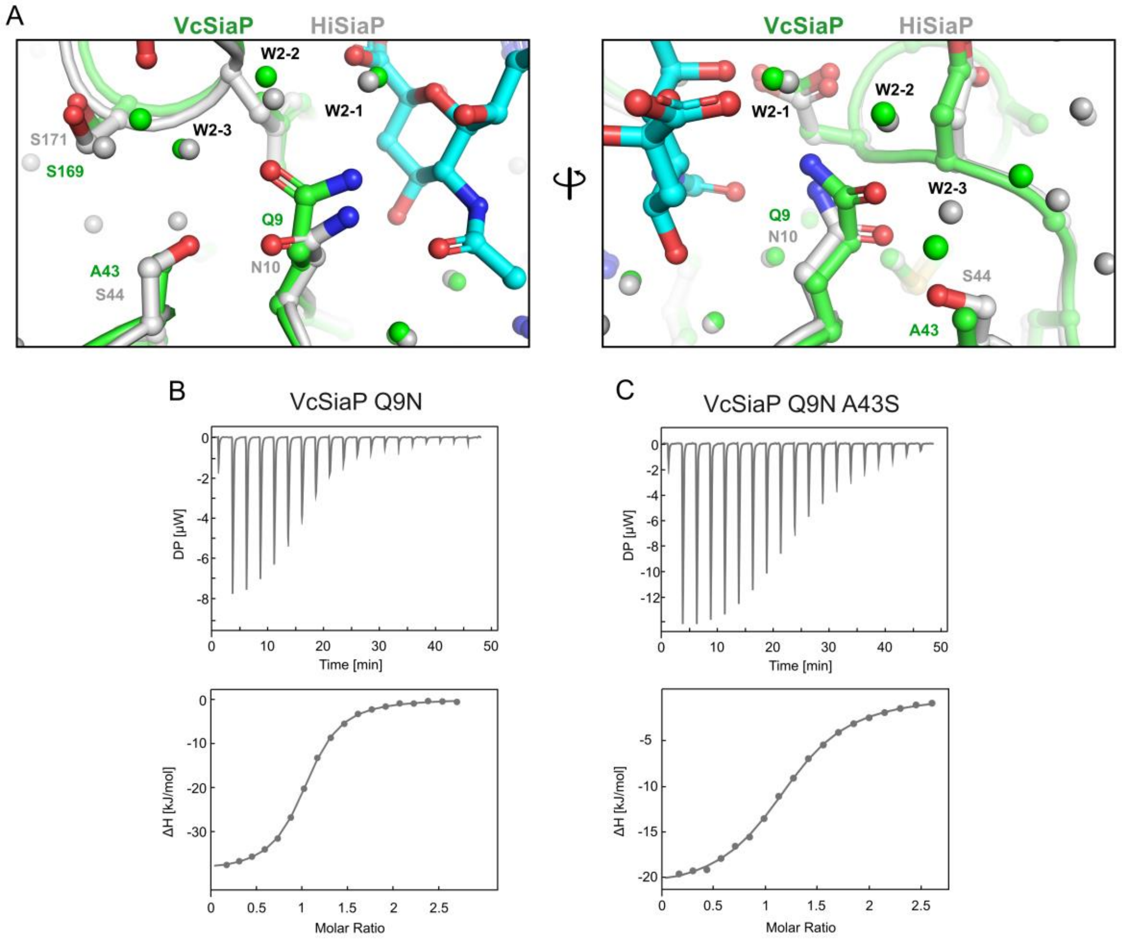


Figure 3-11: Comparison and ITC experiments of mutants from second shell region II. A) Binding site at region II around water molecule 2-1 and VcSiaP glutamine 9 residue with alignment of protein structures from VcSiaP (green) and HiSiaP (grey, PDB-ID: 3B50 (Johnston et al., 2008)). The water molecules are coloured in green for VcSiaP and grey for HiSiaP. The right and left illustrations are rotated views. **B)** Top: ITC experiment with measured differential heat power for titration of sialic acid (1.2 mM) to VcSiaP Q9N mutant (100 µM) at 25 °C (19 injections with 2 µL, except the first with 0.4 µL). Bottom: Automatically calculated binding isotherm from the injection peaks for determination of thermodynamic parameters and binding affinities. **C)** Same as B) but for VcSiaP Q9N A43S mutant.

3.2.2.1 ITC binding studies with VcSiaP mutants of region II

Similar as for region I, the binding behaviour of the VcSiaP second-shell mutants from region II was investigated with ITC. Again, every mutant was analysed in at least three separated runs, one example experiment for every run is shown in Figure 3-11 B and C and the summarized results are listed in Table 3-5. For the mutation of glutamine 9 to the shorter asparagine, a substantial decrease of the substrate affinity by a factor of 16 to 3.2 µM was observed (Table 3-5). Contrary to the thermodynamic behaviour of mutants from the first region, this mutant indicated no change of the reaction enthalpy within the error margin.

Due to the lower binding affinity, the calculated enthalpy ΔG was increased for this mutant and the entropy was also slightly higher than for the wildtype. The overall small changes of the reaction enthalpy and entropy suggested a similar binding event for this mutant and the wildtype protein. This hypothesis is supported by a previous investigation of corresponding HiSiaP N10A mutant in a functional assay (Johnston et al., 2008). The observation that this mutant was nearly fully functional suggests that this position has only a small impact on the overall function of the protein, even though it is located next to the substrate. Additional to the mutation of Q9N, the alanine residue 43 was replaced in a double mutant by a serine as in HiSiaP, to potentially stabilize the water network in region II (Figure 3-11 A). In contrast to the postulated stabilization effect, the binding affinity for the mutant Q9N A43S was further lowered compared to the Q9N single mutant and determined to 27.3 µM, a factor of around 130 weaker than the wildtype (Table 3-5).

Table 3-5: Thermodynamic and binding parameters for VcSiaP mutants in region II. The listed values and errors are averaged from at least three experiments at 25 °C.

	VcSiaP Q9N	**VcSiaP Q9N A43S**
N (sites)	1.02 ± 0.004	1.1 ± 0.012
K_D (nM)	3205 ± 145	27333 ± 1998
ΔH (kJ/mol)	-39.3 ± 0.3	-24.6 ± 0.7
ΔG (kJ/mol)	-31.4	-26.1
-TΔS (kJ/mol)	7.9	-1.5

3.2.2.2 Structural characterization of VcSiaP mutants from region II

To interpret the thermodynamic parameters from the ITC experiments on a molecular level, crystallization trials for the second-shell mutants from this region II were prepared in the same way as for the mutants from region I (Chapter 3.2.1.2). A crystal was obtained for VcSiaP mutant Q9N with a plate-shaped form without specific crystal morphology (Figure 3-12). The diffraction data of this crystal were collected and processed in the same way as for the first three mutants above. An overview of data collection, crystallographic and refinement parameters is listed in Table 3-6. Noticeable, the diffraction data for mutant Q9N were collected with a high resolution of up to 1.18 Å. Thereby, this structure belongs to the top 2.5% of all known protein structures in the protein data bank (4000 out of 159927 at 1.2 Å or higher, PDB, www.rcsb.org, access date: 12.01.2020).

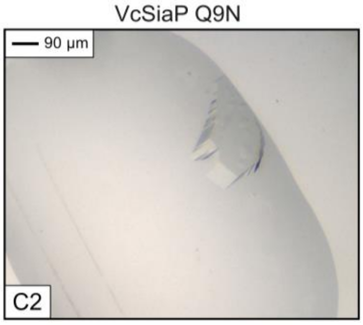


Figure 3-12: Protein crystal of VcSiaP second-shell mutant Q9N. Crystal of VcSiaP Q9N with sialic acid in condition *Morpheus* H4, grown after 55 days (conditions from Molecular Dimension (UK): 0.1 M amino acids (0.2 M DL-glutamic acid monohydrate, 0.2 M DL-alanine, 0.2 M glycine, 0.2 M DL-lysine monohydrochloride, 0.2 M DL-serine), 0.1 M buffer solution (pH 6.5, 1 M, imidazole, MES monohydrate (acid)), 37.5% precipitant mix (25% v/v MPD, 25% PEG 1000, 25% w/v PEG 3350)).

Table 3-6: Data collection, cell parameters and refinement statistics of VcSiaP Q9N with sialic acid. Values in the parentheses represent the shell of highest solution.

	Q9N
Space group	C2
Unit cell (Å)	164.2, 48.6, 107.5
Wavelength (Å)	0.97625
Total reflections	1504799 (52087)
Unique reflections	230190 (10627)
Completeness (%)	99.3 (92.5)
Multiplicity	6.5 (4.9)
Resolution (Å)	88.45-1.18 (1.20-1.18)
I/sigma (I)	14.1 (2.2)
$CC_{1/2}$	0.999 (0.820)
Wilson B-factor (Å^2)	12.58
B-factors (Å^2) (protein/solvent/ligand)	20.3 / 33.5 / 10.1
R_{meas}	0.064 (0.676)
R_{work}	0.173 (0.237)
R_{free}	0.183 (0.254)
RMSD bonds (Å)	0.012
RMSD angles (°)	1.26
Ramachandran favoured/allowed/forbidden	99.2 / 0.8 / 0.0
Unfavourable rotamers (%)	0.37
Molprobity Score (Chen et al., 2010)	1.46

The electron density of mutant VcSiaP Q9N clearly revealed the substrate sialic acid in the binding pocket of each chain in the asymmetric unit and both chains adopt the closed conformation, similar as the wildtype protein (Figure SI-13). For structural characterization, the structure with focus on region II is shown in Figure 3-13 A and a comparison of both protein chains asymmetric unit in Figure 3-13 B. The electron density unambiguously indicated the mutated residue at the expected position as well as several defined water molecules next to the substrate (Figure 3-13 A). In contrast to the structural comparison from region I, the two protein chains and corresponding water molecules for Q9N at region II perfectly fit to each other (Figure 3-13 B).

The structural classification of the mutant reveals the shorter asparagine residue in comparison with the native glutamine in the VcSiaP wildtype structure (Figure 3-13 C). However, the conformation of the side chain in the mutant is slightly different as the corresponding one in HiSiaP (Figure 3-13 D). In both wildtype structures, the glutamine or asparagine interact with the oxygen from the acetylamino group of the sialic acid. In the mutant structure, the distance of Q9N to this position in the substrate was determined to be 4.3 Å and with this, too long for a strong polar interaction (Figure 3-13 E). In contrast, the mutated residue was closer to the hydroxy group at C4 of the substrate and have a polar interaction with this position. For both wildtype proteins, the distances between corresponding residues and the hydroxy group were determined to 4.5 Å and thereby too far away to strongly interact (Figure 3-13 E). An obvious reason for the alternative conformation of mutated asparagine 9 remained elusive from the comparison of the structures. One possible explanation is an interaction with serine 44 in HiSiaP that stabilized the asparagine 10 in a conformation which facilitated the interaction with the substrate. However, the thermodynamic characterisation of VcSiaP Q9N A43S resulted in a much lower binding affinity, which contradicted such a stabilization effect.

Interestingly, the position from the amino group of the wildtype VcSiaP was occupied in the mutant structure by a water molecule that presumably replaced the interaction with the acetylamino group from the substrate (W1, Figure 3-13 E). The inspection of the remaining water network revealed similar positions for the water molecules W2 and W3 as in the VcSiaP wildtype. Consequently, due to the conservation of this water, a slight difference to the HiSiaP water molecules was still observable. Additionally, the water network that was built up in the background indicated completely alternative positions for water W4 and W5.

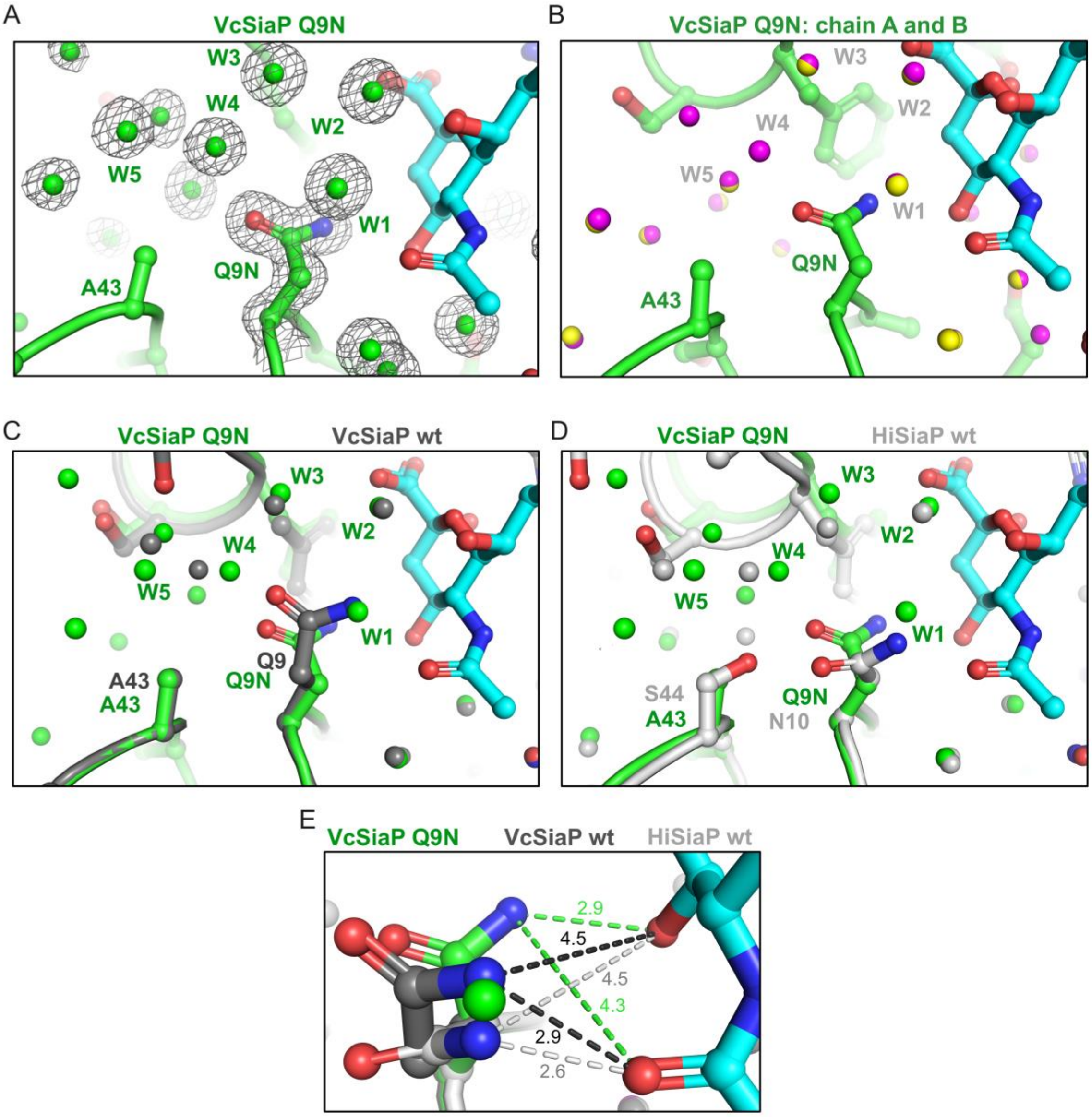


Figure 3-13: Structural details of VcSiaP Q9N. A) Mutant VcSiaP Q9N (green) with detailed view of the mutated region. The electron density of the mutated amino acid and water molecules in the surrounding are illustrated as grey meshes ($2mF_0$-DF_c, sigma level = 1.0). The water is shown as green spheres and the substrate as cyan ball-and-stick model. **B)** Alignment of the two protein chains in the asymmetric unit from VcSiaP Q9N. The water molecules of both chains are highlighted in different colours. **C)** Detailed structural view of the mutated region from protein VcSiaP Q9N (green) and aligned structure of VcSiaP wildtype, substrate bound structure (dark grey) from the previous chapter. **D)** Same as C) but with alignment of HiSiaP wildtype, substrate bound protein (grey, PDB-ID: 3B50 (Johnston et al., 2008)). **E)** Same structures and colour-schemes as in C) and D), with a detailed illustration of the distance between residues and substrate. The distances are shown with lines in corresponding colours, the values were calculated with PyMOL (Schrödinger LLC) and are given in Å.

The combination of the detected binding affinities and the structural analysis indicated an important correlation between the substrate binding and glutamine 9 in VcSiaP, as the affinity of the designed mutant was strongly decreased by a factor of 15. On the other hand, the impact of this residue on the water network seemed to be rather low. Even if the conformation of the residue was changed compared to the wildtype structures, the central waters W1-W3 were still located at the same position. This unchanged water molecules presumably conserve the functionality of the protein during *in vivo* studies with the corresponding mutant HiSiaP N10A (Johnston et al., 2008).

3.2.3 Structural and functional characterization of region III

As third region with variations in the second-shell regions, the water network with central water W3-1 next to C2 of sialic acid was identified (Figure 3-14 A). In VcSiaP, the water is stabilized by a threonine at position 189, similar as in the other two SBPs PmSiaP and FnSiaP. In contrast, HiSiaP is the only structure that contains an alanine in the corresponding position which cannot build up a water bridge to the network. However, in HiSiaP, the water molecule is located at the same position, suggesting additional interactions between the water molecule with the protein. Further, the water molecules that interacts with the central water W3-1 are conserved in VcSiaP and HiSiaP (W3-2 and W3-3, Figure 3-14 A). Because the threonine was the only observable difference between both proteins, only the mutation of threonine 189 to an alanine was designed (T189A). The mutant was purified as described before for the other mutants and the SEC run and SDS-gel indicated a high purity of the protein sample (Figure SI-12).

3.2.3.1 ITC binding studies with VcSiaP mutant of region III

Similar as for the other mutants, the sialic acid binding of VcSiaP T189A was investigated with ITC. The experiments for this mutant revealed a very good quality of the results with clear DP peaks for each substrate injection and a high SNR (Figure 3-14 B). The averaged thermodynamic parameters of three separated experiments for VcSiaP T189A are listed in Table 3-7. Strikingly, the mutation T189A revealed an increase of the binding affinity compared to the wildtype from 203 nM to 49 nM (Table 3-7). Additionally, only for this mutant a significant reaction-favoured change of ΔH from -42.8 kJ/mol to -59.8 kJ/mol was detected. In contrast, the calculation of the entropy ΔS for this reaction indicated a strong increase from 4.5 kJ/mol to 18 kJ/mol. The comparison of the thermodynamic parameters from the single mutant VcSiaP T189A and HiSiaP wildtype protein revealed a high agreement, the binding affinity as well as the reaction enthalpy and the entropy are similar. Consequently, the initially described and observed different binding behaviour between VcSiaP and HiSiaP are presumably caused by different interactions of the water network and second shell residues in this region. However, an obvious reason for this observation could not be determined from the ITC experiments, as the other known SBPs, PmSiaP and FnSiaP, indicated a HiSiaP similar binding affinity but have also a threonine residue at position 189.

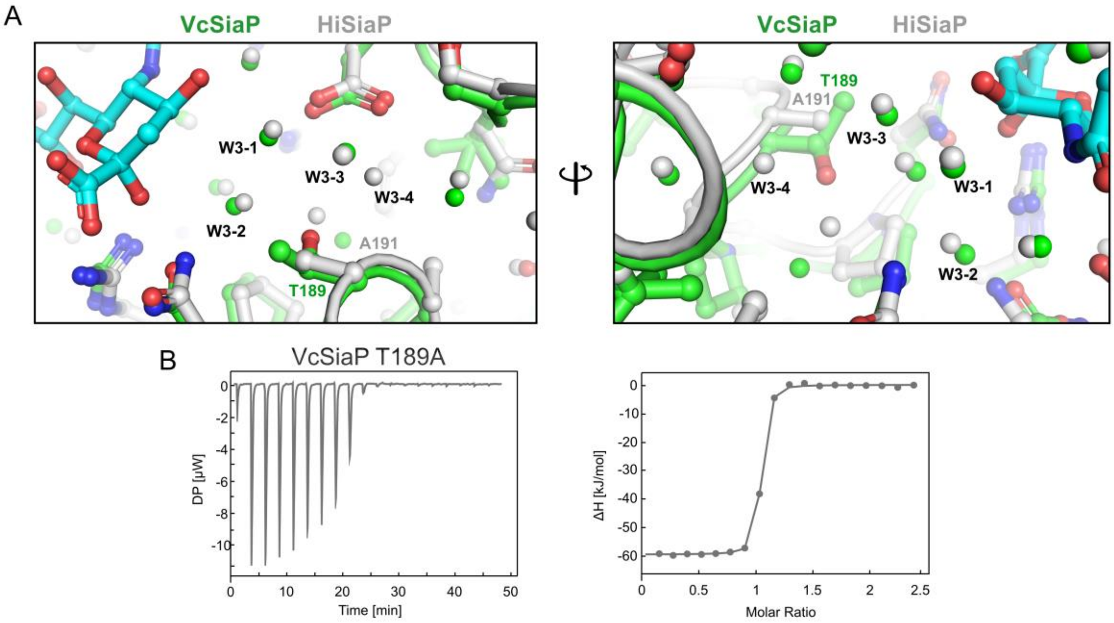


Figure 3-14: Comparison and ITC experiment of a mutant from second shell region III. A) Binding site at region III around water 3-1 and VcSiaP threonine 189 residue with alignment of structures from VcSiaP (green) and HiSiaP (grey, PDB-ID: 3B50 (Johnston et al., 2008)). The water molecules are coloured in green for VcSiaP and grey for HiSiaP. The right and left illustrations are rotated views. **B)** Left: ITC experiment with measured differential heat power for titration of sialic acid (1.2 mM) to VcSiaP T189A mutant (100 µM) at 25 °C (19 injections with 2 µL, except the first with 0.4 µL). Right: Automatically calculated binding isotherm from the injection peaks for determination of thermodynamic parameters and binding affinity.

Table 3-7: Thermodynamic and binding parameters for VcSiaP mutant in region III. The listed values and errors are averaged from at least three experiments at 25 °C.

	VcSiaP T189A
N (sites)	1.03 ± 0.001
K_D (nM)	49.3 ± 7.4
ΔH (kJ/mol)	-59.8 ± 0.2
ΔG (kJ/mol)	-41.8
-TΔS (kJ/mol)	18.0

3.2.3.2 Structural characterization of VcSiaP mutant from region III

To elucidate the reason for the stronger binding of VcSiaP T189A mutant on a molecular level, the protein was co-crystallized with sialic acid in the same way as the other mutants before. Two crystals were observed in a crystallization trial with a trigonal or hexagonal, pyramidal shape (Figure 3-15). The ingredients of the crystallization condition are presented in the figure description. The diffraction data of this crystal were collected and processed similar to the other mutants. The data collection, crystallographic and refinement parameters are listed in Table 3-8. The crystal structure of VcSiaP T189A revealed the closed conformation, as expected from the wildtype structure (Figure SI-13). Further, the electron density clearly indicated the bound sialic acid in the binding pocket, as well as several, highly located water molecules. For detailed characterization of the mutation, the crystal structure with focus on region III is shown in Figure 3-16 A and B. As for all mutants before, the electron density indicated the mutated residue at the expected position as well as several defined water molecules between this residue and the substrate (Figure 3-16 A). The two protein chains and corresponding water molecules at region III overall fit to each other with only minimal variations in the position of the water molecules (Figure 3-16 B).

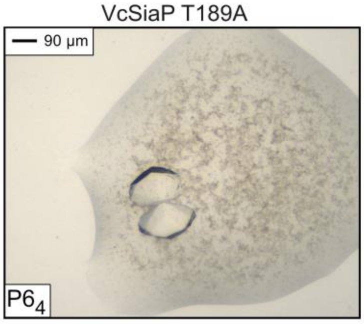


Figure 3-15: Protein crystal of VcSiaP second-shell mutant T189A. Crystal of VcSiaP T189A with sialic acid in condition *JCSG plus* C5, grown after 116-147 days (conditions from Molecular Dimension (UK): 0.8 M Sodium-phosphate monobasic monohydrate, 0.1 M sodium HEPES (pH 7.5)).

The inserted alanine residue was detected at the expected position and revealed no difference to the HiSiaP structure (Figure 3-16 C and D). The only differences between the wildtypes and the mutant structures are slight varied positions in the water molecules W1-W3 and larger deviations in the water molecules that are further away from the substrate, such as W4 and W5 (Figure 3-16 C and D). However, the local shifts for W1-W3, that presumably strongly interact with the substrate, are on the same scale as the differences of the water molecules from the two protein chains from VcSiaP T189A (Figure 3-16 B). Therefore, an impact of these small shifts is probably not responsible for the stronger binding. A meticulous analysis of the environment revealed no further changes between the mutant and the VcSiaP wildtype than the described water molecules 4 and 5. Even if a beneficial influence of these water molecules on the substrate binding is possible, an obvious structural reason for the increased binding affinity of a factor 4 for the mutant could not be described with the crystal structure. Further possible reasons and proposed methods that can be used to study the stronger binding for the mutant VcSiaP T189A are discussed in Chapter 3.4.

Table 3-8: Data collection, cell parameters and refinement statistics of VcSiaP T189A with sialic acid. Values in the parentheses represent the shell of highest solution.

	T189A
Space group	$P6_4$
Unit cell (Å)	117.9, 117.9, 83.4
Wavelength (Å)	1.000031
Total reflections	545290 (30011)
Unique reflections	122931 (14470)
Completeness (%)	94.3 (68.5)
Multiplicity	4.4 (2.1)
Resolution (Å)	48.15-1.75 (1.86-1.75)
I/sigma (I)	13.64 (0.91)
$CC_{1/2}$	99.8 (42.6)
Wilson B-factor (Å^2)	29.11
B-factors (Å^2) (protein/solvent/ligand)	32.02/40.96/24.14
R_{meas}	0.072 (0.932)
Twinning operator	h, -h-k, -l
R_{work}	0.149 (0.267)
R_{free}	0.181 (0.306)
RMSD bonds (Å)	0.006
RMSD angles (°)	1.00
Ramachandran favoured/allowed/forbidden	98.5 / 1.5 / 0
Unfavourable rotamers (%)	0.19
Molprobity Score (Chen et al., 2010)	1.36

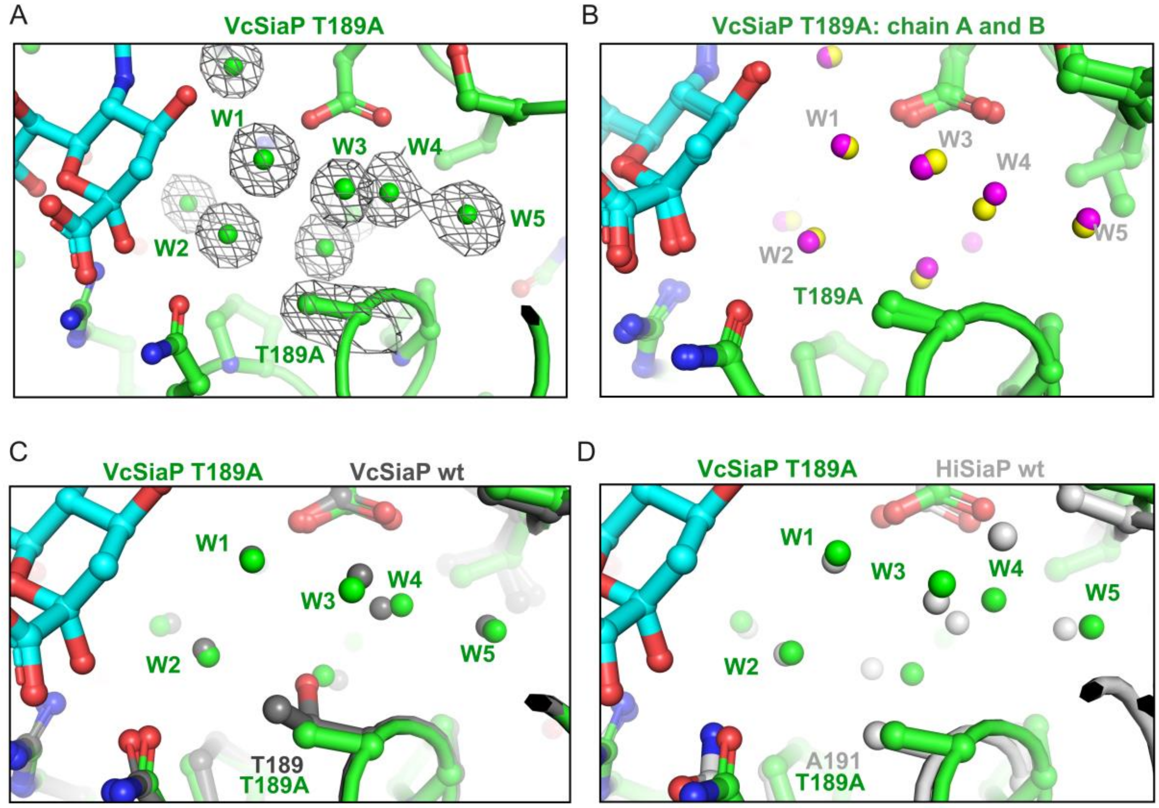


Figure 3-16: Structural details of VcSiaP T189A. A) Mutant VcSiaP T189A (green) with detailed view of the mutated region. The electron density of the mutated amino acid and water molecules in the surrounding are illustrated as grey meshes ($2mF_0$-DF_c, sigma level = 1.0). The water is shown as green spheres and the substrate as cyan ball-and-stick model. **B)** Alignment of the two protein chains in the asymmetric unit from VcSiaP T189A. The water molecules of both chains are highlighted in different colours. **C)** Detailed structural view of the mutated region from protein VcSiaP T189A (green) and aligned structure of VcSiaP wildtype, substrate bound structure (dark grey) from the previous chapter. **D)** Same as C) but with alignment of HiSiaP wildtype, substrate bound protein (grey, PDB-ID: 3B50 (Johnston et al., 2008)).

3.3 Investigation of the open-closed dynamic of VcSiaP

In addition to the characterization of sialic acid binding with thermodynamic and structural methods, the dynamics of the binding process was also investigated with smFRET. The smFRET experiments were performed in cooperation with Christian Gebhardt and Marijn de Boer from the group of Thorben Cordes and they provided the Figure 3-17 and Figure 3-18 (University of Groningen and University of Munich). As already shown with the new substrate bound VcSiaP structure, the binding of substrate involves a large conformational change from a substrate free open (apo) to a substrate bound closed conformation (holo) (Figure 3-3 D). This process was already investigated by our group with PELDOR/DEER distance measurements in frozen solution (Glaenzer et al., 2017). These experiments indicated a clear substrate binding dependency of the conformational change. In the present work, smFRET spectroscopy was used to study the mechanism of the conformational change of the protein by real-time measurements (milliseconds) and under more physiological conditions, meaning in solution and at room temperature. For these experiments, the cysteine double mutant VcSiaP Q54C L173C was used. This mutant has been used in previous studies for several distance measurements and revealed a good combination of label positions, which reports on the conformational change of the protein (Glaenzer et al., 2017).

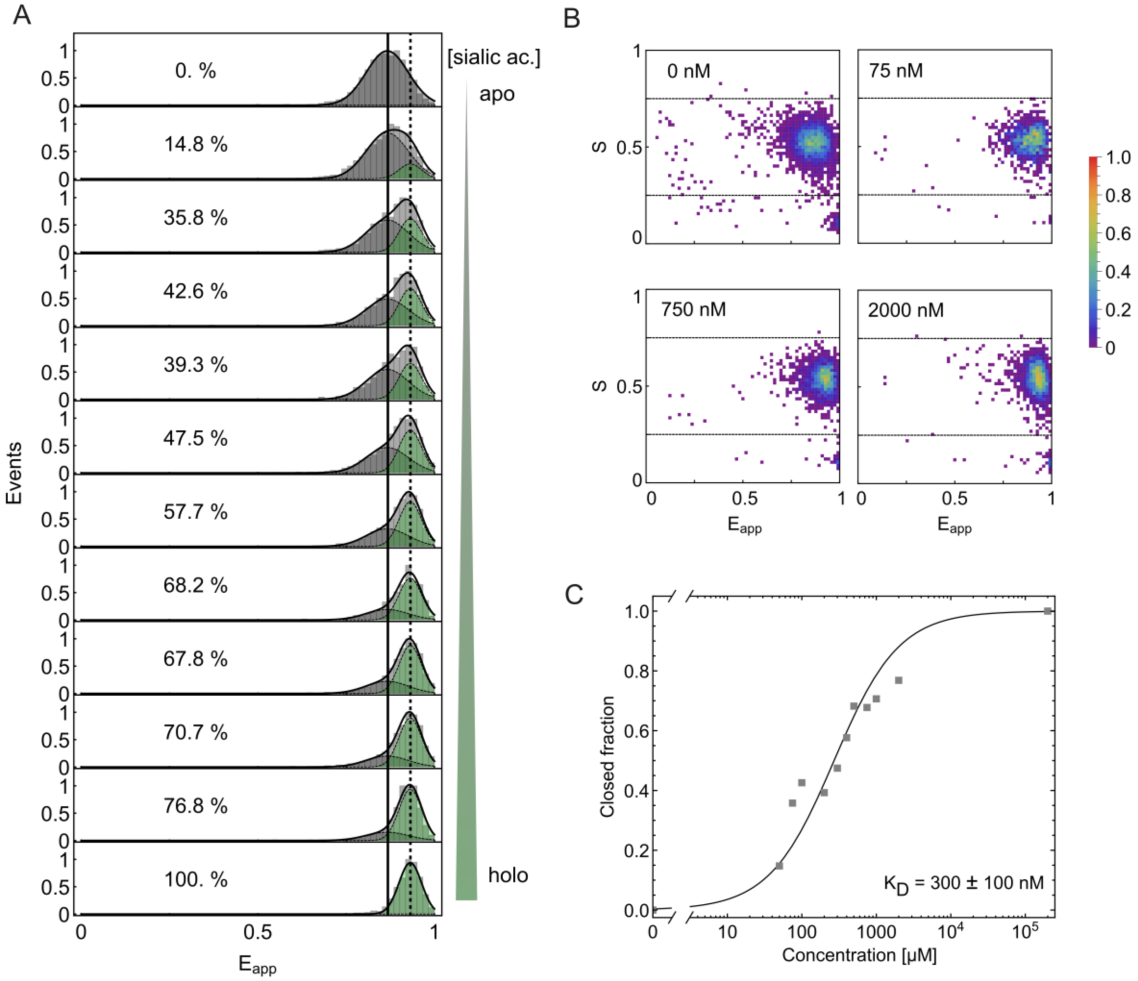


Figure 3-17: Steady-state analysis of sialic acid binding by VcSiaP. A) Apparent FRET efficiency histograms of VcSiaP Q54C/L173C labelled with AlexaFluor 555 and AlexaFluor 647 and supplemented with different concentrations of sialic acid (0-200 µM, top to bottom). The FRET efficiencies for the open and closed state are visualized by a dashed black bar and a black bar, respectively. The fraction of closed conformation proteins (r_C) was determined via fitting of a gaussian curve for each state, the open (grey) and closed (green) state. **B)** Example 2D ES-histograms for four different sialic acid concentrations from A). **C)** Binding curve with fractions of the closed state from A) at different substrate concentration. The figure was provided by the group of Thorben Cordes and modified from (Peter et al., 2021).

The protein was purified with the standard procedure, the final SEC run and SDS-gel are shown in Figure SI-15. For smFRET studies, the mutant was labelled with maleimide-functionalized fluorophores AlexaFluor 555 (donor) and AlexaFluor 647 (acceptor). Overall, the labelling efficiency was higher than 90% with ~1.83 fluorophores per protein and a yield of ~49% of donor-acceptor labelled protein.

In the first experiment, the function of the protein was analysed to confirm the intact substrate binding mechanism of the fluorophore labelled protein. Such an experiment should identify potential disrupting influences of the fluorophores on the protein states, as observed in the methodical comparison in Chapter 2. For this, a titration experiment was performed with several substrate concentrations and the populations of the open and closed state were determined from the resulting histograms (Figure 3-17 A and B). A clear tendency from lower to higher FRET energy populations, depending on the substrate amount, was observed. This indicated the expected conformational change of the protein with a decrease of the distance between the two labels. The fraction of the closed state in combination with the substrate concentration was used to generate a binding curve of VcSiaP and sialic acid from the smFRET measurements (Figure 3-17 C). From this, the substrate affinity was determined to be 300 ± 100 nM, which was in a good agreement with the ITC-based value of 203 ± 17 nM (Chapter 3.1.2). In the error margin of both methods, the attachments of the fluorophores had no effect upon substrate binding. Further, the detected change of the distance between the fluorophores was clearly coupled to the binding event of sialic acid by VcSiaP.

After this initial characterization, the binding mechanism of sialic acid by VcSiaP was studied with smFRET on surface-immobilized proteins. In contrast to the diffusion-based approach, the proteins were immobilized via binding of their His$_6$-tag to a His-specific antibody, which was again fixed via biotin-neutravidin interaction to a glass surface (Figure 3-18 A, Chapter 9.2.4). Afterwards, the conformational change was analysed on single immobilized proteins through detection of changes in the FRET energy on a real-time scale. To characterize substrate binding, the experiment was performed under apo conditions, 500 nM sialic acid (around K_D value) and 1 mM sialic acid (saturated concentration). The FRET-states that were observed with the immobilized protein in the absence and presence of saturated substrate concentration indicated populations for the open and closed conformation and were in accordance with the results from the diffusion-based experiments (Figure 3-18

B and C). For the set-up with substrate concentration around the K_D, a transition between low- and high-FRET efficiency states was observed, fitting to the states that were detected without and with substrate saturation, respectively (Figure 3-18 B, middle). In detail, the high-FRET efficiency for 500 nM substrate was determined to be 0.895 ± 0.010 and the low-FRET efficiency to be 0.791 ± 0.012, while the substrate saturated experiment was determined to be 0.898 ± 0.003 and the substrate free approach to be 0.807 ± 0.003. Therefore, the states that were observed for the 500 nM substrate experiment represent the transition between the open (low FRET energy) and closed (high FRET energy) conformation of the labelled VcSiaP proteins.

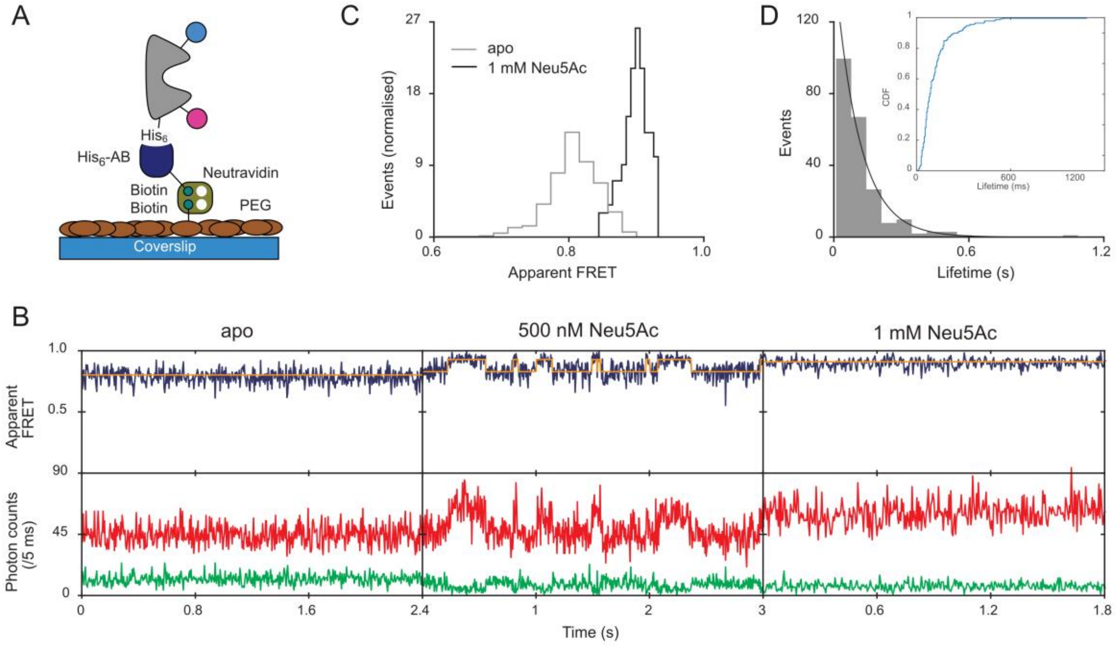


Figure 3-18: Dynamic smFRET studies on immobilized VcSiaP. A) Schematic of the immobilizing strategy of VcSiaP (Q54C L173C) on a surface by using the N-terminal His$_6$-tag. The labelled protein (grey) with two fluorophores (magenta and blue) was attached to a biotinylated His-antibody (dark blue). The antibody was coupled to neutravidin which was previously fixed to the biotin-PEG surface. **B)** Example of detected fluorophore trajectories of immobilized protein from A), labelled with AlexaFluor 555 and AlexaFluor 647 with variations in the substrate concentration as mentioned in the figure. The apparent FRET efficiency is shown at top in blue and the photon counts at the bottom with donor in green and acceptor in red. The orange line in the apparent FRET shows the determined trajectories from the Hidden Markov Model (Rabiner, 1989). **C)** Histograms of the apparent FRET efficiency of all trajectories without (grey, apo) and with substrate (black, 1 mM Neu5Ac, holo). **D)** Histogram and cumulative distribution function (CDF) for lifetime of the closed conformation with high FRET efficiency from trajectories with 500 nM sialic acid (45 traces). The histogram distribution was analysed by an exponential fit (black line). The figure was provided from the group of Thorben Cordes and modified from (Peter et al., 2021).

The trajectories that were observed and assigned to the different states were processed by a Hidden Markov Model (HMM) (Rabiner, 1989). Using this model, the lifetimes of the closed conformation were extracted, plotted and the average lifetime was calculated by an exponential fit (Figure 3-18 D). Thereby, the lifetime of the closed conformation was determined to be 122 ± 16 ms, which means that the substrate bound, closed conformation needed approximately 120 ms to re-open and released the sialic acid from the binding site.

Overall, the trajectories from Figure 3-18 confirmed the previous study in frozen solution by PELDOR spectroscopy (Glaenzer et al., 2017). Considering the detection limit of smFRET in this set-up (around 5 ms), VcSiaP without substrate only exists in the open conformation and the closing just occurs after addition of the substrate. Overall, the smFRET results of the substrate binding by the TRAP transporter SBP VcSiaP support an induced-fit binding mechanism, also known as Venus flytrap mechanism (Mao et al., 1982, Walmsley et al., 1992). The open state of VcSiaP is stabilized without substrate and just by addition of the substrate, a transition from the open to closed conformation is induced. With low substrate concentrations, a re-opening of the protein was observed for the first time, indicating presumably a releasing of the substrate from the protein binding cleft.

3.4 Discussion and conclusion

In this chapter, the high-resolution structure of the substrate bound SBP from the sialic acid TRAP transporter of *V. cholerae*, VcSiaP, was described for the first time. Thereby, the new structure completes the basic structural information about VcSiaP. Furthermore, it is only the second SBP from the class of sialic acid TRAP transporters that is described in the two functional conformations. Based on this new structure, an identical change between the open and closed conformation was observed as for the other TRAP SBP HiSiaP (Johnston et al., 2008, Müller et al., 2006). As previously postulated with experiments on the binding behaviour and from the sequence similarity, the observation of the same structural behaviour supports an equivalent functional mechanism of sialic acid TRAP transporter SBPs (Glaenzer et al., 2017, Johnston et al., 2008, Mulligan et al., 2011).

Despite the identical overall structures of the sialic acid TRAP transporter SBPs, a previous study determined a lower affinity for VcSiaP to sialic acid of a factor 10 compared to three other sialic acid SBPs (Gangi Setty et al., 2014). At that time, a reason for this observation could not be given due to the missing structure of the substrate bound conformation. The successful description of this protein structure in the first part of this chapter allowed an investigation of possible reasons for the different binding affinities. However, the initial repetition of the binding characterization for two wildtype proteins, HiSiaP and VcSiaP, revealed a slightly smaller difference for the binding affinities compared to the previous study from Gangi Setty et al. (Gangi Setty et al., 2014).

These different results might be caused by slight variations in the protein samples or experimental procedures, for example purification processes with more or less impurities, uncertainties in protein concentration determination or a different processing and analysis of the ITC results. Unfortunately, the previous study showed neither the experimental data of the ITC measurements nor steps from the purification process or the purity of the final protein sample (Gangi Setty et al., 2014). Another reason for the differences might be the affinity tag which was not cleaved off in the previous study, according to the described procedure (Gangi Setty et al., 2014). From recent work in our group with a new VcSiaP crystal structure it is known, that VcSiaP can bind short peptides such as the affinity tag (Peter et al., 2021). Consequently, a strong binding peptide might change the observed substrate binding affinity. Further, the expression of the proteins in the previous study was performed in LB medium and not in M9-minimal medium as in the present work. Investigations on the open and closed states of VcSiaP with PELDOR spectroscopy revealed a proportion of around 20% of the closed conformation after protein expression in LB medium and 0% by substitution with M9-minimal medium (personal communication with Janin Glänzer, University of Bonn). The unintended contamination of the protein with the substrate reduces concentration of unbound protein which is available for substrate binding in an ITC experiments and by this, influence the determined binding parameters.

Nonetheless, the binding affinities for both SBPs were slightly different in all studies with the same tendency of weaker binding for VcSiaP. With the new substrate bound VcSiaP structure, an altered water network in the substrate binding cleft was presumed to be responsible for the lower binding affinity. A detailed mutational binding study with ITC experiments and crystal structures of regions, that were postulated to be responsible for the different affinities of the sialic acid SBPs, revealed functional information about the substrate binding:

(I) The interactions between the protein residues and sialic acid, mediated by the water network, are highly complex and strongly interdependent. For example, the mutations Q9N or G71Q in VcSiaP revealed side chain conformations that were not obtained for similar native residues in the HiSiaP wildtype structure. Additionally, the mutated residues partially caused alternative conformations of further residues, for example Q151 in mutant M81A G71Q. Moreover, even a direct interaction of an amino acid residue with the substrate as for Q9N guaranteed no conservation of the conformation. These observations support that the mutated amino acids and presumably all other residues in the environment are in a defined interaction with each other. Interruptions of these interactions through mutations result in a significantly lower binding affinity. A further indication of the complex interaction is the observation that for the only mutant with an increased binding affinity, VcSiaP T189A, no obvious structural change was identifiable. One possible explanation for this stronger affinity is a different stability of the waters in this region, that changes the kinetics of the substrate binding and by this, facilitates the initial interaction between protein and substrate or the conformational change of the protein. To further validate such an effect, surface plasmon resonance (SPR) experiments with the stronger binding T189A mutant might be helpful to study the kinetics of the binding process.

(II) The previous observation of a water network in the substrate binding site of HiSiaP that has a great influence on the binding affinity of sialic acid, was also confirmed for VcSiaP (Darby et al., 2019). A good example is the mutagenesis of VcSiaP A43 to a serine in binding region II, which can only interact with the water network and not directly with the substrate. Even if no crystal structure could be solved for this mutant during the present work, the ITC experiment revealed a significant decrease of the binding affinity which is presumably caused by an interruption of the water network. A similar observation was described previously through mutation of alanine 11 to a charged asparagine in HiSiaP, which caused also an immensely decrease of the binding affinity (Darby et al., 2019). In both proteins, VcSiaP and HiSiaP, the water network is thereby identified as essential for a high substrate affinity. Interestingly, the second shell mutants were still able to bind sialic acid and substrate bound crystal structures were successfully determined for the most mutants, for VcSiaP in the present work and for HiSiaP in a previous study (Darby et al., 2019). In contrast, mutants with changed residues that interact directly with the substrate, such as R145 and R125, were classified as non-binders and no substrate bound crystal structures of these mutants were described until today (Fischer et al., 2010, Glaenzer et al., 2017, Johnston et al., 2008). These differences suggest that the water network in the binding cleft contributes to the previously observed high substrate affinity of TRAP SBPs and that residues in interaction with this network are used for “fine-tuning” of the substrate binding process (Müller et al., 2006). Speculatively on an evolutionary view, this has the advantage that for a new, slightly changed substrate, the amino acids with a direct substrate interaction do not need to be modified. Therefore, the evolutionary modification of second shell interacting residues prevents the formation of non-binding versions of the SBPs.

Overall, even if the results revealed a highly detailed view about the binding mechanism of the substrate on a molecular level, the reasons for the different binding affinities were difficult to interpret and probably more variations than the presented mutants are required. As observed for region I and II, the mutagenesis of single amino acids can easily interrupt the substrate binding processes and a higher stabilization of the substrate was hard to achieve. One possible alternative which might be faster than the production of more mutants and that can directly address single water molecules in the substrate binding cleft, is a computational calculation of the thermodynamic binding parameters. In such calculations, certain water molecules or defined amino acid residues can be changed or deleted and the accompanying influence on the thermodynamic parameters can be calculated. The comparison of thermodynamic parameters from the wildtype or mutant protein structures with the *in silico* modified models can possibly reveal a defined impact of certain molecules or side chains on the substrate binding and fulfil the understanding of the substrate-protein interaction. Further, molecular dynamic simulations (MD) can be used to study a changed kinetic of the substrate binding mechanism or conformational change of the protein, especially for the stronger binding mutant VcSiaP T189A.

In the second part of this chapter, the dynamic of the conformational change from TRAP transporter SBPs was investigated for the first time with smFRET in solution, by tracking the time-resolved open-closed transition of single proteins. Contrary to some SBPs of the ABC transporter and in agreement with previous studies, no substrate-free closed conformation was observed for the TRAP transporter SBP under these conditions (de Boer et al., 2019a, Glaenzer et al., 2017, Tang et al., 2007). Importantly, the smFRET experiment has a time resolution of around 5 ms and fast transitions below this resolution cannot be detected. With the smFRET experiments, a rearrangement from the closed to the open conformation was observed for the first time for TRAP transporter SBPs and revealing again the substrate-dependency of the conformational change (Glaenzer et al., 2017). As mentioned before, all these observations strongly suggest an induced-fit binding mechanism for the TRAP transporter SBPs (Mao et al., 1982, Walmsley et al., 1992). Similar to ABC transporters with periplasmic SBPs, this mechanism is very likely a selection factor for the membrane domains to discriminate between substrate free and substrate bound protein (Doeven et al., 2008). This selection will increase the efficiency of the transporter system, since the binding to the membrane domains is exclusively possible for the substrate bound protein and cannot be occupied by the substrate free state.

Chapter 4

Reconstitution of TRAP transporter membrane domains into membrane mimetic systems

Preamble:

The reconstitutions and characterizations of protein-bounded (MSP) nanodiscs were done with the help of Peer Depping during his Master thesis.

4.1 Working with membrane proteins in aqueous solution

Analysis of genomes from all types of organisms, prokaryotes, archaea and eukaryotes, showed that on average 20-30% of the open-reading frames (ORF) encode for a membrane protein (Wallin and Heijne, 1998). However, the current knowledge about membrane proteins and especially the number of known membrane protein structures do not reflect this high proportion compared to all other proteins. Until today, 153194 protein structures (protein-only structures) are published in the protein data bank (PDB) (www.rcsb.org, access date: 17.02.2021), including 3604 membrane proteins or based on the mpstruc database classification, 4686 membrane protein structures (https://blanco.biomol.uci.edu/mpstruc/, access date: 17.02.2021). The mpstruc database also indicated that only 1220 of the 4686 membrane protein structures are unique structures. Overall, the data revealed a proportion of less than 3% of membrane protein structures on all known protein structures, illustrating the underrepresentation of this type of proteins in terms of structures. Corresponding to estimations, around 1700 unique structures are necessary to classify representatives of each membrane protein family (Oberai et al., 2006, White, 2009). Methodical developments, especially the great improvement of single particle cryo-EM, revolutionized the field of structural biology and led to a high increase of new determined membrane protein structures during the last years (Chapter 7.1) (Hendrickson, 2016). While the description of a single structure is very helpful, membrane transporters rely on conformational changes for transport of a substrate across the cell membrane (Chapter 1.3). Therefore, several structures in different conformations are usually needed to fully explain their function. Accordingly, optimized methods that should improve and simplify the solubilization and handling of membrane proteins in solution are constantly developed. Methods that were used for the TRAP transporter membrane domains in the later part of this chapter are introduced in the following paragraphs.

Membrane proteins are often solubilized and extracted from their native membrane with detergents, by forming stabilizing micelles around the hydrophobic area of the protein (Garavito and Ferguson-Miller, 2001). This method was used for many years as golden standard to receive structural and functional information about membrane proteins and nowadays, many detergents with different properties are commercially available (Stetsenko and Guskov, 2017). However, the detergent belt around the membrane protein differs considerably from the native environment of a lipid bilayer (Zhou and Cross, 2013). Even if membrane proteins can have the same structure after extraction in detergent micelle and their native lipid environment (Broecker et al., 2017), there are many systems known in which the structure, function and conformational dynamic depend on certain protein-lipid interactions (Gupta et al., 2017, Laganowsky et al., 2014, Martens et al., 2016, Martens et al., 2018, Sanders et al., 2018). For a better imitation of the native membrane protein environment, conservation of protein-lipid interactions and improvement of general handling for faster and easier biochemical characterization, several membrane mimetic model systems were developed.

A large step towards this goal was achieved through the development of self-assembling nanodiscs (Bayburt et al., 2002). In such an assembly, two membrane scaffold proteins (MSPs) surround and separate the hydrophobic area of a small lipid bilayer from the aqueous solution. Membrane proteins can be reconstituted into this lipid bilayer, which generate a system that has a high similarity to the native environment of the membrane protein and allows a handling in solution (Bayburt and Sligar, 2003). To increase the application of these protein-bounded nanodiscs, several constructs of MSP for different sizes of nanodiscs and diverse lipid mixes, including a native like *E. coli* lipid mix, are generated and tested (Bayburt and Sligar, 2010, Denisov et al., 2004). These developments led to a widespread application of the MSP-nanodiscs in structural and functional studies of membrane proteins (Denisov and Sligar, 2016). For example, the structure of a fully membrane-embedded secondary transporter in MSP-nanodiscs was recently determined (Arkhipova et al., 2020).

A disadvantage of the protein-bounded nanodiscs is that the membrane proteins need to be extracted from their native membrane with classical detergents, prior to the reconstitution procedure (Ritchie et al., 2009). Further, for the case of membrane proteins that have essential interactions with certain lipids as mentioned above, the reconstitution into an artificial bilayer with a different lipid composition does presumably not fulfil this interaction and may have a disruptive influence on the function of the protein. A promising solution that avoids the extraction of membrane proteins with detergents is the use of a styrene maleic acid (SMA) polymer, which forms nanodiscs like a “cookie-cutter” directly from the native cell membrane, including the membrane proteins (Knowles et al., 2009, Lee et al., 2016a). The polymer is the product from a polymerization of styrene and maleic anhydride, a 2:1 or 3:1 ratio is standardly used for solubilization of membrane proteins (Figure 4-8 A) (Dörr et al., 2016). The solubilization mechanism of the native membrane is presumably based on an initial hydrophobic interaction between the polymer and the lipid head groups, followed by an insertion of the polymer into the hydrophobic core (Scheidelaar et al., 2015). This process destabilizes the membrane and forms the nanodiscs with a belt of polymers. The resulting polymer-bounded nanodiscs have a size comparable to the protein nanodiscs of around 10-15 nm and were already used in several structural and functional investigations (Dörr et al., 2016). For example, the polymer nanodiscs were successfully used for structure determination of membrane proteins with X-ray crystallography or EM (Broecker et al., 2017, Postis et al., 2015, Qiu et al., 2018, Sun et al., 2018). One important observation during investigations of proteins in polymer-bounded nanodiscs was an enrichment of native anionic lipids in the nanodiscs compared to the average lipids in the membrane (Dörr et al., 2014, Prabudiansyah et al., 2015).

A similar behaviour of nanodisc formation was observed for the related diisobutylene maleic acid polymer (DIBMA) (Oluwole et al., 2017a, Oluwole et al. 2017b). The usage of this polymer has the advantage of a lower UV/Vis absorption at protein-detectable wavelength of around 260-280 nm and a higher acceptance of divalent cations in the buffer solution compared to SMA extractions (Oluwole et al., 2017a). However, even though the first reconstitution attempts and physical characterisations of DIBMA nanodiscs are promising, there are currently only a few described applications in the literature (Adão et al., 2020, Gulamhussein et al., 2020, Oluwole et al., 2017a). Furthermore, due to the encouraging results of SMA and DIBMA polymers, several more variations of polymers are recently described (Esmaili et al., 2020, Fiori et al., 2020, Marconnet et al., 2020). However, the experimental experience with these systems is overall still sparse and advantages and compatibilities to a broad spectrum of membrane proteins, as well as bioanalytical techniques, are unknown and have to be examined in the future.

In summary, the use of polymer nanodiscs showed encouraging and promising results on the way to a native membrane mimetic system. On the one hand, the nanodiscs have a high potential to conserve the intact structure and function of the membrane proteins and on the other hand, represent an easy to use, manageable system which can be used in bioanalytical studies. In addition to the use of classical detergents, the reconstitution of TRAP transporter membrane proteins in protein-bounded and polymer-bounded nanodiscs were examined in the following chapter.

4.2 Reconstitution of VcSiaQM and HiSiaQM into detergent micelles

The following chapter presents the initial experiments for production of the two membrane domains VcSiaQM and HiSiaQM of sialic acid TRAP transporter from *V. cholerae* and *H. influenzae*. This includes protein expression, extraction from the cell membrane with classical detergents and protein purification. Several important optimizations of each step are presented to establish a working routine for high protein yields and a good purity of the protein solution.

4.2.1 Expression, extraction and purification of VcSiaQM

The production of detergent solubilized VcSiaQM was started with test-expressions in C43 (DE3) *E. coli* cells. The cells were grown in LB medium and the protein yield was detected with an SDS-PAGE and western blot (His-tag antibody). The expressions were performed with two different temperatures (27 °C and 37 °C), with different concentrations of inductor L(+)-arabinose (50 mg, 100 mg or 250 mg per litre medium) and gel samples were prepared after different expression times (1, 2 and 3 hours). Figure 4-1 illustrates the result of the test-expression with the optimized inductor concentration of 50 mg/L. The protein yield of VcSiaQM was presumably too low for a detection of the protein in the SDS-PAGE, as no defined changes of a protein band intensity before and after induction of the cultures were visible (Figure 4-1 A). But the corresponding western blot indicated a clear band of a His-tagged protein below 25 kDa protein marker band (Figure 4-1 B). This band presumably represents the His-tagged Q-domain (M_W=22.8 kDa) of the TRAP transporter and is in accordance with previous published results (Mulligan et al., 2012). The larger M-domain was not tagged and was hence not detectable in the western blot. However, previous studies detected a strong complex between both domains and the simultaneous expression and presence of both domains in solution was assumed (Mulligan et al., 2012). Overall, the samples at 37 °C indicated a constant protein amount between 1- and 2-hours expression, while for the 25 °C samples the protein amount was slightly increased between 1 and 2 hours (Figure 4-1 B). The total protein amount for the two different temperatures after 2 hours expression was nearly equal. Interestingly, the gel samples that were prepared after 3 hours expression showed no protein bands at all, independently from the temperature that was used. This was possibly caused by a low stability or fast digestion of VcSiaQM in the used *E. coli* cells.

Finally, 37 °C and one hour were chosen as expression parameters and 10 L LB medium were used in a standard experiment for a satisfying protein yield.

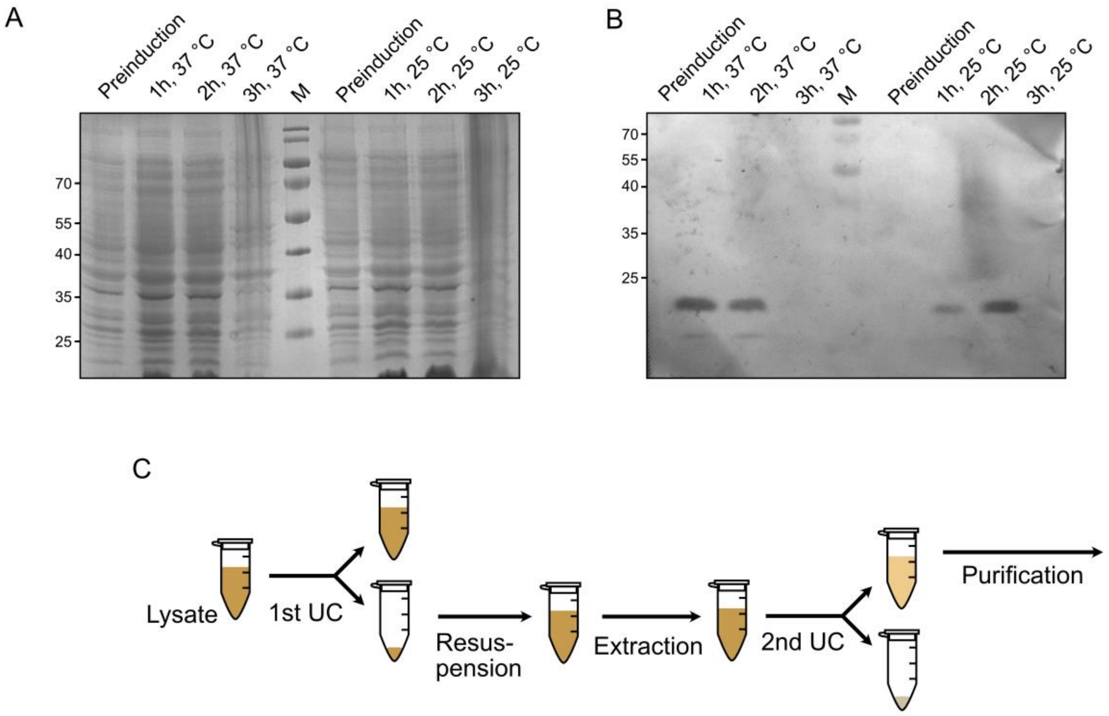


Figure 4-1: Expression and purification of VcSiaQM. A) SDS-PAGE with samples from *E. coli* culture test-expression of VcSiaQM in LB medium. A preinduction sample and three samples with different expression times are shown for two different expression temperatures. **B)** Same samples as A) but after western blotting of the samples with a His-tag specific antibody. **C)** Optimized workflow for preparation of the *E. coli* cell membranes for extraction of VcSiaQM with DDM detergent. The parameters of each step are presented in the method Chapter 9.1.4.2.

A schematic of the extraction procedure is illustrated in Figure 4-1 C. The parameters of each step are described in method Chapter 9.1.4.2. The detergent n-dodecyl-β-D-maltoside (DDM) was used for extraction of the membrane protein from the native membrane. Extraction with the alternative detergent Fos-Choline-14 was also investigated but resulted in a lower yield of target protein and a noticeable increase of the impurities and was therefore not used for further experiments (Figure SI-16).

After extraction of the membrane proteins with DDM, the VcSiaQM protein was further purified via chromatography. During the first purifications of the membrane domains with Ni^{2+}-AC, no VcSiaQM protein bands were found in the corresponding SDS-PAGEs.

Nonetheless, the protein solution was further purified and loaded onto a SEC column. The UV/Vis trace indicated a peak of aggregated protein at around 8 mL and a broad protein peak with two maxima at around 12 mL elution volume (Figure 4-2 A). In some purifications only a protein peak with one maximum was observed, while the reason for this observation remained elusive (for example Figure SI-16). The SDS-PAGE of the peak fractions showed only very low-intensity protein bands (Figure 4-2 B). However, the corresponding western blot with His-tag antibody for the same samples displayed an unambiguous band for the tagged Q-domains (Figure 4-2 C). The single protein band in the western blot at higher molecular weight could not be identified and maybe represented a fraction of VcSiaQM with a changed gel-diffusion behaviour, possibly due to partial unfolding.

To further study the discrepancy between the obvious protein signal in the SEC elution profile and the weak protein band signal on the SDS-PAGE, the SDS-PAGE was repeated without heating of the samples before application on the gel. Indeed, this SDS-PAGE indicated clear protein bands that were not observed before (Figure 4-2 D). The intense band between 15-25 kDa corresponded to the detected band in the western blot and thereby represented the His-tagged Q-domain. The second intense band around 35 kDa indicated presumably the M-domain (M_W=45.46 kDa) of the VcSiaPQM TRAP transporter. The molecular weight shift on the SDS-gel compared to the protein standard is a familiar effect for membrane proteins, which is likely based on an incomplete denaturation of the membrane protein with the detergent SDS (Rath et al., 2009). The protein band for the M-domain has the same intensity pattern in the SEC peak as the Q-domain and the size fits to previously published data (Mulligan et al., 2012). The identical elution behaviour of these two proteins with different masses was likely caused by a stable complex formation between the Q- and M-domain in solution, as previously described for VcSiaQM (Mulligan et al., 2012). On the SDS-gel, denaturing conditions cause the separation of the complex and two protein bands were visible.

Consequently, derived from the observation in this chapter, all SDS-gel samples with the membrane domains in following paragraphs and chapters were not heated. After the optimized gel sample preparation, a new purification trial was performed with a protein elution from Ni^{2+}-AC by using a stepwise increasement of the imidazole concentration (Figure SI-16). The SDS-PAGE of the elution fractions indicated a lot of impurities and only weak protein bands that correspond to VcSiaQM, already in the elution fraction with 31.2 mM imidazole.

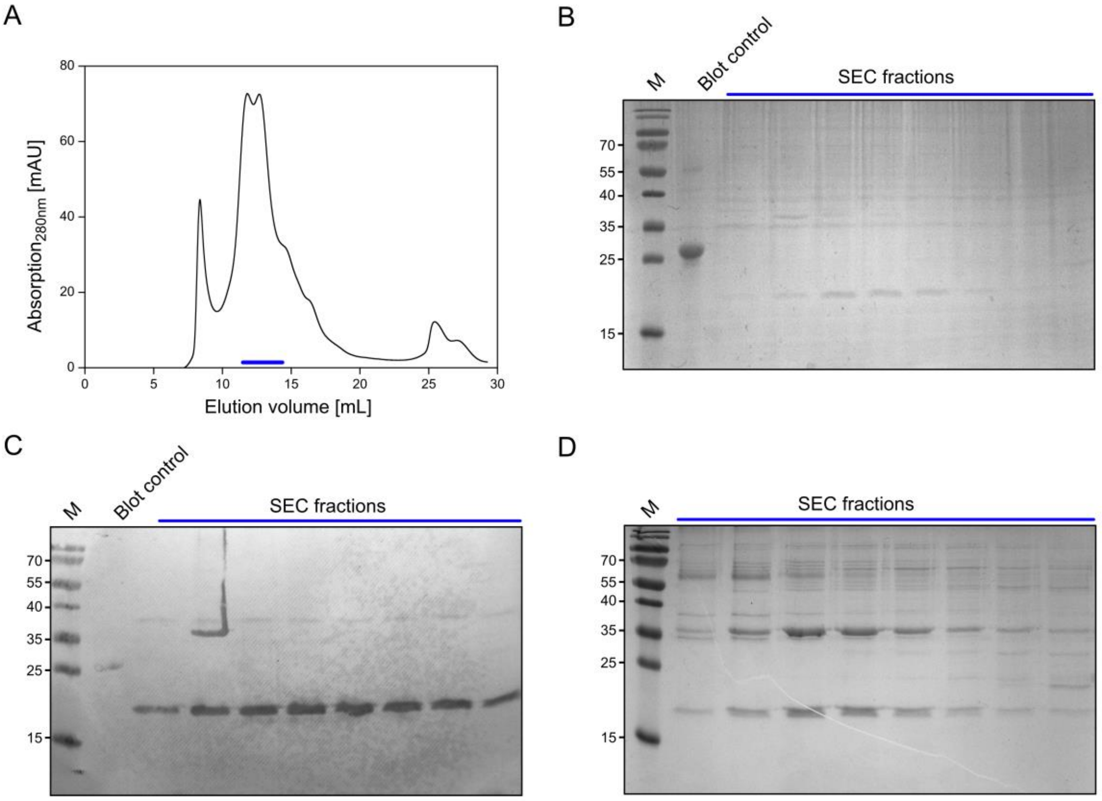


Figure 4-2: Purification of VcSiaQM. A) SEC of concentrated Ni^{2+}-AC elution sample with absorption at 280 nm on an SD 200 10/300 column (standard detergent buffer). The fractions for SDS-PAGE samples are marked in blue. **B)** SDS-PAGE of SEC fractions from the whole protein peak in A), as indicated by the blue line. His-tagged TEV protease was used as western blot control. **C)** Same samples as B) but after western blotting and detection with an His-tag antibody. **D)** SDS-PAGE of SEC fractions from the whole protein peak in A) but without standardly heating of the samples before loading onto the gel.

For optimization of the protein production, the affinity of the protein to the nickel resin was increased to enable a low-imidazole concentration washing step during Ni^{2+}-AC to remove impurities before elution of the target protein. Therefore, the His-tag on the protein was extended with QuickChange mutagenesis by insertion of further histidine residues. Although, the primer was designed for insertion of 4 histidine residues, only a construct with 8 additional residues was produced, maybe through unexpected binding of the primers in the polyhistidine region. Nonetheless, this His_{14}-construct was expressed and a Ni^{2+}-AC was performed with an identical imidazole gradient as before. The eluted fractions were analysed on a western blot to receive a high sensitivity for detection of the target protein VcSiaQM (Figure 4-3 A). The maximum of protein intensity was clearly detected in the fraction with 125 mM imidazole. This shift of the protein to elution at higher imidazole concentrations, allowed an increase of the imidazole concentration in a washing step of the Ni^{2+}-AC and thereby a potentially decrease of impurities.

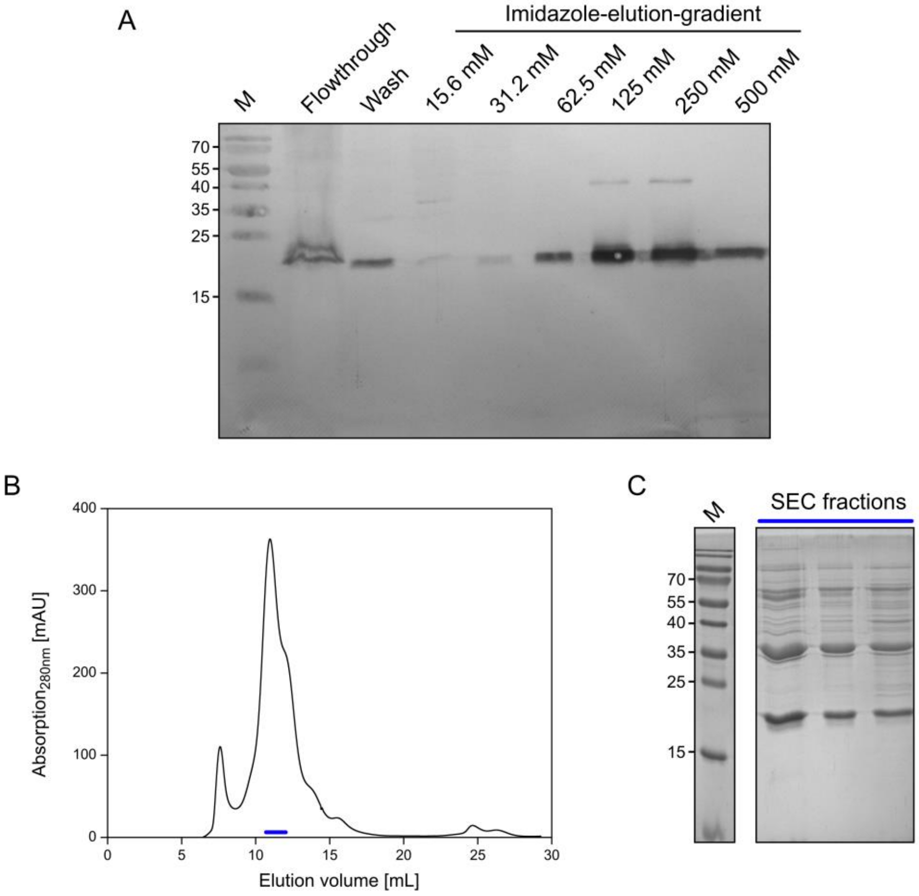


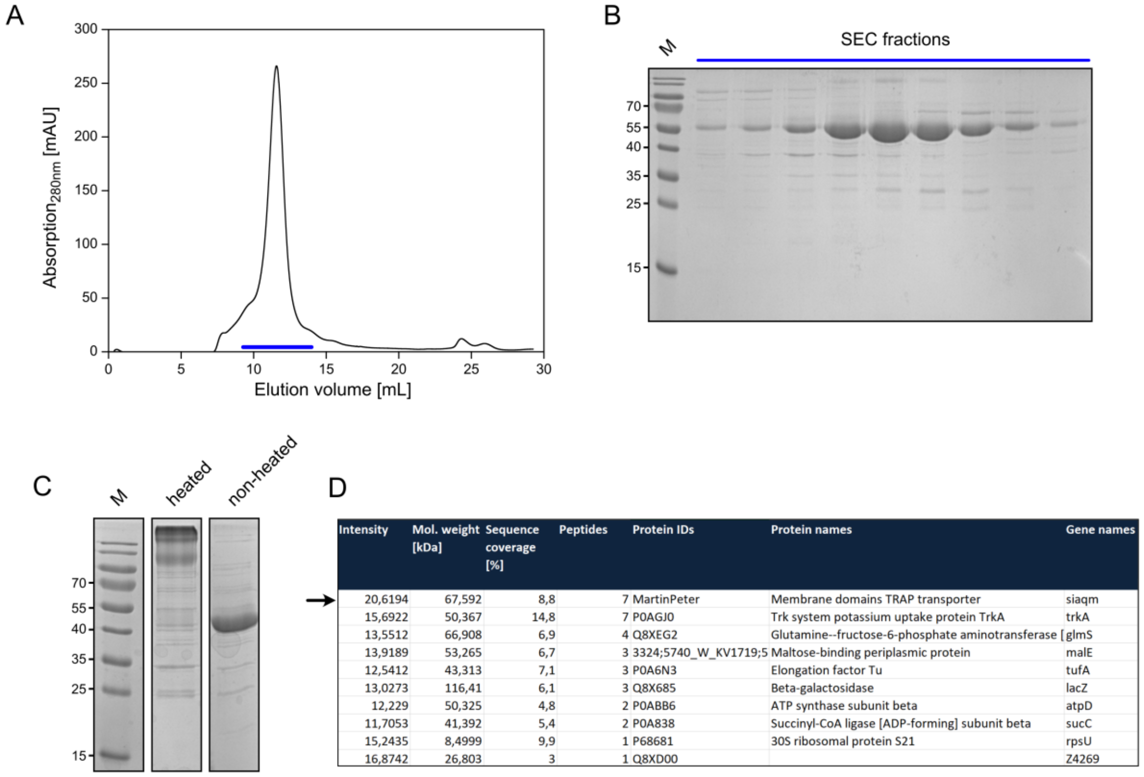
Figure 4-3: Optimizations of VcSiaQM purification. A) Western-Blot with samples from different steps of a Ni^{2+}-AC with His_{14}-tagged VcSiaQM and detection of His-tag antibody. **B)** SEC of concentrated Ni^{2+}-AC elution solution with absorption at 280 nm on an SD 200 10/300 column (standard detergent buffer). The fractions for SDS-PAGE samples are marked in blue. **C)** SDS-PAGE of SEC fractions from the main peak in B), as indicated by the blue line.

Additional purifications with variations of further parameters, such as optimization of the amount of detergent during extraction and varied extraction time, led to a SEC elution profile in Figure 4-3 B. The final optimized parameters for expression, extraction and purification are described in detail in method Chapter 9.1.4.2. Overall, the SDS-PAGE from the SEC fractions showed intense bands for the Q- and M-domains but also impurities that could not be removed during purification (Figure 4-3 C). The final protocol for production of VcSiaQM resulted in a yield of around 1 mg protein from 10 L LB medium.

For all VcSiaQM purifications, an additional protein band in the SDS-gels at around 55-70 kDa were observed, with varied intensity between different purification trials (Figure 4-2 D and Figure 4-3 C). The protein band was investigated with specific detection of His-tagged protein in a western blot to analyse if the protein band results from the unseparated complex of the QM membrane domains. However, the western blot revealed only a weak signal around 35-40 kDa which might be caused by the QM-complex, while the protein band from the SDS-PAGE at around 55-70 kDa was not visible in the western blot and are therefore assigned to a contamination (Figure SI-17).

4.2.2 Expression, extraction and purification of HiSiaQM

Based on the successful purifications of VcSiaQM in the last chapter, the membrane domains of the related sialic acid transporter from *H. influenzae,* HiSiaQM, were also expressed and purified. Here, the optimized parameters and experience from the VcSiaQM experiments, for example detergent concentration, extraction time and useful culture volumes, were directly used for this His_{10}-tagged protein. An elution-profile of a SEC with HiSiaQM protein after Ni^{2+}-AC is illustrated in Figure 4-4 A and showed a defined single peak at around 12 mL elution volume. In comparison with the SEC runs from VcSiaQM, only a small amount of aggregated protein at around 8 mL was detected. The corresponding SDS-gel of the main peak fractions indicated an overall clean solution with only few impurities and a major intense protein band at around 55 kDa protein standard (Figure 4-4 B). In contrast to VcSiaQM, the Q- and M-domains of HiSiaQM are fused together to one peptide chain with a postulated additional transmembrane helix (Figure 1-8) (Kelly and Thomas, 2001). The detected band correlated well with a purification from a previous publication with a HiSiaQM band at around 47.5 kDa standard (Mulligan et al., 2009). Again, the protein samples for the SDS-gel were not heated before they were loaded onto the gel. An example of a heated sample, compared to the same sample without heating, is shown in Figure 4-4 C. Similar to VcSiaP, the heated sample indicated an undefined, smeary protein fraction in the gel, without a clear band. In summary, the purity of the protein solution after SEC, as well as the protein yield for HiSiaQM, was much better and higher than for VcSiaQM. In a standard purification after DDM extraction, 2-3 mg protein was received from 10 L LB medium. For this reason, the following chapters and experimental investigations are more focused on the HiSiaQM membrane domains.



Intensity	Mol. weight [kDa]	Sequence coverage [%]	Peptides	Protein IDs	Protein names	Gene names
20,6194	67,592	8,8	7	MartinPeter	Membrane domains TRAP transporter	siaqm
15,6922	50,367	14,8	7	P0AGJ0	Trk system potassium uptake protein TrkA	trkA
13,5512	66,908	6,9	4	Q8XEG2	Glutamine--fructose-6-phosphate aminotransferase [	glmS
13,9189	53,265	6,7	3	3324;5740_W_KV1719;5	Maltose-binding periplasmic protein	malE
12,5412	43,313	7,1	3	P0A6N3	Elongation factor Tu	tufA
13,0273	116,41	6,1	3	Q8X685	Beta-galactosidase	lacZ
12,229	50,325	4,8	2	P0ABB6	ATP synthase subunit beta	atpD
11,7053	41,392	5,4	2	P0A838	Succinyl-CoA ligase [ADP-forming] subunit beta	sucC
15,2435	8,4999	9,9	1	P68681	30S ribosomal protein S21	rpsU
16,8742	26,803	3	1	Q8XD00		Z4269

Figure 4-4: Purification and characterisation of HiSiaQM. A) SEC run of concentrated Ni^{2+}-AC elution solution on an SD 200 10/300 column (standard detergent buffer). The fractions for SDS-PAGE samples are marked in blue. **B)** SDS-PAGE of SEC fractions in A), as indicated by the blue line. **C)** SDS-gels with HiSiaQM after heated and not-heated sample preparation. **D)** Results of mass identification of main protein band from SDS-gel in B) (performed at CECAD/CMMC Proteomics Facility, University of Cologne). The protein sample in polyacrylamide gel was digested with pepsin and the resulting peptides were analysed with mass spectrometry.

In an additional purification experiment, the HiSiaQM protein after SEC was incubated with TEV protease for a cleavage of the N-terminal affinity tag. Interestingly, even with a high excess of TEV protease and long incubation time, the His-tag could not be cleaved off. This behaviour might be resulting from an inaccessibility of the cleavage site or a reduced activity of TEV protease, caused by the membrane mimetic system (Mohanty et al., 2003, Newby et al., 2009). Nevertheless, the explanation of an inaccessible tag fits to the observed low affinity for the QM-domains to the Ni^{2+}-AC by using a standard His_6-tag and the higher affinities for the use of longer His_{14} (VcSiaQM) and His_{10} (HiSiaQM) tags.

Due to the many following experiments with this protein, the observed protein band needed to be unambiguous identified as HiSiaQM, in addition to the western blots. Therefore, the corresponding band at around 55 kDa protein standard was cut out of the gel and the protein was identified with mass spectrometry at the CECAD/CMMC Proteomics Facility (University of Cologne). The detected peptides after pepsin digestion of the protein are shown in Figure SI-18 and the results of alignment of these peptides with the HiSiaQM sequence and sequences of native *E. coli* proteins are shown in Figure 4-4 D. Seven of the detected peptides were successfully aligned to the HiSiaQM protein sequence with a coverage of 8.8%, proving the presence of the protein HiSiaQM in the SDS-gel sample. The low sequence coverage is an often observed effect for membrane proteins, presumably caused by an unsuccessful digestion of the hydrophobic protein regions (Alfonso-Garrido et al., 2015). Furthermore, as known for the high-sensitive method, additional proteins were identified from the gel piece (Shevchenko et al., 2006). The second most present protein with also seven identified peptides and a sequence coverage of 14.8% was TrkA. TrkA is a regulatory protein of the bacterial potassium uptake system Trk with a mass of 50.3 kDa and were previously detected on SDS-PAGE slightly below 75 kDa protein standard (Zhang et al., 2020). Consequently, large amounts of this protein should be detectable at a higher position in the SDS-gel than HiSiaQM. Therefore, HiSiaQM was successfully identified and purification approaches with similar results in the SEC profile and SDS-PAGE were used for further investigations.

4.3 Incorporation of HiSiaQM in protein-bounded nanodiscs

As membrane mimetic system for investigations in a more native-like environment than in detergent micelles, the membrane protein HiSiaQM was reconstituted into protein-bounded nanodiscs. As mentioned above, this system consists of a lipid-bilayer, which is stabilized by a membrane scaffold protein and in which the membrane protein is reconstituted in a self-assembly approach. In the present work, the shortest construct of the nanodisc belt

MSP1 was used, which creates a nanodiscs with a diameter of around 9.7 nm (Ritchie et al., 2009).

4.3.1 Purification of MSP1

First, the membrane scaffold protein MSP1 was expressed and purified. The procedure was based on the work of the group from Stephen G. Sligar (Bayburt et al., 2002, Ritchie et al., 2009) and details are mentioned in the description of the method in Chapter 9.1.4.3. Figure 4-5 illustrates a final SEC and corresponding SDS-PAGE from a purification experiment. The chromatogram revealed a main peak at around 65 mL, flanked by two smaller peaks at lower and higher elution volume (Figure 4-5 A). The SDS-gel showed an overall clean protein solution already after elution of the Ni^{2+}-AC with intense bands at around 20-25 kDa protein standard and a second band at smaller protein standard (MSP1 M_W=24.6 kDa) (Figure 4-5 B, load). The SDS-gel samples from the SEC fractions revealed the MSP1 in the main peak of the chromatogram. The second band around the 15 kDa protein standard could not be separated during the SEC. The two other SEC peaks, especially the peak at lower elution volumes, indicated the MSP1 in the SDS-gel. The varied elution volumes on the SEC column might be caused by an alternative conformation or oligomerization of MSP1 that was not visible in the gel due to denaturing and reducing conditions. However, the protein from the main peak fractions were used for subsequent nanodiscs reconstitution.

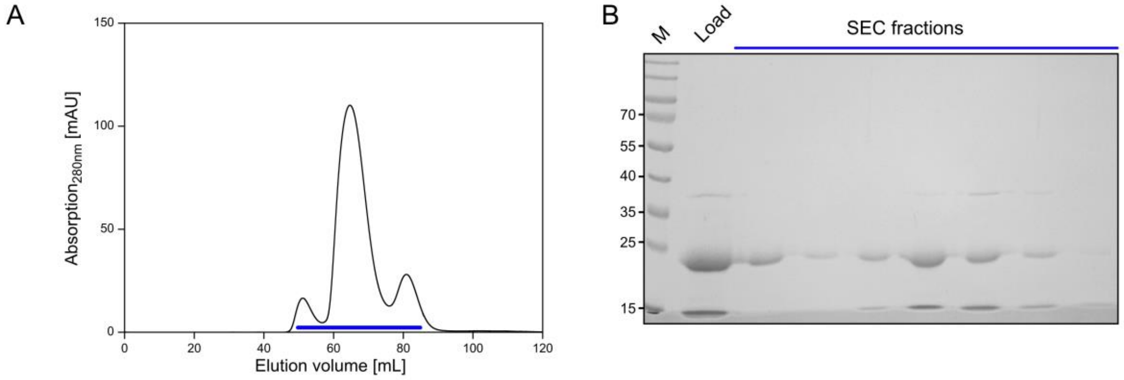


Figure 4-5: Purification of MSP1 for nanodiscs reconstitution. A) SEC run on a HiLoad SD 75 16/600 column with detection of elution at 280 nm absorption. The fractions for SDS-PAGE samples are marked in blue. **B)** SDS-gel with samples from the load of the column (elution of Ni^{2+}-AC) and SEC fractions.

4.3.2 Reconstitution of HiSiaQM in protein-bounded nanodiscs

The reconstitution of the TRAP transporter membrane domains HiSiaQM in MSP-nanodiscs was achieved by optimization of several experimental conditions or compositions together with Peer Deeping during his master thesis. The crucial steps are presented in the following section. The general procedure is based on the work of the group of Stephen G. Sligar (Bayburt et al., 2002, Ritchie et al., 2009). In first test experiments, two reconstitution approaches were prepared with different lipids, DMPC and DOPC (1,2-dimyristoyl-sn-glycero-3-phosphocholine and 1,2-dioleoyl-sn-glycero-3-phospho-choline). These lipids differ in their alkyl chains, DMPC is fully saturated and DOPC has one unsaturated carbon bond in each of the two chains (Figure 4-6 A and B). According to the lipids, a varied component ratio was used, 1:80:0.2 for DMPC and 1:30:0.25 for DOPC (MSP1 to lipids to HiSiaQM molar ratio) (Ritchie et al., 2009). Due to the unsaturated bonds in DOPC, the lipid takes up a larger volume and a smaller ratio should be used. The removal of the detergent and reconstitution of nanodiscs was achieved by an overnight dialysis. Subsequently, the reconstitution trials for both lipids were analysed with a SEC run (Figure 4-6 C and E). Both chromatograms revealed a protein peak at around 12 mL and a shoulder at smaller elution volumes for DOPC, while the DMPC trial showed a prominent peak next to the void volume. The SDS-gels of fractions from both chromatograms revealed two protein bands, one for HiSiaQM at 55 kDa and a second for MSP1 at around 20-25 kDa (Figure 4-6 D and F). Again, the SDS-gel samples were not heated prior to loading on the gel. For both reconstitutions, the HiSiaQM bands showed a maximum of intensity in the SDS-gel at smaller SEC elution volumes, which correspond to higher molecular weights, while the MSP1 was also detectable in the larger elution volumes. This tendency was interpreted as a shift of the HiSiaQM-filled nanodiscs against empty nanodiscs on the SEC column.

Based on the coincidental observation that further DMPC reconstitution experiments with slight changes in the room temperature resulted in a changed SEC elution profile, the experiments were repeated at defined temperatures. The incubation and dialysis of DMPC nanodiscs was performed in an incubator at 26 °C, which is higher than the lipid specific phase transition temperature of 24 °C and should guarantee a fluid-liquid phase of the lipid bilayer (Avanti Polar Lipids, Eeman and Deleu, 2010). In contrast, the reconstitution of DOPC nanodiscs at room temperature was already above the transition temperature of these lipids of -17 °C (Avanti Polar Lipids).

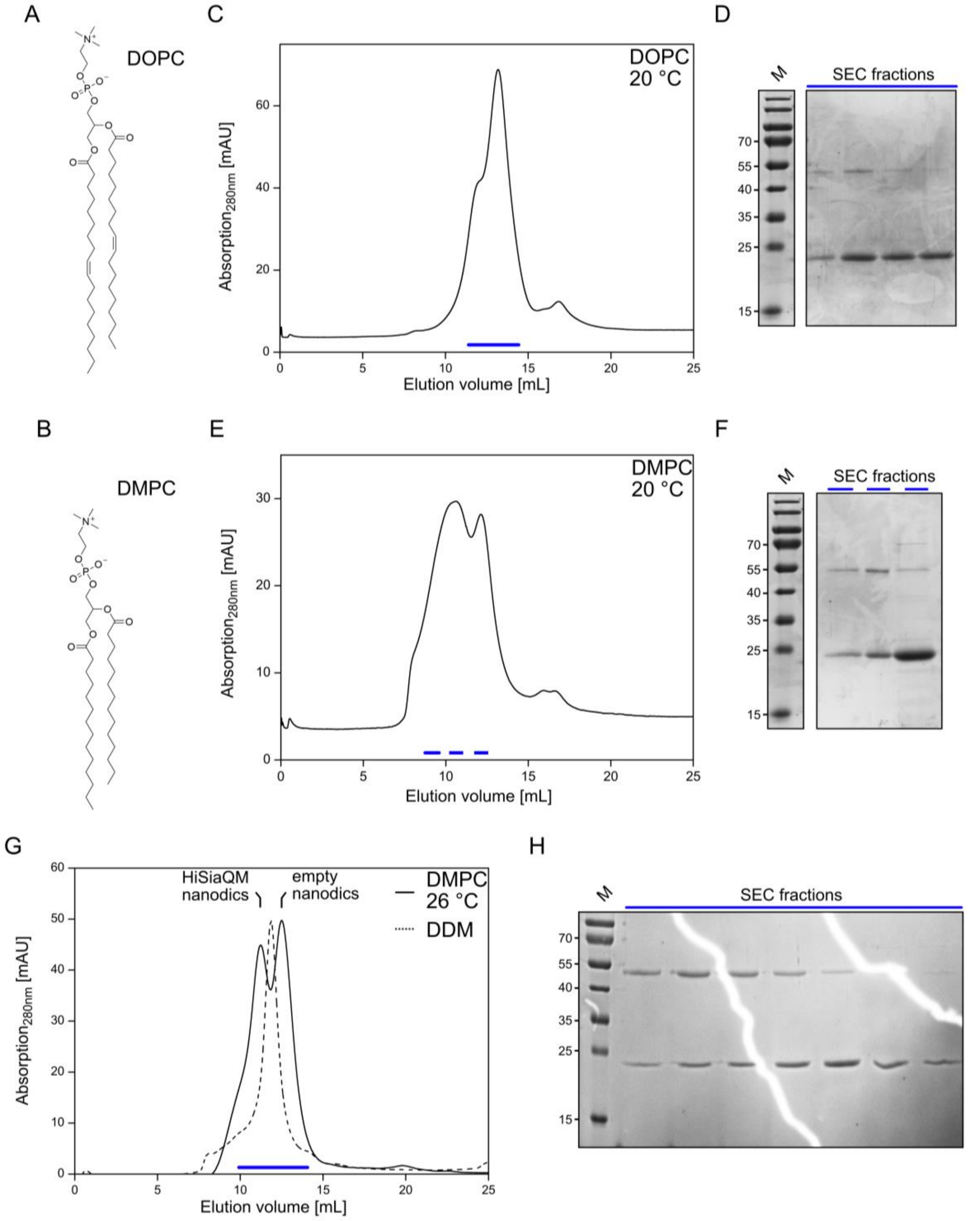


Figure 4-6: Reconstitution approaches of HiSiaQM in protein-bounded nanodiscs. A-B) Chemical structures of DOPC and DMPC lipid. **C)** SEC run of HiSiaQM after reconstitution into MSP-nanodiscs with DOPC lipids on an SD 200 10/300 column (buffer: standard QM-buffer without detergent and glycerol). The fractions for SDS-PAGE samples are marked in blue. **D)** SDS-gel with samples from the SEC eluted fractions in C). **E)** Same as C) but with DMPC lipids and incubation temperature at around 20 °C room temperature. **F)** SDS-gel with samples from the SEC eluted fractions in E). **G)** Same as E) but with incubation temperature of reconstitution mixture and dialysis at 26 °C. Additional, a typical elution profile of a DDM detergent solubilized sample of HiSiaQM on the same column is represented with dashed lines. The intensity is normalized to the maximum of the nanodiscs SEC run and the salt peaks of both chromatograms were used for overlaying. **H)** SDS-gel with samples from the SEC eluted fractions after nanodisc reconstitution in G).

The chromatogram of this experiment is shown in Figure 4-6 G (continuous line) and revealed two protein peaks. The peaks were not baseline separated but the two maxima were distinctly visible. Further, the two peaks were clearly eluted at volumes above 10 mL and were therefore not located in the region of the void volume of around 8 mL. The SDS-PAGE of the SEC fractions indicated a maximum of HiSiaQM protein in the first peak and only small amounts in the second peak (Figure 4-6 H). The maximum of the MSP1 was clearly shifted to higher elution volumes, in direction to the second peak, but was also present in fractions of peak 1. From this result it was postulated that the HiSiaQM-filled nanodiscs were shifted to a smaller elution volume than the empty nanodiscs. Although, this would require a changed hydrodynamic radius between the filled and empty nanodiscs or other specific interactions that resulted in different elution behaviours.

Even if the exact reason for the two peaks remained elusive, the HiSiaQM protein in MSP1-nanodiscs were used for further characterization and experiments in the following chapters. Due to a higher amount of HiSiaQM in the SDS-PAGEs after SEC for the DMPC nanodiscs as for the DOPC nanodiscs, the following reconstitutions were performed with DMPC lipids. Nevertheless, for an additional optimization and increase of the yield, the reconstitution conditions were further tested and modified. For example, the ratio between the MSP1 and lipid towards the amount of HiSiaQM was varied, as well as the total amount of DMPC and sodium cholate in the reconstitution mix (Figure SI-19). One interesting observation was received from a reconstitution approach with high tendency for aggregation. The SEC peak of the aggregated fraction was drastically decreased with repetition of the dialysis overnight (Figure SI-19). Consequently, the dialysis step was increased in time and the number of buffer changes and the final conditions for reconstitution of HiSiaQM in MSP-nanodiscs are described in Chapter 9.1.5.

4.3.3 Verification of HiSiaQM in protein-bounded nanodiscs

In addition to the SEC runs and the corresponding SDS-gels, the reconstituted nanodiscs were characterised with native mass spectrometry and single particle EM. In the present chapter, the EM experiments are just briefly presented for the confirmation of the nanodiscs, a more comprehensive investigation and detailed description is given in Chapter 7.3.

The native mass spectrometry experiments were performed by Kudratullah Karimi from the group of Nina Morgner (University of Frankfurt). In these experiments, the soft ionization technique LILBID (laser induced liquid bead ion desorption) was used that allowed to analyse the mass of structural intact proteins or protein complexes (Peetz et al., 2018). The results for LILBID mass spectrometry with HiSiaQM in MSP-nanodiscs is shown in Figure 4-7 A. The spectrum revealed several peaks which were assigned to components of the reconstitution mix with different charges, including both proteins, HiSiaQM and MSP1. However, some signals were not very sharp and especially at higher mass, the signals were broad and hard to assign to the component. Nonetheless, a signal with corresponding mass for one HiSiaQM and two MSPs at around 119 kDa was detected. The nanodisc belt consists of two MSPs and in the case of a postulated monomeric HiSiaQM protein, this signal confirmed the successful nanodiscs reconstitution (Bayburt et al., 2002, Mulligan et al., 2011). Additionally, the spectrum revealed some species that were assigned to a dimeric HiSiaQM assembly in a nanodiscs. To find an explanation for this observation and to get a more detailed view of the nanodiscs, single particles were investigated with EM.

For initial detection of nanodisc particles with EM, the first peak of a SEC run from a reconstitution approach was investigated after negative staining by the group of Christine Ziegler (University of Regensburg) (Figure 4-7 B, SEC in Figure 7-4 C). The electron micrograph indicated several single particles with a clear visible and well-defined round shape in two preferable orientations, a top and a side-view. The degree of aggregation was low and the particles with a top view orientation had an approximate diameter of 10-11 nm, which correlates well with the literature value of 9.8 nm (Ritchie et al., 2009). The observed stacking of nanodiscs is assigned to an artefact of the negative stain sample preparation and also observed for other MSP-nanodisc samples (personal communication with Veronika Heinz, University of Regensburg). The second peak from a SEC was also investigated with EM after negative staining and similar formed nanodiscs were observed (Figure SI-39). A trend towards slightly smaller particles, probably caused by nanodiscs without protein, was observable and might be the reason for the different elution volumes on the SEC. A more detailed study with EM on HiSiaQM in nanodiscs, including cryoEM samples, is presented in Chapter 7.3 and further confirmed the TRAP transporter membrane domains HiSiaQM embedded in MSP1-bounded nanodiscs.

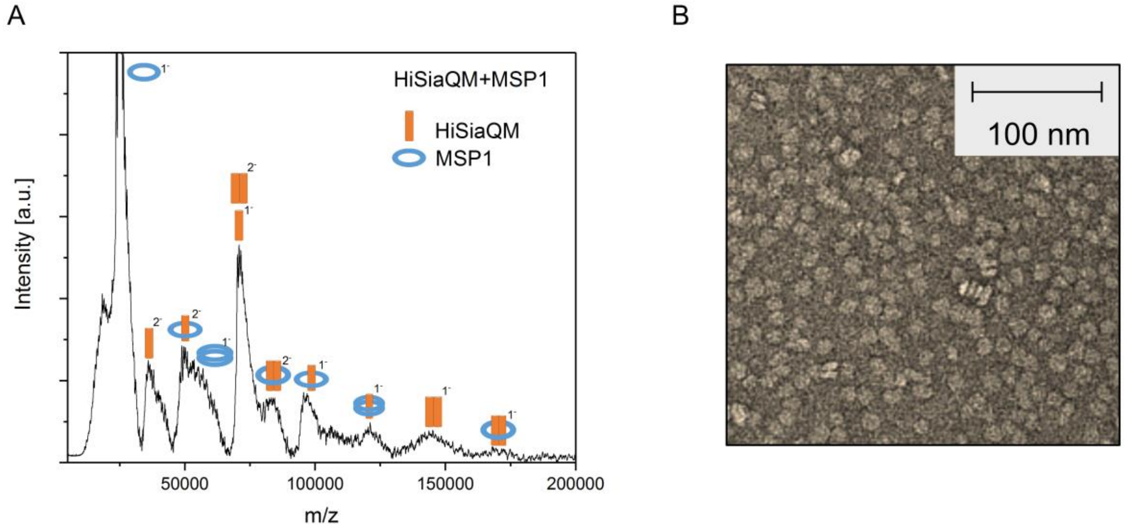


Figure 4-7: Identification of HiSiaQM in protein-bounded nanodiscs. A) Mass spectrometry with LILBID (laser induced liquid bead ion desorption) experiment with HiSiaQM in MSP-nanodiscs, performed by Kudratullah Karimi and Nina Morgner (University of Frankfurt). The MSP1 is illustrated as blue circle and the HiSiaQM as orange bar. The assignment of the mass spectrometry signals is mentioned in the figure with corresponding charge of the biomolecules. **B)** Negative stain EM image of HiSiaQM in MSP1 nanodiscs. A more detailed description of EM with HiSiaQM protein is presented in Chapter 7.3.

4.4 Extraction of HiSiaQM in polymer-bounded nanodiscs

In addition to the detergent micelles and protein-bounded nanodiscs, polymer-bounded nanodiscs as third membrane mimetic system were investigated for solubilization of HiSiaQM. As mentioned before, the advantage of these membrane mimetic system is the conservation of the native environment of the membrane proteins (Dörr et al., 2016).

4.4.1 Extraction and purification

The extraction of HiSiaQM into polymer-bounded nanodiscs was tested by using the two polymers SMA and DIBMA under various conditions (Figure 4-8 A). The concentration of the polymer, as well as the time and temperature of the extraction was varied and the degree of successful solubilized HiSiaQM was determined. For the SMA nanodiscs, the polymer with a 3:1 ratio of styrene to maleic anhydride was used. In the first extractions with SMA, similar conditions as used for the DDM detergent extraction were not successful and neither on an SDS-gel, nor in the elution of a SEC, a specific signal for HiSiaQM was detected. Due to this, the extraction temperature and incubation time were varied in the subsequent experiments. To roughly validate extraction efficiencies, the observation of a changed colouring of the solution and pellet between the two ultracentrifugation steps was used (Figure 4-1 C). In case of DDM detergent extractions, the pellet after the first ultracentrifugation always had a yellow-brown colour which was “bleached” to a nearly white pellet after extraction and the second ultracentrifugation step (Figure 4-8 B). The pellets of extraction with polymers indicated a similar behaviour just for incubation of SMA at 21 °C or 37 °C overnight, with tendency of more “bleaching” to the higher temperature (Figure 4-8 B). The 4 °C approach with SMA, which presents the temperature for DDM extraction, and all three temperature approaches with DIBMA, indicated an unchanged yellow-brown pellet after extraction and second ultracentrifugation.

Additionally, for a more quantitative statement, the supernatant of the second ultracentrifugation step was analysed with a western blot to detect solubilized His-tagged HiSiaQM. As before, the samples were not heated before they were loaded onto the SDS-gel. Samples for the western blot were prepared with an extraction time of only two hours and overnight with varying incubation temperatures. The detected protein bands correlated with the observation of the changed pellet colour (Figure 4-8 C). A DDM extraction at 4 °C was used as a control and indicated the most intense band. Overall, the SMA polymer extraction resulted in higher extraction yields than DIBMA, independent from the temperature change. For both polymers, a small increase in extracted HiSiaQM protein amount was obtained for an increase of the incubation temperature and for longer extraction time.

In another approach, the amounts of polymer during extraction were varied from the initial used 2.5% (v/v). The pellets after the second ultracentrifugation indicated only a “bleaching” for the SMA polymer, similar to the DDM extraction (Figure 4-8 D). For DIBMA extraction, even the solution with 5% (v/v) polymer indicated no change of the pellet colour. The three different concentrations of SMA revealed an increase of the observed effect proportional to the SMA concentration. In contrast to this, the western blot for this test-extraction indicated more intense bands for DIBMA than for SMA extraction (Figure 4-8 E). This could result from sharper bands on the right side of the gel or an experimental error such as an unequal incubation of the western blot during antibody binding or staining. In summary, the results of these test-extractions suggest the successful solubilization of HiSiaQM in polymer-bounded nanodiscs. To further analyse the polymer-bounded nanodiscs, one batch with each polymer was purified after extraction of HiSiaQM with the optimized parameters.

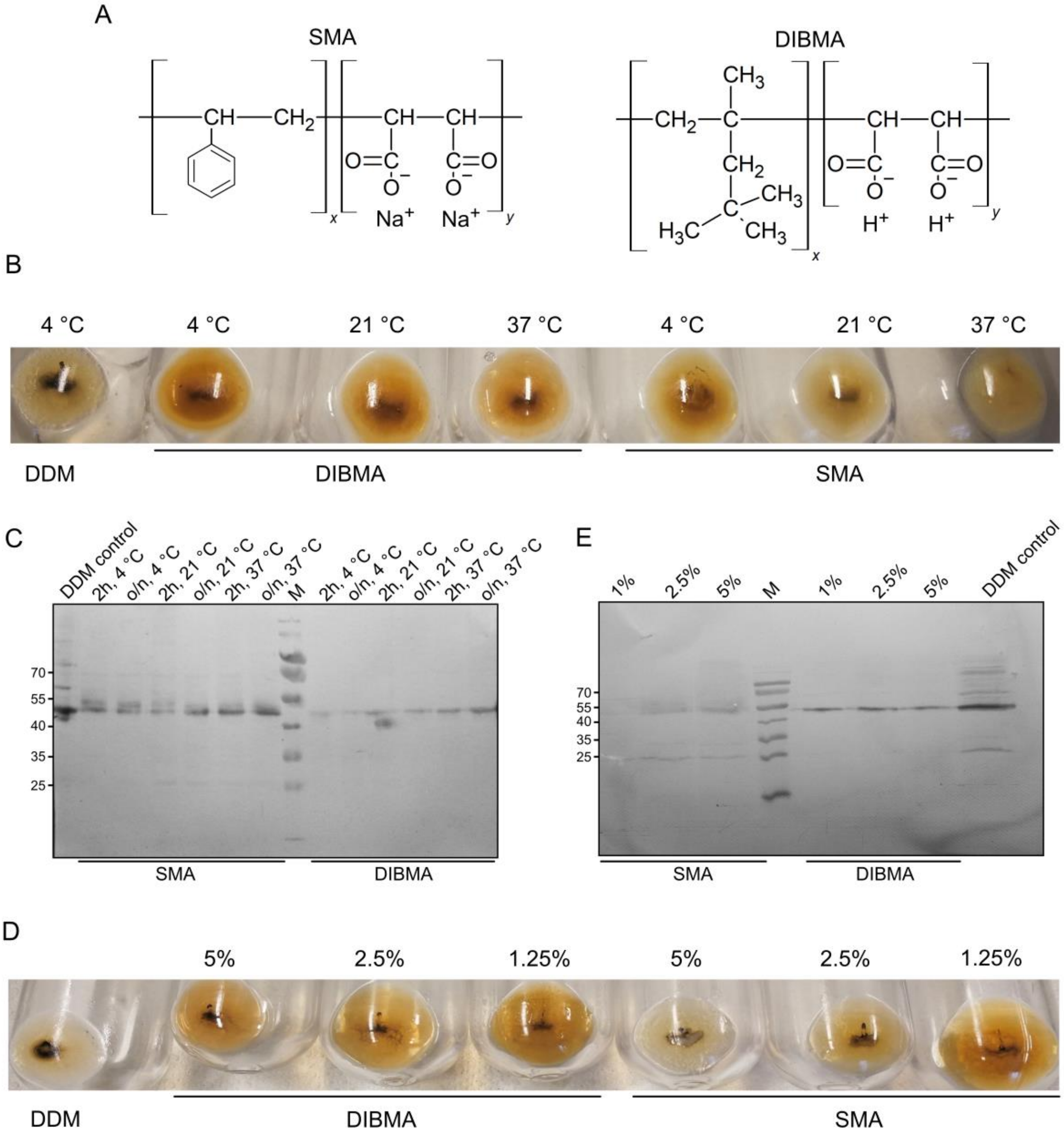


Figure 4-8: Extraction of HiSiaQM in polymer-bounded nanodiscs. A) Chemical structures of the SMA (styrene maleic acid) and DIBMA (diisobutylene maleic acid) polymers that were used for extraction tests (Dörr et al., 2016, Oluwole et al., 2017a). **B)** Pellets after the second ultracentrifugation step (compare Figure 4-1 C) after extraction of HiSiaQM with SMA or DIBMA overnight at different temperatures and 2.5% (v/v) polymer. As control, a DDM extracted sample was simultaneously prepared. **C)** Western blot from supernatant fractions in B), after incubation with SMA and DIBMA and second ultracentrifugation. For each temperature, a sample after 2 hours and overnight incubation were analysed and detected with anti-His antibody. **D)** Same as B) but with different concentrations of polymers instead of different temperatures. **E)** Same as C) but for supernatant fractions from D).

Based on the detergent purification, the polymer-bounded nanodiscs were firstly purified via Ni^{2+}-AC without imidazole in the washing buffer and subsequent loaded onto a SEC column. The resulting SEC elution profiles for both polymers are shown in Figure 4-9 A and C. The SEC run of SMA extracted sample indicated an intensive, monodisperse peak at around 11 mL elution volume with little shoulders at smaller and higher elution volumes. As expected for SMA-nanodiscs, the ratio between absorption of 280 and 254 nm changed significantly in comparison to standard proteins or DDM solubilized HiSiaQM, due to the UV/Vis absorption of the polymer (Oluwole et al., 2017a). The SDS-gel of the SEC fractions from the detected peak in Figure 4-9 B indicated two bands at around 55 kDa, whereas the lower one was assigned with the help of western blots to the target protein HiSiaQM (Figure 4-10 A).

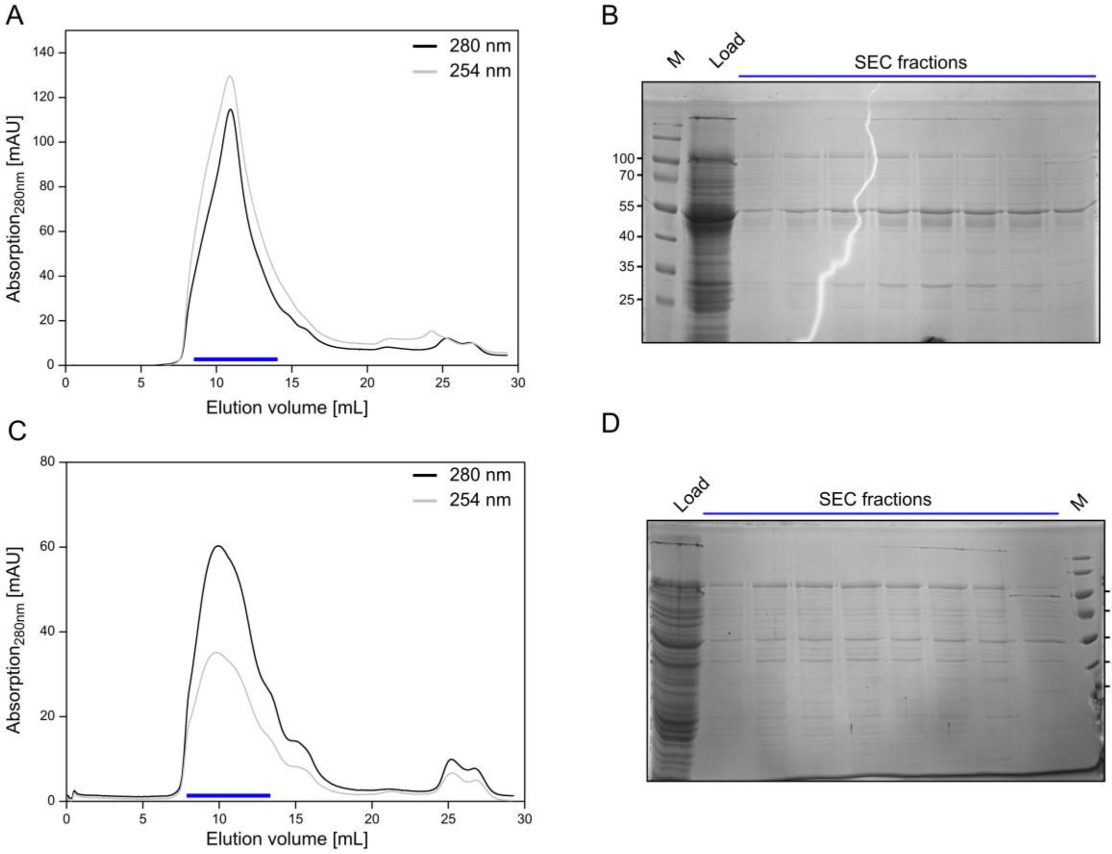


Figure 4-9: SEC runs with HiSiaQM in polymer-bounded nanodiscs. A) SEC of a SMA polymer extracted HiSiaQM sample after Ni^{2+}-AC elution with detected absorption at 280 nm (black) and 254 nm (grey) on an SD 200 10/300 column (standard QM buffer without DDM). The fractions for SDS-PAGE samples are marked in blue. **B)** SDS-PAGE of SEC fractions from the main peak in A), as indicated by the blue marker. **C)** Same as A) but with HiSiaQM extracted with DIBMA polymer. **D)** SDS-PAGE of SEC fractions from the main peak in C), as indicated by the blue marker.

The SEC profile of the DIBMA extraction indicated a broader peak with a maximum at around 10 mL and several shoulders at higher and lower elution volumes (Figure 4-9 C). In contrast to the SMA-nanodiscs, this approach showed a detergent-like ratio between absorption at 280 nm and 254 nm, due to the absence of styrene in the DIBMA polymer. However, the total intensity of absorption in SEC elution was lowered by a factor of 2, compared to the SMA extraction. Further, the SDS-gel of the SEC fractions indicated several different protein bands and only weak bands for HiSiaQM were detectable. Overall, the SEC peaks for both nanodisc preparations were slightly broader than the DDM extracted protein peaks (Figure 4-4 A), but especially the SMA polymer experiment indicated a successful extraction and purification of HiSiaQM in this nanodiscs.

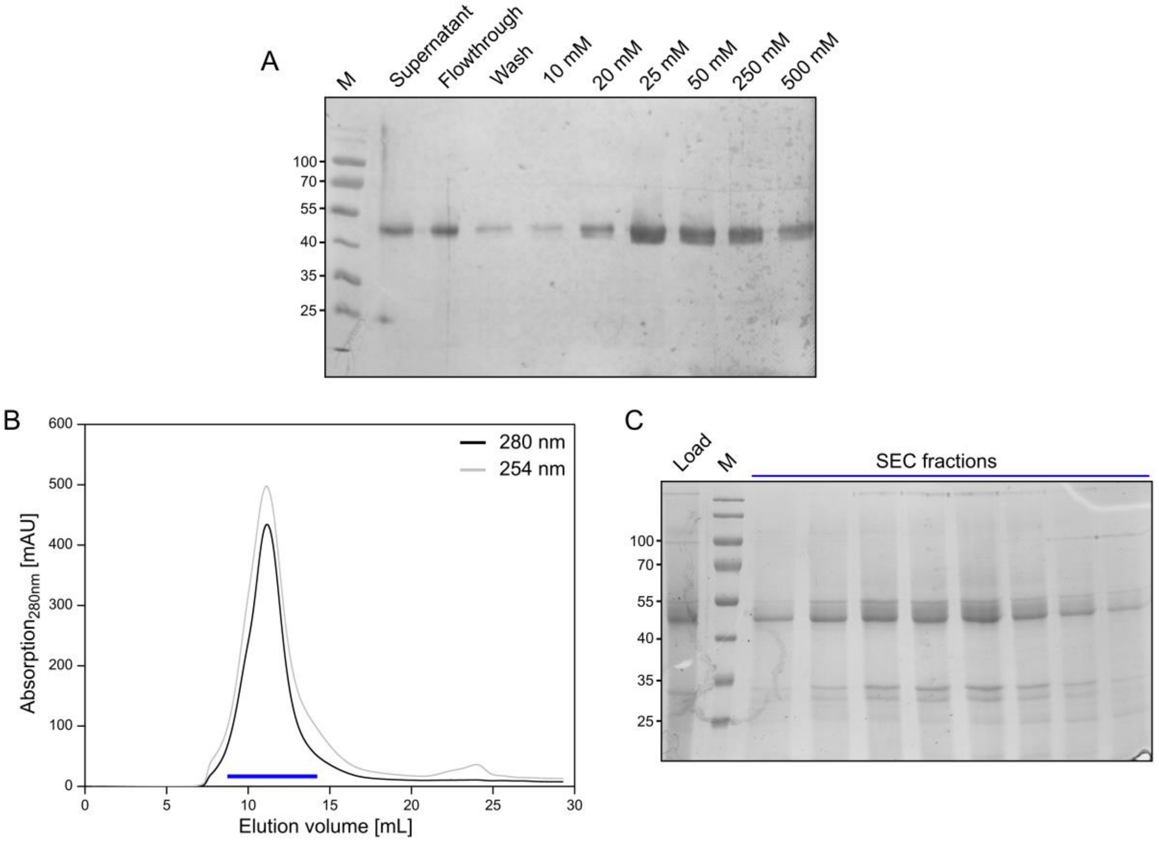
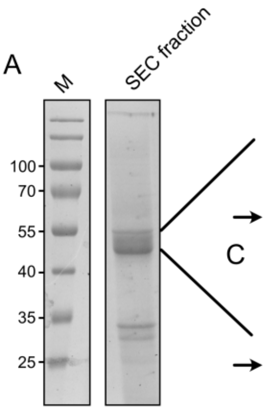


Figure 4-10: Purification optimization of HiSiaQM in SMA-nanodiscs. A) Western blot of fractions from a Ni^{2+}-AC with elution of the protein with an imidazole gradient, detected with a His-tag antibody. The fractions and concentrations are mentioned in the figure. **B)** SEC of HiSiaQM in SMA nanodiscs after Ni^{2+}-AC with absorption at 280 nm (black) and 254 nm (grey) on an SD 200 10/300 column (standard QM buffer without detergent). The fractions for SDS-PAGE samples are marked in blue. **C)** SDS-PAGE of SEC fractions from the main peak in B), indicated by the blue line.

Due to a higher quality and quantity of SMA-extracted HiSiaQM, this polymer was used for further experiments. To optimize the yield and reduce the impurities, the binding strength of HiSiaQM in SMA-nanodiscs to Ni^{2+}-AC resins was analysed with an imidazole elution gradient and a subsequent western blot (Figure 4-10 A). The flowthrough fraction of the Ni^{2+}-AC column after batch-binding indicated a similar protein concentration as the supernatant fraction (Figure 4-10 A). This might result from a low binding affinity or a high protein excess with a fully bound Ni^{2+}-AC column. The following elution fractions with a stepwise increase of the imidazole concentration showed a clear elution of HiSiaQM in protein-bounded nanodiscs at around 20 mM and 25 mM imidazole. Consequently, HiSiaQM in SMA-nanodiscs showed a lower affinity to the column matrix than the same protein stabilized in DDM detergent micelles, which could be treated with 25 mM imidazole without elution of the target protein. The decreased affinity was possibly caused by a lowered accessibility of the His_{10}-tag in the native nanodiscs. Nonetheless, a SEC run of a large-scale approach resulted in an intense, monodisperse peak (Figure 4-10 B). The corresponding SDS-gel visualized a defined and intense peak for HiSiaQM between 40 and 55 kDa but also several other proteins in high amounts that could not be separated during purification (Figure 4-10 C).

4.4.2 Verification of HiSiaQM in polymer-bounded nanodiscs

The specific detection of HiSiaQM with a western blot in the supernatant, after extraction with the polymers and the second ultracentrifugation, was one evidence for the successful stabilization of the membrane domains in the nanodiscs. To further characterise the protein bands after SEC in the SDS-gel and to identify the target protein, the two bands in the region of the 55 kDa protein standard were analysed with mass spectrometry (Figure 4-11 A) (core facility mass spectrometry, University of Bonn). Similar to the identification of DDM solubilized protein, the protein bands were digested with pepsin and the resulting peptides identified via mass spectrometry. The alignment of these peptides with the HiSiaQM sequence and an *E. coli* protein data bank confirmed the target protein in both gel fractions (Figure 4-11 B and C, black arrows). The abundance of HiSiaQM was higher in the lower peak, as postulated by the detection in a western blot. The protein cytochrome bo_3 ubiquinol oxidase was identified as the main additional protein, which might be also the main component of the second band at 55 kDa. This *E. coli* protein is natively located in the cytoplasmic membrane of the bacteria and seems to be co-extracted with the SMA polymer into nanodiscs (Yap et al., 2010).



B

Description	Coverage [%]	Peptides	Peptide spectrum match	Unique Peptides	AAs	MW [kDa]	Abundances
Cytochrome bo(3) ubiquinol oxidase subunit 1	9	9	85	9	663	74,3	1,1E+08
Cytochrome bo(3) ubiquinol oxidase subunit 2	18	7	27	7	315	34,9	4,5E+06
PTS system fructose-specific EIIB'BC component	18	10	25	10	563	57,5	3,0E+06
ATP synthase subunit alpha	18	11	18	11	513	55,2	2,6E+06
Outer membrane protein assembly factor BamA	4	3	13	3	810	90,5	2,2E+06
Protein YdgA	22	10	15	10	502	54,7	2,1E+06
Pyruvate/proton symporter BtsT	3	2	5	2	716	77,3	1,4E+06
Ribose import ATP-binding protein RbsA	18	8	18	8	501	55	1,3E+06
Sialic acid TRAP transporter HiSiaQM	6	4	8	4	634	69,8	1,1E+06
Methyl-accepting chemotaxis protein I	7	4	8	4	551	59,4	6,2E+05

C

Description	Coverage [%]	Peptides	Peptides spectrum match	AAs	MW [kDa]	Abundances
Cytochrome bo(3) ubiquinol oxidase subunit 1	14	12	73	663	74,3	7,2E+07
Sialic acid TRAP transporter HiSiaQM	12	8	23	634	69,8	7,8E+06
Tryptophanase	14	8	21	471	52,7	3,6E+06
Cytochrome bo(3) ubiquinol oxidase subunit 2	26	7	25	315	34,9	3,4E+06
NADH-quinone oxidoreductase subunit F	14	7	23	445	49,3	2,1E+06
PTS system glucose-specific EIICB component	10	6	15	477	50,6	2,0E+06
Pyruvate dehydrogenase E1 component	5	5	7	887	99,6	1,9E+06
Phosphoglucosamine mutase	14	5	8	445	47,5	1,8E+06
Aerobic glycerol-3-phosphate dehydrogenase	13	5	8	501	56,7	1,8E+06
D-amino acid dehydrogenase	18	8	14	432	47,6	1,6E+06

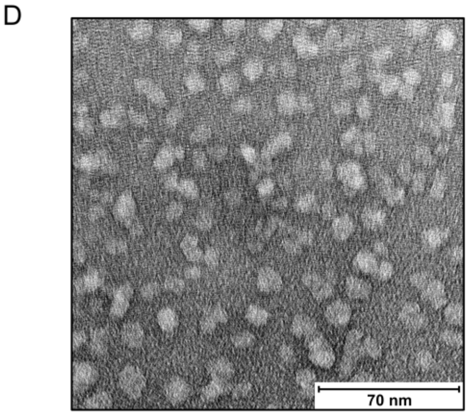


Figure 4-11: Identification of HiSiaQM in polymer-bounded nanodiscs. A) Single SDS-PAGE fraction with sample from the main peak of the SEC in Figure 4-10 B. The two intense protein bands at around 55 kDa protein standard were used for identification. **B)** Protein identification with mass spectrometry from the corresponding protein band in A). The samples were digested with pepsin and the resulting peptides were analysed with mass spectrometry and aligned to protein sequences (core facility mass spectrometry, University of Bonn). **C)** Same as B) but for other protein band in A). **D)** EM micrograph of HiSiaQM in SMA polymer stabilized nanodiscs after uranyl acetate staining (negative stain).

As second method for identification of nanodiscs, the samples were analysed with negative stain EM in the group of Elmar Behrmann (Center of Advanced European Studies and

Research (caesar), Bonn). The protein containing fraction after SEC indicated a similar appearance as the protein-bounded nanodiscs from Chapter 4.3.3 (Figure 4-11 D). The sample revealed single particles with a nanodisc-shaped form and a low degree of aggregation. The size of the nanodiscs was determined to around 10 nm in diameter, which fits nicely to values from previous studies with also around 10 nm (Jamshad et al., 2015).

4.5 Discussion and conclusion

The sections above present the successful production sialic acid TRAP transporter membrane domains for handling in aqueous solution and provided thereby a basis for several bioanalytical studies. The expression, extraction and purification of the membrane domains VcSiaQM and HiSiaQM were only published once before from our cooperation group of Gavin H. Thomas (University of York) by using the classical extraction method of detergent micelles (Mulligan et al., 2009, Mulligan et al., 2012).

For the extraction and purification with DDM detergent, the HiSiaQM protein was received with a significantly higher yield and higher purity than VcSiaQM by using the identical procedure. The exact reason for this observation remained unknown, especially because the proteins have a sequence identity of 45.5% and similarity of 64.1% (EMBOSS Needle Pairwise Alignment (Madeira et al., 2019)). A hypothetical reason for a lower yield might be the interaction between the Q- and M-domains in the two transporters. Even if a strong interaction between the two protein chains to one complex is presumed for VcSiaQM (Mulligan et al., 2012), the fusion of the two domains in one peptide chain as for HiSiaQM can possibly result in a higher stability, especially after extraction from the native environment. The quality and quantity of the HiSiaQM membrane domains solutions after purification were satisfying for subsequent structural and functional studies such as X-ray crystallography trials.

To expand the toolkit of solubilization techniques and to pave the way for functional studies in native-like environments, for the first time HiSiaQM was successfully reconstituted into lipid bilayer nanodiscs. During investigations of these nanodiscs, the identification of HiSiaQM with native mass spectrometry showed hints for a dimerized protein. Until today, the membrane domains are postulated to be functionally active as a monomer, comparable to classical secondary transporter (Mulligan et al., 2011). However, the corresponding signals were rather weak and the reliability is rather low and must be validated with care. Nonetheless, for further bioanalytical studies, the possibility of a dimerization should be considered. Another reason for the detected dimerization might be not based on the function but rather on the random reconstitution of two HiSiaQM proteins into one nanodisc.

Additionally, the HiSiaQM TRAP transporter was also extracted and solubilized with polymer-bounded nanodiscs. Recently it was shown that the extraction with SMA polymers is sensitive to many parameters, such as polymer concentration, temperature, incubation time, ionic strength and pH (Kopf et al., 2020). Also, for HiSiaQM, the extraction was only observed after increasing of the temperature from 4 °C to 21 °C and for an even better extraction to 37 °C. Nonetheless, these results represent the first approach in which TRAP transporter proteins were extracted in their conserved native lipid bilayer. During test-experiments, the use of SMA polymer revealed a higher extraction efficiency of the target protein than the DIBMA-polymer. This observation is in consistency with a recent study, in which the extraction of other membrane proteins with DIBMA also showed a lower yield and larger amounts of impurities than the extractions with SMA (Gulamhussein et al., 2020).

Compared to the detergent purification, also the protein solution of SMA polymer-bounded nanodiscs after purification revealed higher concentrations of impurities. The insufficient purification possibly results from a co-extraction of other membrane proteins in the same nanodiscs as HiSiaQM. Such an assembly makes it impossible to separate the components with the used chromatography steps without breaking off the nanodiscs. Alternatively, the impurities might be present in individual nanodiscs which have a similar size as the HiSiaQM nanodiscs and hinder a sufficient separation. To further enhance the purification, several changes like other kinds of tags, alternative tag positions or additionally purification steps should be investigated.

The outcome of this chapter was directly used in the work that is presented in the following chapters with functional and structural studies on the TRAP transporter membrane domains. Thereby, the investigations benefit from the high yields and good purity that were achieved in the described experiments above. The purified HiSiaQM protein in different membrane mimetic systems were used for functional studies with EPR spectroscopy (Chapter 5), to generate and identify specific binding nanobodies (Chapter 6) and to investigate the structure of the proteins (Chapter 7). For the polymer-bounded nanodiscs, unknown limitations and incompatibilities are observed in the following chapters and summarized in Chapter 8. A further interesting analysis of the SMA-nanodiscs with HiSiaQM will be the determination of the co-isolated lipids. As observed in other studies before, the nanodiscs can reveal an enrichment of specific lipids, which were postulated as important for the membrane proteins as they represent their native environmental lipids (Dörr et al., 2014, Prabudiansyah et al., 2015). For TRAP transporters, none of such protein-lipid interactions were determined before. Therefore, the successful detection of specific lipids can reveal a useful comparison to classical secondary transporters and information about the importance of the native lipid bilayer for the protein.

Chapter 5

Functional studies of TRAP transporter membrane domains with EPR spectroscopy

Preamble:

All EPR experiments were performed at the Institute for Physical and Theoretical Chemistry. The EPR spectrometers from the group of Olav Schiemann were used for the experiments.

5.1 Overview about PELDOR spectroscopy on membrane proteins

As described in Chapter 4.1, the number of determined membrane protein structures has strongly increased during the last 10 years and the demand of a functional description arose with every new structure (Hendrickson, 2016). As presented in the following paragraphs, several membrane protein structures were therefore supported with orthogonal methods such as PELDOR spectroscopy to gain insights into their functional behaviour (Mchaourab et al., 2011, Wunnicke and Hänelt, 2017) (Chapter 2.1).

A major advantage of PELDOR spectroscopy is that normally only the spin centres are spectroscopically detectable and the rest of the protein or the sample environment, especially the membrane mimetic system, are "invisible". Therefore, if the distance between two centres is in a measurable range, standardly around 2-8 nm (Jeschke, 2012) (Chapter 2.1), the size of the investigated biological assembly has no direct influence on the experimental result. These facts are beneficial for PELDOR spectroscopy studies on membrane proteins in solution. Most PELDOR spectroscopy studies on membrane proteins are performed in detergent micelles, but during the last years, also more native-like systems were successfully used with PELDOR spectroscopy, such as lipid bilayer systems of bicelles or protein-bounded nanodiscs (Evans et al., 2020, Kapsalis et al., 2019, Ward et al., 2014). Further, PELDOR experiments were performed in large liposomes, which allowed to imitate cells with a distinct inner and outer environment (Joseph et al., 2014, Zou et al., 2009).

An important PELDOR study on membrane proteins was the first measurement of a transporter system that was embedded in the outer membrane of bacterial cells and thereby in its native environment (Joseph et al., 2015). The labelling was performed via site-directed spin labelling (SDSL) of artificially introduced cysteines. Due to no native cysteines in outer membrane proteins, the labelling procedure of the target transporter system in the bacterial membrane was still selective (Joseph et al., 2015, Merianos et al., 2000). Even if this experiment was a breakthrough for measurements in physiological relevant environments, the procedure is only feasible for a limited number of membrane proteins in the outer membrane of bacteria. Solubilization techniques such as polymer-bounded nanodiscs (Chapter 4.1), which conserve the native environment and allow a simultaneous handling of the membrane proteins in aqueous solution, might be a helpful alternative. However, until today, PELDOR spectroscopy studies in polymer-bounded nanodiscs were only performed with membrane proteins that were first solubilized with detergent and subsequently reconstituted into the nanodiscs (Orwick-Rydmark et al., 2012, Sahu et al., 2013). Nevertheless, the results from these studies promote distance measurements of a membrane protein with PELDOR spectroscopy in their native environment in polymer-bounded nanodiscs.

Because membrane transporters typically undergo large conformational rearrangements during the transport process, one aim that is often addressed with PELDOR spectroscopy is the description of certain steps of the transport process and a postulation of a transport cycle. Until today, several membrane transporters were successfully studied and revealed detailed insights into the function during transport process, for example peptide export by the ABC exporter McjD (Bountra et al., 2017), the ATP-hydrolysis cycle of ABC transporter LmrA (Hellmich et al., 2012) or the interaction between the SBP and TMD of the MalE-FGK$_2$ transporter (Böhm et al., 2013). Especially for secondary active transporters, PELDOR experiments revealed new insights into the coupling mechanisms of LeuT and the MATE antiporter NorM (multidrug and toxic compound extrusion transporter) (Claxton et al., 2010, Claxton et al., 2018). Additional to these investigations on the transport cycle, PELDOR spectroscopy was also used to examine the interaction and influence of lipid onto the dynamic behaviour of the protein (Martens et al., 2016, Ward et al., 2014). Taken together, the presented studies revealed a broad application of PELDOR spectroscopy experiments for functional studies on various membrane proteins in different membrane mimetic systems, with interesting new details of their function.

One important requirement for a reliable interpretation of measured PELDOR distances on the function of a protein is a comparison of the distances with a protein structure (Chapter 2). Until today, no protein structure of the TRAP transporter membrane domains was experimentally determined and only two models are available (Mulligan et al., 2012, Ovchinnikov et al., 2014). Additionally, also many functional characteristics of the TRAP transporter membrane domains are unknown, such as binding of the substrate, interaction with the SBP or influence of the lipids in the environment. Due to the unknown structure, the PELDOR studies in the following chapter are more focused on qualitative changes of the detected PELDOR time traces upon addition of a compound than on a structural interpretation of distinct conformational changes. Furthermore, the PELDOR results can be possibly used for a comparison and experimental validation of the two available protein models.

5.2 Development of a spin labelling strategy for HiSiaQM

All functional studies of the TRAP transporter membrane domains with EPR spectroscopy in the following chapter were performed with the cysteine-reactive nitroxide spin label MTSSL. If not mentioned otherwise, the HiSiaQM protein was extracted and stabilized in DDM detergent micelles for all experiments.

5.2.1 Initial spin labelling approaches

As a first step in preparation of a spin labelled variant of TRAP transporter membrane domains, the binding of MTSSL to the wildtype proteins VcSiaQM and HiSiaQM was analysed. Both proteins have several native cysteines, 6 cysteines in VcSiaQM and 5 in HiSiaQM. According to the sequence-based TMH-topology models from Mulligan et al. (Mulligan et al., 2012), the cysteines in VcSiaQM are all located in TMHs, whereas HiSiaQM has two cysteines in transmembrane-helices and three cysteines in cytoplasmic loops (Figure 5-1 A and B, red circles). Using these models, a reliable prognosis of label accessible cysteines is not possible, even if several cysteines are predicted in the membrane region and presumably have no or only a low accessibility. To analyse the labelling of the native cysteines, the wildtype proteins were therefore treated with the spin label during Ni^{2+}-AC elution, as described in the standard labelling procedure for soluble proteins (Chapter 9.1.6.1). For VcSiaQM, in one spin labelling experiment only the half protein sample was used for labelling, while the other one was separated and used to identify aggregation or other changes of the protein that occurs due to the labelling addition. Both protein samples were loaded onto a SEC column and the chromatograms revealed a similar elution profile without a change of aggregation peak between the untreated and spin labelled sample (Figure SI-20).

In the next step, labelling approaches of both proteins after standard purification were analysed with X-band cw-EPR at room temperature to investigate the labelling procedure. The measuring parameters such as protein concentration or number of averaged scans were changed and optimized during the first cw-EPR experiments and the spectra in Figure 5-1 and Figure 5-2 must be compared with caution. Nonetheless, they were used for a first estimation of the quality of the labelling procedure on the membrane domains.

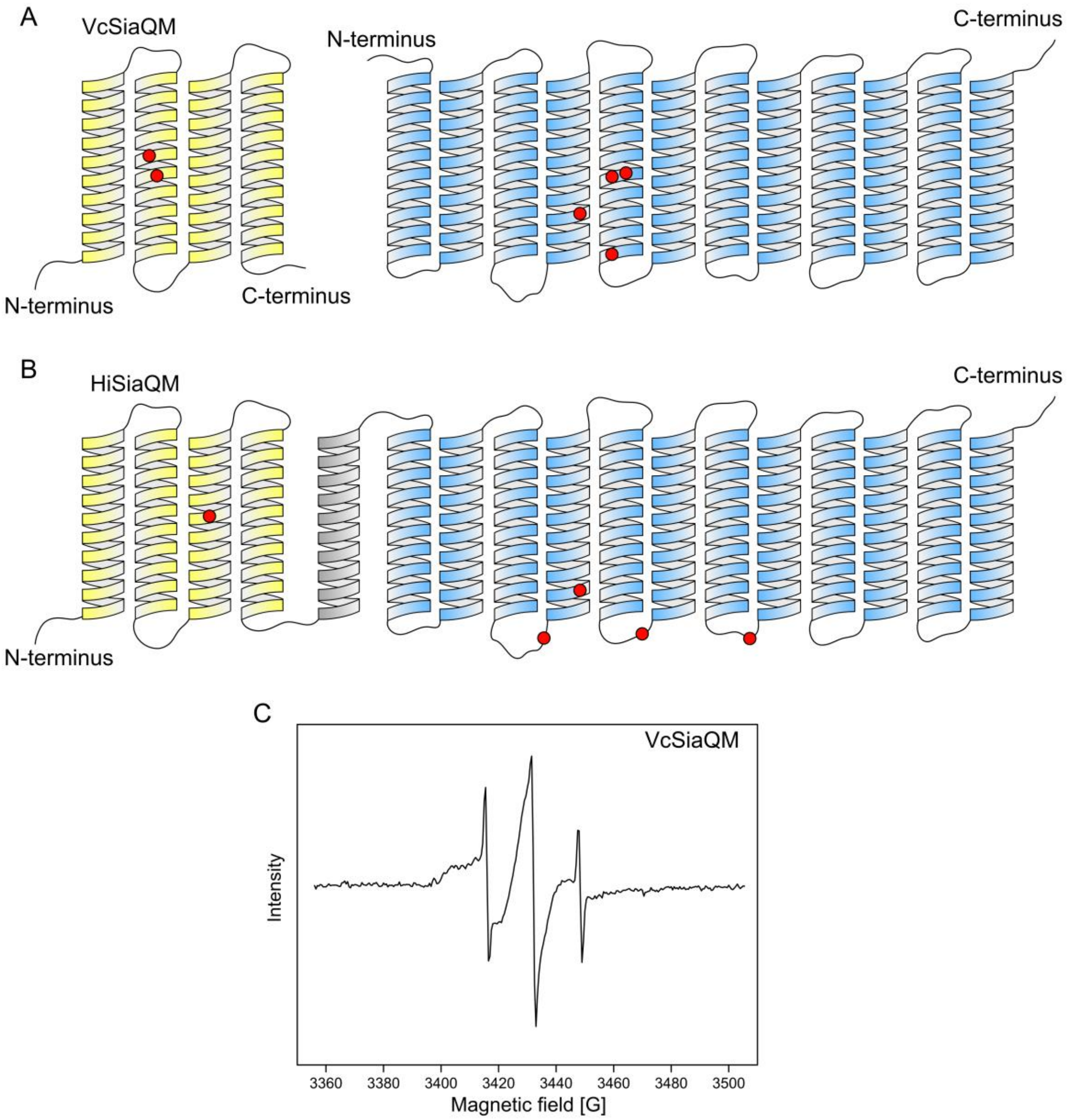


Figure 5-1: Initial labelling of TRAP transporter membrane domains. A) TMH-topology prediction of the VcSiaQM domains, the Q-domain is represented in yellow and the M-domain in blue (Mulligan et al., 2011). The postulated positions of cysteines are marked with red circles. **B)** Same as A) but for HiSiaQM with a connecting TMH in grey. **C)** X-band RT cw-EPR spectrum (normalized) of VcSiaQM wildtype after labelling with cysteine reactive nitroxide spin label MTSSL (200 µM protein, 500 scans). The labelling was performed with the standard protocol for soluble proteins (Chapter 9.1.6.1). Comparable labelling of HiSiaQM is shown in Figure 5-2 A.

In the spectra of both proteins, a clear signal of triplets for a nitroxide spin label was observed (Figure 5-1 C and Figure 5-2 A). However, both spectra consist of at least two different kinds of triplets, a more isotropic one and a more anisotropic one. These features are caused by two different subsets of spin labels in the solution with a different rotation correlation time (NejatyJahromy and Schubert, 2014). The isotropic spectrum indicated a subset with a higher mobility and is presumably caused by a free, unbound spin label and the anisotropic with a lower mobility might be result from a protein-bounded label.

The cw-EPR spectra of both membrane proteins suggest that at least one native cysteine was accessible for the label and also that excess, unbound label was not successfully separated from the protein with a SEC run. Consequently, prior to PELDOR spectroscopy for functional studies, these two features needed to be either clarified or prevented. The unbound spin label should be removed from the protein solution to reduce intermolecular spin label interactions. Additionally, to correlate a subsequent PELDOR experiment with the function of the protein, the spin labels need to be introduced into the protein at known positions. Therefore, different ways to enhance the selectivity of labelling, such as unnatural amino acids or mutations of native cysteines were attempted and described in the following chapters. Due to a simultaneous obtained higher protein yield for HiSiaQM than for VcSiaQM (Chapter 4.2), the further experiments were focused on the membrane domains HiSiaQM.

5.2.2 Optimization of the spin labelling procedure

During the optimization of the labelling selectivity in the following chapters, the labelling procedure was simultaneously customized to enhance the labelling efficiency. The free excess label was successfully separated from the sample by changing the labelling step from the elution step of the Ni^{2+}-AC to a washing step. After this label-washing step, the Ni^{2+}-AC column was washed in an additionally step with standard buffer to remove the unbound label. After elution of the protein and a subsequent SEC run, the cw-EPR spectrum revealed a clear change to the initial recorded spectrum (Figure 5-2 A and B). The spectrum had only an anisotropic signal that is typical for a nitroxide spin label attached to a biomolecule, while the isotropic signal of a highly mobile spin label species was not detected. This showed that in contrast to a SEC run, a washing step on the Ni^{2+}-AC column was able to separate the free label from the protein solution.

By using this labelling procedure, the incubation time between the spin label and the target protein was strongly reduced and also the concentration of spin label in the washing step was lower than in the elution step. These factors resulted in a decreased labelling efficiency, as directly visible from the cw-EPR spectra in Figure 5-2 B. To increase the labelling efficiency, the procedure was further modified in several steps and trials. In the final protocol, the HiSiaQM protein was eluted from the Ni^{2+}-AC after a washing step with reducing agent and directly supplemented with MTSSL. After incubation overnight at 4 °C and a buffer change to remove imidazole, a second Ni^{2+}-AC was performed to separate the free excess label from the protein. The protein on the Ni^{2+}-AC column was washed with standard buffer, eluted in the similar way as before and finally purified with SEC.

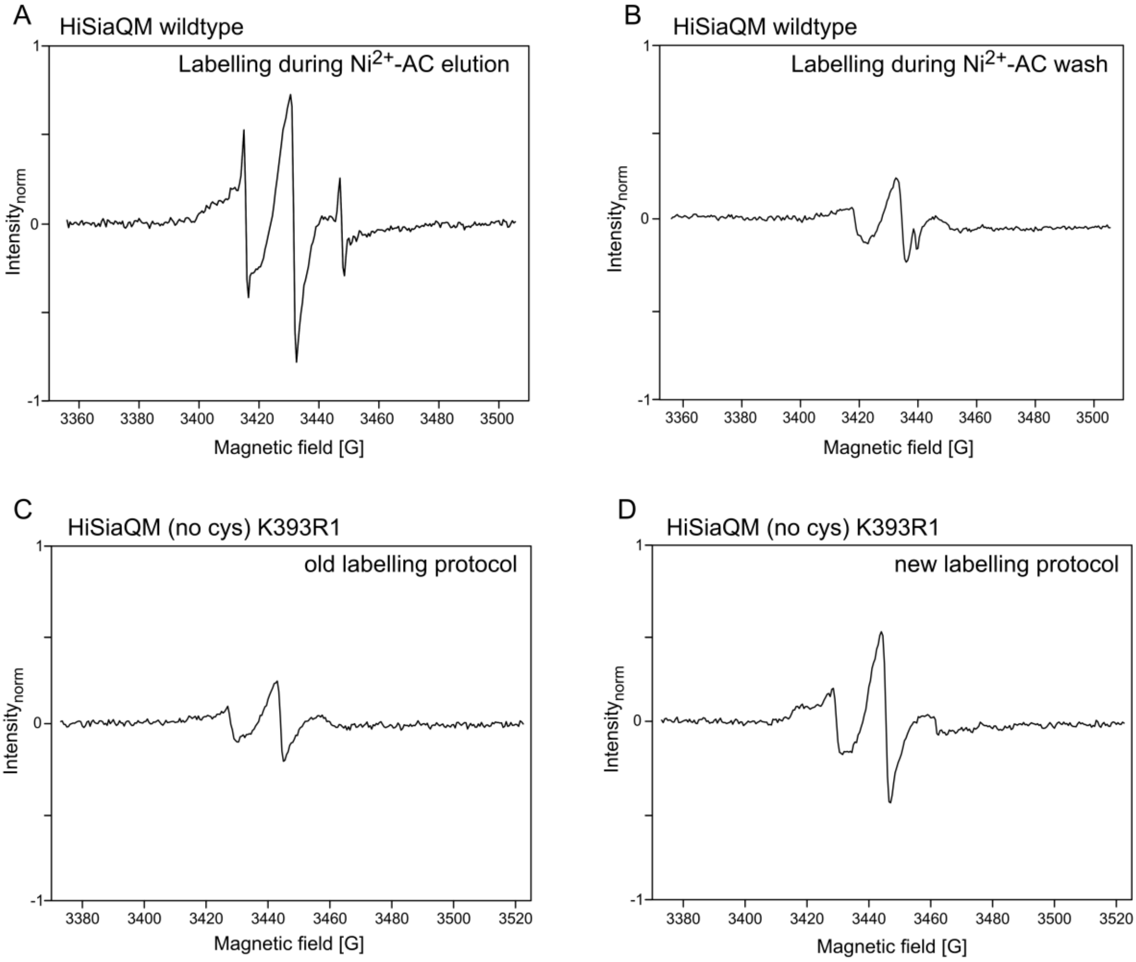


Figure 5-2: Optimization of spin labelling procedure for HiSiaQM. A) X-band RT cw-EPR spectrum of initial labelling with standard labelling protocol for soluble proteins after elution of Ni^{2+}-AC (143 µM protein, 350 scans). All four spectra from the figure are normalized on the same scale. **B)** X-band RT cw-EPR spectrum with labelling during Ni^{2+}-AC washing step (105 µM protein, 900 scans). The reason for the small signal at 3440 G was possibly caused by an unknown artefact. **C)** X-band RT cw-EPR spectrum of a HiSiaQM single cysteine mutant with standard labelling procedure analogous to B) (50 µM protein, 350 scans). **D)** X-band RT cw-EPR spectrum of same mutant as C) but with optimized labelling procedure (50 µM protein, 350 scans).

The labelling protocol was initially applied to the single cysteine mutant K393C of HiSiaQM (introduced in Chapter 5.3). For a direct comparison, the same single cysteine mutant was labelled with the old procedure. The cw-EPR spectra of both samples, with label addition in the elution (old protocol) or in a washing step (new protocol) are shown in Figure 5-2 C and D and revealed a clearly higher intensity of the cw-EPR spectrum with the optimized labelling procedure. A similar effect was observed in Q-band field-sweep spectra, performed prior to each PELDOR measurement, with a significantly increased echo signal intensity (Figure SI-21).

5.2.3 Selective spin labelling of HiSiaQM

The labelling experiments with the HiSiaQM wildtype protein resulted in an unselective binding of MTSSL, presumably due to the native cysteines. In a first attempt to increase the selectivity, the incorporation of unnatural, noncanonical amino acids (UAA) into the HiSiaQM protein sequence was tested, together with Mariam Abd El Fatah during her bachelor thesis. The UAAs have the advantage that they can be introduced into a protein with only minimal invasive changes of the sequence and subsequently labelled with a spin centre via a specific functional group (Fleissner et al., 2009, Kucher et al., 2017). Alternatively, UAAs are described that directly contain a spin centre and need no further modification or labelling after reconstitution into the protein sequence (Schmidt et al., 2014). To receive a double spin labelled HiSiaQM mutant for PELDOR distance measurements, only two amino acids need to be varied and the remaining protein, especially cysteine residues that are presumably relevant for function, can remain unchanged. For introduction of UAAs into the protein sequence during expression, an amber stop codon (TAG) was mutated at desired position. A suppressor plasmid with an UAA-specific tRNA/synthetase gene was co-transformed into the *E. coli* cells and should guarantee the incorporation of the UAA during expression of the protein (Fleissner et al., 2009, Wang et al., 2001).

For the membrane domains HiSiaQM, two amber codon constructs were initially created to decrease the risk of accidental choice of a functionally relevant position. The amino acids E121 and E312 were chosen due to their polarity and an accompanying possible higher chance of surface accessibility than for hydrophobic residues that might be covered by the detergent. Further, these amino acids have been postulated to be located in regions between a TMH and an extramembrane loop, which should represent a more rigid state of the labelling positions (Mulligan et al., 2012).

The incorporation of two UAAs, L-4-acetylphenylalanine (pAcF) and N-propagyl-lysine (PrK), in the HiSiaQM constructs were analysed in several test-expressions (Chatterjee et al., 2013, Nguyen et al., 2009, Young et al., 2011). Experimental conditions such as the UAA concentrations, expression temperature and incubation time were varied. Additionally, different available suppressor plasmids were used for the corresponding unnatural amino acids, pEVolAcF and pUltraCNF for pAcF and pEVolPylRSAF and pEVolPylRSWT for PrK (Chatterjee et al., 2013, Nguyen et al., 2009, Plass et al., 2011, Young et al., 2011). For none of the tested expression conditions, a full-length protein was observed on a His-tag specific western blot. Only smaller protein fragments were detectable, presumably corresponding to a truncated protein (not shown). If the necessary concentration of UAAs or tRNA is not present during expression, the amber codon would result in a terminated protein. Overall, the usage of UAAs for selective spin labelling was not successful and could not be used for labelling of HiSiaQM.

As an alternative attempt for selective labelling of HiSiaQM, the initially used site-directed spin labelling of cysteines was modified. Based on the unselective labelling reaction of the wildtype protein with five native cysteines (Figure 5-2 A), these cysteines were mutated individually to amino acids that are not able to bind MTSSL. The constructed mutants were all expressed and purified, and the accessibility of the remaining cysteines for the spin label MTSSL was investigated through labelling. The final purification step (SEC run and corresponding SDS-PAGE) of each HiSiaQM mutant is shown in Figure SI-22 and the corresponding cw-EPR spectra for detection of the nitroxide spin label in Figure 5-3 A. Clearly, all cw-EPR spectra from mutants with a remaining cysteine in the protein sequence revealed an immobilized spin label which was presumably bound to the biomolecule. In contrast, the cysteine-free mutant of HiSiaQM indicated indeed only a minimal amount of spin label signal with a significant decrease of the signal intensity compared to the other mutants (Figure 5-3 A). The cysteine-free mutant was purified and labelled in a second batch to exclude experimental errors. Importantly, the corresponding cw-EPR spectrum of this second batch showed a similar result as the first protein batch. The comparison of the cw-EPR spectra for the mutants with one and without cysteines suggested a labelling of cysteine 94. However, a mutant with single change of this cysteine 94 showed also a successful binding of MTSSL to HiSiaQM (Figure SI-23). These spin labelling experiments revealed that more than one of the native cysteines in HiSiaQM is accessible and can bind the nitroxide spin label. The cysteine-free mutant is thus the only observed construct that is suitable for selective spin labelling.

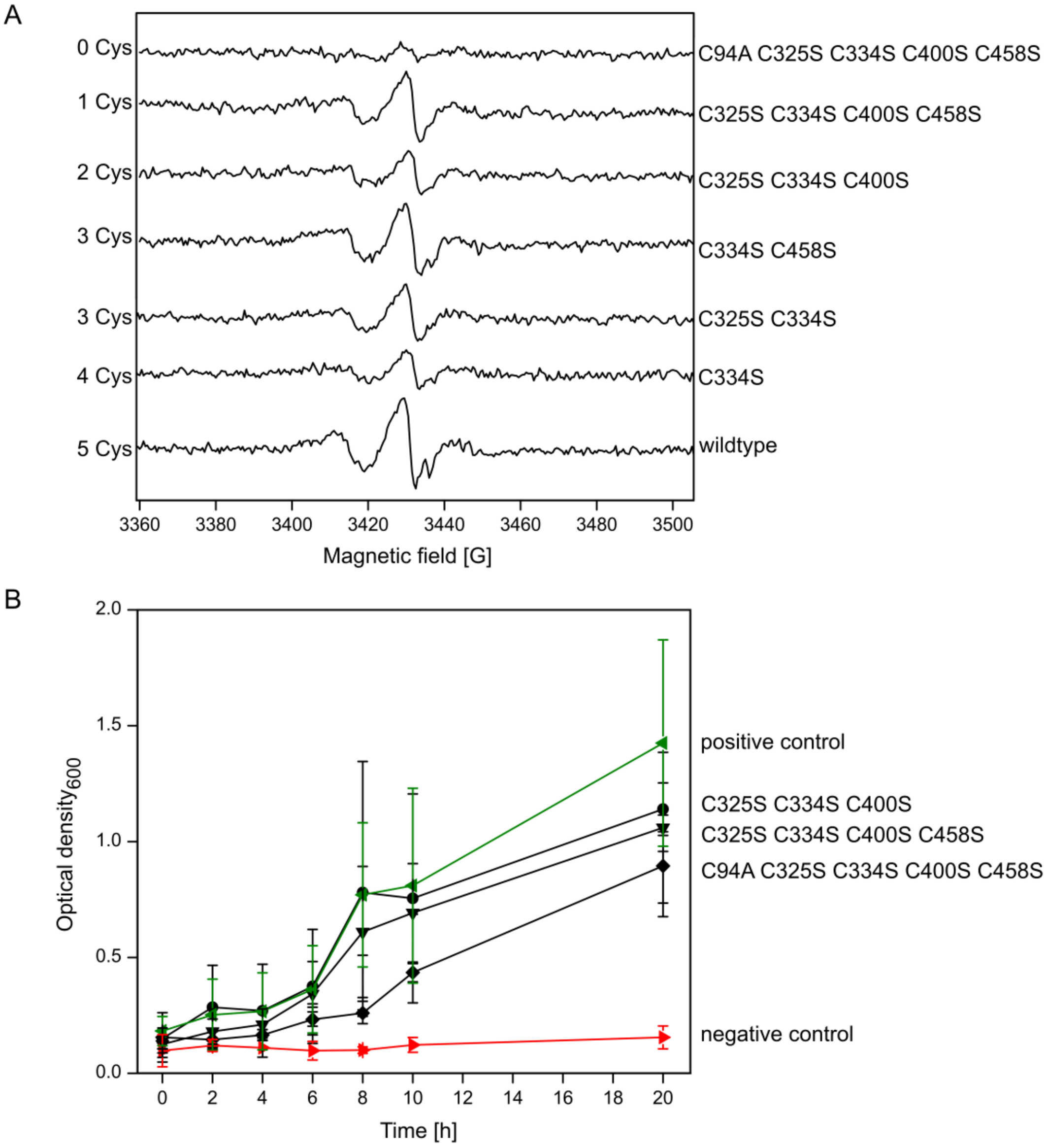


Figure 5-3: Mutations of HiSiaQM for selective labelling. A) Stacked X-band RT cw-EPR spectra from HiSiaQM wildtype and mutants with various number of native cysteines. The mutated residues are given next to the spectra. All spectra are illustrated on the same scale and were recorded under the same conditions (350 scans, 70 µM protein). **B)** Functional growth assay of *ΔnanT E. coli* cells (SEVY3) with the TRAP transporter HiSiaPQM as wildtype (green) and with several mutated native cysteines (black) (Severi et al., 2010). The cells were grown in an environment with sialic acid as single carbon source and the growth was detected with the optical density at 600 nm. A negative control cell culture without TRAP transporter is shown in red. A detailed description of the assay is given in Chapter 9.2.1 and Figure 6-12.

Before this HiSiaQM construct could be used for further applications, especially for functional studies, the functionality of the protein after mutagenesis of the native cysteines was confirmed. For this, a TRAP transporter specific *in vivo* cell growth assay from the lab of Gavin H. Thomas (University of York) was used (see Chapter 9.2.1 and Figure 6-12 for a detailed description). The central component of this assay is a modified *E. coli* strain (SEVY3), which cell growth depends on the uptake of the nutrient sialic acid from the environment with the artificial introduced sialic acid TRAP transporter HiSiaPQM (Severi et al., 2010). Mutagenesis of essential amino acids in the TRAP transporter proteins will consequently result in a non-observable growth of the cell culture (Fischer et al., 2015). To analyse the function of the cysteine-free HiSiaQM construct, cysteine mutations were introduced into the TRAP transporter HiSiaPQM gene for the growth assay. The growth rates of cysteine-mutants, including the non-cysteine mutants, as well as for positive and negative control cultures, are shown in Figure 5-3 B. Except the negative control without a transporter system for sialic acid, all cell cultures showed an increase of the cell density. Even though, the growth of the cysteine mutants is slightly decreased, particularly for the cysteine-free mutant, the TRAP transporter mutants were still functionally active and the cysteine-free construct was used as starting point for selective spin labelling approaches of HiSiaQM.

5.3 Spin labelling of HiSiaQM cysteine mutants for PELDOR spectroscopy

In the following chapters, several cysteine mutants of HiSiaQM were designed and investigated with EPR spectroscopy. All cysteine mutations were inserted into the cysteine-free construct C94A C325S C334S C400S C458S of HiSiaQM. For the sake of a clearer illustrations and a better readability, only the inserted artificial cysteines are mentioned below.

5.3.1 Single cysteine mutants of HiSiaQM

Based on the cysteine-free HiSiaQM construct, artificial cysteines were mutated into the protein to generate several single cysteine mutants. The positions of these cysteines were selected on the basis of the TMH-topology model (Mulligan et al., 2012). They were introduced into postulated loops between two transmembrane helices on both predicted sides of the membrane protein. The choice of mutated amino acids was based on the hydrophilic properties of the residues and polar amino acids were used which should guarantee a high solvent accessibility and no steric hindrance by the detergent micelle. The locations of the mutated residues were distributed over the whole protein in both domains to generate useful sets of labelling locations that can be later used as starting point for the construction of double mutants for distance measurements.

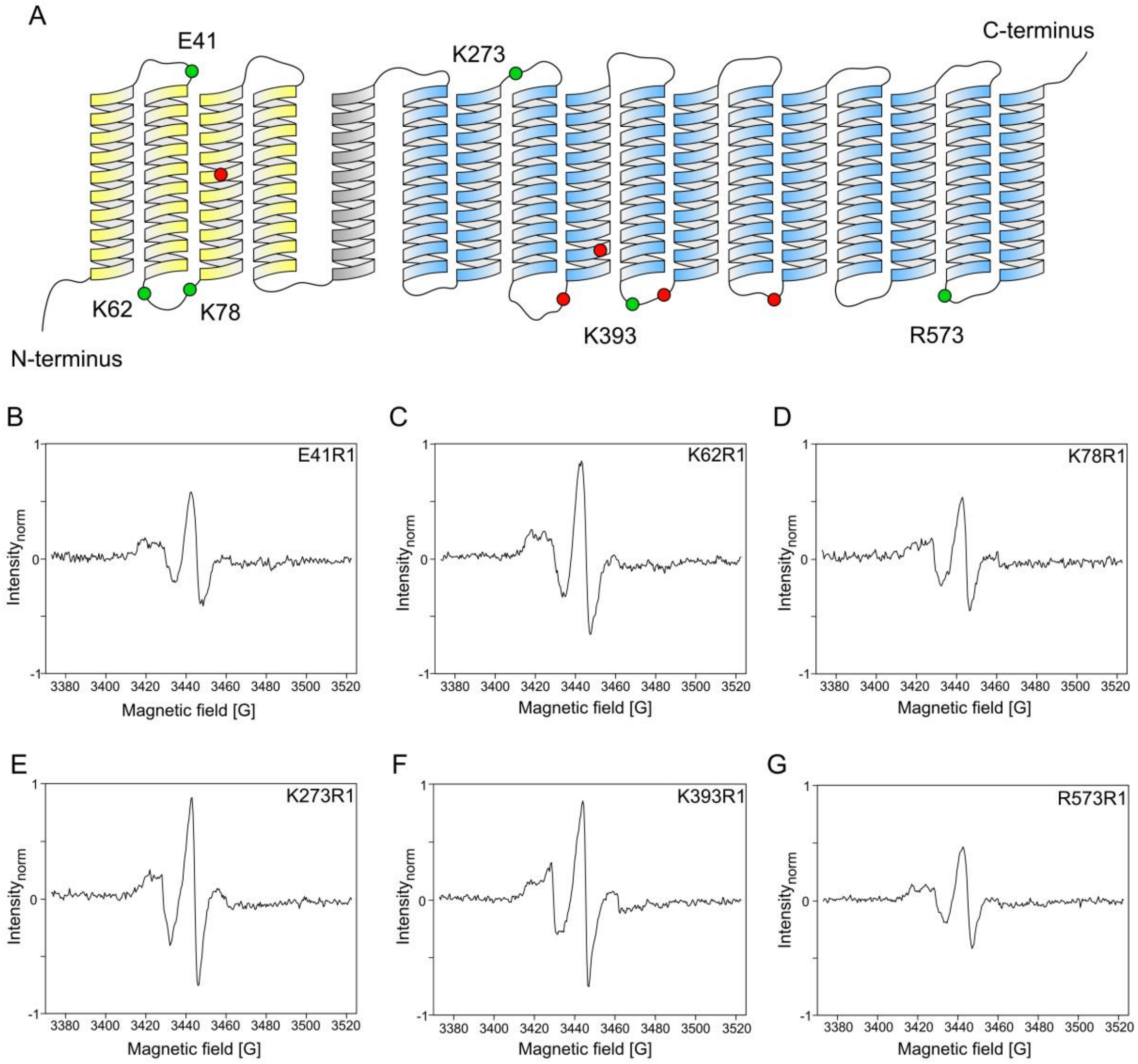


Figure 5-4: Spin labelling of HiSiaQM single cysteine mutants. A) TMH-topology prediction of HiSiaQM, the Q-domain is represented in yellow and the M-domain in blue (Mulligan et al., 2012). The postulated positions of mutated, native cysteines are marked with red circles and for new, introduced cysteines with green circles. **B-G)** X-band RT cw-EPR spectra of single cysteine mutants at mentioned positions. The optimized spin labelling procedure from Chapter 5.2 was used for all mutants. All cw-EPR spectra were recorded with the same experimental parameters, similar protein concentrations and normalized on the same scale.

Overall, six single cysteine mutants of HiSiaQM were constructed (Figure 5-4 A, green circles). The mutants were expressed, purified and labelled with the above described, optimized labelling procedure. The SEC profile of each mutant, as well as a corresponding SDS-PAGE gel, are illustrated in Figure SI-24. For the most labelled single cysteine mutants, except K273R1, the SEC runs revealed a lowered protein yield compared to a HiSiaQM wildtype purification. Additionally, the mutant K393R1 showed an increase of aggregated protein in the SEC elution profile, which was separated during the SEC run. Nevertheless, all SDS-gels indicated bands for HiSiaQM without significant amounts of impurities.

Afterwards, the spin labelled HiSiaQM single cysteine mutants were all analysed with cw-EPR spectroscopy (Figure 5-4 B-G). The spectra indicated a signal of a nitroxide spin label with clear anisotropic triplet peaks and by this, suggested a successful binding of the MTSSL to the introduced cysteines. However, the intensity of the detected signal was rather low and due to the low SNR, a confident quantitative analysis for determination of labelling efficiency was not possible. Though, the cw-EPR spectra of different cysteine mutants revealed small changes of the signal, especially in the spectral region at around 3420 G. These features were presumably caused by a varied mobility of the label at these positions and support the successful labelling procedure at different positions in the protein (NejatyJahromy and Schubert, 2014).

5.3.2 Double cysteine mutants of HiSiaQM

As proposed in previous studies on TRAP transporters, HiSiaQM is presumed to be functionally active as single protein and not fused into dimeric or oligomeric complex (Mulligan et al., 2011). Consequently, to investigate the function of the protein by PELDOR spectroscopy, double spin labelled mutants were required. For this, the designed and characterised label positions from the single cysteine mutants were used as starting point and a variety of double mutants were constructed. The double mutants are presented in Figure 5-5 A on the TMH topology prediction with black arrows between the involved cysteines (Mulligan et al., 2012). Double mutants with label positions on different sides of the membrane, with cysteines in only the Q- or M-domains and between both domains were constructed.

Overall, six double cysteine mutants were prepared and labelled during purification, in the same way as the single cysteine mutants. The final SEC and SDS-gel of each mutant are shown in Figure SI-25 and confirmed the successful purification with very few contaminations in the final sample. However, all four double mutants with a cysteine at position 393 showed again a significant amount of aggregated protein in the SEC run, in contrast to the other mutants. This observation, in combination with similar results from the K393R1 single mutant, suggested an important functionality of lysine 393 for the stability of the protein.

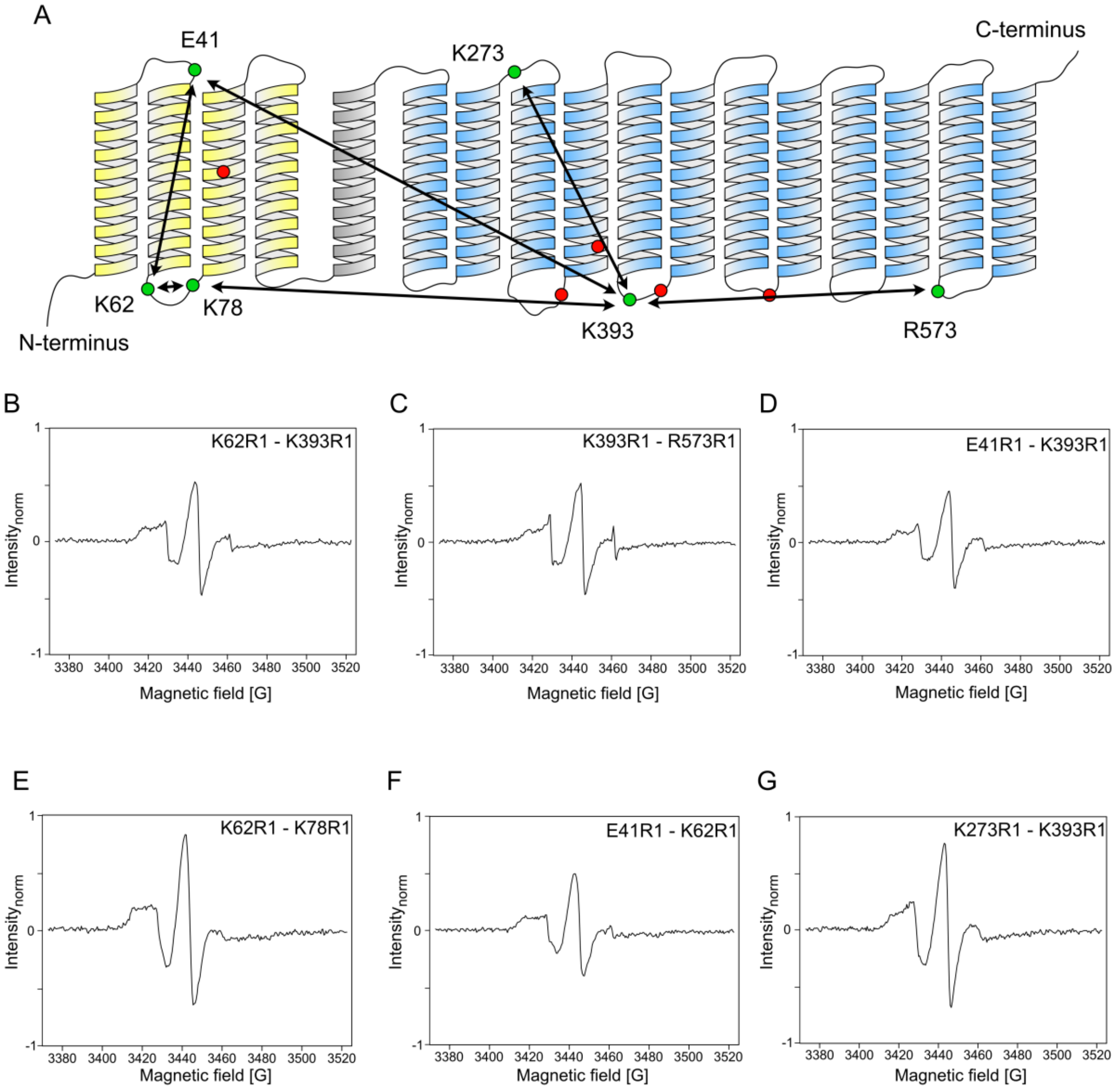


Figure 5-5: Spin labelling of HiSiaQM double cysteine mutants A) TMH-topology prediction of HiSiaQM, the Q-domain is represented in yellow and the M-domain in blue (Mulligan et al., 2012). The postulated positions of mutated, native cysteines are marked with red circles and for new, introduced cysteines with green circles. Constructed double cysteine mutants for distance measurements are presented with black arrows. **B-G)** X-band RT cw-EPR spectra of double cysteine mutant at mentioned positions. The optimized spin labelling procedure from Chapter 5.2 was used. All cw-EPR spectra were recorded with the same experimental parameters, similar protein concentrations and normalized on the same scale.

After separation of the aggregates with SEC, all mutants were concentrated and analysed with cw-EPR spectroscopy. The spectra of all six double mutants revealed anisotropic signals for immobilized, protein-bound nitroxide spin labels (Figure 5-5 B-G). Again, the signal intensities were rather low, even though increased SNRs were observed in comparison to the single mutants. These higher labelling efficiencies were likely caused by the second labelling positions. However, the background from the double mutant spectrum was still too noisy for an unambiguous determination of the label quantification. An estimation with low experimental certainty revealed labelling efficiencies of around 40-50% for the double mutants. Analogous to the single cysteine mutants, qualitative differences of the cw-EPR spectra were observable in the first peak at around 3420 G. These features indicated again a varied mobility of the spin labels at the different cysteine positions. Overall, the use of artificially introduced cysteines allowed the selective spin labelling of HiSiaQM with MTSSL with low, but useful labelling efficiencies, and the subsequent analysis of the mutants with PELDOR spectroscopy.

5.3.3 Spin labelled HiSiaQM in protein-bounded nanodiscs

For several membrane proteins in other studies, essential interactions of the protein with membrane lipids could been proven (Gupta et al., 2017, Laganowsky et al., 2014, Martens et al., 2016, Martens et al., 2018, Sanders et al., 2018). In Chapter 4.3, the successful reconstitution of HiSiaQM into MSP-nanodiscs with a lipid bilayer was presented. In the following chapter, this membrane mimetic system was combined with spin labelled double cysteine mutants, to study a potential influence of the environment, the detergent micelle or lipid bilayer, towards the conformational state of HiSiaQM.

For this, the single spin labelled mutant K273R1 and the double mutant E41R1-K235R1 were used. The latter one was an additionally prepared mutant that was also used for other studies. Both spin labelled mutants were reconstituted into the nanodiscs according to the standard protocol (Chapter 9.1.5). The SEC runs are shown in Figure 5-6 and indicated the previously described two peaks that are typical for HiSiaQM reconstitutions (Figure 4-6). Notably, the intensity of the first peak is slightly decreased, particularly for the double mutant. The SDS-PAGE (Figure SI-26) revealed the expected two bands for HiSiaQM and MSP1 and therefore the fractions from the first peak were concentrated and investigated with cw-EPR spectroscopy. Anisotropic nitroxide signals were observable but the comparison with cw-EPR spectra for same mutants before reconstitution revealed a strongly decreased signal intensity and lower SNR of the spectra (Figure 5-6). One possible reason for this observation was a different protein concentration between HiSiaQM in the nanodiscs and in detergent micelles. The membrane scaffold proteins and the DMPC lipids, as well as the presence of empty nanodiscs, hindered an accurate determination of the HiSiaQM concentration in the nanodisc sample. However, slight amounts of immobilized spin labels were detected and the reconstituted HiSiaQM mutants were investigated in the following chapters with PELDOR spectroscopy.

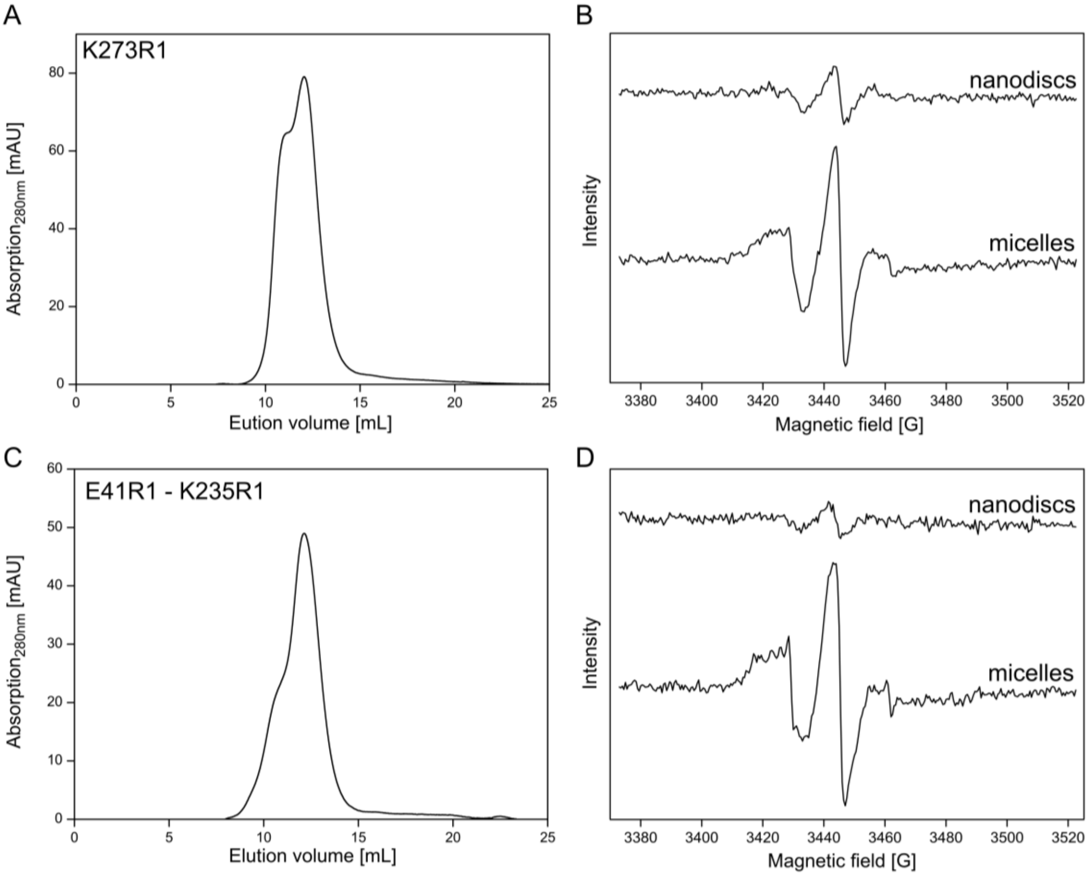


Figure 5-6: Reconstitution of spin labelled HiSiaQM in protein-bounded nanodiscs. A) SEC run of HiSiaQM K273R1 after reconstitution into MSP-nanodiscs with DMPC lipids on an SD 200 10/300 column (buffer: standard QM-buffer without detergent and glycerol). **B)** X-band RT cw-EPR spectra of the first protein-nanodisc peak in A) (top) and the same mutant in DDM micelles (bottom). The same experimental spectrometer settings were used and the spectra are illustrated on the same scale. **C-D)** Same as A-B) but for HiSiaQM E41R1 K235R1 double mutant.

One interesting observation was made during reconstitution of two other spin labelled mutants into MSP-nanodiscs (K393R1 and K273R1-K393R1). After standard overnight dialysis, a white precipitant was visible in the protein solution for both mutants. The precipitant was separated by centrifugation and the supernatant applied onto a SEC column. The resulting chromatograms revealed only one peak, corresponding to the second peak of the wildtype protein reconstitution (Figure SI-26). Additionally, the SDS-gels indicated just an observable protein band for MSP1, which suggested the aggregation of the spin labelled HiSiaQM mutants during reconstitution (Figure SI-26). These observations are in agreement with the previously detected high degree of aggregation of HiSiaQM mutants with mutated lysine residue 393. Due to an unknown effect of this mutation onto the protein, the protein constructs with mutation K393C were further analysed and interpreted with caution.

5.3.4 Spin labelling of HiSiaQM in polymer-bounded nanodiscs

In addition to the protein-bounded nanodiscs, the HiSiaQM membrane domains were also successfully reconstituted into SMA polymer-bounded nanodiscs (Chapter 4.4). As mentioned before, the great advantage of the polymer-based extraction is the solubilization of membrane proteins in their native lipid bilayer. To investigate a possible influence of the native lipid bilayer to the conformational state of HiSiaQM with PELDOR spectroscopy, HiSiaQM cysteine mutants were reconstituted into SMA-nanodiscs and incubated with the spin label MTSSL. Noticeable, no spin labelling experiments or PELDOR spectroscopy measurements with SMA polymer extracted membrane proteins were published until today. Based on the labelling experiences with DDM detergent micelles, two double cysteine mutants were expressed and extracted into SMA polymer-bounded nanodiscs and labelled with the optimized labelling strategy during purification (K62C-K78C and E41C-K62C). The SEC runs of both mutants and SDS-gels are illustrated in Figure SI-27 and revealed a typical result with a few impurities in the sample that were already observed and characterized for HiSiaQM wildtype experiments before (Chapter 4.4).

The cw-EPR spectra of both double spin labelled mutants showed the typical signal for an immobilized nitroxide spin label (Figure 5-7 A). The intensities of the signals were slightly lower than for the same mutants in detergent micelles (Figure 5-7 B). Possible explanations for this might be the quite high amounts of impurities or a deviating protein concentration in the nanodisc samples, due to UV/Vis absorption of SMA at 280 nm (Oluwole et al., 2017a). Because of the high amounts of impurities, the SMA-extraction and labelling procedure was repeated for the cysteine-free HiSiaQM mutant (SEC and SDS-gel in Figure SI-27). Surprisingly, the cw-EPR spectrum showed a clear, anisotropic nitroxide spin label spectrum, similar to the double cysteine mutants (Figure 5-7 A). To rule out experimental errors, a control experiment with a second batch of cysteine-free HiSiaQM was repeated, resulting in a very similar cw-EPR spectrum. In contrast, the same HiSiaQM cysteine-free mutant in detergent micelles showed no spin label signal (Figure 5-7 B). These results suggested a non-selective labelling of the SMA polymer-bounded nanodiscs, possibly caused by an interaction of the spin label with impurities or the native lipid bilayer.

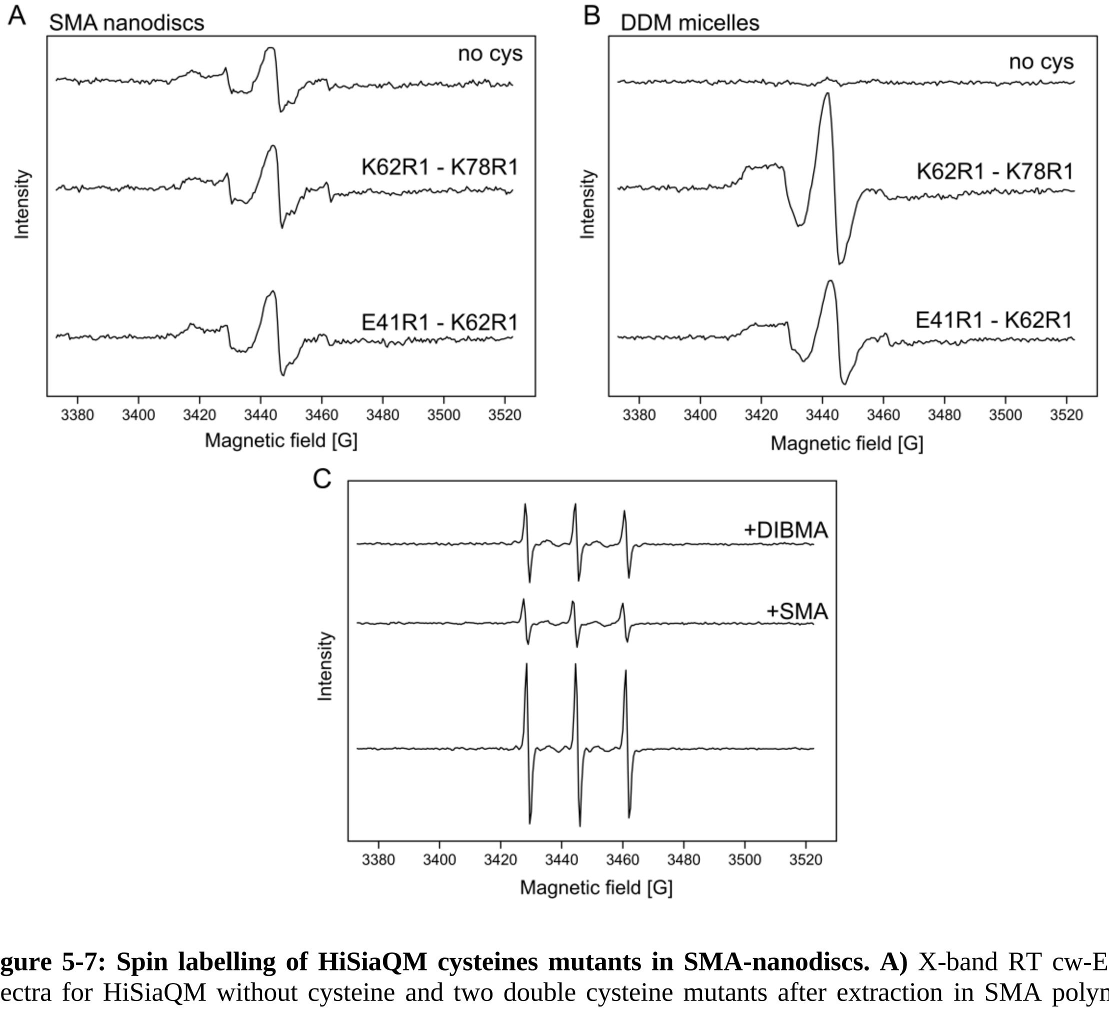


Figure 5-7: Spin labelling of HiSiaQM cysteines mutants in SMA-nanodiscs. A) X-band RT cw-EPR spectra for HiSiaQM without cysteine and two double cysteine mutants after extraction in SMA polymer nanodiscs. The proteins were labelled with same procedure as DDM detergent samples. The spectra were detected with equal experimental set-up and illustrated on the same scale. **B)** Same as A) but with HiSiaQM solubilized in DDM micelles. **C)** X-band RT cw-EPR spectra of a 100 µM MTSSL solution without additives and after addition of 6.6% of SMA and DIBMA polymer with the same experimental set-up and scale.

During investigations of SMA-nanodisc samples with cw-EPR spectroscopy, a decrease of the spin label signal intensity over time was detected. To validate this observation, a control measurement with unbound MTSSL in solution without and with addition of the polymers SMA and DIBMA were performed. The cw-EPR spectra revealed an isotropic signal for all three samples with a significant decrease of signal intensity for samples with supplemented SMA or DIBMA polymers (Figure 5-7 C). The reason for this influence of the two polymers on the spin label signal in a cw-EPR experiment remained elusive. As mentioned before, there are currently no studies available in which polymer extracted membrane proteins are labelled and analysed with cw-EPR or PELDOR spectroscopy. Overall, due to an unselective labelling and unknown influence of the polymer on the spin label MTSSL, HiSiaQM in polymer-bounded nanodiscs were not used for PELDOR spectroscopy.

5.4 PELDOR spectroscopy on HiSiaQM

After selective spin labelling, the HiSiaQM mutants were analysed with PELDOR spectroscopy. As described before, due to the unknown structure of the protein, the distance measurements were more focused on qualitative changes of the detected PELDOR time traces than on a structural interpretation of distinct conformational changes. The experiments were performed in detergent micelles, measurements in protein-bounded nanodiscs are mentioned.

5.4.1 PELDOR measurements on HiSiaQM

The first PELDOR experiments were performed with single spin labelled mutants of HiSiaQM to investigate a possible dimerization or oligomerization of the membrane domains. The PELDOR time traces of six single mutants are presented in Figure 5-8. All spectra indicated an initial decrease of the signals but no clear oscillations. Interestingly, the decrease of the time traces was more pronounced for the single mutants located in the Q-domain (E41R1, K62R1 and K78R1) than for M-domain mutations (K273R1, K393R1 and R573R1), resulting in a higher modulation depth of the first three mutants.

The observed non-oscillation PELDOR time traces may result from an intermolecular coupling or background decay without a defined distance between two or more spin labels.

Another possible explanation, especially for the PELDOR time traces of spin labels in the Q-domain, might be indeed a flexible HiSiaQM dimer or oligomer with a broad distribution of distances between the spin labels. However, with the detected PELDOR time traces, a possible dimerization or oligomerization of HiSiaQM in solution can be neither reliably confirmed, nor negated.

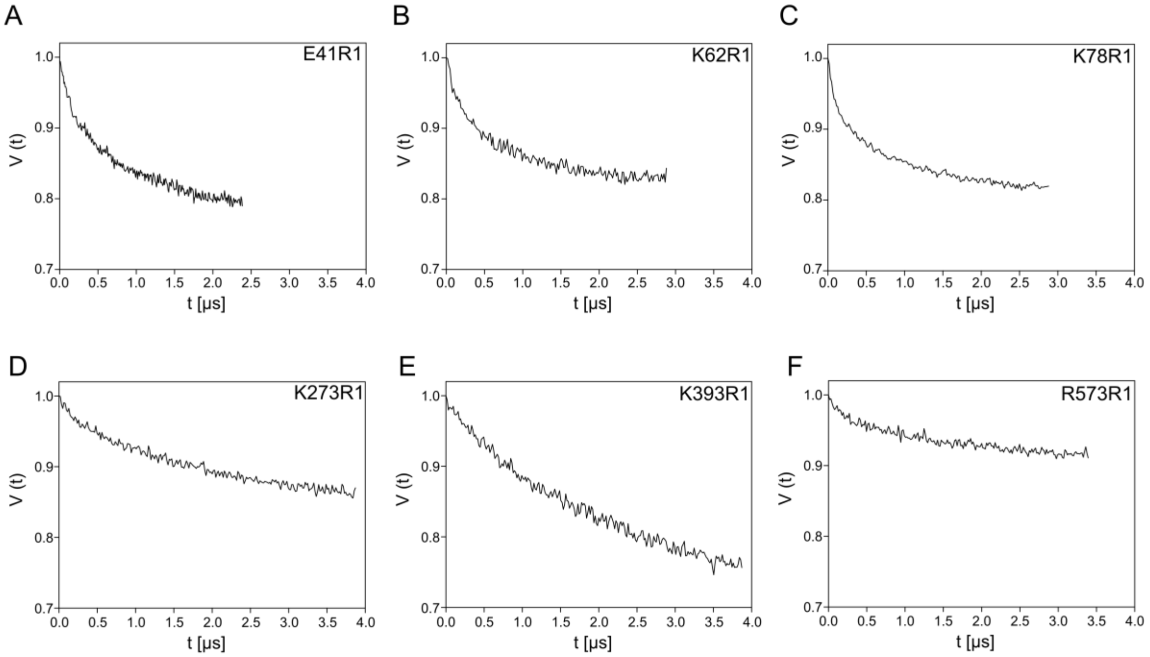


Figure 5-8: PELDOR time traces for HiSiaQM single mutants. A-F) Raw Q-band PELDOR time traces for HiSiaQM single cysteine mutants after labelling with MTSSL, corresponding to the cw-EPR samples in Figure 5-4. All PELDOR samples were prepared and measured with the same concentration and in the same way.

Following these experiments, the double spin labelled HiSiaQM mutants were investigated. The raw PELDOR time traces of six double mutants are presented in Figure 5-9 and revealed significant changes for different label positions. For example, for three mutants (K393R1-R573R1, K62R1-K78R1 and E41R1-K62R1) an initial fast decay of the PELDOR time traces was observed, indicating a rather short distance between the spin labels. Interestingly, the two spin labels of each of the three mutants are located in the same domain of HiSiaQM. The other three mutants, either with spin labels in the same (K273R1-K393R1) or different domains (K62R1-K393R1 and E41R1-K393R1), showed a long decay of the PELDOR time trace. Again, this could be result from a background decay of intermolecular coupled spins or from a broad distribution of distances between the two introduced spin labels. For all six PELDOR time traces, the modulation depth was around 10%-30% and by this, significantly deeper than for the single mutants (Figure 5-8 and Figure 5-9). However, none of the HiSiaQM double mutants showed a clear oscillation in the PELDOR time traces. Especially for the first three mentioned double mutants, K393R1-R573R1, K62R1-K78R1 and E41R1-K62R1, this might be caused by a high conformational flexibility between the two spin label positions in HiSiaQM. For the other three mutants, it remained elusive if the signal results from a background decay or indicates a long distance with large oscillation frequency, which are not resolved due to the length of the time trace.

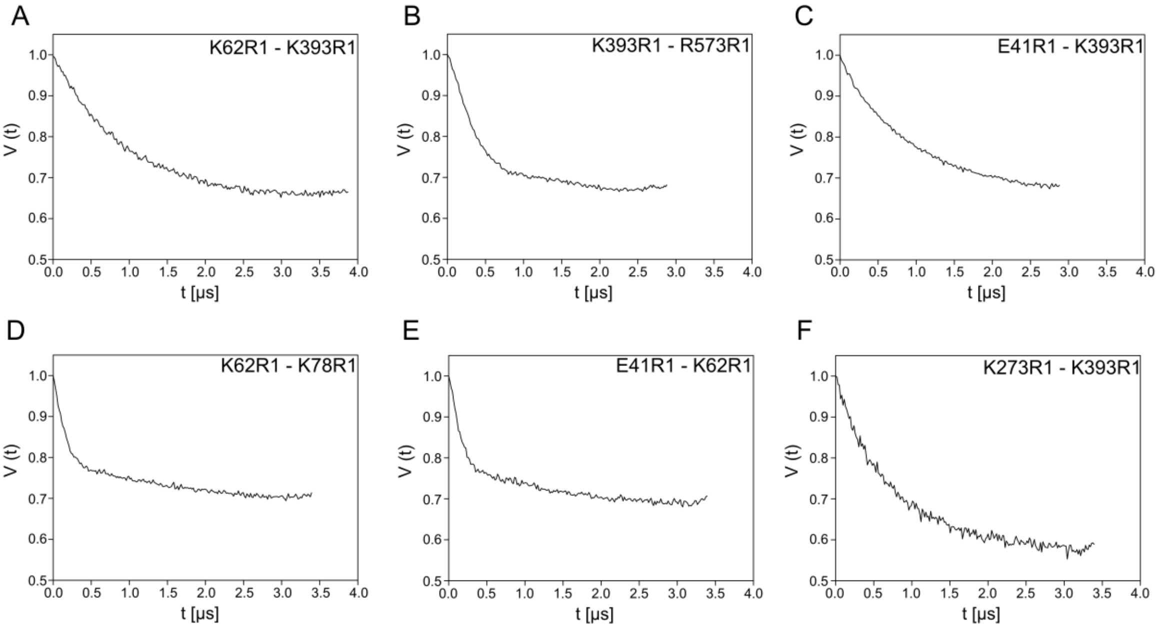


Figure 5-9: PELDOR time traces for HiSiaQM double mutants. A-F) Raw Q-band PELDOR time traces for HiSiaQM double cysteine mutants after labelling with spin label MTSSL, corresponding to the cw-EPR samples in Figure 5-5. All PELDOR samples were prepared and measured with the same concentration in the same way.

For comparison of the PELDOR results with the protein model, distance distributions were calculated with DeerAnalysis (Jeschke et al., 2006) (Figure 5-10 B and C). Further, *in silico* distance distributions were simulated with determination of spin label ensembles on the protein model from Ovchinnikov et al. with the mtsslWizard software (Figure 5-10 A and

C) (Hagelueken et al., 2015, Ovchinnikov et al., 2014). The *in silico* distances were additional used for simulations of PELDOR time traces, to validate the oscillation that would be expected for the theoretical distance (Figure 5-10 D, calculated by Dinar Abdullin, University of Bonn).

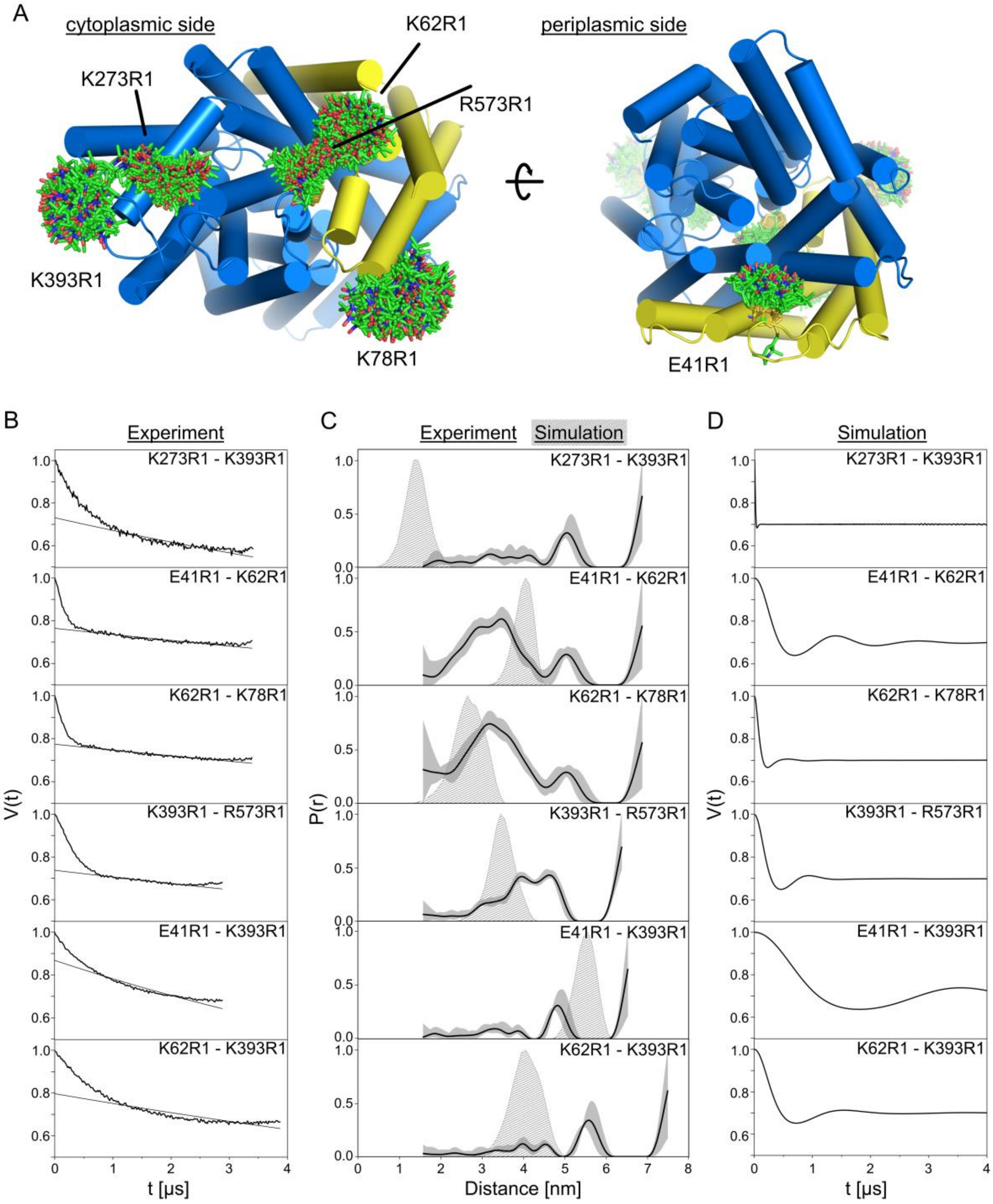


Figure 5-10: Comparison of PELDOR distances with HiSiaQM model. A) Model of HiSiaQM after Ovchinnikov et al. (Ovchinnikov et al., 2014) with views from the cytoplasmic (left) and periplasmic (right)

side. The Q-domain is presented in yellow and the M-domain in blue. Spin label ensembles of MTSSL (green) at the cysteine-mutated positions were calculated with mtsslWizard (Hagelueken et al., 2015). **B)** Raw Q-band PELDOR time traces for HiSiaQM double cysteine mutants, the background-corrections are indicated as black lines and were calculated with DeerAnalysis (Jeschke et al., 2006). **C)** PELDOR distance distributions (black lines) with validation (grey areas) from experiments in B), calculated with DeerAnalysis (Jeschke et al., 2006). The *in silico* distances (grey-white strips) were calculated with the mtsslWizard (Hagelueken et al., 2015) by using the protein model in A). **D)** Simulations of PELDOR time traces from the *in silico* distance distributions from C) (calculated by Dinar Abdullin, University of Bonn).

As expected from the experimental time traces without oscillations, the PELDOR distance distributions were quite broad. For example, for the three mutants K393R1-R573R1, K62R1-K78R1 and E41R1-K62R1, the experimental distances were in a range of 2-5 nm. Nevertheless, the comparison of the simulated distances from the protein model showed a good fit to these experimental data (Figure 5-10 C). But importantly, due to the broad distributions, the comparison of the protein model with the experimental data must be validated with care. For the other three mutants, K273R1-K393R1, K62R1-K393R1 and E41R1-K393R1, the experimental PELDOR distance distributions were even broader and more unspecific, due to the long decay of the time traces. According to the protein model, mutant K273R1-K393R1 indicated a distance that would be shorter than 2 nm and by this in the limiting distance region of PELDOR spectroscopy (Jeschke, 2012). For mutant E41R1-K393R1, the simulated distance was quite large and it might be possible, that the PELDOR time trace was too short to describe the oscillation for such a distance in a sufficient manner. However, as mentioned before, it remained elusive if the experimental PELDOR time traces for these three double mutants are caused by a background decay, a too small time scale of the experimental time trace or a very high structural flexibility of the protein. A reliable comparison between the experimental results and the protein model was therefore not possible.

5.4.2 PELDOR on HiSiaQM with addition of transporter important components

To investigate the transport mechanism by changes of the PELDOR time traces, the PELDOR experiments of certain HiSiaQM spin labelled mutants were performed after supplementation of compounds that potentially interact with the membrane domains.

For classical secondary transporters, the binding of their substrate by the transmembrane domains was observed in binding studies and in crystal structures (Reyes et al., 2009, Weyand et al., 2008). However, for SBP-dependent TRAP transporters, a binding of the substrate was only described for the SBP and not for the membrane domains (Johnston et al., 2008, Müller et al., 2006). To investigate such an interaction for HiSiaQM and a possibly accompanying change or stabilization of a conformational state, the double mutants E41R1-K62R1 and K393R1-R573R1 were treated with an excess of Neu5Ac. These two mutants were chosen, due to their presence in both membrane domains and the previously observed distance features in their PELDOR time traces. For the PELDOR experiments after addition of Neu5Ac, both mutants again revealed a fast decay of the PELDOR time traces. However, the comparison with the measurements without substrate showed no significant change, suggesting no changes of conformational states through the presence of the substrate (Figure 5-11 A and B).

As known from previous studies, TRAP transporters have an additional SBP which is essential for substrate transport and postulated to deliver the substrate to the transmembrane domains through formation of a tripartite complex (Forward et al., 1997, Mulligan et al., 2011). Importantly, the formation of the tripartite complex is based on assumptions such as the mechanisms of ABC transporters and has never been observed experimentally before, including experiments in our group (personal communication with Janin Glänzer, University of Bonn). To investigate the formation of a tripartite complex and similar as before, an accompanying change of stabilization of a conformational state of HiSiaQM, PELDOR experiments after addition of HiSiaP were performed. As before, the two double mutants E41R1-K62R1 and K393R1-R573R1 were used. The PELDOR time traces are shown in Figure 5-11 C and D and revealed again no significant changes that would represent another distance between the spin labels.

To exclude a tripartite complex formation without conformational rearrangement of the spin labelled regions in the two HiSiaQM mutants, single spin labelled HiSiaQM mutants were supplemented with a single spin labelled HiSiaP protein. In the case of a defined tripartite complex formation of the transporter domains, a single distance should be obtained in the PELDOR experiment. For this approach, a single cysteine HiSiaP mutant with a mutational introduced cysteine into the backbone helix of the protein at position K254C were prepared and labelled, similar as for the homologous protein VcSiaP in a previous study (Peter et al., 2021) (SEC, SDS-PAGE and cw-EPR spectrum are shown in Figure SI-28). Due to a second labelled species, the PELDOR time trace changed but no modulation or oscillation was detected that would be expected for a defined distance (Figure 5-11 E and F).

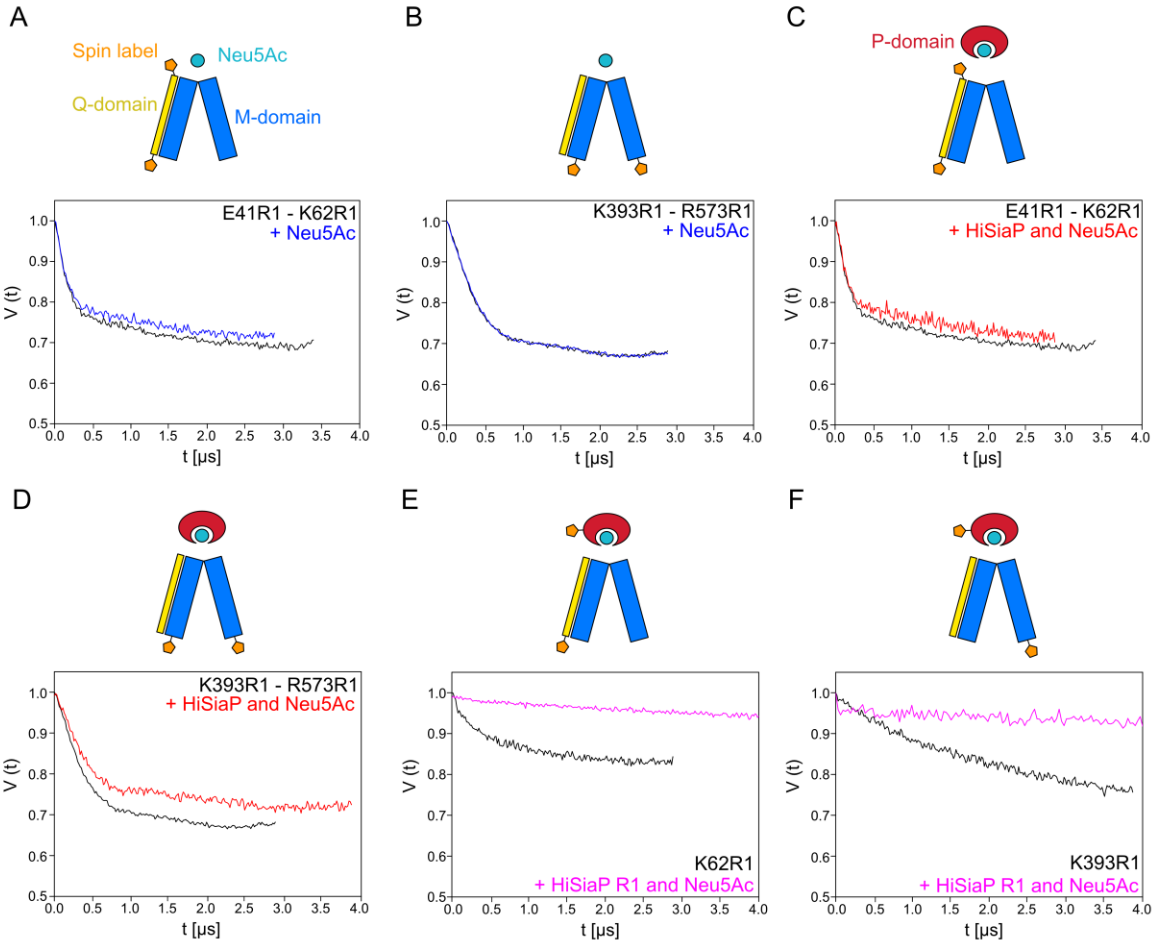


Figure 5-11: PELDOR time traces for labelled HiSiaQM with additives. A) Raw Q-band PELDOR time traces for HiSiaQM mutant E41R1-K62R1 without additives (black) and after supplementation of native transporter substrate sialic acid (blue) (20 mM sialic acid). A schematic of the transporter components that were present in the experiment are given above. **B)** Same as A) but for other cysteine mutant and with 1 mM sialic acid. **C)** Same as A) but with addition of HiSiaP wildtype protein (red) in 1:1 ratio (1 mM sialic acid). **D)** Same as C) but for cysteine double mutant K393R1-R573R1. **E)** Raw Q-band PELDOR time traces for HiSiaQM mutant K62R1 without (black) and with addition of single spin labelled mutant HiSiaP K254R1 (magenta) in 1:1 ratio (1 mM sialic acid). **F)** Same as E) but for HiSiaQM K393R1.

In the next experiments, the influence of the membrane mimetic system was investigated in a similar way, by analysing changes in the PELDOR time traces for different measurements. Because the spin labelling of HiSiaQM in polymer-nanodiscs with native membranes was unselective, the focus was set on a comparison between HiSiaQM in detergent micelles and after reconstitution into an artificial lipid bilayer in protein-bounded nanodiscs. For this experiment, the single spin labelled mutant K273R1 and the double mutant E41R1-K235R1 were used (Figure 5-6). The PELDOR time traces of the double spin labelled mutant in detergent micelles and protein-bounded nanodiscs are shown in Figure 5-12 A and revealed a comparable result with no changes that would refer to a changed conformational state. In contrast, for the single spin labelled mutant, the PELDOR time traces showed indeed small differences with a changed background between the two membrane mimetic systems (Figure 5-12 B). However, the time traces for the single mutant in both membrane mimetic systems yielded no certain or defined signals that indicated a changed conformational state as a consequence of another membrane mimetic system around the membrane protein. Possible reasons for the different time traces are discussed in the following Chapter 5.5.

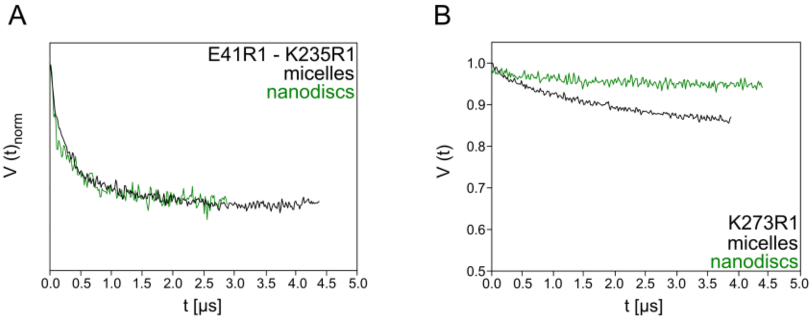


Figure 5-12: PELDOR time traces for labelled HiSiaQM in detergent micelles and protein-bounded nanodiscs. A) Raw Q-band PELDOR time traces for HiSiaQM mutant E41R1-K235R1 in DDM detergent micelles (black) and after reconstitution in protein-bounded nanodiscs (green). For comparison, the background of the PELDOR time traces were normalized to the same scale. **B)** Same as A) but with HiSiaQM single cysteine mutant K273R1 and without normalization of the background.

5.5 Discussion and conclusion

The sections above report the progress from an unselectively labelled HiSiaQM protein towards several PELDOR measurements with the membrane domains under several conditions and in two membrane mimetic systems. Overall, a selective labelling procedure with good labelling efficiency was established and several mutants were analysed with PELDOR spectroscopy for measurements of distances between spin labels.

5.5.1 PELDOR spectroscopy uncovers a high structural flexibility of HiSiaQM

The recorded PELDOR time traces of the HiSiaQM double mutants revealed broad distributions of the spin label distances, visible as no or only a low oscillation of the PELDOR time traces. These results suggest a high structural flexibility of the membrane protein in solution with no specific conformational state. The analysed mutants further showed partially differences in the time traces, suggesting a different conformational flexibility between the labelled positions for selected double mutants. Such a structural flexibility and the resulting broad distance distributions for membrane proteins were observed several times before in different PELDOR experiments, for example for the secondary multidrug transporter MdfA (Yardeni et al., 2019), the secondary MATE transporter NorM (Claxton et al., 2018), the ABC transporter McjD (Bountra et al., 2017), the secondary sugar transporter vSGLT (Paz et al., 2018) and for bacteriorhodopsin (Orwick-Rydmark et al., 2012). Especially for secondary membrane transporter, such a high structural flexibility is considered to be necessary for the function of the protein (Boudker and Verdon, 2010, Oldham et al., 2007, Thomas and Tampé, 2020). According to these models, the substrate translocation in secondary transporters includes at least a rearrangement of the transmembrane helices between an outward open to an inward open conformation. If none of these conformational states are stabilized in solution, the protein is presumably able to switch uncontrolled between these states. The positioning of the spin labels for distance measurements in such regions will result in broad distance distribution.

One additional effect for membrane proteins that can cause or promote a destabilization of defined conformational states is the change of the environment from the native lipid bilayer to a membrane mimetic system. In a recent study on secondary transporters, a conserved cytoplasmic network of amino acids was identified that is responsible for the change between conformational states of the proteins (Martens et al., 2018). In that work, nanodisc preparations showed an interaction of lipids with this network, resulting in a strong influence on the conformational changes. The strength of interaction and influence depended also on the type of lipid, PE (phosphatidylethanolamine) for example showed a higher effect on the conformational change from the outward-faced to the inward-faced conformation than PC (phosphatidylcholine) (Martens et al., 2018). This study suggests that the removal of the native lipids can have an enormous effect on the stabilization of conformational states and encourage uncontrolled switching between these states, as presumably observed with

PELDOR spectroscopy for HiSiaQM. The previous study may also explain the unchanged PELDOR time traces after reconstitution of HiSiaQM into protein-bounded nanodiscs. Because the influence of the lipids on the conformational states changed for different lipids, the use of only one kind of lipid for the HiSiaQM nanodiscs might not fulfil such a necessary interaction.

The observed changed background for the PELDOR time trace of single spin labelled HiSiaQM in detergent micelles and protein-bounded nanodiscs possibly results from different sizes of the membrane mimetic systems. Due to this, the distance between two membrane proteins is different for the two systems, including the distance of coupled intermolecular spin labels. For the case of the double spin labelled mutant, the PELDOR signal is mainly caused by an intramolecular coupling and no changes were detectable. Even if no structural features could be obtained from the PELDOR experiments in different membrane mimetic systems, especially the single labelled mutant indicated the successful reconstitution of HiSiaQM into the nanodiscs.

5.5.2 Influence of additives on the conformational states of HiSiaQM

Because broad distance distributions were expected for the PELDOR experiments on HiSiaQM, one focus was set on changes in the detected distance upon addition of transporter-important components. The addition of the substrate sialic acid revealed no changes of the conformational state, which could result from either no binding of the substrate through HiSiaQM or no influence on the conformational state of the protein. A missing, direct binding of the substrate by the membrane domains might be an interesting observation that would indicate a different behaviour between TRAP transporters and classical secondary transporters, which are known to directly bind the substrate (Reyes et al., 2009, Weyand et al., 2008). This particularity might explain the essential presence of all three TRAP transporter components and highlights the SBP-dependence for substrate translocation (Forward et al., 1997, Mulligan et al., 2011). However, for a reliable hypothesis, the unable substrate binding through the membrane domains need to be confirmed with further binding experiments, for example with ITC or SPR measurements.

The possible inability of direct interaction between sialic acid and the membrane domains promotes the idea of a tripartite complex formation of all transporter domains, as postulated for the function of TRAP transporter in previous work (Mulligan et al., 2011). However, with different set-ups regarding the spin label position, no formation of a tripartite complex was observed in the present PELDOR experiments. As mentioned above, these observations are in accordance with previous studies, in which the tripartite complex was never observed and described before by using other methods for example cw-EPR spectroscopy, microscale thermophoresis or analytical SEC (personal communication with Janin Glänzer, University of Bonn). One possible factor that can decrease the tripartite complex formation is the high structural flexibility of the membrane domains in the detergent micelles, as observed in the PELDOR experiments without additives. It might be possible, that the SBP can only interact with a certain conformation of HiSiaQM that is present in the native membranes but underrepresented or even not exist in the detergent solubilized protein. However, the reasons for the undetected tripartite complex might be manifold and for following studies, many more experimental parameters should be modified to possibly promote the complex formation, for example another membrane mimetic system or the presence of a sodium-ion gradient. Additionally, the complex formation should also be simultaneously analysed with alternative binding methods, to exclude an influence of the modified cysteines and introduced spin labels.

5.5.3 Outlook

The PELDOR experiments on the TRAP transporter membrane domains revealed new insights into the conformational states of HiSiaQM and also showed the uncertainty that arises from such distance determining experiments without a known protein structure. Therefore, a main goal prior to following PELDOR experiments should be the determination of the protein structure. This will not only allow a better cross-validation of the PELDOR results but also enhance the prediction of suitable labelling sites. Essential parameters such as label accessibility or measurable distance between two label positions can be determined on a protein structure with a much higher certainty as with the TMH topology model. As shown in a previous study on a secondary transporter, the choice of labelling positions for PELDOR experiments can have a crucial influence on the detected distance between the spin labels (Claxton et al., 2018). In this study on another membrane protein, the change of labelling positions by just a few amino acid numbers resulted in significant different detected distance distributions and by this, allowed a more sensitive investigation of conformational changes. Alternatively, without the availability of a protein structure, the preparation of more mutants can be helpful to get a mutant by chance with a more defined distance and an accompanying sharper distance distribution in the PELDOR experiment.

As mentioned in the discussion above, the analysis of the PELDOR experiments suggests several experiments that might be interesting for further studies of the TRAP transporter membrane domains. One possibility for example is a further variation of additives or conditions such as the salt concentration or pH-value, that showed a strong influence on the conformational states of another secondary transporter in a recent study (Claxton et al., 2018). Another central point that should be considered is the impact of the membrane mimetic system on the protein. For further studies, the use of a variation of artificial lipids in protein-bounded nanodiscs can be crucial to analyse the function of HiSiaQM. Another promising approach regarding the environment of the membrane protein, that would possibly circumvent challenges from the detergent micelles or protein-bounded nanodiscs approaches, are polymer-bounded nanodiscs. However, these reconstitutions showed an inexplicable incompatibility with the nitroxide spin label MTSSL in the cw-EPR experiments. One alternative is the use of other types of spin labels such as trityl radicals, which are more inert and might facilitate functional EPR studies in this system (Fleck et al., 2020). If successful, these experiments would present the first functional PELDOR studies on solubilized membrane proteins in their native lipid bilayer and can importantly increase the knowledge about protein-lipid interaction.

Chapter 6

Nanobodies for TRAP transporter membrane domains

Preamble:

The immunization of an alpaca was carried out by Paul-Albert König and Jan Tödtmann from the nanobody core facility of the University of Bonn. Further, the facility performed the phage display, as well as the first identification steps of the nanobodies. The SPR measurements were performed with Karl Gatterdam from the Institute of Structural Biology (ISB), University of Bonn.

6.1 Nanobodies - small helpers for structural biology

A promising tool for protein structure elucidation that was successfully used in several studies during the last years are nanobodies (Muyldermans, 2013). Nanobodies are functional parts from an antigen-binding domain of heavy-chain antibodies that can be standardly generated within months through immunization of a camelid with the target antigen protein (Pardon et al., 2014). In general, nanobodies are small and stable proteins with a molecular weight of around 15 kDa and they have a high affinity and specificity in the nanomolar range to the target protein (Pardon et al., 2014).

During the last years, nanobodies have an increasing field of application in research, diagnostics and therapeutics, for example fluorescent nanobodies for detection of antigens in living cells, marking of epidermal growth factor receptor proteins from cancer cells for a positron emission tomography (PET) detection or neutralization of bacterial toxins (Hussack et al., 2011a, Rothbauer et al., 2006, Vosjan et al., 2011). In the last months, the urgent need and search for drugs against SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2) is also partially focused on nanobodies and in several studies, an influence and inhibiting effect of nanobodies against specific target proteins from the virus could be observed (Hanke et al., 2020, Huo et al., 2020, Koenig et al., 2021, Wrapp et al., 2020, Xiang et al., 2020). In addition to these applications, the specific binders are often used as supporting tools for structure determination of proteins and were already successfully applied to membrane proteins for X-ray crystallography (Rasmussen et al., 2011b, Rasmussen et al., 2011a). The advantage of nanobodies during structure elucidation is their ability to bind specific epitopes and by this, support the formation and stabilization of weakly populated conformational states or complexes (Pardon et al., 2014). For membrane proteins, especially fully membrane embedded secondary active transporters in a large detergent micelle or nanodiscs reconstitution, the binding of nanobodies also increases the region for crystal forming intermolecular interactions and by this, enhance the crystallization procedure to yield higher quality diffraction data.

Further, nanobodies are also helpful for structural studies with cryo-EM and several high-resolution EM protein structures in complex with a nanobody or a nanobody-derived binder were recently described (Hofmann et al., 2019, Laverty et al., 2019, Uchański et al., 2021). In one of these examples, a nanobody facilitated the structural description of the ABC export transporter TmrAB in eight different conformational states (Hofmann et al., 2019). The specific binding of the nanobody was used to break a twofold pseudo symmetry of the two functional domains that hindered the processing towards a protein model. The determination of the different conformational states allowed a detailed description of the function during transport of this membrane protein. In another examples, an enlarged nanobody construct, named megabody, was used to solve a high-resolution structure of a β3-GABA$_A$ receptor with cryo-EM (Uchański et al., 2021). The specific binding-megabody can facilitate to overcome the size-limitation of cryo-EM and to prevent a possible preferentially particle orientation during vitrification (Uchański et al., 2021).

Despite the application of nanobodies for structure determination, they can be used for functional investigations of membrane proteins. As presented before, nanobodies are for example helpful during PELDOR spectroscopy studies as target of the necessary spin label, such that the investigated membrane protein does not have to be modified (Galazzo et al., 2020). Another interesting application of nanobodies regarding membrane proteins is a disrupting effect onto the transport function, as observed for the vitamin B12 ABC transporter BtuCD-F and a vesicular glutamate secondary transporter (VGLUT) (Mireku et al., 2017, Schenck et al., 2017). Here, the addition of a specific nanobody revealed a reduced transport of the substrate through blocking of conformational states or essential interactions and can be thereby potentially used as compound to selectively inhibit the transport process.

6.2 Initial identification and production of nanobodies for HiSiaQM

6.2.1 HiSiaQM purification for alpaca immunization

This chapter describes the way towards obtaining of specific nanobodies from an alpaca for the TRAP transporter membrane domains HiSiaQM. An overview about the entire procedure is given in Figure 6-1. The immunization of an alpaca for generation of nanobodies was accomplished with the HiSiaQM wildtype protein. The protein was recombinantly expressed in LB medium, extracted and solubilized with DDM detergent and purified after the established standard protocol using Ni^{2+}-AC and SEC (Chapter 9.1.4.2). The protein was concentrated to 3.5 mg/mL, flash frozen in 11 aliquots with 60 µL and provided to the nanobody core facility (University of Bonn).

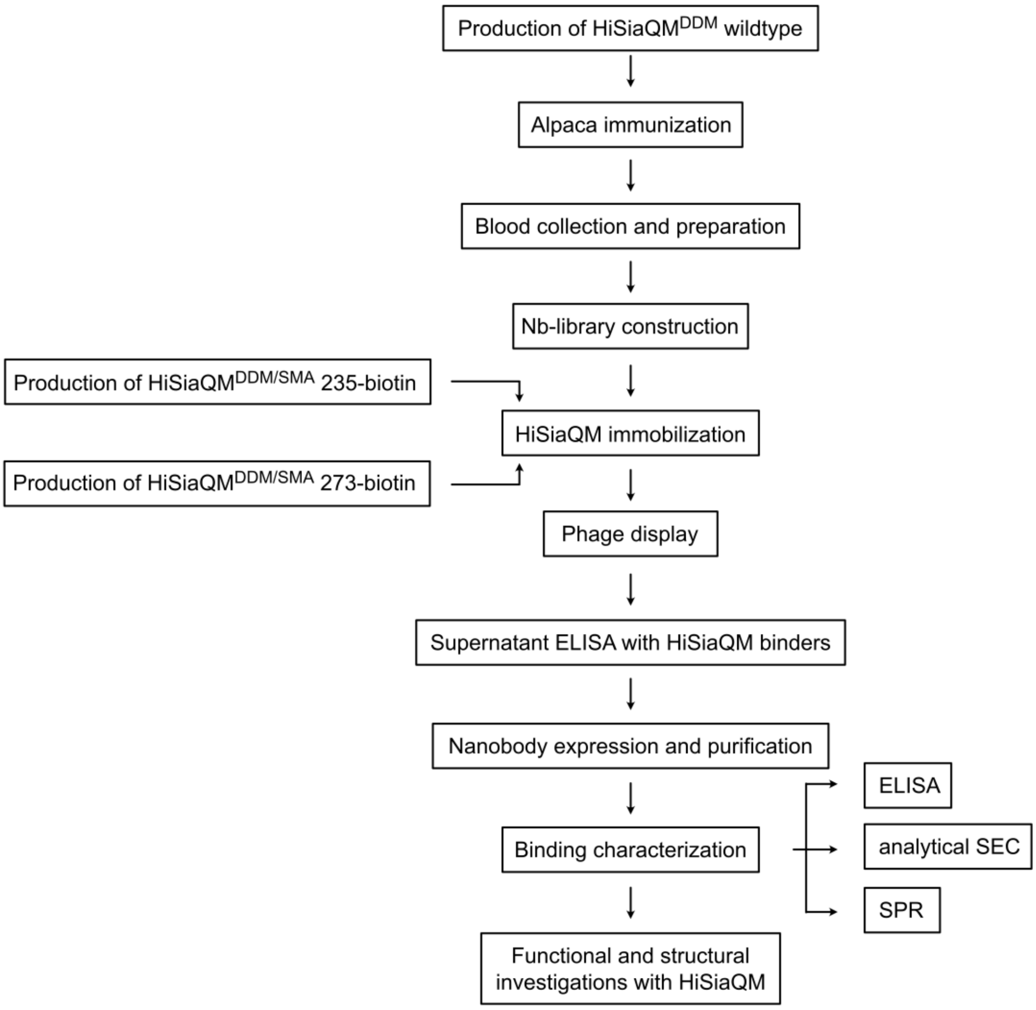


Figure 6-1: Overview about process of generating nanobodies. The figure was adapted and modified after (Dmitriev et al., 2016) and (Pardon et al., 2014).

For identification of binding nanobodies in the following chapters, the HiSiaQM protein, had to be immobilized on several stationary phases as a nanobody-antigen. In the standard procedure for soluble proteins, the antigens are often randomly biotinylated using unselectively labelling of lysines and subsequent immobilized on a streptavidin coated surface (Schmidt et al., 2016a). The unselective labelling of presumably several lysines on the surface of the protein provides a mostly successful immobilization in different orientations that ensure the accessibility of many possible nanobody binding sites. Unfortunately, due to the membrane mimetic system, the solvent accessible regions for membrane proteins are limited, especially for small, fully membrane embedded secondary transporter such as TRAP transporters.

Therefore, for all immobilization steps, two single cysteine HiSiaQM mutants, that were designed in the previous PELDOR experiments (Chapter 5.3.1), were selectively labelled with a maleimide-functionalized biotin reagent. Noticeable, these single cysteine-mutants were based on the cysteine-free HiSiaQM in which five native cysteines were mutated to an alanine or a serine (Chapter 5.3.1). However, the native cysteines are mainly predicted in the transmembrane regions (Figure 5-1) and a binding of the nanobodies to these positions, that might be disrupted in the cysteine-free mutant, is rather unlikely. The positions of the artificially introduced cysteines were selected on the basis of the two available topology models of HiSiaQM, which can be distinguished by the direction of the second transmembrane helix of the M-domain (Figure 6-2 A) (Mulligan et al., 2012, Ovchinnikov et al., 2014). The two cysteines were positioned in the loop before and after the second transmembrane helix, creating the single cysteine mutants K235C and K273C, respectively. Even though the positions were chosen using a model, the cysteine positions on different sides of the models should guarantee that every side was once available for immobilization and once as free side for nanobody binding (Figure 6-2 B). Nevertheless, even though the exact orientation of the membrane protein was unknown, a nanobody that is selected with this method may contain information about a side-selectivity.

The two single cysteine mutants K235C and K273C were purified with solubilization in DDM detergent micelles and once with extraction in SMA-nanodiscs and labelled with the biotin reagent during purification as described in Chapter 9.1.6.3. The final SECs and corresponding SDS-gels are shown Figure 6-2 C. The chromatograms of the detergent solubilized mutants revealed sharp, monodisperse peaks at 11 mL elution volume, similar to the wildtype. Small amounts of aggregated proteins were obtained in the void volume and successfully separated from the protein peak. The SDS-PAGEs of the fractions from the main peak indicated a high purity of the target HiSiaQM protein. The SECs for HiSiaQM in polymer-bounded nanodiscs showed a high amount of aggregated protein but also a well separated protein peak at around 11 mL elution volume (Figure 6-2 C). While the protein concentration for the mutant K235-biotin was too low for detection on an SDS-PAGE and for further use, the SDS-PAGE for mutant K273-biotin indicated an expected protein band around 55 kDa.

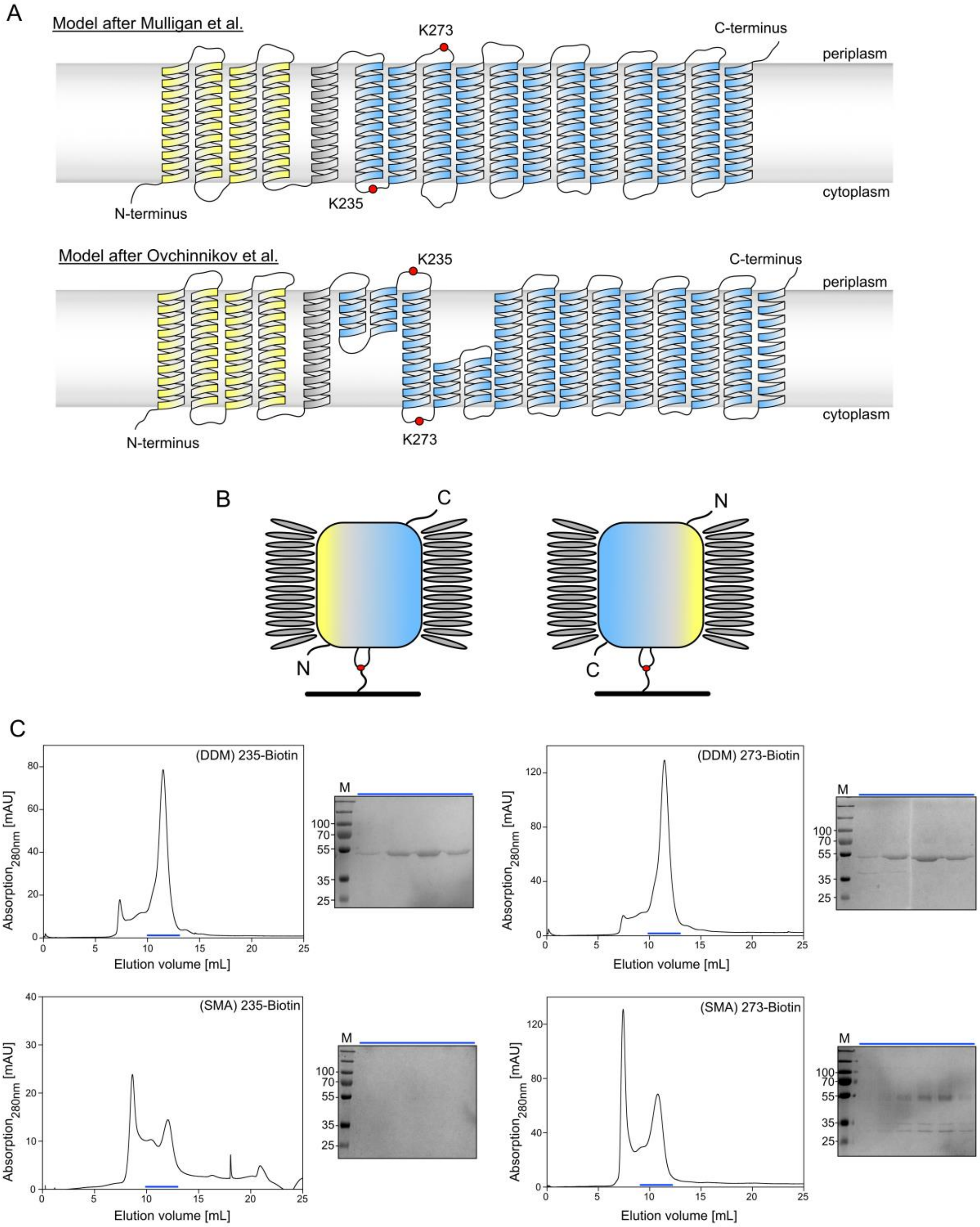


Figure 6-2: Immobilization strategy for HiSiaQM and purified mutants. A) TMH-topologies of HiSiaQM after Mulligan et al. (Mulligan et al., 2012) (top) and after Ovchinnikov et al. (Ovchinnikov et al., 2014) (bottom). The Q-domain is coloured in yellow, the M-domain in blue and the connecting TMH in grey. The positions of the artificial introduced cysteines for biotinylation are marked in red. **B)** Immobilization strategy of the protein HiSiaQM on a surface with periplasmic side (side of N-terminus, left) and with cytoplasmic side (side of C-terminus, right) by using selective biotinylated single cysteine mutant. The DDM-detergent is presented in grey oval circles and the colour of the QM-domain is similar to A). **C)** SEC runs of biotinylated HiSiaQM mutants K235C (left) and K273C (right) in DDM detergent micelles (top) and polymer-bounded nanodisc (bottom) (SD 200 10/300). The corresponding SDS-gels are shown on the right side of the SEC.

6.2.2 Identification of binding nanobodies

After immunization of an alpaca, collection and preparation of a blood sample and construction of a nanobody library, a phage display was the first step in isolation and identification of HiSiaQM specific nanobodies (performed by the nanobody core facility, University of Bonn) (Figure 6-1). For panning of the phages, the biotinylated HiSiaQM mutants K235-biotin and K273-biotin in detergent micelles were immobilized as antigens on a streptavidin-coated surface. After the first panning, around 60 phage-infected colonies were obtained for the antigen HiSiaQM K235-biotin and 96 colonies for HiSiaQM K273-biotin (Figure 6-3 A). Due to the low number of colonies in both approaches, no second round for further selection was carried out.

All observed colonies from both antigens were cultivated in small-scale approaches and nanobodies were expressed with partially leakage into the supernatant of these cultures. The supernatants were subsequently used for an enzyme-linked immunosorbent assay (ELISA). For each supernatant, five ELISAs with different antigens were performed, as illustrated schematically in Figure 6-3 B. The specific biotinylated mutants were immobilized on a streptavidin-coated surface and the HiSiaQM wildtype sample and GST (control) were added directly on the ELISA plate without further modification. The mutant HiSiaQM K235-biotin in SMA-bounded nanodiscs were not used, due to the low yield of protein after purification. The number of binding hits between the nanobodies in the supernatants and the different HiSiaQM antigens are shown in Figure 6-3 B. Overall, 31 wells from different supernatants were identified with positive binding hits. The nanobody plasmids for all hits were analysed via sequencing for the determination of the corresponding nanobody sequence. The resulting sequences are listed in Figure SI-29 in connection with the HiSiaQM antigens that were used for the panning and ELISA. The 31 unique sequences were clustered into 9 different nanobody groups, based on a sequence difference of 3% (Figure SI-29).

The positive hits for the different ELISA antigens clearly showed the advantage of the single biotinylated HiSiaQM mutants in contrast to the wildtype protein. For the wildtype, only four hits with three different sequences were identified, while the use of mutant K273-biotin in detergent micelles resulted in a total number of 20 hits. The small number of hits for the wildtype protein can be caused by at least two possible reasons, a low quantity of immobilized protein on the surface or a preferred immobilization orientation, which result in a blocking of the nanobody epitopes. But importantly, the wildtype protein also revealed binding for one nanobody (5), that was not observed with one of the other immobilization strategies. This strongly supports that the simultaneous use of the wildtype protein was also helpful to obtain a maximum number of potential nanobodies. Interestingly, the HiSiaQM 235-biotin mutant in DDM detergent micelles showed no positive hits in both ELISAs, even though several colonies after phage display were observed for this mutant. Possibly, this effect was caused by an experimental error or an unfavoured effect of this HiSiaQM mutant such as a fast aggregation, but a concrete explanation remained elusive.

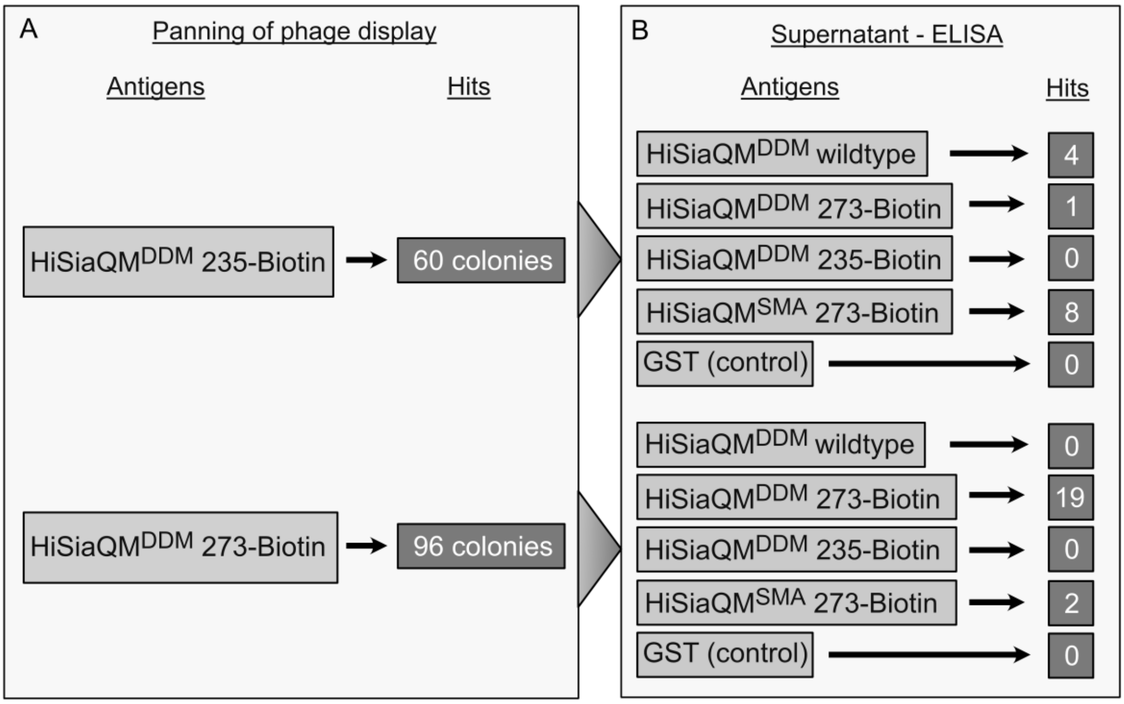


Figure 6-3: Templates and hits of pannings and ELISAs. A) The templates that are used for two different panning approaches are shown in bright grey and the hits from each template in dark grey. **B)** The different antigens that were used in a supernatant ELISA with the positive hits from the panning are shown in bright grey and the number of positive hits for each ELISA is shown in dark grey. The sequence of each hit is shown in Figure SI-29 in connection with the used HiSiaQM mutants.

In contrast, for the HiSiaQM (DDM) antigen 273-biotin in the ELISAs, a clear dependence of the identified nanobodies for the different panning antigens was observed. For the combination of HiSiaQM 273-biotin (DDM) in the panning and subsequent ELISA, a high number of 19 hits was detected. On the other hand, from the colonies after HiSiaQM (DDM)

235-biotin panning, only one hit was observable in the ELISA with HiSiaQM (DDM) 273-biotin (Figure 6-3). However, because no binding hits in the ELISA for the antigen HiSiaQM (DDM) 235-biotin were observed after panning with the same antigen, the reliability of a side-selective nanobody identification is quite low.

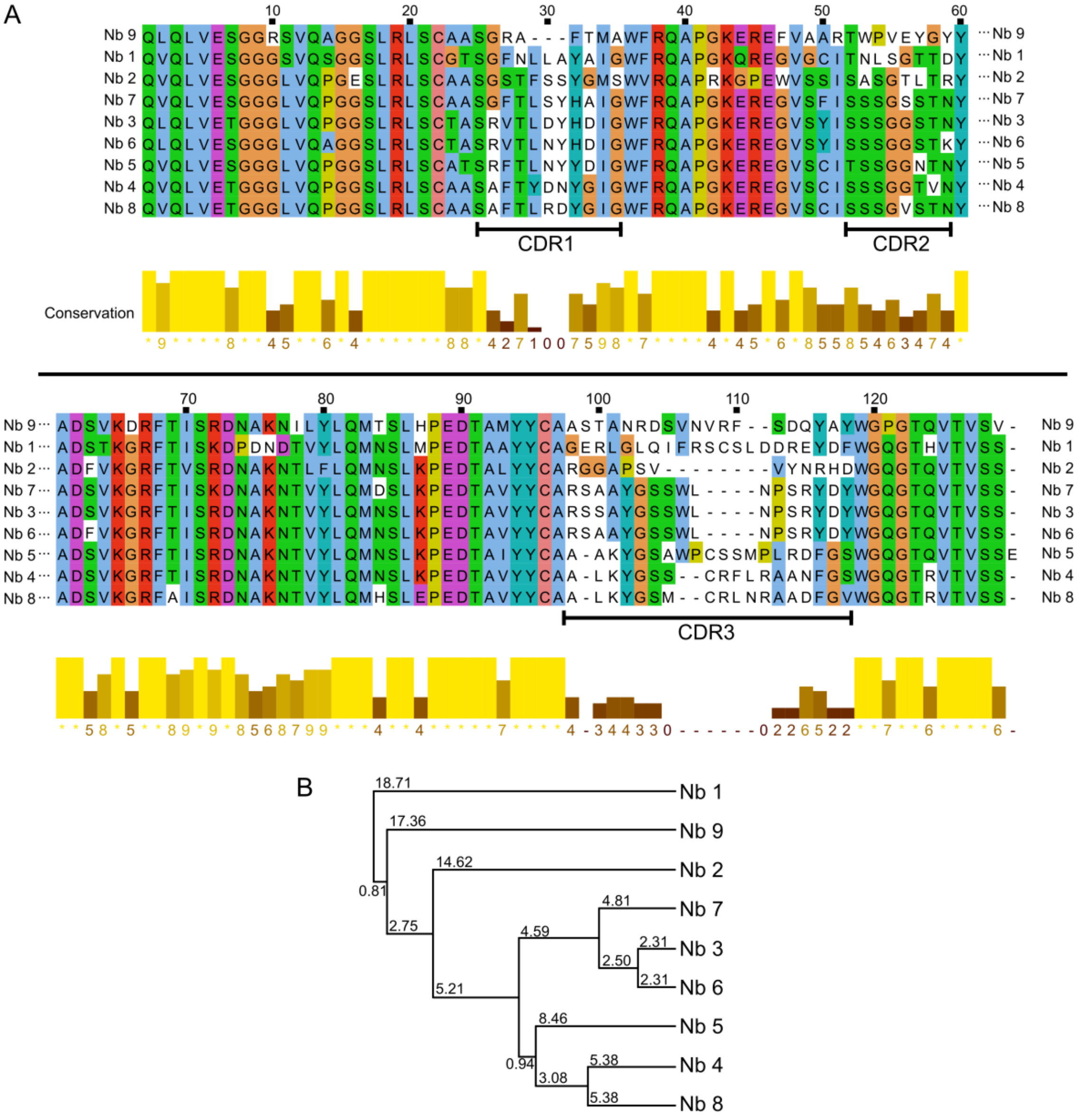


Figure 6-4: Identified nanobodies for HiSiaQM. A) Sequence alignment and conservation of the identified nanobodies (Nb), calculated with Clustal Omega (Sievers et al., 2011) and illustrated with Jalview (Waterhouse et al., 2009). The three complementarity-determining regions (CDRs) are marked. **B)** Average distance tree of the nine nanobodies from A), calculated with Jalview by percentage identity (Waterhouse et al., 2009).

The usage of different membrane mimetic systems for the HiSiaQM antigens revealed some inconsistent binding hits in the ELISAs, compared to the corresponding mutant in detergent micelles. However, the characterization of nanobody binding to HiSiaQM in polymer-bounded nanodiscs in the following chapters indicated an incompatibility of this membrane mimetic system with nanobodies, summarized and discussed in Chapter 6.5.2. Therefore, a reliable interpretation of the ELISA results with this antigen is not feasible. Importantly, all ten hits for the polymer-bounded nanodisc antigen were also observed with the detergent-solubilized antigens (Figure 6-3 and Figure SI-29).

Nevertheless, several nanobodies with variations in their sequences were successfully identified as specific TRAP transporter binders. The results from the use of different antigens during the identification procedure support the advantage of several constructs and immobilization strategies of membrane proteins to generate a useful set of nanobodies. Furthermore, the identification of unique nanobodies with different antigens provided nanobodies with diverse binding properties. From each of the nine clustered nanobody groups, one representative nanobody was chosen for further investigations (nanobody (Nb) 1-9, Figure 6-4 A). The comparison of the nanobody sequences clearly revealed the three nanobody-specific complementarity-determining regions (CDRs), which are mainly responsible for the binding of the antigen (Figure 6-4 A) (Dmitriev et al., 2016). The phylogeny of the nanobodies in Figure 6-4 B represents the sequence similarity of the nanobodies and was used for a comparison with subsequent analysed binding behaviour, such as competitive binding studies for an epitope mapping (Chapter 6.3.4).

6.2.3 Expression and purification of nanobodies

For a more detailed analysis of the binding between the membrane protein HiSiaQM and the identified nanobodies, the latter ones were expressed and purified in separated small-scale approaches. The purity of the final protein fraction from each nanobody was checked via SDS-PAGE (Figure 6-5). All fractions indicated a clean protein band of around 15 kDa (molecular weights ranking from 13.5-15.6 kDa). The yields of all nanobodies were sufficient for first binding experiments.

For further functional and structural investigations, especially for crystallization attempts, higher amounts of the nanobodies were required. Therefore, after initial characterization of the binding in Chapter 6.3, the nanobodies 1, 3, 4, 5, 6 and 9 with an unambiguous binding to HiSiaQM were expressed in approaches with 1 L medium. Despite higher yields, the protein was purified with a SEC column, which usually increase the purity of the protein sample compared to the small-scale approach. The SDS-gels of SEC fractions from all six purified nanobodies are presented in Figure SI-30. All nanobodies were obtained with only small impurities and with a yield of around 10-40 mg per L medium.

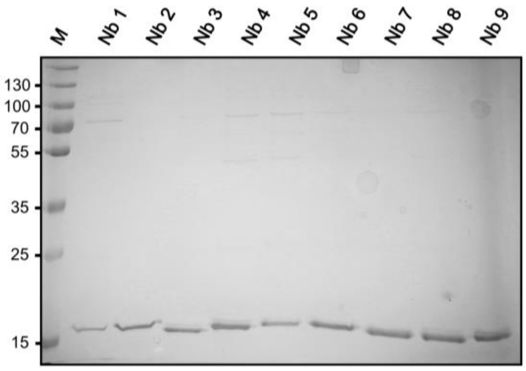


Figure 6-5: Small scale purification of nanobodies. SDS-PAGE of the nine identified nanobodies for HiSiaQM after small-scale expression and purification (1 µg protein was used for each fraction).

6.3 Binding characterization of HiSiaQM nanobodies

6.3.1 Clean-up ELISA

After first identification of the binding nanobody in the previous chapter, a second ELISA was performed with purified nanobodies. This experiment should confirm or negate the binding of nanobodies from the previous experiments and in case of positive hits, should give a first impression of the binding strength. As before, three HiSiaQM antigens (HiSiaQM (DDM) K237-biotin, HiSiaQM (DDM) K235-biotin and HiSiaQM (SMA) K273-biotin) were immobilized onto an ELISA plate surface via streptavidin-biotin interaction.

Figure 6-6 A illustrates the result of the ELISA for the antigen HiSiaQM 273-biotin in detergent micelles after incubation with the nine purified nanobodies. To estimate the binding strength, each nanobody was used in a concentration gradient. For five nanobodies (Nb 3, 4, 5, 6 and 9), a binding was observed with clear dependence of the signal intensity on the concentration (Figure 6-6 A). For comparison and quantification, the absorption of each well in an ELISA plate was measured and plotted against the nanobody concentration (Figure 6-6 B-D). The ELISAs for the different HiSiaQM antigens showed only a detectable binding of the nanobodies for HiSiaQM K273-biotin mutant in detergent micelles. For this antigen, if the highest concentration of the nanobody is assumed as saturated signal, the affinity for nanobody 5 and 9 can be roughly estimate to a low nM-range binding affinity. Similar to the first ELISA experiments with nanobodies from a culture supernatant, the antigen HiSiaQM K235-biotin in DDM indicated no binding of any of the nine nanobodies, even if nanobody 1 and 5 are exclusively obtained from the panning approach with this functionalized HiSiaQM mutant. Also, for the antigen HiSiaQM K273-biotin in SMA polymer nanodiscs, no binding of any nanobody was detected in the ELISA, presumably due to the incompatibility that is described in the following chapters (Figure 6-6 D).

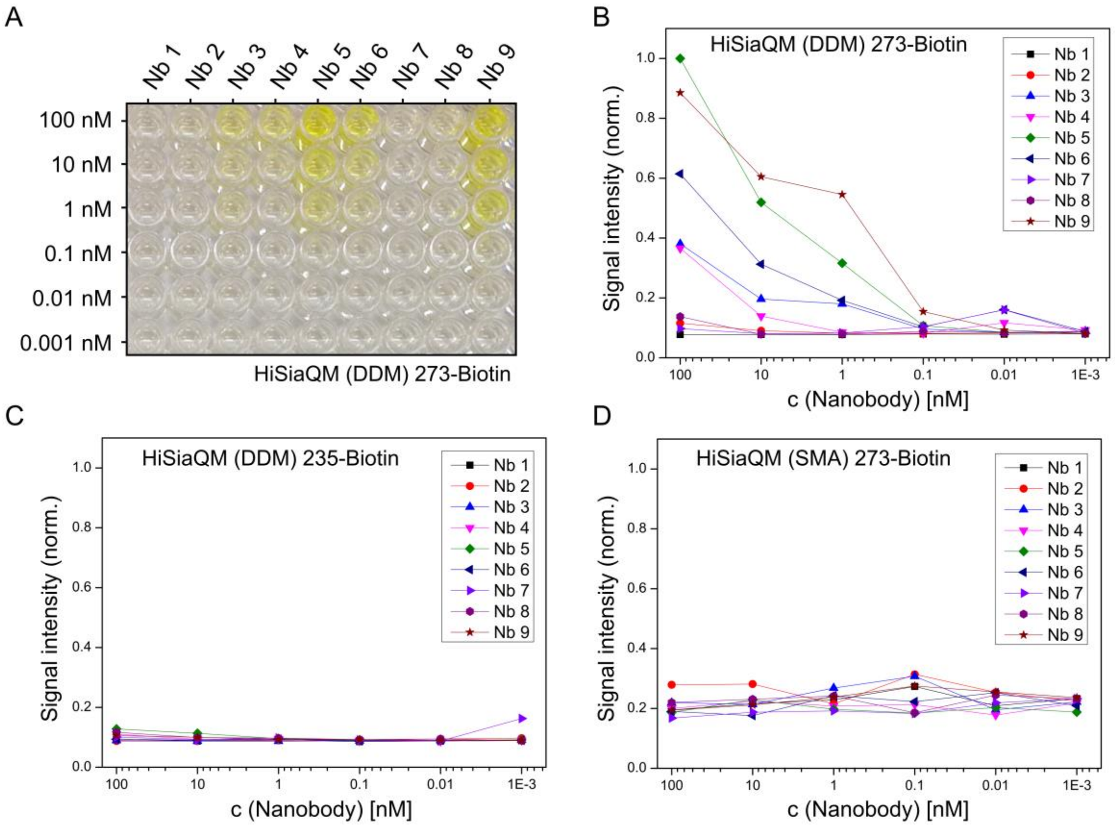


Figure 6-6: Clean-up ELISA with purified nanobodies. A) ELISA plate with 9 different nanobodies in dilution series from top to bottom with the antigen HiSiaQM K273-biotin in detergent micelles. The yellow colour results from the TMB blotting solution and indicated positive binding of nanobodies to the antigen. **B)** Signal intensity of the ELISA plate from A) at 450 nm in a dilution series with six different nanobody concentrations for each of the 9 nanobodies. **C)** Same as B) but with antigen HiSiaQM K235-biotin in DDM detergent micelles. **D)** Same as B) but with antigen HiSiaQM K273-biotin in SMA-bounded nanodiscs.

Nonetheless, the ELISAs with purified nanobodies confirmed the binding of 5 nanobodies, while the remaining 4 nanobodies showed no interaction. The results of the two different biotinylated HiSiaQM proteins in detergent micelles were consistent with the results from the first ELISA experiment. Only the antigen HiSiaQM (DDM) 273-biotin showed a binding of nanobodies, while the mutant 235-biotin from the other side of the membrane protein indicated no nanobody binding at all. This might result from an inaccessibility of the nanobody binding regions in the later mutant or another unwanted effect, such as protein aggregation for the mutant HiSiaQM (DDM) 235-biotin.

6.3.2 Analytical SEC

The binding of nanobodies to HiSiaQM was further investigated with analytical SEC. This method has the advantage that no binding partner need to be immobilized and the binding can be analysed with freely diffusing proteins. Consequently, all epitopes should be available for binding, similar to the conditions during immunization of the alpaca, and no binding should be inhibited or disrupted through blocking or destroying of an epitope. Further, this study allows the use of the wildtype HiSiaQM protein instead of the single-cysteine mutants. These mutants were modified at six different positions, the deleted 5 native cysteines and the introduced single artificial cysteine, which can disturb the structure of an epitope. On the other hand, the analytical SEC gives only a rough tendency about a binding process and the accurate prediction of a binding strength from the elution chromatogram is hard to achieve.

All analytical runs were performed on an Agilent HPLC system with an automatic sample loader to guarantee a high reproducibility of the loaded quantities. First, all proteins were analysed with the identical set-up in a single component run with same quantities as in the subsequent mixed sample runs. The resulting single component chromatograms were used as standards for comparison and detection of differences in the two component runs with mixed nanobody and HiSiaQM. In the two component experiments, a 1:1 ratio between nanobody and HiSiaQM was applied onto the SEC-column (except nanobody 5 with molar ratio of 1:0.84).

6.3.2.1 SEC with detergent-solubilized HiSiaQM and nanobodies

The chromatograms for all nanobody runs with HiSiaQM solubilized in DDM-detergent micelles are shown in Figure 6-7 (blue). The single component run with HiSiaQM showed a protein peak with two maxima at 18-19 min retention time. Even though the protein was previously purified with the standard protocol, such a double peak was never observed in the purification SEC or in any other purification of HiSiaQM before (Chapter 4.2). Due to the usage of an analytical column and another chromatography system in the present experiments, the reason for this observation was not further investigated.

The single component runs with each of the nine nanobodies resulted in a sharp, monodisperse peak at around 27 min retention time, except for nanobody 5 (Figure 6-7 E, orange). The protein signal for this nanobody indicated a bulky shape and was shifted to a higher retention time, into the region of the total column volume time of 34 min. Importantly, also the SEC during purification of this nanobody showed such a behaviour (not shown). However, the SDS-PAGE of the purification of nanobody 5 showed a clear protein band at expected position and also the clean-up ELISA showed a distinctly binding. The unexpected behaviour of nanobody 5 during SEC are therefore presumably caused by an interaction with the SEC column or an untypical structural feature.

The SEC elution profiles of the two component runs with HiSiaQM and each nanobody are shown in overlay with the single component runs (Figure 6-7 A-J, magenta). To confirm a nanobody binding, an unspecific nanobody with another antigen was used as negative, non-binding control (Figure 6-7 J, Nb X, provided by Anja Kopp, University of Bonn). For a better validation, this nanobody was used in two different molar ratios. As expected for the case of no binding, the Nb X peak at 27 min showed no or only a small decrease between the single and two component runs. Even for a low amount of Nb X, the small nanobody peak was unchanged after addition of HiSiaQM (Figure 6-7 J, dotted line). In contrast, for all other nanobodies, the detected signal of the nanobody showed a clear intensity decrease after incubation with HiSiaQM, compared with single nanobody runs. Especially nanobody 1 with a rather small peak in the single component run indicated an unambiguous effect, no nanobody peak was obtained after incubation with HiSiaQM (Figure 6-7 A). In contrast, nanobody 8 showed the qualitatively smallest decrease of the nanobody peak (Figure 6-7 H).

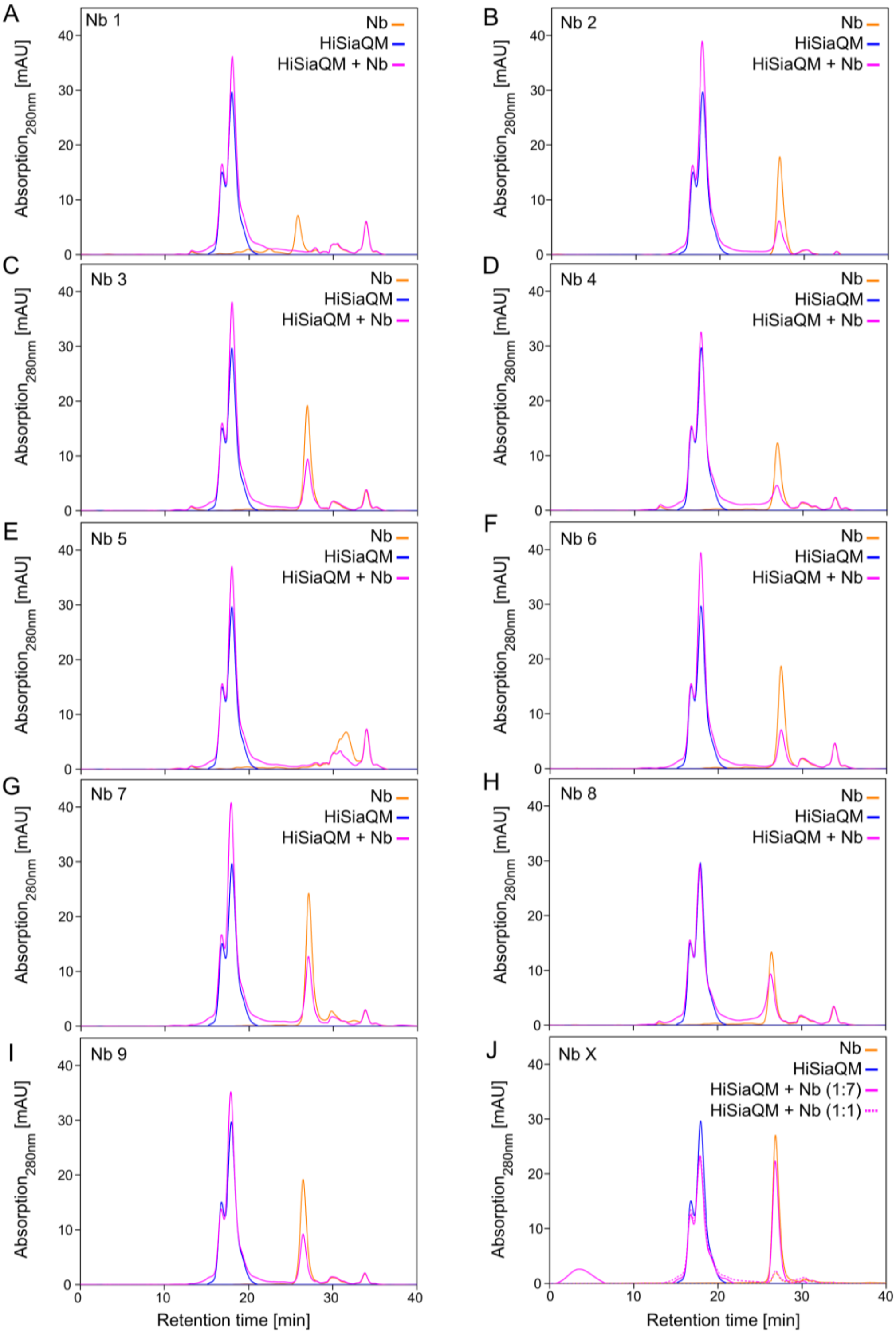


Figure 6-7: SEC-analysis of nanobody binding with detergent-solubilized HiSiaQM. A-I) SEC runs on an Agilent HPLC system (SD 200 3.2/300, buffer: 50 mM KH_2PO_4 (pH 7.8), 200 mM NaCl, 0.035% DDM). The experiments with single nanobody are shown in orange, the runs with HiSiaQM wildtype in DDM in blue and the runs of both components are shown in magenta. The mixed sample were prepared with the same amount of each protein as used in the single component runs. Both proteins were mixed in a molar ratio of 1:1 (except Nb 5, with a molar ratio of 1:0.84, due to low concentration of nanobody). **J)** Same as A-I) but with nanobody (Nb X) for another antigen than HiSiaQM. The dotted line presents the run with 1:1 mix of HiSiaQM and nanobody and the solid line a 1:7 ratio of HiSiaQM to Nb X.

As a consequence of the decreased nanobody peak intensity, the absorption peak at position for HiSiaQM showed a clear increase for the most nanobodies after mixing of HiSiaQM and nanobody, compared to the HiSiaQM single run. Only for the two runs of HiSiaQM with nanobody 4 and 8, no or only a small increase was detected. For the control nanobody X, the HiSiaQM peaks were not increased between the single and two component runs. Contrary, the HiSiaQM peak is slightly decreased for this run which was presumably caused by a small peak broadening. The observation of a decreased absorption for the nanobody peak and an increased absorption for the HiSiaQM peak in the two component runs suggests the successful binding of the nanobody to HiSiaQM and the elution of the complex from the SEC column. Indeed, a shift of the retention time for the HiSiaQM peak, caused by the HiSiaQM-nanobody complex, was not obtained, presumably due to the small size of the nanobodies in comparison to the HiSiaQM protein in micelles (15 kDa and 140 kDa, respectively). Nonetheless, the analytical SEC-runs showed evidence of binding between the wildtype HiSiaQM in detergent micelles and all nine nanobodies. However, for the nanobodies 4, 5 and 8, the binding is ambiguous and could not be confirmed with certainty from these experiments. On the other hand, the nanobodies 1, 2, 3, 6, 7 and 9, showed a clear change of the elution profile of the SEC runs and were classified as presumably good HiSiaQM binding nanobodies.

The comparison of these binding results from analytical SEC with the ELISA from Chapter 6.3.1 revealed a nice fit of successful binding for nanobody 3, 6, and 9 to HiSiaQM. The three nanobodies 1, 2 and 7 showed no binding in the ELISA but a clear tendency of a HiSiaQM-nanobody complex in the analytical SEC. In contrast, for the nanobodies 4 and 5, only the ELISAs revealed a clear binding. For nanobody 8, none of the two methods revealed an obvious binding process, which classifies this nanobody as unspecific or low affinity binder. To further characterise the nanobodies and to quantify the binding behaviour, SPR experiments are present in Chapter 6.3.3.

6.3.2.2 SEC with HiSiaQM in polymer-bounded nanodiscs and nanobodies

In analogy to the ELISA experiments, the analytical SEC runs were repeated with HiSiaQM solubilized in SMA polymer-bounded nanodiscs. By using the same set-up as before for detergent solubilized protein, the SEC runs revealed undefined elution profiles with absorption peaks at higher volumes than the column volume. A change of the set-up from the Agilent HPLC system to an FPLC Äkta pure system allowed the detection of the protein at expected elution volume of around 1 mL by using the same column (Figure SI-31). The first test runs on the Äkta pure system with the identical column as before revealed the disadvantage that the samples were loaded manually, and that the system is not customized for such small volumes as the Agilent HPLC. The manual load of the sample had a lower reproducibility than the automatic sample application on the HPLC machine and the intensity of absorption from all samples were significantly lower than expected from the experiments before (Figure SI-31). Further, two HiSiaQM single component test-runs with the same amount of protein resulted in broad elution profiles with up to 3 peaks and with a strongly varied absorbance intensity (Figure SI-31).

Overall, the analytical SEC studies with SMA polymer-bounded nanodiscs revealed some unexpected incompatibilities that were not described until today. Due to these uncertainties, analytical SEC runs of HiSiaQM in polymer-bounded nanodiscs could not be used for the investigation of nanobody binding. Possible reasons for these incompatibilities and unexpected behaviour of polymer-bounded nanodiscs with the HPLC system and analytical SEC are discussed in Chapter 6.5.

6.3.3 Single cycle kinetics experiments with SPR

Surface-plasmon resonance (SPR) was used as third method to investigate and characterise the nanobody binding to the TRAP transporter membrane domains HiSiaQM. In contrast to the other two methods, ELISA and analytical SEC that were used before, SPR allows not only the qualitative confirmation of binding, but also reveals information about the strength and kinetics of the binding process. For an SPR experiment, one binding partner is immobilized on a surface and the interaction with another protein or compound, which is rinsed over the immobilized protein, is detected. Analogous to the panning and ELISA experiments, the two biotinylated mutants of HiSiaQM K235-biotin and K273-biotin in detergent micelles were used for immobilization and to investigate side-specificity of the nanobodies (Figure 6-2). The proteins were immobilized on a streptavidin-coated SPR-chip (Chapter 9.2.6). In addition to the detergent stabilized proteins, HiSiaQM K273-biotin in SMA polymer-bounded nanodiscs was also investigated. However, it was not possible to immobilize this protein on the surface and no useful SPR-signal was obtained. Therefore, similar to the analytical SEC experiments, binding studies for the nanobodies with SPR was not feasible for HiSiaQM in polymer-bounded nanodiscs.

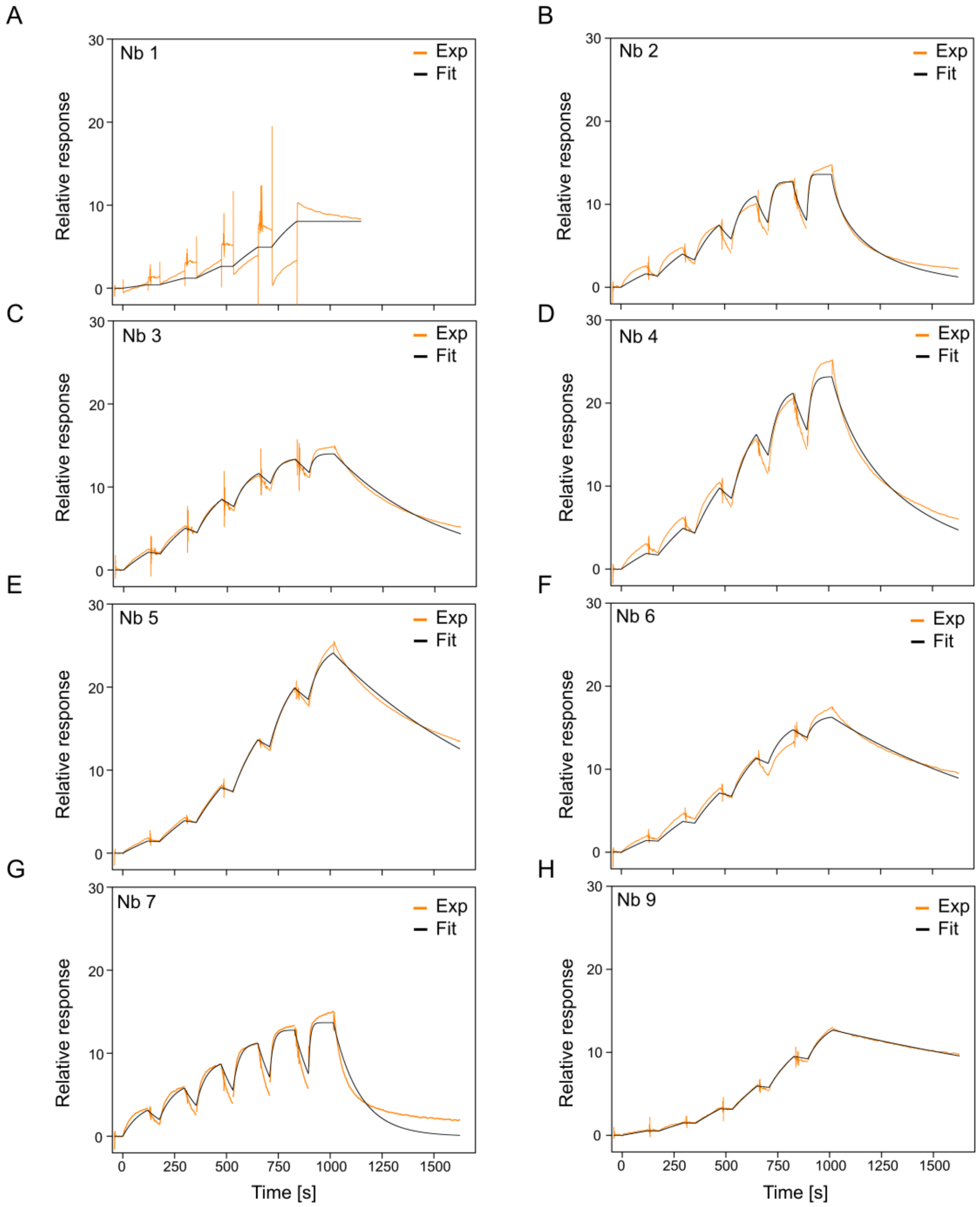


Figure 6-8: SPR experiments of HiSiaQM K273-biotin and nanobodies. A-H) Sensorgrams of the SPR single cycle kinetics experiments for HiSiaQM nanobody with HiSiaQM K273-biotin in DDM detergent micelles as immobilized binding target. The relative response of the surface is shown as orange line, the fit to the experimental data as black line.

The SPR sensorgrams of immobilized HiSiaQM K273-biotin in detergent micelles are shown in Figure 6-8. Due to the chip design with eight flow cells and the observation of no or even low binding for nanobody 8 in the previous studies, this nanobody was not analysed in the SPR experiments. The other nanobodies were injected in single cycle kinetics experiments on immobilized HiSiaQM in different flow cells. For this, the concentration of the nanobodies was increased for each injection with only short, incomplete dissociation steps. All nanobodies, except nanobody 1, showed a clear binding to the immobilized HiSiaQM (Figure 6-8, orange). The single cycle sensorgrams of the binding nanobodies revealed a clear association signal, partially a short steady state and a dissociation after injection with larger observed time scale for the last injection. Further, all sensorgrams for the seven nanobodies revealed defined characteristics of the association and dissociation during and after injection. For example, nanobodies 2, 4 and 7 showed a fast increase of the response after injection and a fast decrease after injection. This suggested a faster binding kinetics than for example for nanobodies 5, 6 or 9. Especially for nanobody 9, the dissociation after the last injection revealed a slow dissociation of the nanobody from HiSiaQM. The unsuccessful binding of nanobody 1 was in good agreement with the accomplished ELISAs. Interestingly, this nanobody 1 derived from one of the two nanobody groups (1 and 5), for which no positive binding was observed with HiSiaQM K273-biotin in DDM in the first ELISA.

To describe and analyse the SPR data, the single cycle kinetics sensorgrams were fitted automatically with the SPR software (Figure 6-8). The dissociation constant K_D and the kinetic parameters of association (k_{on}) and dissociation (k_{off}) are summarized in Table 6-1. For all seven nanobodies with an observed binding, the dissociation constants were determined to high affinities with low or sub nM concentrations. Nanobody 3, 6 and 9 were identified as the strongest binders with high affinities of 640, 640 and 360 pM, respectively. Interestingly, these were also the only nanobodies for which a binding was observed in all binding studies, in the ELISAs, analytical SEC and SPR. Nevertheless, also nanobody 4 with the weakest binding affinity of around 10 nM was classified as binder with a high affinity and is for example still higher than the affinity between TRAP transporter SBPs and their substrate sialic acid of around 50-200 nM (Figure 3-4). The kinetic parameters for the eight nanobodies revealed large differences for the eight nanobodies, suggesting different binding behaviours. For example, nanobodies 2, 4 and 7 revealed a fast dissociation from the immobilized HiSiaQM membrane domains, also clearly visible from the decrease of the response after final injection. In contrast, the other nanobodies, especially nanobodies 6 and 9, showed a clearly slower dissociation after binding to HiSiaQM (Table 6-1).

Table 6-1: Dissociation constants and kinetic parameters of SPR experiments with eight nanobodies and immobilized HiSiaQM 273-biotin.

	K_D [nM]	k_{on} [$M^{-1}s^{-1}$]	k_{off} [s^{-1}]
Nanobody 2	9.83	$1.94 \cdot 10^6$	$19.0 \cdot 10^{-3}$
Nanobody 3	0.64	$3.01 \cdot 10^6$	$1.91 \cdot 10^{-3}$
Nanobody 4	10.40	$0.47 \cdot 10^6$	$4.92 \cdot 10^{-3}$
Nanobody 5	1.06	$1.01 \cdot 10^6$	$1.07 \cdot 10^{-3}$
Nanobody 6	0.64	$1.54 \cdot 10^6$	$0.98 \cdot 10^{-3}$
Nanobody 7	4.88	$1.62 \cdot 10^6$	$7.92 \cdot 10^{-3}$
Nanobody 9	0.36	$1.27 \cdot 10^6$	$0.46 \cdot 10^{-3}$

The SPR experiments were repeated with the second mutant, HiSiaQM K235-biotin in detergent micelles, to also present the other side of the membrane domains for binding, as postulated by the two protein models in Figure 6-2. The SPR sensorgrams of the single cycle kinetics experiments for the same eight nanobodies as before revealed a strongly changed behaviour compared to the first immobilized HiSiaQM mutant. Only for nanobodies 4 and 5, a binding to immobilized HiSiaQM K235-biotin was observed (Figure 6-9). Additionally, just nanobody 4 revealed a sensorgram with good enough quality for fitting and analysing of binding parameters. Nanobody 5 indicated an observable binding process for the first injections but an undefined response signal for the later injections. This overall reduced binding of nanobodies to HiSiaQM K235-biotin compared to K273-biotin was in agreement with both ELISAs, in which only positive hits were detected for K273-biotin. The dissociation constant for nanobody 4 to HiSiaQM K235-biotin was determined to be around 27 nM and was thereby still high, but also 1.5 times weaker than for the other HiSiaQM mutant (Table 6-2). The kinetics of the binding process of nanobody 4 were in the same range for the two different immobilized HiSiaQM mutants. The overall comparably binding characteristic for this nanobody, by using two different immobilized HiSiaQM proteins, suggests the same binding mechanism to both HiSiaQM mutants.

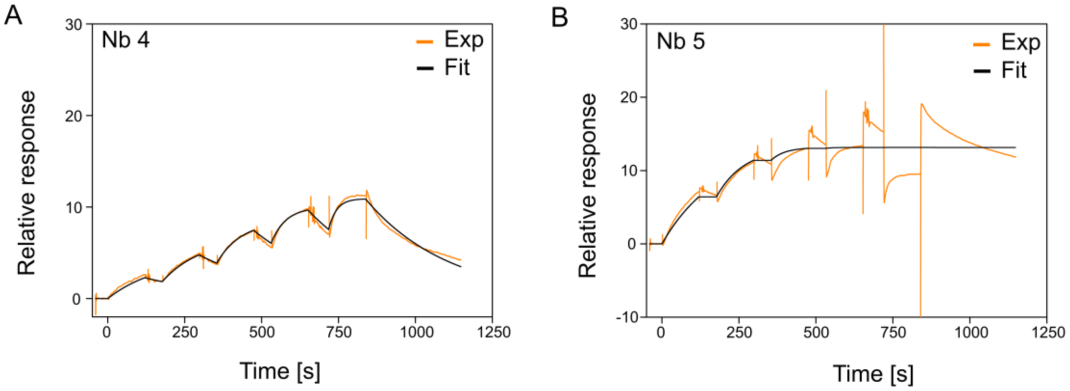


Figure 6-9: SPR experiments of HiSiaQM 235-biotin and nanobodies. A-B) Sensorgrams of the SPR single cycle kinetics experiments for HiSiaQM nanobodies with HiSiaQM K235-biotin in DDM detergent micelles as immobilized binding target. The relative response of the surface is shown as orange line, the fit to the experimental data as black line.

Table 6-2: Dissociation constants and kinetic parameters of SPR experiment with one nanobody and immobilized HiSiaQM 235-biotin.

	K_D [nM]	k_{on} [$M^{-1}s^{-1}$]	k_{off} [s^{-1}]
Nanobody 4	26.8	$0.14 \cdot 10^{6}$	$3.68 \cdot 10^{-3}$

6.3.4 Binding regions of nanobodies on HiSiaQM

After verification and classification of the binding between HiSiaQM and several nanobodies in the SPR experiments, the binding nanobodies were used to determine relative binding regions on HiSiaQM. For these measurements, only the immobilized HiSiaQM K273-biotin mutant was used, as for this mutant several binding nanbodies were observed. For the determination of the binding regions, the nanobodies were applied in a defined way on HiSiaQM to investigate competitive binding processes. First, one nanobody was applied onto the immobilized HiSiaQM until a steady state of binding was observed. Afterwards, a second nanobody was injected into the same flow cell under constant injection of the first nanobody. The signal of the sensorgram during application of the second nanobody indicates, if the first applied nanobody blocks the binding region of the second nanobody or if both nanobodies can simultaneously bind to HiSiaQM. By using each of the seven nanobodies as first and as second nanobody and measure all possible combinations of the nanobodies, an overview about relative binding regions on HiSiaQM was created.

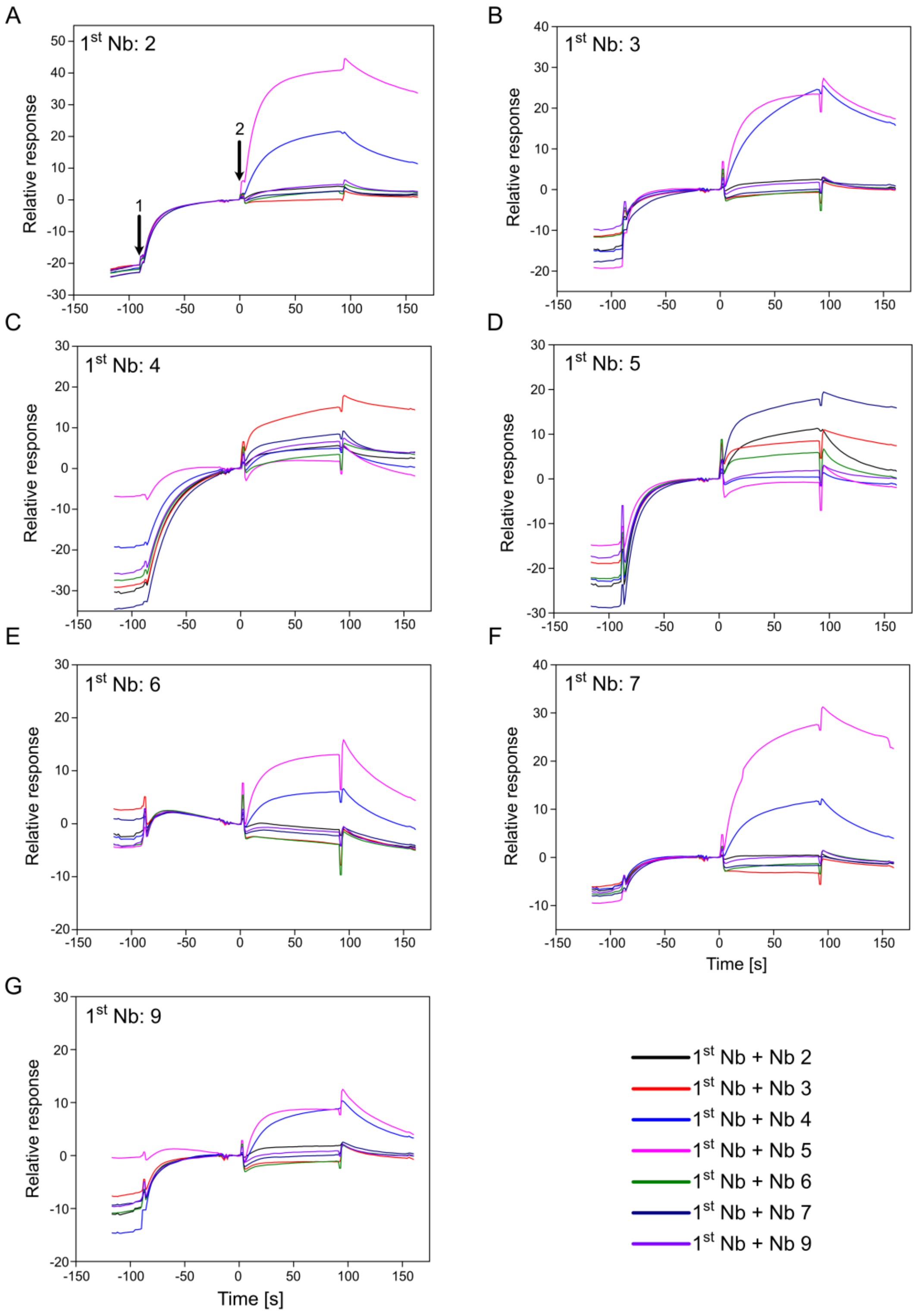
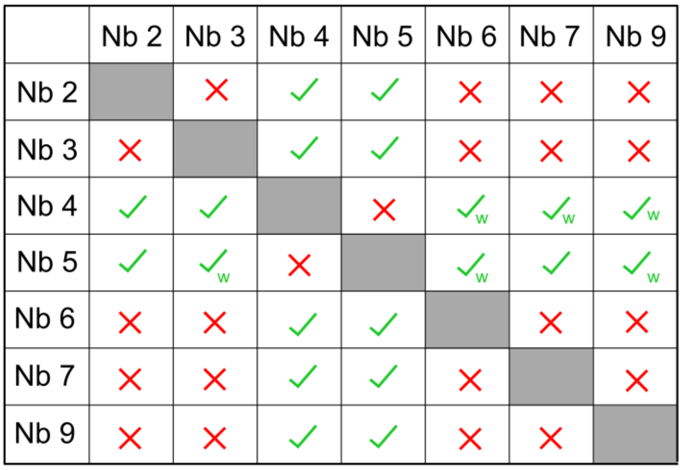


Figure 6-10: SPR-measurements of competitive binding experiments. A) SPR sensorgrams with nanobody 2 in the first injection. The different colours of the lines correspond to the nanobody from the second injection, as shown in the bottom right corner of the figure. The time points at which injection of first and second nanobody started are highlighted with black arrows. **B-G)** Same as A) but with different saturated nanobodies in the first injection, as mentioned in the figure.

The competetive binding experiments for all nanobody combinations are illustrated in Figure 6-10. The sensorgrams for nanobody 2 as first nanobody are good examples, in which different effects after application of the second nanobodies were observed (Figure 6-10 A). After binding of nanobody 2 in the first application, nanobodies 4 and 5 showed an association to the immobilized HiSiaQM protein in the second injection. In contrast, the signal after injection of the other four nanobodies remained unchanged. These observations suggest that nanobody 2 blocked the binding sites for the nanobodies 3, 6, 7 and 9 on HiSiaQM. From these results it could be already predicted that nanobodies 2, 3, 6, 7 and 9 bind to a similar region and that binding of nanobody 4 and 5 are independent from nanobody 2. In accordance to that, for example the measurements with nanobody 5 at first appplied binder indicated a binding for nanobodies 2, 3, 6 and 7 and no interaction between HiSiaQM and nanobody 4 in the second injection (Figure 6-10 D).

The data of all competitive binding experiments from Figure 6-10 are summarized in Figure 6-11. It allows a fast determination of the relative binding regions that are targeted by the nanobodies and which nanobodies can be combined for simultaneously binding. Overall, the seven nanobodies that were identified as binders for HiSiaQM were assigned to two different binding regions. The first region is targeted by the nanobodies 2, 3, 6, 7 and 9 and the second region by nanobodies 4 and 5. This result fits to the sequence identity of all nanobodies in Figure 6-4, in which nanobodies 4 and 5 showed a high sequence similarity and the remaining nanobodies are more different. Thereby, for these two nanobodies, the binding to one specific epitope can be assumed. The other five nanobodies have larger sequence differences and it cannot be excluded if these nanobodies bind to one or more epitopes in the same protein region and by this, blocking the binding position for the other nanobody.

The results from this binding map fits also nicely to the previous side-specific single cycle kinetics SPR experiments. Only for nanobodies 4 and 5, which presumably target the same epitope on HiSiaQM, a binding was observed for the HiSiaQM K235-biotin mutant (Figure 6-9). Therefore, the functionalization or immobilization of HiSiaQM at position K235 presumably destroyed or disrupted the accessibility of the binding region for nanobodies 2, 3, 6, 7 and 9. A possible binding model for the HiSiaQM specific nanobodies which summarizes and describes the observations is given in the discussion (Chapter 6.5).



	Nb 2	Nb 3	Nb 4	Nb 5	Nb 6	Nb 7	Nb 9
Nb 2	▨	✗	✓	✓	✗	✗	✗
Nb 3	✗	▨	✓	✓	✗	✗	✗
Nb 4	✓	✓	▨	✗	✓$_{w}$	✓$_{w}$	✓$_{w}$
Nb 5	✓	✓$_{w}$	✗	▨	✓$_{w}$	✓	✓$_{w}$
Nb 6	✗	✗	✓	✓	▨	✗	✗
Nb 7	✗	✗	✓	✓	✗	▨	✗
Nb 9	✗	✗	✓	✓	✗	✗	▨

Figure 6-11: Binding map of HiSiaQM nanobodies. Summarized, observed binding in the competitive experiments from Figure 6-10 are marked as green hook and no detectable binding as red cross. For cases of weak binding, the combinations are marked with a green "w".

6.4 Functional studies of TRAP transporters with nanobodies

6.4.1 Sialic acid dependent *in vivo* growth assay

The nanobodies that were produced and characterised as binders for the TRAP transporter membrane domains HiSiaQM were further used to study an influence through binding on the function of the transporter. For this purpose, an *in vivo* cell growth assay was used, which were designed and provided from the group of Gavin H. Thomas (University of York) (Chapter 9.2.1). This assay was already used in Chapter 5.2.3 to analyse the function of HiSiaQM cysteine mutants and in previous publications (Fischer et al., 2015, Severi et al., 2010). This assay is based on a modified *E. coli* strain with knockout of the native sialic acid transporter nanT (*E. coli ΔnanT,* also named SEVY3). Consequently, if only sialic acid is provided to the cells as carbon source for the generation of energy, the *E. coli ΔnanT* strain is not able to grow. If an external, artificial sialic acid transporter, such as the TRAP transporter HiSiaPQM, is provided to the cells, they can incorporate sialic acid, metabolise it and show a normal growth (Severi et al., 2010).

In the present study, the cells were supplemented with the TRAP transporter HiSiaPQM and with one of the transporter-specific nanobody (Figure 6-12). For this purpose, the nanobody genes were cloned into the alternative expression vector pET28a with kanamycin resistance that allowed a co-transformation with the HiSiaPQM gene containing vector with ampicillin resistance into the SEVY3 cells. To investigate a possible influence of the nanobodies with binding to a specific side of the membrane protein, each nanobody gene was cloned into the pET28a vector without and with pelB signal sequence for export in the periplasm after expression. Thus, in separated experiments, the nanobodies were either accumulated in the cytoplasm or periplasm and could interact selectively with both sides of HiSiaQM (Figure 6-12).

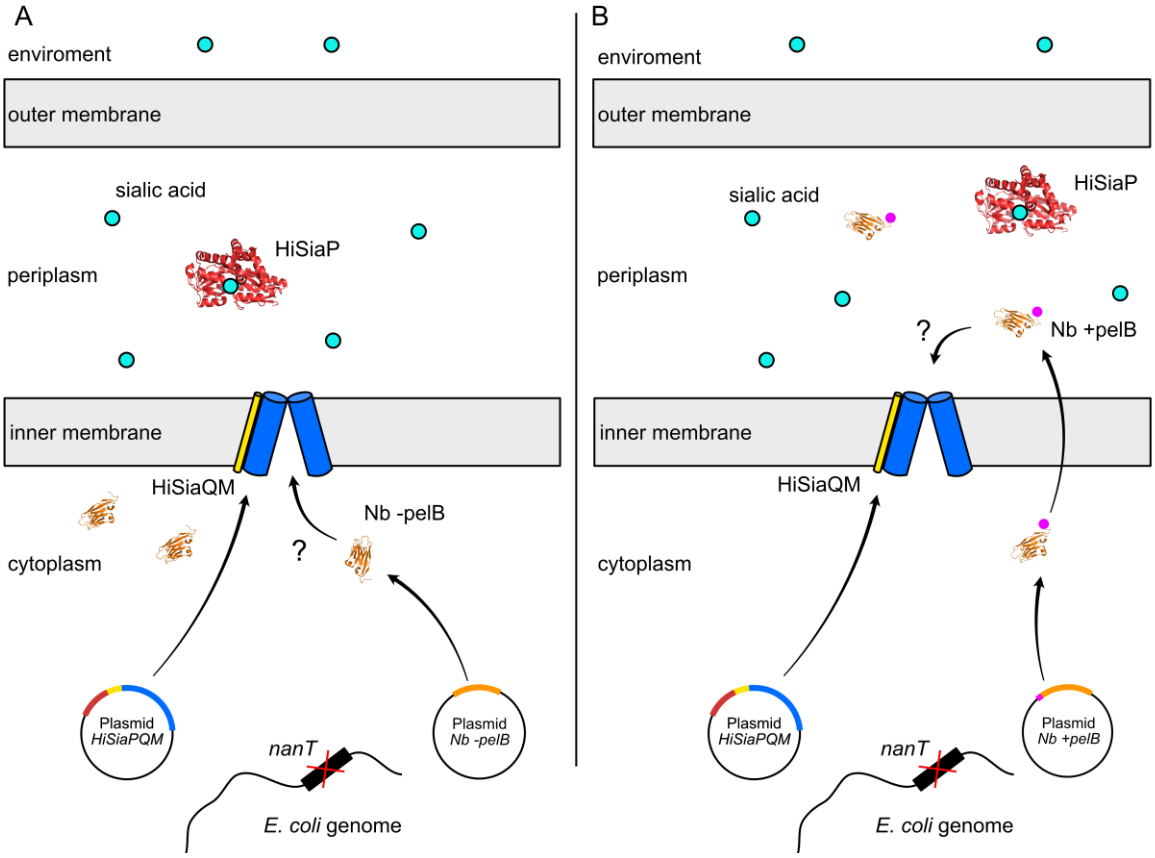


Figure 6-12: Schematic of TRAP transporter growth assay with nanobodies. A) A *ΔnanT E. coli* strain was used and supplemented with a HiSiaPQM transporter plasmid for sialic acid dependent growth of the cells. To investigate the influence of nanobodies, a nanobody plasmid was co-transformed into the cells without pelB signal sequence (-). Consequently, the nanobodies are accumulated in the cytoplasm (HiSiaP is illustrated in red, HiSiaQM in blue and yellow, sialic acid in cyan, the membranes in grey. Nanobody PDB-ID: 5U64 (Koch et al., 2017), HiSiaP PDB-IDs: 2CEY (Müller et al., 2006) and 3B50 (Johnston et al., 2008)). **B)** Same as A) but with nanobody fused to a pelB signal sequence (magenta) for export and accumulation in the periplasm.

The sialic acid depended growth assay of SEVY3 cells can be performed on agar plates or in liquid medium cultures. Due to a possible constant detection and quantification of the cell growth at several time periods in liquid medium cultures, these experiments provided a highly sensitive read-out for a changed growth behaviour. The cell culture growth was analysed over several hours with measurements of the optical density at 600 nm and optimal time points for monitoring of high growth rates were initially tested in several experiments. As a positive control, a culture with HiSiaPQM and an empty pET28a vector without nanobody sequence was prepared. As before for the analytical SEC (Chapter 6.3.2.1), the unspecific nanobody X was also cloned into the vector with and without the pelB sequence and used as control sample. A negative control culture without HiSiaPQM and without nanobody was additionally used to confirm the absence of any other nutrient than sialic acid and the success of the washing step for removing of LB medium from the precultures. The growth experiments were repeated for each nanobody and with comparable results but due to changed starting points of growth, the data were not combined in one data set.

The results of the detected cell growth for SEVY3 cells with HiSiaPQM and different nanobodies over several hours are illustrated in Figure 6-13 (repeating experiments in Figure SI-34). The positive control (green) showed a constant cell growth between 12-20 hours after preparation of the cultures, while the optical density of the negative control (red) did not increase during the entire experiment. Importantly, the control Nb X demonstrated for both constructs, with and without pelB sequence, a comparable cell growth to the positive sample. This observation confirmed no general influence of the nanobody expression or unspecific nanobody interactions on the growth rates of the cells. For all nine nanobodies that were not fused to the pelB sequence and are therefore located in the cytoplasm, the detected growth curves were comparable to the positive sample and the control nanobody X (Figure 6-13 A). Interestingly, for most cell cultures with nanobodies that were exported into the periplasm, a significant change in the growth curves compared to the control samples was observed (Figure 6-13 B, Figure SI-34). Except nanobody 1 and 8, all other seven nanobodies showed a clearly reduced growth of the cell cultures between 12-20 hours after preparation. The strongest effect was obtained for nanobody 7, followed by nanobody 6. Even after 24 hours, the optical density of cell cultures with these nanobodies was significantly lower than for all other nanobodies.

However, for all cell cultures with reduced growth rates, an increase of the optical density was visible after several hours. The detection of cell growth for a longer time-period is shown in Figure SI-35 and support this observation. Thereby, none of the nanobody supplemented cultures indicated no cell growth, as for the negative control sample. This behaviour was also observed in a comparable agar plate-based growth assay, which gives just a yes or no answer about the cell growth and indicated cell growth upon presence of any nanobody (Figure SI-32 and Figure SI-33).

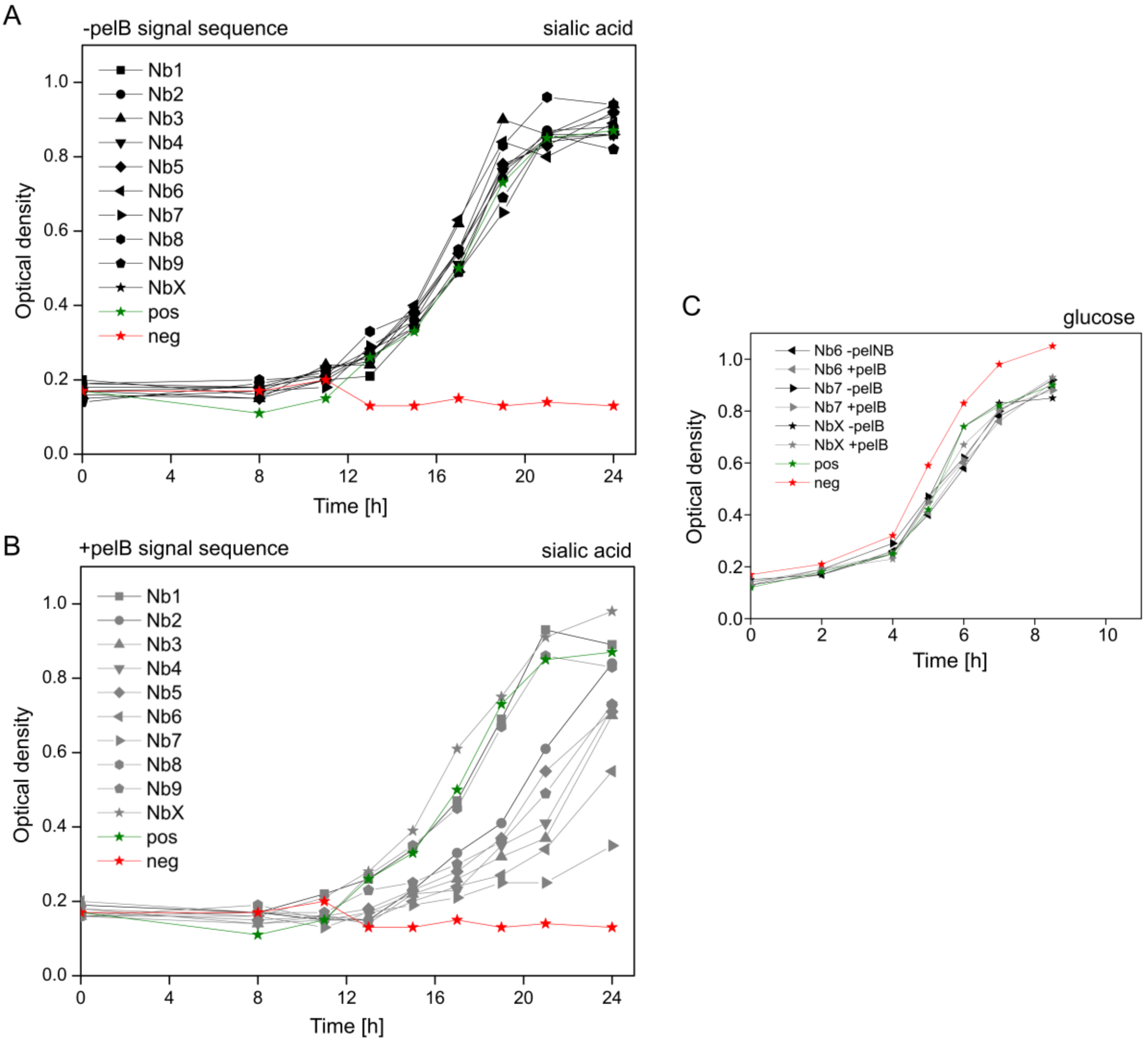


Figure 6-13: TRAP transporter growth assay with addition of nanobodies. A) Growth curves of *ΔnanT E. coli* cells in M9-minimal medium supplemented with sialic acid. Cells with the sialic acid TRAP transporter HiSiaPQM and nanobodies without pelB signal sequence (-pelB) for cytoplasm accumulation after expression (black) (Nb X is a TRAP transporter unspecific nanobody). For positive control (green), a cell culture without nanobody was used and for negative control (red), no sialic acid transporter or nanobody was transformed into the cells (just empty ampicillin vector for resistance). **B)** Same as A) but the nanobodies were fused to the pelB sequence (+pelB) for transport into the periplasm (grey). **C)** Growth rates of selected cell cultures from A) and B) with strongest inhibition, supplemented with glucose to validate sialic acid-independent growth.

Figure 6-13 B showed that the two cell cultures with pelB-fused nanobody 1 and 8 had no alternated or inhibited effects on the growth rates and were comparable to the positive control. These results fit to the binding characterisation studies from the previous chapters. The two nanobodies showed overall the lowest binding affinity to the TRAP transporter HiSiaQM, for example nanobody 1 in the SPR measurements (Chapter 6.3.3) or nanobody 8 in the analytical SEC experiments (Chapter 6.3.2) and both nanobodies in the ELISAs (Chapter 6.3.1). Therefore, the observation of no influence of these two nanobodies to the function of the TRAP transporter, compared to the other nanobodies, are presumably caused by no or only weak binding.

To ensure that the inhibiting effect was truly related on an interaction between the nanobodies and the TRAP transporter, a control experiment was performed with the two strongest inhibiting nanobodies. Therefore, the cell growth experiments were repeated with replacement of the sialic acid by glucose. Under this condition, each of the *E. coli* cultures should be able to grow, independent from the function of the TRAP transporter. The cell growths for these measurements are illustrated in Figure 6-13 C and indicated no deviation of the nanobody supplemented cultures from the positive control. Due to the glucose, also the negative control cell culture showed a cell growth and indicated simultaneously the strongest growth, presumably since no kanamycin was added and no expression of TRAP transporter or nanobodies. Overall, these results excluded any toxic or general, unspecific effects of the strongest inhibiting nanobodies onto the cell cultures. It suggests the idea of a specific inhibition of the TRAP transporters function which results in a significantly lower sialic acid uptake into the cells.

The experiments from the last paragraphs demonstrated a significant influence of several nanobodies to the function of the TRAP transporter HiSiaPQM. The reduced cell growth in the presence of seven nanobodies, that were previously characterised as strong binders for HiSiaQM, suggests a direct influence of the nanobody binding on the function of the transporter with a decreasing effect on the activity. Further, since the reduced cell growth was exclusively observed after export of nanobodies into the periplasm of the cells, the binding epitopes are presumably located on this periplasmic side of HiSiaQM. A detailed discussion of the binding and a hypothetical binding model for the nanobody-based inhibition of TRAP transporters are presented in the discussion (Chapter 6.5).

6.4.2 Analysis of tripartite complex formation upon nanobody binding

In previous studies, nanobodies were described that can stabilize distinct conformations of membrane proteins, for example an active state of β_2 adrenoceptor or the outward-open conformation of secondary transporter LacY (Rasmussen et al., 2011a, Smirnova et al., 2014). For TRAP transporters, the previous PELDOR experiments suggested a high structural flexibility of the membrane domains in solution (Chapter 5.5.1), which was postulated as one possible reason for the undescribed tripartite complex between the SBP and the membrane domains. In the present chapter, analytical SEC was used to investigate, if the nanobody binding promotes a conformation of HiSiaQM that is able to form a P-QM complex, which was postulated as essential for the function of the transporter (Forward et al., 1997, Mulligan et al., 2011).

The analytical SEC runs were performed in a similar way as before for the nanobody binding experiments, with investigation of changes in the elution profiles between single and mixed component runs. For a specific detection, the HiSiaP single cysteine mutant K254C was previously labelled with the fluorophore fluorescein-maleimide, after standard expression and purification (SEC and SDS-gel are shown Figure SI-36). The single component runs of both transporter components, HiSiaQM and HiSiaP, are shown in Figure 6-14 A and B. For HiSiaP, the absorption of fluorescein at 495 nm was detected additionally to the protein absorption at 280 nm. This facilitated the detection of a changed elution behaviour, especially a shift of the specific absorption into the HiSiaQM protein peak region. Subsequently to the single runs, HiSiaQM was combined with a nanobody and the labelled HiSiaP, and the mixture was loaded onto the analytical SEC. For the experiments in Figure 6-14, a sialic acid supplemented buffer was used (1 mM, ratios of components are given in description of Figure 6-14). Overall, the absorption at 495 nm for all SEC runs with HiSiaQM and different nanobodies revealed no differences for the eluted volume of HiSiaP (Figure 6-14 C). The repetition of the experiments without sialic acid in the buffer showed the same result (Figure SI-37). Consequently, the binding of the nanobodies stabilized no conformation of HiSiaQM that promotes the formation of the tripartite complex with HiSiaP.

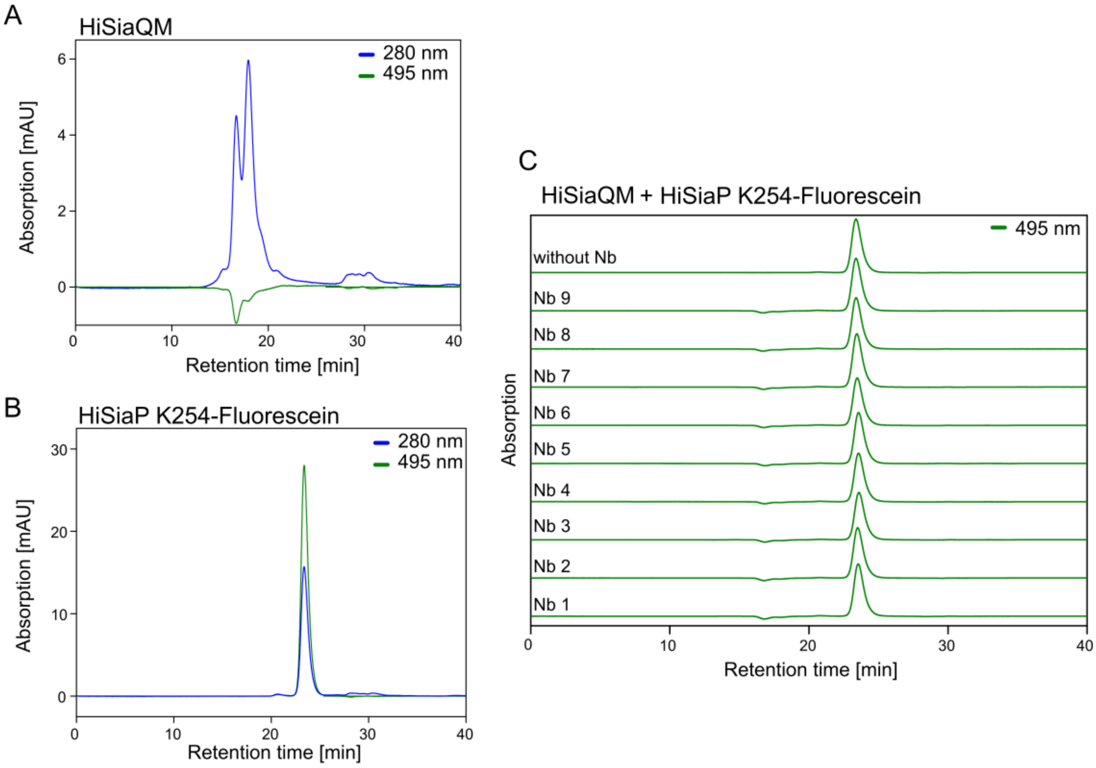


Figure 6-14: Interaction of P- and QM-domains upon addition of nanobodies. A) SEC run of HiSiaQM with detection of elution absorption at 280 nm and 495 nm (SD 200 3.2/300, Agilent HPLC system). The standard protein DDM buffer was supplemented with 1 mM sialic acid (experiments without sialic acid are shown in Figure SI-37). **B)** Same as A) but with HiSiaP K254-fluorescein. **C)** Same set-up as A) but just the absorption at 495 nm is shown. The chromatograms correspond to a mixed sample of HiSiaQM and HiSiaP K254-fluorescein (top) and addition of each nanobody (1-9). The ratio between the components were chosen to be 1:1:2 for HiSiaQM to nanobody to HiSiaP.

6.5 Discussion and conclusion

6.5.1 Confirmation of HiSiaQM model from nanobody binding

The results from the side-selective nanobody binding characterisation and from the *in vivo* assay can be combined and used for experimental validation of the current protein models of HiSiaQM (Figure 6-2). The *in vivo* experiments indicated only an observable effect if the nanobodies are present on the periplasmic side of the transporter protein. This observation strongly suggests that the binding regions of all nanobodies are located on this side of the protein. Additionally, the binding studies with SPR experiments and ELISAs indicated a much better binding for all nanobodies, visible as higher number of binding nanobodies and higher affinities, during the immobilization of HiSiaQM K273-biotin compared to HiSiaQM K235-biotin.

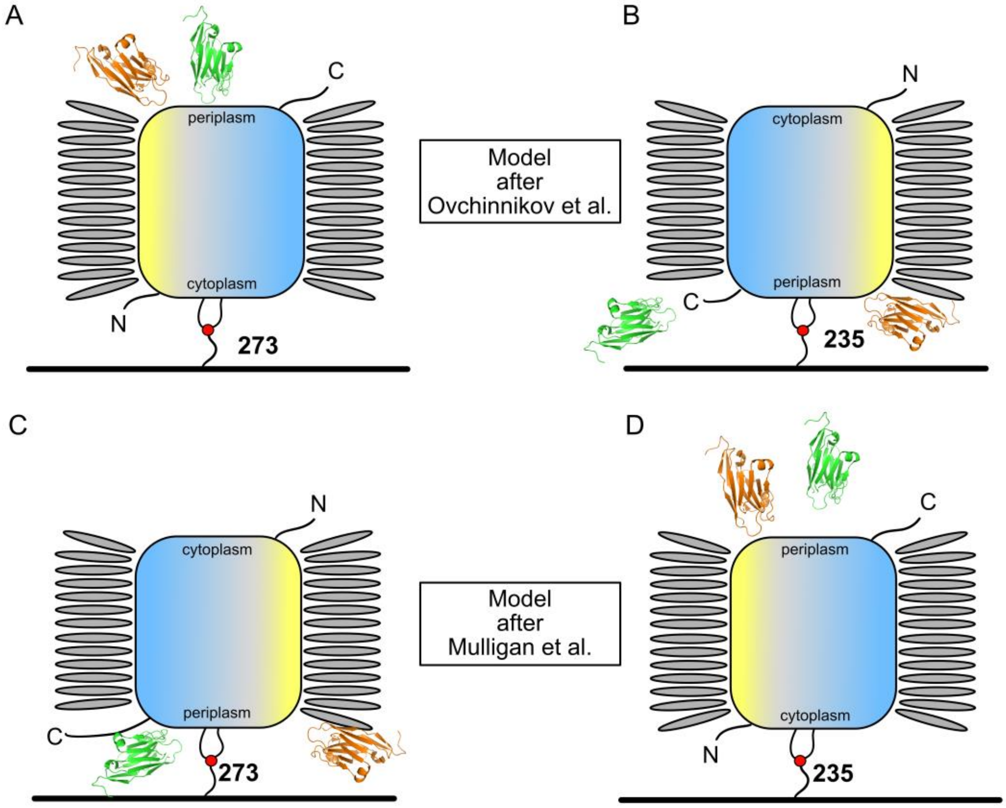


Figure 6-15: Immobilization variants and conformation of HiSiaQM model through nanobody binding. A) HiSiaQM model from Ovchinnikov et al. (Ovchinnikov et al., 2014), immobilized on a surface via position K273. The periplasmic side is shown in red and nanobody binding is assumed on this side (nanobodies for two different binding regions in orange and green, PDB-ID: 5U64 (Koch et al., 2017)). The detergent micelle is indicated in grey. **B)** Same model as A) but immobilized with HiSiaQM K235-biotin. **C-D)** Same as A-B) but with protein model for HiSiaQM from Mulligan et al. (Mulligan et al., 2011), in which the biotin-positions are postulated the other way around.

The two protein models with both immobilization possibilities are schematically shown in Figure 6-15. Importantly for the model validation, all nanobody binding regions are presumed to be located on the periplasmic side of the protein, the side of the HiSiaQM C-terminus. In this case, the observed better binding for immobilized HiSiaQM K273-biotin supports the protein model of Ovchinnikov et al. (Ovchinnikov et al., 2014). In this model, the K273 is located on the cytoplasmic side and an immobilization of HiSiaQM with this position will provide a high accessibility of the periplasmic side for nanobody binding (Figure 6-15 A). In contrast, for the protein model from Mulligan et al. (Mulligan et al., 2011), the immobilization of HiSiaQM K235-biotin will provide a high accessibility of the periplasmic side. However, for this immobilization strategy, only for two nanobodies a weaker binding was observed in the SPR experiments. Consequently, this lower accessibility of binding regions for mutant HiSiaQM K235-biotin are in agreement with the model from Ovchinnikov et al. (Ovchinnikov et al., 2014), in which the periplasmic side is nearly located at the immobilization surface (Figure 6-15 B). As seen from the determination of binding regions with SPR experiments, only the binding region for nanobody 4 and 5 is accessible by immobilization of this mutant, while the other one for nanobody 2, 3, 6, 7 and 9 is inaccessible or disrupted through the introduced biotin-label. Importantly, these conclusions, as well as the protein model, are based on several assumptions and must be interpreted with care. Nevertheless, no further information about the HiSiaQM structure is known until today and the combination of several results from this chapter represents the first experimental confirmation of one HiSiaQM model structure.

6.5.2 Failure of nanobody binding studies with HiSiaQM in SMA-bounded nanodiscs

Due to a previous described important impact of the hydrophobic environment onto a membrane proteins function and structure, the characterization of the TRAP transporter nanobodies were additionally performed with HiSiaQM in polymer-bounded nanodiscs (Chapter 4.4) (Gupta et al., 2017, Laganowsky et al., 2014, Martens et al., 2016, Martens et al., 2018, Sanders et al., 2018). Overall, even if some nanobodies were identified as binders to HiSiaQM in SMA-nanodiscs in the first ELISA, the subsequent binding studies with different methods were mostly unsuccessful and uninterpretable.

The reason for the incompatibilities of HiSiaQM in SMA-nanodiscs remained unclear, it might be due to the physically properties of the polymer belt around the nanodiscs or the native membrane components inside the nanodisc. The current knowledge about this relatively new system, particularly the handling and behaviour during different bioanalytical experiments, is still sparse, as already mentioned in Chapter 4. For example, SPR experiments of SMA-nanodiscs with native membranes are not described until today. Experiments in this work indicated an unsuccessful injection and immobilization of HiSiaQM in polymer-nanodiscs on the SPR-chip, by using comparable parameters and conditions as for the detergent samples. A similar effect was observed for the analytical SEC on the Agilent HPLC machine. Possible reasons for this incompatibility are stainless steel HPLC components, which are removed in a bio-inert version of the HPLC machine. The change to a bioinert FPLC system partially solved this problem but the eluted absorption intensities still fluctuated strongly. Additionally, the washing procedure of the SEC column after several HiSiaQM-nanodisc runs indicated a high amount of eluted protein that presumably results from HiSiaQM-nanodiscs which stick on the SEC column. Overall, the experimental observations in bioanalytical studies with HiSiaQM in SMA-nanodiscs revealed great differences to the detergent samples that hinder the successful characterization of nanobody binding to HiSiaQM in this second membrane mimetic system.

Additional to these methodical problems, it was observed that mixing of HiSiaQM in SMA-nanodiscs and one of the nanobody in high concentrations led to an instantaneous precipitation of the sample. This was not only observable for all nine TRAP transporter-specific nanobodies after addition to HiSiaQM, but also for the addition of these nanobodies to other SMA-nanodiscs reconstituted membrane proteins. Similar to the technically problems, this behaviour was never described before. The aggregation with various combinations of nanobodies and SMA-nanodisc samples indicated a general intolerance of both components. It could be possible, that either the nanobodies disrupt the nanodisc-assembly and the membrane protein aggregates after decomposition of the nanodisc or that the surface of the SMA-nanodiscs, in particularly the polymer itself, interacts with the nanobodies and encourages their aggregation. As alternative to the SMA polymer, the studies can be repeated in further experiments with DIBMA polymer nanodiscs with changed physical properties (Chapter 4.1) (Oluwole et al., 2017a). However, the initial test-extractions with this polymer indicated lower yields and higher impurities for HiSiaQM and should be therefore improved before such experiments (Chapter 4.4).

6.5.3 Hypothetical mechanism for the nanobody-inhibition of TRAP transporters

The addition of several strong binding nanobodies to the periplasmic side of the TRAP transporter showed an inhibiting influence on the transporter function in an *in vivo* cell growth assay. A comparable inhibition through nanobodies was already observed for the transporters BtuCD-F for vitamin B12 (Mireku et al., 2017) and VGLUT for glutamate (Schenck et al., 2017) and through an antibody fragment for Mn(II) MntABC transporter (Ahuja et al., 2015). However, for the BtuCD-F and MntABC transporters, the specific binders target the SBPs and block the initial substrate binding process and presumably have no interaction with the membrane domains, in contrast to the TRAP transporter nanobodies from the present work. On the basis of the hypothetical transport mechanism of TRAP transporters, several different inhibiting mechanisms are conceivable (Figure 6-16) (Mulligan et al., 2011): (I) the coupling of sodium-ions is disrupted, (II) the nanobody blocks the channel for substrate translocation, (III) a conformational state of QM-domains which for example cannot interact with the P-domain is induced and stabilized or (IV) the nanobody disturb the P-QM interaction.

The mechanism of an inhibiting effect through nanobody binding on the entrance and coupling of a sodium-ion (I) into the membrane protein cannot be exclude from presented experiments (Figure 6-16 A). However, since an inhibiting effect was obtained for nanobodies of both binding regions, it is unlikely that both regions are connected to sodium-ion interacting residues in HiSiaQM. The same conclusion can be used for mechanism (II), the blocking of the substrate translocation pathway (Figure 6-16 B). However, this mechanism presumably involves more amino acid residues than for the sodium-ion interaction, which increases the probability of an inhibiting effect through nanobodies that bind to different regions. The third potential inhibiting mechanism is based on a stabilization of a certain conformation of HiSiaQM that is functionally inactive (Figure 6-16 C). As mentioned above, such an influence of binding nanobodies were described before for other membrane proteins (Rasmussen et al., 2011a, Smirnova et al., 2014). However, to validate such an effect of the TRAP transporter nanobodies, further experiments are necessary. In this context, one interesting experiment might be the addition of the nanobodies during PELDOR distance measurements of HiSiaQM, as these experiments without any addition revealed a high structural flexibility.

The fourth inhibiting mechanism of disrupting the SBP-membrane domains interaction has presumably the lowest specificity of the nanobody binding regions and is therefore probably the most obvious one (IV) (Figure 6-16 D). Even if nanobodies are comparable small proteins, they still have the half of the size of the P-domain of TRAP transporters (around 15 kDa to 33 kDa, respectively). If the affinity for the SBP to the membrane domains in TRAP transporters is assumed in a comparable range to ABC transporter systems with low nanomolar or picomolar concentrations, the binding strength is in a similar range as for the nanobodies (Lewinson et al., 2010, Qasem-Abdullah et al., 2017, Vigonsky et al., 2013). But importantly, until today the binding affinity of TRAP transporter SBPs is not described which might be due to a low affinity or other reasons such as inconvenient buffer or solubilization conditions. Nonetheless, such a competing binding situation between a nanobody and the SBP presumably lowers the formation of the tripartite complex between SBP and membrane domains and lead to a reduced transport of the substrate. This influence might just reduce the functionality and not totally inhibit the transporter, which can also explain the slow, but detectable growth of the cell cultures after a longer time. Even if this fourth mechanism is the most obvious one, the mentioned inhibition of the secondary transporter VGLUT without an SBP suggests that also the first three mechanisms are possible (Schenck et al., 2017).

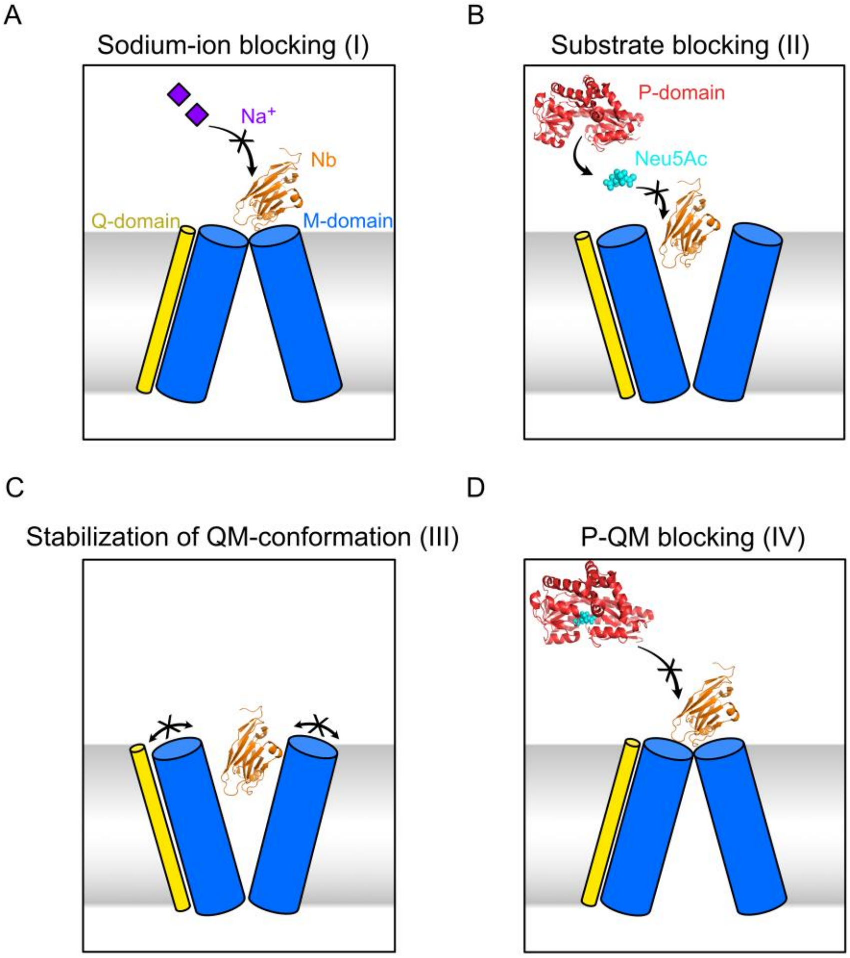


Figure 6-16: Hypothetical mechanisms of nanobody-based inhibition of TRAP transporters. A-D) Four possible inhibiting mechanism through nanobody binding (PDB-ID: 5U64, orange (Koch et al., 2017)). The QM-domains are presented in yellow and blue, the P-domains (open: PDB-ID 2CEY (Müller et al., 2006), closed: PDB-ID: 3B50 (Johnston et al., 2008)) in red with the substrate in cyan. The sodium-ions are shown as purple squares.

6.5.4 Outlook

The nanobodies that were generated and characterised in this chapter are a useful tool for many further investigations of the TRAP transporter membrane domains. They have already been used in first trials for structural investigations with X-ray crystallography to enhance necessary crystal contacts or in cryo-EM to increase the particle size and improve the data processing (Chapter 7). For this purpose, one strong binding nanobody was also successfully enlarged to a 60 kDa megabody, which is presented and characterized in Chapter 7.6. Since the structure of TRAP transporter membrane domains is still unknown, the nanobodies strongly increase the chance of a structure determination for both methods.

Further, the observed nanobody-based inhibition of the TRAP transporters function represents an interesting starting point for development of specific drugs against *H. influenzae*. As described in the introduction, the sialic acid uptake with the TRAP transporter HiSiaPQM is coupled to the virulence of the pathogen *H. influenzae* (Severi et al., 2005, Severi et al., 2007). Therefore, a compound that inhibits the TRAP transporter HiSiaPQM has the potential to act as a bacteria specific, bacteriostatic antibiotic or represents a starting point for development of specific drugs. The nanobodies from the present work are the first described compounds that showed an inhibiting effect on the function of the TRAP transporter HiSiaPQM.

In terms of antibiotic compounds, one advantage of the inhibiting nanobodies is that they just need to overpass the outer membrane into the periplasm and no crossing of the inner membrane. One possible mechanism for crossing the outer membrane is the incorporation of the nanobodies into liposomes and a later fusion of the liposomes with the outer membrane. Such experiments were already successfully performed for the delivering of an antibiotic into the periplasm (Wang et al., 2016). This procedure can presumably ensure a high concentration of nanobodies in the periplasm without any modification of the nanobody sequence, such as uptake-mediating peptides (cell-penetrating peptides, CPP or antimicrobial peptides, AMP) that are often fused to molecules for crossing of the inner membrane (Derakhshankhah and Jafari, 2018, Lee et al., 2019, Reinhardt and Neundorf, 2016). Alternatively, phage therapy with bioengineered phages can be used to deliver the nanobody genes into a bacterium (Altamirano and Barr, 2019). Doing this, the bacteria themselves express the nanobodies, similar as the *E. coli* cells in the growth assay. Since the nanobodies are still large molecules compared to the most antibiotic compounds, the description of the nanobody-inhibiting mechanism on a molecular level with HiSiaQM-nanobody complex structures can serve as starting point for the design of small molecules with similar inhibiting properties. Further to the strategies for accumulation in the periplasm, the combination of nanobodies with different binding regions, especially those with the strongest inhibiting effect, should be tested. If these nanobodies bind different epitopes or have different inhibiting mechanisms, the combined usage might increase the overall inhibiting effect. Additional, if available, HiSiaP specific nanobodies can be used in combination with the presented HiSiaQM nanobodies, which will probably further increase an inhibiting effect.

Chapter 7

Structural investigations of TRAP transporter membrane domains

Preamble:

The cryo-EM sample preparations and data processing were performed in collaboration with Veronika Heinz and Christine Ziegler from the University of Regensburg. The cryo-EM data sets were recorded at the Center for Integrative and Translational Bioimaging (Bettina Böttcher) of the University of Würzburg and with Alexandre Durand at the IGBMC (Institute of Genetics and Molecular and Cellular Biology) in Strasbourg (FR).

7.1 X-ray crystallography and cryo-EM for membrane proteins

The most important and often most challenging task during studies of proteins is the description of their molecular structure. Such a structure allows to analyse and postulate specific amino acid interactions and binding characteristics of functional domains, and often immensely increase the understanding of the function of a protein. Until today, the world of macromolecular structures is mainly built up from three methods, X-ray crystallography (154264 structures, 88.4% from all structures), nuclear magnetic resonance (NMR, 13285 structures, 7.6%) and electron microscopy (EM, 6989 structures, 4.0%) (PDB, www.rcsb.org, access date 17.02.2021). While each method has its own advantages and disadvantages, the overall importance of structure determining methods for the understanding of nearly all processes in living cells is emphasized when considering the Nobel prices in this field. Every of the three methods or pioneering results which are closely related to a protein structure determination were awarded with at least one Nobel price, X-ray crystallography for example in 1946, 1962, 1964, 1972, 2006, 2009 and 2012, NMR in 2002 and cryo-EM in 2017 (Nobel Media AB 2020).

The first membrane protein structures were solved around 40 years ago, at low resolution of around 7 Å in 1975 and at high resolution of around 3 Å in 1985 with X-ray crystallography (Deisenhofer et al., 1985, Henderson and Unwin, 1975). Until today, this method is mainly responsible for the exponential growth of described structures in the last years, including the description of 50% of all unique membrane protein structures between 2011-2016 (Hendrickson, 2016). X-ray crystallography protein structures allow a detailed insight into the biomolecules with high-resolutions and by this, provide a fundamental basis for functional description. Due to the high importance, the method is steadily further developed and modified regarding the crystallization process with automatization, the improvement of data collection in resolution and time saving and the simplification of structure refinement and building. One crucial and often limiting step for X-ray crystallography of proteins is the preparation of protein crystals, which rely on defined intermolecular distances between protein chains to build up a 3-dimensional crystal. These intermolecular interactions are especially low for membrane proteins due to the essential presence of a membrane mimetic system around the hydrophobic protein regions. This might be one of the reasons, why despite the large numbers of membrane protein encoding ORFs in all types of organisms less than 3% of all protein structures are assigned to membrane proteins (Chapter 4.1).

The two other commonly used methods, NMR and cryo-EM, avoid the limiting step of protein crystallization. However, also these methods have important limitations and challenges for protein structure determination. For example, NMR is limited in the size of the biomolecule with a maximum molecular weight of around 30-50 kDa (subunit size) (Gauto et al., 2019). For the most membrane proteins, this limitation represents an exclusion criterion and therefore less than 1% of the NMR structures in the PDB are related to membrane proteins (112 of 13285, PDB, www.rcsb.org, access date: 17.02.2021). Additionally, these structures are often just smaller domains or helices and not complete membrane protein structures.

This proportion of membrane protein structures to the overall determined structures by a single method changed drastically for the method of cryo-EM. In this, around 19% of the published structures are related to the class of membrane proteins (1311 of 6989, PDB, www.rcsb.org, access date: 17.02.2021). Especially the development during the last decades in hardware and in software for data processing, led to an important improvement of the resolution in EM (Frank, 2017, Merk et al., 2016, Nogales, 2016). Through these developments, cryo-EM has become a competitive method in the field of structural biology in the last years. Protein EM-structures can be described nowadays with high resolution, for examples with up to 1.22 Å of a soluble protein or 1.9 Å for a membrane protein (Flores et al., 2020, Nakane et al., 2020). Further, the successful detection of smaller single particles enables the investigation of a larger number of proteins with cryo-EM and today, structure determination of proteins smaller than 100 kDa is possible (Herzik et al., 2019).

Importantly for membrane proteins, cryo-EM is compatible with the commonly used membrane mimetic systems, with detergent micelles and the presented lipid bilayer nanodiscs in both forms, protein- and polymer-bounded (Arkhipova et al., 2020, Sun et al., 2018, Tang et al., 2019). The usage of nanodiscs in cryo-EM has also the advantage that the size of the particles is increased by a molecular weight of around 100 kDa (Efremov et al., 2017). This can be a crucial modification, as these larger particles are more easy to detect and was already successfully applied for small transporters with no or only small outer-membrane domains (Arkhipova et al., 2020, Gao et al., 2016). Additional to the methodical properties of cryo-EM, which represent a high compatibility for structure determination of membrane proteins, the method also facilitates the observation of several protein conformations in a short time or even in the same experiment (Arkhipova et al., 2020, Hofmann et al., 2019). Especially the recent work of Hofmann et al. (Hofmann et al., 2019) revealed an unprecedented insight into the dynamic of an ABC transporter with description of eight different conformational states at high resolution. As mentioned before, conformational changes are generally essential for transport of a substrate and a simplified and faster detection of different conformational states strongly supports a better description of the full function.

In the present work, a combined approach of both methods, X-ray crystallography and cryo-EM, was used to study the protein structure of the HiSiaQM membrane protein. Since both methods rely on a different experimental preparation and diverse requirements on the protein, this combined approach should increase the probability to determine the structure of the membrane proteins.

7.2 Preparation and diffraction of HiSiaQM crystals

In the following chapter, approaches with varied conditions and intermediate goals for determination of the HiSiaQM protein structure with X-ray crystallography are presented. For all crystallization trials with HiSiaQM, β-DDM detergent was used standardly for extraction and purification and was replaced in a washing step during Ni^{2+}-AC with α-DDM. The use of this alternative DDM conformer should improve the crystallization of the TRAP transporter membrane domains, as reported by Nien-Jen Hu (National Chung Hsing University, Taiwan, personal communication).

An example purification of HiSiaQM wildtype protein in detergent micelles for crystallization is illustrated in Figure 7-1, based on the optimized protocol that was elaborated in Chapter 4.2. The protein was purified as described in Chapter 9.1.4.2 by using 25 mM imidazole during the washing step on the Ni^{2+}-AC for reduction of impurities and contaminants. During this step, the detergent was changed to α-DDM. The subsequent SEC elution profile illustrated a main protein peak for HiSiaQM at around 12 mL elution volume and small amounts of aggregated protein that were separated (Figure 7-1 A). The protein peak was analysed via SDS-PAGE, which confirmed a high quality of pure HiSiaQM protein with minimal contaminations (Figure 7-1 B). The corresponding SEC fractions were combined, concentrated to around 15 mg/mL and used for crystallization.

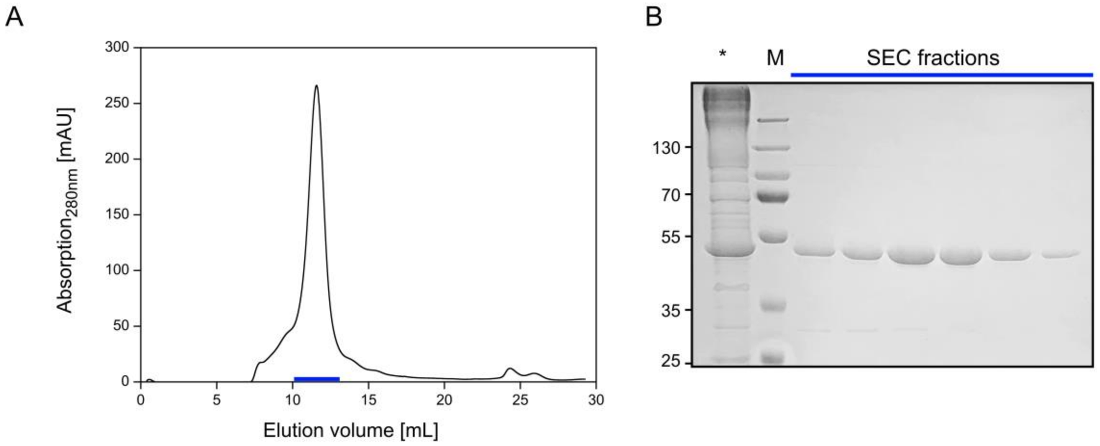


Figure 7-1: Purification of HiSiaQM for crystallization. A) SEC run of HiSiaQM after Ni^{2+}-AC on an SD 200 10/300 column (buffer: 50 mM KH_2PO_4 (pH 7.8), 200 mM Tris, 10% Glycerol, 0.035% α-DDM). **B)** SDS-PAGE of HiSiaQM purification with fractions of the Ni^{2+}-AC elution (*), the protein standard (M) and from the SEC run in A).

7.2.1 Crystallization of transporter domains

As a first crystallization attempt, the membrane domains HiSiaQM were used without any supplements. In addition to all available standard protein crystallization screens (see Chapter 10.1), commercial screens that are optimized for membrane proteins, *MemGold* and *MemGold II* (both Molecular Dimensions, UK) were used. In several conditions, crystallization was observed after a few weeks. Especially the *PACT* screen showed a large number of hits (Figure 7-2 A-C). Most of the observed crystals had a prism or rod shape and were large enough for diffraction experiments. The diffraction images of three crystals with the highest resolution are shown in Figure 7-2. Unambiguously, the recorded diffraction spots revealed the evidence of protein crystals and excluded salt crystals. Even if diffraction spots were clearly visible, the resolution of the diffraction spots was detected at around 20 Å, which was far too low for a further processing and determination of a structure.

As a further approach, the membrane domains HiSiaQM were co-crystallized with HiSiaP and the native substrate sialic acid. Doing this, some crystallization conditions might promote the formation and crystallization of the tripartite complex. Albeit the complex was never observed experimentally before, the crystal lattice might stabilize the weak tripartite complex. The P-domain in complex with the DDM-stabilized membrane domains will increases the number of possible intermolecular interactions, which are necessary for formation of well-ordered protein crystals.

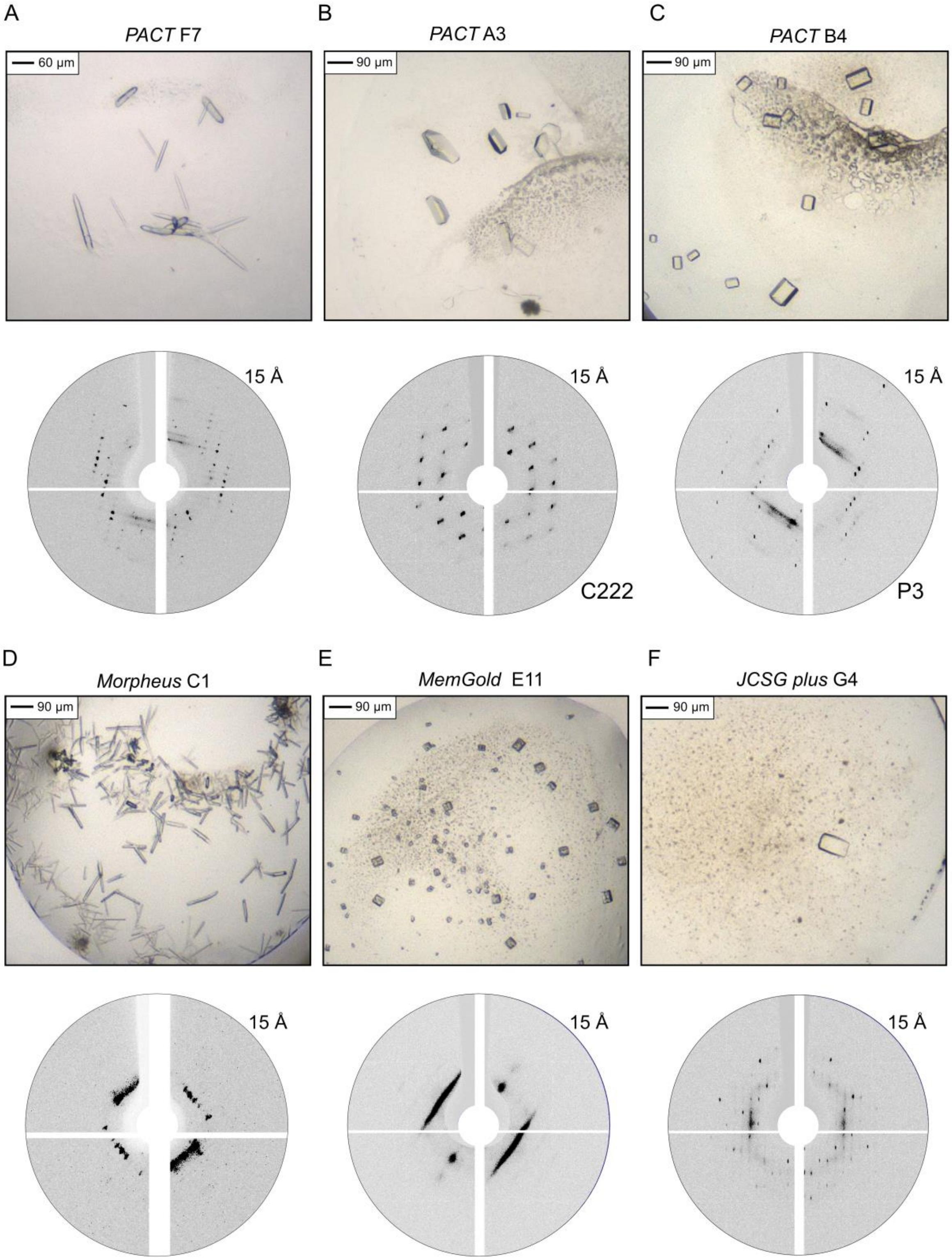


Figure 7-2: Crystals and diffraction images of HiSiaQM. A-C) Top: Conditions with resulting crystals in a sitting drop, imaged shortly before collecting for diffraction. Bottom: Diffraction images, the limiting, outer circle represents a resolution of 15 Å. In the case of a possible indexing, the resulting space group of the diffraction images are given. **D-F)** Same as A-C) but with HiSiaQM-HiSiaP mixture.

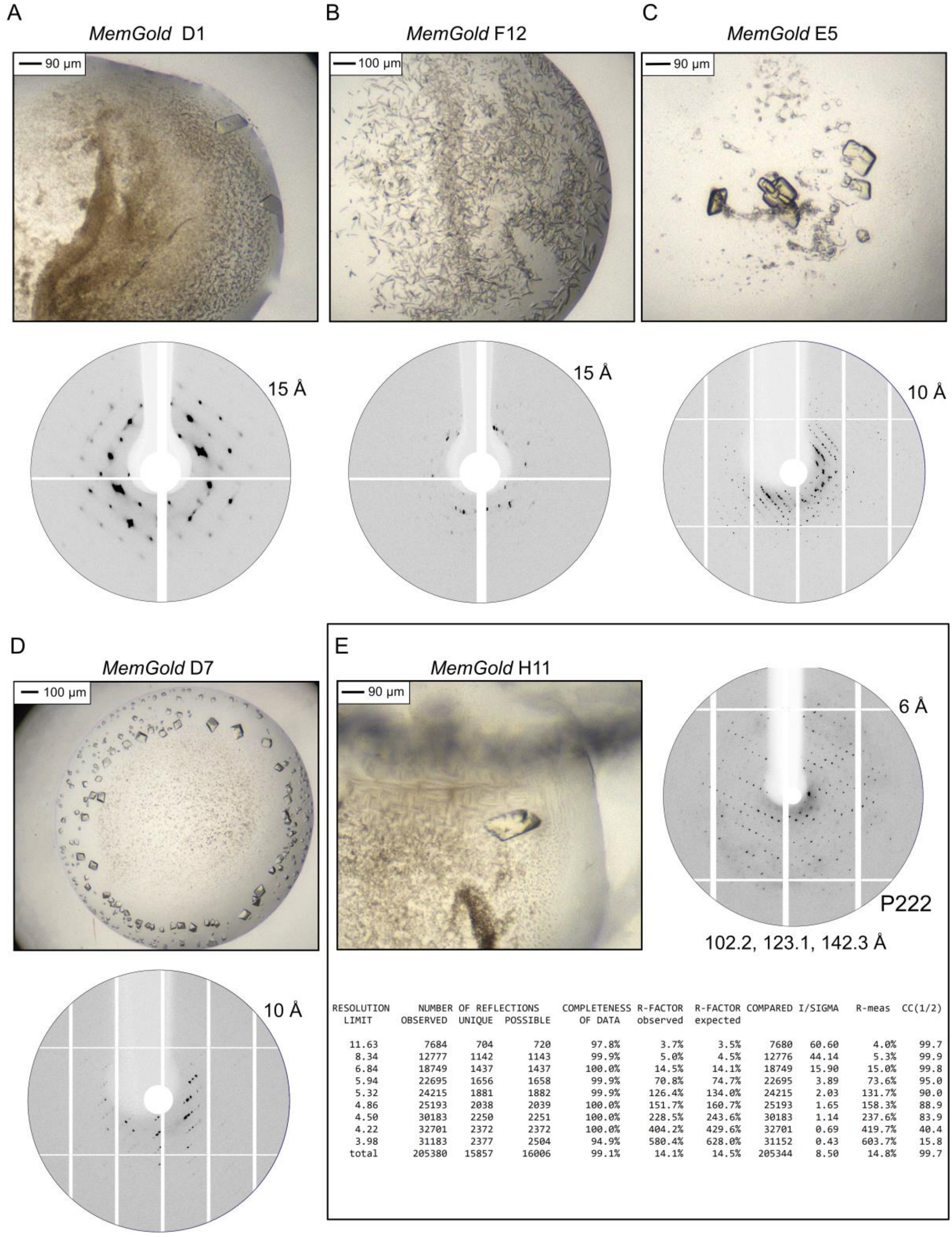
In addition to increased intermolecular interactions, the structure of the P-domain had already been described in different conformational states on an atomic scale. Therefore, if the P-domain is part of the crystal, it can be used as a search model for molecular replacement. As before, crystallization was obtained in several conditions and example crystals are illustrated in Figure 7-2 D-E. Most of the crystals have a similar rod or a prism shape as the HiSiaQM crystals before. The diffraction resolution of these crystals was quite low with diffraction spots at a maximum resolution of 20-25 Å. Due to the simultaneous progress in other crystallization trials in the following chapters, these crystallization trials were not further optimized or investigated to get better diffraction crystals.

7.2.2 Crystallization of HiSiaQM with nanobodies

The generation of HiSiaQM specific nanobodies (Chapter 6) enabled the co-crystallization of the membrane domains HiSiaQM with a specific binder. The advantages of specific binders are similar to those of the tripartite complex from the paragraph above. The binders represent an outer-membrane domain and increase the possibility of intermolecular interactions. In contrast to the SBP, the nanobodies were successfully characterized as strong binders with high affinities in the nanomolar range (Chapter 6). The possibility of a HiSiaQM-nanobody complex in the crystallization trial is therefore high. Further, the structure of the nanobodies is known and they can potentially be used for molecular replacement during phasing. Due to the detection and description of nanobodies for two different binding regions (Chapter 6.3.4), binders for different regions on HiSiaQM were simultaneously used to increase the mentioned advantages.

For co-crystallization, the membrane domains were supplemented with the strong binding nanobody 3 and in another approach with nanobody 3 and 4, as these two nanobodies were identified as binders for two different regions. Growth of single crystals was observed in several conditions and screens, particularly with the most crystals in the *MemGold* screen (Figure 7-3). The shape of the observed crystals from this protein mixture varied for different conditions. For example, the crystals in condition *MemGold* D7 had a bipyramidal shape, while condition *MemGold* F12 crystallized as needles and *MemGold* E5 and H11 with an unregular plate (Figure 7-3). The diffraction images of two nanobodies revealed an overall higher resolution than with one nanobody. Additionally, the resolution of these crystals was higher than in the previous experiments without additives or addition of HiSiaP. The best diffracting crystal in *MemGold* H11 was recorded in a recent diffraction experiment with a resolution of up to 5 Å (Figure 7-3 E). This crystal showed by far the best resolution from all crystals that were investigated and allowed for the first time a further processing. The parameters of the crystal after indexing and integration are shown in Figure 7-3 E. From the diffraction spots, an orthorhombic crystal system was determined with unit cell parameters of 102.2, 123.1 and 142.3 Å. The calculation of the cell content resulted in an estimated solvent content of 72.6% and 45.1% for a monomer or a dimer of HiSiaQM-nanobody complex in the unit cell, respectively (calculated with MATTPROB (Kantardjieff and Rupp, 2003, Matthews, 1968, Weichenberger and Rupp, 2014)). However, the protein crystal was measured recently and the phases for calculation of the electron density and model building could not be reconstructed until today.

In contrast, the other crystallization conditions from Figure 7-3 with resolutions up to 10-15 Å were already optimized each in 48 new solutions by variation of the initial pH-value and the salt and precipitant concentration. For optimization of *MemGold* D1, 10 conditions with possible protein crystals were detected. The shape of most crystals was similar to the original prism-shaped crystals in *MemGol*d D1, even though also crystals were formed with a bipyramidal-shape (Figure SI-38). The diffraction images revealed an increase of the diffraction resolution compared to the initial experiment. However, the best crystals showed a resolution of around 10-11 Å. For optimization of condition *MemGold* F12, crystals were only obtained at an incubation temperature of 4 °C. Similar to the original condition, the crystals have a rod-shaped appearance. The diffraction of these crystals was limited by a maximum resolution of 10 Å (Figure SI-38). Further, the conditions *MemGold* E5 and *MemGold* D7 were optimized and crystals were obtained for each optimization after several days (Figure SI-38). The observed crystals had a similar crystal shape as the corresponding initial conditions. The diffraction images revealed an increase of the resolution to 8 Å, which is still too low for determination of the protein structure. Overall, the crystallization trials with the TRAP transporter specific nanobodies showed an important increase of the quality of the protein crystals with better resolutions of the diffraction images than in crystallization trials without nanobodies. Nonetheless, the resolution of all crystals was low and further possibilities to receive higher diffraction crystals and to facilitate the processing procedure are discussed in Chapter 7.7.



```
RESOLUTION     NUMBER OF REFLECTIONS    COMPLETENESS R-FACTOR  R-FACTOR COMPARED I/SIGMA   R-meas  CC(1/2)
  LIMIT     OBSERVED  UNIQUE  POSSIBLE     OF DATA   observed  expected

   11.63        7684     704       720       97.8%       3.7%      3.5%     7680   60.60     4.0%    99.7
    8.34       12777    1142      1143       99.9%       5.0%      4.5%    12776   44.14     5.3%    99.9
    6.84       18749    1437      1437      100.0%      14.5%     14.1%    18749   15.90    15.0%    99.8
    5.94       22695    1656      1658       99.9%      70.8%     74.7%    22695    3.89    73.6%    95.0
    5.32       24215    1881      1882       99.9%     126.4%    134.0%    24215    2.03   131.7%    90.0
    4.86       25193    2038      2039      100.0%     151.7%    160.7%    25193    1.65   158.3%    88.9
    4.50       30183    2250      2251      100.0%     228.5%    243.6%    30183    1.14   237.6%    83.9
    4.22       32701    2372      2372      100.0%     404.2%    429.6%    32701    0.69   419.7%    40.4
    3.98       31183    2377      2504       94.9%     580.4%    628.0%    31152    0.43   603.7%    15.8
   total      205380   15857     16006       99.1%      14.1%     14.5%   205344    8.50    14.8%    99.7
```

Figure 7-3: Crystals and diffraction images of HiSiaQM with nanobodies 3 and 4. A-D) Top: Conditions with resulting crystals in a sitting drop, imaged shortly before collecting for diffraction. Bottom: Diffraction images, the limiting, outer circles represent the resolution of the diffraction image. **E)** Highest resolution crystal of HiSiaQM with nanobody 3 and 4 from crystallization condition *MemGold* H11. The recorded diffraction images for this crystal were processed with XDS (Kabsch, 2010) and the results are listed.

7.3 EM of HiSiaQM in protein-bounded nanodiscs

In addition to the structural investigations with X-ray crystallography, it was also attempted to determine the structure of HiSiaQM with cryo-EM. The SEC runs, EM sample preparations, data recording and processing were performed in cooperation with Veronika Heinz and further members from the group of Christine Ziegler from the University of Regensburg.

7.3.1 Sample preparation for EM

Immediately before the preparation of the EM samples in the following chapters, the protein samples were loaded on a SEC column. This procedure was used to guarantee that any traces of aggregates were removed from the sample and to yield particles with a homogenous size. The first attempt of a sample preparation for structure determination with cryo-EM was performed with HiSiaQM in the protein-bounded nanodiscs (Chapter 4.3). The reconstitution should create particles with a size of ~10 nm that are large enough for investigations with cryo-EM (Arkhipova et al., 2020, Efremov et al., 2017, Ritchie et al., 2009).

The reconstitution approaches and SEC runs for initial EM sample preparation are shown in Figure 7-4. To receive a high protein concentration after SEC without a concentration step that was assumed to promote formation of new aggregates, a large amount of protein was applied onto the SEC column. For this, around 10 HiSiaQM standard purifications, each from protein expression in 10 L LB medium, were pooled and used in three batches for nanodisc reconstitutions. The SDS-gel of the three batches before SEC revealed a low quantity of impurities and clear protein bands for HiSiaQM and MSP1 (Figure 7-4 A). The SEC of the first batch resulted in a broad elution peak with unsuccessful separated maxima, as observed for reconstitutions before (Figure 7-4 B, compare Chapter 4.3). Same result was obtained for batch number 3 (Figure 7-4 D). In contrast, batch 2 showed the peak with two separated maxima, as expected from previous experiments, with also clear separation of a small fraction of aggregates at around 7.5 mL (Figure 7-4 C). The only differences between the three reconstitution batches were the use of separately purified MSP1. However, since all MSPs were all purified in the identical way with equal quality checks, the reason for the different SEC profiles remained unknown.

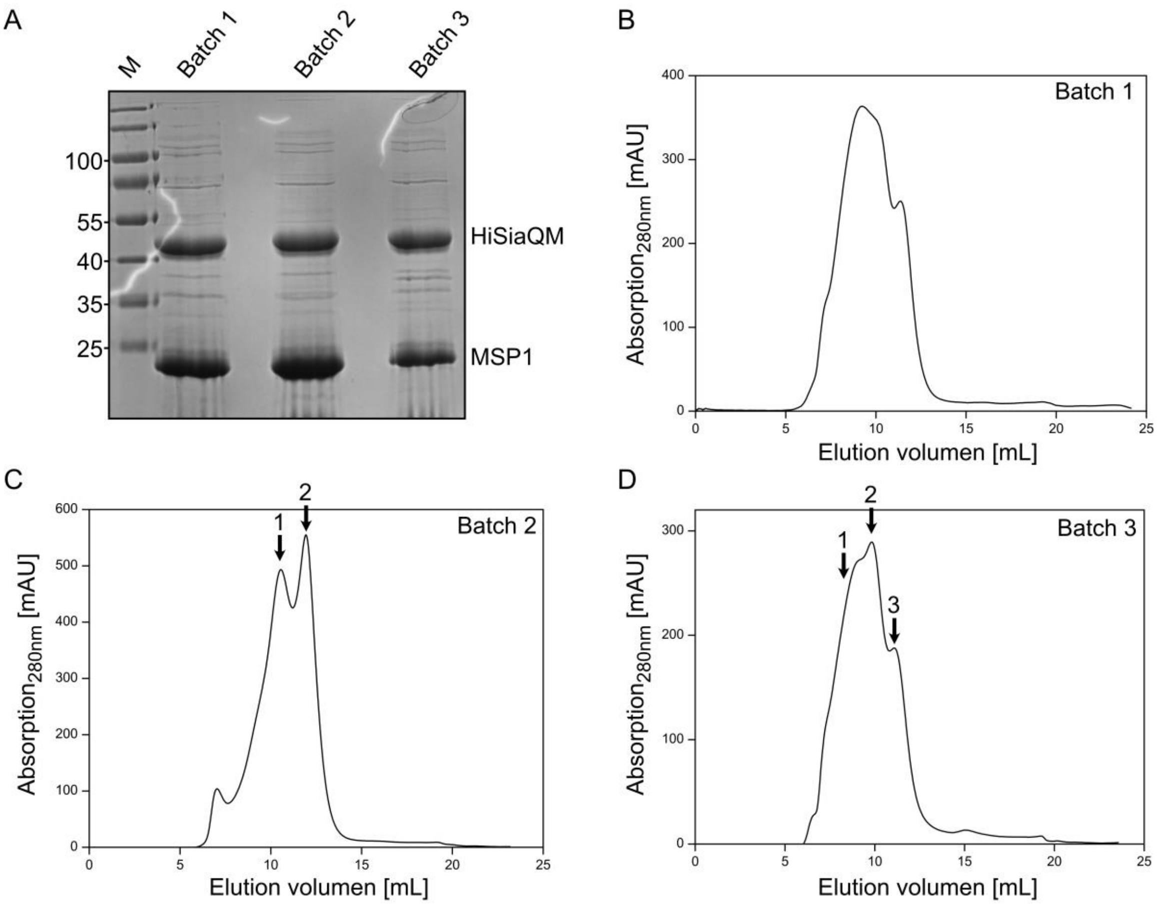


Figure 7-4: MSP-nanodisc reconstitutions of HiSiaQM for EM. A) SDS-PAGE with samples of three separated nanodisc reconstitution batches in large scale approaches. The two main bands in each sample correspond to HiSiaQM and MSP1 (around 50 kDa and 25 kDa, respectively). **B)** SEC run of reconstitution after dialysis from batch 1 (column: SD 200 10/300, buffer: 50 mM Tris (pH 7.8), 50 mM NaCl). **C-D)** Same as A) but for the other batches. Fractions of the SEC that were used for EM sample preparations are marked.

7.3.2 Negative stain EM of HiSiaQM in MSP-nanodiscs

Initially, negative stain EM was used to characterize the protein samples concerning particle size and shape, particle concentration and impurities with aggregated fragments. Due to the inconsistent SEC elution profiles, different fractionated samples from SEC runs of batch 2 and 3 were investigated with EM after negative staining (Figure 7-4 C and D). Representative images that were recorded for the different eluted fractions are illustrated in Figure 7-5 for batch 2 and Figure SI-39 for batch 3. The images from the first protein peak of batch 2 after a 1:100-dilution revealed detectable, single particles in a satisfying concentration and a low degree of aggregation (Figure 7-5 A). The particles have an overall homogeneous size distribution and a shape that was expected for a nanodiscs sample. The size of the particles was estimated to around 11-13 nm, which correlates quite well with the described size for MSP1-nanodiscs of 9.8 nm (Ritchie et al., 2009). The images also revealed a preferable top-view orientation of the nanodiscs, but side-views were also clearly observable. Also, a few side-views were visible with formed dimers or also higher oligomers up to 4-5 single nanodiscs, by sticking the top and bottom of the nanodiscs together (Figure 7-5 A). This observation is presumably caused by the nanodiscs assembly under the negative stain conditions, as samples with other proteins also showed this effect (personal communication with Veronika Heinz, University of Regensburg).

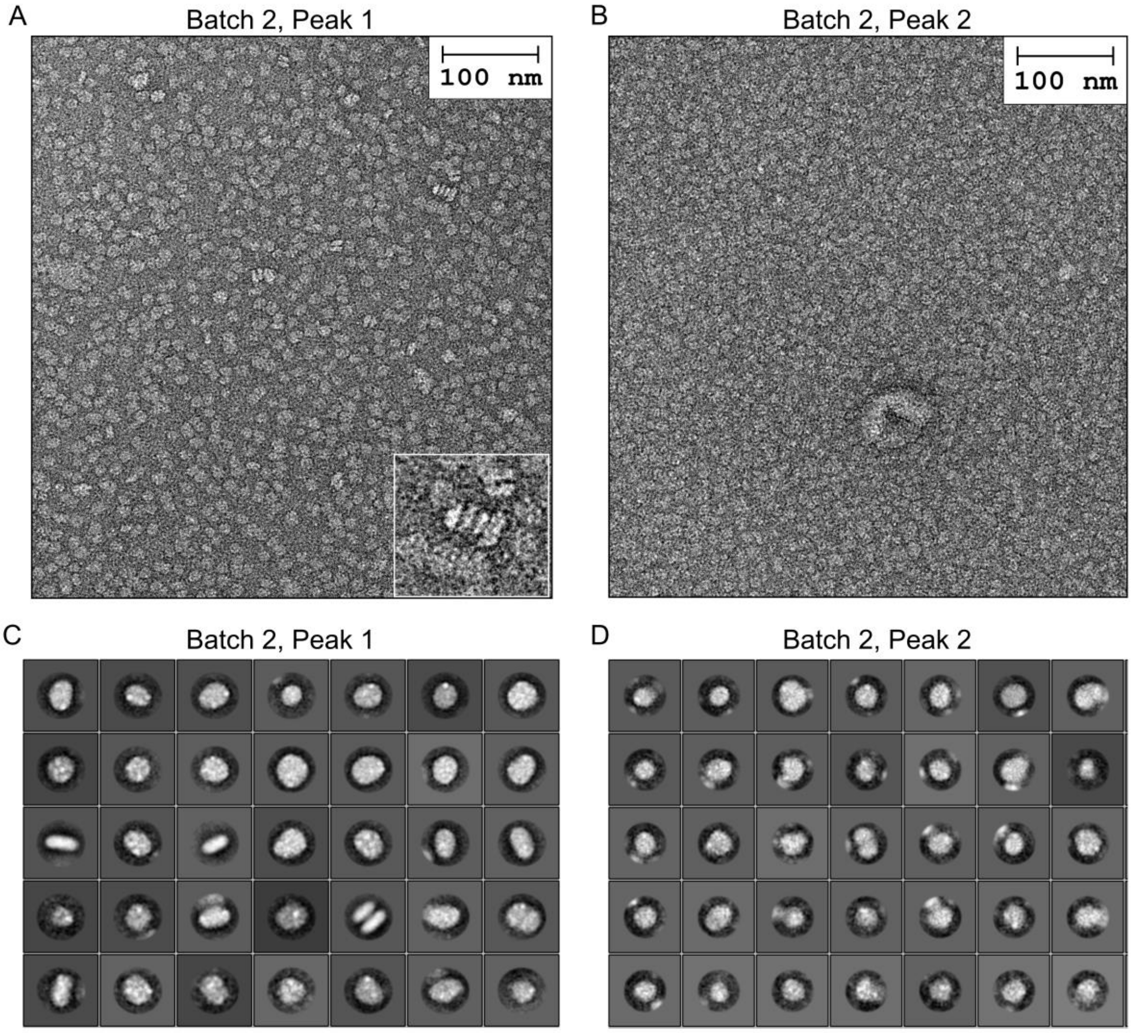


Figure 7-5: Negative stain EM of HiSiaQM in MSP-nanodiscs. A) Negative stain EM image of HiSiaQM reconstituted into MSP-nanodiscs. The sample was taken from the first SEC-fraction in Figure 7-4 C and diluted by a factor of 100. The solution was stained with 2% uranyl acetate on glow charged grids. **B)** Same as A) but for the second SEC-fraction in Figure 7-4 C. **C)** 2D classes from several negative stain EM images from protein sample that was used in A). **D)** Same as C) but with protein sample from B).

The negative stain EM image for the second SEC protein peak of batch 2 revealed similar particles which seems to be slightly smaller than for the first SEC fraction (Figure 7-5 B). An obvious difference in the EM images between the two SEC fractions was the absence of dimerized or oligomerized particles in the second peak sample. The negative stain EM images for SEC fractions from batch 3 revealed a similar result as for batch 2 and a reason for the different SEC elution profiles was not visible from the single particle observation. For a more detailed characterization of the nanodiscs from Figure 7-5 A and B, the particles from recorded images were assigned to 2D classes (Figure 7-5 C and D). This classification for particles from the first SEC fraction from batch 2 showed preferable top-view orientation, but also several classes with particles in a side-view orientation were found (Figure 7-5 C). Additionally, 2D classes with dimeric nanodisc particles, as visible in the EM images, were obtained. Due to the resolution of negative stain EM of around 10-20 Å, the detection of the membrane protein or further structural details were not feasible with this dataset (Gallagher et al., 2019).

7.3.3 Cryo-EM with HiSiaQM in MSP-nanodiscs

Judged by the characterization of the HiSiaQM-nanodiscs sample with negative stain EM, the sample was of high purity and homogeneity. Thus, the first SEC fraction of SEC run from batch 2 was prepared by the group of Christine Ziegler for cryo-EM by vitrification with a Vitrobot plunge freezer (ThermoFisher Scientific) (experimental parameters in Table 9-9). For this, the sample was carefully concentrated to estimated 2000 mAU to receive an optimal concentration of the protein, after verification with negative stain EM that this step caused no aggregation of the sample.

The cryo-EM samples were analysed on a Titan-Krios G3 electron microscope, equipped with a Falcon III camera (Bettina Böttcher, University of Würzburg, experimental parameters in Table 9-9). An example electron micrograph from the vitrified HiSiaQM nanodiscs sample is shown in Figure 7-6 A. The image revealed detectable particles in a satisfying concentration, with numerous visible and distinguishable single particles. Importantly, no larger aggregates were visible on the micrographs that reduced the amount of detectable single particles. The detected particles showed a similar shape as before in the negative stain EM images and indicated the successful vitrification of the nanodiscs. Further, the micrographs revealed several orientations of the single particles, top- and side-views were already visible from the unprocessed images.

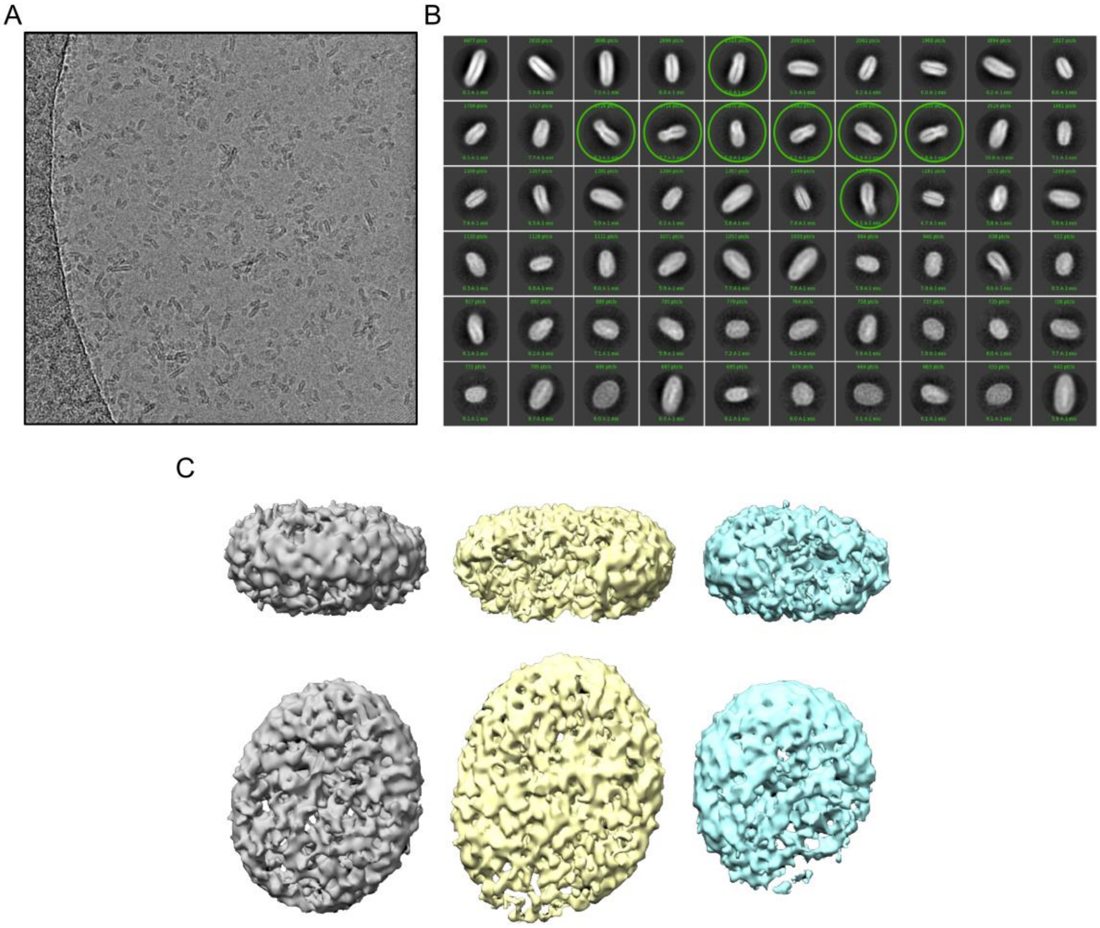


Figure 7-6: Cryo-EM of HiSiaQM in protein-bounded nanodiscs. A) Electron micrograph of HiSiaQM in MSP-nanodiscs from first SEC-peak of reconstitution batch 2 (Figure 7-4). **B)** 2D-classes of nanodisc particles of electron micrographs from same sample as A), processed with cryoSPARC (Punjani et al., 2017). **C)** 3D reconstitutions of single particle classes from B) in side-view (top) and top-view (bottom) orientations. The groups include 47647 (grey), 29143 (yellow) and 23226 (blue) individual observed particles.

Overall, 1759 micrographs were recorded and 1454 were successfully processed for further analysis. From these images, 631911 single particles were identified and localized in an autopicking process with an average of 435 particles per image. The assignment of particles in 2D classes was performed with the software program cryoSPARC (Punjani et al., 2017). After 10 rounds of classifications, 100043 particles were assigned in 100 classes with well-defined particles in each class (Figure 7-6 B). The 2D classes revealed a preferred side-view orientation of the nanodiscs in the cryo-EM sample, in contrast to the preferred top-view orientation in the negative stain EM samples. However, also some 2D classes with top-view orientated particles were observed in the vitrified sample. Interestingly, some 2D classes with side-view particles revealed regions with little indentation and at these positions, also a slightly higher density inside the nanodiscs (Figure 7-6 B, green circles). As final processing step, the identified particles were used for a 3D ab initio construction in 3 classes. The resulting 3D models that were obtained are illustrated in Figure 7-6 C. Overall, the models clearly indicated the expected nanodisc shape. For some positions of the models, especially the grey coloured model, a high density around the nanodisc was visible, possibly caused by the MSP1 belt. The inspection of the lipid bilayer and protein region of the nanodiscs revealed only unspecific density with no detectable structural motifs.

The structure of the HiSiaQM protein in the nanodiscs could not be reconstituted, mainly because of two problems. First, the particle alignment was focused on the nanodisc shape since no other structural features were visible. This is especially the case for membrane proteins that are fully-embedded in the membrane and have no large outer-membrane regions, as postulated for HiSiaQM (Mulligan et al., 2011). Such an alignment results in a low resolution of nanodiscs embedded molecules with an unspecific density. The second problem might be the indistinguishability of nanodiscs with and without HiSiaQM in the lipid bilayer. The mixing of these filled and empty nanodiscs reduces the overall resolution of embedded protein. In the following chapters, several modifications and variations of these initial EM experiment were performed that should help to overcome these problems and limitations.

7.4 Nanobodies as tool for HiSiaQM structural analysis with EM

To avoid the mentioned problems from the initial cryo-EM experiments, strong binding nanobodies from Chapter 6 were used as specific HiSiaQM binders. The binding of the nanobodies was presumed in the characterization experiments to be on the solvent accessible periplasmic side of HiSiaQM (Chapter 6). Therefore, a nanobody that binds to the HiSiaQM-MSP-nanodisc assembly should be visible as an outer membrane domain. By this it can be probably used to identify HiSiaQM-nanodiscs and to improve the particle alignment during classification.

7.4.1 Sample preparation for EM

A reconstituted HiSiaQM-nanodisc sample was supplemented with a 5 times molar excess of nanobody 6 which has a binding affinity of 640 pM to HiSiaQM (Chapter 6.3.3). The high excess guaranteed that every HiSiaQM protein was bound by a nanobody. To remove excess nanobody and aggregates, a SEC run was performed before EM sample preparation (Figure 7-7). The first peak at 2 mL elution volume corresponds to HiSiaQM and the second peak at 2.6 mL to the excess nanobody. The HiSiaQM-nanodisc fraction eluted in a peak with only one maximum and not in a double peak, as observed before for several other reconstitutions (Chapter 4.3). This presumably results from the usage of a Superose 6 5/150 SEC-column with a changed separation behaviour than the previously used SD 200 10/300 column. Nevertheless, the nanodisc peak was obtained at the expected position with high intensity and separated from aggregates and unbound nanobodies, and was therefore used for preparation of cryo-EM samples. As before for vitrification, the sample was slightly concentrated to the half of original volume to receive particles with a theoretical absorption of 2000 mAU. The experimental parameters for vitrification are listed in Table 9-10.

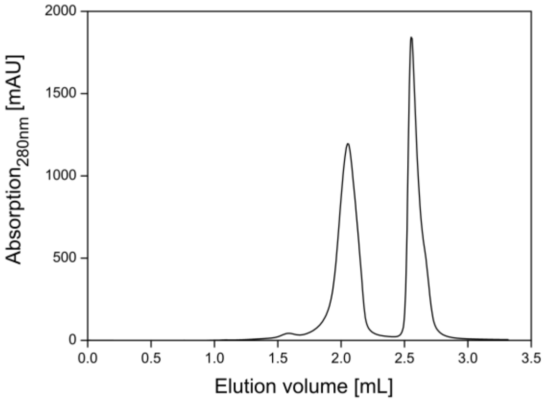


Figure 7-7: MSP-nanodisc reconstitution of HiSiaQM for EM with nanobody. SEC run of HiSiaQM in MSP-nanodiscs after incubation with nanobody 6 in 5x molar excess. A Superose 6 increase 5/150 column was used for this experiment (Äkta micro, buffer: 20 mM Tris (pH 8.3), 100 mM NaCl).

7.4.2 Cryo-EM of HiSiaQM with nanobodies

The cryo-EM electron micrographs of HiSiaQM with nanobody 6 were recorded with a similar set-up as before on a Titan-Krios G3 electron microscope with a Falcon III camera

(Table 9-10, Bettina Böttcher, University of Würzburg). One exemplary electron micrograph of the vitrified sample is shown in Figure 7-8 A. As before, the image revealed an overall good quality of the sample, without any contaminations or artifacts. Single nanodisc particles were clearly visible in a satisfying concentration. In the first processing step, 141660 particles were picked from 500 collected images with an average of 286 particles per image.

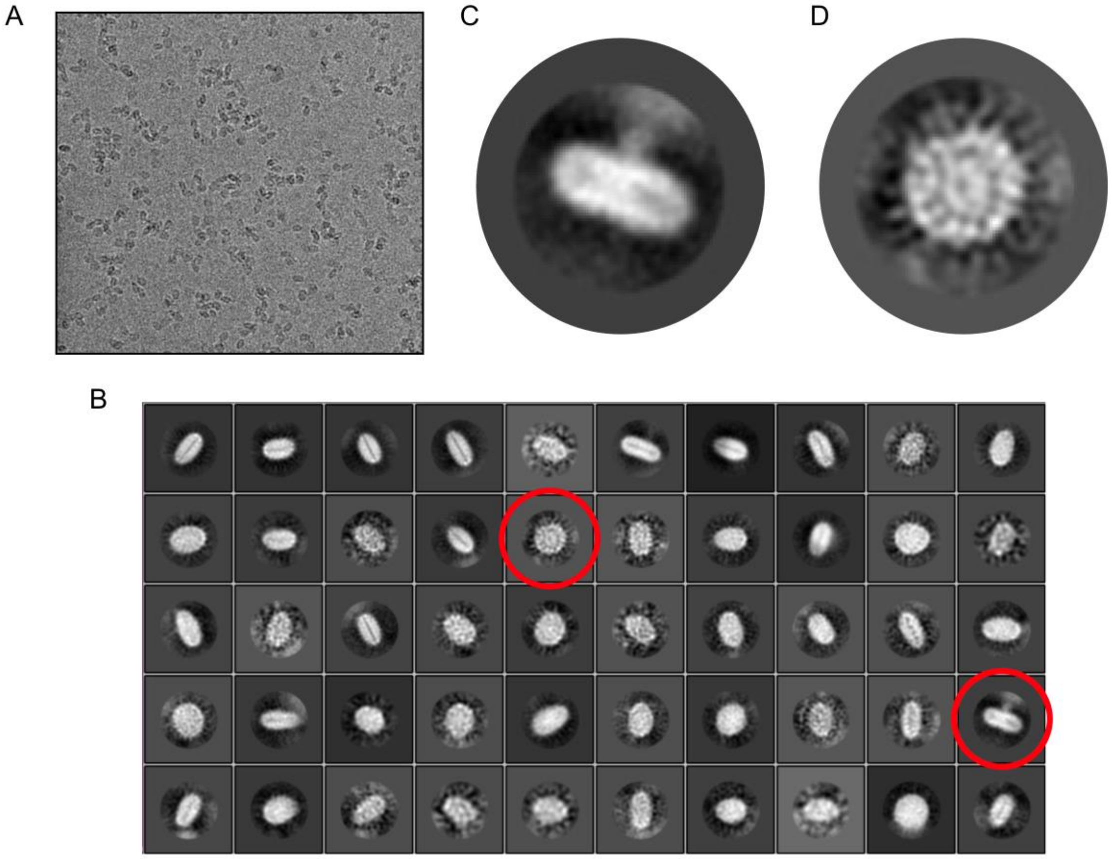


Figure 7-8: Cryo-EM of HiSiaQM with nanobody. A) Electron micrograph from vitrified HiSiaQM in MSP-nanodiscs with addition of nanobody 6 from first SEC peak in Figure 7-7. **B)** 2D classification of nanodisc particles and 8 rounds of classifications with RELION (Scheres, 2012). **C-D)** Larger representation of two 2D classes from B) (red circles) with presumably structural features of HiSiaQM and nanobody.

The particles were assigned into 2D classes and after 8 rounds of classification, 35052 well defined nanodisc particles were obtained in 2D classes (Figure 7-8 B). On a first view, these particles shared a high similarity to the first recorded cryo-EM data set for HiSiaQM without nanobodies. Again, both orientations of the nanodiscs, side- and top-views, were visible with larger proportion of top-view orientations than for the first sample (Figure 7-6 B). On closer inspection, a few 2D classes contained some additional structural features (Figure 7-8 C and D). For example, one class of particles with a side-view orientation indicated an additional density on the side of the nanodiscs. In the case of nanobody binding to HiSiaQM in a nanodisc, such an observation was expected. Additionally, the class indicated a slightly changed form compared to side-view particles from the first 2D classes and a higher density contrast inside the nanodisc, as already observed in the previous dataset (Figure 7-8 C). Further, also a few top-view particle classes seemed to indicate a structural motif inside the nanodisc (Figure 7-8 D). However, as before, the further processing of the 2D classes was unsuccessful and revealed no interpretable density inside the lipid bilayer, as well as no 3D model with a detectable density for the nanobody on the side of the nanodisc. This might be presumably caused by the same reasons as for the first dataset, the insufficient separating of HiSiaQM-filled and empty nanodiscs and no or only two less structural features that can be used for alignment of the membrane protein inside the nanodiscs.

7.5 Design and characterization of a TRAP transporter megabody

The small molecular weight and size of nanobodies of around 15 kDa are an advantage for many applications and are also reasons for their high conformational stability and resistance against external factors such as temperature changes or protease digestion (Dumoulin et al., 2002, Hussack et al., 2011b). However, as marker for cryo-EM, large molecules with high binding affinities and antigen specificities, similar as for nanobodies, are needed. One possible alternative are classical antibodies with a size of around 150 kDa, but these molecules are more difficult to produce and more expensive than nanobodies, which can be easily expressed in prokaryotic expression systems with high yields as described in Chapter 6.2.3 (Steeland et al., 2016, Pardon et al., 2014). To prevent such problems or high costs, the group of Jan Steyart (University of Brussels) developed a nanobody-derived chimeric protein, called megabody (Uchański et al., 2021). These proteins are built up from a nanobody, which was ideally already characterized as strong binder to the antigen, and a larger bacterial protein with an accessible β-turn between two β-sheets. Here, the β-turn adhesin domain from *Helicobacter pylori*, called HopQ, was fused to the β-turn between the first and second β-sheet of the nanobody to produce a rigid bridge between the two proteins (Figure 7-9 A) (Uchański et al., 2021).

7.5.1 Construction and production of nanobody-derived megabody

The cloning routine in this work is based on Uchański et al. (Uchański et al., 2019, Uchański et al., 2021) and is shown in Figure SI-40. First, a synthetical gene was designed (Figure SI-40, ordered at GeneArt, Regensburg) with the first conserved β-sheet of nanobodies, followed by the C-terminal part of HopQ, a linker which fused this C-terminal part to the following N-terminal part and two SapI cloning sites. The gene was cloned into the nanobody standard vector pHEN6 with N-terminal pelB signal sequence and C-terminal His_6-tag. The resulting construct can serve as a backbone for every nanobody to generate the corresponding HopQ megabody (Figure SI-40). For the present work, HiSiaQM nanobody 3 was cloned into the megabody-backbone vector via the two restriction sites for SapI. The resulting construct, megabody 3, was verified via sequencing.

The protocol for expression and purification of the megabody is based on the protocol for nanobodies and the published work of Uchański et al. (Uchański et al., 2019, Uchański et al., 2021) (Chapter 9.1.4.5). Due to a sparse description of the purification procedure in the literature, the following steps are described in detail. The postinduction and supernatant fractions in Figure 7-9 B indicated a strong band at around 55 kDa, corresponding probably to the megabody protein with a molecular weight of around 60 kDa. Because the protein was purified for the first time, the Ni^{2+}-AC was treated with a step wise increase of the imidazole concentration in six elution fractions. The SDS-PAGE of these fractions indicated again a main, intensive protein band at around 55 kDa protein marker (Figure 7-9 B). All fractions contained also small quantities of impurities at around 35 and 25 kDa that had the same elution characteristics on the imidazole gradient as the main protein. The following SEC showed a monodisperse, sharp single peak in the elution profile with a maximum at around 85 mL elution volume (Figure 7-9 C). The analysis of the peak with an SDS-PAGE indicated again a strong band for the protein at around 55 kDa, even though the contaminations from the step before were still present. Due to the different mass of the two bands and the unsuccessful separation during SEC, as well as the same elution characteristics from the Ni^{2+}-AC, the two bands were either fragments of the larger protein or have a strong interaction with this protein. However, for the use of the megabody in cryo-EM experiments as specific binder, the quality of the sample after SEC was sufficient and the protein fractions were combined and concentrated for analysis.

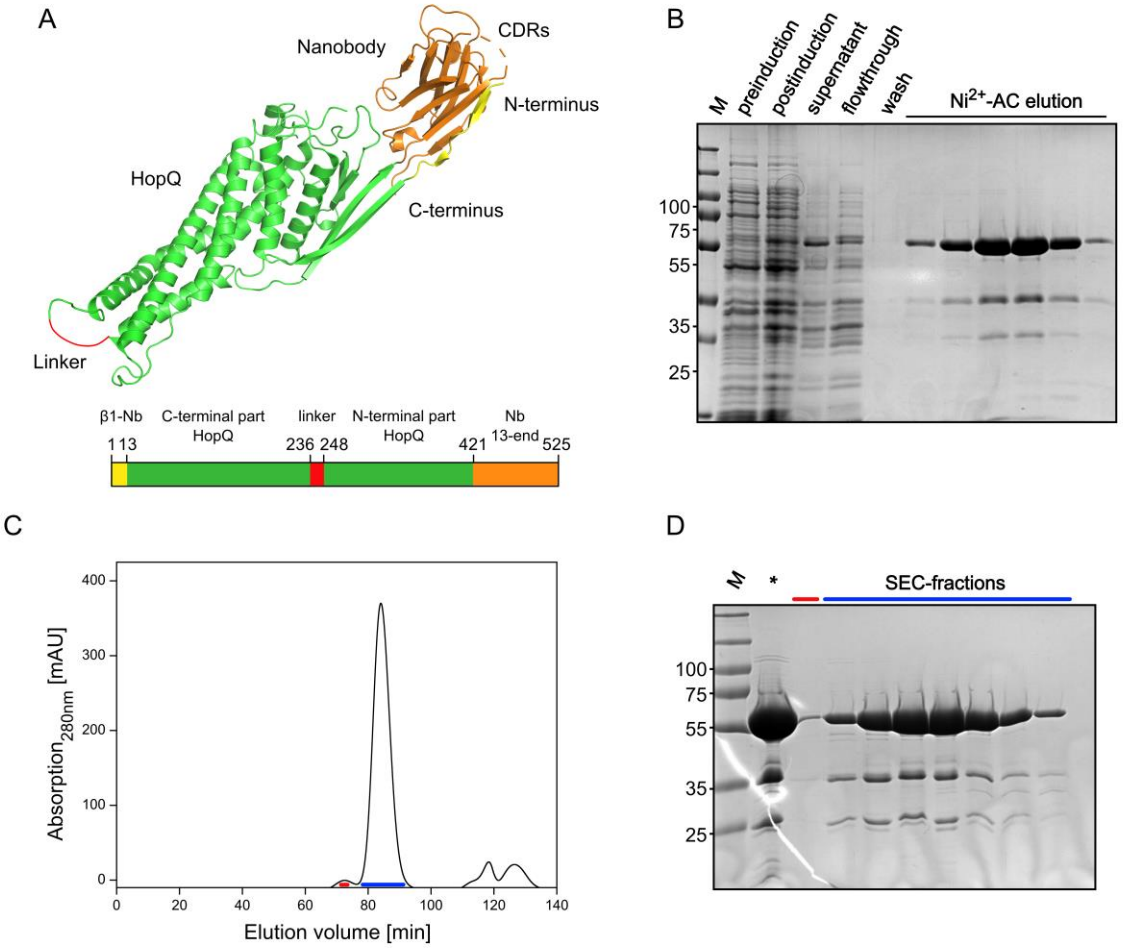


Figure 7-9: Purification results of nanobody-derived megabody 3. A) Protein structure and topology of sequence of nanobody-derived megabody, based on the bacterial protein HopQ. The nanobody is shown in yellow and orange, the HopQ in green and the linker between C- and N-terminus of HopQ in red (PDB-ID: 6QD6 (Uchański et al., 2021)). **B)** SDS-PAGE with samples from expression, lysis and Ni^{2+}-AC. The elution of the Ni^{2+}-AC was performed using a stepwise imidazole gradient, with 15.6, 31.3, 62.5, 125, 250 and 500 mM imidazole. **C)** SEC run with absorption at 280 nm (HiLoad SD 75 16/600, standard buffer) **D)** SDS-PAGE of SEC run fractions from C) (* = loaded sample).

7.5.2 Binding characterization of the megabody

After construction, expression and purification of the nanobody-derived megabody 3, the binding to HiSiaQM was analysed in a similar way as for the nanobodies (Chapter 6.3).

First, the binding was studied with analytical SEC runs. The experiments were accomplished on an HPLC machine, by using the HiSiaQM wildtype antigen solubilized in detergent micelles and in protein-bounded nanodiscs. Due to the troubles with HiSiaQM in SMA-nanodiscs and the uninterpretable elution profiles from previous experiments, this solubilization technique was not used for the binding studies (Chapter 6.5.2). The chromatograms of the analytical SEC runs, with single protein loads and mixing of HiSiaQM and megabody 3, are shown in Figure 7-10 A and B for both solubilization methods. In the single component runs, the HiSiaQM in detergent micelles eluted with the same double peak profile as before and HiSiaQM in protein-bounded nanodiscs in a more defined single peak, with a small shoulder to higher elution volumes (Figure 7-10 A and B, blue chromatograms).

The mixture of megabody and HiSiaQM in both membrane mimetic systems indicated a clear decrease of the absorption of the megabody peak at around 21 min retention time, compared to the single megabody run. Additionally, the absorption intensity of the HiSiaQM peak at earlier elution times was increased for both membrane mimetic systems. In contrast to the analytical SEC runs with nanobodies, both chromatograms showed a slightly shift of the first peak to earlier elution volumes, compared to the HiSiaQM single run (Figure 7-10 A and B). In case of DDM-stabilized HiSiaQM, the shift was observed with around 0.035 column volumes (CV) (1.2 min, 84 µL) and for the MSP-nanodisc sample, a difference of 0.018 CVs (0.6 min, 42 µL) was detected. The fact that such a shift was just visible for experiments with the megabody is presumably caused by the larger size and molecular weight of around 60 kDa for the megabody compared to 15 kDa of the nanobodies. Overall, the changed intensities of the peaks and the elution time shift of the first peak suggested the complex formation of megabody 3 and HiSiaQM in both membrane mimetic systems. The modification of the nanobody 3 to the megabody 3 was therefore presumably still able to bind the HiSiaQM antigen.

Beside the binding characterization with analytical SEC, the binding of megabody 3 was also analysed with SPR, similar as for the nanobodies. Due to the results from corresponding nanobody 3, which indicated only an interaction with HiSiaQM K273-biotin and no binding to HiSiaQM K235-biotin, only the first mutant was used as antigen (Chapter 6.3.3). The sensorgram and the calculated fit are shown in Figure 7-10 C and revealed a defined binding of the megabody to HiSiaQM K273-biotin. The binding affinity was detected to be decreased from 0.64 nM for the nanobody 3 to 15.1 nM for the megabody 3 by a factor of 24.

The association and dissociation of the megabody indicated a similar behaviour between the mega- and nanobody, both processes were just slightly faster for the megabody than for the nanobody (Figure 7-10 C and Table 6-1).

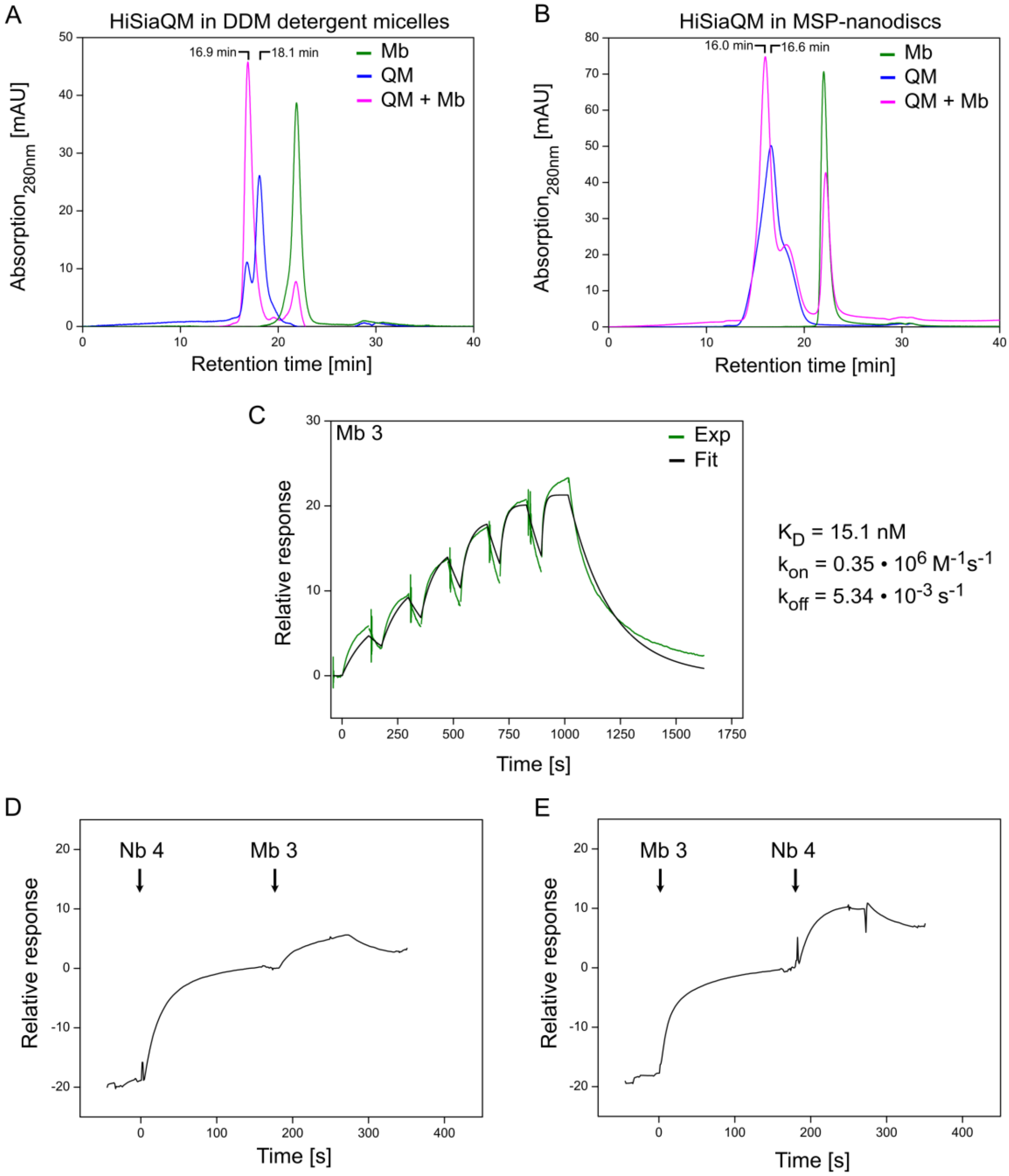


Figure 7-10: Binding studies with HiSiaQM and megabody 3. A) SEC runs with detected absorbance of the elution at 280 nm by using an SD 200 3.2/300 column on an Agilent HPLC system (buffer: 50 mM KH_2PO_4 (pH 7.8), 200 mM NaCl, 0.035% DDM). The experiment with single megabody 3 is shown in green, the single run with HiSiaQM (wildtype, DDM) in blue and the experiment with mixed sample of HiSiaQM and megabody 3 is shown in magenta. The mixed sample were prepared with the same amount of each protein as used in the single component run. Both proteins were mixed in a molar ratio of 1:1. **B)** Same as A) but with MSP-

nanodiscs stabilized HiSiaQM. The same buffer was used but without addition of DDM. **C)** Sensorgram of the SPR single cycle kinetics experiment for megabody 3 with immobilized HiSiaQM (DDM) 273-biotin. The relative response of the surface is shown as green line, the fit to the experimental data as black line. The binding parameters from the fit are mentioned on the left. **D)** SPR sensorgram of competitive binding experiment with nanobody 4 used in the first injection and as saturated component and megabody 3 used in the second injection. **E)** Same as D) but with megabody 3 in the first and nanobody 4 in the second injection.

In analogy to the competitive experiments with nanobodies, a similar experiment was performed with megabody 3 to investigate if the increasing size of one binding partner leads to the blocking of a binding region (Chapter 6.3.4). Due to the fact, that megabody 3 has the same CDRs as nanobody 3, the information from the previous experiments with identified binding regions were used for the current approach. The competitive binding of nanobody 4 and megabody 3 was confirmed by using both components once for the first and once for the second injection (Figure 7-10 D and E). As clearly visible from the SPR sensorgrams, both combinations showed a successful binding after the second injection, similar to the competitive experiment with the two corresponding nanobodies. Consequently, the increase of the size from a nano- to a megabody had presumably no influence on the binding of nanobodies to the other binding region. In summary, the designed and produced megabody 3 contained similar binding properties as the corresponding nanobody 3. The larger protein was also characterized as strong binder to the membrane domains of TRAP transporter HiSiaQM and was used in the following chapter for cryo-EM studies.

7.6 Megabody-bound HiSiaQM for EM structure determination

Similar as for the HiSiaQM-nanobody cryo-EM experiments, the megabody was used as HiSiaQM specific binder that should represent a large outer membrane domain on the nanodiscs. By this, the identification of HiSiaQM-nanodiscs should be improved and larger structural features can be probably used as alignment marker for the classification of the particles.

7.6.1 Sample preparation for EM

The HiSiaQM membrane domains were supplemented with a two times excess of megabody 3. The subsequent chromatogram of the SEC run for EM sample preparation is shown in Figure 7-11 and had 4 peaks. The first peak was caused by an experimental error, the fraction collector of the FPLC system was accidentally switched on at this point. The double peak at around 1.3 mL elution volume revealed the characteristic that was standardly observed for HiSiaQM in MSP-nanodiscs (Chapter 4.3). The last peak was assigned to excess, unbound megabody in the sample. As for previous experiments, the fractions from the first maximum in the double peak were used for preparation of the cryo-EM sample, the experimental parameters for vitrification are listed in Table 9-11.

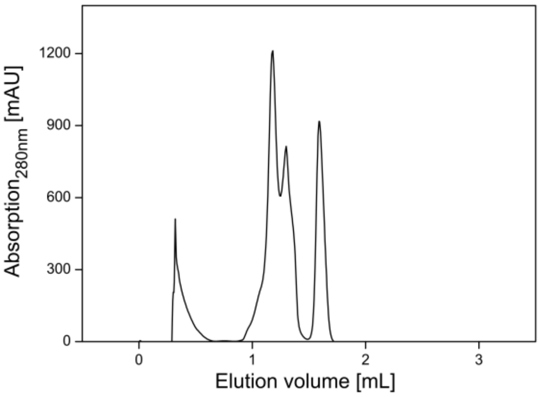


Figure 7-11: SEC of HiSiaQM with megabody 3. SEC run of HiSiaQM in MSP-nanodiscs after incubation with megabody 3 in 2 times molar excess. A SD 200 3.2/300 column on an Äkta micro system was used (buffer: 20 mM Tris (pH 7.5), 100 mM NaCl).

7.6.2 Cryo-EM of HiSiaQM with megabody

The cryo-EM images for HiSiaQM in combination with megabody 3 were recorded on a Titan Krios with a K3 camera at the IGBMC in Strasbourg (FR) (Table 9-11). Because a first test-processing of the data during recording indicated a detectable density on the side of the nanodiscs, a large data set with 7263 electron micrographs was collected. An example micrograph is presented in Figure 7-12 A. As before, nanodisc single particles were visible in both orientations, in side- and top-views. However, on many micrographs, the particle concentration was very high and regions with very close or overlapping particles were detected.

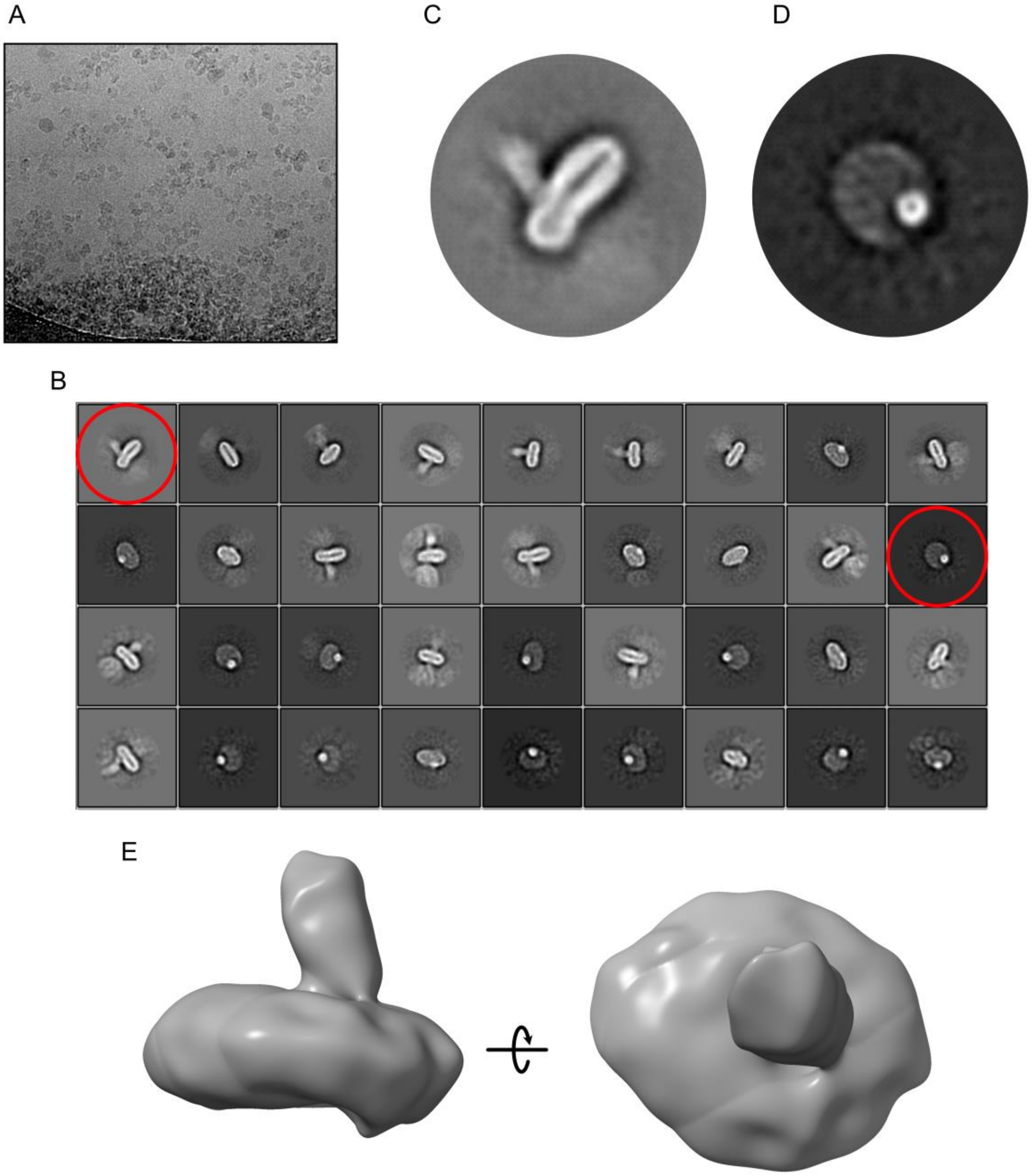


Figure 7-12: Cryo-EM results of HiSiaQM with megabody 3. A) Original micrograph from vitrified HiSiaQM in MSP-nanodiscs with addition of megabody 3. **B)** 2D classification of single nanodisc particles with RELION (Scheres, 2012). The top- and side-view orientations were initially processed separately and merged into one data set for the presented final 2D classification. **C)** Selected 2D class with particles in side-view orientation from B). **D)** Selected 2D class with particles in top-view orientation from B). **E)** 3D model from selected 2D class in B) in two orientations, side- and top-view. The nanodiscs, as well as an attached large fragment are clearly visible in the reconstructed model with around 15 Å resolution.

In contrast to previous processing procedures, in this experiment the two preferable orientations of the nanodiscs were initially handled separately for efficient picking of the particles and an easier processing during 2D classification. For both pathways, with top- and side-view orientation, more than 1 million particles were identified and processed in at least four rounds of 2D classification. After separation of obviously empty nanodiscs or unusable junk-particles, the two sets of particles were merged and assigned in one combined 2D classification (Figure 7-12 B). The finally obtained 2D classes in the side-view orientations revealed a clear identifiable density on the side of the nanodiscs (Figure 7-12 B and C). The size of the density was clearly larger than for the nanobody 2D class in Figure 7-8 C and showed higher resolved structural features. Additionally, this presumably megabody bound nanodiscs revealed a higher density in the nanodiscs at the binding position of the megabody, likely caused by the HiSiaQM protein. For the first time, this cryo-EM 2D classes presumably revealed a view on the membrane domains of TRAP transporters at low resolution. A comparison with current protein models is discussed in Chapter 7.7.

In the processed 2D classes, also top-view orientations were observed for the first time with a significant change in the detected density compared with similar orientations from previous experiments. Many particles with a small, high density spot in the nanodiscs region were identified (Figure 7-12 B and D). Since the nanodiscs and megabody-HiSiaQM complex had a strongly differing thickness for observation in the top-view orientation, the contrast between these components is higher than for the side-view particles. For some top-view classes, low resolution structural features were presumably detectable in the predicted megabody-HiSiaQM density. As for the side-view particles, further comparisons are presented in the discussion (Chapter 7.7). One interesting observation in the top-view classes was a preferred position of the HiSiaQM-megabody complex on the outer edge of the nanodiscs, in direct contact with the MSP1 belt (Figure 7-12 D). Such an observation was not described before and might be caused by an interaction between the HiSiaQM and the MSP1, especially in the hydrophobic region of both proteins.

Overall, around 50 000 particles with a clearly observable bound megabody were received from the micrographs and assigned to the 2D classes in Figure 7-12 B. The selected particles were used for a reconstitution of a 3D model (Figure 7-12 E). This model indicated a clear density on the side of the nanodiscs, that is caused by the bound megabody. However, a further processing towards a protein model with higher resolution was not feasible until today with the standard procedure. Even if the large megabody allowed the unambiguous identification of HiSiaQM-nanodiscs, the particle alignment was presumably mainly focused on the nanodiscs and by this, lowers the resolution of the protein inside the nanodiscs.

7.7 Discussion and conclusion

In the chapter above, several experiments for determination of the protein structure of the TRAP transporter membrane domains HiSiaQM were described. Even if important steps and intermediate goals were reached and presented, the structure could not yet be described at an atomic scale. However, important intermediate steps were reached with both methods, which pave the way for a determination of the structure in further experiments.

First, the protein structure was tried to solve with X-ray crystallography via protein crystallization. For many different conditions and prepared approaches, crystals were obtained but for the most crystals, only low-resolution diffraction images could be collected. The use of HiSiaQM specific nanobodies for co-crystallization improved the protein crystals significantly. Presumably due to larger contact areas between the HiSiaQM-nanobody complex, diffraction resolutions of up to 5 Å were obtained. Especially the result from the diffraction experiment with crystal of condition *MemGold* H11 are promising for further investigations. In following experiments, this crystallization condition should be optimized and alternative techniques for the reconstruction of the phases should be tested.

To receive crystals of HiSiaQM with a higher resolution, several variations should be tested in following experiments. One important optimization will be the use of smaller detergents, such as OG or DM (n-octyl-β-D-glucoside, n-decyl-β-D-maltoside) (Newby et al., 2009). The solubilization of HiSiaQM with the latter detergent was already confirmed in a SEC run in previous experiments (Figure SI-41). Both alternative detergents form smaller micelles around the membrane protein which might increase the intermolecular interaction regions and by this, presumably promote the formation of crystals with higher resolution (Stetsenko and Guskov, 2017). Also, the constructed megabody can be used for co-crystallization to further increase the contact areas. After construction of another megabody with a nanobody from the other binding region on HiSiaQM and a confirmed simultaneous binding of both megabodies, the solvent accessible regions might be immensely enlarged through the attachment of two megabodies. On the other hand, another known megabody with a molecular weight of 100 kDa, based on the glucosidase YgjK, can be constructed for HiSiaQM nanobodies and used for co-crystallization (Uchański et al., 2021). If also for this construct the simultaneous binding of two different megabodies to HiSiaQM is possible, the soluble protein proportion of the complex will be around 75%. Even if the crystallization is often a trial and error process, these variations in combination with the experience and intermediate goals can be used for further crystallization approaches of TRAP transporter membrane domains HiSiaQM.

In the second part of this chapter, it was tried to determine the protein structure of HiSiaQM with cryo-EM. Overall, the sample preparation and detection of single nanodisc particles were successful for several experiments. The processing of the first cryo-EM experiments with HiSiaQM in nanodiscs failed due to unsuccessful separation of full and empty nanodiscs and alignment of the single particles. To solve these problems, HiSiaQM specific nano- and megabodies were used for an easier identification of particles in subsequent cryo-EM experiments. This allowed for the first time the detection of nanodiscs with HiSiaQM embedded in the lipid bilayer and to distinguish between filled and empty nanodiscs. A rough alignment of a protein model for HiSiaQM with a nanobody on the corresponding 2D class is shown in Figure 7-13 A. The HiSiaQM-nanobody model was constructed with the postulated HiSiaQM model from Ovchinnikov et al. (Ovchinnikov et al., 2014) and with a nanobody structure, tentatively positioned to one side of the model membrane protein. By adjustment of the HiSiaQM size with the nanodisc thickness as reference, the protein model and especially the nanobody fits very nicely to the experimental density (Figure 7-13 A and B). For the top-view orientation, the models were used in the same scale as adjusted for the side-view orientation. The turned model with direction of view through the translocation pathway fits also nicely to the detected bright spot in the nanodiscs. A similar result was obtained for the overlaying of a HiSiaQM-megabody model, constructed and aligned in a similar way as the HiSiaQM-nanobody model (Figure 7-13 C and D).

However, even if several 2D classes with obvious HiSiaQM-megabody complexes in different orientations were observed, the data need to be further processed to obtain a 3D structural model at high-resolution. Especially the nanodisc-assembly that was initially used to facilitate data collecting and processing, hinders the detection of HiSiaQM in the lipid bilayer. One possibility to solve this alignment problem is the use of masks during 2D classification and a 3D multi body refinement. Doing this, the focus of the particles is tried to set on the HiSiaQM-megabody complex and the surrounded nanodisc should not be considered. Even if such a procedure is time-consuming and could not be processed during the present work, this might be a promising starting point for subsequent processing attempts.

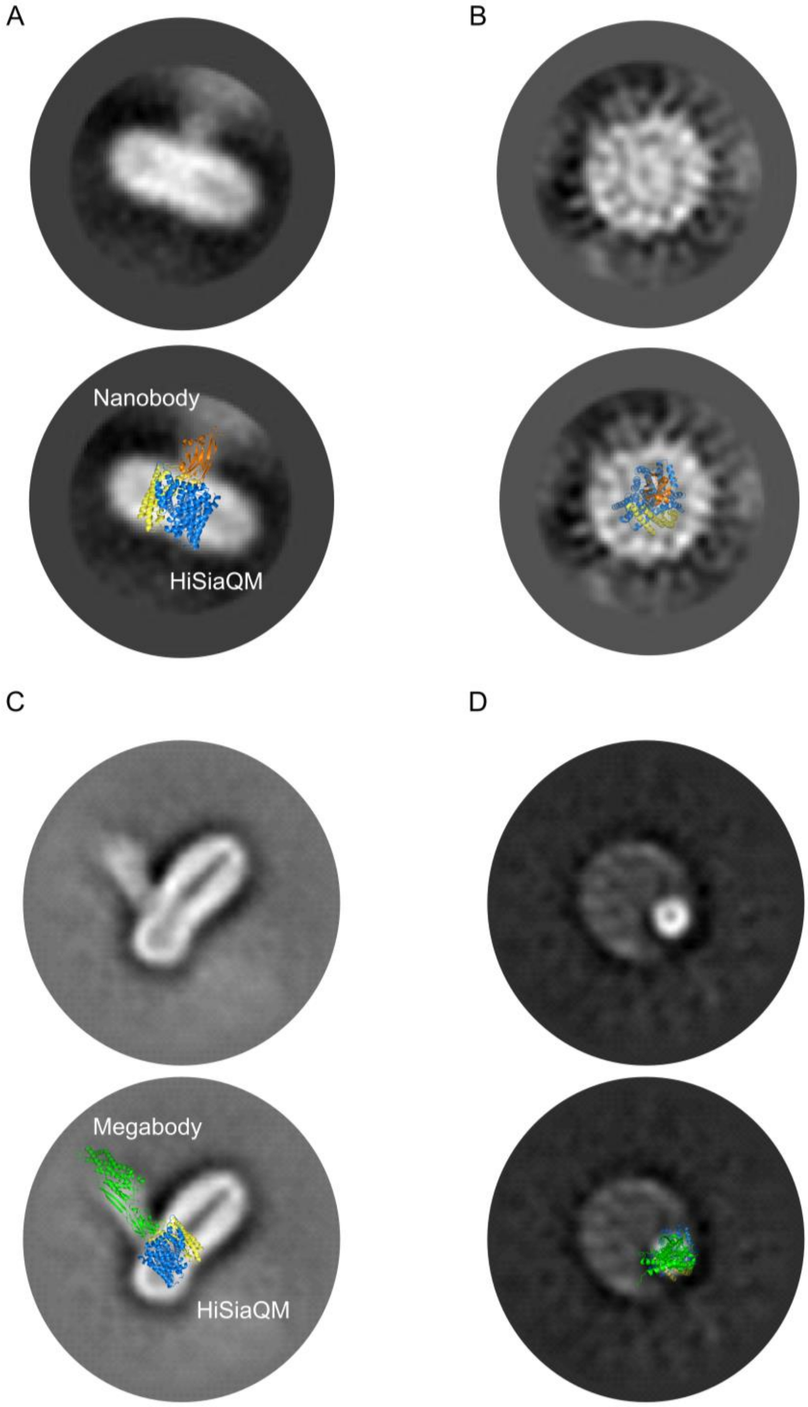


Figure 7-13: Model of HiSiaQM with specific binders in cryo-EM 2D classes. A) Top: 2D class from Figure 7-8. Bottom: Same 2D class with manual alignment of structural model of HiSiaQM (blue and yellow) and nanobody (orange) ((Ovchinnikov et al., 2014) and PDB-ID: 5U64 (Koch et al., 2017)). The thickness of the nanodisc was used as reference for the size of the membrane protein. **B)** Top: 2D class from Figure 7-8. Bottom: Same 2D class with alignment of same protein model as in A). The scale of the model was adapted from A) and inverted by 90°. **C)** Top: 2D class from Figure 7-12. Bottom: Same 2D class with manual alignment of structural model of HiSiaQM (blue and yellow, same as A) and megabody (green, PDB-ID: 6QD6 (Uchański et al., 2021)). The thickness of the nanodisc was used as reference for the size of the protein model **D)** Top: 2D class from Figure 7-12. Bottom: Same 2D class with manual alignment of structural model as in C). The scale of the model was adapted from C) and inverted by 90°. Importantly, the 2D classes in A-B) and C-D) are from other data sets with different scales.

Additional to the processing optimizations, several experimental parameters and conditions can be optimized for further cryo-EM measurements. As mentioned before, the quality of the detected particles on the original micrographs was rather low. A repetition of sample preparation for a higher quality cryo-EM sample might increase the outcoming protein model. Furthermore, a large number of empty nanodiscs without bound megabody was picked and if not successfully separated from filled nanodiscs, decrease the quality and resolution of a 3D protein structure. In further experiments, affinity pull-downs might be useful to separate the empty nanodiscs and detect only HiSiaQM particles in the cryo-EM images. As for the crystallization experiments, the simultaneous use of a second megabody, maybe with the alternative YgjK protein construct, can facilitate the particle processing (Uchański et al., 2021). Additional, HiSiaQM with two bound megabodies might be presumably large enough to study the structure in detergent micelles. A reconstitution into the nanodiscs will be therefore not necessary anymore and the HiSiaQM protein might be directly detectable outside a lipid bilayer.

8 Summary and conclusion

In the first results chapter of this work, two important and widely used methods for integrative structural biology, PELDOR and smFRET spectroscopy, were cross-validated in a comprehensive experimental comparison. For this, three SBP model systems were used and analysed with both methods in several distance measurements during functional, conformational changes. Overall, both methods were in good agreement but for several measurements, misleading results were detected in form of inconsistency to the crystal structures. Diverse reasons, such as protein-label interactions or artifacts caused by the addition of cryo-protectants were found to be responsible for these deviations. Moreover, this comparison revealed that a combination of both methods is a promising approach to mutually confirm the measured distances.

The following chapters were focused on functional and structural studies of TRAP transporters. For the substrate-binding protein VcSiaP, its structure in a substrate-bound state was described for the first time. It is only the second sialic acid TRAP SBP that was structurally classified in different conformations. A mutational study was used to gain insights into the substrate-binding process of this protein with five new mutant protein structures and thermodynamic characterisation. A complex water network between the substrate and protein was identified as an affinity determinant, similar as previously described for the homologous protein HiSiaP. Furthermore, smFRET was used to observe the dynamics of the binding process and the results supported the postulated substrate-induced, “Venus-fly-trap” conformational change of VcSiaP.

The investigation of the membrane domains from TRAP transporters was the focus of the main part of this work. Still today, the work with membrane proteins for analytical studies is challenging, due to the special requirements of the hydrophobic proteins on their environment. A reliable and efficient purification routine via detergent solubilization was established, which resulted in high-quality samples of the membrane protein. In addition, recently developed new methods for membrane protein reconstitution, protein- and polymer-bounded nanodiscs, were established for TRAP transporter membrane domains. These methods allowed the application of subsequent analytical studies on TRAP transporters in native-like lipid bilayers. Several studies with the polymer-bounded nanodiscs revealed surprising and unknown differences compared to the other membrane mimetic systems. The TRAP transporters in these assemblies were incompatible with the nitroxide spin label for PELDOR spectroscopy, analytical SEC on an HPLC device or supplementation with nanobodies. The reasons for these complications might be manifold and remained elusive during the present work. Overall, the studies showed that a simple replacement of classical detergents with the polymers is not straight forward and that the compatibility and behaviour for undescribed bioanalytical experiments have to be studied in detail.

Nevertheless, the TRAP transporter in the other two membrane mimetic systems, detergent micelles and protein-bounded nanodiscs, were used in further functional studies, which benefited from the established protocols for purification. For PELDOR spectroscopy studies, several cysteine-mutants were successfully constructed and purified, and a useful strategy for specific spin labelling with satisfying labelling efficiencies for the TRAP transporter membrane domains HiSiaQM was achieved. The PELDOR spectroscopy results indicated a high conformational flexibility of the membrane domains in detergent micelles, which resulted in unspecific and broad distance distributions. Further, the addition of transporter specific components, such as the SBP, revealed no interactions to the TRAP transporter membrane domains. Particularly the membrane domains environment was postulated to be very important for the conformational state and suggestions for further functional studies regarding this influence were presented.

For further functional and structural studies, specific nanobodies for TRAP transporter membrane domains HiSiaQM were generated and characterized as strong binders with nano- to picomolar affinities. Through combination of binding and functional studies, the nanobody binding allowed the experimental validation of the current theoretical HiSiaQM structure models. Afterwards, one nanobody was engineered to a megabody with similar binding characteristics and both kinds of TRAP transporter specific binders were used for structure determination experiments. In both used methods, X-ray crystallography and cryo-EM, the specific binders increased the quality of the results significantly. Especially the last cryo-EM experiments with a megabody allowed the first ever identification of the TRAP transporter membrane proteins at low-resolution. These studies pave the way with important intermediate results for an imminent full structural description of TRAP transporter membrane proteins.

Another important result was obtained during studies on the influence of nanobody binding on the function of TRAP transporter membrane domains. Several nanobodies revealed an inhibiting effect on the substrate translocation in *in vivo* studies. To date, this represents the first observation of a specific inhibition of TRAP transporters and constitutes a promising fundament for the development of compounds with antibiotic properties. The introduction chapter of this thesis explained the growing danger of worldwide antibiotic resistance, including *H. influenzae*, which was also assigned from the WHO to the list of bacteria with urgently need of new antibiotics. Therefore, if the inhibition of the HiSiaPQM TRAP transporter with nanobodies or nanobody-derived small compounds is confirmed for *H. influenzae*, the results of this work represent a highly important discovery for potentially new therapeutic strategies against the pathogen *H. influenzae*.

9 Methods

9.1 Synthetical and preliminary methods

This chapter describes all methods and work routines that were used in this thesis to produce the target proteins, either as native proteins or with desired modification. The expression, extraction and purification of TRAP transporter membrane domains were optimised during this work and the optimized protocols are presented. Differences in the methods, which result from the change of laboratories during the experimental work, are marked as IPTC (Institute for Physical and Theoretical Chemistry) or ISB (Institute of Structural Biology).

9.1.1 Molecular genetics and cloning

9.1.1.1 Polymerase Chain Reaction (PCR) with QuickChange mutagenesis

The variant of PCR QuickChange mutagenesis described by Liu and Naismith (Liu and Naismith, 2008) was used to mutate proteins, especially with site-directed mutagenesis for labelling experiments or functional studies of proteins. The modification of interest was introduced into the protein encoding gene by using a pair of primers with a defined change in the base sequence in a polymerase chain reaction. The standard PCR QuickChange reaction mix is shown in Table 9-1 and the temperature-time protocol in Table 9-2.

Table 9-1: Reaction mix for PCR QuickChange mutagenesis.

Component	Stock concentration	Volume / Quantity
DNA-template	variable	Variable / 50-100 ng
Pfu-polymerase	5 units/µL	0.5 µL / 2.5 units
Buffer	10x	5 µL
dNTPs	10 mM	2 µL / 0.4 mM
Primer – forward	100 µM	0.5 µL / 1 µM
Primer – reverse	100 µM	0.5 µL / 1 µM
Water		Fill up to 50 µL

Table 9-2: Protocol for PCR QuickChange mutagenesis.

Number of cycles	Temperature [°C]	Time [s]	Process
1	95	300	Initial denaturation
3	95	60	Denaturation
	$T_{\text{m, overlap}}-5$	60	Annealing
	72	900	Elongation
15	95	60	Denaturation
	$T_{\text{m, non-overlap}}-5$	60	Annealing
	72	900	Elongation
2	95	60	Denaturation
	43	60	Annealing
	72	900	Elongation

9.1.1.2 Polymerase Chain Reaction (PCR)

Standard polymerase chain reaction was used to amplify DNA-fragments, such as genes or part of a gene, from a plasmid to insert it into another expression vector. Recognition sites for restriction enzymes were introduced at the ends of the linear DNA strands via non-binding regions of the primers. The reaction mix for a PCR is shown in Table 9-3 and the temperature-time protocol in Table 9-4. Table 10-15 lists the primer and the corresponding annealing temperatures.

Table 9-3: Reaction mix for standard PCR.

Component	Stock concentration	Volume / Quantity
DNA-template	variable	Variable / 50-100 ng
OptiTaq-polymerase	5 units / µL	0.2 µL / 1 units
Buffer	10x	5 µL
dNTPs	10 mM	0.4 µL / 80 µM
Primer – forward	100 µM	0.2 µL / 0.4 µM
Primer – reverse	100 µM	0.2 µL / 0.4 µM
Water		Fill up to 20 µL

Table 9-4: Protocol for standard PCR.

Number of cycles	Temperature [°C]	Time [s]	Process
1	94	120	Initial denaturation
30	94	15	Denaturation
	T_m -5	15	Annealing
	72	45	Elongation
1	72	480	Final elongation

9.1.1.3 Agarose gel-electrophoresis

Agarose gel-electrophoresis was used to identify DNA fragments after a PCR QuickChange mutagenesis or to purify DNA fragments from a PCR or restriction digest. The agarose gel was prepared by adding 1% agarose in TAE buffer and heated up until the agarose was fully solved. Around 25-50 mL of this liquid mixture, depending on the gel chamber, was immediately supplemented with 0.1% ethidium bromide (3,8-Diamino-5-ethyl-6-phenyl-phenanthridinium bromide) solution (IPTC) or the same amount of PEQgreen dye (ISB) for detection of DNA bands in the gel via UV light absorption. The mixture was transferred into a gel chamber with a comb to create sample wells and cooled down. After this, the chamber was put into an electrophoresis cell, filled with TAE buffer. The DNA samples for the gel were mixed with 6x DNA sample buffer and pipetted into sample wells, together with 2 µL of a DNA standard marker. To identify DNA after a QuickChange mutagenesis, 5 µL of PCR sample was supplemented with the DNA sample buffer and loaded onto the gel, while for a separation of different DNA fragments the whole samples were loaded onto the gel. For electrophoresis, the gel was run with constant voltage of 100 V for 30 min. The migration of DNA fragments was analysed and documented via UV light absorption of the DNA-dye by a gel imaging system or a bench-top UV light table. If necessary, the detected DNA fragments were extracted from the agarose gel after electrophoresis. Therefore, the required DNA band was cut out of the gel and the DNA was extracted from the agarose gel with the *ExtractMe DNA clean-up & gel-out* kit, according to the manufactures' manual. The concentration of the final DNA elution was measured with a NanoDrop 2000 spectrophotometer at wavelength of 260/280 nm and stored at -20 °C until further processing.

9.1.1.4 Restriction digestions of DNA

Restriction enzymes were primarily used in this work for digestion of unmutated DNA-templates from a PCR QuickChange mutagenesis. The restriction endonuclease DpnI recognizes the methylated palindromic sequence G^{m6}ATC and cuts a DNA plasmid into DNA fragments. Due to the fact, that only the template DNA of a PCR is methylated through previous transformation into *E. coli* cells, the DNA template is digested and the PCR product remains intact. For digestion, 45 µL of the DNA sample were mixed with 5 µL Tango buffer (IPTC) or 5 µL CutSmart buffer (ISB) and 0.5 µL DpnI enzyme. The solution was incubated for at least one hour at 37 °C. Subsequently, the restriction enzyme as well as the other PCR ingredients were removed with PCR purification kit (ExtractMe DNA clean-up & gel-out (ISB) or QIAquick PCR purification (IPTC)), following the manufactures' manual.

Further, endonucleases were used for DNA cleavage at desired, specific positions, for example to cut off an insert from a DNA plasmid or to prepare a linear DNA fragment for ligation. With the used endonucleases, overhangs of the DNA are produced which can be used for ligation (Chapter 9.1.1.5). All enzymes that were used for specific DNA cleavage are used in combination with the corresponding buffer. The time, temperature and pipetting schematic for incubation of DNA with the enzymes was done according to the manufactures' protocol (New England Biolabs). If necessary, enzymes were heat-deactivated after restriction, also following the manufactures' protocol. After restriction, linear DNA fragments were purified using a PCR purification kit (ExtractMe DNA clean-up & gel-out) and cleaved DNA plasmids were separated, analysed and purified by agarose gel electrophoresis.

Analytical restriction was performed to check the successful ligation of an insert into a vector, after transformation and mini prepping of the DNA. For this, 3.6 µL of DNA sample was supplemented with 0.2 µL of each corresponding enzyme and 1 µL CutSmart buffer and incubated for 30 min at 37 °C. The result was analysed through separation of the digestion fragments by agarose gel electrophoresis.

9.1.1.5 Ligation of DNA

DNA fragments were ligated by using a T4 DNA ligase. This ligase from bacteriophage T4 catalyses the fusion of two sticky or blunt ends of DNA through formation of a phosphodiester bond. Due to the use of DNA with sticky ends, the ligation of an insert into a new vector can be controlled. For a typical reaction, 50 ng of digested vector DNA was mixed with digested DNA insert in a molar ratio of 1:3. The mixture was supplemented with 2 µL T4 DNA ligase buffer and 0.2 µL T4 DNA ligase and filled up with water to a total volume of 20 µL. Ligation was performed at 16 °C overnight for at least 14 hours and on the next day, the ligase was inactivated by incubation at 65 °C for 10 min.

9.1.1.6 Transformation of DNA plasmids into bacteria

Transformation of DNA plasmids into bacterial cells was performed for separation of single plasmids after a PCR QuickChange mutagenesis or ligation and for recombinant expression of proteins in *E. coli* cells. Independently from the used cell lines, all plasmids were transformed into self-produced, chemically competent *E. coli* cells via heat shock. Previously prepared 50 µL aliquots of bacteria, that were stored at -80 °C, were thawed on ice for 30 min before supplementation with 3 µL of prechilled DNA plasmid. After 20 min incubation on ice, the tubes were heat-shocked in a water bath at 42 °C for 45 s and immediately put back on ice for 2 min. Then, 1 mL of LB medium was added to the bacteria and incubated for 1 h at 37 °C under constant shaking.

For transformation of DNA into expression cells, 100 µL of the cell suspension were directly streaked out on a LB-agar plate with corresponding antibiotic, depending on the resistance gene of the transformed plasmid. For transformation of a PCR product, the bacteria were centrifuged for 5 min at 4000 rpm and 900 µL of the supernatant were discarded. The bacteria pellet was resuspended in the remaining liquid and streaked out on a LB-agar plate with corresponding antibiotics. The agar plates were incubated overnight for at least 12 h at 37 °C to observe growth of single colonies. Afterwards, the agar plates were stored at 4 °C for several weeks of for further use, single colonies were transferred with a pipette tip into LB medium supplemented with appropriate antibiotics and let grown overnight at 37 °C. For a long-time storage of bacteria after successful transformation and as helpful starting point for later expression approaches, glycerol stocks were prepared from each transformation for protein expression or growth experiments. For this, 800 µL of the overnight culture was mixed with 200 µL sterile glycerol in a 2 mL Eppendorf tube and stored at -80 °C.

9.1.1.7 Isolation of plasmid DNA and sequencing

After modification of DNA, for example PCR mutagenesis of or ligation of a gene into a new vector, the resulting DNA was verified via sequencing. Therefore, overnight cultures of single colonies on an agar plate after a transformation were prepared by using 10 mL LB medium (IPTC) or 4 mL LB medium (ISB) with appropriate antibiotics. The plasmid DNA was isolated using the GeneJET Plasmid kit (IPTC) or ExtratMe Plasmid mini kit (ISB) and the manufactures’ protocol. After isolation of the DNA, the concentration was determined with a NanoDrop spectrophotometer by measuring the absorption of the sample at 260/280 nm. Aliquots of the samples were prepared and used for external sequencing of the plasmid DNA by Microsynth SEQLAB (IPTC) or Eurofins Genomics (ISB). The provided results were analysed with the Geneious software program regarding the sequence of the bases and the quality of the sequence signal.

9.1.2 General methods for protein biochemistry

9.1.2.1 SDS-PAGE

The standard method SDS-PAGE (sodium dodecyl sulfate polyacrylamide gelelectrophoresis) was used to analyse proteins by separation of their size. More explicit, SDS-PAGE was used in this work to follow and verify the expression of protein, monitor the purity of a sample during different purification steps or to use it for functional analysis such as the formation of protein complexes in solution.

In this work, only the method of discontinuous electrophoresis with two different gel regions (stacking and separating gel) was used. The composition of each gel is shown in Table 9-5. First, the ingredients of the separating gel were combined, poured between two glass plates and topped with 70% ethanol. After polymerization, the ethanol was discarded and the components of the stacking gel were pipetted into one tube, mixed and pipetted on the separating gel between the two glass plates. For sample application, a comb with 10 or 15 sample wells was put into the stacking gel and removed after polymerization. The gels were directly used for electrophoresis or stored in a wet paper at 4 °C for several days. For SDS-PAGE, the gel is clamped into a gel holder and sink into SDS running buffer in an electrophoresis chamber.

To analyse the proteins in an SDS-PAGE, the samples were treated with SDS sample buffer (1:4 ratio of SDS sample buffer to protein sample) previous to the electrophoresis. Doing this, the structure of the proteins was denatured with β-mercaptoethanol and detergent SDS from the sample buffer and the intrinsic charges were masked by the anionic detergent SDS. Apart from SDS-PAGE samples containing the membrane domains of TRAP transporters, all other protein samples were heated up to 95 °C for 5 min after addition of the SDS sample buffer to aid the denaturation process. For electrophoresis, the prepared protein samples were transferred into the sample wells in the stacking gel. A protein marker was pipetted into a free sample well to estimate the size of protein bands and the electrophoresis was processed for around 50 min at 175 V (IPTC) or 35 min at 200 V (ISB), until the blue bromophenol dye reached the bottom of the gel.

Afterwards, the gel was taken out of the glass plates and stained with gentle shaking in Coomassie staining solution for 10 min. The unspecific colouring of the gels background was removed by boiling the gel in 700 mL water for 5-10 min (IPTC) or incubating the gel with warm SDS destain solution under gentle shaking for around 10-20 min (ISB). Finally, if coloured protein bands were clearly visible, the result was analysed and documented with an imaging system.

Table 9-5: Composition of gels for SDS-PAGE.

Gel	Ingredients	Volume
Separating gel 15%	Acrylamide	2.63 mL
	Separating gel buffer	1.46 mL
	H_2O	1.16 mL

	TEMED	1.76 µL
	APS (10%)	58.88 µL
Separating gel 12%	Acrylamide (30%)	2.1 mL
	Separating gel buffer	1.46 mL
	H_2O	1.69 mL
	TEMED	1.76 µL
	APS (10%)	58.88 µL
Stacking gel 5%	Acrylamide (30%)	225 µL
	Stacking gel buffer	175 µL
	H_2O	900 µL
	TEMED	1.3 µL
	APS (10%)	13.25 µL

9.1.2.2 Western Blot

Western Blots were used for detection of specific proteins after SDS-PAGE. This method was mainly used to analyse different steps of the expression, extraction and purification of the TRAP transporter membrane domains, due to low quantity of the membrane protein. All western blots in this work were performed with a primary antibody against a His$_6$ affinity tag and a goat anti-mouse HRP conjugated second antibody.

Previous to western blotting, an SDS-PAGE with protein samples that should be detected with a western blot was performed as described in Chapter 9.1.2.1. In addition to the samples of interest, 100 ng of a positive control such as TEV protease or the P-domain of TRAP transporters were loaded into a sample well. These proteins had a known running behaviour on SDS-PAGE and western blot and the required epitope for the primary antibody. At the IPTC, a PVDF (polyvinylidene fluoride) membrane with the size of the polyacrylamide gel was activated with methanol for a few seconds and then washed in 10 mL water for 5 min. Afterwards, all components for a western blot, the PVDF membrane, the unstained SDS-PAGE gel and four pieces of blotting paper were equilibrated in transfer buffer for 15 min. Alternatively, at the ISB, the components (nitrocellulose membrane) were directly incubated with the corresponding transferbuffer. After that step, the components were assembled in a way that the SDS-PAGE gel and the blotting membrane are stacked onto each other and were covered with two blotting papers on both sides. The construct was put into a semi-dry blotting cell with the blotting membrane facing the cathode and the SDS-PAGE gel facing the anode. Air bubbles between the different layers were removed by hand. The separated proteins from the SDS-PAGE gel were blotted onto the membrane at constant amperage of

300 mA with a maximum voltage of 25 V for 50 min (IPTC) or at constant amperage of 160 mA for 45 min (ISB).

The blotting membrane was transferred into a blocking solution with 5% milk powder in TBST (IPTC) or PBST (ISB) for one hour at room temperature with gentle shaking to avoid unspecific binding of antibodies. Before incubation of the primary antibody, the blotting membrane was washed three times for 10 min with 10 mL TBST (IPTC) or PBST (ISB). The first antibody was diluted 1:1000 in 5 mL TBST (IPTC) or PBST (ISB) and added to the membrane for incubation of 90 min at room temperature or preferable incubation overnight at 4 °C, both with gentle shaking. As next step, the membrane was again washed in the same way as before and the secondary antibody in a 1:5000 dilution in 10 mL TBST (IPTC) or PBST (ISB) was added onto the membrane. After incubation of one hour under gentle shaking at room temperature, the membrane was finally washed in the same way as before and the final washing solution was discarded. The secondary antibody was detected in two different ways: At the IPTC, the blot was developed with 10 mL TMB solution under gentle shaking, until blue spots were clearly visible (10-30 min). The staining was stopped by rinsing the membrane with water, the membrane was carefully dried with a heat gun and documented with an imaging system. At ISB, the blotting membrane was developed with a western blotting detection reagent and documented with an imaging system.

9.1.2.3 Concentration of proteins

Protein solutions were concentrated for different applications by using a centrifugation filter with specific size of filter pores. Depending on the volume of protein solution and final desired volume, different dimensions of centrifugation filters were used (2 mL, 6 mL or 20 mL). Previously to the sample, the filter was centrifuged for 1 min with small amounts of the corresponding sample buffer for equilibration. Afterwards, the protein solution was loaded onto the filter and centrifuged corresponding to the manufactures’ instructions. In general, a MWCO (molecular weight cut off) was chosen that the target protein remained in the upper solution (Table 9-6). For example, the P-domains of TRAP transporter were concentrated with a MWCO of 30 kDa and the QM-domains with 100 kDa, due to their membrane mimetic system around the membrane protein. The concentration process was checked continuously and the protein solution above the filter was regularly mixed by pipetting.

The final concentration of protein solutions was determined via a NanoDrop spectrophotometer by measuring the absorption of the solution at a wavelength of 280 nm. The concentration was automatically determined from the NanoDrop software by using the protein-specific extinction coefficient (Table 9-6). The concentrated protein was directly used for further purification or experiments or flash-frozen and stored at -80 °C.

9.1.2.4 Chromatography of proteins

Different kinds of chromatography methods were used in this work for purification of proteins. This chapter gives an overview about the two often used chromatography methods Ni^{2+}-affinity-chromatography (AC) and size-exclusion chromatography (SEC) and the general handling of the columns. The use of the chromatography methods for different proteins as well as specific characteristics are described in Chapter 9.1.4

Ni^{2+}-affinity chromatography

Ni^{2+}-AC was used for purification of His-tagged proteins by specific binding between the histidines and the Ni-cations on the stationary phase of the column. In this work, Ni^{2+}-AC was exclusively performed using gravity flow columns. For a standard approach, about 3 mL of Ni-NTA agarose beads suspension in ethanol-water solution was transferred into the column and the storage buffer was discarded after running through the column. Before applying to the protein, the Ni-NTA beads were equilibrated with at least 15 times column volume (CV) of sample buffer which correspond to the protein. The equilibration was stopped when 2 times CV of the buffer remains and the Ni-NTA beads were resuspended in the buffer and pipetted to the protein solution. After incubation at variable time and temperature, the suspension was transferred back onto the column and the flowthrough was running through the column. If necessary, all steps were performed under constant cooling with a special cooling jacket or in a cold room. Some solutions, such as the flowthrough or wash buffers without any supplements, were carefully pulled through the column with a syringe. After elution of the protein, the Ni-NTA agarose beads were washed with 10 CV of 2 M imidazole solution (pH 8) to remove unspecific binder. Afterwards, the beads were washed with 10 CV water and 10 CV 20% ethanol (p.a.) and stored in the ethanol solution at 4 °C. After 5-10 times of use, the beads were deep cleaned with a standard protocol by stripping of the Ni-cations from the beads with EDTA (ethylenediaminetetraacetic acid), different washing steps and reloaded of the beads with a $NiSO_4$ solution.

Size-exclusion chromatography

A SEC was used as standard method for the purification of proteins and was often performed as last step of a purification process. In addition to preparative application of SEC, it was also used as an analytical method to characterize proteins regarding their hydrodynamical radius or to study protein complexes. The SEC columns consist of a stationary phase with different pore sizes, which produce a specific diffusion volume for each molecule which is applied onto the column, depending on the hydrodynamic radius. Therefore, a mixture of proteins or oligomers can efficiently be separated by their specific size in the solution. In this work, SEC columns with cross-linked agarose material were used with a specific pore size, depending on the target proteins. All preparative SEC runs were performed on a Äkta FPLC systems or for analytical runs on an Agilent HPLC system. The SEC runs were performed at room temperature (IPTC) or at 4 °C (ISB) after the manufactures' manuals, especially regarding the maximal flow rates and column pressures. The columns were equilibrated with 1.2 CV of the corresponding protein buffer before the protein was loaded onto the column and after use, the buffer was replaced by 1.2 CV 20% ethanol (p.a) (IPTC) or water (ISB) for storage of the column.

9.1.3 Generation of nanobodies

Nanobodies were produced for structural and functional investigations of the sialic acid TRAP transporter membrane domains from *Haemophilus influenzae*. The first steps, involving the immunization of alpacas, generation of the nanobody library and phage display were carried out by the nanobody core facility of the University of Bonn, as partially described in Schmidt et al. (Schmidt et al., 2016a). The following chapter presents only the main steps and procedures that are useful for understanding of the corresponding results chapter.

9.1.3.1 Immunization and nanobody library

As a template for the nanobody production, 700 µL of HiSiaQM wildtype protein with a total concentration of 3.512 mg/mL was prepared. In the last concentration step, the remaining concentrated protein solution was diluted around a factor 40 with HEPES standard buffer (50 mM HEPES, pH 7.8, 200 mM NaCl, 0.035% DDM) and the solution was concentrated again to 700 µL. The alpaca *Picasso* was used for immunization with HiSiaQM for generation of heavy chain-only antibodies. At the moment of immunization, there was no evidence that the protein sample had any negative effect on the animal. The injections into the alpaca were repeated several times for a few weeks. After immunization, blood of the animal was collected and processed, to receive nanobody encoding DNA sequences. The nanobody sequences were amplified, ligated into a vector and transformed into electrocompetent *E. coli* TG1 cells to create a nanobody library. The cells were streaked out onto an agar plate and all colonies were harvest and stored as glycerol stock.

9.1.3.2 Phage display and panning

Phage display was used to identify HiSiaQM binding nanobodies from the created nanobody library. Therefore, a culture with the nanobody library in *E. coli* TG1 cells were infected with helper phages to produce phages with the nanobodies as fusion proteins on their surface. To select HiSiaQM binders, two single cysteine mutants K235C and K273C were biotinylated (Chapter 9.1.6.3) and immobilized via the label on a streptavidin-coated surface. The phages were added and incubated with the HiSiaQM coated surface for panning and non-binders were removed in a washing step. The remained phages were eluted and used for a new infection of *E. coli* cells. Due to the small number of colonies observed on an agar plate, only one round of panning was performed with the two HiSiaQM constructs.

9.1.3.3 ELISA and identification

The infected and observed *E. coli* single colonies were picked and grown in a 96-well plate for an ELISA. The cultures were induced with IPTG for expression of nanobodies. For the ELISA of each of the two panning approaches (K235C and K273C), five antigens were coated onto different ELISA plates. Standard plates were used for non-biotinylated proteins and streptavidin-coated plates for biotinylated proteins. The antigens HiSiaQM wildtype (DDM), HiSiaQM 273C-biotin (DDM), HiSiaQM 235C-biotin (DDM), HiSiaQM 275C-biotin (SMA) and GST were used for each of the two panning approaches. Each ELISA-plate was incubated with supernatants from the nanobody expressions and the binding of nanobodies were detected with antibodies against the nanobody fused HA-tag, that led to a colouring of TMB solution. For wells with positive binding hits, the corresponding nanobody-DNA plasmid was extracted and sequenced for identification of nanobodies. The nanobody sequence of each hit was compared to all other nanobody sequences in a phylogenetic tree and sequences with a difference identity of around 3% were clustered in one nanobody group for further steps. From each group, one nanobody was produced in a small-scale expression and purification as described in Chapter 9.1.4.4. The purified nanobodies were used for a second ELISA, again with the same five antigens and immobilization strategies as used for the first ELISA.

9.1.4 Protein expression and purification

This chapter is subdivided into the different proteins that were expressed and purified for this work. As already mentioned above, optimized and finally used protocols are summarized in this chapter and important developments, variations or tests are presented in the corresponding result chapters. Table 9-6 summarizes physical properties of the different wildtype proteins and Table 9-7 the molecular characteristics that were used for standard procedure, such as tags that were fused to the genes.

Table 9-6: Physical properties of purified proteins without mutations. The parameters were calculated with ProtParam (Gasteiger et al., 2005). The modifications are specified in Table 9-7.

Protein	UniProt	modification	M_W [Da]	ε_{280} [$M^{-1}cm^{-1}$]	pI
HiSiaP	P44542	-	34165	23840	5.80
		+ His_6-TEV	37463	29800	5.88
HiSiaQM	P44543	-	67593	66600	9.18
		+ His_{10}-TEV	69872	68090	9.09
VcSiaP	Q9KR64	-	33755	36900	5.17
		+ His_6-TEV	37040	42860	5.39

VcSiaQM	Q9KR65	-	64941	50880	7.64
	Q9KR66	+ His_{14}-TEV	69323	56840	6.89
MSP1		+ His_6	24738	23950	6.02
Nanobody 1		+ LPETG-His_6	15429	20190	5.72
Nanobody 2		+ LPETG-His_6	13459	24075	9.07
Nanobody 3		+ LPETG-His_6	15154	30035	8.00
Nanobody 4		+ LPETG-His_6	14961	21680	8.99
Nanobody 5		+ LPETG-His_6	15545	27180	7.88
Nanobody 6		+ LPETG-His_6	15171	30035	8.94
Nanobody 7		+ LPETG-His_6	14944	28545	8.00
Nanobody 8		+ LPETG-His_6	15001	20190	8.44
Nanobody 9		+ LPETG-His_6	15313	28545	8.01
Megabody 3		+ LPETG-His_6	60296	64790	8.43

Table 9-7: Molecular modifications of proteins for purification.

Protein	Vector	Tags
VcSiaP	pBADHisTEV	His_6 $^{\text{N-term}}$ TEV cleavage site $^{\text{N-term}}$
HiSiaP	pBADHisTEV	His_6 $^{\text{N-term}}$ TEV cleavage site $^{\text{N-term}}$
VcSiaQM	pBADHisTEV	His_{14} $^{\text{N-term}}$ TEV cleavage site $^{\text{N-term}}$
HiSiaQM	pBADHisTEV	His_{10} $^{\text{N-term}}$ TEV cleavage site $^{\text{N-term}}$
MSP1	pET28a	His_6 $^{\text{N-term}}$
Nanobodies 1-9	pHEN6	pelB signal peptide $^{\text{N-term}}$ His_6 $^{\text{C-term}}$ LPETG-sortase motif $^{\text{C-term}}$
Megabody 3	pHEN6	pelB signal peptide $^{\text{N-term}}$ His_6 $^{\text{C-term}}$ LPETG-sortase motif $^{\text{C-term}}$

9.1.4.1 SBPs of TRAP transporter

The substrate-binding proteins of TRAP transporters, VcSiaP and HiSiaP, and several mutants of these proteins, were expressed and purified with a standard protocol (Glaenzer et al., 2017). Importantly, all SiaP proteins were expressed in M9-minimal medium. This is necessary to avoid contamination of the protein with the substrate sialic acid from nutrient medium such as LB medium, which leads to a significant proportion of around 20% of substrate-bound SiaP (personal communication with Janin Glänzer, University of Bonn). This chapter describes the standard protocol, variations such as labelling procedures during the purification are mentioned in following chapters.

For protein production, the pBADHisTEV vectors with corresponding *siap* genes were transformed into C43 *E. coli* cells (Chapter 9.1.1.6). An overnight culture of 10 mL LB medium with ampicillin (100 µg/mL) was prepared single colony cells from an agar-plate or from a previous prepared frozen glycerol stock. After incubation overnight at 37 °C and 180 rpm, 5 mL of the preculture was used to inoculate a second preculture in a baffled flask with 100 mL LB medium and ampicillin (100 µg/mL). The culture was again incubated at same conditions as before for 8 hours. To remove traces of LB medium, the second preculture was washed with M9-minimal medium before preparing a main culture. Therefore, the culture was centrifuged for 15 min at 3000 rcf in two falcon tubes, the supernatants were discarded and each pellet was resuspended in 50 mL M9-minimal medium without any supplements. The centrifugation and resuspension steps were repeated two more times by using fresh M9-minimal medium for each step.

For protein expression, each litre of M9-minimal medium, supplemented with ampicillin (100 mg/L), sterile $MgSO_4$ (2 mM) and $CaCl_2$ (0.1 mM), was inoculated with 10 mL of the washed preculture cells. To produce sufficient quantities of proteins, 3-6 L of main cultures were prepared at the same time. The main cultures were grown for 14-16 h overnight at 37 °C and 180 rpm. The optical density of the cultures at 600 nm (OD_{600}) were measured with a NanoDrop spectrophotometer and the protein expression were induced at an optical density of 0.6-1 by adding 500 mg/L L(+)-arabinose. The induced cultures were incubated for 3 h at 37 °C and 180 rpm. To harvest the *E. coli* cells after expression, the cultures were transferred into 1 L centrifugation bottles and centrifugated for 20 min at 4000 rcf (rotor: Beckman JLA-8.1000). The supernatants were discarded and the cell pellets were scratched and pooled into a 50 mL falcon tube. After weighing the tube, the pellet was directly used for purification of the protein or stored at -80 °C until further use.

As first step for purification, the cells were lysed to extract the protein from the cytoplasm into a buffer solution. Therefore, the cell pellet was resuspended under gentle shaking in 5 times excess of standard P-buffer (w/v). The resuspended cells were lysed twice under pressure of 30 kpsi with a cell disruptor (IPTC) or with a sonicator at 40% amplitude for 5 min with 10 s pulses (5 s breaks) (ISB). During sonification of the cells, the solution was constantly chilled on ice to prevent strong heating. The lysate was subsequently centrifuged for 20 min with 75000 rcf at 4 °C to remove insoluble cell components from the solution (rotor: Beckmann JA-25.50). The supernatant was mixed for Ni^{2+}-AC with Ni-NTA beads, equilibrated with standard P-buffer. For binding of the His-tagged protein to the column, the suspension was incubated with gentle rotation for 1 h at room temperature. After that, the solution was transferred to a bench-top column, the flow-through was discarded and the Ni-NTA-beads were washed with 100 mL standard P-buffer. Finally, the protein was eluted from the Ni-NTA-beads with 15 mL of elution P-buffer.

Following this step, the protein was purified with an anion-exchange chromatography by using an ENrich Q 10/100 or a HiTrap Q XL column on a Äkta chromatography system. The columns were equilibrated with 5 CV of No-salt buffer, 5 CV High-salt buffer and again 5 CV No-salt buffer. The elution fraction from Ni^{2+}-AC was diluted with No-salt buffer to 50-100 mL to reduce the salt concentration. The protein sample was loaded onto the column and eluted with a salt gradient by constant increase of the High-salt buffer percentage. The elution was monitored by measuring the absorbance at 280 nm and collected by an automatic fraction collector. Proteins, for which the affinity tag should be cleaved off after Ni^{2+}-AC, were incubated with TEV protease (1:100, TEV protease to SiaP) overnight at 4 °C or at room temperature for 2 hours. An inverse Ni^{2+}-AC was performed in the similar way as before, but with the target, cleaved protein in the flowthrough fraction. Due to the efficient separation of further contaminants with the inverse Ni^{2+}-AC, no anion-exchange chromatography was performed for cleaved proteins.

As final step, the protein was purified with SEC by using a HiLoad SD 75 16/600 or a HiLoad SD 200 16/600 column, equilibrated with standard P-buffer. If necessary, the protein solution was concentrated to 5 mL before application onto the column and the elution was again monitored with absorbance at 280 nm and collected in fractions. To monitor the decrease of impurities in the protein sample as well as the identification of SiaP protein in the elution fraction, SDS-PAGEs were performed for each purification step. The protein containing fractions were combined, the solution was concentrated to around 20 mg/mL and directly used for investigations or snap frozen in 100 µL aliquots and stored at -80 °C.

9.1.4.2 Membrane domains of TRAP transporter

The expression and purification processes were based on the work of Mulligan et al. (Mulligan et al., 2009, Mulligan et al., 2012), but many steps were modified and ultimately differed considerably from the previous published routines.

For recombinant protein expression of the membrane domains HiSiaQM and VcSiaQM, the corresponding pBADHisTEV vectors with *siaqm* genes were transformed into MC1061 *E. coli* cells (Chapter 9.1.1.6). Around 120 mL LB medium of a preculture with ampicillin (100 µg/mL) was inoculated with a single colony from an agar-plate or a glycerol stock and incubated overnight at 37 °C and 180 rpm shaking. On the next day, a main culture was prepared with 1 L LB medium, supplemented with standard concentration of ampicillin and 10 mL of preculture. For useful amounts of proteins, 10 L LB medium were used at the same time. The cells were grown at 37 °C with 180 rpm for 2-3 h to an OD_{600} of 1 and the protein expression was induced by adding 50 mg/L L(+)-arabinose. After expression for 2 h at 37 °C and 180 rpm, the cells were harvested with centrifugation in 1 L bottles for 20 min at 4000 rcf (rotor: Beckman JLA-8.1000). The cell pellets were separated from the supernatant and combined into a 50 mL falcon tube. After weighing of the tube, the pellet was directly used for purification of the protein or stored at –80 °C until further processing.

For purification, a cell pellet was resuspended in 5 times excess of standard QM-buffer (w/v) under gentle shaking. The suspension was lysed with either a cell disruptor system and pressure (IPTC) or with a sonicator (ISB). In the first method, the cell suspension was lysed in three repetitions with a pressure of 30 kpsi. The lysate after each run was collected in a glass bottle on ice to guarantee constant cooling. For sonification, the cell suspension was divided into fractions with maximum 35 mL and each fraction was sonicated 5 times for 1 min at 40% amplitude. After every minute, the suspension was mixed by inverting the tube. As for the cell disruptor, the tubes were constantly cooled during sonification. In the next step, the lysate was centrifuged with an ultracentrifuge at 300 000 rcf (rotor: Beckman 50.2 Ti) for 1 h at 10 °C. The supernatants were discarded and the membrane pellets were pooled and resuspended with a homogenizer in 50 mL extraction buffer. For extraction of the membrane proteins from the native membrane, the solution was incubated overnight under gentle shaking at 4 °C for the detergent extraction or at 37 °C for extraction with a polymer. On the next day, the solution was centrifuged in a second ultracentrifugation run (300 000 rcf, 1h, 10 °C) to separate the extracted membrane proteins from insoluble cell components.

As first chromatography step, the supernatants from ultracentrifugation were incubated with Ni-NTA agarose beads, equilibrated in standard QM-buffer, for 2h at 4 °C under gentle rotation. The suspension was loaded onto a gravity bench-top column at 4 °C or in an ice rack and the flowthrough was discarded. The Ni-NTA beads were washed with 100 mL QM-washing buffer for detergent or polymer extracted protein and the protein was eluted with 15 mL QM-elution buffer. For further purification, the elution was concentrated to 500 µL and loaded onto an equilibrated SEC column SD 200 10/300 or SD increase 200 10/300. The elution with QM-standard buffer was monitored at 280 nm and fractions were collected in an automatic fraction collector with 500 µL fraction sizes. Purification process and identification of QM-domains containing fractions were inspected with SDS-PAGE. SEC-fractions with QM-domains were pooled and concentrated to around 20 mg/mL and directly used for further experiments or snap-frozen and stored at -80 °C.

For SMA-stabilized QM-domains, the determination of the protein concentration was hard to determine, due to a strong absorbance of the polymer at 280 nm (Oluwole et al., 2017a). For this work, the concentration was calculated with an extinction coefficient of 69.09 and a molecular weight of 69782 Da.

9.1.4.3 Membrane scaffold proteins

Expression and purification of membrane scaffold protein 1 (MSP1) was performed following published protocols with slight customizations (Bayburt et al., 2002, Ritchie et al., 2009).

For expression, the pET28a vector with *msp1* gene was transformed into BL21 *E. coli* cells and a preculture in 150 mL LB medium, supplemented with kanamycin (50 µg/mL), was inoculated with a single colony from an agar-plate or a glycerol stock. The culture was incubated for 4-6 h at 37 °C and 180 rpm until a cell density with an OD_{600} of 0.6-0.8 was reached. The culture was then stored overnight at 4 °C without shaking. A main culture was prepared with 1 L TB medium, supplemented with kanamycin (50 µg/mL) and 20 mL of the prechilled preculture. For sufficient protein yields, 6 L of main culture were prepared in baffled flasks and incubated for around 4 h at 37 °C and 180 rpm, until an OD_{600} of 2.5-3.0 was reached. Protein expression was induced by adding 1 mM IPTG to each culture and the cells were incubated for 1 h at 37 °C and then for 3.5 h at 28 °C, the whole time shaking with 180 rpm. Afterwards, the cells were harvested by centrifugation in 1 L centrifugation bottles at 4000 rcf for 15 min (rotor: Beckman JLA-8.1000) and the cell pellets were scratched and transferred into two 50 mL falcon tube. The tubes were stored at -80 °C or directly used for purification.

For lysis of the cells, the pellets were resuspended in 2.5 excess of lysis buffer (w/v) under gentle shaking. Afterwards, the cells were lysed with two runs in a cell disruptor with 30 kpsi or sonicated for 5 min at 40% amplitude with 10 s on and 5 s off pulses. For separation of insoluble cell components, the lysate was centrifuged for 30 min with 30 000 rcf at 10 °C (rotor: Beckman JA-25.50). The supernatant was mixed with Ni-NTA agarose beads, equilibrated with basic buffer and incubated for 1 h at room temperature under gentle rotation. After transfer of the Ni-beads to a bench-top gravity column, the flowthrough was discarded and the column was washed with 12 mL of wash buffer 1 and 2, the standard buffer and finally with wash buffer 3. Finally, the protein was eluted from the Ni-NTA agarose beads with 3 times 10 mL of elution buffer. After inspection of the Ni^{2+}-AC fractions with SDS-PAGE, the protein containing elution fractions were combined and concentrated to a total volume of 5 mL. This protein solution was further purified with SEC by using a HiLoad SD 75 16/600, equilibrated with standard buffer. Elution was monitored with absorption at 280 nm and an SDS-PAGE was run to identify MSP-containing fractions. These fractions were combined, concentrated to around 10 mg/mL and flash-frozen in liquid nitrogen for storage at -80 °C.

9.1.4.4 Nanobodies

Nanobodies were expressed and purified in two different ways, a small-scale or a standard approach with higher yields. Both protocols were provided by the group of Florian Schmidt (University of Bonn) and the small-scale approach was performed at the laboratories of the nanobody core facility from the University of Bonn. For both different methods, a pHEN6 vector with *nanobody* genes were transformed into *E. coli* WK6 cells (Chapter 9.1.1.6).

Small-scale approach

For the main culture, 50 mL TB medium were supplemented with 100 µg/ml ampicillin and inoculated with 1.25 mL of an overnight preculture in LB medium. The culture was grown at 37 °C at 180 rpm until an OD_{600} of 0.6 and induced with final concentration of 1 mM IPTG. The temperature was decreased to 30 °C and the cultures were incubated overnight for protein expression. Afterwards, the cultures were harvested by centrifugation for 10 min at 3000 rcf and 4 °C in 50 mL falcon tubes and the pellets were directly used for periplasmic extraction.

To extract the nanobodies from the periplasm, the cell pellets were resuspended in 750 µL TES buffer and incubated under gentle shaking for 1 h at 4 °C before addition of 3.5 mL 0.25x TES buffer and then incubated at the same conditions overnight. Insoluble cell fragments were removed by centrifugation in 15 mL falcon tubes for 10 min at 3000 rcf and 4 °C. For purification, a Ni-NTA spin column was used and every centrifugation step was performed at 700 rcf for 2 min. The Ni-NTA columns were separated from the storage solution by centrifugation and the beads were equilibrated with 500 µL standard buffer. Subsequently, the beads were resuspended into the extracted protein supernatant and incubated for 30 min at 4 °C under gentle shaking for binding. Afterwards, the suspension was centrifuged to remove non-binding proteins and the beads were washed by centrifugation for three times with 500 µL wash buffer. After transfer of the columns into new collection tubes, the proteins were eluted with three times 200 µL elution buffer. For further purification and buffer change, the eluted nanobodies were pipetted onto an equilibrated PD MiniTrap G-25 column and eluted after the manufactures' instruction with bench top gravity flow. The resulting protein samples were collected, the concentrations were determined with a NanoDrop spectrophotometer and the samples were stored at -80 °C.

High-yield approach

For higher-yield and higher level of purity of nanobodies for further analytical studies, a preculture with 25 mL LB medium, 100 µg/mL ampicillin and a single colony from an agar plate or a glycerol stock was grown overnight at 37 °C and 180 rpm. On the next day, a main culture of 1 L TB medium with appropriate antibiotics was inoculate with 25 mL of the preculture and grown at 37 °C and 180 rpm, until an OD_{600} of 0.6. Protein expression was induced with addition of IPTG to a total concentration of 1 mM. The cultures were incubated overnight at 30 °C at 180 rpm and harvested on the next morning by centrifugation in 1 L centrifugation bottles at 5000 rpm and 4 °C for 20 min (rotor: Beckman JLA-8.1000). The pellets were collected in 50 mL falcon tubes and stored at -80 °C until further processing.

The purification of nanobodies was started with extraction of the periplasmic fraction. The cell pellets were resuspended in 15 mL TES buffer for 1 h under gentle shaking at 4 °C. The solutions were fractionated into two 50 mL falcon tubes and each fraction was diluted with 35 mL 0.25x TES buffer. For extraction, the suspension was incubated overnight at 4 °C under gentle shaking. On the next day, the solutions were centrifuged at 8000 rcf for 40 min (rotor: Beckman JA-25.50) and the supernatant was filtered through a 0.45 µm sterile filter into a new falcon tube. For affinity chromatography, the Ni-NTA agarose beads were equilibrated with 0.25x TES buffer, transferred to the supernatants and incubated for 1 h with gentle shaking at 4 °C. After pipetting the suspension to a bench-top gravity column, the flowthrough was discarded and the column was washed with 50 mL of Nb-wash buffer. The protein was eluted after incubation of the beads for 30 min at 4 °C with 10 mL elution buffer. For the next step, the elution was concentrated to a total volume of 5 mL and loaded onto a SEC HiLoad SD 75 16/600 on an Äkta system, equilibrated in QM-buffer standard. The absorbance of the elution was monitored at 280 nm and 2 mL fractions of the elution were collected. Finally, an SDS-PAGE was performed to identify fractions with nanobodies and to observe the purification process. Fractions with nanobody were combined, concentrated to around 20 mg/mL and snap-frozen for storage at -80 °C.

9.1.4.5 Megabodies

Megabodies were expressed and purified by partially following the procedure from previous publication (Uchański et al., 2019, Uchański et al., 2021) and the protocol for nanobody expression and purification (Chapter 9.1.4.4). The construction of the megabody gene, consisting of a fusion between HopQ protein and a nanobody, is presented in Chapter 7.5. For expression, the *megabody* gene was cloned into a pHEN6 vector, with fusion to a pelB signal sequence, and transformed into WK6 *E. coli* cells.

For expression, a preculture was prepared with 50 mL LB medium, 100 µg/mL ampicillin and a single colony of an agar plate or a glycerol stock. After incubation overnight at 37 °C and 180 rpm, 2 L of TB medium were used for one main culture. Each litre was supplemented with appropriate antibiotics and 25 mL of the preculture. The cultures were grown at 37 °C and 180 rpm until an OD_{600} of 0.8 and induced with 1 mM IPTG at final concentration in each litre. The protein was expressed overnight at 28 °C under constant shaking of 180 rpm. To harvest the cells, the cultures were transferred into 1 L centrifugation bottles and centrifuged for 20 min with 4000 rpm (rotor: Beckman JLA-8.1000). The supernatant was discarded and the cells were transferred into two 50 mL falcon tubes.

For lysis of the cells by periplasm extraction, the pellets were resuspended in 15 mL TES buffer and incubated for 1 h at 4 °C under gentle shaking. After completely resuspension, 35 mL of 0.25x TES buffer was added to each falcon tube and the extraction of the periplasm was performed with overnight incubation at 4 °C under gentle shaking. On the next day, the solutions were centrifuged for 40 min with 8000 rcf (rotor: Beckman JA-25.50) at 10 °C to remove cell fragments and the supernatant was filtered with a sterile 0.45 µm filter into a suspension of Ni-NTA agarose beads, equilibrated with 0.25x TES buffer. The solution was incubated at 4 °C for 1 h under gentle shaking and then transferred to a bench-top gravity column. The flowthrough was discarded and the beads were wash-ed with 100 mL standard buffer, before the protein was eluted with 15 mL elution buffer. For further purification, the protein was loaded onto an SD 200 16/600 SEC column on an Äkta system, equilibrated with standard buffer. The elution of the SEC-column was monitored with absorbance at 280 nm and collected in 2 mL fractions. After verification of purification process and megabody containing peaks, the corresponding fractions were combined and concentrated to around

30 mg/mL. The final protein solution was snap-frozen and stored at -80 °C until further experiments.

9.1.5 Reconstitution of nanodiscs

This chapter summarizes the developed reconstitution of HiSiaQM membrane domains into standardly used DMPC-MSP1 nanodiscs. Variations and different tests are shown in the result Chapter 4.3 and were optimised by Peer Depping during his master thesis. The procedure of reconstitution is based on previous published work (Bayburt et al., 2002, Ritchie et al., 2009).

As first step for reconstitution, a lipid buffer containing 50 mM DMPC lipid and 100 mM Na-cholate was prepared in QM-standard buffer without detergent. The volume of the buffer depends on the amount of QM-domains that was used for the reconstitution and the final concentration of Na-cholate in the reconstitution mix, which should be higher than 14 mM. The lipid buffer was prepared in a glass tube and put into an ultrasonic water bath, until all components were completely solved. Afterwards, the lipid buffer was supplemented with purified MSP1 and HiSiaQM, in a molar ratio of 1:80:0.2 (MSP1:DMPC:HiSiaQM). The solution was diluted by adding standard QM-buffer to final concentrations of 11 mM DMPC and 22 mM Na-cholate. Preferable, the concentration-range should be between 6 mM/12 mM to 15 mM/30 mM. The reconstitution solution was incubated for 2 h at 26 °C with gentle shaking of 60 rpm. To remove DDM detergent and Na-cholate from the solution, a dialysis tube (5-10 cm, 6000-8000 MWCO) was equilibrated in standard QM-buffer for 5 min and the reconstitution solution was pipetted into the tube. The solution was dialysed in 1000 times excess of standard QM-buffer at 26 °C under gentle mixing and the buffer was refreshed three times after 2 h incubation. The final dialysis step was performed overnight for at least 16 h. To separate the QM-domains in the nanodiscs from other ingredients and educts of the incubation solution and to verify the reconstitution into the nanodiscs, a SEC run was performed after reconstitution. Depending on the application of the reconstitution, an SD 200 10/300, Superose 6 5/150 or an SD 200 3.2/300 was used. The elution was monitored with absorbance of the solution at 280 nm and detected peaks were analysed with an SDS-PAGE. Corresponding fractions were combined, concentrated and snap-frozen for storage at -80 °C or directly used for analytics. The concentration of the nanodiscs were determined with a molecular weight of 119 kDa (HiSiaQM: 67.593 kDa and 2 times MSP1: 2 x 24.608 kDa) and an extinction coefficient of 115380 (HiSiaQM: 67480 and MSP1: 23950).

9.1.6 Labelling of proteins

During this work, many proteins were labelled with several types of label for different applications. Independent from the kind of label or protein, all labels were introduced into the proteins via surface-accessible cysteines. To prevent intermolecular dimerization or oligomerization of proteins by formation of disulfide-bridges between the cysteines, all proteins were treated with reducing agents such as TCEP or DTT (Tris-(2-carboxyethyl)-phosphine and Dithiothreitol) before the labels were added. An overview about the labels that were used for smFRET or PELDOR spectroscopy distance measurements is presented in Figure 2-2.

9.1.6.1 Spin labelling

In this work, the frequently used methanethiosulfonate-functionalized nitroxide spin label MTSSL was used to prepare the proteins for EPR spectroscopy measurements (Berliner et al., 1982). The labels were attached to the proteins via site-directed spin labelling (Altenbach et al., 1989, Hubbell et al., 1996). Here, the spin labels are bind to cysteine residues and thereby allow a specific introduction of labels into the macromolecular structure.

For spin labelling of stable and soluble proteins such as HiSiaP, MalE or SBD2, the proteins were previous purified and stored in buffer with a reducing agent such as TCEP or DTT. Alternatively, the proteins were incubated for at least 30 min with a reducing buffer prior to the following labelling procedure. To remove the reducing agent, a PD-10 desalting column was performed after the manufactures’ instructions or alternatively a SEC run, both with standard buffer for the corresponding protein. The eluted protein fraction was immediately supplemented with MTSSL in a five-time molar excess per cysteine residue and incubated for 60 min at room temperature or at 4 °C with gentle shaking. To remove unbound protein, the label solution was again loaded onto another PD-10 column or SEC column. The protein containing elution fractions were concentrated to around 20 mg/mL. Labelling of the membrane domains of TRAP transporter were more demanding and in the present work, a specific labelling procedure were developed. Individual steps and their impact on the labelling result are presented in detail in Chapter 5.3.

At final step after each site-directed spin labelling approach, the successful labelling of the proteins and the remove of unbound, free label was investigated with cw-EPR spectroscopy (Chapter 9.2.2).

9.1.6.2 Fluorophore labelling

Analogous to spin labelling of proteins, the attachment of fluorophores to proteins were also performed via site-directed labelling of surface-accessible cysteines. All fluorophores were functionalized with a cysteine-reactive maleimide group.

The procedure of fluorophore labelling for a HiSiaP single cysteine mutant with fluorescein was similar to the site-directed spin labelling. The purified protein was loaded onto an equilibrated SD 75 10/300 column to remove the reducing agent. The elution fraction was immediately supplemented with the fluorophore dye in 4 times excess. The solution was incubated overnight at 4 °C under gentle shaking. To remove free label on the next day, the protein solution was loaded onto a PD-10 column and the protein elution fraction was collected.

For FRET measurements, the proteins were labelled by Christian Gebhardt and other group members of Thorben Cordes at the University of Munich (de Boer et al., 2019b, Gouridis et al., 2015, Peter et al., 2020, Peter et al., 2021). The proteins were incubated with 1mM DTT reducing agent and then immobilized on a Ni sepharose 6 fast flow resin (GE healthcare). The fluorophores were dissolved in 1 mL labelling buffer (50 mM Tris-HCl, pH 7.4-8.0, 50 mM KCl, 5% glycerol) and the resin was incubated at 4 °C with a molar ratio of label to protein of 20:1. Double cysteines mutants for FRET measurements were incubated with the two different FRET pair fluorophores at the same time. After 2-4 h labelling, the column was washed with 3 mL labelling buffer 1 and 2 (50 mM Tris-HCl, pH 7.4-8.0, 150 mM KCl, 50% glycerol) to remove excess label. The protein was eluted with elution buffer (50 mM Tris-HCl, pH 7.4-8.0, 150 mM KCl, 5% glycerol, 500 mM imidazole). All fluorophore labelled proteins were finally loaded onto a SEC SD 75 10/300 column, equilibrated in corresponding protein buffer. During elution, the absorbance at 280 nm and specific absorbances of the attached fluorophores were detected to verify the successful labelling procedure. The protein containing elution fractions were combined, concentrated and snap-frozen for storage at -80 °C or directly used for investigations.

9.1.6.3 Biotin labelling

For immobilization of the QM-domains of TRAP transporters on streptavidin-coated surfaces, a biotin-label was attached to the protein. The biotin-label was used as maleimide-functionalized linker label and attached to two single cysteine mutants of HiSiaQM (K235C and K273C). The labelling procedure was adapted from the experience and optimization of site-directed spin labelling of HiSiaQM with MTSSL (Chapter 5.3).

The labelling of HiSiaQM was performed during the purification step with Ni^{2+}-AC (Chapter 9.1.4.2). After the protein was bound to the Ni-NTA agarose beads and washed with 100 mL imidazole-buffer, the beads were washed with 50 mL reducing-buffer to prevent intermolecular disulfide bridges. Afterwards, the reducing buffer was removed from the column with 50 mL standard buffer and the protein was eluted with 15 mL elution buffer, supplemented with 100 times excess of biotin-maleimide label. For high labelling efficiencies, the solution was incubated overnight under gentle shaking at 4 °C. In a final step, the protein was further purified and the unbound excess label was removed by a SEC run on an SD 200 10/300 column with standard buffer.

9.2 Analytical methods

9.2.1 *In vivo* functional TRAP transporter assay

A bacterial growth assay was used to investigate the function of TRAP transporters, regarding the influence of different mutations in the protein sequence or the addition of transporter-specific components. Central component of this assay is a modified *E. coli* cell strain, BW25113 Δ*nanT* (also named SEVY3), provided by the group of Gavin H. Thomas (University of York) (Fischer et al., 2015, Severi et al., 2010). In these *E. coli* cells, the gene *nanT* of the native sialic acid transporter is knocked out. This transporter is responsible for sialic acid uptake and predicted to be the only uptake route for *E. coli* cells (Vimr et al., 2004, Vimr and Troy, 1985). Furthermore, from these studies it is also known that native *E. coli* cells can catabolise exogenous sialic acid with a sialic acid aldolase (nanA) to pyruvate and the amino sugar ManNAc and by this, use sialic acid as carbon source. the bacteria BW25113 Δ*nanT* are not able growth in an environment with sialic acid as only carbon source, due to the missing uptake transporter (Fischer et al., 2015, Severi et al., 2010). However, if a sialic acid transporter is introduced artificial into these cells, the cells can growth by using this alternative uptake route. Therefore, the function of the transporter is directly coupled to the growth of the cells. The system can be used as bacterial *in vivo* growth assay for sialic acid transporter such as the TRAP transporter HiSiaPQM and the influence of different mutations or additives to the transporters function can be determine under native-like conditions. To provide the TRAP transporter system to the modified *E. coli* cells, a pwks30 plasmid with the genes *hisiap* and *hisiaqm* (named pES7, provided by Gavin H. Thomas, University of York) was transformed into the cells. To analyse the cell growth, different kinds of sialic acid-only experimental set-ups were used and the methods protocol is presented in the following paragraphs. In every approach, a negative control with an empty pwks30 vector and a positive control with HiSiaPQM transporter without modifications were prepared for validation of the results.

9.2.1.1 Preparation of cell cultures

As first step, the desired pES7 plasmids with and without modifications were transformed into the assay cells via heat-shock. For addition of another plasmid, cells after pES7 transformation were prepared as competent cell stock and used for co-transformation of a second plasmid with different antibiotic resistance. The competent SEVY3 cells were prepared after the laboratory standard procedure. After transformation of the desired plasmids into the uptake assay cells, a single colony was picked from the agar plate and used to inoculate 5 mL LB medium and incubate overnight at 37 °C and 180 rpm. On the next morning, the cells were harvested by centrifugation for 15 min at 4000 rpm. The cell pellet was resuspended in the same tube in 10 mL M9-minimal medium without any additives and again harvested by an analogous centrifugation step. This washing procedure was repeated one more time and the final pellet was resuspended in 5 mL M9-minimal medium, prechilled on ice. The washed cell culture was directly used for growth assay cultures on M9-minimal agar plates or in liquid medium, as presented in the following chapters.

9.2.1.2 M9-minimal agar plates growth assay

The preparation of M9-minimal medium agar plates is one possibility to analyse the growth of the sialic acid transporter dependent *E. coli* cells, as shown in previous studies by Severi et al. (Severi et al., 2010). To prepare the plates, 250 mL of M9-minimal medium without glycerol were supplemented with 3.75 g agar, 25 µL $CaCl_2$ (1 M) and 500 µL $MgSO_4$ (1 M) and heated in a microwave until all components were dissolved. For two plates, 30 mg of sialic acid and glucose were separately added to each 1 mL M9-minimal medium and supplemented with 30 µL of appropriate antibiotics. The glucose solution was used to prepare positive control plates, on which the cells can use glucose as energy source and every approach should be able to growth. After cooling of the M9-agar to a warm solution, 30 mL were supplemented with 1mL of the antibiotic-sugar mix and plates with sialic acid or glucose were prepared. The plates were cooled down and solidified at room temperature. To test the effect of induction of the HiSiaPQM transporter and to analyse leaky expression, 15 µL of IPTG (1M) was diluted in 70 µL M9-minimal medium and pipetted onto one of the agar plates and streaked out. As final step, the washed cell cultures were added to the agar plate. Therefore, 5 µL from up to six different cell cultures were distributed on the outer edge and streaked out towards the middle point. The agar plates were incubated for 16 h to 40 h at 37 °C and the cell growth were observed regularly.

9.2.1.3 M9-minimal medium growth assay

The second method to detect the sialic acid-dependent cell growth is with liquid M9-minimal medium growth. This method is more time-consuming but allows an accurate tracking of the growth rate, while the agar plate method ended in absolute results. For the liquid medium approach, the required amount of M9-minimal medium was prepared with appropriate antibiotics and 2 mM $MgSO_4$ and 0.1 mM $CaCl_2$. 1 mM IPTC was added for induction of HiSiaPQM on pwks30 vector and the solution was supplemented with 1:1000 of sialic acid. Again, for a positive control the same amount of glucose was added to the medium.

After mixing, the solution was transferred into a 24 well-plate for high throughput approach or into 50 mL baffled flasks for higher volume approach. The medium was inoculated with 1:100 – 1:33 dilution of the washed *E. coli* precultures, while the optimal amount of culture was individually validated for each new growth experiment. The cultures were incubated at 37 °C under shaking (24 well plates: 70 rpm; baffled flasks: 130 rpm) and the cell growth were analysed in regular intervals. The optical density was measured with absorption at 600 nm (OD_{600}) with a NanoDrop spectrophotometer. Again, the optimal intervals to measure the growth rate of the cells must be validated for each new experimental series. Overall, the growth of the cells was observed for around 12-20 h, until no further increase of the cell density was observed.

9.2.2 Continuous wave-EPR spectroscopy

EPR spectroscopy is a method which allows to investigate paramagnetic centres, for example to study their type, dynamics and interactions with their environment (Schiemann, 2012). One field of application that increased since several years is the use of EPR spectroscopy for integrative structural biology to study biomolecules, such as proteins or oligonucleotides (Jeschke, 2018c, Schiemann and Prisner, 2007). For this, many different EPR experiments and set-ups were developed and are still under constant research. In the present work, the nowadays standard set-ups to study proteins with EPR spectroscopy are used. In all experiments, the cysteine-reactive nitroxide spin label MTSSL was used and attached to the proteins via site-directed spin labelling (Chapter 2.1.1 and Chapter 9.1.6.1) (Altenbach et al., 1989, Hubbell et al., 1996). Two different experimental set ups were used, cw-EPR (continuous wavelength) and the pulsed-EPR experiment PELDOR/DEER (pulsed electron electron double resonance) (Chapter 9.2.3).

In a cw-EPR experiment, a paramagnetic sample is illuminated by microwaves of a continuous wavelength. By applying a magnetic field to the sample, the energy levels of the unpaired electrons split up (Zeeman effect), according to the strength of the magnetic field. Through constant increasing of the magnetic field strength, at a certain point the differential energy between the electron spin energy levels is equal to the energy amount of the incoming microwaves. This point is known as resonance condition and can be expressed as $h\nu = \Delta E = g_e \mu B$, where h is the Planck constant, ν the wavelength of the microwaves, E the energy, g_e the g factor of the electron and μ the Bohr magneton (Schiemann, 2012). The left term describes the energy of the microwaves and the right term the energy between the electron spins energy levels. If the strength of magnetic field is at this certain point, the microwaves are absorbed and the electron spins are excited to the higher energy level. This absorption is detected by the spectrometer and represented as absorption line in dependence of the magnetic field. For technical reasons, the first deviation of the absorption is recorded and displayed (Schiemann, 2012).

A resulting cw-EPR spectrum can give important information for site-directed spin labelling experiments of biomolecules (NejatyJahromy and Schubert, 2014). For example, the existence and type of paramagnetic centre can be determined, as well as the stability of the free electron, especially in samples with various conditions such as for in cell measurements (Fleck et al., 2020). Further, the shape of the spectrum can confirm the successful attachment of a nitroxide spin label to a large molecule and detect free, unbound spin label in the sample (NejatyJahromy and Schubert, 2014). Not mentioned in this chapter or used in this work, a cw-EPR spectrum can also be used to determine the distance between two free electrons in a sample, due to the coupling of their spins. However, the measured distance for nitroxide spin labels is limited to a maximum of around 20 Å and therefore, pulsed-EPR experiments (Chapter 9.2.3) are more often used for distance measurements in integrative structural biology (Schiemann, 2012).

9.2.2.1 Preparation of samples

For cw-EPR measurement, usually 15 µL protein solutions with a concentration of 25-100 µM protein were used. If necessary, the solutions were diluted with the corresponding protein buffer. The solutions were soaked up into a 10 µL glass capillary excluding air bubbles in the capillary and one side of the capillary were sealed with superglue. The glue was completely dried at room temperature for at least 20 min and subsequent, the capillary was carefully put into a quartz glass X-band tube for measurement.

9.2.2.2 Measurement and Quantification

The cw-EPR spectra were recorded at a X-band bench-top EMXnano EPR spectrometer from Bruker (US), equipped with a built-in resonator and by using the appropriate Xenon software. All cw-EPR spectra were recorded at room temperature with a microwave power of 2.51 mW. In the beginning, a cw-EPR spectrum with a single scan was recorded to optimise the experimental parameters such as modulation amplitude, video amplifier gain, time constant, conversion time and resolution (points per G). By default, for cw-EPR spectra of the double spin labelled substrate-binding proteins (HiSiaP, MalE and SBD2) and the TRAP transporter membrane proteins, the magnetic field was set to a center field of around 3448 G and X-band microwave frequencies around 9.63 GHz were used. The attenuation of the microwave power was set to 16 dB and the receiver gain was adjusted to receive a high signal sensitivity. For recording of cw-EPR spectra, the sweep width of the magnetic field was adjusted to 150 G with a sweep time of 10.03 s. The time constant was 20.48 ms with 1 G of modulation amplitude. To obtain a good signal-to-noise ratio, 350 single scans were averaged for every cw-EPR measurement. The quantification of the spin labels from the cw-EPR spectra was determined with the Xenon software tool by double integration of the cw-EPR spectrum.

9.2.3 PELDOR spectroscopy

In comparison to cw-EPR spectroscopy, pulsed EPR experiments allow to determine much weaker couplings between magnetic moments of two electron spins or an electron and a nuclear spin (Jeschke, 2018a, Schiemann, 2012). There are several pulsed EPR techniques for various paramagnetic centres, for example double-quantum coherence EPR (DQC), relaxation induced dipolar modulation enhancement (RIDME), single frequency technique for refocusing dipolar couplings (SIFTER) and pulsed electron electron double resonance (PELDOR), also known as DEER (double electron electron resonance) (Abdullin and Schiemann, 2020, Schiemann und Prisner, 2007). Nowadays, the PELDOR experiment is the most frequently used method for distance measurements in biomolecules and only this experiment was used in this work.

For a pulsed EPR experiment, the labelled protein is introduced into a constant magnetic field which lead to a certain orientation of the magnetic moments of the free electrons. With microwave pulses at certain frequency and defined strength or length (90° (π/2) or 180° (π)), the magnetization can be manipulated such that even weak dipolar couplings between remote spin centres can be measured (Schiemann, 2012). For setting up a PELDOR experiment, in the first step a 2-pulse experiment was applied onto the sample and the maximum of echo intensity was determined by recording a field sweep spectrum (Figure 2-1 C). For this, the echo intensity was recorded under constant change of the magnetic field. A typical field sweep spectrum for a double nitroxide labelled biomolecule at Q-band frequency is shown in Figure 2-1 C and reveals different orientations of the label with regard to the magnetic field. In the PELDOR experiment, these subsets can be used as pump and observer spins (Figure 2-1 C and E). After the field sweep experiment, a 2pESEEM spectrum (electron spin echo envelope modulation) was recorded to analyse the relaxation of the spins. As final preliminary experiment, a 3-pulse nutation sequence was performed to define the optimal pump pulse length. Afterwards, the defined parameters were used to set up the PELDOR pulse sequence (Figure 2-1 E, parameters are given in Chapter 9.2.3.2). Figure 2-1 E shows a typically detected spectrum of a PELDOR experiment, a PELDOR time-trace. After background correction of this signal, the distance distribution of the spins can be calculated from this PELDOR time-trace (Figure 2-1 G) (Jeschke, 2018a).

9.2.3.1 Preparation of samples

The following chapter describes the standard protocol for a PELDOR sample preparation, variations are mentioned in correlation with the experimental result. First, the spin labelled protein was prechilled or defrost for 30 min of ice. Depending on the designed experiment, additives were supplemented to the protein solution in desired concentration and incubated for 10 min on ice. The solution was diluted to a protein concentration of 15-25 µM in 80 µL total volume with prechilled PELDOR buffer and 50% deuterated ethylenglycole-d_6. The solution was mixed carefully by pipetting and transferred into a 3 mm quartz EPR Q-band tube. After the solution was put down to the bottom of the tube, the tube was flash-frozen in liquid nitrogen and immediately measured without thawing or stored in liquid nitrogen.

9.2.3.2 Instrumentation and Measurement

The PELDOR experiments were performed on an ELEXSYS E580 pulsed spectrometer equipped with an ER5106QT-2 Q-band resonator, both from Bruker (US). A continuous flow helium cryostat (CF935) with a liquid helium dewar was connected to the spectrometer and an ITC 502 temperature control system was used for stable cooling conditions. The microwaves were applied into the spectrometer with a 150 W TWT-amplifier (Applied Systems Engineering, US). The spectrometer was controlled and data were recorded with manufactures own software.

For the PELDOR experiment, the temperature of the resonator was cooled down to 50 K with liquid helium. The strength of the magnetic field was set to the maximum of echo intensity from the field sweep spectrum with fixed frequency ν_B for the pump pulse. The observer frequency ν_A was set 80 MHz lower than the pump pulse frequency. The short repetition time was typically set to 1000 µs and the length of π/2-pulse was set to 12 and π-pulse to 24 ns. The pump pulse was set to length of maximum of inversion from 3-pulse nutation experiment (typically around 16 ns). The time τ_1 was set to 260 ns and τ_2 to values depending on the result from the 2pESEEM experiment. The time traces were recorded with a video amplifier bandwidth of 20 MHz and an amplifier gain of 42-45 dB. For suppression of deuterium modulation, 8 spectra with variable τ_1 times ($\Delta\tau$=16ns) were combined. Standardly, several hundred spectra were recorded and averaged, to get a sufficient SNR.

9.2.3.3 Data processing and analysis

The resulting PELDOR time traces were processed with the MATLAB program DeerAnalysis (Jeschke et al., 2006). To eliminate intermolecular distances from the spectrum, the PELDOR time traces were first background corrected. The corrected time traces were used for calculation of distance distributions. The influence of the background-correction on the distances was analysed with the evaluation tool of the program and added for validation of the results into the distance distributions.

9.2.3.4 Label positions and distance predictions

If protein structures were available, possible labelling positions for PELDOR distance measurements were initially identified with PyMOL plugin mtsslWizard (Hagelueken et al., 2015) or the server-based version mtsslSuite (Hagelueken, University of Bonn, www.mtsslsuite.isb.ukbonn.de). Difference distance matrices were prepared with this software to identify regions with high structural changes between two protein conformations. Possible regions were inspected on the proteins structure with consideration of surface accessibility and functional relevance. In each region, one residue was chosen for site-directed spin labelling. Subsequently, the selected residues were labelled *in silico* for each protein conformation with MTSSL by using again the mtsslWizard software (Hagelueken et al., 2015). The software was used to calculate the *in silico* distance distributions between the label ensembles and distributions for different conformations were compared regarding a distance change. In addition to this preliminary work, the *in silico* distance distributions were used to compare the experimental PELDOR distance distribution of the corresponding double mutants with the predictions based on the crystal structure.

If no protein structure was available, possible labelling positions were determined from other parameters such as conservation of amino acids. For the determination of label positions for the TRAP transporter membrane proteins, hydrophobic properties of side chains and topology models with predictions of membrane regions were used. Further details for determination of labelling sites on HiSiaQM are presented in Chapter 5.3.

9.2.4 Single molecule FRET spectroscopy

FRET spectroscopy (Förster resonance energy transfer) is a method which allows to determine distances on a nanometer scale by measure a radiation-free energy transfer from a dipole-dipole coupling between two different fluorophores (Gust et al., 2014). Due to this, FRET is often used for distance measurements on biomolecules to analyse conformational changes, comparable to pulsed EPR spectroscopy. For FRET experiments in this work, double fluorophore labelled proteins were used with known and commercially available donor-acceptor FRET pairs (Table 9-8).

In a FRET experiment, the valence electrons of the donor-fluorophore are excited with light of their specific absorption wavelength. Instead of emission of light and returning to the ground state, a part of the energy is transferred to a nearby acceptor-fluorophore. Consequently, the valence electrons of the acceptor-fluorophore are changed into an excited state. The relaxation of these electrons leads to the emission of light that can be detected as fluorescence signal at acceptor specific emission wavelength. The efficiency of the energy transfer E_{FRET}, also called FRET efficiency, between the donor and the acceptor fluorophore is defined as $E_{FRET} = R_0^6/(R_0^6+r^6)$ and depends on the experimental distance r between the two fluorophores (Lakowicz, 2006). The connection between the efficiency of energy transfer and the distance r is shown in Figure 2-1 H, the efficiency decreases with an increase of the radius. The term R_0 describes the Förster radius, a fluorophore pair-specific term. This term defines the distance between the donor and acceptor at a FRET efficiency of 50% and is described as $R_0 = 0.211\ (\kappa^2 n^{-4} Q_D J(\lambda))^{1/6}$, where κ is the relative orientation between the donor emission-dipole and the acceptor absorption-dipole, n the refractive index of the medium, Q_D the fluorescence quantum yield of the donor and J the overlap of the donor emissions-spectrum and the acceptor absorption-spectrum (Lakowicz, 2006). In summary, if the fluorophore parameters of the given equations are known, the distance between two fluorophores can be calculated through experimental determination of the FRET efficiency (Figure 2-1 F and H). The FRET efficiency can in general be determined in different ways, often with calculation from detected fluorescence intensities of the donor and acceptor (Lakowicz, 2006). Nowadays, FRET measurements allow the detection and investigation of single spectroscopically active molecules (Roy et al., 2008).

9.2.4.1 Labels for single molecule FRET

Organic dyes are nowadays often used for single molecule FRET and a large selection of labels with different spectrophotometric properties is commercially available, ranging from blue to red light wavelength region (Gust et al., 2014). In general, fluorophores for single molecule FRET should fulfil the following requirements: good brightness, high photostability, small size, high water-soluble and tolerance with bio-conjugation chemistry (Roy et al., 2008). Furthermore, the combination of two fluorophores to a FRET pair should have a large separation between the emission spectra of the donor and acceptor dyes and similar quantum yields and detection efficiencies (Roy et al., 2008). For attachment of the fluorophores to proteins, site-directed labelling of cysteines is often use and divers functionalized dyes are commercially available. The fluorophores that are used in this work are presented in Figure 2-2 and Table 9-8, the labelling procedure is described in Chapter 9.1.6.2.

9.2.4.2 Preparation of smFRET experiment

The smFRET experiments that are described in this thesis were performed by Christian Gebhardt and Marijn de Boer, both from the group of Thorben Cordes (University of Groningen and University of Munich). Depending on the study, the smFRET experiments were performed with surface-immobilized and free-diffusing proteins (Figure 2-1 C). Conformational changes were analysed with free-diffusing proteins, while tracking of time-resolved dynamics were studied with surface-immobilized proteins. The sample preparation, experimental set-up and measurement conditions for both methods are described in the following paragraphs.

The surface-based experiments were carried out on a home-built confocal scanning microscope at room temperature, the set-up of this microscope was described by Gouridis et al. (Gouridis et al., 2015). The preparation, experimental set-up and technical details were previously described by de Boer et al. (de Boer et al., 2019b, de Boer et al., 2019a). For the surface-immobilization of the labelled protein, freshly deep cleaned microscope coverslides were incubated with a mixture of PEG-silane (2-[methoxy(polyethleneoxy)-6,9-propyl]trimethoxysilane) and Biotin-PEG-silane (MW 3400) in toluene at 55 °C overnight. On the next day, the surface was incubated with 0.2 mg/mL neutravidin in standard buffer (50 mM Tris-HCl, pH 7.4, 50 mM KCl) for 10 min and the unbound neutravidin proteins were washed away with the same buffer. Afterwards the bound neutravidin was treated with a solution of 1 nM biotinylated penta-His antibody in standard buffer for 10 min. As last step, the fluorophore labelled His-tagged proteins were immobilized on the surface by incubation of a 5-100 pM protein solution for 5 min in standard buffer and unbound protein was washed away with standard buffer. The immobilizing strategy is outlined in Figure 3-18 A. The surface-based smFRET experiments were performed without substrate, with a high excess of substrate and 500 nM of substrate Neu5Ac. The last concentration was around the K_D of the protein (Figure 3-4) and represented a condition with high deficit of substrate. The FRET efficiency time traces, resulting from the experiment, were processed and analysed as published before with Hidden Markov Model (HMM) (de Boer et al., 2019a, Rabiner, 1989). In the cases of 1 mM Neu5Ac and absence of Neu5Ac, the time-traces were fitted with one state trajectories and in the substrate-deficit experiment, with two states trajectories.

For solution-based smFRET experiments, a homebuilt confocal ALEX (alternating-laser excitation) microscope was used as described by Kapanidis et al. (Kapanidis et al., 2004). For sample preparation, 100 µL solution with labelled protein was diluted to a concentration of 50 pM with buffer (50 mM Tris-HCl, pH 7.4, 50 mM KCl) and was pipetted to a coverslip. The technical details of the equipment are described in Peter et al. (Peter et al., 2020). After identification of single molecule events, the data analysis was performed with previous published software package (Gouridis et al., 2015). The final, processed data were plotted in ES-histograms (Figure 2-1 G) (Kapanidis et al., 2004). The double labelled events were extracted from these by selecting $0.25<S<0.75$ and the FRET efficiencies for distance determinations were extracted with Gaussian distribution fits.

9.2.4.3 Distance predictions

For validation and comparison of experimental determined FRET distances with protein structures, the distances for selected labelled residues were determined *in silico* with the FRET-restrained positioning and screening (FPS) method (Kalinin et al., 2012). With this program, the accessible volume (AV) of the labels at certain positions on the protein surface were calculated with three geometric parameters, the linker-length, the linker-width and the ellipsoid volume of the fluorophore (Table 9-8). The AVs of combined double label mutants were used to determine the average distance between the labels, based on the spherical distribution on the proteins surface.

Table 9-8: Geometric parameters for *in silico* predictions of different FRET labels. The data from the first dye (*) were taken from (Gebhardt et al., 2020) and for the last three dyes ($^+$) from (Kalinin et al., 2012). The table was adapted from (Peter et al., 2020).

Label	Linker length [Å]	W [Å]	R1 [Å]	R2 [Å]	R3 [Å]
AlexaFluor 555 – C2 maleimide*	21	4.5	8.8	4.2	1.5

AlexaFluor 647 – C2 maleimide$^+$	21	4.5	11	4.7	1.5
TMR$^+$	12	4.5	6	4.2	1.5
Cy5.5$^+$	21	4.5	11	3	1.5

9.2.5 Isothermal titration calorimetry (ITC)

Isothermal titration calorimetry (ITC) is a method to analyse the thermodynamics of a reaction between interacting molecules and is often used to characterize the reaction between a biomacromolecule and a smaller compound such as ligands or antibodies (Wiseman et al., 1989). Furthermore, it can be used to analyse interaction between two or more macromolecules in complex formation or the reverse dissociation of a macromolecular complex (Velazquez-Campoy et al., 2015). An advantage of ITC is the use of label- and modification-free reactants and that the components are present in solution and not immobilized to a surface, as in SPR experiments (Chapter 9.2.6).

During the mentioned reactions in biomolecules, non-covalent bonds are formed or rearranged while the components interact with each other. The energy required for this process is exchanged with the environment in the form of heat. In an exothermic reaction, the heat is released while an endothermic reaction absorbs the heat from the environment. Using a closed system like a calorimeter, the heat exchange of the system with the environment can be measured and used for calculation of thermodynamic parameters. Due to the fact, that the energy in this closed system can only be exchanged as heat and that the energy of a closed system is constant, the change of free enthalpy ΔG (Gibbs energy) during the reaction at constant temperature can be defined as $\Delta G = \Delta H - T\Delta S = -RT \ln(K)$ (Du et al., 2016). In case of a binding process, K is the binding affinity constant of the reaction and describes the equilibrium of the chemical reaction. The enthalpy ΔH corresponds to the amount of heat of the reaction. In the case of a protein-ligand binding process, this term includes the formation and disruption of noncovalent interactions, for example all polar and nonpolar interactions (Du et al., 2016). The entropy term ΔS of a biomolecular binding process can be described as $\Delta S = \Delta S_{solv} + \Delta S_{conf} + \Delta S_{r/t}$ (Du et al., 2016). The first term ΔS_{solv} includes the solvent entropy changes that results from solvent rearrangements due to a change of surface charge or accessibility of protein regions upon ligand binding. ΔS_{conf} describes the change of the conformational freedom of protein and ligand upon binding and the term $\Delta S_{r/t}$ describes the change of translational and rotational degrees of freedom (Du et al., 2016). In an ITC experiment, ΔH and K are determined directly from the experimental result and the thermodynamic parameters ΔG and ΔS can be calculated with the given equation (Du et al., 2016, Perozzo et al., 2004).

The calorimeter exists of two identical cells, a reference cell filled with water and a sample cell. In the case of a protein-ligand binding experiment, the sample cell is typically loaded with the protein solution. A syringe with a small stirrer is filled with the second reaction partner, often a ligand, substrate or other specific binders, and dipped into the sample cell. For the measurement, both cells were slightly and sensitively heated up to a constant temperature. Under stirring, defined volumes from the syringe were injected into the protein sample. In the case of interaction between the two components, the sample cell absorbed or released energy in form of heat while the protein and ligand reacted with each other. As a consequence of the heat change, the ITC instrument varied the power of the heater from the sample cell to keep the same temperature in the sample and reference cell. The difference in the power of the heaters (DP) from both cells were detected and used for analysis of the thermodynamic parameters. For a classical binding event, a high DP is obtained for the first peaks, suggesting a large amount of heat change in the sample cell due to complete reaction of both components. After decrease of the DP, the last peaks are optimally significant smaller, due to no longer available reaction partner of the injected component (Blume, 2012).

The reaction heat from each injection was determined by integration of the differential heat power peak. In a calorimeter, the reaction heat is equal to the enthalpy ΔH and a binding isotherm can be created by plotting the normalized reaction enthalpy against the molar ratio of the reaction partners and fitting of a binding model against the data points. Using the binding isotherm, the reaction enthalpy ΔH can be determined from the height of the sigmoidal curve and the binding affinity from the slope (Blume, 2012). Furthermore, the stoichiometry of the reacting partners can be analysed from the molar ratio value at the midpoint of the isotherm. As mentioned above, the measured and determined parameters can be used to calculate the free enthalpy and entropy of the reaction.

The measurements were performed on a MicroCal PEAQ-ITC instrument from Malvern Panalytical (Germany) with the corresponding software. The software was also used to design the experiments, meaning to calculate optimized component concentrations, and to analyse the results. The reference cell was filled with filtered water and the solution was replaced in regular intervals. In the case of sialic acid titration to the soluble domain of the TRAP transporters HiSiaP and VcSiaP, the ligand is dissolved in the protein buffer. After dilution of the protein and ligand to the desired concentration with standard protein buffer, the sample cell was washed 5 times with 300 µL of the protein standard buffer. Especially the last washing step was performed with prevention of air bubbles. 300 µL of the protein sample was loaded carefully into the cell, also with prevention of any remaining air in the cell. The excess of sample on the top of the cell was removed and the syringe was automatically loaded with the prepared ligand solution. After dipping the syringe into the sample cell, the measurement was started. All measurements were performed at 25 °C and with 13, 19 or 27 injections of 3, 2 or 1 µL, respectively. The excess protein solution from the loading sample was used to verify the protein concentration with a NanoDrop spectrophotometer and this concentration was used for determining of the thermodynamic parameters.

9.2.6 Surface plasmon resonance

Surface plasmon resonance (SPR) is a method that was firstly applied for studies of biomolecules in the early 1980s and is widely used for analysis of the interaction, in particular the binding behaviour, of biomolecules (Liedberg et al., 1983). An SPR binding experiment can provide information about the binding, specificity, kinetics and affinity between interacting (bio-)molecules (Nguyen et al., 2015, Oesterhelt 2012). Based on these experimental properties, the SPR technique has several applications for example screening for drug discovery or characterization of antibody-antigen interactions or other specific binders (Della Pia and Martinez, 2015, Nguyen et al., 2015, Oesterhelt 2012).

The principle of these technique is based on the conditions of resonance between a surface plasmon and photon of light (Oesterhelt 2012). The surface plasmon is build up from an electron gas on a metal-dielectric interface (often gold surface), by excitation with photons of a light source. The resonance condition of the light by the metal electron gas depends on the angle of the incoming light and the refractive index of the medium on the other side of the metal film. If the experimental set-up matches the resonance condition, most of the light is coupled into the plasmon and just a small amount is reflected. At this condition, the electromagnetic field intensity of the plasmon can reach up to 300 nm from the metals surface

(Oesterhelt, 2012). A binding process on the metal surface, which is in the region of the plasmon, results in the change of the resonance condition. The coupling of the incoming light to the plasmon changes and this can be measured by a change of the resonance angle of the reflected light, described as resonance units (RU). This unit is proportional to the amounts of molecules, bound to the surface (Oesterhelt, 2012).

In a SPR instrument, one side of the chip is facing to a flow cell and can be probed with different molecules. By this, association and dissociation of molecules on the surface as well as interactions between surface-coupled and free molecules can be analysed. In affinity studies, often one binding partner (ligand) is immobilized on the surface and the other partner (analyte) is flushed through the flow cell to the surface and interaction between both components is detected. Due to many applications of the SPR technique, different modifications of the surface were developed and chips with several modified surfaces are commercially available (Nguyen et al., 2015, O'Shannessy D. et al., 1992). For SPR studies in the present work, the ligand was specifically biotinylated and immobilized on a streptavidin surface.

The SPR experiments were performed with a Sensor Chip SA on a Biacore^{TM} 8K instrument, both from GE healthcare life sciences (US) and the related experimental and analytical software. On the chip, streptavidin is covalently bound to a carboxymethyldextran matrix on a layer of gold. The system, flow cells and surface of the chip were equilibrated with standard QM protein buffer with DDM detergent. Afterwards, the ligand HiSiaQM was injected into the flow cells and bound to the streptavidin surface until a constant response signal was reached. The single cycle kinetics strategy was used for the binding experiments of specific analytes (Karlsson et al., 2006). This experiment allows the determination of binding affinity and kinetics between two molecules in just one flow cell with short association and incomplete dissociation steps and without regeneration steps. The analyte was applied successively in six injections into the flow cell with doubling of the ligand concentration of each injection step. The optimal analyte concentration was estimated with the expected binding affinity or adjusted based on previous experiments. After the last injection, the regeneration was performed for a longer time and partially observed until complete regeneration. To exclude effects of injection of the analyte into the flow cells such as buffer variations or analyte-surface interaction, each injection was also applied into a flow cell without ligand. The difference between the reference and the ligand-bound flow cell resulted in a SPR sensorgram that was automatically analysed by the software with a single cycle binding-kinetic fit. From this, the binding affinities (K_D) and kinetics of the association and dissociation (k_{on} and k_{off}, respectively) were calculated.

If different analytes are found that bind to the same ligand, SPR measurements can be used to determine the binding regions, also named epitope mapping (Abdiche et al., 2009, Fägerstam et al., 1990, Thomsen and Gurevich, 2018). This experiment allows the identification of analytes that can bind the ligand at the same time or mutually disturb the binding. The identification of binding regions or epitopes is based on a competitive binding experiment. In analogy to the standard binding experiment, a ligand was bound to the surface and an analyte was applied in the flow cell onto the surface. If binding was observed and a constant response signal was reached, the second analyte was applied into the flow cell and the interaction of this analyte with the ligand-analyte complex was investigated. In the case of not detectable binding, the binding region was presumably blocked by the first analyte. To support a competitive binding behaviour, another run was performed where the second analyte was firstly applied to the flow cell and then the first analyte was flushed into the flow cell. If both experiments showed no binding in the second injection, the analytes blocked each others binding regions and if binding are visible, the binding regions were presumably separate on the antigens surface.

9.2.7 X-ray crystallography

X-ray crystallography is the most used method for determination of macromolecular structures and until this year, around 89% of all known macromolecular structures in the protein data bank (PDB, www.rcsb.org, access: 17.02.2021) are solved by this method. The first and often critical step for this method is the production of protein crystals with usable size and sufficient X-ray diffraction quality. In a single protein crystal, the biomolecules form defined and repetitive intermolecular interactions and by this, set up repeating units which form a single crystal with a spatial periodicity. These interactions are necessary and due to the complex surface of proteins regarding polarity, charge and shape, they are not predictable. Hence, commercial screens are available with a great diversity of solutions with different compositions, which should support the formation of intermolecular interactions and already turned out to promote crystallization in many trials. For the formation and growing of protein crystals, several methods are available (Chayen and Saridakis, 2008). In this work, the widely used vapor diffusion method is used in which the concentration of all components in protein solution, including the supplemented crystallization screen ingredients, are increased slowly over time through evaporation of the solvent molecules. Due to the increasing concentrations, the protein is delivered to a supersaturated solution in which the spontaneous nucleation of crystals can occur.

To determine the structure of the macromolecules, the crystal is irradiated with an X-ray beam of monochromatic wavelength. The electromagnetic waves interact with the electrons from the atoms in the crystal and are diffracted. The diffracted X-rays interfere with each other and constructive interference from a crystal can be detected as diffraction pattern. The condition for a reflection can be described by Bragg's law $2d \sin(\theta) = n\lambda$, where d is the distance between Bragg planes, θ the angle of incident X-ray beam, n the diffraction order (integer) and λ the wavelength of the X-ray beam (Rupp, 2010). The diffraction pattern of a single crystal therefore contains information about the structure and content of the crystallographic unit cell. Consequently, the diffraction pattern can be used to reconstruct a model of the proteins that build up the crystal. For this, the amplitude and the phase of each diffraction spot must be defined. The amplitude of each reflection can be calculated from the intensity of the detected pattern, but the phases are not detectable. To reconstruct the phases, different methods such as isomorphous replacement, multiple anomalous dispersion and molecular replacement can be used. In the present work, protein structures with a high sequence identity to the crystallized protein were available and the phases were reconstructed with molecular replacement. The phases were used for construction of the electron density map, in which a protein model can be built in and refined under consideration of different quality factors until a protein model is observed that fits to the diffraction data.

9.2.7.1 Preparation of protein crystals and diffraction experiments

In all crystallization trials, the sitting drop method was used for preparation of protein crystals. For initial screening of crystal forming conditions, different commercially available protein crystallization screens were used with a variety in each condition regarding buffer system, pH-value, salt and precipitant composition. The two screens *MemGold* I and *MemGold* II were exclusively used for membrane protein crystallization. The screens were aliquoted in a 96 deep well block and stored at 4 °C until further use or at -20 °C for long term storage. Crystallization plates were automatically prepared in 96 MRC-well plates with a Crystal Gryphon LCP device (Art Robbins Instruments, US). First, 50 µL of each of the 96 crystallization screen conditions were pulled up and 25 µL were pipetted into each reservoir well. Then, 0.1 µL were pipetted into the sitting drop-well and the remaining solution was added to the same reservoir well as before. Immediately, 0.1 µL of the protein solution were dropped into each sitting drop-well and the plate was sealed airtightly with an adhesive film. If necessary, the protein solution was previously supplemented with additives or compounds such as substrate or nanobodies. The sitting drop-wells were controlled under a light microscope and in the case of two single drops in one of the 96 wells, the plate was centrifuged for several seconds at 1000 rpm. For formation of protein crystals, the plates were standardly stored in a Rock Imager 1000 (Formulatrix, US) excluding light, physical stress like vibrations or bumps and at constant temperature of 20 °C. Each sitting drop were imaged in fixed intervals and the images were regularly checked for formation of crystals.

Detected crystals in sitting drops were either directly mounted for a diffraction experiment or an optimization screen was prepared. Concerning the latter, the optimization screens were designed with an institute-internal program. In this, the condition of the initial crystallization hit was varied concerning the ingredients concentrations and the pH-value and new conditions were randomly defined by the program. Stock solutions of the ingredients were prepared freshly with water and filtered through a 0.2 µm sterile filter. The optimized conditions were pipetted from the stock solution with an epMotion 5073m robot (Eppendorf, Germany). Crystal plates with optimization screens were prepared in the same way as for the initial plates. For diffraction analysis of observed crystals, the crystals were mounted from the sitting drop to a cryo-loop of fitting size. To avoid the formation of ice crystals during freezing of the crystal in liquid nitrogen, a cryo-protectant was used. If no protectant was included in the screen condition by default, the cryo-protectant solution was prepared from the corresponding screen condition, supplemented with 35% glycerol. After collecting of the crystal in the loop, the loop was shortly dipped into the cryo-protectant solution and directly frozen in liquid nitrogen afterwards. Until transport and measurement, the crystals were stored in liquid nitrogen.

The diffraction experiments were performed at the Swiss Light Source (SLS) at the Paul Scherrer Institute (Villigen, Switzerland) with beamline PX or at PETRA III at the

Helmholtz Research centre DESY *(Deutsches Elektronen Synchrotron)* (Hamburg, Germany) with beamline P13. At SLS, an EIGER 16M detector was available for diffraction observing and at DESY a PILATUS 6M detector. The proteins were automatically mounted in the X-ray beam under constant cooling with a cold nitrogen stream. After alignment of the protein crystal in the X-ray beam path, two test-diffraction images were recorded with a rotation of the crystal by 90° between both images. The diffraction images were directly inspected regarding resolution of diffraction and clearness of reflections. In the case of successful indexing of the diffraction reflexes from the software, the suggested strategy for complete diffraction recording was used for data collection. In the case of failing indexing, the diffraction was recorded over an angle of ϕ=180° by rotating of the crystal of 0.1° per image. If necessary, the recorded diffraction images were inspected with the software programs ALBULA (Dectris) or iMosflm (Battye et al., 2011).

9.2.7.2 Data processing and structure determination

For structure building, the data sets with recorded diffraction images were firstly processed with the *XDS* program suite (Kabsch, 2010): First a possible detector surface distortion is corrected by *XYCORR* and then the background of the images is estimated for a classification of pixels to the background or a diffraction spot with *INIT*. In the next steps, the program *COLSPOT* identify and locate strong diffraction spots in a subset of the data images and the program *IDXREF* uses these spots and experimental parameters to determine parameters of the crystal lattice like the unit cell and space group. Before analysis of all spots, the program *DEFPIX* further specify the background and regions for detecting spots and the program *INTEGRATE* finally determines the intensity of reflections at defined regions. The last program *CORRECT* further processes the detected intensities with various corrections factors, refine the initial classified unit cell parameters and gives information about the quality and completeness of the data set. Finally, for further processing, the program *XDSCONV* was used to convert the reflection-intensity files from the previous processing into a format for structure determination.

After initial processing of the diffraction data sets, the phases for reconstitution of an electron density map were calculated with related protein structures via molecular replacement. This was done with the *Phaser* program (McCoy et al., 2007) from the program suite

PHENIX (Python-based hierarchical environment for integrated xtallography) (Liebschner et al., 2019). For the processed diffraction data set of wildtype structure of VcSiaP with sialic acid, the HiSiaP protein structure in substrate-bound conformation (PDB-ID: 3B50, (Johnston et al., 2008)) was used for MR. For the different VcSiaP mutants with substrate, the previously solved and refined VcSiaP structure was used. After calculation of the phases and electron density map, a model of the crystallized protein was built into the electron density map. The resulting protein structure, as well as the electron density, were examined with the molecular graphics program *Coot* (Emsley et al., 2010). Alternatively, the data were processed and validated with CCP4 (Collaborative Computational Project, Number 4, 1994), especially if processing with the initial programs failed.

The refinement of the protein structure was performed with the program *phenix.refine* (Afonine et al., 2012) by optimizing the agreement of the model parameters to the diffraction data and minimizing the difference between both. The progress of the refinement process is monitored by calculation of the R_{work} factor, which represents differences between the actual model and experimental data:

$$R_{work} = \frac{\sum \left| F_{obs\,(h)} - F_{calc} \right|}{\sum F_{obs\,(h)}}$$

F_{obs} are the experimental, observed amplitudes and F_{calc} the calculated amplitudes from the model. Simultaneously, a R_{free} factor is calculated for cross validation and by this, to prevent an overfitting of both amplitudes (Rupp, 2010). For this, 5% of the observed diffraction data are removed before the first refinement and not utilized for refinement steps but used for calculation of the R_{free} factor after each refinement (Rupp, 2010):

$$R_{free} = \frac{\sum \left| F_{obs\,(h \neq work)} - F_{calc} \right|}{\sum F_{obs\,(h \neq work)}}$$

After each automated refinement step, the values of the R-factors were checked and the model was graphically analysed with *Coot* (Emsley et al., 2010). During inspection, the model was further fitted to the electron density map by hand, considering the stereochemistry of the peptide backbone and side chains. Further, mutations and small molecules, which were visible in the electron density map, were added to the model. The models were continuously controlled with *MolProbity* (Chen et al., 2010), by analysis of colliding regions, unfavoured bond lengths and angles, as well as Ramachandran outliners.

9.2.8 EM

Electron microscopy was used for structural investigations of the TRAP transporter membrane domains. Since recent years, cryo-electron microscopy (cryo-EM) is used as standard method for structural determination of biomolecules (Frank, 2017, Merk et al., 2016, Nogales, 2016). In this method, transmission electron microscopies (TEM) are used with similarities to conventional light microscopies. In a nutshell, the electrons are emitted from a source and focused on a sample through electromagnetic lenses. Depending on the sample and preparation method, some electrons are blocked from the sample and some electrons are transmitted through the sample. The transmitted electrons pass additional electromagnetic lenses and focused on a detector behind the sample.

Two different sample preparation methods were used in this work, negative staining and vitrification in ice (cryo). Negative staining was used for initial screening of the electron microscopy samples, due to the advantage of fast sample preparation procedure (Gallagher et al., 2019, Thompson et al., 2016). Additionally, the images have a higher contrast than vitrified EM samples, due to the treatment with a heavy metal which interferes strong with the electrons. The negative staining sample is therefore well suited to obtain the heterogeneity and purity of the sample, directly after the last purification step (see next paragraphs). However, electron microscopy with negative staining has structural resolution limit of 10-20 Å and a structural description at near-atomic resolution is thereby not possible (Gallagher et al., 2019). For this approach, the protein samples were prepared with vitrification in aqueous solution. The advantage in this preparation method is the absence of any staining component and that the image is produced by the macromolecules itself (Engelhardt, 2012). This allows atomic resolution and 3D structural information calculation in a native aqueous environment. Requirement for this preparation method is the fixation of the biomolecules in amorphic ice to prevent the samples from being destroyed by the formation of ice crystals. This vitrification process is accomplished with high freezing rates of ~10^6 °C/s of the samples and a permanent handling of the sample at low temperatures (Thompson et al., 2016).

9.2.8.1 Sample preparation

All final sample SEC runs and the EM sample preparations were performed with Veronika Heinz from the group of Christine Ziegler (University of Regensburg). For the electron microscopy measurements, the membrane transporter HiSiaQM was reconstituted into MSP-nanodiscs for several reasons. On the one hand, this system imitates the native environment for the membrane domains much better than a detergent micelle. On the other hand, the nanodisc particles are easier to detect with the processing software, since the nanodiscs have an additional size of around 100 kDa molecular weight (Efremov et al., 2017). One important requirement for a protein sample for the EM preparation is to have as few aggregates as possible. For this, every EM sample was prepared with a new batch of reconstituted HiSiaQM protein in nanodiscs. The solution from the reconstitution-overnight dialysis (Chapter 9.1.5) was immediately used or concentrated to around 100 - 200 µL, depending on the amount of initially used protein. Afterwards, the solution was loaded onto a SEC column to separate aggregates from the desired protein peak. For the first EM samples, an SD 200 10/300 column was used, while for the following preparations a Superose 6 5/150 or an SD 200 3.2/300 was used. The latter two columns led to a decreased dilution of the protein sample and with smaller amounts of proteins, similar peak intensities were obtained. The column and buffers are mentioned in correlation with the resulting chromatograms in Chapter 7. The elution of the column was observed with a wavelength at 280 nm and the eluate was fractionated collected. The fractions from the expected proteins peaks were directly used for electron microscopy sample preparation.

As mentioned before, the protein samples were firstly analysed in a scanning electron microscopy with negative staining, to check the presence of single particles and degree of aggregation. For this, the eluted fractions were diluted according to the recorded chromatogram to around 2-5 mAU with corresponding SEC buffer. A carbon EM grid was treated for a few seconds in a plasma glow discharge instrument for an optimal attachment of the molecules on the surface, shortly before use. From the protein peak, 3 µL of the diluted protein solution was pipetted on the EM grid and incubated for 1 min. The excess buffer was carefully removed with filter paper and the grid was rinsed with water and 3 µL of 2% uranyl acetate solution was pipetted onto the grid. After incubation of 30 s, the solution was again removed with a filter paper and the grid was stored for measurement at room temperature. The EM measurements with negative stained samples were performed from the group of Christine Ziegler (University of Regensburg).

For higher resolution EM measurements, the protein was fixed in amorphic ice on an EM grid without staining. Therefore, the protein sample needed a concentration of around 2000 mAU in the elution fraction from the SEC run. If necessary, the protein fraction was slightly concentrated to this absorption unit, after no formation of aggregation through this step was confirmed with previous experiments. For vitrification of the protein sample, the Vitrobot system from ThermoFisher Scientific (US) was used. The protein was pipetted in the Vitrobot on a Quantifoil R1.2/1.3 EM grid (Quantifoil, DE) and vitrified into liquid ethane with a blot time of 5 or 6 s at 4 °C and 100% humidity. The samples were stored and transported in liquid nitrogen. The cryo-EM experiments were performed in the laboratory of Bettina Böttcher, University of Würzburg at a Titan Krios (300) with a Falcon III camera or at the Institute of Genetics and Molecular and Cellular Biology (IGBMC) in Strasbourg, with a K3 camera. The two software programs RELION (Scheres, 2012) or cryoSPARC (Punjani et al., 2017) were used for processing of the images, identification of particles via autopicking and preparation of 2D classes.

Table 9-9: Parameters for cryo-EM of HiSiaQM in MSP-nanodiscs, Chapter 7.3.

Grid type	Quantifoil R1.2/1.3
Blot time	5 sec
Blot force	0
Temperature / humidity	4 °C / 100%
Sample concentration	~ 2000 mAU (concentrated)
Sample buffer	50 mM Tris (pH 7.8), 50 mM Tris
Electron microscope (keV)	Titan Krios (300)
Camera	Falcon III (linear mode)
Mode	Nanoprobe, parallel
Cs	2.7
C2	70 µm
OA	100 µm

Spot size	5
Beam diameter	1.0 µm
Magnification	75000x
Pixel size	1.0635 Å
Total e- dose	85.04 e^{-}/ $Å^{2}$
Dose rate	18.64 pe^{-}/px/sec
Exposure time	5.16 s
Recorded micrographs	1759

Table 9-10: Parameters for cryo-EM of HiSiaQM in MSP-nanodiscs with nanobody, Chapter 7.4.

Grid type	Quantifoil R1.2/1.3
Blot time	5 sec
Blot force	0
Temperature / humidity	4 °C / 100%
Sample concentration	~ 2000 mAU (concentrated)
Sample buffer	20 mM Tris (pH 8.3), 100 mM NaCl
Electron microscope (keV)	Titan Krios (300)
Camera	Falcon III (counting mode)
Mode	Nanoprobe, parallel
Cs	2.7
C2	70 µm
OA	100 µm
Spot size	5
Beam diameter	1.0 µm
Magnification	75000x
Pixel size	1.0635 Å

Table 9-11: Parameters for cryo-EM of HiSiaQM in MSP-nanodiscs with megabody, Chapter 7.6.

Grid type	Quantifoil R1.2/1.3
Blot time	5 sec
Blot force	0
Temperature / humidity	4 °C / 100%
Sample concentration	1200 mAU
Sample buffer	20 mM Tris (pH 8.3), 100 mM NaCl
Electron microscope	Titan Krios
Camera	K3 (counting mode)
Illumination area	1000 nm
Beam size	1015 nm
C2	70 µm
OA	1

Spot size	6
Magnification	81000x
Pixel size	0.862 Å
Total e- dose	58.61 e^-/ $Å^2$
Dose rate	21.81 pe^-/px/sec
Exposure time	1.997 s
Recorded micrographs	7263

10 Materials

10.1 Consumables

The following chapter lists all commonly used consumables like reagents, chemicals, labels and enzymes and their suppliers. Due to a change of the laboratory during the practical work for this thesis, some materials are received from more than one supplier and marked with IPTC (Institute for Physical and Theoretical Chemistry) or ISB (Institute of Structural Biology). All materials were stored as described in the corresponding producer's manual.

Table 10-1: General chemicals and consumables.

Chemicals	Supplier or Manufacturer
6x DNA sample buffer	New England Biolabs
Acrylamide (Rotiphorese Gel 30)	Carl Roth
Agar	AppliChem (IPTC) Sigma-Aldrich (ISB)
Agarose	Carl Roth
Ampicillin, sodium salt	Carl Roth
APS (Ammonium peroxydisulfate)	Carl Roth
$CaCl_2$	Carl Roth
Coomassie Brilliant Blue R-250	Carl Roth
D(+)-Glucose	Carl Roth
D(+)-Maltose monohydrate	Carl Roth
D_2O	Deutero GmbH
DIBMA (Sokalan CP 9)	BASF
DM (n-decyl-β-D-maltoside)	Cube Biotech
DMPC (1,2-Dimyristoyl-sn-glycero-3-phosphocholine)	Tokyo Chemical Industry (TCI)
DOPC (1,2 Dioleyl-sn-glycero-3-phosphocholine)	Tokyo Chemical Industry (TCI)
dNTPs	New England Biolabs

EDTA (Ethlyenediaminetetraacetic acid)	Carl Roth
Ethidiumbromide	AppliChem
Ethyleneglycol D_6	Deutero GmbH
FOS-Choline-14	Cube Biotech
Glycerol	Carl Roth
Glycine	Carl Roth
Imidazole	Carl Roth
IPTG (Isopropyl-β-D-1-thiogalactopyranoside)	Carbolution (IPTC) Carl Roth (ISB)
Kanamycin sulfate	Carl Roth
KCl	Carl Roth
KH_2PO_4	Carl Roth
L(+)-arabinose	Carl Roth
L-glutamine	Carl Roth
$MgSO_4$	Carl Roth
Milk powdered	Carl Roth
N-acetylneuraminic acid	Carbosynth Limited
NaCl	Carl Roth
NaH_2PO_4	Carl Roth
Natrium-Cholate	Carl Roth
NH_4Cl (ammonium chloride)	Carl Roth
$NiSO_4$ (nickel sulfate)	ThermoFisher Scientific
PEQ Green	VWR
6x-His Tag antibody (primary)	Pierce (IPTC) Invitrogen (ISB)
Goat anti-mouse HRP conjugated antibody (secondary)	Pierce (IPTC) Invitrogen (ISB)
SMA (SL25010 S25)	Polyscope Polymers
Sodium dodecylsulfate (SDS), 20%	AppliChem
Tris-(2-carboxyethyl)-phosphine	Carl Roth
Tetramethylethylenediamine	Carl Roth
TMB blotting solution (1-Step Ultra)	ThermoFisher Scientific
Tris	Carl Roth
Tween 20	Carl Roth
Yeast extract	Carl Roth
α-DDM (n-dodecyl- α -D-maltoside)	Anatrace
β-DDM (n-dodecyl-β-D-maltoside)	Carl Roth / Biosynth Chemistry & Biology
β -Mercaptoethanol	AppliChem

Table 10-2: Labels

Labels	Supplier or Manufacturer
MTSSL (Methanethiosulfonate spin label)	Toronto Research Chemicals
Fluorescein-5-maleimide	ThermoFisher Scientific
Maleimide-PEG2-Biotin (EZ-Link^{TM})	ThermoFisher Scientific

Table 10-3: Marker and Standards

Marker	Supplier or Manufacturer
GeneRuler DNA Ladder 1kB	ThermoFisher Scientific (IPTC) Carl Roth (ISB)
100 bp DNA Ladder	Carl Roth (ISB)
PageRuler Plus Prestained Protein Ladder 10-250 kDa	ThermoFisher Scientific (IPTC)
PageRuler Prestained Protein Ladder 10-180 kDa	ThermoFisher Scientific (ISB)

Table 10-4: Enzymes and buffers

Enzymes and buffers	Type of enzyme	Supplier or Manufacturer
Pfu	Polymerase	ThermoFisher Scientific and Carl Roth
OptiTaq	Polymerase	Roboklon
T4 DNA ligase	Ligase	New England Biolabs
DpnI	Restriction	ThermoFisher Scientific (IPTC) New England Biolabs (ISB)
NcoI	Restriction	New England Biolabs
BstEII	Restriction	New England Biolabs
SapI	Restriction	New England Biolabs
XhoI	Restriction	New England Biolabs
BamHI	Restriction	New England Biolabs
TEV	Protease	Self-made
Pfu buffer		ThermoFisher Scientific and Carl Roth
CutSmart buffer		New England Biolabs

Tango buffer	ThermoFisher Scientific
Polymerase buffer B	New England Biolabs

Table 10-5: Crystallization screens

Crystallization Screens	Supplier or Manufacturer
JCSG plus	Jena Bioscience
PACT	Molecular Dimensions
Morpheus	Molecular Dimensions
Ligand Friendly	Molecular Dimensions
LMB	Molecular Dimensions
ProPlex	Molecular Dimensions
MemGold I	Molecular Dimensions
MemGold II	Molecular Dimensions

Table 10-6: Further consumables and general materials.

Material	Usage	Supplier or Manufacturer
96-well plate (2mL)	Chromatography and Crystallization	VWR
MRC plate	Crystallization	Jena Bioscience
Sealing foil	Crystallization	Molecular Dimensions
X-band tube	EPR spectroscopy	Wilmad-LabGlass
X-band capillary, 10 µL	EPR spectroscopy	Hirschmann Laborgeräte
Q-band tube	EPR spectroscopy	Wilmad-LabGlass
Ni-NTA resin	Protein purification	ThermoFisher Scientific
Vivaspin™ concentrator	Protein purification	Sartorius (IPTC)
Amicon®Ultra concentrator	Protein purification	Merck (ISB)
QIAquick PCR Purification Kit	DNA-extraction	Qiagen
ExtractMe DNA clean-up & gel-out kit	DNA-extraction	BLIRT S.A.
GeneJET Plasmid Miniprep Kit	DNA-extraction	ThermoFisher Scientific
ExtratMe Plasmid mini kit	DNA-extraction	BLIRT S.A.

PVFD membranes	Western Blot	Carl Roth (IPTC)
Nitrocellulose membranes	Western Blot	VWR (ISB)
ECL™ Start Western Blotting Detection Reagent	Western Blot	GE healthcare
SuperClear™ 24 well plates	Bacteria culture	Jena Bioscience
Snake Skin® dialysis tubing, 6000-8000 MWCO	Dialysis	ThermoFisher Scientific

10.2 Solutions and Buffers

Unless otherwise stated, all solutions and buffers in this thesis were prepared with deionised and filtered water from a water purification system. Buffers for chromatography application were filtered and degassed under vacuum in an ultrasonic water bath. Buffers containing glycerol were only prepared in small-scales and discarded after a few days of storage. At the Institute for Physical and Theoretical Chemistry, the buffers were stored bench-top at room temperature, at the Institute of Structural Biology the buffers were stored at 4 °C. All buffers are summarized and listed due to the protein and application they were mainly used for. Because of that a few buffers were firstly used and optimized during development of purification routines the tables only include the optimized standard buffers.

Table 10-7: Buffers for P-domain of TRAP transporter.

Buffer	Ingredients
Standard	50 mM Tris (pH 8), 200 mM NaCl
Reduction	50 mM Tris (pH 8), 200 mM NaCl, 1 mM TCEP
Elution	50 mM Tris (pH 8), 200 mM NaCl, 500 mM Imidazole
No-salt	50 mM Tris (pH 8)
High-salt	50 mM Tris (pH 8), 1 M NaCl

Table 10-8: Buffers for detergent stabilized QM-domain of TRAP transporter.

Buffer	Ingredients
Standard 1	50 mM KH_2PO_4 (pH 7.8), 200 mM NaCl, 20% Glycerol, *
Standard 2	50 mM KH_2PO_4 (pH 7.8), 200 mM NaCl, *
Extraction (DDM)	50 mM KH_2PO_4 (pH 7.8), 200 mM NaCl, 20% Glycerol, 1.5% DDM
Extraction (SMA)	50 mM KH_2PO_4 (pH 7.8), 200 mM NaCl, 20% Glycerol, 5% SMA
Wash (DDM)	50 mM KH_2PO_4 (pH 7.8), 200 mM NaCl, 20% Glycerol, 22 mM Imidazole, *
Wash (SMA)	50 mM KH_2PO_4 (pH 7.8), 200 mM NaCl, 20% Glycerol, 10 mM Imidazole
Elution	50 mM KH_2PO_4 (pH 7.8), 200 mM NaCl, 20% Glycerol, 250 mM Imidazole, *
Label-elution	50 mM KH_2PO_4 (pH 7.8), 200 mM NaCl, 20% Glycerol, 250 mM Imidazole, *

* For protein samples in DDM detergent, all buffers after extraction procedure were supplemented with 0.035% DDM.

Table 10-9: Buffers for MSP1, after (Ritchie et al., 2009).

Buffer	Ingredients
Lysis	Na_3PO_4 (pH 7.4), 1% Triton X, 0.001% Nuclease
Standard	40 mM Tris (pH 8.0), 300 mM NaCl
Wash 1	40 mM Tris (pH 8.0), 300 mM NaCl, 1% Triton X
Wash 2	40 mM Tris (pH 8.0), 300 mM NaCl, 50 mM Natrium-Cholate
Wash 3	40 mM Tris (pH 8.0), 300 mM NaCl, 50 mM Imidazole
Elution	40 mM Tris (pH 8.0), 300 mM NaCl, 300 mM Imidazole
Dialysis/No-salt	40 mM Tris (pH 8)
High-salt	40 mM Tris (pH 8.0), 1 M NaCl
Chromatography	10 mM Tris (pH 7.4), 100 mM NaCl, 1 mM EDTA

Table 10-10: Buffers for Nano- and Megabodies.

Buffer	Ingredients
TES extraction	200 mM Tris (pH 8), 0.65 mM EDTA, 500 mM Sucrose
Wash	50 mM Tris (pH 7.5), 150 mM NaCl, 10 mM Imidazole
Elution	50 mM Tris (pH 7.5), 150 mM NaCl, 500 mM Imidazole
Mb - chromatography	10 mM Tris (pH 7.3), 140 mM NaCl

Table 10-11: General buffers and solutions

Buffer/Solution	Usage	Ingredients
Sample buffer (4x)	SDS-PAGE	240 mM Tris (pH 6.8), 40% Glycerol, 8% SDS, 5% β-Mercaptoethanol, 0.04% bromophenol blue
Running buffer	SDS-PAGE	25 mM Tris (pH 8.3), 192 mM Glycine, 0.1% SDS
Stain solution – Coomassie	SDS-PAGE	45% Methanol (v/v), 10% Acetic acid (v/v), 2.5 g/L Coomassie Brilliant Blue R-250
Destain solution	SDS-PAGE	10% Ethanol (v/v), 10% Acetic acid (v/v)
Separating gel buffer	SDS-PAGE	1.5 M Tris (pH 8.8), 10 mM SDS
Stacking gel buffer	SDS-PAGE	0.5 M Tris (pH 6.8), 10 mM SDS
Transferbuffer (IPTC)	Western Blot	25 mM Tris (pH 8.3-8.5), 192 mM Glycine, 0.05% (v/v) SDS
Transferbuffer (ISB)	Western Blot	50 mM Tris (pH 8.3-8.5), 40 mM Glycine, 1 mM SDS, 20% (v/v) Methanol
TBST	Western Blot	50 mM Tris (pH 7.5), 150 mM NaCl, (v/v) 0.1% Tween 20
PBST	Western Blot	Roti®fair (Carl Roth) PBS, (v/v) 0.05% Tween 20
TAE buffer	Agarose gel	40 mM Tris (pH 8) 20 mM Acetic acid, 1 mM EDTA
PELDOR buffer	EPR spectroscopy	100 mM TES (pH 7.5) 100 mM NaCl, in D_2O

10.3 Media

Table 10-12: Media for protein expression.

Media	Ingredients
LB	10 g/L Tryptone, 5 g/L Yeast extract, 10 g/L NaCl, (pH 7)
M9-minimal	13.4 g/L Na_2HPO_4, 6 g/L KH_2PO_4, 1 g/L NaCl, 2 g/L NH_4Cl, 11 mg/L $CaCl_2$, 240 mg/L $MgSO_4$ (pH 7.4)
TB	12 g/L Tryptone, 24 g/L Yeast extract, 12.54 g/L K_2HPO_4, 231 g/L KH_2PO_4, (pH 7.2)

10.4 Bacterial strains

Table 10-13: *E. coli* strains. The competent cells were prepared in the laboratory according to a standard protocol for chemical competent cells.

Strains	Use
MC1061	Expression of HiSiaQM
DH5α	Transformation
C43 (DE3)	Expression of VcSiaP, VcSiaQM
BL21 (DE3)	Expression of HiSiaP
β10	Transformation
BW25113 (SEVY3)	Growth Assay
WK6	Expression of Nano- and Megabody

10.5 Plasmids and Oligonucleotides

Table 10-14. Plasmids with characteristics that were used in this work.

Plasmids	Characteristics	Source
pBADHisTEV	His_6 $^{N\text{-term}}$ TEV cleavage site $^{N\text{-term}}$ Ampicillin resistance pBAD promotor	Huanting Liu (University of St. Andrews)
pWKS30	Ampicillin resistance Lac promotor	Gavin H. Thomas (University of York)
pHEN6	pelB signal peptide $^{N\text{-term}}$	Nanobody Core Facility (University of Bonn)

	His$_6$ $^{\text{C-term}}$ LPETG-sortase motif $^{\text{C-term}}$ Ampicillin resistance Lac promotor	
pET28a	Kanamycin resistance Lac promotor	Novagen

Table 10-15: Oligonucleotides for QuickChange mutagenesis. The oligonucleotides were designed with the software Geneious and synthesized by SEQLAB sequence laboratories or Metabion international AG.

Primer	Sequence
VcSiaQM 14His	GTGATGGTGATGGTGATGGTGATGGTGATGGTAGTAGCT ATCACCATCACCATCACGATTACGACATCCCAACGACC
HiSiaQM E312Amber	GGTCAATTGTAGATCAAAGCAATGCGTGATGCTGGTTATG CTTTGATCTACAATTGACCTAAGCCCCCCGCATCAG
HiSiaQM E121Amber	TCCATTTCTAAGAAAGCCTCCTAAGGCATCAATAGGGAA GGCATTTCTTAGAAATGGATTTTCGCAGCACTGCCTG
HiSiaQM C94A	GTTAAAAGTACGAATACCGAAATGAATAAATAAGAAAATTGCTATAAAGAC GAAAATTCACAAATACGTTTGTACAATTATTAGTCTTTATAGCAATTTTCTTATT
HiSiaQM C325S/C334S	AATGCGTGATGCTGGTTATGACGATGATATTTCGGGAGGAAT CCAATAATAGAAGAAGCAGCAGTAATTCCTCCCGAAATATCATC
HiSiaQM C400S	TACACCAAAAGCTACGAGAGAACAACTTTCGAGCAGC AGATTGCCCAAAAAGATTGTTTAAAGCTGCTCGAAAGTTGTT
HiSiaQM C458S	ACTACGCCCGTAATTGCCATTGCTTCTATCGAACTATTAAA GTGTATAAAGAATTAACCTTAAAAAGCTTATTTAATAGTTCGATAGAAGCA
HiSiaQM K62C	GGCATTAGATTAGTTAAAAAATCAATAAATACGTGTTCTTGGCATCTCACAG GGTATGTTGGGTATCAGCGTTGCTGTGAGATGCCAAGAAC
HiSiaQM K78C	ATAAAGACTAATAATTGTACAAACGTATTTGTGAATTTTCTGATACATTCGGGCAT CAAGAACACGTATTTATTGATTTTTTAACTAATCTAATGCCCGAATGTATCAGAAA
HiSiaQM E41C	CCACGTAAACAAATAAGAGCTTGGCGAGTTCACAACTCCAAAT CGCCAAGTTTTTCATTCTCCGTTAATTTGGAGTTGTGAACTCG
HiSiaQM K393C	TTAAAGCTGCTCGAAAGTTGTTCTCTCGTAGCACATGGTGTAC CATTGCAAAAAAACGAGGTTATCCACGTACACCATGTGCTACG
HiSiaQM R573C	CTTTGGTTACCGTGGAAACTGACATATTTCCTACACAAGCAACAACAA CCACCAATGGGAATGGCTCTCTTTGTTGTTGCTTGTGTAGGAA
HiSiaP K254C	CAAAATATCACACTAAATTATTCGTAGATGGAGAGTGTGATTTAGTC CACGCCTTGTTTTTCAAAGAATGTGACTAAATCACACTCTCC
HiSiaP Q55C	TGCAAAGGTAAAGTCGAGAGAACCGTCTTTTAAGCATTTTAACATTG GTTCACAATTAGGTGATGACCGTGCAATGTTAAAATGCTTAAAAGAC
HiSiaP L175C	TTGACCATCGACGGCATTGGTTTGTAACGCACAATAAACTTC ATCACCAACACCAATGGCATTTTCTGAAGTTTATTGTGCGTTACAA
HiSiaQM K273C	CCATTCCTCCTGTGTAATGACCGAGTAAGGCACAAGCAAAAT AGGTGGGATTACCGAACGTATTTTTAATTTTGCTTGTGCCTTACTC
HiSiaQM E235C	GTAATGGGAAGCTGTCTAGGCTATAGACTAATTTACAAGTTGCGG ATGACACGTTGGAACGTCGTAAATGCCGCAACTTGTAAATTAGTC
HiSiaP K112C	GTTATAAGCTTGGGAAAGTAAAGTCACGCCAAGATCACAATCCATTTT TATTCGATACAGAATTCGGTAAAGATTTAATTAAAAAAATGGATTGTGATCTTGGC
HiSiaP E228C	TAGTAAGCAACGAGACTTATAAAGAACTCCCTTGTGATCTTC GGCAGCATCTTTTACGACTTTTTGAAGATCACAAGGGAG
HiSiaP D58C	AGCAGATTCTGCAAAGGTAAAGTCGAGAGAACCACATTTTAATTG GGTGATGACCGTGCAATGTTAAAACAATTAAAATGTGGTTCTCTC
HiSiaP A134C	GCTTATAACGGAACTCGCCAAACGACTTCAAATCGTTGTATCAACAG CACGAAGTTTTAAGCCTTTCATATCTGCAATACTGTTGATACAACGATTTGAAG
VcSiaP Q9N	CGAGTTATACTCTACAGACCCCACAGAAGCGTTCATCCC GGGGCGACGACTTTAAAGATGGGGATGAACGCTTCTG
VcSiaP T189A	CGTAGATGGGCAAGAAAACCCGTTACCAGCAATTAAAAC CTTTTGCACTTCATAGAACTTCATCGTTTTAATTGCTGGTAACG

VcSiaP G71Q	CTCGATATAACTTATGCTGAGTTTGGTCGTATGCAGCTTTG CGCTTCTGCTCGCGGTATCCAAAGCTGCATACGAC
VcSiaP E78A	GTATGCAGCTTTGGATACCGCGAGCAGCAGCGGTC CGCGACATAAGGGAGCGCGACCGCTGCTGCTC
VcSiaP M81A	GTCAAAATCTTTCGCGACATAAGGGAGCGCGACCG GATACCGCGAGCAGAAGCGGTCGCGCTCCC
VcSiaP A43S	GCATGGCACGATCATCACCAAGCTGGCTGCTTGG GAGAGATCAAACTCGCTTTGTACCCAAGCAGCCAGCTT

Table 10-16: Oligonucleotides for general cloning.

Primer	Sequence
Nanobody XhoI (rv)	CAGTACCTCGAGTGAGGAGACGGTGACCTG
Nanobody BamHI (fw)	GATCCGGGATCCATGCAGGTGCAGCTCGTGGA
Nanobody BamHI pelB (fw)	GATCCGGGATCCATGAAATACCTATTGCCTACGGC
Nanobody Amplification SapI	GATCCTGCTCTTCCGGGTCTCTGAGACTCTCCT CCTGCAGCTCTTCTTGGGTCCCCTGGCCC

10.6 Instruments and Columns

Table 10-17: General laboratory instruments

Instruments	Name/Type	Company
Ultracentrifuge	Optima XPN	Beckman Coulter
Ultracentrifuge Rotor	Type 50.2 Ti Fixed-Angle	Beckman Coulter
Water Purification System	Milli-Q Direct	Merck Millipore
Cell disruptor	Cell disruptor benchtop	Constant Systems Limited
Cell sonicator	SONOPULS mini20	Bandelin electronic
PCR Thermocycler	Mastercycler Nexus SX1	Eppendorf
Crystallization Robot	Crystal Gryphon LCP	Art Robbins Instruments
Crystallization Imager	Rock Imager 1000	Formulatrix
Pipetting Robot	epMotion 5073m	Eppendorf
FPLC	Äkta avant	GE Healthcare
FPLC	Äkta pure	GE Healthcare
FPLC	Äkta prime	GE Healthcare
HPLC	1260 Infinity II	Agilent
ITC	MicroCal PEAQ-ITC	Malvern Panalytical
SPR	Biacore 8K	GE Healthcare
Spectrophotometer	NanoDrop DN-2000C	ThermoFisher Scientific
Crystal-Loops		MiTeGen

Gel imager (ISB)	ChemiDoc XRS+	Bio-Rad Laboratories
Gel imager (IPTC)		Peqlab
Incubators (ISB)	Multitron	Infors HT
	Minitron	Infors HT
Incubator (IPTC)	Ecotron	Infors HT
	I26	New Brunswick Scientific
Ultrasonic bath	USC100T	VWR

Table 10-18: Columns

Columns	Company
HiLoad 16/600 Superdex 75 prep grade	GE Healthcare Europe GmbH
HiLoad 16/600 Superdex 200 prep grade	GE Healthcare Europe GmbH
Superdex 200 10/300 GL	GE Healthcare Europe GmbH
Superdex 200 Increase 10/300 GL	GE Healthcare Europe GmbH
Superose 6 Increase 5/150 GL	GE Healthcare Europe GmbH
Superdex 200 Increase 3.2/300 GL	GE Healthcare Europe GmbH
ENrich^{TM} Q 10/100	Bio-Rad Laboratories GmbH
HiTrap Q XL	GE Healthcare Europe GmbH
PD-10	GE Healthcare Europe GmbH
PD MiniTrap G-25	GE Healthcare Europe GmbH

Table 10-19: Instruments for EPR spectroscopy.

Instruments	Company
EPR spectrometer Elexsys E580	Bruker
EPR spectrometer EMXnano	Bruker
Temperature Control System ITC 502	Oxford Instruments
Helium Flow Cryostat	Oxford Instruments
TWT-amplifier	Applied Systems Engineering
Q-band resonator ER 5106QT-2	Bruker

10.7 Software

Table 10-20: Software and Program tools that were used for experimental design, data recording, handling, processing, illustrating and validation.

Software	Developer/Company
MATLAB R2015b	MathWorks
DeerAnalysis 2018	Jeschke et al., 2006
Xenon nano 1.2a.2	Bruker
Xepr 2.6b.151	Bruker
mtsslWizard	Hagelueken et al., 2015
Microcal PEAQ-ITC control V1.22	Malvern Panalytical
Microcal PEAQ-ITC analysis V1.22	Malvern Panalytical
Geneious Prime (ISB) Geneious 6.1.8 (IPTC)	Geneious
ProtParam	Expasy, Swiss Institute of Bioinformatics (Gasteiger et al., 2005)
OriginPro 8G	OriginLab
FPS	Kalinin et al., 2012
Biacore Insight Evaluation	GE healthcare
LigPlot	Laskowski and Swindells, 2011
COOT 0.8.9.2	Emsley et al., 2010
ALBULA Viewer	Dectris
CCP4i	(Collaborative Computational Project, Number 4, 1994)
iMosflm	Battye et al., 2011
Phaser	McCoy et al., 2007
Phenix	Liebschner et al., 2019
PyMOL Version 2.1.0	Schrödinger LLC
XDS	Kabsch, 2010
RELION	Scheres, 2012
cryoSPARC	Punjani et al., 2017

References

Abdiche, Y. N.; Malashock, D. S.; Pinkerton, A.; Pons, J.; Exploring blocking assays using Octet, ProteOn, and Biacore biosensors. *Analytical Biochemistry,* **2009,** *386* (2), 172–180.

Abdullin, D.; Duthie, F.; Meyer, A.; Müller, E. S.; Hagelueken, G.; Schiemann, O.; Comparison of PELDOR and RIDME for Distance Measurements between Nitroxides and Low-Spin Fe(III) Ions. *The Journal of Physical Chemistry B,* **2015,** *119* (43), 13534–13542.

Abdullin, D.; Schiemann, O.; Pulsed Dipolar EPR Spectroscopy and Metal Ions: Methodology and Biological Applications. *ChemPlusChem,* **2020,** *85* (2), 353–372.

Abramson, J.; Smirnova, I.; Kasho, V.; Verner, G.; Kaback, H. R.; Iwata, S.; Structure and mechanism of the lactose permease of Escherichia coli. *Science,* **2003,** *301* (5633), 610–615.

Ackermann, K.; Wort, J.; Bode, B.; Nanomolar Pulse Dipolar EPR Spectroscopy in Proteins Using Commercial Labels and Hardware. *ChemRxiv,* **2020**, DOI: doi.org/10.26434/chemrxiv.13370924.v1.

Ada, G. L.; French, E. L.; Purification of bacterial neuraminidase (receptor-destroying enzyme). *Nature,* **1959,** *183* (4677), 1740–1741.

Adão, R.; Cruz, P. F.; Vaz, D. C.; Fonseca, F.; Pedersen, J. N.; Ferreira-da-Silva, F.; Brito, R. M. M.; Ramos, C. H. I.; Otzen, D.; Keller, S.; Bastos, M.; DIBMA nanodiscs keep α-synuclein folded. *Biochimica et Biophysica Acta – Biomembranes,* **2020,** *1862* (9), 183314.

Afonine, P. V.; Grosse-Kunstleve, R. W.; Echols, N.; Headd, J. J.; Moriarty, N. W.; Mustyakimov, M.; Terwilliger, T. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D.; Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallographica. Section D, Biological crystallography,* **2012,** *68,* 352–367.

Ahuja, S.; Rougé, L.; Swem, D. L.; Sudhamsu, J.; Wu, P.; Russell, S. J.; Alexander, M. K.; Tam, C.; Nishiyama, M.; Starovasnik, M. A.; Koth, C. M.; Structural analysis of bacterial ABC transporter inhibition by an antibody fragment. *Structure,* **2015,** *23* (4), 713–723.

Aitha, M.; Moritz, L.; Sahu, I. D.; Sanyurah, O.; Roche, Z.; McCarrick, R.; Lorigan, G. A.; Bennett, B.; Crowder, M. W.; Conformational dynamics of metallo-β-lactamase CcrA during catalysis investigated by using DEER spectroscopy. *Journal of Biological Inorganic Chemistry,* **2015,** *20* (3), 585–594.

Akiyama, N.; Takeda, K.; Miki, K. Crystal structure of a periplasmic substrate-binding protein in complex with calcium lactate. *Journal of Molecular Biology,* **2009,** *392* (3), 559–565.

Albers, S.-V.; Elferink, M. G.; Charlebois, R. L.; Sensen, C. W.; Driessen, A. J. M.; Konings, W. N.; Glucose Transport in the extremely thermoacidophilic Sulfolobus solfataricus involves a high-affinity membrane-integrated binding protein. *Journal of Bacteriology,* **1999,** *181* (14), 4285–4291.

Alexander, C.; Rietschel, E. T.; Bacterial lipopolysaccharides and innate immunity. *Journal of Endotoxin Research,* **2001,** *7* (3), 167–202.

Alexander, N. S.; Stein, R. A.; Koteiche, H. A.; Kaufmann, K. W.; Mchaourab, H. S.; Meiler, J.; RosettaEPR: rotamer library for spin label structure and dynamics. *PloS One,* **2013,** *8* (9), e72851.

Alfonso-Garrido, J.; Garcia-Calvo, E.; Luque-Garcia, J. L.; Sample preparation strategies for improving the identification of membrane proteins by mass spectrometry. *Analytical and Bioanalytical Chemistry,* **2015,** *407* (17), 4893–4905.

Ali, M.; Nelson, A. R.; Lopez, A. L.; Sack, D. A. Updated global burden of cholera in endemic countries. *PloS Neglected Tropical Diseases,* **2015,** *9* (6), e0003832.

Allen, S.; Zaleski, A.; Johnston, J. W.; Gibson, B. W.; Apicella, M. A.; Novel sialic acid transporter of Haemophilus influenzae. *Infection and Immunity,* **2005,** *73* (9), 5291–5300.

Almagro-Moreno, S.; Boyd, E. F.; Sialic acid catabolism confers a competitive advantage to pathogenic vibrio cholerae in the mouse intestine. *Infection and Immunity,* **2009,** *77* (9), 3807–3816.

Altamirano, F. L. G.; Barr, J. J.; Phage therapy in the postantibiotic era. *Clinical Microbiology Reviews,* **2019,** *32* (2), e00066-18

Altenbach, C.; Flitsch, S. L.; Khorana, H. G.; Hubbell, W. L.: Structural studies on transmembrane proteins. 2. spin labeling of bacteriorhodopsin mutants at unique cysteines. *Biochemistry,* **1989,** *28* (19), 7806–7812.

Angata, T.; Varki, A.; Chemical diversity in the sialic acids and related alpha-keto acids: an evolutionary perspective. *Chemical Reviews,* **2002,** *102* (2), 439–469.

Arkhipova, V.; Guskov, A.; Slotboom, D. J.; Structural ensemble of a glutamate transporter homologue in lipid nanodisc environment. *Nature Communications,* **2020,** *11*, 998.

Avanti Polar Lipids. Phase transition temperatures for glycerophospholipids. https://avantilipids.com/tech-support/physical-properties/phase-transition-temps, accessed 26.11.2020, **2020**.

Azarkh, M.; Okle, O.; Eyring, P.; Dietrich, D. R.; Drescher, M.; Evaluation of spin labels for in-cell EPR by analysis of nitroxide reduction in cell extract of Xenopus laevis oocytes. *Journal of Magnetic Resonance,* **2011,** *212,* 450–454.

Barbour, M. L.; Conjugate vaccines and the carriage of Haemophilus influenzae type b. *Emerging Infectious Diseases* **1996,** *2* (3), 176–182.

Battye, T. G. G.; Kontogiannis, L.; Johnson, O.; Powell, H. R.; Leslie, A. G. W. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallographica Section D, Biological crystallography,* **2011,** *67,* 271–281.

Bayburt, T. H.; Grinkova, Y. V.; Sligar, S. G.; Self-Assembly of Discoidal Phospholipid Bilayer Nanoparticles with Membrane Scaffold Proteins. *Nano Letters* **2002,** *2* (8), 853–856.

Bayburt, T. H.; Sligar, S. G.; Self-assembly of single integral membrane proteins into soluble nanoscale phospholipid bilayers. *Protein Science,* **2003,** *12* (11), 2476–2481.

Bayburt, T. H.; Sligar, S. G.; Membrane protein assembly into nanodiscs. *FEBS letters,* **2010,** *584* (9), 1721–1727.

Beckers, M.; Drechsler, F.; Eilert, T.; Nagy, J.; Michaelis, J.; Quantitative structural information from single-molecule FRET. *Faraday Discussions,* **2015,** *184,* 117–129.

Berliner, L. J.; Grundwald, J.; Hankovszky, O.; Hideg, K.; A novel reversible thiol-specific spin label: papain active site labeling and inhibition. *Analytical Biochemistry,* **1982,** *119,* 450–455.

Berntsson, R. P.-A.; Smits, S. H. J.; Schmitt, L.; Slotboom, D.-J.; Poolman, B. A structural classification of substrate-binding proteins. *FEBS letters,* **2010,** *584* (12), 2606–2617.

Blume, A.; Mikrokalorimetrie: Isothermal titration calorimetry (ITC). In *Bioanalytik,* 3. Auflage; Lottspeich, F., Engels, J. W.; Springer Spektrum: Berlin, Heidelberg, **2012**; pp 67–75.

Böhm, S.; Licht, A.; Wuttge, S.; Schneider, E.; Bordignon, E.; Conformational plasticity of the type I maltose ABC importer. *Proceedings of the National Academy of Sciences of the United States of America,* **2013,** *110* (14), 5492–5497.

Bordenave, T.; Helle, M.; Beau, F.; Georgiadis, D.; Tepshi, L.; Bernes, M.; Ye, Y.; Levenez, L.; Poquet, E.; Nozach, H.; Razavian, M.; Toczek, J.; Stura, E. A.; Dive, V.; Sadeghi, M. M.; Devel, L.; Synthesis and in Vitro and in Vivo Evaluation of MMP-12 Selective Optical Probes. *Bioconjugate Chemistry,* **2016,** *27* (10), 2407–2417.

Bouchet, V.; Hodd, D. W.; Li, J.; Brisson; J-R.; Randle, G. A.; Martin, A.; Li, Z.; Goldstein, R.; Schweda, E. K. H.; Pelton, S. I.; Richards, J. C.; Moxon, E. R.; Host-derived sialic acid is incorporated into Haemophilus influenzae lipopolysaccharide and is a major virulence factor in experimental otitis media. *Proceedings of the National Academy of Sciences of the United States of America,* **2003,** *100* (15), 8898-8903.

Boudker, O.; Verdon, G.; Structural perspectives on secondary active transporters. *Trends in Pharmacological Sciences,* **2010,** *31* (9), 418–426.

Bountra, K.; Hagelueken, G.; Choudhury, H. G.; Corradi, V.; El Omari, K.; Wagner, A.; Mathavan, I.; Zirah, S.; Yuan Wahlgren, W.; Tieleman, D. P.; Schiemann, O.; Rebuffat, S.; Beis, K.; Structural basis for antibacterial peptide self-immunity by the bacterial ABC transporter McjD. *The EMBO Journal,* **2017,** *36* (20), 3062–3079.

Brautigam, C. A.; Deka, R. K.; Schuck, P.; Tomchick, D. R.; Norgard, M. V.; Structural and thermodynamic characterization of the interaction between two periplasmic Treponema pallidum lipoproteins that are components of a TPR-protein-associated TRAP transporter (TPAT). *Journal of Molecular Biology,* **2012,** *420* (1-2), 70–86.

Broecker, J.; Eger, B. T.; Ernst, O. P.; Crystallogenesis of membrane proteins mediated by polymer-bounded lipid nanodiscs. *Structure,* **2017,** *25* (2), 384–392.

Cassel, D.; Pfeuffer, T.; Mechanism of cholera toxin action: covalent modification of the guanyl nucleotide-binding protein of the adenylate cyclase system. *Proceedings of the National Academy of Sciences of the United States of America,* **1978,** *75* (6), 2669–2673.

Cellular gatekeepers. *Nature Structural & Molecular Biology,* **2016,** *23* (6), 463. DOI: 10.1038/nsmb.3246.

Chatterjee, S. N.; Chaudhuri, K.; Lipopolysaccharides of Vibrio cholerae. I. Physical and chemical characterization. *Biochimica et Biophysica Acta,* **2003,** *1639* (2), 65–79.

Chatterjee, A.; Sun, S. B.; Furman, J. L.; Xiao, H.; Schultz, P. G.; A versatile platform for single- and multiple-unnatural amino acid mutagenesis in Escherichia coli. *Biochemistry,* **2013,** *52* (10), 1828–1837.

Chayen, N. E.; Saridakis, E.; Protein crystallization: from purified protein to diffraction-quality crystal. *Nature Methods,* **2008,** *5* (2), 147–153.

Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C.; MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D, Biological Crystallography,* **2010,** *66,* 12–21.

Chowdhury, N.; Norris, J.; McAlister, E.; Lau, S. Y. K.; Thomas, G. H.; Boyd, E. F.; The VC1777-VC1779 proteins are members of a sialic acid-specific subfamily of TRAP transporters (SiaPQM) and constitute the sole route of sialic acid uptake in the human pathogen Vibrio cholerae. *Microbiology,* **2012,** *158* (Pt 8), 2158–2167.

Chowdhury, F. R.; Nur, Z.; Hassan, N.; Seidlein, L. von; Dunachie, S.; Pandemics, pathogenicity and changing molecular epidemiology of cholera in the era of global warming. *Annals of Clinical Microbiology and Antimicrobials,* **2017,** *16* (1), 10.

Claxton, D. P.; Jagessar, K. L.; Steed, P. R.; Stein, R. A.; Mchaourab, H. S.; Sodium and proton coupling in the conformational cycle of a MATE antiporter from Vibrio cholerae. *Proceedings of the National Academy of Sciences of the United States of America,* **2018,** *115* (27), E6182-E6190.

Claxton, D. P.; Quick, M.; Shi, L.; Carvalho, F. D. de; Weinstein, H.; Javitch, J. A.; Mchaourab, H. S.; Ion/substrate-dependent conformational dynamics of a bacterial homolog of neurotransmitter:sodium symporters. *Nature Structural & Molecular Biology,* **2010,** *17* (7), 822–829.

Collaborative Computational Project, Number 4.; The CCP4 suite: programs for protein crystallography. *Acta Crystallographica Section D, Biological Crystallography,* **1994,** *50* (Pt 5), 760–763.

Condemine, G.; Berrier, C.; Plumbridge, J.; Ghazi, A.; Function and expression of an N-acetylneuraminic acid-inducible outer membrane channel in Escherichia coli. *Journal of Bacteriology,* **2005,** *187* (6), 1959–1965.

Cooper, G. M.; *The Cell. A molecular approach - Cell Membranes;* 2nd edition, Sunderland (MA): Sinauer Associates, **2000**. Cell Membranes, Available from: https://www.ncbi.nlm.nih.gov/books/NBK9928/ (accessed 02.11.2020).

Corfield, T.; Bacterial sialidases - roles in pathogenicity and nutrition. *Glycobiology,* **1992,** *2* (6), 509–521.

Coulton, J. W.; Chin, A. C.; Vachon, V.; Recombinant porin of Haemophilus influenzae type b. *The Journal of Infectious Diseases,* **1992,** *165* (suppl 1), 188–191.

Cox, A. D.; Hood, D. W.; Martin, A.; Makepeace, K. M.; Deadman, M. E.; Li, J.; Brisson, J.-R.; Moxon, E. R.; Richards, J. C.; Identification and structural characterization of a sialylated lacto-N-neotetraose structure in the lipopolysaccharide of Haemophilus influenzae. *European Journal of Biochemistry,* **2002,** *269* (16), 4009–4019.

Cuneo, M. J.; Changela, A.; Miklos, A. E.; Beese, L. S.; Krueger, J. K.; Hellinga, H. W.; Structural analysis of a periplasmic binding protein in the tripartite ATP-independent transporter family reveals a tetrameric assembly that may have a role in ligand transport. *The Journal of Biological Chemistry,* **2008,** *283* (47), 32812–32820.

Cunningham, T. F.; Putterman, M. R.; Desai, A.; Horne, W. S.; Saxena, S.; The double-histidine Cu^{2+} -binding motif: A highly rigid, site-specific spin probe for electron spin resonance distance measurements. *Angewandte Chemie International Edition,* **2015,** *54* (21), 6330-6334.

Dalsgaard, A.; Forslund, A.; Sandvang, D.; Arntzen, L.; Keddy, K.; Vibrio cholerae O1 outbreak isolates in Mozambique and South Africa in 1998 are multiple-drug resistant, contain the SXT element and the addA2 gene located on class 1 integrons. *Journal of Antimicrobial Chemotherapy,* **2001,** *48* (6), 827–838.

Darby, J. F.; Hopkins, A. P.; Shimizu, S.; Roberts, S. M.; Brannigan, J. A.; Turkenburg, J. P.; Thomas, G. H.; Hubbard, R. E.; Fischer, M.; Water networks can determine the affinity of ligand binding to proteins. *Journal of the American Chemical Society,* **2019,** *141* (40), 15818–15826.

Davidson, A. L.; Shuman, H. A.; Nikaido, H.; Mechanism of maltose transport in Escherichia coli: Transmembrane signaling by periplasmic binding proteins. *Proceedings of the National Academy of Sciences,* **1992,** *89* (6), 2360–2364.

Davidson, A. L.; Dassa, E.; Orelle, C.; Chen, J.; Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiology and Molecular Biology Reviews,* **2008,** *72* (2), 317-364.

de Boer, M.; Gouridis, G.; Muthahari, Y. A.; Cordes, T.; Single-molecule observation of ligand binding and conformational changes in FeuA. *Biophysical Journal,* **2019a,** *117* (9), 1642–1654.

de Boer, M.; Gouridis, G.; Vietrov, R.; Begg, S. L.; Schuurman-Wolters, G. K.; Husada, F.; Eleftheriadis, N.; Poolman, B.; McDevitt, C. A.; Cordes, T.; Conformational and dynamic plasticity in substrate-binding proteins underlies selective transport in ABC importers. *eLife,* **2019b**, DOI: 10.7554/eLife.44652.

Deisenhofer, J.; Epp, O.; Miki, K.; Michel, H.; Structure of the protein subunits in the photosynthetic reaction centre of Rhodopseudomonas viridis at 3Å resolution. *Nature,* **1985,** *318,* 618–624.

Deka, R. K.; Brautigam, C. A.; Goldberg, M.; Schuck, P.; Tomchick, D. R.; Norgard, M. V.; Structural, bioinformatic, and in vivo analyses of two Treponema pallidum lipoproteins reveal a unique TRAP transporter. *Journal of Molecular Biology,* **2012,** *416* (5), 678–696.

DeLeo, F. R.; Hinnebusch, B. J.; A plague upon the phagocytes. *Nature medicine,* **2005,** *11* (9), 927-928.

Della Pia, E. A.; Martinez, K. L.; Single domain antibodies as a powerful tool for high quality surface plasmon resonance studies. *PloS one,* **2015,** *10* (3), e0124303.

Denisov, I. G.; Grinkova, Y. V.; Lazarides, A. A.; Sligar, S. G.; Directed self-assembly of monodisperse phospholipid bilayer nanodiscs with controlled size. *Journal of the American Chemical Society,* **2004,** *126* (11), 3477–3487.

Denisov, I. G.; Sligar, S. G.; Nanodiscs for structural and functional studies of membrane proteins. *Nature Structural & Molecular Biology,* **2016,** *23* (6), 481–486.

Derakhshankhah, H.; Jafari, S.; Cell penetrating peptides: A concise review with emphasis on biomedical applications. *Biomedicine & pharmacotherapy,* **2018,** *108,* 1090–1096.

Dimura, M.; Peulen, T. O.; Hanke, C. A.; Prakash, A.; Gohlke, H.; Seidel, C. A. M.; Quantitative FRET studies and integrative modeling unravel the structure and dynamics of biomolecular systems. *Current Opinion in Structural Biology,* **2016,** *40,* 163–185.

Dmitriev, O. Y.; Lutsenko, S.; Muyldermans, S.; Nanobodies as probes for protein dynamics in vitro and in cells. *The Journal of Biological Chemistry,* **2016,** *291* (8), 3767–3775.

Doeven, M. K.; van den Bogaart, G.; Krasnikov, V.; Poolman, B.; Probing receptor-translocator interactions in the oligopeptide ABC transporter by fluorescence correlation spectroscopy. *Biophysical Journal,* **2008,** *94* (10), 3956–3965.

Dörr, J. M.; Koorengevel, M. C.; Schäfer, M.; Prokofyev, A. V.; Scheidelaar, S.; van der Cruijsen, E. A. W.; Dafforn, T. R.; Baldus, M.; Killian, J. A.; Detergent-free isolation, characterization, and functional reconstitution of a tetrameric K^+ channel: The power of native nanodiscs. *Proceedings of the National Academy of Sciences of the United States of America,* **2014,** *111* (52), 18607–18612.

Dörr, J. M.; Scheidelaar, S.; Koorengevel, M. C.; Dominguez, J. J.; Schäfer, M.; van Walree, C. A.; Killian, J. A.; The styrene-maleic acid copolymer: a versatile tool in membrane research. *European Biophysics Journal,* **2016,** *45* (1), 3–21.

Dromigny, J. A.; Rakoto-Alson, O.; Rajaonatahina, D.; Migliani, R.; Ranjalahy, J.; Mauclere, P.; Emergence and rapid spread of tetracycline-resistant Vibrio cholerae strains, Madagascar. *Emerging Infectious Diseases,* **2002,** *8* (3), 336–338.

Du, X.; Li, Y.; Xia, Y.-L.; Ai, S.-M.; Liang, J.; Sang, P.; Ji, X.-L.; Liu, S.-Q.; Insights into protein-ligand interactions: Mechanisms, models, and methods. *International Journal of Molecular Sciences,* **2016,** *17* (2), 144.

Dumoulin, M.; Conrath, K.; van Meirhaeghe, A.; Meersman, F.; Heremans, K.; Frenken, L. G. J.; Muyldermans, S.; Wyns, L.; Matagne, A.; Single-domain antibody fragments with high conformational stability. *Protein Science,* **2002,** *11* (3), 500–515.

Edmond, K.; Clark, A.; Korczak, V. S.; Sanderson, C.; Griffiths, U. K.; Rudan, I.; Global and regional risk of disabling sequelae from bacterial meningitis: a systematic review and meta-analysis. *The Lancet Infectious Disease,* **2010,** *10* (5), 317–328.

Eeman, M.; Deleu, M.; From biological membranes to biomimetic model membranes. *Biotechnologie, Agronomie, Société et Environnement,* **2010,** *14* (4), 719–736.

Efremov, R. G.; Gatsogiannis, C.; Raunser, S.; Lipid Nanodiscs as a Tool for High-Resolution structure determination of membrane proteins by single-particle cryo-EM. *Methods in Enzymology,* **2017,** *594,* 1–30.

Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K.; Features and development of Coot. *Acta Crystallographica Section D, Biological Crystallography,* **2010,** *66,* 486–501.

Engelhardt, H.; Elektronenmikroskopie. In *Bioanalytik,* 3. Auflage; Lottspeich, F., Engels, J. W., Eds.; Springer Spektrum: Berlin, Heidelberg, **2012**; pp 527–563.

Esmaili, M.; Brown, C. J.; Shaykhutdinov, R.; Acevedo-Morantes, C.; Wang, Y. L.; Wille, H.; Gandour, R. D.; Turner, S. R.; Overduin, M.; Homogeneous nanodiscs of native membranes formed by stilbene-maleic-acid copolymers. *Nanoscale,* **2020,** *12* (32), 16705–16709.

Ethayathulla, A. S.; Yousef, M. S.; Amin, A.; Leblanc, G.; Kaback, H. R.; Guan, L.; Structure-based mechanism for Na^+/melibiose symport by MelB. *Nature Communications,* **2014,** *5,* 3009.

Evans, E. G. B.; Morgan, J. L. W.; DiMaio, F.; Zagotta, W. N.; Stoll, S.; Allosteric conformational change of a cyclic nucleotide-gated ion channel revealed by DEER spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America,* **2020,** *117* (20), 10839–10847.

Fägerstam, L. G.; Frostell, A.; Karlsson, R.; Kullman, M.; Larsson, A.; Malmqvist, M.; Butt, H.; Detection of antigen—antibody interactions by surface plasmon resonance, Application to epitope mapping. *Journal of Molecular Recognition,* **1990,** *3* (5/6), 208–214.

Fiori, M. C.; Zheng, W.; Kamilar, E.; Simiyu, G.; Altenberg, G. A.; Liang, H.; Extraction and reconstitution of membrane proteins into lipid nanodiscs encased by zwitterionic styrene-maleic amide copolymers. *Scientific reports,* **2020,** *10,* 9940.

Fischer, M.; Zhang, Q. Y.; Hubbard, R. E.; Thomas, G. H.; Caught in a TRAP: substrate-binding proteins in secondary transport. *Trends in Microbiology,* **2010,** *18* (10), 471–478.

Fischer, M.; Hopkins, A. P.; Severi, E.; Hawkhead, J.; Bawdon, D.; Watts, A. G.; Hubbard, R. E.; Thomas, G. H.; Tripartite ATP-independent periplasmic (TRAP) transporters use an arginine-mediated selectivity filter for high affinity substrate binding. *The Journal of Biological Chemistry,* **2015,** *290* (45), 27113–27123.

Flach, C.-F.; Qadri, F.; Bhuiyan, T. R.; Alam, N. H.; Jennische, E.; Lönnroth, I.; Holmgren, J.; Broad up-regulation of innate defense factors during acute cholera. *Infection and Immunity,* **2007,** *75* (5), 2343–2350.

Fleck, N.; Heubach, C. A.; Hett, T.; Haege, F. R.; Bawol, P. P.; Baltruschat, H.; Schiemann, O.; SLIM: A short-linked, highly redox-stable trityl label for high-sensitivity in-vell EPR distance measurements. *Angewandte Chemie International Edition,* **2020,** *59* (24), 9767–9772.

Fleck, N.; Heubach, C.; Hett, T.; Spicher, S.; Grimme, S.; Schiemann, O.; Ox-SLIM: Synthesis of and site-specific labelling with a highly hydrophilic trityl spin label. *Chemistry - A European Journal,* **2021**, DOI: https://doi.org/10.1002/chem.202100013.

Fleissner, M. R.; Bridges, M. D.; Brooks, E. K.; Cascio, D.; Kálai, T.; Hideg, K.; Hubbell, W. L.; Structure and dynamics of a conformationally constrained nitroxide side chain and applications in EPR spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America,* **2011,** *108* (39), 16241–16246.

Fleissner, M. R.; Brustad, E. M.; Kálai, T.; Altenbach, C.; Cascio, D.; Peters, F. B.; Hideg, K.; Peuker, S.; Schutz, P. G.; Hubbell, W. L.; Site-directed spin labeling of a genetically encoded unnatural amino acid. *Proceedings of the National Academy of Sciences of the United States of America,* **2009,** *106* (51), 21637–21642.

Flores, J. A.; Haddad, B. G.; Dolan, K. A.; Myers, J. B.; Yoshioka, C. C.; Copperman, J.; Zuckerman, D. M.; Reichow, S. L.; Connexin-46/50 in a dynamic lipid environment resolved by CryoEM at 1.9 Å. *Nature Communications,* **2020,** *11,* 4331.

Forrest, L. R.; Tavoulari, S.; Zhang, Y.-W.: Rudnick, G.; Honig, B.; Identification of a chloride ion binding site in Na^+/Cl^--dependent transporters. *Proceedings of the National Academy of Sciences of the United States of America,* **2007,** *104* (31), 12761–12766.

Forrest, L. R.; Krämer, R.; Ziegler, C.; The structural basis of secondary active transport mechanisms. *Biochimica et Biophysica Acta,* **2011,** *1807* (2), 167–188.

Forward, J. A.; Behrendt, M. C.; Wyborn, N. R.; Cross, R. Kelly, D. J.; TRAP transporters: a new family of periplasmic solute transport systems encoded by the dctPQM genes of Rhodobacter capsulatus and by homologs in diverse Gram-negative bacteria. *Journal of Bacteriology,* **1997,** *179* (17), 5482–5493.

Fox, K. L.; Cox, A. D.; Gilbert, M.; Wakarchuk, W. W.; Li, J.; Makepeace, K.; Richards, J. C.; Moxon, E. R.; Hood, D. W.; Identification of a bifunctional lipopolysaccharide sialyltransferase in Haemophilus influenzae: incorporation of disialic acid. *The Journal of Biological Chemistry,* **2006,** *281* (52), 40024–40032.

Frank, J.; Advances in the field of single-particle cryo-electron microscopy over the last decade. *Nature Protocols,* **2017,** *12,* 209–212.

Fulyani, F.; Schuurman-Wolters, G. K.; Zagar, A. V.; Guskov, A.; Slotboom, D.-J.; Poolman, B.; Functional diversity of tandem substrate-binding domains in ABC transporters from pathogenic bacteria. *Structure,* **2013,** *21* (10), 1879–1888.

Galazzo, L.; Meier, G.; Timachi, M. H.; Hutter, C. A. J.; Seeger, M. A.; Bordignon, E.; Spin-labeled nanobodies as protein conformational reporters for electron paramagnetic resonance in cellular membranes. *Proceedings of the National Academy of Sciences of the United States of America,* **2020,** *117* (5), 2441–2448.

Galen, J. E.; Ketley, J. M.; Fasano, A.; Richardson, S. H.; Wasserman, S. S.; Kaper, J. B.; Role of Vibrio cholerae neuraminidase in the function of cholera toxin. *Infection and Immunity,* **1992,** *60* (2), 406–415.

Gallagher, J. R.; Kim, A. J.; Gulati, N. M.; Harris, A. K.; Negative-stain transmission electron microscopy of molecular complexes for image analysis by 2D class averaging. *Current Protocols in Microbiology,* **2019,** *54* (1), e90.

Gangi Setty, T.; Cho, C.; Govindappa, S.; Apicella, M. A.; Ramaswamy, S.; Bacterial periplasmic sialic acid-binding proteins exhibit a conserved binding site. *Acta Crystallographica Section D, Biological Crystallography,* **2014,** *70* (Pt 7), 1801–1811.

Gao, X.; Lu, F.; Zhou, L.; Dang, S.; Sun, L.; Li, X.; Wang, J.; Shi, Y.; Structure and mechanism of an amino acid antiporter. *Science,* **2009,** *324* (5934), 1565–1568.

Gao, Y.; Cao, E.; Julius, D.; Cheng, Y.; TRPV1 structures in nanodiscs reveal mechanisms of ligand and lipid action. *Nature,* **2016,** *534* (7607), 347–351.

Garavito, R. M.; Ferguson-Miller, S.; Detergents as tools in membrane biochemistry. *The Journal of Biological Chemistry,* **2001,** *276* (35), 32403–32406.

Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M. R.; Appel, R. D.; Bairoch, A.; Protein identification and analysis tools on the ExPasy server. In *The Proteomics Protocols Handbook*; Walker, J. M., Editors; Humana Press, Springer Science+Business Media LLC, **2005**; pp 571–607.

Gauto, D. F.; Estrozi, L. F.; Schwieters, C. D.; Effantin, G.; Macek, P.; Sounier, R.; Sivertsen, A. C.; Schmidt, E.; Kerfah, R.; Mas, G.; Colletier, J.-P.; Güntert, P.; Favier, A.; Schoehn, G.; Schanda, P.; Boisbouvier, J.; Integrated NMR and cryo-EM atomic-resolution structure determination of a half-megadalton enzyme complex. *Nature Communications,* **2019,** *10,* 2697.

Gebhardt, C.; Lehmann, M.; Reif, M. M.; Zacharias, M.; Cordes, T.; Molecular and spectroscopic characterization of green and red cyanine fluorophores from the Alexa Fluor and AF series. *BioRxiv,* **2020**. DOI: https://doi.org/10.1101/2020.11.13.381152.

Georgieva, E. R.; Roy, A. S.; Grigoryants, V. M.; Borbat, P. P.; Earle, K. A.; Scholes, C. P.; Freed, J. H.; Effect of freezing conditions on distances and their distributions derived from double electron electron resonance (DEER): a study of doubly-spin-labeled T4 lysozyme. *Journal of Magnetic Resonance,* **2012,** *216,* 69–77.

Gill, D. M.; The arrangement of subunits in cholera toxin. *Biochemistry,* **1976,** *15* (6), 1242–1248.

Glaenzer, J.; Peter, M. F.; Thomas, G. H.; Hagelueken, G.; PELDOR spectroscopy reveals two defined states of a sialic acid TRAP transporter SBP in solution. *Biophysical Journal,* **2017,** *112* (1), 109–120.

Glaenzer, J.; Peter, M. F.; Hagelueken, G.; Studying structure and function of membrane proteins with PELDOR/DEER spectroscopy – The crystallographers' perspective. *Methods,* **2018,** *147,* 163-175.

Glass, R. I.; Hug, I.; Alim, A. R. M. A.; Yunus, M.; Emergence of multiply antibiotic-resistant Vibrio cholerae in Bangladesh. *The Journal of Infectious Diseases,* **1980,** *142* (6), 939–942.

Goldfarb, D., Stoll, S.; *EPR Spectroscopy. Fundamentals and methods,* 1st edition; John Wiley & Sons Ltd, UK, **2018**.

Gonin, S.; Arnoux, P.; Pierru, B.; Lavergne, J.; Alonso, B.; Sabaty, M.; Pignol, D.; 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. *BMC Structural Biology,* **2007,** *7* (11).

Gouridis, G.; Schuurman-Wolters, G. K.; Ploetz, E.; Husada, F.; Vietrov, R.; de Boer, M.; Cordes, T.; Poolman, B.; Conformational dynamics in substrate-binding domains influences transport in the ABC importer GlnPQ. *Nature Structural & Molecular Biology,* **2015,** *22* (1), 57–64.

Grammann, K.; Volke, A.; Kunte, H. J.; New type of osmoregulated solute transporter identified in halophilic members of the bacteria domain: TRAP transporter TeaABC mediates uptake of ectoine and hydroxyectoine in Halomonas elongata DSM 2581^T. *Journal of Bacteriology,* **2002,** *184* (11), 3078–3085.

Granoff, D. M.; Assessing efficacy of Haemophilus influenzae type b combination vaccines. *Clinical Infectious Diseases,* **2001,** *33* (suppl 4), 278–287.

Gulamhussein, A. A.; Uddin, R.; Tighe, B. J.; Poyner, D. R.; Rothnie, A. J.; A comparison of SMA (styrene maleic acid) and DIBMA (di-isobutylene maleic acid) for membrane purification. *Biochimica et Biophysica Acta, Biomembranes,* **2020,** *1862* (7), 183281.

Gupta, K.; Donlan, J. A. C.; Hopper, J. T. S.; Uzdavinys, P.; Landreh, M.; Struwe, W. B.; Drew, D.; Baldwin, A. J.; Stansfeld, P. J.; Robinson, C. V.; The role of interfacial lipids in stabilizing membrane protein oligomers. *Nature,* **2017,** *541*, 421–424.

Gust, A.; Zander, A.; Gietl, A.; Holzmeister, P.; Schulz, S.; Lalkens, B.; Tinnefeld, P.; Grohmann, D.; A starting point for fluorescence-based single-molecule measurements in biomolecular research. *Molecules,* **2014,** *19* (10), 15824–15865.

Haensch, S.; Bianucci, R.; Signoli, M.; Rajerison, M.; Schultz, M.; Kacki, S.; Vermunt, M.; Weston, D. A.; Hurst, D.; Achtman, M.; Carniel, E.; Bramanti, B.; Distinct clones of Yersinia pestis caused the black death. *PLoS Pathogens,* **2010,** *6* (10), e1001134.

Hagelueken, G.; Ward, R.; Naismith, J. H.; Schiemann, O.; MtsslWizard: In silico spin-labeling and generation of distance distributions in PyMOL. *Applied Magnetic Resonance,* **2012,** *42*, 377–391.

Hagelueken, G.; Abdullin, D.; Schiemann, O.; mtsslSuite: Probing biomolecular conformation by spin-labeling studies. *Methods in Enzymology,* **2015,** *563*, 595–622.

Hanke, L.; Vidakovics Perez, L.; Sheward, D. J.; Das, H.; Schulte, T.; Moliner-Morro, A.; Corcoran, M.; Achour, A.; Karlsson Hedestam, G. B.; Hällberg, B. M.; Murrell, B.; McInerney, G. M.; An alpaca nanobody neutralizes SARS-CoV-2 by blocking receptor interaction. *Nature Communications,* **2020,** *11*, 4420.

Harris, J. B.; LaRocque, R. C.; Qadri, F.; Ryan, E. T.; Calderwood, S. B.; Cholera. *The Lancet,* **2012,** *379* (9835), 2466–2476.

Hellenkamp, B.; Schmid, S.; Doroshenko, O.; Opanasyuk, O.; Kühnemuth, R.; Rezaei Adariani, S.; Ambrose, B.; Aznauryan, M.; Barth, A.; Birkedal, V.; Bowen, M. E.; Chen, H.; Cordes, T.; Eilert, T.; Fijen, C.; Gebhardt, C.; Götz, M.; Gouridis, G.; Gratton, E.; Ha, T.; Hao, P.; Hanke, C. A.; Hartmann, A.; Hendrix, J.; Hildebrandt, L. L.; Hirschfeld, V.; Hohlbein, J.; Hua, B.; Hübner, C. G.; Kallis, E.; Kapanidis, A. N.; Kim, J.-Y.; Krainer, G.; Lamb, D. C.; Lee, N. K.; Lemke, E. A.; Levesque, B.; Levitus, M.; McCann, J. J.; Naredi-Rainer, N.; Nettels, D.; Ngo, T.; Qiu, R.; Robb, N. C.; Röcker, C.; Sanabria, H.; Schlierf, M.; Schröder, T.; Schuler, B.; Seidel, H.; Streit, L.; Thurn, J.; Tinnefeld, P.; Tyagi, S.; Vandenberk, N.; Vera, A. M.; Weninger, K. R.; Wünsch, B.; Yanez-Orozco, I. S.; Michaelis, J.; Seidel, C. A. M.; Craggs, T. D.; Hugel, T.; Precision and accuracy of single-molecule FRET measurements-a multi-laboratory benchmark study. *Nature Methods,* **2018,** *15* (9), 669–676.

Hellmich, U. A.; Lyubenova, S.; Kaltenborn, E.; Doshi, R.; van Veen, H. W.; Prisner, T. F.; Glaubitz, C.; Probing the ATP hydrolysis cycle of the ABC multidrug transporter LmrA by pulsed EPR spectroscopy. *Journal of the American Chemical Society,* **2012,** *134* (13), 5857–5862.

Henderson, R.; Unwin, P. N. T.; Three-dimensional model of purple membrane obtained by electron microscopy. *Nature,* **1975,** *257*, 28–32.

Hendrickson, W. A.; Atomic-level analysis of membrane-protein structure. *Nature Structural & Molecular Biology,* **2016,** *23* (6), 464–467.

Herzik, M. A. Jr.; Wu, M.; Lander, G. C.; High-resolution structure determination of sub-100 kDa complexes using conventional cryo-EM. *Nature Communications,* **2019,** *10*, 1032.

Hetherington, S. V.; Patrick, C. C.; Hansen, E. J.; Outer membrane protein binding sites of complement component 3 during opsonization of Haemophilus influenzae. *Infection and Immunity,* **1993,** *61* (12), 5175-5163.

Hofmann, S.; Januliene, D.; Mehdipour, A. R.; Thomas, C.; Stefan, E.; Brüchert, S.; Kuhn, B. T.; Geertsma, E. R.; Hummer, G.; Tampé, R.; Moeller, A.; Conformation space of a heterodimeric ABC exporter under turnover conditions. *Nature,* **2019,** *571* (7766), 580–583.

Hollenstein, K.; Frei, D. C.; Locher, K. P.; Structure of an ABC transporter in complex with its binding protein. *Nature,* **2007,** *446* (7132), 213–216.

Hood, D. W.; Cox, A. D.; Gilbert, M.; Makepeace, K.; Walsh, S.; Deadman, M. E.; Cody, A.; Martin, A.; Mansson, M.; Schweda, E. K. H.; Brisson, J.-R.; Richards, J. C.; Moxon, E. R.; Wakarchuk, W. W.; Identification of a lipopolysaccharide α-2,3-sialyltransferase from Haemophilus influenzae. *Molecular Microbiology,* **2001,** *39* (2), 341–350.

Hood, D. W.; Makepeace, K.; Deadman, M. E.; Rest, F. R.; Thibault, P.; Martin, A.; Richards, J. C.; Moxon, E. R.; Sialic acid in the lipopolysaccharide of Haemophilus influenzae: Strain distribution, influence on serum resistance and structural characterization. *Molecular Microbiology,* **1999,** *33* (4), 679–692.

Hood, D. W.; Randle, G.; Cox, A. D.; Makepeace, K.; Li, J.; Schweda, E. K. H.; Richards, J. C.; Moxon, E. R.; Biosynthesis of cryptic lipopolysaccharide glycoforms in Haemophilus influenzae involves a mechanism similar to that required for O-antigen synthesis. *Journal of Bacteriology,* **2004,** *186* (21), 7429–7439.

Hu, D.; Liu, B.; Feng, L.; Ding, P.; Guo, X.; Wang, M.; Cao, B.; Reeves, P. R.; Wang, L.; Origins of the current seventh cholera pandemic. *Proceedings of the National Academy of Sciences of the United States of America,* **2016,** *113* (48), E7730-E7739.

Hubbell, W. L.; Mchaourab, H. S.; Altenbach, C.; Lietzow, M. A.; Watching proteins move using site-directed spin labeling. *Structure,* **1996,** *4* (7), 779–783.

Hughes, L. D.; Rawle, R. J.; Boxer, S. G.; Choose your label wisely: water-soluble fluorophores often interact with lipid bilayers. *PloS One,* **2014,** *9* (2), e87649.

Hunte, C.; Screpanti, E.; Venturi, M.; Rimon, A.; Padan, E.; Michel, H.; Structure of a Na^+/H^+ antiporter and insights into mechanism of action and regulation by pH. *Nature,* **2005,** *435* (7046), 1197–1202.

Huo, J.; Le Bas, A.; Ruza, R. R.; Duyvesteyn, H. M. E.; Mikolajek, H.; Malinauskas, T.; Tan, T. K.; Rijal, P.; Dumoux, M.; Ward, P. N.; Ren, J.; Zhou, D.; Harrison, P. J.; Weckener, M.; Clare, D. K.; Vogirala, V. K.; Radecke, J.; Moynié, L.; Zhao, Y.; Gilbert-Jaramillo, J.; Knight, M. L.; Tree, J. A.; Buttigieg, K. R.; Coombes, N.; Elmore, M. J.; Carroll, M. W.; Carrique, L.; Shah, P. N. M.; James, W.; Townsend, A. R.; Stuart, D. I.; Owens, R. J.; Naismith, J. H.; Neutralizing nanobodies bind SARS-CoV-2 spike RBD and block interaction with ACE2. *Nature Structural & Molecular Biology,* **2020,** *27* (9), 846–854.

Hussack, G.; Arbabi-Ghahroudi, M.; van Faassen, H.; Songer, J. G.; Ng, K. K.-S.; MacKenzie, R.; Tanha, J.; Neutralization of Clostridium difficile toxin A with single-domain antibodies targeting the cell receptor binding domain. *The Journal of Biological Chemistry,* **2011a,** *286* (11), 8961–8976.

Hussack, G.; Hirama, T.; Ding, W.; MacKenzie, R.; Tanha, J.; Engineered single-domain antibodies with high protease resistance and thermal stability. *PloS One,* **2011b,** *6* (11), e28218.

Hvorup, R. N.; Goetz, B. A.; Niederer, M.; Hollenstein, K.; Perozo, E.; Locher, K. P.; Asymmetry in the structure of the ABC transporter-binding protein complex BtuCD-BtuF. *Science,* **2007,** *317* (5843), 1387–1390.

Igarashi, R.; Sakai, T.; Hara, H.; Tenno, T.; Tanaka, T.; Tochio, H.; Shirakawa, M.; Distance determination in proteins inside Xenopus laevis oocytes by double electron-electron resonance experiments. *Journal of the American Chemical Society,* **2010,** *132* (24), 8228–8229.

Islam, M. S.; Mahmud, Z. H.; Ansaruzzaman, M.; Faruque, S. M.; Talukder, K. A.; Qadri, F.; Alam, M.; Islam, S.; Bardhan, P. K.; Mazumder, R. N.; Khan, A. I.; Ahmed, S.; Iqbal, A.; Chitsatso, O.; Mudzori, J.; Patel, S.; Midzi, S. M.; Charimari, L.; Endtz, H. P.; Cravioto, A.; Phenotypic, genotypic, and antibiotic sensitivity patterns of strains isolated from the cholera epidemic in Zimbabwe. *Journal of Clinical Microbiology,* **2011,** *49* (6), 2325–2327.

Jamshad, M.; Grimard, V.; Idini, I.; Knowles, T. J.; Dowle, M. R.; Schofield, N.; Sridhar, P.; Lin, Y.; Finka, R.; Wheatley, M.; Thomas, O. R. T.; Palmer, R. E.; Overduin, M.; Govaerts, C.; Ruysschaert, J.-M.; Edler, K. J.; Dafforn, T. R.; Structural analysis of a nanoparticle containing a lipid bilayer used for detergent-free extraction of membrane proteins. *Nano Research,* **2015,** *8* (3), 774-789.

Jassoy, J. J.; Berndhäuser, A.; Duthie, F.; Kühn, S. P.; Hagelueken, G.; Schiemann, O.; Versatile trityl spin labels for nanometer distance measurements on biomolecules in vitro and within cells. *Angewandte Chemie International Edition,* **2017,** *56* (1), 177-181.

Jennings, H. J.; Capsular polysaccharides as human vaccines. *Advances in Carbohydrate Chemistry and Biochemistry.* **1983,** *41*, 155–208.

Jermyn, W. S.; Boyd, E. F.; Characterization of a novel Vibrio pathogenicity island (VPI-2) encoding neuraminidase among toxigenic Vibrio cholerae isolates. *Microbiology,* **2002,** *148* (11), 3681–3693.

Jeschke, G.; DEER distance measurements on proteins. *Annual Review of Physical Chemistry,* **2012,** *63*, 419–446.

Jeschke, G.; Conformational dynamics and distribution of nitroxide spin labels. *Progress in Nuclear Magnetic Resonance Spectroscopy,* **2013,** *72*, 42–60.

Jeschke, G.; Dipolar Spectroscopy - Double-resonance methods. In *EPR Spectroscopy: Fundamentals and methods,* 1st edition; Goldfarb, D., Stoll, S., Editors; John Wiley & Sons: UK, **2018a**; pp 401–423.

Jeschke, G.; MMM: A toolbox for integrative structure modeling. *Protein science,* **2018b,** *27* (1), 76–85.

Jeschke, G.; The contribution of modern EPR to structural biology. *Emerging Topics in Life Sciences,* **2018c,** *2* (1), 9–18.

Jeschke, G.; Chechik, V.; Ionita, P.; Godt, A.; Zimmermann, H.; Banham, J.; Timmel, C. R.; Hilger, D.; Jung, H.; DeerAnalysis2006 -a comprehensive software package for analyzing pulsed ELDOR data. *Applied Magnetic Resonance,* **2006,** *30,* 473–498.

Jeschke, G.; Polyhach, Y.; Distance measurements on spin-labelled biomacromolecules by pulsed electron paramagnetic resonance. *Physical Chemistry Chemical Physics,* **2007,** *9* (16), 1895–1910.

Johnston, J. W.; Coussens, N. P.; Allen, S.; Houtman, J. C. D.; Turner, K. H.; Zaleski, A.; Ramaswamy, S.; Gibson, B. W.; Apicella, M. A.; Characterization of the N-acetyl-5-neuraminic acid-binding site of the extracytoplasmic solute receptor (SiaP) of nontypeable Haemophilus influenzae strain 2019. *The Journal of Biological Chemistry,* **2008,** *283* (2), 855–865.

Jones, P. A.; Samuels, N. M.; Phillips, N. J.; Munson, R. S.; Bozue, J. A.; Arseneau, J. A.; Nichols, W. A.; Zaleski, A.; Gibson, B. W.; Apicella, M. A.; Haemophilus influenzae type b strain A2 has multiple sialyltransferases involved in lipooligosaccharide sialylation. *The Journal of Biological Chemistry,* **2002,** *277* (17), 14598–14611.

Joseph, B.; Korkhov, V. M.; Yulikov, M.; Jeschke, G.; Bordignon, E.; Conformational cycle of the vitamin B_{12} ABC importer in liposomes detected by double electron-electron resonance (DEER). *The Journal of Biological Chemistry,* **2014,** *289* (6), 3176–3185.

Joseph, B.; Sikora, A.; Bordignon, E.; Jeschke, G.; Cafiso, D. S.; Prisner, T. F.; Distance measurement on an endogenous membrane transporter in E. coli cells and native membranes using EPR spectroscopy. *Angewandte Chemie International Edition,* **2015,** *54* (21), 6196–6199.

Kabsch, W.; XDS. *Acta Crystallographica Section D, Biological Crystallography,* **2010,** *66* (Pt 2), 125–132.

Kalinin, S.; Peulen, T.; Sindbert, S.; Rothwell, P. J.; Berger, S.; Restle, T.; Goody, R. S.; Gohlke, H.; Seidel, C. A. M.; A toolkit and benchmark study for FRET-restrained high-precision structural modeling. *Nature Methods,* **2012,** *9* (12), 1218–1225.

Kantardjieff, K. A.; Rupp, B.; Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Science,* **2003,** *12,* 1865–1871.

Kapanidis, A. N.; Lee, N. K.; Laurence, T. A.; Doose, S.; Margeat, E.; Weiss, S.; Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules. *Proceedings of the National Academy of Sciences of the United States of America,* **2004,** *101* (24), 8936–8941.

Kapsalis, C.; Wang, B.; El Mkami, H.; Pitt, S. J.; Schnell, J. R.; Smith, T. K.; Lippiat, J. D.; Bode, B. E.; Pliotas, C.; Allosteric activation of an ion channel triggered by modification of mechanosensitive nano-pockets. *Nature Communications,* **2019,** *10,* 4619.

Karaolis, D. K. R.; Somara, S.; Maneval Jr, D. R.; Johnson, J. A.; Kaper, J. B.; A bacteriophage encoding a pathogenicity island, a type-IV pilus and a phage receptor in cholera bacteria. *Nature,* **1999,** *399,* 375–379.

Karlsson, R.; Katsamba, P. S.; Nordin, H.; Pol, E.; Myszka, D. G.; Analyzing a kinetic titration series using affinity biosensors. *Analytical Biochemistry,* **2006,** *349* (1), 136–147.

Kelly, D. J.; Thomas, G. H.; The tripartite ATP-independent periplasmic (TRAP) transporters of bacteria and archaea. *FEMS Microbiology Reviews,* **2001,** *25* (4), 405–424.

Kimura, A.; Hansen, E. J.; Antigenic and phenotypic variations of Haemophilus influenzae type b lipopolysaccharide and their relationship to virulence. *Infection and Immunity,* **1986,** *51* (1), 69–79.

Klare, J. P.; Steinhoff, H.-J.; Spin labeling EPR. *Photosynthesis Research,* **2009,** *102* (2-3), 377–390.

Klein, E.; Smith, D. L.; Laxminarayan, R.; Hospitalizations and deaths caused by methicillin-resistant Staphylococcus aureus, United States, 1999-2005. *Emerging Infectious Diseases,* **2007,** *13* (12), 1840–1846.

Klein, E. Y.; Sun, L.; Smith, D. L.; Laxminarayan, R.; The changing epidemiology of methicillin-resistant Staphylococcus aureus in the United States: a national observational study. *American Journal of Epidemiology,* **2013,** *177* (7), 666–674.

Knowles, T. J.; Finka, R.; Smith, C.; Lin, Y.-P.; Dafforn, T.; Overduin, M.; Membrane proteins solubilized intact in lipid containing nanoparticles bounded by styrene maleic acid copolymer. *Journal of the American Chemical Society,* **2009,** *131* (22), 7484–7485.

Koch, K.; Kalusche, S.; Torres, J. L.; Stanfield, R. L.; Danquah, W.; Khazanehdari, K.; Briesen, H. von; Geertsma, E. R.; Wilson, I. A.; Wernery, U.; Koch-Nolte, F.; Ward, A. B.; Dietrich, U.; Selection of nanobodies with broad neutralizing potential against primary HIV-1 strains using soluble subtype C gp140 envelope trimers. *Scientific Reports,* **2017,** *7* (1), 8390.

Koenig, P.-A.; Das, H.; Liu, H.; Kümmerer, B. M.; Gohr, F. N.; Jenster, L.-M.; Schiffelers, L. D. J.; Tesfamariam, Y. M.; Uchima, M.; Wuerth, J. D.; Gatterdam, K.; Ruetalo, N.; Christensen, M. H.; Fandrey, C. I.; Normann, S.; Tödtmann, J. M. P.; Pritzl, S.; Hanke, L.; Boos, J.; Yuan; Zhu, X.; Schmid-Burgk, J. L.; Kato, H.; Schindler, M.; Wilson, I. A.; Geyer, M.; Ludwig, K. U.; Hällberg, B. M.; Wu, N. C.; Schmidt, F. I.; Structure-guided multivalent nanobodies block SARS-CoV-2 infection and suppress mutational escape. *Science,* **2021**. DOI: 10.1126/science.abe6230.

Kolker, E.; Makarova, K. S.; Shabalina, S.; Picone, A. F.; Purvine, S.; Holzman, T.; Cherny, T.; Armbruster, D.; Munson, R. S.; Kolesov, G.; Frishman, D.; Galperin, M. Y.; Identification and functional analysis of 'hypothetical' genes expressed in Haemophilus influenzae. *Nucleic Acids Research,* **2004,** *32* (8), 2353–2361.

Kopf, A. H.; Dörr, J. M.; Koorengevel, M. C.; Antoniciello, F.; Jahn, H.; Killian, J. A.; Factors influencing the solubilization of membrane proteins from Escherichia coli membranes by styrene-maleic acid copolymers. *Biochimica et Biophysica Acta – Biomembranes,* **2020,** *1862* (2), 183125.

Korkhov, V. M.; Mireku, S. A.; Locher, K. P.; Structure of AMP-PNP-bound vitamin B_{12} transporter BtuCD-F. *Nature,* **2012,** *490* (7420), 367–372.

Krainer, G.; Hartmann, A.; Schlierf, M.; farFRET: Extending the range in single-molecule FRET experiments beyond 10 nm. *Nano Letters,* **2015,** *15* (9), 5826–5829.

Krishnamurthy, H.; Piscitelli, C. L.; Gouaux, E.; Unlocking the molecular secrets of sodium-coupled transporters. *Nature,* **2009,** *459* (7245), 347–355.

Kucher, S.; Korneev, S.; Tyagi, S.; Apfelbaum, R.; Grohmann, D.; Lemke, E. A.; Klare, J. P.; Steinhoff, H.-J.; Klose, D.; Orthogonal spin labeling using click chemistry for in vitro and in vivo applications. *Journal of Magnetic Resonance,* **2017,** *275,* 38–45.

Kuhlmann, S. I.; Terwisscha van Scheltinga, A. C.; Bienert, R.; Kunte, H.-J.; Ziegler, C.; 1.55 Å structure of the ectoine binding protein TeaA of the osmoregulated TRAP-transporter TeaABC from Halomonas elongata. *Biochemistry,* **2008,** *47* (36), 9475–9485.

Laganowsky, A.; Reading, E.; Allison, T. M.; Ulmschneider, M. B.; Degiacomi, M. T.; Baldwin, A. J.; Robinson, C. V.; Membrane proteins bind lipids selectively to modulate their structure and function. *Nature,* **2014,** *510* (7503), 172–175.

Lakowicz, J. R.; *Principles of Fluorescence Spectroscopy*, 3rd Edition; Springer: New York, **2006**.

Lambert, M.-L.; Suetens, C.; Savey, A.; Palomar, M.; Hiesmayr, M.; Morales, I.; Agodi, A.; Frank, U.; Mertens, K.; Schumacher, M.; Wolkewitz, M.; Clinical outcomes of health-care-associated infections and antimicrobial resistance in patients admitted to European intensive-care units: a cohort study. *The Lancet Infectious Diseases,* **2011,** *11* (1), 30–38.

Laskowski, R. A.; Swindells, M. B.; LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *Journal of Chemical Information and Modeling,* **2011,** *51* (10), 2778–2786.

Laverty, D.; Desai, R.; Uchański, T.; Masiulis, S.; Stec, W. J.; Malinauskas, T.; Zivanov, J.; Pardon, E.; Steyaert, J.; Miller, K. W.; Aricescu, A. R.; Cryo-EM structure of the human α1β3γ2 GABA_A receptor in a lipid bilayer. *Nature,* **2019,** *565* (7740), 516–520.

Lecher, J.; Pittelkow, M.; Zobel, S.; Bursy, J.; Bönig, T.; Smits, S. H. J.; Schmitt, L.; Bremer, E.; The crystal structure of UehA in complex with ectoine-A comparison with other TRAP-T binding proteins. *Journal of Molecular Biology,* **2009,** *389* (1), 58–73.

Lee, S. C.; Knowles, T. J.; Postis, V. L. G.; Jamshad, M.; Parslow, R. A.; Lin, Y.-P.; Goldman, A.; Sridhar, P.; Overduin, M.; Muench, S. P.; Dafforn, T. R.; A method for detergent-free isolation of membrane proteins in their local lipid environment. *Nature Protocols,* **2016a,** *11* (7), 1149–1162.

Lee, T. C.; Kang, M.; Kim, C. H.; Schultz, P. G.; Chapman, E.; Deniz, A. A.; Dual unnatural amino acid incorporation and click-chemistry labeling to enable single-molecule FRET studies of p97 folding. *Chembiochem,* **2016b,** *17* (11), 981–984.

Lee, H.; Lim, S. in; Shin, S.-H.; Lim, Y.; Koh, J. W.; Yang, S.; Conjugation of cell-penetrating peptides to antimicrobial peptides enhances antibacterial activity. *ACS Omega,* **2019,** *4* (13), 15694–15701.

Lewinson, O.; Lee, A. T.; Locher, K. P.; Rees, D. C.; A distinct mechanism for the ABC transporter BtuCD-BtuF revealed by the dynamics of complex formation. *Nature Structural & Molecular Biology,* **2010,** *17* (3), 332–338.

Li, M.; Kotetishvili, M.; Chen, Y.; Sozhamannan, S.; Comparative genomic analyses of the vibrio pathogenicity island and cholera toxin prophage regions in nonepidemic serogroup strains of Vibrio cholerae. *Applied and Environmental Microbiology,* **2003,** *69* (3), 1728–1738.

Licht, A.; Schneider, E.; ATP binding cassette systems: structures, mechanisms, and functions. *Central European Journal of Biology,* **2011,** *6* (5), 785–801.

Liebschner, D.; Afonine, P. V.; Baker, M. L.; Bunkóczi, G.; Chen, V. B.; Croll, T. I.; Hintze, B.; Hung, L. W.; Jain, S.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R. D.; Poon, B. K.; Prisant, M. G.; Read, R. J.; Richardson, J. S.; Richardson, D. C.; Sammito, M. D.; Sobolev, O. V.; Stockwell, D. H.; Terwilliger, T. C.; Urzhumtsev, A. G.; Videau, L. L.; Williams, C. J.; Adams, P. D.; Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallographica Section D, Structural Biology,* **2019,** *75* (Pt 10), 861–877.

Liedberg, B.; Nylander, C.; Lundstrom, I.; Surface plasmon resonance for gas detection and biosensing. *Sensors and Actuators,* **1983,** *4*, 299–304.

Lillington, J. E. D.; Lovett, J. E.; Johnson, S.; Roversi, P.; Timmel, C. R.; Lea, S. M.; Shigella flexneri Spa15 crystal structure verified in solution by double electron electron resonance. *Journal of Molecular Biology,* **2011,** *405* (2), 427–435.

Liu, H.; Naismith, J. H.; An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnology,* **2008,** *8* (91).

Locher, K. P.; Mechanistic diversity in ATP-binding cassette (ABC) transporters. *Nature Structural & Molecular Biology,* **2016,** *23* (6), 487–493.

Lodish, H., Berk, A., Zipursky, S. L. et al.; Molecular Cell Biology. 4th edition. New York: W. H. Freeman; **2000**. Section 5.3, Biomembranes: Structural Organization and Basic Functions. https://www.ncbi.nlm.nih.gov/books/NBK21583/ (accessed 02.11.2020).

Madeira, F.; Park, Y. M.; Lee, J.; Buso, N.; Gur, T.; Madhusoodanan, N.; Basutkar, P.; Tivey, A. R. N.; Potter, S. C.; Finn, R. D.; Lopez, R.; The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research,* **2019,** *47* (W1), W636-W641.

Mandrell, R. E.; McLaughlin, R.; Kwaik, Y. A.; Lesse, A.; Yamasaki, R.; Gibson, B.; Spinola, S. M.; Apicella, M. A.; Lipooligosaccharides (LOS) of some Haemophilus species mimic human glycosphingolipids, and some LOS are sialylated. *Infection and Immunity,* **1992,** *60* (4), 1322–1328.

Mao, B.; Pear, M. R.; McCammon, J. A., Quiocho, F. A.; Hinge-bending in L-arabinose binding protein. The “Venus’s-flytrap” model. *The Journal of Biological Chemistry,* **1982,** *257* (3), 1131–1133.

Marconnet, A.; Michon, B.; Le Bon, C.; Giusti, F.; Tribet, C.; Zoonens, M.; Solubilization and stabilization of membrane proteins by cycloalkane-modified amphiphilic polymers. *Biomacromolecules,* **2020,** *21* (8), 3459–3467.

Marion, C.; Aten, A. E.; Woodiga, S. A.; King, S. J.; Identification of an ATPase, MsmK, which energizes multiple carbohydrate ABC transporters in Streptococcus pneumoniae. *Infection and Immunity,* **2011,** *79* (10), 4193–4200.

Martens, C.; Stein, R. A.; Masureel, M.; Roth, A.; Mishra, S.; Dawaliby, R.; Konijnenberg, A.; Sobott, F.; Govaerts, C.; Mchaourab, H. S.; Lipids modulate the conformational dynamics of a secondary multidrug transporter. *Nature Structural & Molecular Biology,* **2016,** *23* (8), 744–751.

Martens, C.; Shekhar, M.; Borysik, A. J.; Lau, A. M.; Reading, E.; Tajkhorshid, E.; Booth, P. J.; Politis, A.; Direct protein-lipid interactions shape the conformational landscape of secondary transporters. *Nature Communications,* **2018,** *9*, 4151.

Masoud, H.; Moxon, E. R.; Martin, A.; Krajcarski, D.; Richards, J. C.; Structure of the variable and conserved lipopolysaccharide oligosaccharide epitopes expressed by Haemophilus influenzae serotype b strain Eagan. *Biochemistry,* **1997,** *36* (8), 2091–2103.

Matthews, B. W.; Solvent content of protein crystals. *Journal of Molecular Biology,* **1968,** *33*, 491–497.

McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J.; Phaser crystallographic software. *Journal of Applied Crystallography,* **2007,** *40* (Pt 4), 658–674.

Mchaourab, H. S.; Steed, P. R.; Kazmier, K.; Toward the fourth dimension of membrane protein structure: insight into dynamics from spin-labeling EPR spectroscopy. *Structure,* **2011,** *19* (11), 1549–1561.

Meinert, C.; Senger, J.; Witthohn, M.; Wübbeler, J. H.; Steinbüchel, A.; Carbohydrate uptake in Advenella mimigardefordensis strain DPN7T is mediated by periplasmic sugar oxidation and a TRAP-transport system. *Molecular Microbiology,* **2017,** *104* (6), 916–930.

Merianos, H. J.; Cadieux, N.; Lin, C. H.; Kadner, R. J.; Cafiso, D. S.; Substrate-induced exposure of an energy-coupling motif of a membrane transporter. *Nature Structural Biology,* **2000,** *7,* 205–209.

Merk, A.; Bartesaghi, A.; Banerjee, S.; Falconieri, V.; Rao, P.; Davis, M. I.; Pragani, R.; Boxer, M. B.; Earl, L. A.; Milne, J. L. S.; Subramaniam, S.; Breaking cryo-EM resolution barriers to facilitate drug discovery. *Cell,* **2016,** *165* (7), 1698–1707.

Mireku, S. A.; Sauer, M. M.; Glockshuber, R.; Locher, K. P.; Structural basis of nanobody-mediated blocking of BtuF, the cognate substrate-binding protein of the Escherichia coli vitamin B12 transporter BtuCD. *Scientific reports,* **2017,** *7* (1), 14296.

Mohanty, A. K.; Simmons, C. R.; Wiener, M. C.; Inhibition of tobacco etch virus protease activity by detergents. *Protein Expression and Purification,* **2003,** *27* (1), 109–114.

Morris, J. G. Jr; Cholera and other types of vibriosis: a story of human pandemics and oysters on the half shell. *Clinical Infectious Diseases,* **2003,** *37,* 272–280.

Moxon, E. R.; Deich, R. A.; Connelly, C.; Cloning of chromosomal DNA from Haemophilus influenzae: Its use for studying the expression of type b capsule and virulence. *The Journal of Clinical Investigation,* **1984,** *73* (2), 298–306.

Müller, A.; Severi, E.; Mulligan, C.; Watts, A. G.; Kelly, D. J.; Wilson, K. S.; Wilkinson, A. J.; Thomas, G. H.; Conservation of structure and mechanism in primary and secondary transporters exemplified by SiaP, a sialic acid binding virulence factor from Haemophilus influenzae. *The Journal of Biological Chemistry,* **2006,** *281* (31), 22212–22222.

Mulligan, C.; Kelly, D. J.; Thomas, G. H.; Tripartite ATP-independent periplasmic transporters: application of a relational database for genome-wide analysis of transporter gene frequency and organization. *Journal of molecular microbiology and biotechnology,* **2007,** *12* (3-4), 218–226.

Mulligan, C.; Geertsma, E. R.; Severi, E.; Kelly, D. J.; Poolman, B.; Thomas, G. H.; The substrate-binding protein imposes directionality on an electrochemical sodium gradient-driven TRAP transporter. *Proceedings of the National Academy of Sciences of the United States of America,* **2009,** *106* (20), 1778–1783.

Mulligan, C.; Fischer, M.; Thomas, G. H.; Tripartite ATP-independent periplasmic (TRAP) transporters in bacteria and archaea. *FEMS Microbiology Reviews,* **2011,** *35* (1), 68–86.

Mulligan, C.; Leech, A. P.; Kelly, D. J.; Thomas, G. H.; 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. *The Journal of Biological Chemistry,* **2012,** *287* (5), 3598–3608.

Muschielok, A.; Andrecka, J.; Jawhari, A.; Brückner, F.; Cramer, P.; Michaelis, J.; A nanopositioning system for macromolecular structural analysis. *Nature Methods,* **2008,** *5* (11), 965–971.

Muyldermans, S.; Nanobodies: natural single-domain antibodies. *Annual Review of Biochemistry,* **2013,** *82,* 775–797.

Nakane, T.; Kotecha, A.; Sente, A.; McMullan, G.; Masiulis, S.; Brown, P. M. G. E.; Grigoras, I. T.; Malinauskaite, L.; Malinauskas, T.; Miehling, J.; Uchański, T.; Yu, L.; Karia, D.; Pechnikova, E. V.; Jong, E. de; Keizer, J.; Bischoff, M.; McCormack, J.; Tiemeijer, P.; Hardwick, S. W.; Chirgadze, D. Y.; Murshudov, G.; Aricescu, A. R.; Scheres, S. H. W.; Single-particle cryo-EM at atomic resolution. *Nature,* **2020,** *587* (7832), 152–156.

NejatyJahromy, Y.; Schubert, E.; Demystifying EPR: A rookie guide to the application of electron paramagnetic resonance spectroscopy on biomolecules. *Progress in Biological Sciences*, **2014,** *4* (2), 133–152.

Newby, Z. E. R.; O'Connell, J. D.; Gruswitz, F.; Hays, F. A.; Harries, W. E. C.; Harwood, I. M.; Ho, J. D.; Lee, J. K.; Savage, D. F.; Miercke, L. J. W.; Stroud, R. M.; A general protocol for the crystallization of membrane proteins for X-ray structural investigation. *Nature Protocols,* **2009,** *4* (5), 619–637.

Nguyen, D. P.; Lusic, H.; Neumann, H.; Kapadnis, P. B.; Deiters, A.; Chin, J. W.; Genetic encoding and labeling of aliphatic azides and alkynes in recombinant proteins via a pyrrolysyl-tRNA synthetase/tRNA$_{CUA}$ pair and click chemistry. *Journal of the American Chemical Society,* **2009,** *131* (25), 8720–8721.

Nguyen, H. H.; Park, J.; Kang, S.; Kim, M.; Surface plasmon resonance: a versatile technique for biosensor applications. *Sensors,* **2015,** *15* (5), 10481–10510.

Nicolaou, K. C.; Rigol, S.; A brief history of antibiotics and select advances in their synthesis. *The Journal of Antibiotics,* **2018,** *71* (2), 153–184.

Nobel Media AB **2020**. All Nobel Prizes in Chemistry. https://www.nobelprize.org/prizes/lists/all-nobel-prizes-in-chemistry/ (accessed 18.11. 2020).

Nogales, E.; The development of cryo-EM into a mainstream structural biology technique. *Nature Methods,* **2016,** *13* (1), 24–27.

Norcross, S.; Sunderraj, A.; Tantama, M.; pH- and temperature-dependent peptide binding to the Lactococcus lactis oligopeptide-binding protein A measured with a fluorescence anisotropy assay. *ACS Omega,* **2019,** *4* (2), 2812–2822.

North, R. A.; Horne, C. R.; Davies, J. S.; Remus, D. M.; Muscroft-Taylor, A. C.; Goyal, P.; Wahlgren, W. Y.; Ramaswamy, S.; Friemann, R.; Dobson, R. C. J.; "Just a spoonful of sugar…": import of sialic acid across bacterial cell membranes. *Biophysical Reviews,* **2018,** *10* (2), 219–227.

Oberai, A.; Ihm, Y.; Kim, S.; Bowie, J. U.; A limited universe of membrane protein families and folds. *Protein science,* **2006,** *15* (7), 1723–1734.

Oesterhelt, F.; Protein-Protein-Wechselwirkung: Plasmonenspektroskopie (surface plasmon resonance). In *Bioanalytik,* 3. Auflage; Lottspeich, F., Engels, J. W., Editors.; Springer Spektrum: Berlin, Heidelberg, 2012; pp 443–445.

Oikawa, H.; Takahashi, T.; Kamonprasertsuk, S.; Takahashi, S.; Microsecond resolved single-molecule FRET time series measurements based on the line confocal optical system combined with hybrid photodetectors. *Physical chemistry chemical physics,* **2018,** *20* (5), 3277–3285.

Okamoto, K.; Hibino, K.; Sako, Y.; In-cell single-molecule FRET measurements reveal three conformational state changes in RAF protein. *Biochimica et Biophysica Acta. General subjects,* **2020,** *1864* (2), 129358.

Oldham, M. L.; Khare, D.; Quiocho, F. A.; Davidson, A. L.; Chen, J.; Crystal structure of a catalytic intermediate of the maltose transporter. *Nature,* **2007,** *450* (7169), 515–521.

Oluwole, A. O.; Danielczak, B.; Meister, A.; Babalola, J. O.; Vargas, C.; Keller, S.; Solubilization of membrane proteins into functional lipid-bilayer nanodiscs using a diisobutylene/maleic acid copolymer. *Angewandte Chemie International Edition,* **2017a,** *56* (7), 1919–1924.

Oluwole, A. O.; Klingler, J.; Danielczak, B.; Babalola, J. O.; Vargas, C.; Pabst, G.; Keller, S.; Formation of lipid-bilayer nanodiscs by diisobutylene/maleic acid (DIBMA) copolymer. *Langmuir,* **2017b,** *33* (50), 14378–14388.

Omasits, U.; Ahrens, C. H.; Müller, S.; Wollscheid, B.; Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics,* **2014,** *30* (6), 884–886.

Orwick-Rydmark, M.; Lovett, J. E.; Graziadei, A.; Lindholm, L.; Hicks, M. R.; Watts, A.; Detergent-free incorporation of a seven-transmembrane receptor protein into nanosized bilayer lipodisq particles for functional and biophysical studies. *Nano Letters,* **2012,** *12* (9), 4687–4692.

O’Shannessy, D.; Brigham-Burke, M.; Peck, K.; Immobilization chemistries suitable for use in the BIAcore surface plasmon resonance detector. *Analytical Biochemistry,* **1992,** *205* (1), 132–136.

Ovchinnikov, S.; Kamisetty, H.; Baker, D.; Robust and accurate prediction of residue-residue interactions across protein interfaces using evolutionary information. *eLife,* **2014**, DOI: 10.7554/eLife.02030.

Overington, J. P.; Al-Lazikani, B.; Hopkins, A. L.; How many drug targets are there? *Nature Reviews Drug discovery,* **2006,** *5* (12), 993–996.

Pan American Health Organization / World Health Organization. Epidemiological Update: Cholera. 6. August 2018, Washington, D.C. PAHO/WHO. **2018**.

Pardon, E.; Laeremans, T.; Triest, S.; Rasmussen, S. G. F.; Wohlkönig, A.; Ruf, A.; Muyldermans, S.; Hol, W. G. J.; Kobilka, B. K.; Steyaert, J.; A general protocol for the generation of nanobodies for structural biology. *Nature Protocols,* **2014,** *9* (3), 674–693.

Paz, A.; Claxton, D. P.; Kumar, J. P.; Kazmier, K.; Bisignano, P.; Sharma, S.; Nolte, S. A.; Liwag, T. M.; Nayak, V.; Wright, E. M.; Grabe, M.; Mchaourab, H. S.; Abramson, J.; Conformational transitions of the sodium-dependent sugar transporter, vSGLT. *Proceedings of the National Academy of Sciences of the United States of America,* **2018**, *115* (12), E2742-E2751.

Peetz, O.; Hellwig, N.; Henrich, E.; Mezhyrova, J.; Dötsch, V.; Bernhard, F.; Morgner, N.; LILBID and nESI: Different native mass spectrometry techniques as tools in structural biology. *Journal of the American Society for Mass Spectrometry,* **2018,** *30*, 181–191.

Perozzo, R.; Folkers, G.; Scapozza, L.; Thermodynamics of protein-ligand interactions: history, presence, and future aspects. *Journal of Receptor and Signal Transduction research* **2004,** *24* (1-2), 1–52.

Perry, R. D.; Fetherston, J.; Yersinia pestis - Etiologic agent of plague. *Clinical Microbiology Reviews,* **1997,** *10* (1), 35–66.

Peter, M. F.; Untersuchung der Konformationsänderung von substratbindenden Proteinen aus TRAP-Transporters mit PELDOR-Spektroskopie und Röntgenstrukturanalyse. Master Thesis; Rheinische Friedrich-Wilhelms-Universität, Bonn (University of Bonn), **2016**.

Peter, M. F.; Tuukkanen, A. T.; Heubach, C. A.; Selsam, A.; Duthie, F. G.; Svergun, D. I.; Schiemann, O.; Hagelueken, G.; Studying conformational changes of the Yersinia type-III-secretion effector YopO in solution by integrative structural biology. *Structure,* **2019,** *27* (9), 1416-1426.

Peter, M. F.; Gebhardt, C.; Mächtel, R.; Glaenzer, J.; Thomas, G. H.; Cordes, T.; Hagelueken, G.; Cross-validation of distance measurements in proteins by PELDOR/DEER and single-molecule FRET. *BioRxiv,* **2020**, DOI: https://doi.org/10.1101/2020.11.23.394080

Peter, M. F.; Gebhardt, C.; Glaenzer, J.; Schneberger, N.; de Boer, M.; Thomas, G. H.; Cordes, T.; Hagelueken, G.; Triggering closure of a sialic acid TRAP transporter substrate binding protein through binding of natural of artificial substrates. *Journal of Molecular Biology,* **2021,** *433* (3), 166756.

Phillips, N. J.; Apicella, M. A.; McLeod Griffiss, J.; Gibson, B. W.; Structural studies of the lipooligosaccharides from Haemophilus influenzae type b strain A2. *Biochemistry,* **1993,** *32* (8), 2003–2012.

Pittman, M.; Variation and type specificity in the bacterial species Hemophilus influenzae. *The Journal of Experimental Medicine,* **1931,** *53* (4), 471–493.

Plass, T.; Milles, S.; Koehler, C.; Schultz, C.; Lemke, E. A.; Genetically encoded copper-free click chemistry. *Angewandte Chemie International Edition,* **2011,** *50* (17), 3878–3881.

Plumbridge, J.; Vimr, E.; Convergent pathways for utilization of the amino sugars N-acetylglucosamine, N-acetylmannosamine, and N-acetylneuraminic acid by Escherichia coli. *Journal of Bacteriology,* **1999,** *181* (1), 47–54.

Post, D. M. B.; Mungur, R.; Gibson, B. W.; Munson, R. S.; Identification of a novel sialic acid transporter in Haemophilus ducreyi. *Infection and Immunity,* **2005,** *73* (10), 6727–6735.

Postis, V.; Rawson, S.; Mitchell, J. K.; Lee, S. C.; Parslow, R. A.; Dafforn, T. R.; Baldwin, S. A.; Muench, S. P.; The use of SMALPs as a novel membrane protein scaffold for structure study by negative stain electron microscopy. *Biochimica et Biophysica Acta,* **2015,** *1848* (2), 496–501.

Potapov, A.; Yagi, H.; Huber, T.; Jergic, S.; Dixon, N. E.; Otting, G.; Goldfarb, D.; Nanometer-scale distance measurements in proteins using Gd^{3+} spin labeling. *Journal of the American Chemical Society,* **2010,** *132* (26), 9040–9048.

Prabudiansyah, I.; Kusters, I.; Caforio, A.; Driessen, A. J. M.; Characterization of the annular lipid shell of the Sec translocon. *Biochimica et Biophysica Acta,* **2015,** *1848* (10 Pt A), 2050–2056.

Punjani, A.; Rubinstein, J. L.; Fleet, D. J.; Brubaker, M. A.; cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods,* **2017,** *14* (3), 290–296.

Qasem-Abdullah, H.; Perach, M.; Livnat-Levanon, N.; Lewinson, O.; ATP binding and hydrolysis disrupt the high-affinity interaction between the heme ABC transporter HmuUV and its cognate substrate-binding protein. *The Journal of Biological Chemistry,* **2017,** *292* (35), 14617–14624.

Qiu, W.; Fu, Z.; Xu, G. G.; Grassucci, R. A.; Zhang, Y.; Frank, J.; Hendrickson, W. A.; Guo, Y.; Structure and activity of lipid bilayer within a membrane-protein transporter. *Proceedings of the National Academy of Sciences of the United States of America,* **2018,** *115* (51), 12985–12990.

Qu, Y.; Li, Y.; Tan, X.; Zhai, W.; Han, G.; Hou, J.; Liu, G.; Song, Y.; Liu, Y.; Synthesis and characterization of hydrophilic trityl radical TFO for biomedical and biophysical applications. *Chemistry-A European Journal,* **2019,** *25* (33), 7888–7895.

Quintero, M. J.; Montesinos, M. L.; Herrero, A.; Flores, E.; Identification of genes encoding amino acids permeases by inactivation of selected ORFs from the Synechocystis genomic sequence. *Genome research,* **2001,** *11* (12), 2034–2040.

Quiocho, F. A.; Higgins, C. F.; Atomic structures of periplasmic binding proteins and the high-affinity active transport systems in bacteria (and discussion). *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences,* **1990,** *326* (1236), 341-352.

Quiocho, F. A.; Spurlino, J. C.; Rodseth, L. E.; Extensive features of tight oligosaccharide binding revealed in high-resolution structures of the maltodextrin transport/chemosensory receptor. *Structure,* **1997,** *5* (8), 997–1015.

Rabiner, L. R.; A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE,* **1989,** *77* (2), 257–286.

Rabus, R.; Jack, D. L.; Kelly, D. J.; Saier, M. H. JR; TRAP transporters: an ancient family of extracytoplasmic solute-receptor-dependent secondary active transporters. *Microbiology,* **1999,** *145,* 3431–3445.

Ramadas, K.; Petersen, G. M.; Heiner, D. C.; Ward, J. I.; Class and subclass antibodies to Haemophilus influenzae type b capsule: Comparison of invasive disease and natural exposure. *Infection and Immunity,* **1986,** *53* (3), 486–490.

Rasmussen, S. G. F.; Choi, H.-J.; Fung, J. J.; Pardon, E.; Casarosa, P.; Chae, P. S.; Devree, B. T.; Rosenbaum, D. M.; Thian, F. S.; Kobilka, T. S.; Schnapp, A.; Konetzki, I.; Sunahara, R. K.; Gellman, S. H.; Pautsch, A.; Steyaert, J.; Weis, W. I.; Kobilka, B. K.; Structure of a nanobody-stabilized active state of the β_2 adrenoceptor. *Nature,* **2011a,** *469* (7329), 175–180.

Rasmussen, S. G. F.; Devree, B. T.; Zou, Y.; Kruse, A. C.; Chung, K. Y.; Kobilka, T. S.; Thian, F. S.; Chae, P. S.; Pardon, E.; Calinski, D.; Mathiesen, J. M.; Shah, S. T. A.; Lyons, J. A.; Caffrey, M.; Gellman, S. H.; Steyaert, J.; Skiniotis, G.; Weis, W. I.; Sunahara, R. K.; Kobilka, B. K.; Crystal structure of the β_2 adrenergic receptor-Gs protein complex. *Nature,* **2011b,** *477* (7366), 549–555.

Rath, A.; Glibowicka, M.; Nadeau, V. G.; Chen, G.; Deber, C. M.; Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proceedings of the National Academy of Sciences of the United States of America,* **2009,** *106* (6), 1760–1765.

Reginsson, G. W.; Kunjir, N. C.; Sigurdsson, S. T.; Schiemann, O.; Trityl radicals: spin labels for nanometer-distance measurements. *Chemistry - A European Journal*, **2012,** *18* (43), 13580–13584.

Reinhardt, A.; Neundorf, I.; Design and application of antimicrobial peptide conjugates. *International Journal of Molecular Sciences*, **2016,** *17* (5), 701.

Ressl, S.; van Terwisscha Scheltinga, A. C.; Vonrhein, C.; Ott, V.; Ziegler, C.; Molecular basis of transport and regulation in the Na^+/betaine symporter BetP. *Nature*, **2009,** *458* (7234), 47–52.

Reyes, N.; Ginter, C.; Boudker, O.; Transport mechanism of a bacterial homologue of glutamate transporters. *Nature*, **2009,** *462* (7275), 880–885.

Risberg, A.; Masoud, H.; Martin, A.; Richards, J. C.; Moxon, E. R.; Schweda, E. K. H.; Structural analysis of the lipopolysaccharide oligosaccharide epitopes expressed by a capsule-deficient strain of Haemophilus influenzae Rd. *European Journal of Biochemistry*, **1999,** *261* (1), 171–180.

Ritchie, T. K.; Grinkova, Y. V.; Bayburt, T. H.; Denisov, I. G.; Zolnerciks, J. K.; Atkins, W. M.; Sligar, S. G.; Chapter 11 - Reconstitution of membrane proteins in phospholipid bilayer nanodiscs. *Methods in Enzymology*, **2009,** *464*, 211–231.

Rodionov, D. A.; Hebbeln, P.; Gelfand, M. S.; Eitinger, T.; Comparative and functional genomic analysis of prokaryotic nickel and cobalt uptake transporters: evidence for a novel group of ATP-binding cassette transporters. *Journal of Bacteriology*, **2006,** *188* (1), 317–327.

Rodionov, D. A.; Hebbeln, P.; Eudes, A.; ter Beek, J.; Rodionova, I. A.; Erkens, G. B.; Slotboom, D. J.; Gelfand, M. S.; Osterman, A. L.; Hanson, A. D.; Eitinger, T.; A novel class of modular transporters for vitamins in prokaryotes. *Journal of Bacteriology*, **2009,** *191* (1), 42–51.

Rosa, L. T.; Bianconi, M. E.; Thomas, G. H.; Kelly, D. J.; Tripartite ATP-independent periplasmic (TRAP) transporters and tripartite tricarboxylate transporters (TTT): From uptake to pathogenicity. *Frontiers in Cellular and Infection Microbiology*, **2018,** *8* (33).

Rosadini, C. V.; Kagan, J. C.; Early innate immune responses to bacterial LPS. *Current Opinion in Immunology*, **2017,** *44*, 14–19.

Rothbauer, U.; Zolghadr, K.; Tillib, S.; Nowak, D.; Schermelleh, L.; Gahl, A.; Backmann, N.; Conrath, K.; Muyldermans, S.; Cardoso, M. C.; Leonhardt, H.; Targeting and tracing antigens in live cells with fluorescent nanobodies. *Nature Methods*, **2006,** *3* (11), 887–889.

Roy, R.; Hohng, S.; Ha, T.; A practical guide to single-molecule FRET. *Nature Methods,* **2008,** *5* (6), 507–516.

Rupp, B.; *Biomolecular crystallography. Principles, practice, and application to structural biology;* Garland Science, Taylor &Francis Group, LLC, New York, **2010**.

Sahin-Tóth, M.; Karlin, A.; Kaback, H. R.; Unraveling the mechanism of the lactose permease of Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America,* **2000,** *97* (20), 10729–10732.

Sahu, I. D.; McCarrick, R. M.; Troxel, K. R.; Zhang, R.; Smith, H. J.; Dunagan, M. M.; Swartz, M. S.; Rajan, P. V.; Kroncke, B. M.; Sanders, C. R.; Lorigan, G. A.; DEER EPR measurements for membrane protein structures via bifunctional spin labels and lipodisq nanoparticles. *Biochemistry,* **2013,** *52* (38), 6627–6632.

Saier, M. H. JR; A functional-phylogenetic classification system for transmembrane solute transporters. *Microbiology and Molecular Biology Reviews,* **2000,** *64* (2), 354–411.

Sanders, M. R.; Findlay, H. E.; Booth, P. J.; Lipid bilayer composition modulates the unfolding free energy of a knotted α-helical membrane protein. *Proceedings of the National Academy of Sciences of the United States of America,* **2018,** *115* (8), E1799-E1808.

Sauer, M.; Hofkens, J.; Enderlein, J.; *Handbook of fluorescence spectroscopy and imaging: from ensemble to single molecules;* John Wiley & Sons, 2010.

Saurin, W.; Hofnung, M.; Dassa, E.; Getting in or out: Early segregation between importers and exporters in the evolution of ATP-binding cassette (ABC) transporters. *Journal of Molecular Evolution,* **1999,** *48* (1), 22–41.

Scheepers, G. H.; Nijeholt, J. A. L. A.; Poolman, B.; An updated structural classification of substrate-binding proteins. *FEBS Letters,* **2016,** *590* (23), 4393–4401.

Scheidelaar, S.; Koorengevel, M. C.; Pardo, J. D.; Meeldijk, J. D.; Breukink, E.; Killian, J. A.; Molecular model for the solubilization of membranes into nanodisks by styrene maleic acid copolymers. *Biophysical Journal,* **2015,** *108* (2), 279–290.

Schenck, S.; Kunz, L.; Sahlender, D.; Pardon, E.; Geertsma, E. R.; Savtchouk, I.; Suzuki, T.; Neldner, Y.; Štefanić, S.; Steyaert, J.; Volterra, A.; Dutzler, R.; Generation and characterization of anti-VGLUT nanobodies acting as inhibitors of transport. *Biochemistry,* **2017,** *56* (30), 3962–3971.

Scheres, S. H. W.; RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *Journal of Structural Biology,* **2012,** *180*, 519–530.

Schiemann, O.; Magnetische Resonanzspektroskopie von Biomolekülen: EPR-Spektroskopie an biologischen Systemen. In *Bioanalytik,* 3. Auflage; Lottspeich, F., Engels, J. W., Editors.; Springer Spektrum: Berlin, Heidelberg, **2012**; pp 509–525.

Schiemann, O.; Prisner, T. F.; Long-range distance determinations in biomacromolecules by EPR spectroscopy. *Quarterly Reviews of Biophysics,* **2007,** *40* (1), 1–53.

Schmidt, M. J.; Borbas, J.; Drescher, M.; Summerer, D.; A genetically encoded spin label for electron paramagnetic resonance distance measurements. *Journal of the American Chemical Society,* **2014,** *136* (4), 1238–1241.

Schmidt, F. I.; Lu, A.; Chen, J. W.; Ruan, J.; Tang, C.; Wu, H.; Ploegh, H. L.; A single domain antibody fragment that recognizes the adaptor ASC defines the role of ASC domains in inflammasome assembly. *The Journal of Experimental Medicine,* **2016a,** *213* (5), 771–790.

Schmidt, T.; Wälti, M. A.; Baber, J. L.; Hustedt, E. J.; Clore, G. M.; Long distance measurements up to 160 Å in the GroEL tetradecamer using Q-Band DEER EPR spectroscopy. *Angewandte Chemie International Edition,* **2016b,** *55* (51), 15905–15909.

Schmidt, T.; Jeon, J.; Okuno, Y.; Chiliveri, S. C.; Clore, G. M.; Submillisecond freezing permits cryoprotectant-free EPR double electron-electron resonance spectroscopy. *Chemphyschem,* **2020,** *21* (12), 1224–1229.

Schneider, E.; Hunke, S.; ATP-binding-cassette (ABC) transport systems: Functional and structural aspects of the ATP-hydrolyzing subunits/domains. *FEMS Microbiology Reviews,* **1998,** *22* (1), 1–20.

Schweda, E. K. H.; Brisson, J.-R.; Alvelius, G.; Martin, A.; Weiser, J. N.; Hood, D. W.; Moxon, E. R.; Richards, J. C.; Characterization of the phosphocholine-substituted oligosaccharide in lipopolysaccharides of type b Haemophilus influenzae. *European Journal Biochemistry,* **2000,** *267* (12), 3902–3913.

Severi, E.; Hood, D. W.; Thomas, G. H.; Sialic acid utilization by bacterial pathogens. *Microbiology,* **2007,** *153* (Pt 9), 2817–2822.

Severi, E.; Hosie, A. H. F.; Hawkhead, J. A.; Thomas, G. H.; Characterization of a novel sialic acid transporter of the sodium solute symporter (SSS) family and in vivo comparison with known bacterial sialic acid transporters. *FEMS Microbiology Letters,* **2010,** *304* (1), 47–54.

Severi, E.; Randle, G.; Kivlin, P.; Whitfield, K.; Young, R.; Moxon, R.; Kelly, D.; Hood, D.; Thomas, G. H.; Sialic acid transport in Haemophilus influenzae is essential for lipopolysaccharide sialylation and serum resistance and is dependent on a novel tripartite ATP-independent periplasmic transporter. *Molecular Microbiology,* **2005,** *58* (4), 1173–1185.

Shaffer, P. L.; Goehring, A.; Shankaranarayanan, A.; Gouaux, E.; Structure and mechanism of a Na^+-independent amino acid transporter. *Science,* **2009,** *325* (5943), 1010–1014.

Shakhnovich, E. A.; King, S. J.; Weiser, J. N.; Neuraminidase expressed by Streptococcus pneumoniae desialylates the lipopolysaccharide of Neisseria meningitidis and Haemophilus influenzae: A paradigm for interbacterial competition among pathogens of the human respiratory tract. *Infection and Immunity,* **2002,** *70* (12), 7161–7164.

Sharff, A. J.; Rodseth, L. E.; Spurlino, J. C.; Quiocho, F. A.; Crystallographic evidence of a large ligand-induced hinge-twist motion between the two domains of the maltodextrin binding protein involved in active transport and chemotaxis. *Biochemistry,* **1992,** *31* (44), 10657-10663.

Shaw, J. G.; Kelly, D. J.; Binding protein dependent transport of C4-dicarboxylates in Rhodobacter capsulatus. *Archives of Microbiology,* **1991,** *155,* 466–472.

Shevchenko, A.; Tomas, H.; Havlis, J.; Olsen, J. V.; Mann, M.; In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature Protocols,* **2006,** *1* (6), 2856–2860.

Shi, Y.; Common folds and transport mechanisms of secondary active transporters. *Annual Review of Biophysics,* **2013,** *42,* 51–72.

Shilton, B. H.; The dynamics of the MBP-MalFGK$_2$ interaction: A prototype for binding protein dependent ABC-transporter systems. *Biochimica et Biophysica Acta,* **2008,** *1778* (9), 1772–1780.

Shimanda, T.; Arakawa, E. Itoh, K.; Okitsu, T. Matsushima, A.; Asai, Y.; Yamai, S.; Nakazato, T.; Balakrish Nair, G.; John Albert, M.; Takeda, Y.; Extended serotyping scheme for Vibrio cholerae. *Current Microbiology,* **1994,** *28,* 175–178.

Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Thompson, J. D.; Higgins, D. G.; Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology,* **2011,** *7,* 539.

Sillanaukee, P.; Pönniö, M.; Jääskeläinen, I. P.; Occurrence of sialic acids in healthy humans and different disorders. *European Journal of Clinical Investigations,* **1999,** *29* (5), 413–425.

Smirnova, I.; Kasho, V.; Jiang, X.; Pardon, E.; Steyaert, J.; Kaback, H. R.; Outward-facing conformers of LacY stabilized by nanobodies. *Proceedings of the National Academy of Sciences of the United States of America,* **2014,** *111* (52), 18548–18553.

Spicher, S.; Abdullin, D.; Grimme, S.; Schiemann, O.; Modeling of spin-spin distance distributions for nitroxide labeled biomacromolecules. *Physical chemistry chemical physics,* **2020,** *22* (42), 24282–24290.

Spicher, S.; Grimme, S.; Robust atomistic modeling of materials, organometallic, and biochemical systems. *Angewandte Chemie International Edition,* **2020**, *59* (36), 15665-15673.

Steeland, S.; Vandenbroucke, R. E.; Libert, C.; Nanobodies as therapeutics: big opportunities for small antibodies. *Drug Discovery Today,* **2016,** *21* (7), 1076–1113.

Steinberg, E. B.; Greene, K. D.; Bopp, C. A., Cameron, D. N.; Wells, J. G.; Mintz, E. D.; Cholera in the United States, 1995-2000: Trends at the end of the twentieth century. *The Journal of Infectious Diseases,* **2001,** *184* (6), 799–802.

Stetsenko, A.; Guskov, A.; An overview of the top ten detergents used for membrane protein crystallization. *Crystals,* **2017,** *7* (7), 197.

Sukupolvi-Petty, S.; Grass, S.; St Geme, J. W.; The Haemophilus influenzae Type b hcsA and hcsB gene products facilitate transport of capsular polysaccharide across the outer membrane and are essential for virulence. *Journal of Bacteriology,* **2006,** *188* (11), 3870–3877.

Sun, C.; Benlekbir, S.; Venkatakrishnan, P.; Wang, Y.; Hong, S.; Hosler, J.; Tajkhorshid, E.; Rubinstein, J. L.; Gennis, R. B.; Structure of the alternative complex III in a supercomplex with cytochrome oxidase. *Nature,* **2018,** *557* (7703), 123–126.

Sutcliffe, I. C.; Russell, R. R. B.; Lipoproteins of Gram-positive bacteria. *Journal of Bacteriology,* **1995,** *177* (5), 1123–1128.

Tacket, C. O.; Taylor, R. K.; Losonsky, G.; Lim, Y.; Nataro, J. P.; Kaper, J. B.; Levine, M. M.; Investigation of the roles of toxin-coregulated pili and mannose-sensitive hemagglutinin pili in the pathogenesis of Vibrio cholerae O139 infection. *Infection and Immunity,* **1998,** *66* (2), 692–695.

Tang, C.; Schwieters, C. D.; Clore, G. M.; Open-to-closed transition in apo maltose-binding protein observed by paramagnetic NMR. *Nature,* **2007,** *449* (7165), 1078–1082.

Tang, X.; Chang, S.; Luo, Q.; Zhang, Z.; Qiao, W.; Xu, C.; Zhang, C.; Niu, Y.; Yang, W.; Wang, T.; Zhang, Z.; Zhu, X.; Wei, X.; Dong, C.; Zhang, X.; Dong, H.; Cryo-EM structures of lipopolysaccharide transporter $LptB_2FGC$ in lipopolysaccharide or AMP-PNP-bound states reveal its transport mechanism. *Nature Communications*, **2019,** *10*, 4175.

The UniProt Consortium; UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Research*, **2019,** *47* (D1), D506-D515.

Thomas, C.; Tampé, R.; Structural and mechanistic principles of ABC transporters. *Annual Review of Biochemistry*, **2020,** *89*, 605–636.

Thomas, G. H.; Southworth, T.; León-Kempis, M. R.; Leech, A.; Kelly, D. J.; Novel ligands for the extracellular solute receptors of two bacterial TRAP transporters. *Microbiology*, **2006,** *152* (Pt 1), 187–198.

Thomas, G. H.; On the pull: periplasmic trapping of sugars before transport. *Molecular Microbiology*, **2017,** *104* (6), 883–888.

Thompson, R. F.; Walker, M.; Siebert, C. A.; Muench, S, P.; Ranson, N, A.; An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology. *Methods*, **2016,** *100*, 3–15.

Thomsen, L.; Gurevich, L.; A surface plasmon resonance assay for characterisation and epitope mapping of anti-GLP-1 antibodies. *Journal of Molecular Recognition*, **2018,** *31* (8), e2711.

Toseland, C. P.; Fluorescent labeling and modification of proteins. *Journal of Chemical Biology*, **2013,** *6* (3), 85–95.

Tsukanov, R.; Tomov, T. E.; Berger, Y.; Liber, M.; Nir, E.; Conformational dynamics of DNA hairpins at millisecond resolution obtained from analysis of single-molecule FRET histograms. *The Journal of Physical Chemistry B*, **2013,** *117* (50), 16105–16109.

Turk, D. C.; The pathogenicity of Haemophilus influenzae. *Journal of Medical Microbiology*, **1984,** *18* (1), 1-16.

Tuukkanen, A. T.; Spilotros, A.; Svergun, D. I.; Progress in small-angle scattering from biological solutions at high-brilliance synchrotrons. *IUCrJ*, **2017,** *4* (Pt 5), 518–528.

Uchański, T.; Masiulis, S.; Fischer, B.; Kalichuk, V.; Wohlkönig, A.; Zögg, T.; Remaut, H.; Vranken, W.; Aricescu, A. R.; Pardon, E.; Steyaert, J.; Megabodies expand the nanobody toolkit for protein structure determination by single-particle cryo-EM. *bioRxiv*, **2019**, DOI: 10.1101/812230.

Uchański, T.; Masiulis, S.; Fischer, B.; Kalichuk, V.; López-Sánchez, U.; Zarkadas, E.; Weckener, M.; Sente, A.; Ward, P.; Wohlkönig, A.; Zögg, T.; Remaut, H.; Naismith, J. H.; Nury, H.; Vranken, W.; Aricescu, A. R.; Pardon, E.; Steyaert, J.; Megabodies expand the nanobody toolkit for protein structure determination by single-particle cryo-EM. *Nature Methods,* **2021**, *18*, 60-68.

Vagenende, V.; Yap, M. G. S.; Trout, B. L.; Mechanisms of protein stabilization and prevention of protein aggregation by glycerol. *Biochemistry,* **2009,** *48* (46), 11084–11096.

van der Heide, T.; Poolman, B.; ABC transporters: one, two or four extracytoplasmic substrate-binding sites? *EMBO reports,* **2002,** *3* (10), 938–943.

Velazquez-Campoy, A.; Leavitt, S. A.; Freire, E.; Characterization of protein-protein interactions by isothermal titration calorimetry. Chapter 11. In *Protein-Protein Interactions: Methods and applications*: *Methods in Molecular Biology,* Meyerkord, C. L., Fu, H., Editors.; Springer Science+Buisness Media: New York, **2015,** *1278,* 183–204.

Ventola, C. L.; The Antibiotic Resistance Crisis, Part 1: Causes and threats. *Pharmacy & Therapeutics,* **2015,** *40* (4), 277-283.

Vera, L.; Stura, E. A.; Strategies for protein cryocrystallography. *Crystal Growth & Design,* **2014,** *14* (2), 427–435.

Vetting, M. W.; Al-Obaidi, N.; Zhao, S.; San Francisco, B.; Kim, J.; Wichelecki, D. J.; Bouvier, J. T.; Solbiati, J. O.; Vu, H.; Zhang, X.; Rodionov, D. A.; Love, J. D.; Hillerich, B. S.; Seidel, R. D.; Quinn, R. J.; Osterman, A. L.; Cronan, J. E.; Jacobson, M. P.; Gerlt, J. A.; Almo, S. C.; Experimental strategies for functional annotation and metabolism discovery: targeted screening of solute binding proteins and unbiased panning of metabolomes. *Biochemistry,* **2015,** *54* (3), 909–931.

Vigonsky, E.; Ovcharenko, E.; Lewinson, O.; Two molybdate/tungstate ABC transporters that interact very differently with their substrate binding proteins. *Proceedings of the National Academy of Sciences of the United States of America,* **2013,** *110* (14), 5440–5445.

Vimr, E. R.; Troy, F. A.; Identification of an inducible catabolic system for sialic acids (nan) in Escherichia coli. *Journal of Bacteriology,* **1985,** *164* (2), 845–853.

Vimr, E. R.; Lawrisuk, L.; Galen, J.; Kaper, J. B.; Cloning and expression of the Vibrio cholerae neuraminidase gene nanH in Escherichia coli. *Journal of Bacteriology,* **1988,** *170* (4), 1495-1504.

Vimr, E. R.; Aaronson, W.; Silver, R. P.; Genetic analysis of chromosomal mutations in the polysialic acid gene cluster of Escherichia coli K1. *Journal of Bacteriology,* **1989,** *171* (2), 1106–1117.

Vimr, E.; Lichtensteiger, C.; Steenbergen, S.; Sialic acid metabolism's dual function in Haemophilus influenzae. *Molecular Microbiology,* **2000,** *36* (5), 1113–1123.

Vimr, E. R.; Kalivoda, K. A.; Deszo, E. L.; Steenbergen, S. M.; Diversity of microbial sialic acid metabolism. *Microbiology and Molecular Biology Reviews,* **2004,** *68* (1), 132–153.

Vosjan, M. J. W. D.; Perk, L. R.; Roovers, R. C.; Visser, G. W. M.; Stigter-van Walsum, M.; van Bergen En Henegouwen, P. M. P.; van Dongen, G. A. M. S.; Facile labelling of an anti-epidermal growth factor receptor Nanobody with ^{68}Ga via a novel bifunctional desferal chelate for immuno-PET. *European Journal of Nuclear Medicine and Molecular Imaging,* **2011,** *38* (4), 753–763.

Wahlgren, W. Y.; Dunevall, E.; North, R. A.; Paz, A.; Scalise, M.; Bisignano, P.; Bengtsson-Palme, J.; Goyal, P.; Claesson, E.; Caing-Carlsson, R.; Andersson, R.; Beis, K.; Nilsson, U. J.; Farewell, A.; Pochini, L.; Indiveri, C.; Grabe, M.; Dobson, R. C. J.; Abramson, J.; Ramaswamy, S.; Friemann, R.; Substrate-bound outward-open structure of a Na$^+$-coupled sialic acid symporter reveals a new Na$^+$ site. *Nature Communications,* **2018,** *9,* 1753.

Wales, T. E.; Engen, J. R.; Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spectrometry Reviews,* **2006,** *25* (1), 158–170.

Wallin, E.; von Heijne, G.; Genome-wide analysis of integral membrane proteins from eubacterial, achaean, and eukaryotic organisms. *Protein Science,* **1998**, *7* (4), 1029–1038.

Walmsley, A. R.; Shaw, J. G.; Kelly, D. J.; The mechanism of ligand binding to the periplasmic C4-dicarboxylate binding protein (DctP) from Rhodobacter capsulatus. *The Journal of Biological Chemistry,* **1992,** *267* (12), 8064–8072.

Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G.; Expanding the genetic code of Escherichia coli. *Science,* **2001,** *292* (5516), 498–500.

Wang, Z.; Ma, Y.; Khalil, H.; Wang, R.; Lu, T.; Zhao, W.; Zhang, Y.; Chen, J.; Chen, T.; Fusion between fluid liposomes and intact bacteria: study of driving parameters and in vitro bactericidal efficacy. *International Journal of Nanomedicine,* **2016,** *11,* 4025–4036.

Ward, R.; Pliotas, C.; Branigan, E.; Hacker, C.; Rasmussen, A.; Hagelueken, G.; Booth, I. R.; Miller, S.; Lucocq, J.; Naismith, J. H.; Schiemann, O.; Probing the structure of the mechanosensitive channel of small conductance in lipid bilayers with pulsed electron-electron double resonance. *Biophysical Journal,* **2014,** *106* (4), 834–842.

Watanabe, A.; Choe, S.; Chaptal, V.; Rosenberg, J. M.; Wright, E. M.; Grabe, M.; Abramson, J.; The mechanism of sodium and substrate release from the binding pocket of vSGLT. *Nature,* **2010,** *468* (7326), 988–991.

Waterhouse, A. M.; Procter, J. B.; Martin, D. M. A.; Clamp, M.; Barton, G. J.; Jalview Version 2 - a multiple sequence alignment editor and analysis workbench. *Bioinformatics,* **2009,** *25* (9), 1189–1191.

Weichenberger, C. X.; Rupp, B.; Ten years of probabilistic estimates of biocrystal solvent content: new insights via nonparametric kernel density estimate. *Acta Crystallographica Section D, Biological Crystallography,* **2014,** *70* (Pt 6), 1579–1588.

Weyand, S.; Shimamura, T.; Yajima, S.; Suzuki, S.; Mirza, O.; Krusong, K.; Carpenter, E. P.; Rutherford, N. G.; Hadden, J. M.; O‘Reilly, J.; Ma, P.; Saidijam, M.; Patching, S. G.; Hope, R. J.; Norbertczak, H. T.; Roach, P. C. J.; Iwate, S.; Henderson, P. J. F.; Cameron, A. D.; Structure and molecular mechanism of a nucleobase-cation-symport-1 family transporter. *Science,* **2008,** *322* (5902), 703–709.

White, S. H.; Biophysical dissection of membrane proteins. *Nature,* **2009,** *459* (7245), 344–346.

Winnen, B.; Hvorup, R. N.; Saier, M. H. JR; The tripartite tricarboxylate transporter (TTT) family. *Research in Microbiology,* **2003,** *154* (7), 457–465.

Wirth, C.; Condemine, G.; Boiteux, C.; Bernèche, S.; Schirmer, T.; Peneff, C. M.; NanC crystal structure, a model for outer-membrane channels of the acidic sugar-specific KdgM porin family. *Journal of Molecular Biology,* **2009,** *394* (4), 718–731.

Wiseman, T.; Williston, S.; Brandts, J. F.; Lin, L.-N.; Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Analytical Biochemistry,* **1989,** *179,* 131-137.

World Health Organization; The World Health Report 2004. Changing history, **2004.**

World Health Organization; The Immunological Basis for Immunization Series, Modul 9: Haemophilus influenzae type b, **2007.**

World Health Organization; Haemophilus influenzae type b (Hib) vaccination position paper – September 2013. *Weekly epidemiological record,* **2013,** *88* (39), 413–428.

World Health Organization; Prioritization of pathogens to guide discovery, research and development of new antibiotics for drug-resistant bacterial infections, including tuberculosis, **2017a**.

World Health Organization; Cholera, 2016. *Weekly epidemiological record,* **2017b,** *92* (36), 521–536.

World Health Organization; WHO publishes list of bacteria for which new antibiotics are urgently needed. https://www.who.int/news/item/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed (accessed: 22.10.2020), **2017c**.

World Health Organization; Antibiotic Resistance. https://www.who.int/news-room/fact-sheets/detail/antibiotic-resistance (accessed: 21.10.2020), **2020a**.

World Health Organization; Global Antimicrobial Resistance and Use Surveillance System (GLASS) report, **2020b**.

World Health Organization; Global tuberculosis report 2020, **2020c**.

World Health Organization; The top 10 causes of death. www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death (accessed: 06.01.2020), **2020d**.

Wrapp, D.; Vlieger, D. de; Corbett, K. S.; Torres, G. M.; Wang, N.; van Breedam, W.; Roose, K.; van Schie, L.; Hoffmann, M.; Pöhlmann, S.; Graham, B. S.; Callewaert, N.; Schepens, B.; Saelens, X.; McLellan, J. S.; Structural basis for potent neutralization of Betacoronaviruses by single-domain camelid antibodies. *Cell,* **2020,** *181* (5), 1004-1015.

Wunnicke, D.; Hänelt, I.; The synergetic effects of combining structural biology and EPR spectroscopy on membrane proteins. *Crystals,* **2017,** *7* (4), 117.

Xiang, Y.; Nambulli, S.; Xiao, Z.; Liu, H.; Sang, Z.; Duprex, W. P.; Schneidman-Duhovny, D.; Zhang, C.; Shi, Y.; Versatile and multivalent nanobodies efficiently neutralize SARS-CoV-2. *Science,* **2020,** *370* (6523), 1479-1484.

Yagi, H.; Banerjee, D.; Graham, B.; Huber, T.; Goldfarb, D.; Otting, G.; Gadolinium tagging for high-precision measurements of 6 nm distances in protein assemblies by EPR. *Journal of the American Chemical Society,* **2011,** *133* (27), 10418–10421.

Yamashita, A.; Singh, S. K.; Kawate, T.; Jin, Y.; Gouaux, E.; Crystal structure of a bacterial homologue of Na^+/Cl^--dependent neurotransmitter transporters. *Nature,* **2005,** *437* (7056), 215–223.

Yang, N. J.; Hinner, M. J.; Getting across the cell membrane: an overview for small molecules, peptides, and proteins. *Methods in Molecular Biology,* **2015,** *1266,* 29–53.

Yap, L. L.; Lin, M. T.; Ouyang, H.; Samoilova, R. I.; Dikanov, S. A.; Gennis, R. B.; The quinone-binding sites of the cytochrome bo_3 ubiquinol oxidase from Escherichia coli. *Biochimica et Biophysica Acta,* **2010,** *1797* (12), 1924–1932.

Yardeni, E. H.; Bahrenberg, T.; Stein, R. A.; Mishra, S.; Zomot, E.; Graham, B.; Tuck, K. L.; Huber, T.; Bibi, E.; Mchaourab, H. S.; Goldfarb, D.; Probing the solution structure of the E. coli multidrug transporter MdfA using DEER distance measurements with nitroxide and Gd(III) spin labels. *Scientific reports,* **2019,** *9*, 12528.

Yernool, D.; Boudker, O.; Jin, Y.; Gouaux, E.; Structure of a glutamate transporter homologue from Pyrococcus horikoshii. *Nature,* **2004,** *431*, 811–818.

Yin, D. M.; Hammler, D.; Peter, M. F.; Marx, A.; Schmitz, A.; Hagelueken, G.; Inhibitor-directed spin labelling-A high precision and minimally invasive technique to study the conformation of proteins in solution. *Chemistry – A European Journal,* **2018,** *24* (26), 6665–6671.

Young, D. D.; Young, T. S.; Jahnz, M.; Ahmad, I.; Spraggon, G.; Schultz, P. G.; An evolved aminoacyl-tRNA synthetase with atypical polysubstrate specificity. *Biochemistry,* **2011,** *50* (11), 1894–1900.

Zhang, H.; Pan, Y.; Hu, L.; Hudson, M. A.; Hofstetter, K. S.; Xu, Z.; Rong, M.; Wang, Z.; Prasad, B. V. V.; Lockless, S. W.; Chiu, W.; Zhou, M.; TrkA undergoes a tetramer-to-dimer conversion to open TrkH which enables changes in membrane potential. *Nature Communications,* **2020,** *11*, 547.

Zhou, H.-X.; Cross, T. A.; Influences of membrane mimetic environments on membrane protein structures. *Annual Review of Biophysics,* **2013,** *42*, 361–392.

Zomot, E.; Bendahan, A.; Quick, M.; Zhao, Y.; Javitch, J. A.; Kanner, B. I.; Mechanism of chloride interaction with neurotransmitter:sodium symporters. *Nature,* **2007,** *449* (7163), 726–730.

Zou, P.; Bortolus, M.; Mchaourab, H. S.; Conformational cycle of the ABC transporter MsbA in liposomes: detailed analysis using double electron-electron resonance spectroscopy. *Journal of Molecular Biology,* **2009,** *393* (3), 586–597.

Zwahlen, A.; Winkelstein, J. A.; Moxon, E. R.; Surface determinants of Haemophilus influenzae pathogenicity: Comparative virulence of capsular transformants in normal and complement-depleted rats. *Journal of Infectious Diseases,* **1983,** *148* (3), 385-394.

Supporting Information

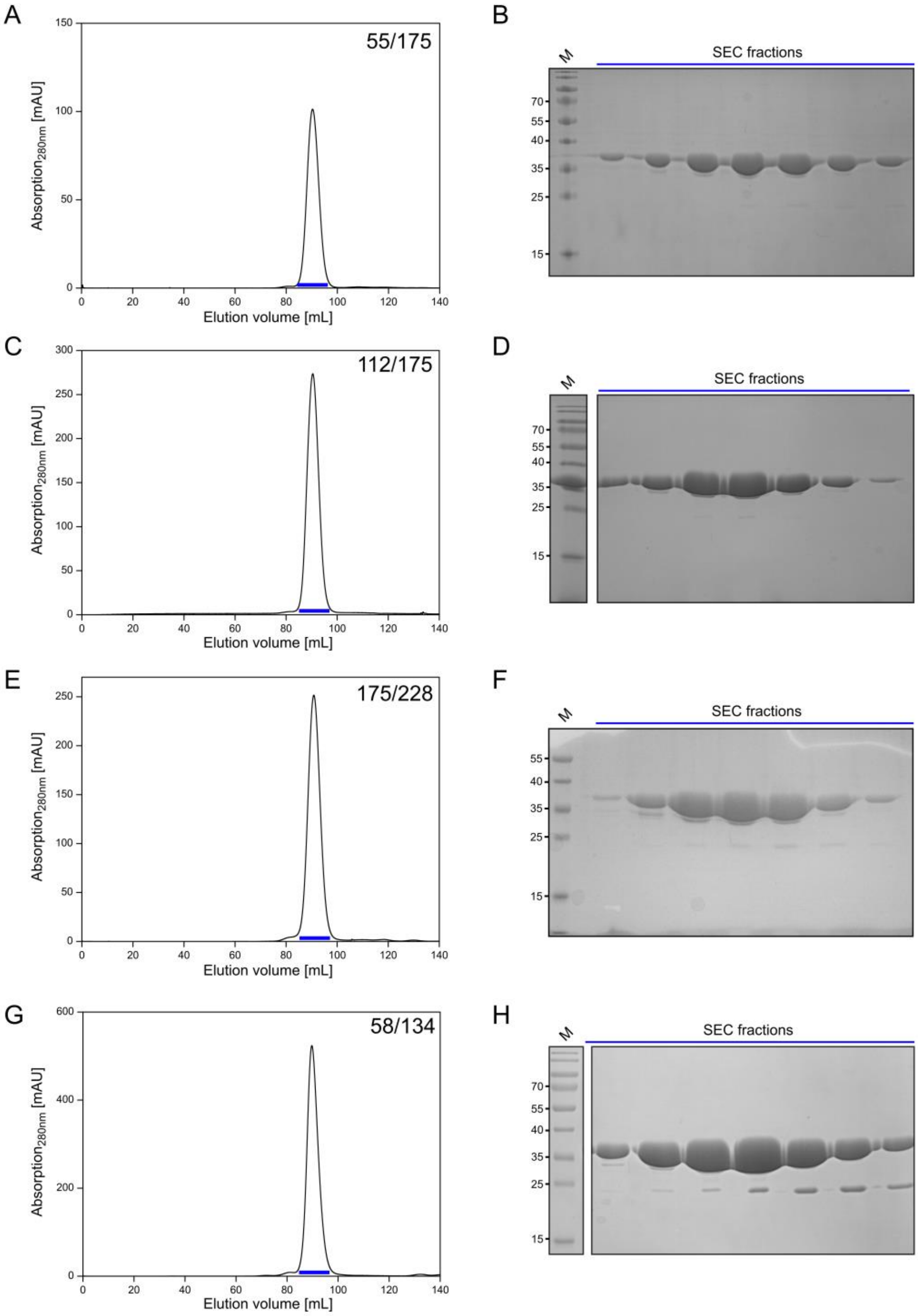


Figure SI-1: Purification of HiSiaP double cysteine mutants. A) SEC run of HiSiaP 55C/175C on a HiLoad SD 200 16/600 column (buffer: 50 mM Tris (pH 8), 50 mM NaCl, 1 mM TCEP). **B)** SDS-gel with samples from the SEC eluted fractions in A). **C-G)** Same as A) and B) but for different HiSiaP mutants as mentioned in the figure.

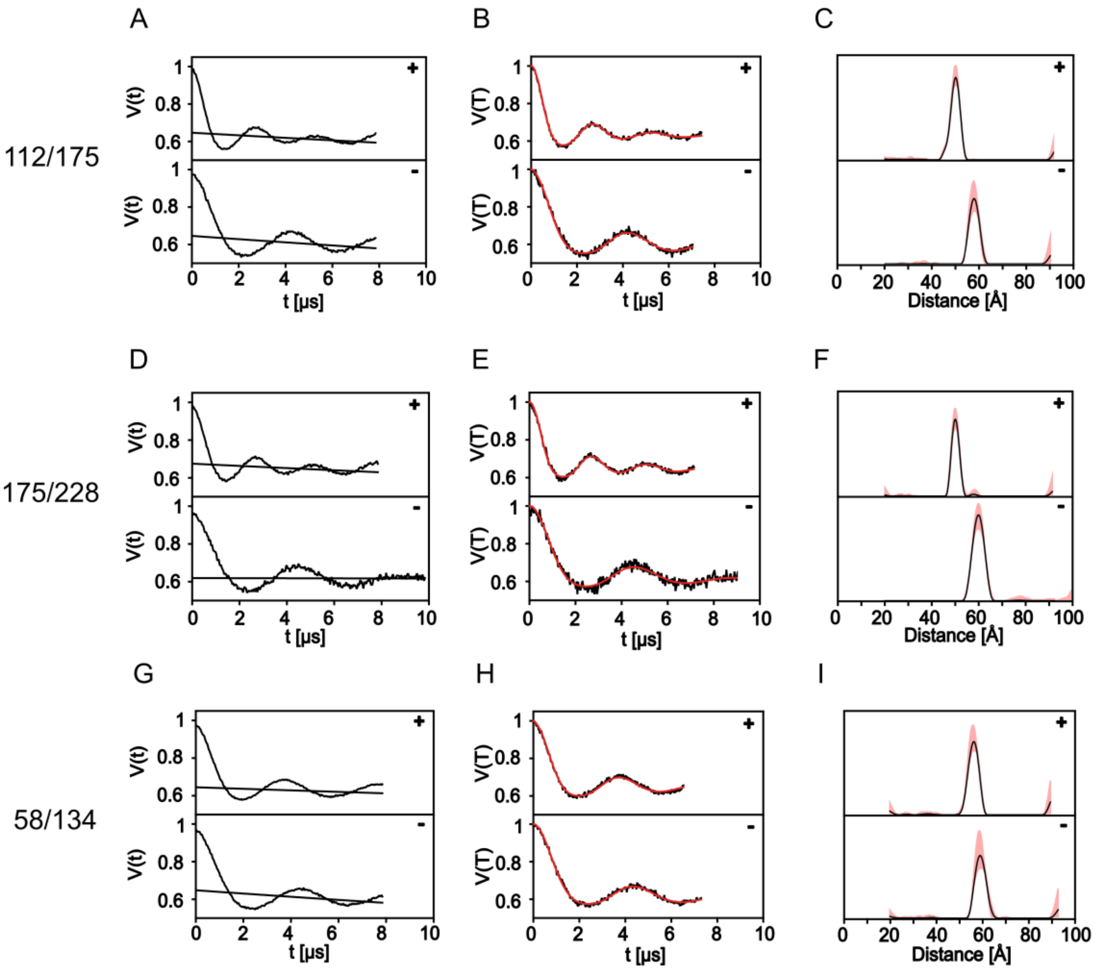


Figure SI-2: PELDOR data of HiSiaP mutants. A, D, G) Raw PELDOR time traces for apo (-) and holo (+) measurements of HiSiaP double spin labelled mutants, as mentioned in the figure. The background correction is indicated as black line. **B, E, H)** Background-corrected PELDOR time traces and fit of the signals (red). **C, F, I)** Distance distributions from PELDOR time traces with validation of the distribution (red). All data were processed with DeerAnalysis (Jeschke et al., 2006). The figure was modified from (Peter et al., 2020).

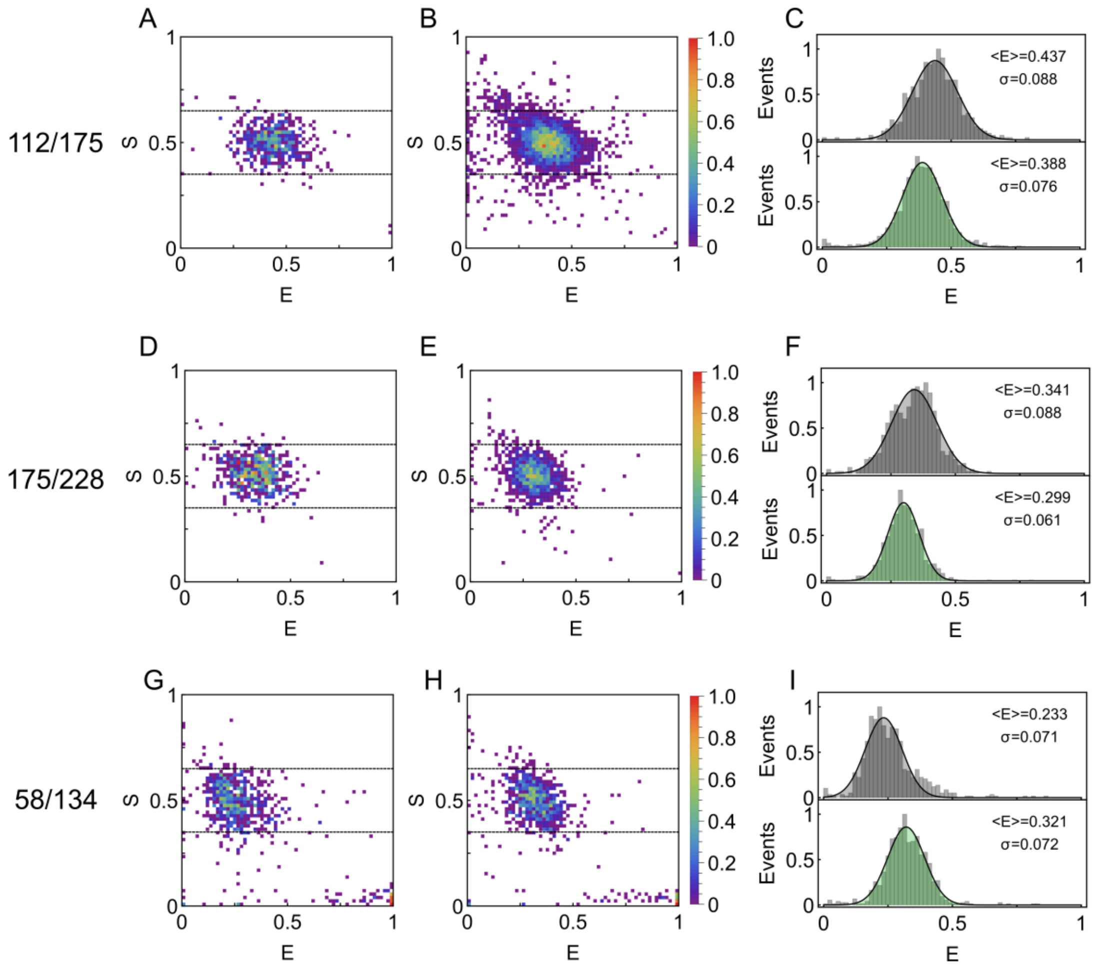


Figure SI-3: smFRET data of HiSiaP mutants. A, D, G) ES-2D-Histograms of HiSiaP mutants in apo conditions. **B, E, H)** Same as A, D, G) but in holo conditions. **C, F, I)** 1D-E-Histograms from the ES-data for apo (grey) and holo (green), fitted with a 1D-Gaussian distribution. The mean E and standard deviation σ are given. The figure was modified from (Peter et al., 2020).

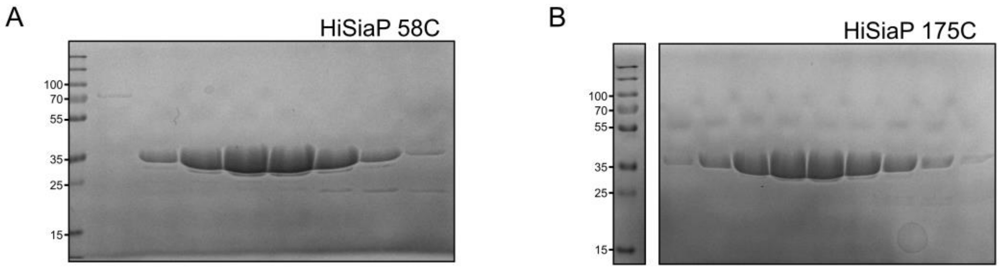


Figure SI-4: Purification of HiSiaP single cysteine mutants. A) SDS-gel with fractionated samples from a final SEC run of HiSiaP single cysteine mutant 58C. **B)** Same as A) but for mutant HiSiaP 175C.

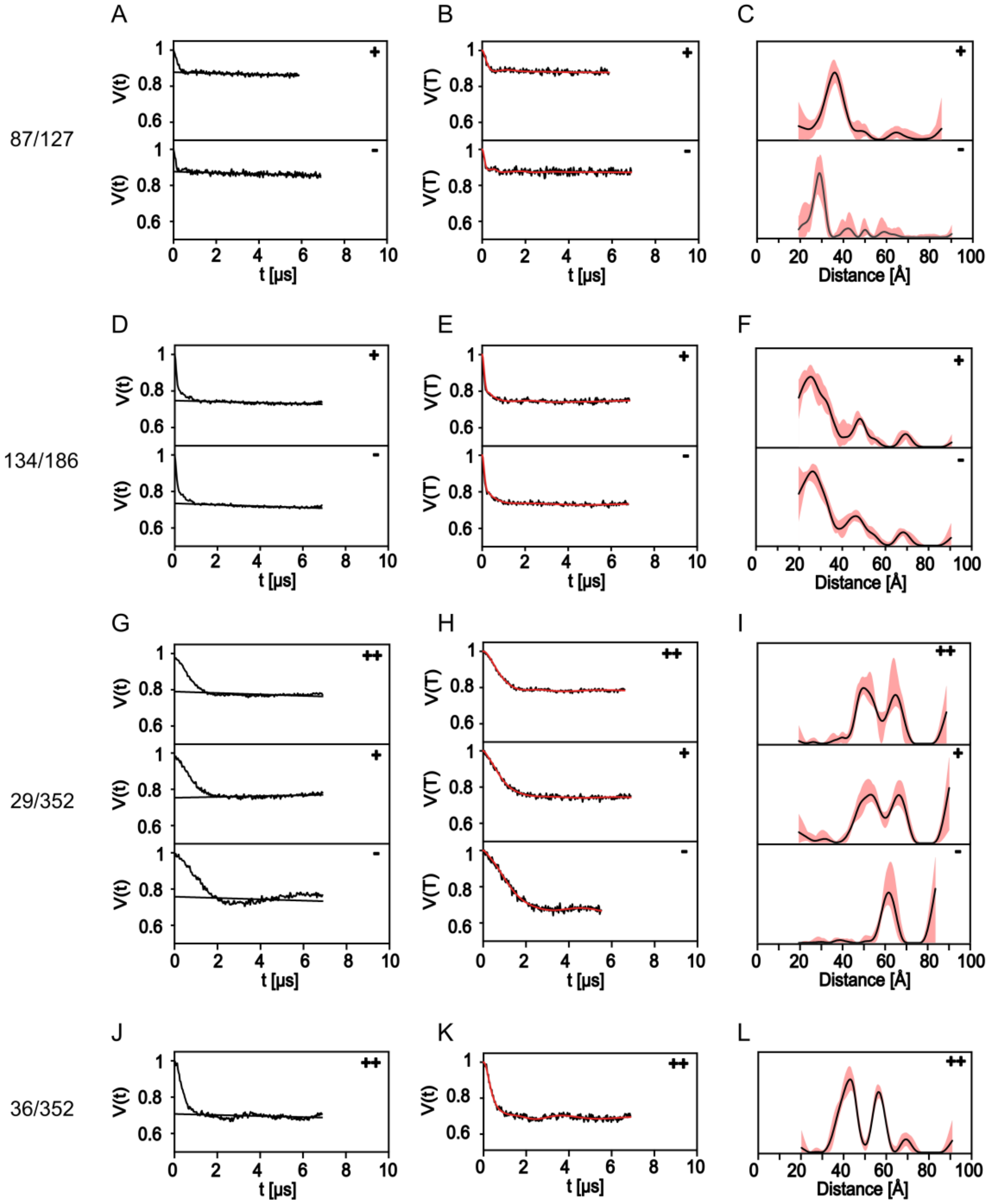


Figure SI-5: PELDOR data of MalE mutants. A, D, G, J) Raw PELDOR time traces for apo (-) and holo (+, 1 mM; ++, 10 mM maltose) measurements of MalE double spin labelled mutants. The background-correction is indicated as black line. Further apo and holo measurements for MalE 36/352 are shown in Figure 2-11. **B, E, H, K)** Background-corrected PELDOR time traces and fit of the signals (red). **C, F, I, L)** Distance distributions from PELDOR time traces with validation of the distribution (red). All data were processed with DeerAnalysis (Jeschke et al., 2006). The figure was modified from (Peter et al., 2020).

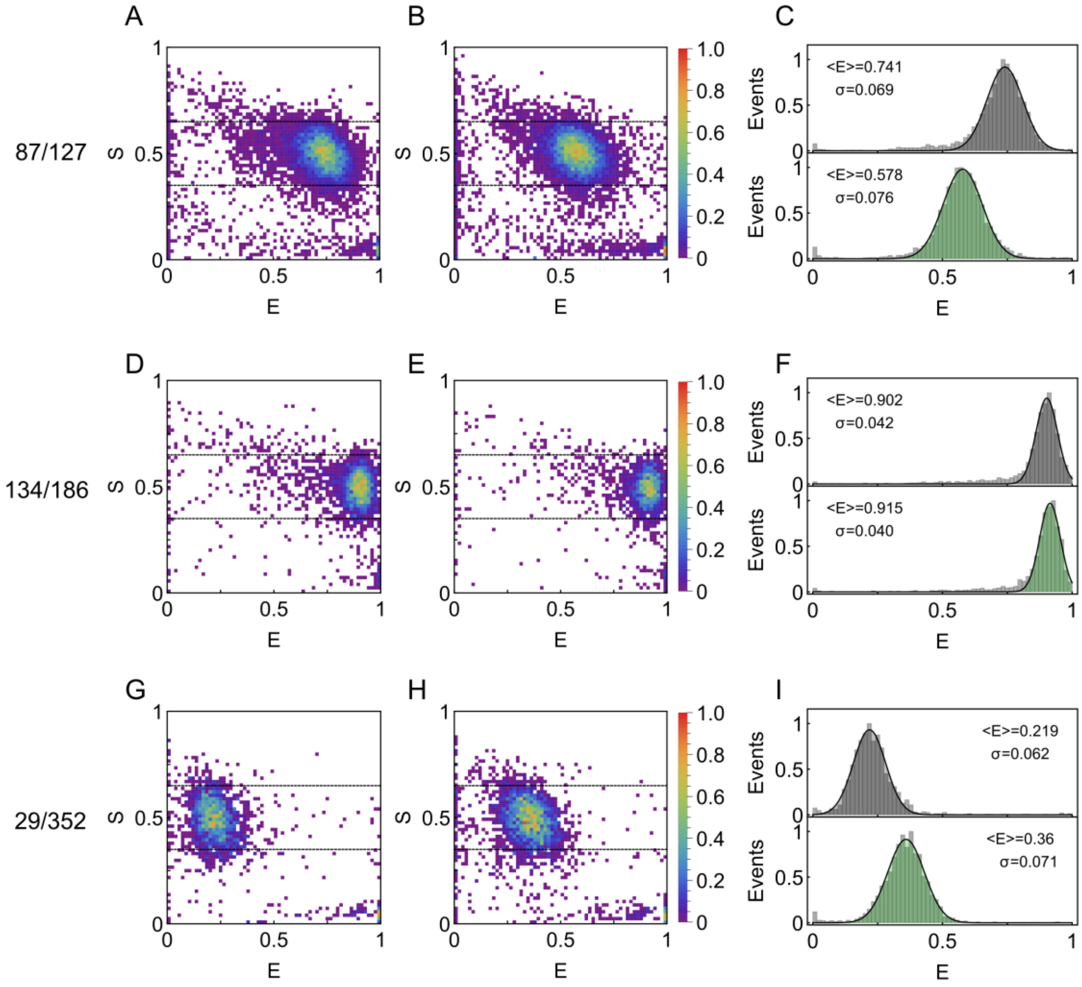


Figure SI-6: smFRET data of MalE mutants. A, D, G) ES-2D-Histograms of MalE mutants in apo conditions. **B, E, H)** Same as A, D, G) but in holo conditions. **C, F, I)** 1D-E-Histograms from the ES-data for apo (grey) and holo (green), fitted with a 1D-Gaussian distribution. The mean E and standard deviation σ are given. The figure was modified from (Peter et al., 2020).

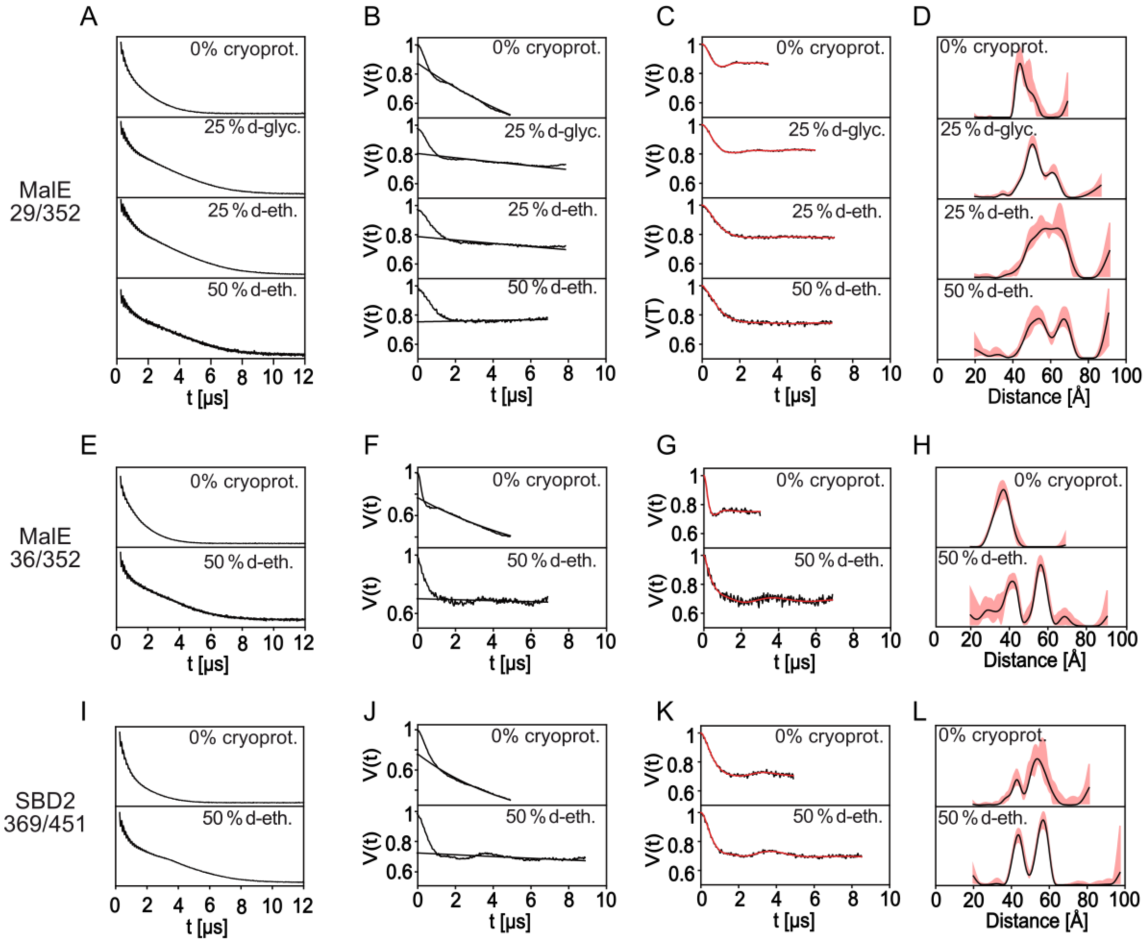


Figure SI-7: Data of PELDOR experiments with variations of cryoprotectant. A, E, I) 2PESEEM spectra for different mutants of MalE and SBD2. The variations of cryoprotectants are mentioned in the figure, 50% d-ethylene glycol was standardly used. **B, F, J)** Raw PELDOR time traces for the measurements with variations of cryoprotectant, the background-correction is indicated as black line. **C, G, K)** Background-corrected PELDOR time traces and fit of the signals (red). **D, H, L)** Distance distributions from PELDOR time traces with validation of the distribution (red). All data were processed with DeerAnalysis (Jeschke et al., 2006). The figure was modified from (Peter et al., 2020).

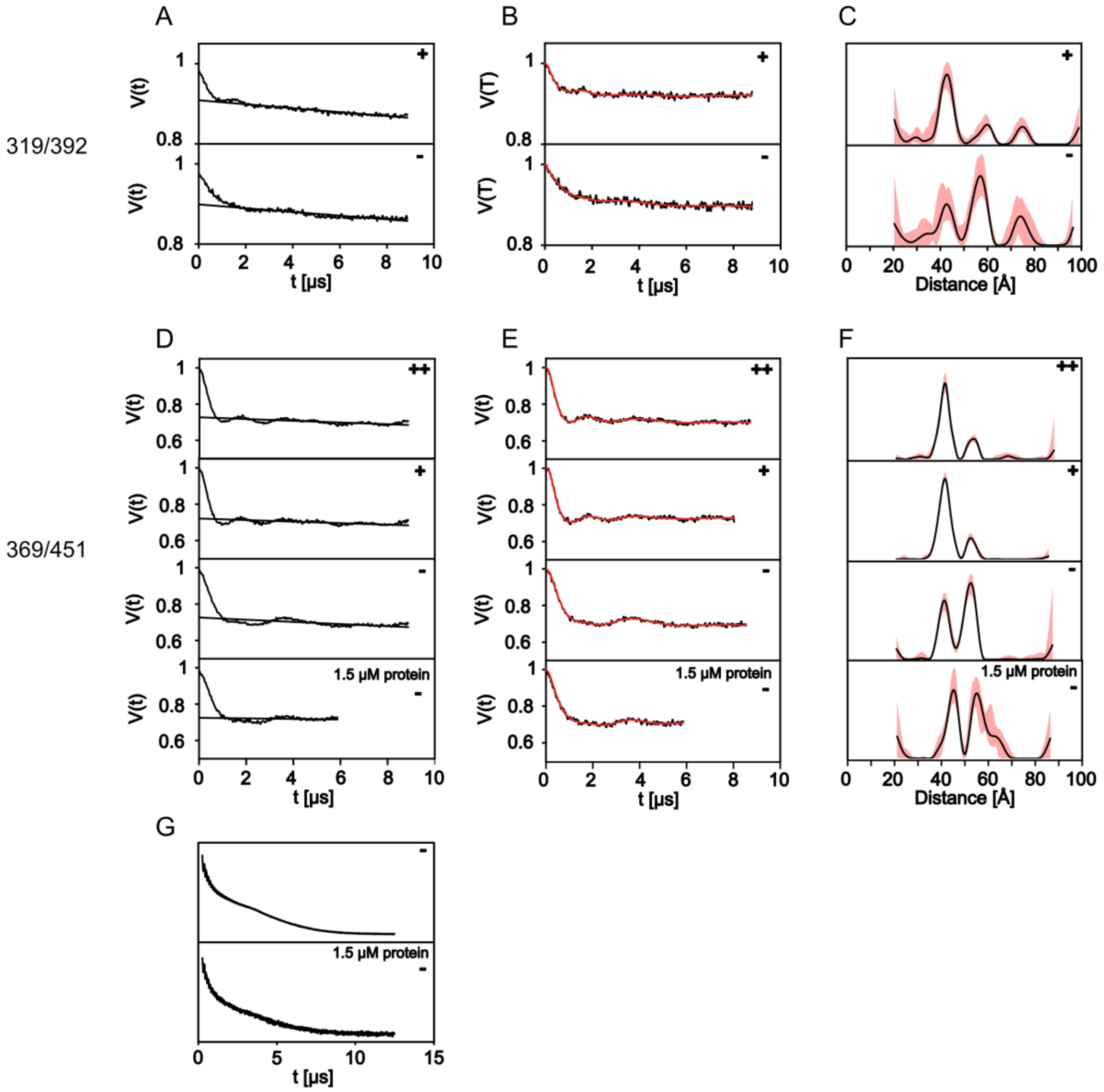


Figure SI-8: PELDOR data of SBD2 mutants. A, D) Raw PELDOR time traces for apo (-) and holo (+, 1 mM; ++, 10 mM maltose) measurements for MalE double spin labelled mutants. The background-correction is indicated as line. **B, E)** Background-corrected PELDOR time traces and fit of the signals (red). **C, F)** Distance distributions from PELDOR time traces with validation of the distribution (red). **G)** 2PESEEM spectra of apo measurement for mutant 369/451 for 15 µM and 1.5 µM protein concentration. All data were processed with DeerAnalysis (Jeschke et al., 2006). The figure was modified from (Peter et al., 2020).

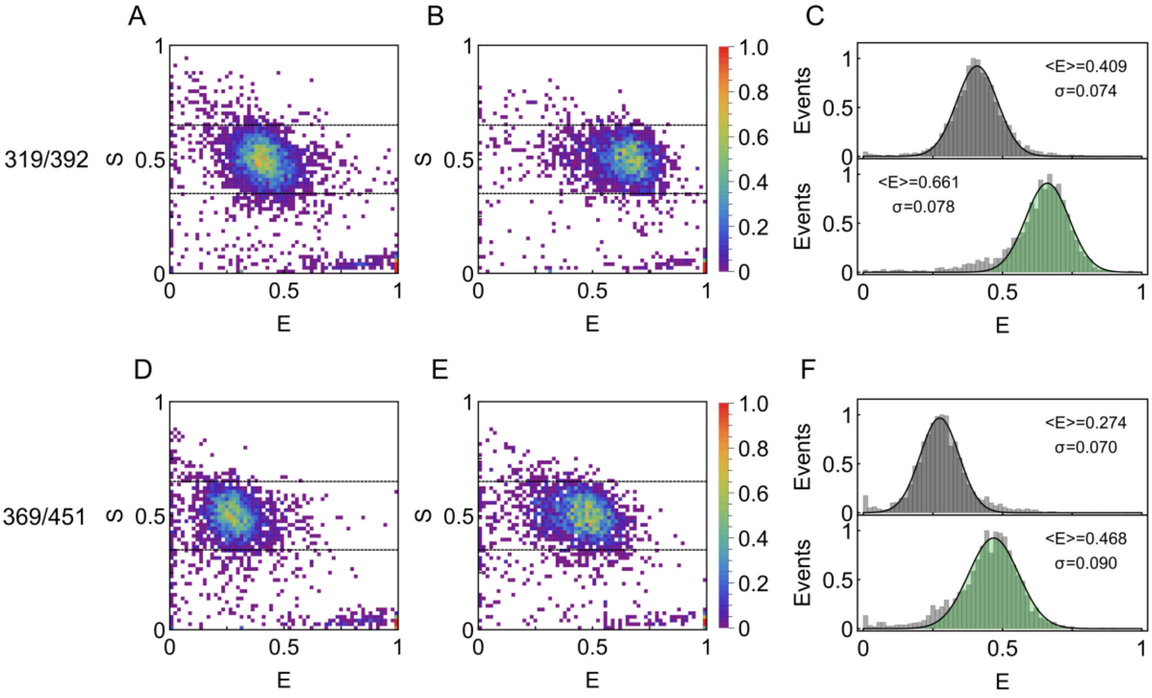


Figure SI-9: smFRET data of SBD2 mutants. A, D) ES-2D-Histograms of SBD2 mutants in apo conditions. **B, E)** Same as A, D) but in holo conditions. **C, F)** 1D-E-Histograms from the ES-data for apo (grey) and holo (green), fitted with a 1D-Gaussian distribution. The mean E and standard deviation σ are given in the figure. The figure was modified from (Peter et al., 2020).

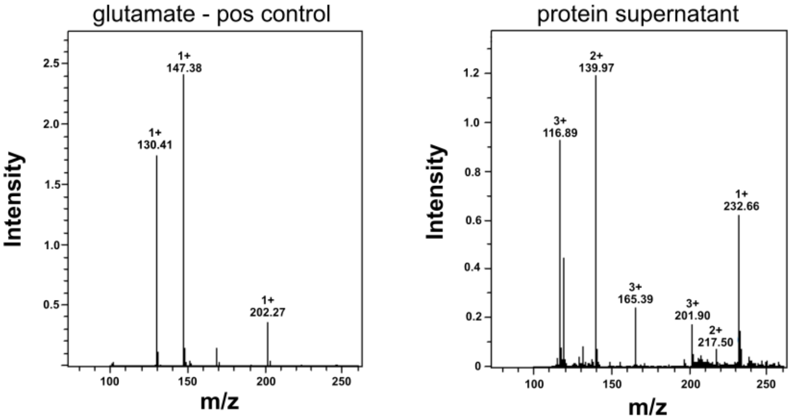


Figure SI-10: LC-MS for detection of glutamine. A) LC-MS spectrum of glutamine in standard protein buffer as positive control ($MW_{glutamine}$: 146.15 g/mol). **B)** LC-MS spectrum of supernatant after heat-precipitation and centrifugation of a SBD2 protein sample. The measurements were performed with the help of Frank Eggert and Stephanie Kath-Schorr, University of Bonn. The figure was modified from (Peter et al., 2020).

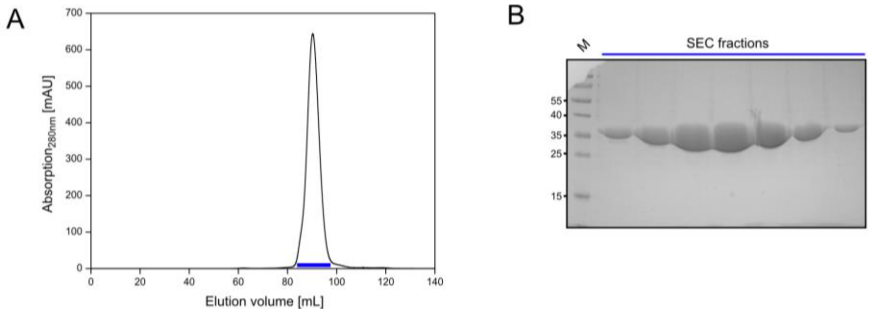


Figure SI-11: Purification of HiSiaP wildtype. A) SEC run of HiSiaP wildtype on a HiLoad SD 200 16/600 column with standard buffer. **B)** SDS-gel with samples from the SEC eluted fractions in A).

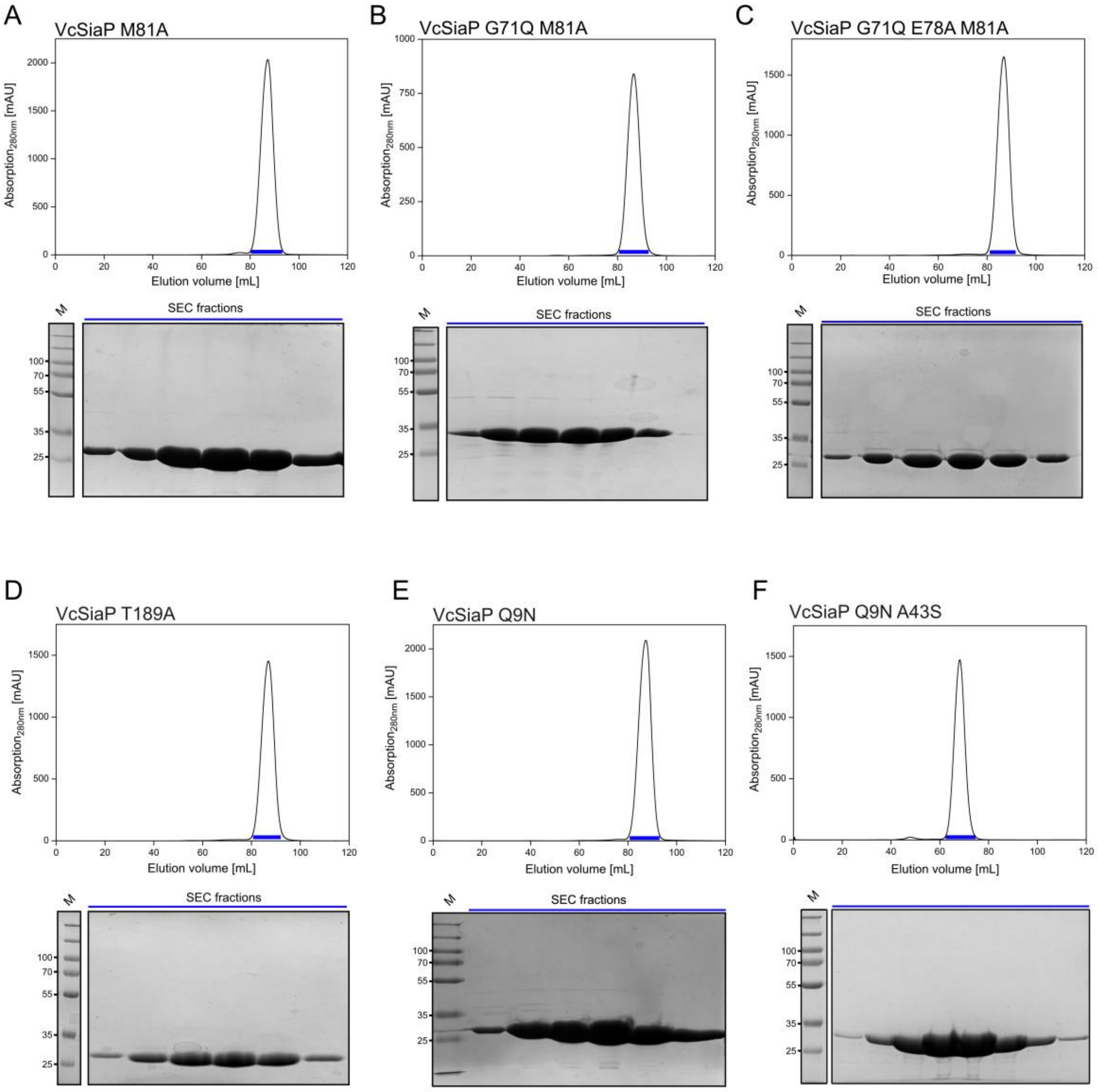


Figure SI-12: Purification of VcSiaP second-shell mutants. A) Top: SEC run of VcSiaP mutant M81A on a HiLoad SD 200 16/600 column with standard buffer. Bottom: SDS-gel with samples from the SEC eluted fractions. **B-F)** Same as A) but with other VcSiaP mutants as indicated in the figure. A HiLoad SD 75 16/600 column was used for SEC in F).

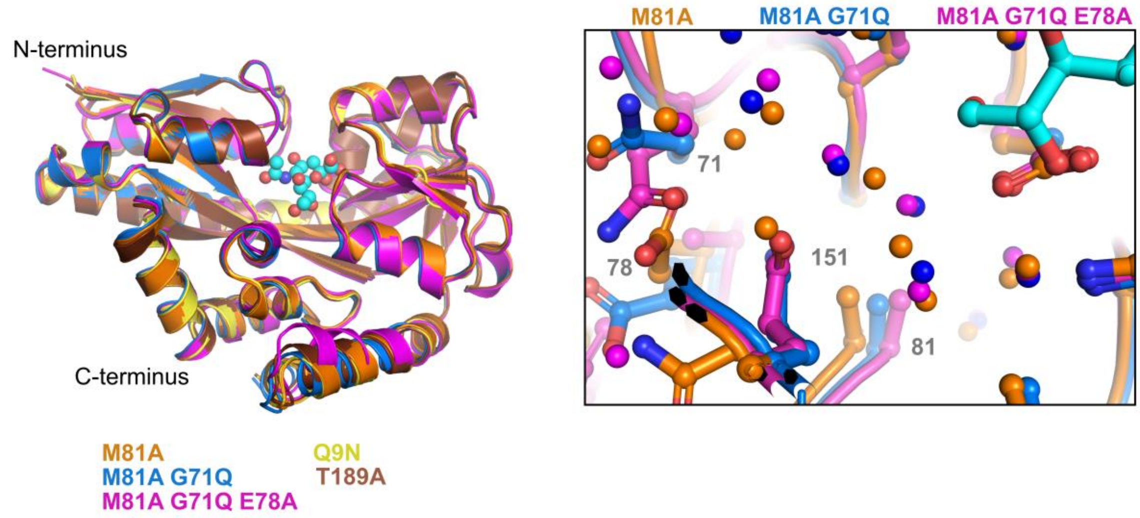


Figure SI-13: Comparison of VcSiaP second-shell mutants. A) Alignment of VcSiaP second shell mutants on the N-terminus fold domain from Chapter 3. Chain A was used from each structure for alignment. **B)** Alignment of the three mutant structures from region I with focus on the mutated region.

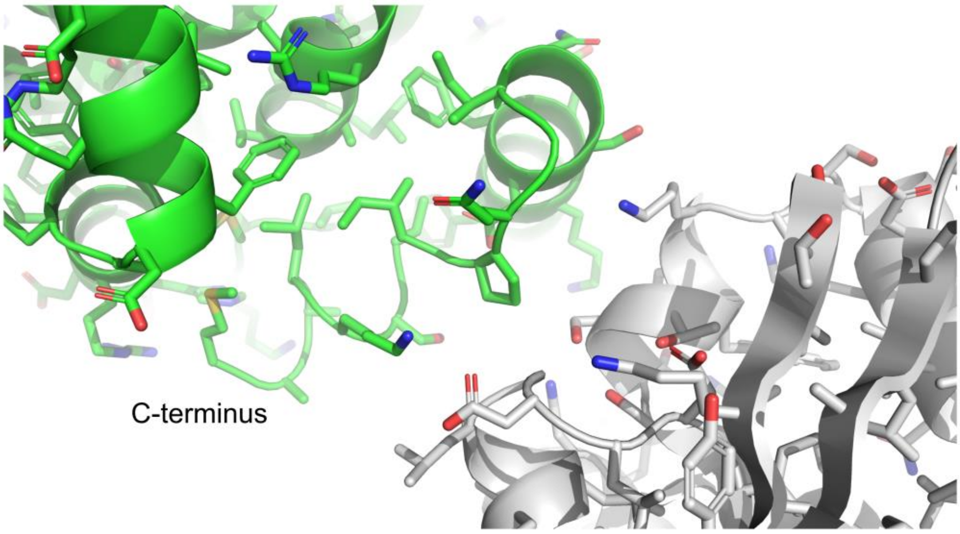


Figure SI-14: Crystal contact of HiSiaP M81A G71Q E78A. Interacting region between two HiSiaP proteins (green and grey) in the crystal structure of HiSiaP M81A G71Q E78A.

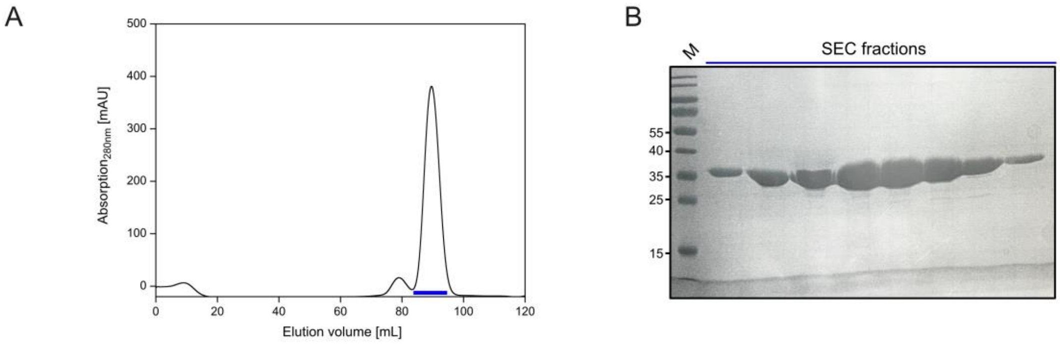


Figure SI-15: Purification of VcSiaP 54C 173C. A) SEC run of VcSiaP 54C 173C on a HiLoad SD 200 16/600 column (buffer: 50 mM Tris (pH 8), 50 mM NaCl, 1 mM TCEP). **B)** SDS-gel with samples from the SEC eluted fractions in A).

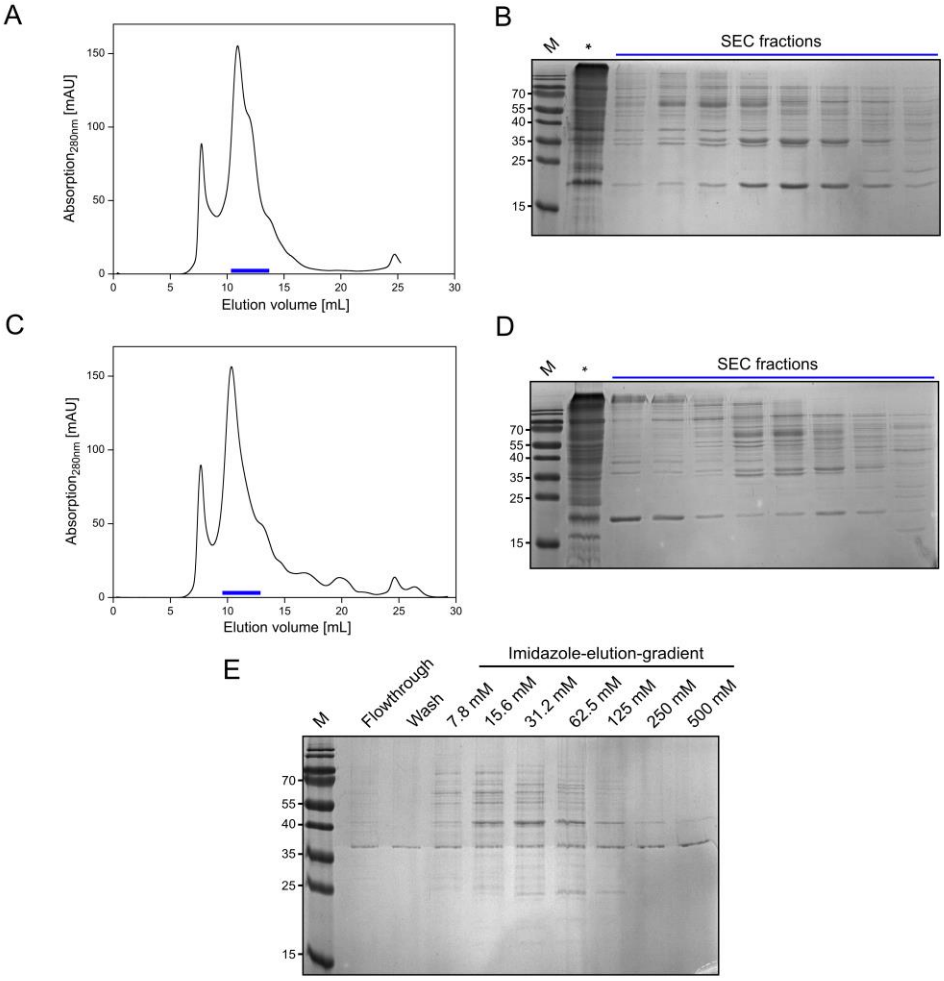


Figure SI-16: Extraction and Purification of VcSiaQM. A) SEC run of VcSiaQM on an SD 200 10/300 column after extraction with DDM (standard QM-buffer with DDM). **B)** SDS-gel with samples from the SEC eluted fractions in A) (* loaded sample). **C-D)** Same as A) but after extraction and with stabilization of VcSiaQM with FOS-Choline-14. **E)** SDS-PAGE with samples of different steps of a Ni^{2+}-AC from VcSiaQM purification with DDM detergent. The elution was performed with an imidazole concentration gradient, as mentioned in the figure.

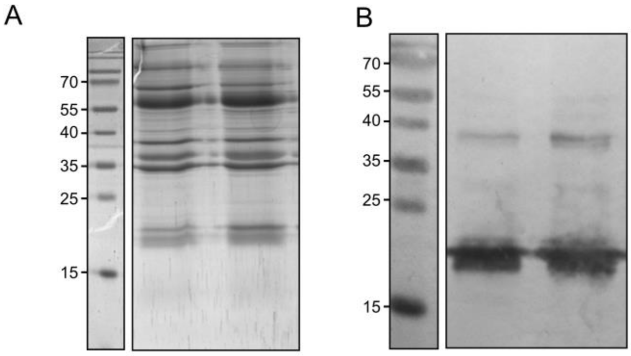
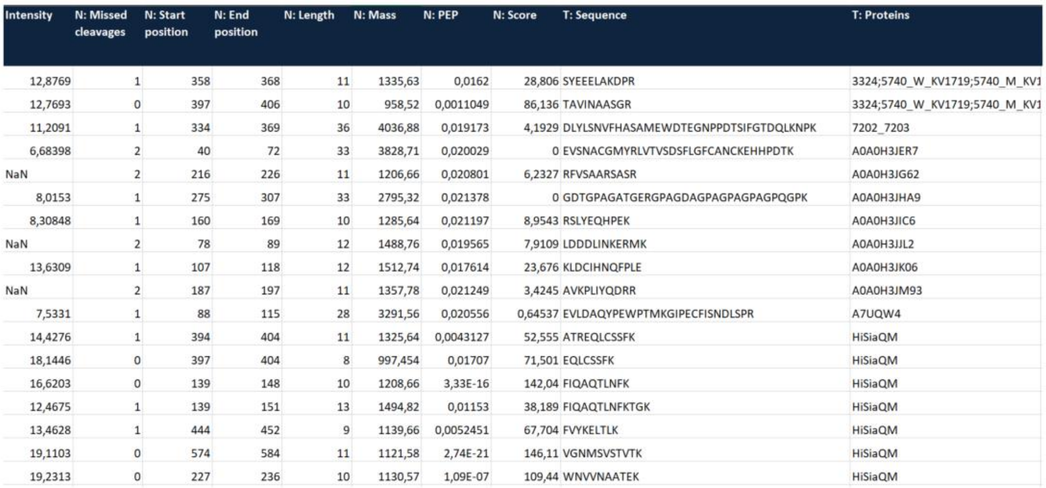


Figure SI-17: VcSiaQM complex formation. A) SDS-PAGE with samples from SEC fractions after VcSiaQM purification in DDM detergent. **B)** Same samples as in A) but detected with His-specific antibodies on a western blot.



Intensity	N: Missed cleavages	N: Start position	N: End position	N: Length	N: Mass	N: PEP	N: Score	T: Sequence	T: Proteins
12,8769	1	358	368	11	1335,63	0,0162	28,806	SYEEELAKDPR	3324;5740_W_KV1719;5740_M_KV1
12,7693	0	397	406	10	958,52	0,0011049	86,136	TAVINAASGR	3324;5740_W_KV1719;5740_M_KV1
11,2091	1	334	369	36	4036,88	0,019173	4,1929	DLYLSNVFHASAMEWDTEGNPPDTSIFGTDQLKNPK	7202_7203
6,68398	2	40	72	33	3828,71	0,020029	0	EVSNACGMYRLVTVSDSFLGFCANCKEHHPDTK	A0A0H3JER7
NaN	2	216	226	11	1206,66	0,020801	6,2327	RFVSAARSASR	A0A0H3JG62
8,0153	1	275	307	33	2795,32	0,021378	0	GDTGPAGATGERGPAGDAGPAGPAGPAGPQGPK	A0A0H3JHA9
8,30848	1	160	169	10	1285,64	0,021197	8,9543	RSLYEQHPEK	A0A0H3JIC6
NaN	2	78	89	12	1488,76	0,019565	7,9109	LDDDLINKERMK	A0A0H3JJL2
13,6309	1	107	118	12	1512,74	0,017614	23,676	KLDCIHNQFPLE	A0A0H3JK06
NaN	2	187	197	11	1357,78	0,021249	3,4245	AVKPLIYQDRR	A0A0H3JM93
7,5331	1	88	115	28	3291,56	0,020556	0,64537	EVLDAQYPEWPTMKGIPECFISNDLSPR	A7UQW4
14,4276	1	394	404	11	1325,64	0,0043127	52,555	ATREQLCSSFK	HiSiaQM
18,1446	0	397	404	8	997,454	0,01707	71,501	EQLCSSFK	HiSiaQM
16,6203	0	139	148	10	1208,66	3,33E-16	142,04	FIQAQTLNFK	HiSiaQM
12,4675	1	139	151	13	1494,82	0,01153	38,189	FIQAQTLNFKTGK	HiSiaQM
13,4628	1	444	452	9	1139,66	0,0052451	67,704	FVYKELTLK	HiSiaQM
19,1103	0	574	584	11	1121,58	2,74E-21	146,11	VGNMSVSTVTK	HiSiaQM
19,2313	0	227	236	10	1130,57	1,09E-07	109,44	WNVVNAATEK	HiSiaQM

Figure SI-18: Mass identification of HiSiaQM. Detected peptides of the main protein band from mass identification in Figure 4-4. Identification was performed from the CECAD/CMMC Proteomics Facility (University of Cologne).

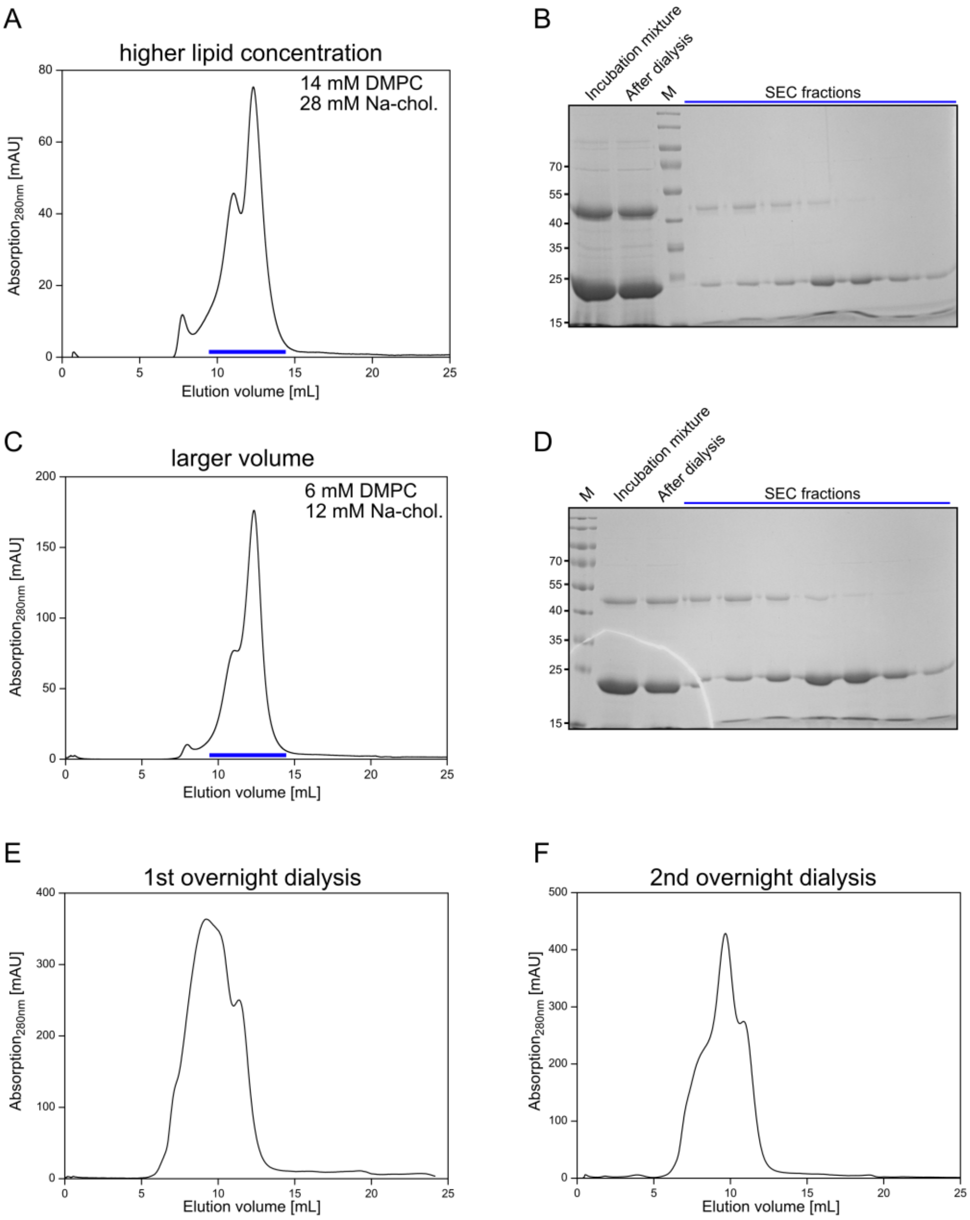


Figure SI-19: MSP-reconstitution of HiSiaQM. A) SEC run of HiSiaQM after MSP-nanodisc reconstitution with varied DMPC/Na-cholate concentration (here 14/28 mM, standard 6/12 mM) on an SD 200 10/300 column (standard QM-buffer without DDM). **B)** SDS-gel with samples from the reconstitution dialysis and SEC eluted fractions in A). **C-D)** Same as A-B) but with 2.5 times larger reconstitution volume. **E)** SEC run of HiSiaQM after MSP-nanodisc reconstitution at standard conditions on an SD 200 10/300 column (standard QM-buffer without DDM). **F)** Same protein reconstitution as in E) but with two nights of dialysis.

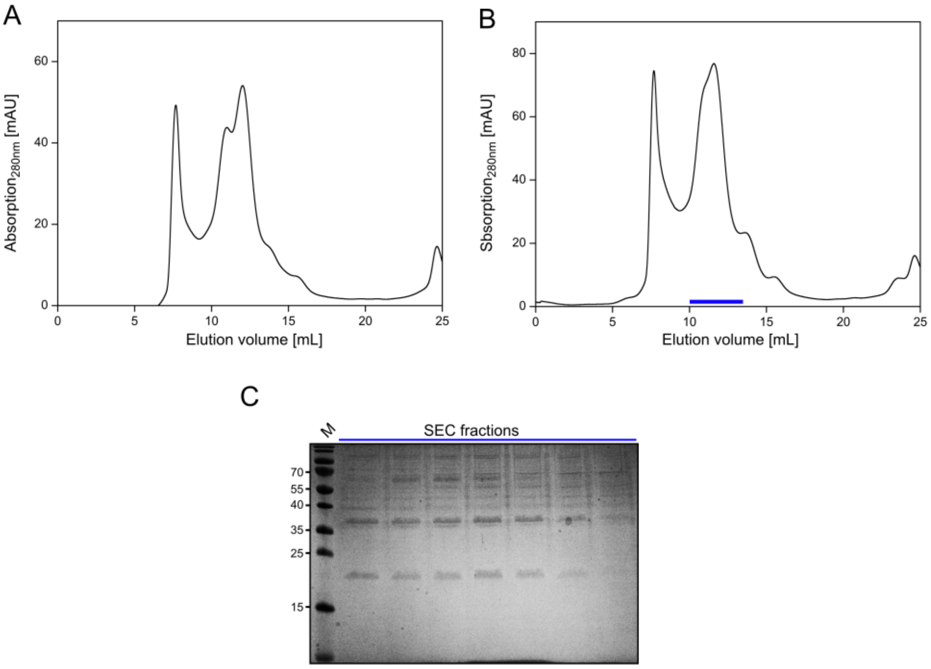


Figure SI-20: Spin labelling of VcSiaQM. A) SEC run of VcSiaQM on an SD 200 10/300 column after extraction with DDM (standard QM-buffer with DDM). **B)** Same as A) but after incubation of VcSiaQM with spin label MTSSL. **C)** SDS-gel with samples from the SEC eluted fractions in B).

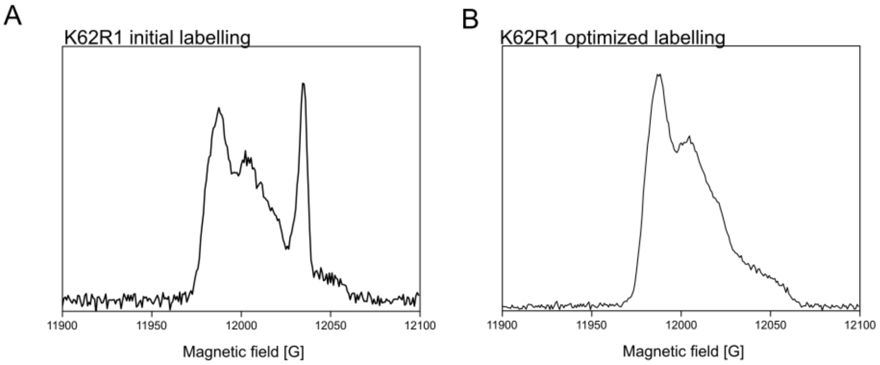


Figure SI-21: Optimization of HiSiaQM spin labelling. A) 2PESEEM spectrum of labelled HiSiaQM K62R1 mutant with initial labelling strategy. **B)** Same as A) but with optimized, new labelling strategy.

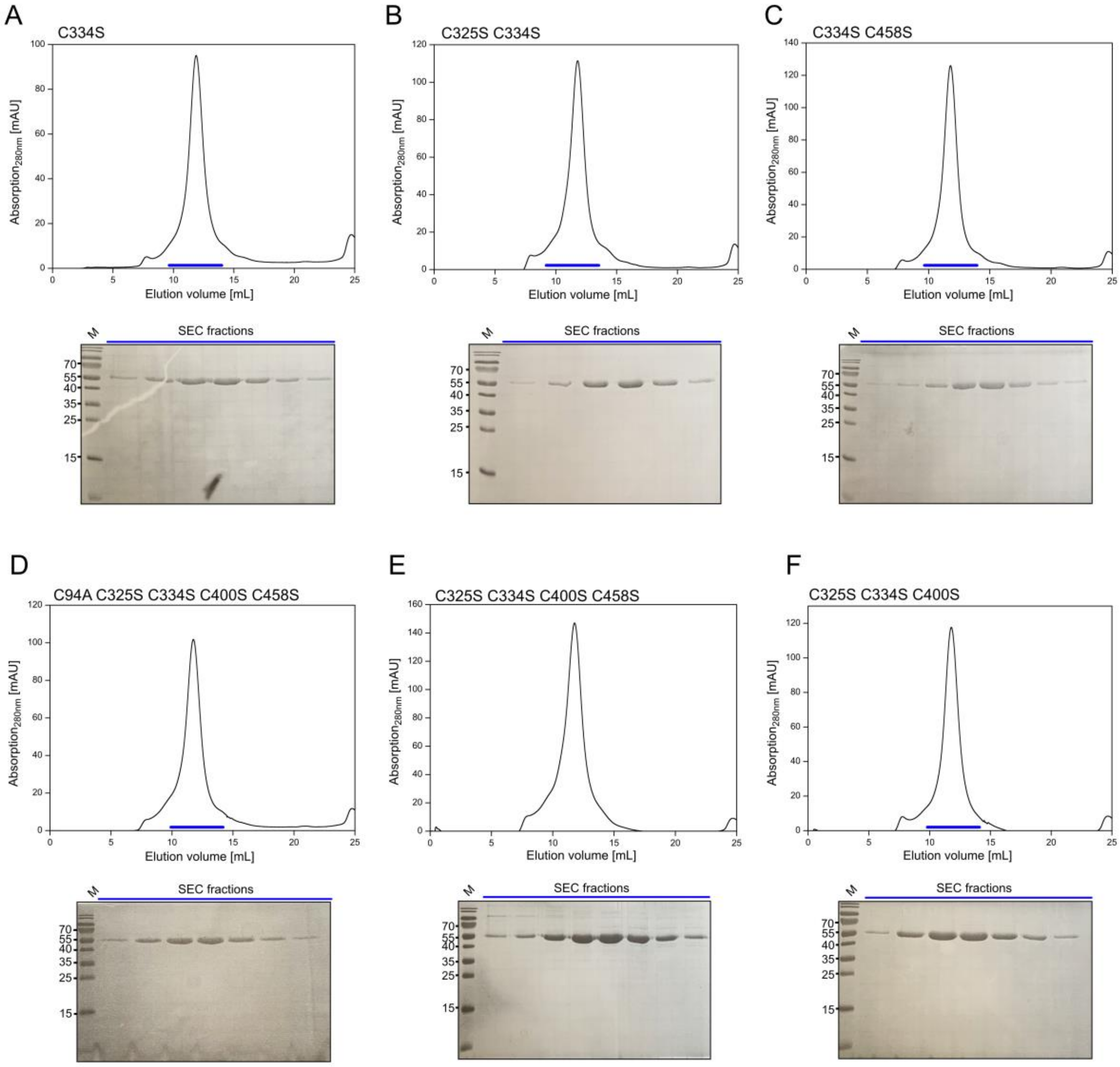


Figure SI-22: Cysteine-deletion mutants of HiSiaQM. A) SEC run of HiSiaQM C334S on an SD 200 10/300 column with standard DDM protein buffer after labelling with MTSSL (top). SDS-gel with samples from the SEC eluted fractions (bottom). **C-F)** Same as A) but for other HiSiaQM mutants.

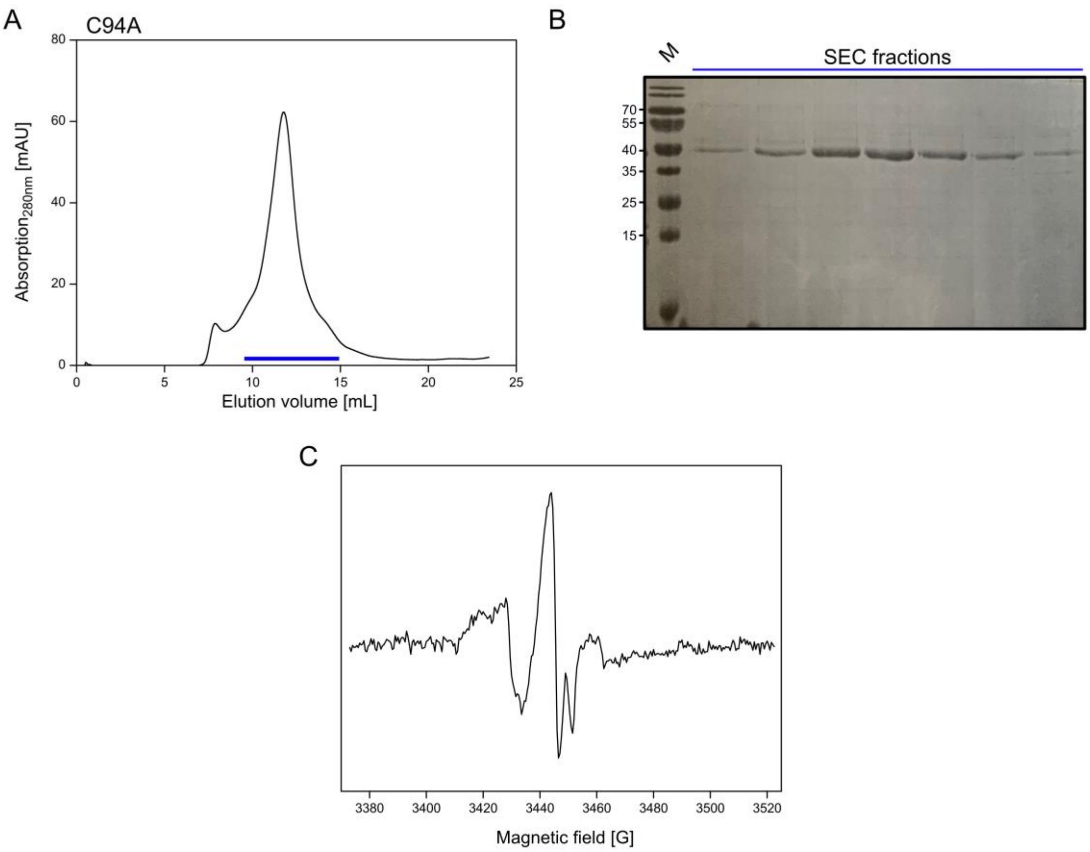


Figure SI-23: Spin labelling approach of HiSiaQM C94A. A) SEC run of HiSiaQM C94A after incubation with MTSSL on an SD 200 10/300 column with standard buffer, supplemented with DDM. **B)** SDS-gel with samples from the SEC eluted fractions in A). **C)** X-band RT cw-EPR spectrum of a HiSiaQM C94A. The spectrum was recorded as for the other mutants in Figure 5-3.

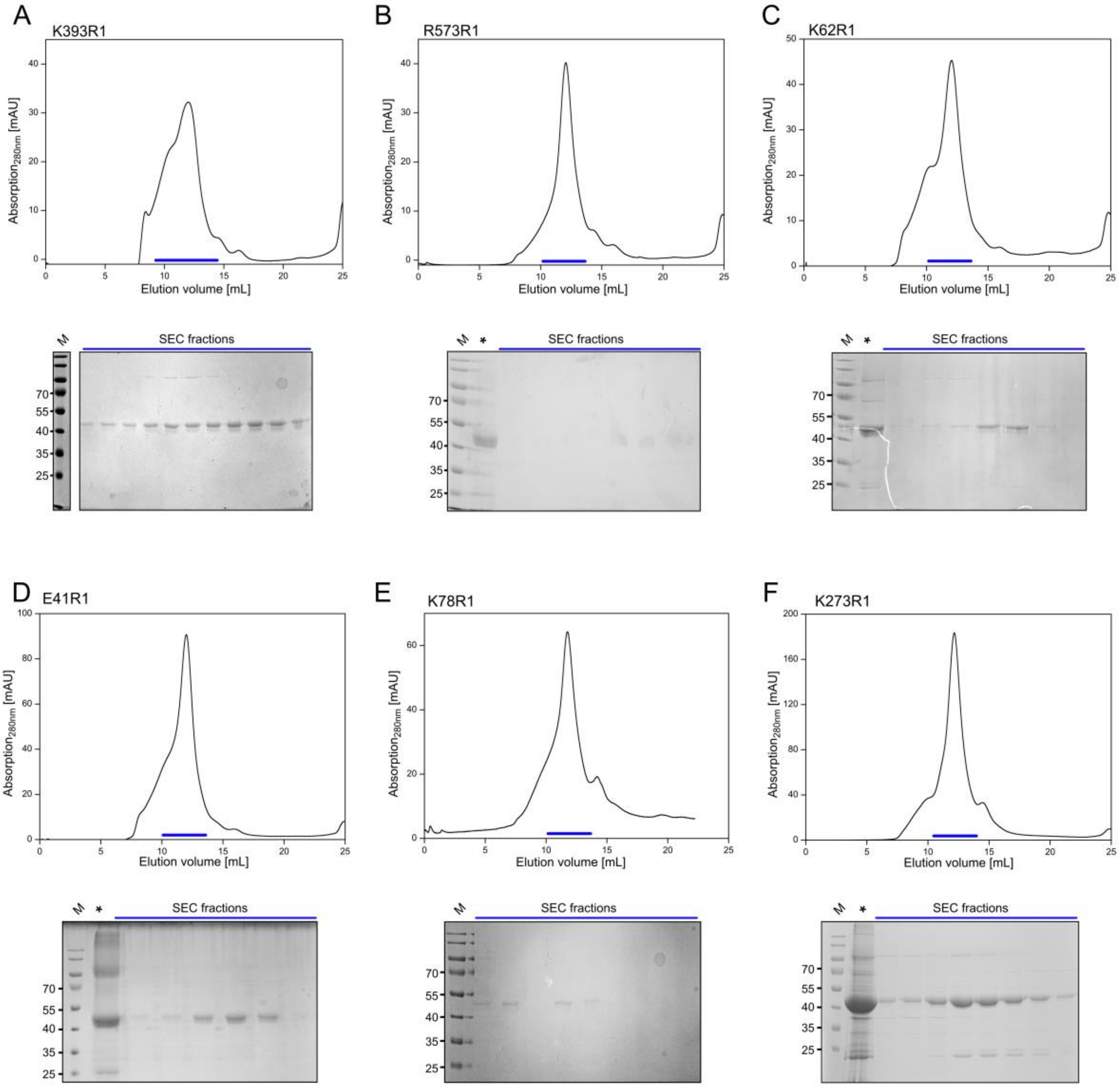


Figure SI-24: Purification of single cysteine mutants of HiSiaQM. A) SEC run of HiSiaQM single cysteine mutant K393R1 on an SD 200 10/300 column with standard DDM protein buffer, after labelling with MTSSL (top). SDS-gel with samples from the SEC eluted fractions (* loaded sample) (bottom). **C-F)** Same as A) but for other labelled single cysteine mutants.

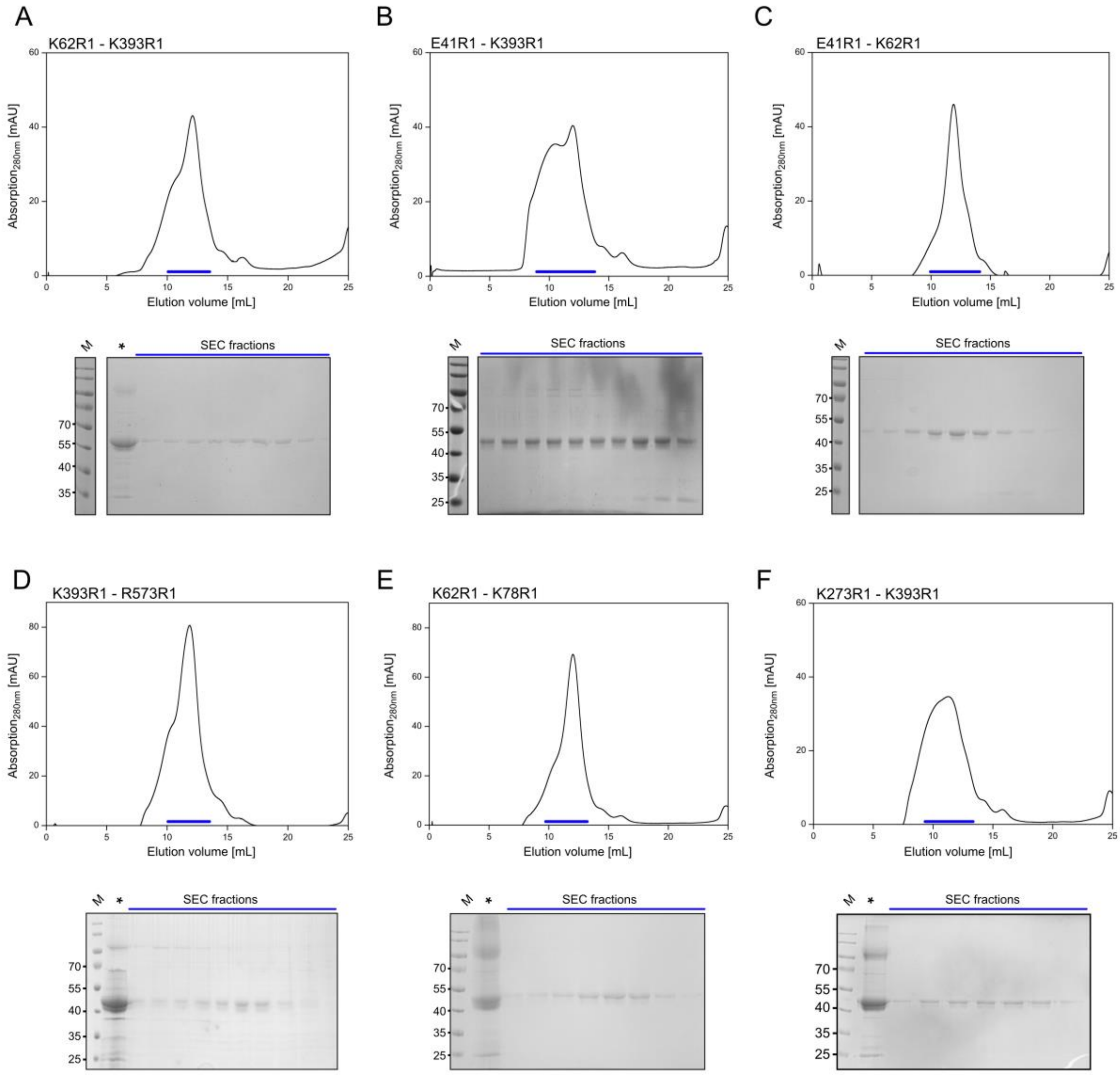


Figure SI-25: Purification of double cysteine mutants of HiSiaQM. A) SEC run of HiSiaQM double mutant K62R1-K393R1 on an SD 200 10/300 column with standard DDM protein buffer, after labelling with MTSSL (top). SDS-gel with samples from the SEC eluted fractions (* loaded sample) (bottom). **C-F)** Same as A) but for other labelled double mutants.

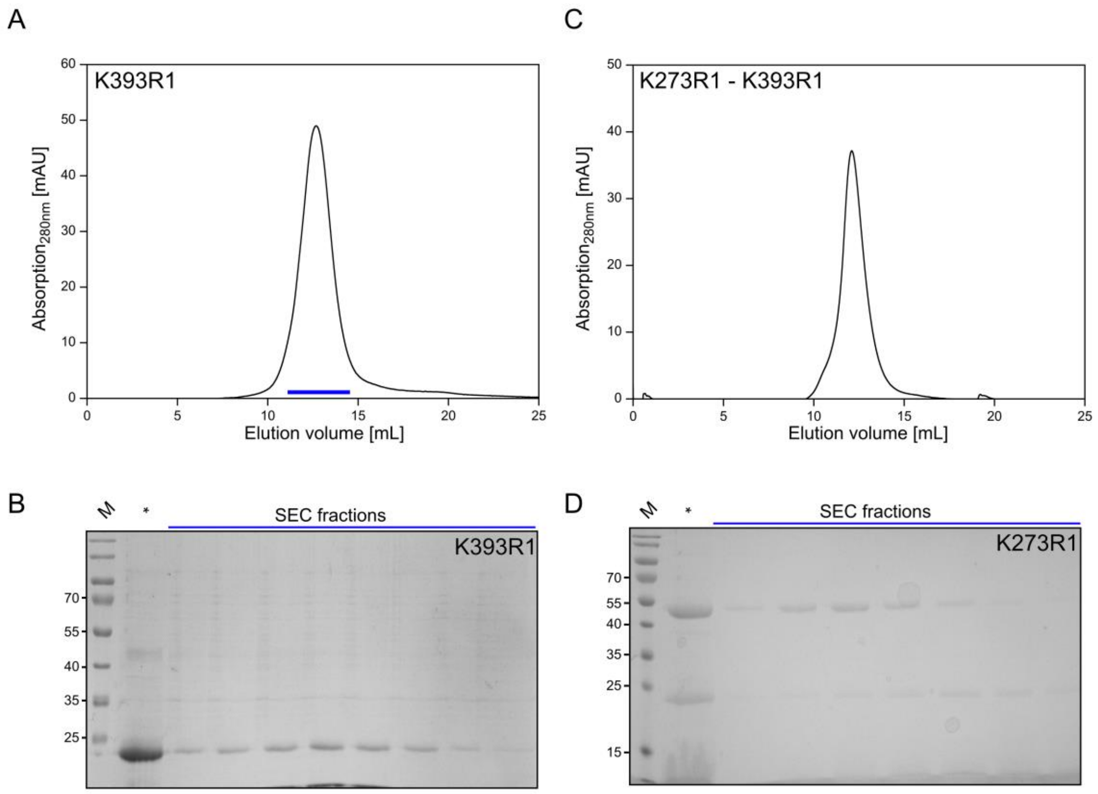


Figure SI-26: MSP-nanodisc reconstitution with spin labelled HiSiaQM. A) SEC run of labelled HiSiaQM K393R1 after reconstitution in MSP-nanodiscs on an SD 200 10/300 column with standard buffer, supplemented with DDM. **B)** SDS-gel with samples from the SEC eluted fractions in A) (* loaded sample). **C)** Same as A) but with labelled HiSiaQM K273R1-K393R1. **D)** SDS-gel with samples from the SEC eluted fractions from Figure 5-6 A (* loaded sample).

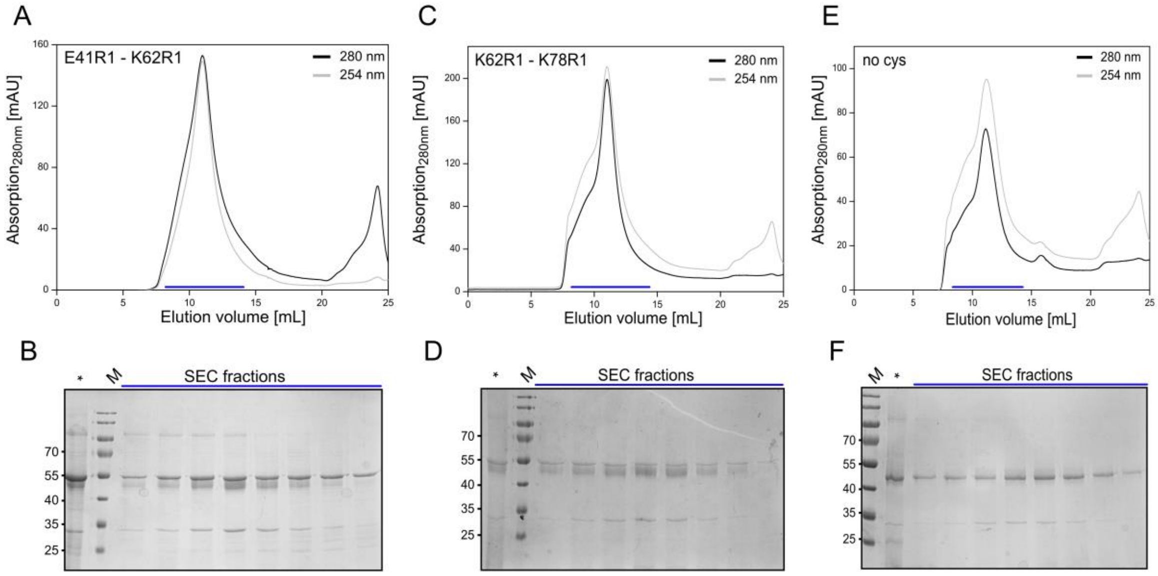


Figure SI-27: Purifications of HiSiaQM in SMA-nanodiscs. A) SEC run of polymer-bounded HiSiaQM E41R1-K62R1 after incubation with MTSSL on an SD 200 10/300 column with standard buffer, without DDM. Additionally, the absorption of the eluted solution at 254 nm is given. **B)** SDS-gel with samples from the SEC eluted fractions in A) (* loaded sample). **C-D)** Same as A) and B) but for polymer-bounded HiSiaQM K62R1-K78R1. **E-D)** Same as A) and B) but for MTSSL supplemented non-cysteine mutant HiSiaQM C94A C325S C334S C400S C458S.

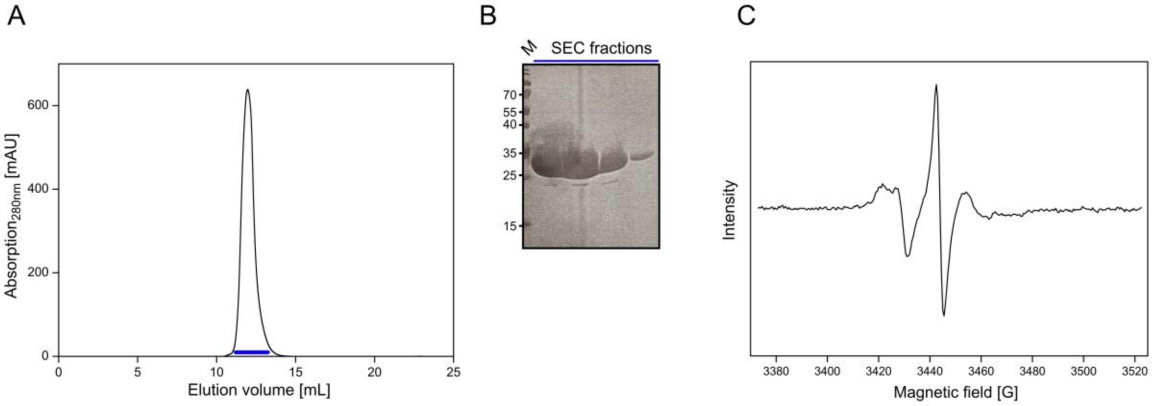


Figure SI-28: Purification and spin labelling of HiSiaP K254C. A) SEC run of HiSiaP K254R1 on an SD 75 10/300 with standard buffer. **B)** SDS-gel with samples from the SEC eluted fractions in A). **C)** X-band RT cw-EPR spectrum of HiSiaP K254R1 mutant from A) and B). The spectrum was recorded with 25 µM protein concentration and 350 scans, as described in the methods.

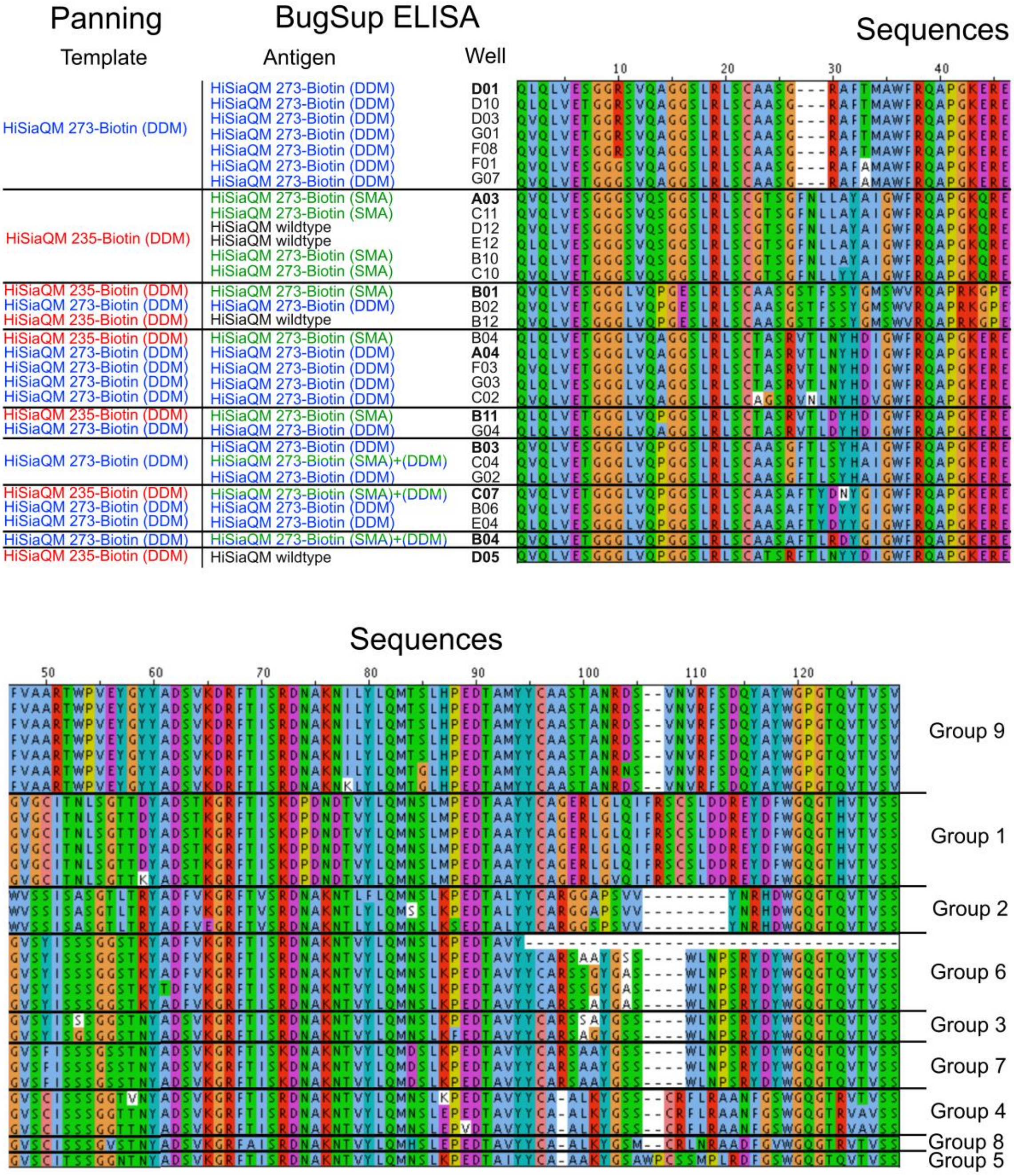


Figure SI-29: Identified nanobodies for HiSiaQM. A) Sequence alignment of the identified nanobodies, calculated with Clustal Omega (Sievers et al., 2011) and illustrated with Jalview (Waterhouse et al., 2009). The templates and antigens that were used for identification of the corresponding nanobodies are mentioned. The nanobody that was used as representative for each group is highlighted with bold letters in the corresponding well.

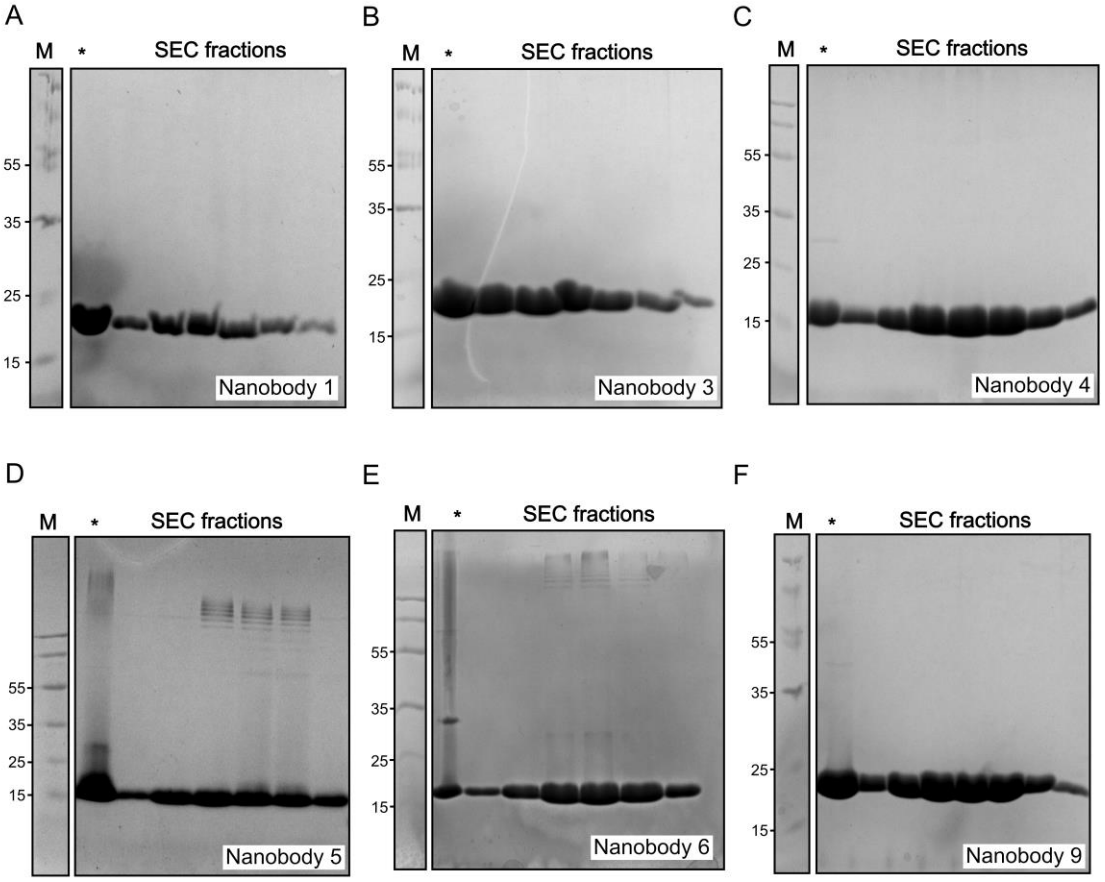


Figure SI-30: Large scale purification of nanobodies. A-F) SDS-gel with samples from the SEC loaded sample (*) and the eluted fractions.

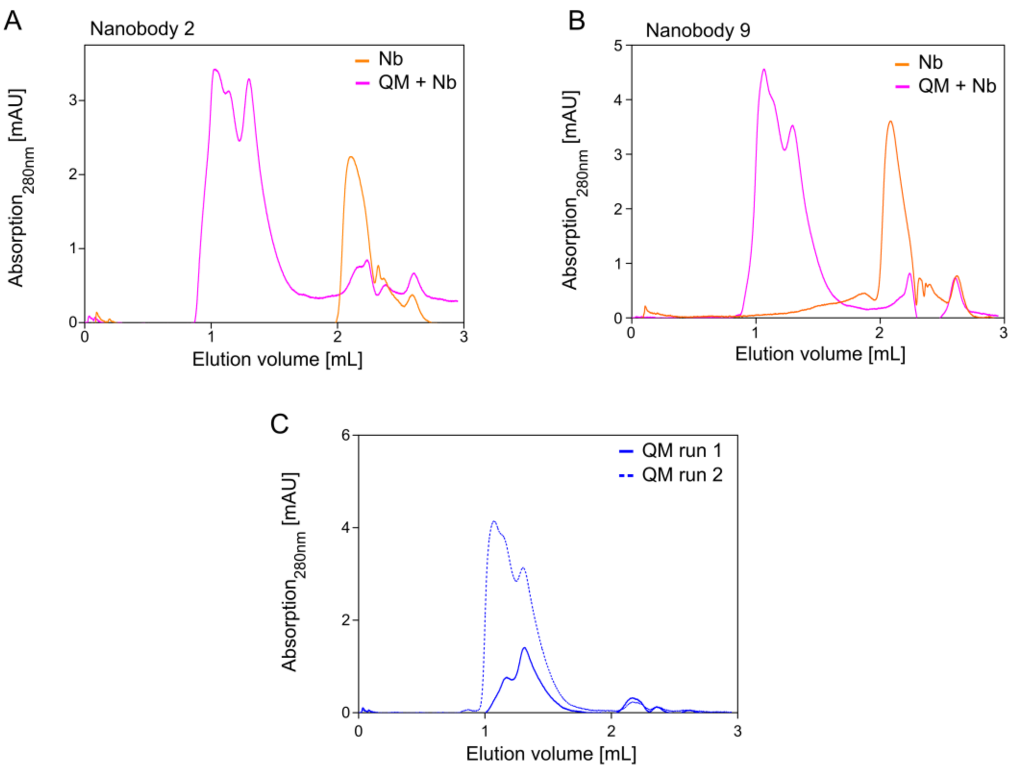


Figure SI-31: SEC-analysis of nanobody binding with HiSiaQM in SMA polymer nanodiscs. A-B) SECs of single component runs with two different nanobodies (orange) and after addition of HiSiaQM in SMA-nanodiscs (magenta) (SD 200 3.2/300 column, ÄKTA pure device). **C)** Two SEC runs of HiSiaQM in SMA-nanodiscs with identical amounts of loaded protein (SD 200 3.2/300 column, ÄKTA pure device).

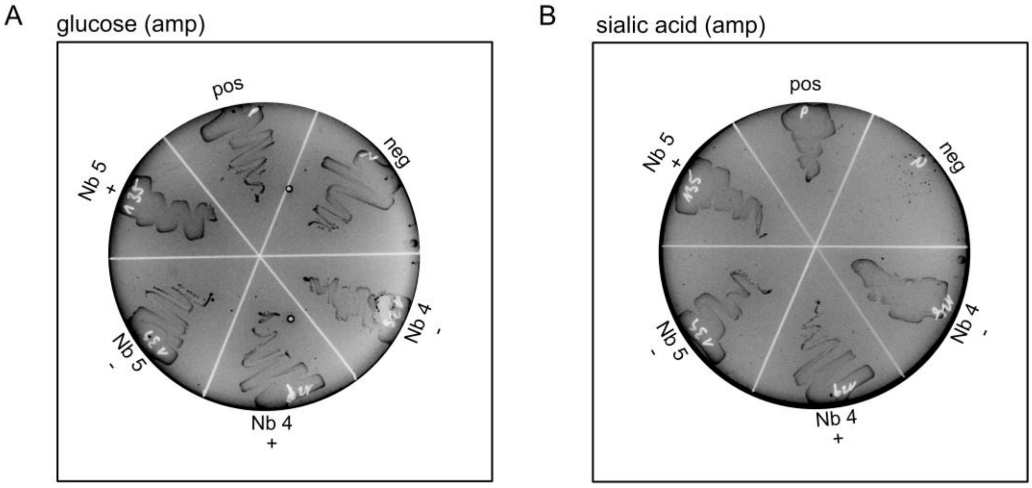


Figure SI-32: TRAP transporter growth assay with nanobodies on agar plates. A) Plates with M9-minimal agar and glucose as nutrient, supplemented with antibiotic ampicillin. The positive (pos) and negative sample (neg) are shown and the nanobody with or without pelB signal sequence are marked with + or -, respectively. **B)** Same as A) but with sialic acid as nutrient source.

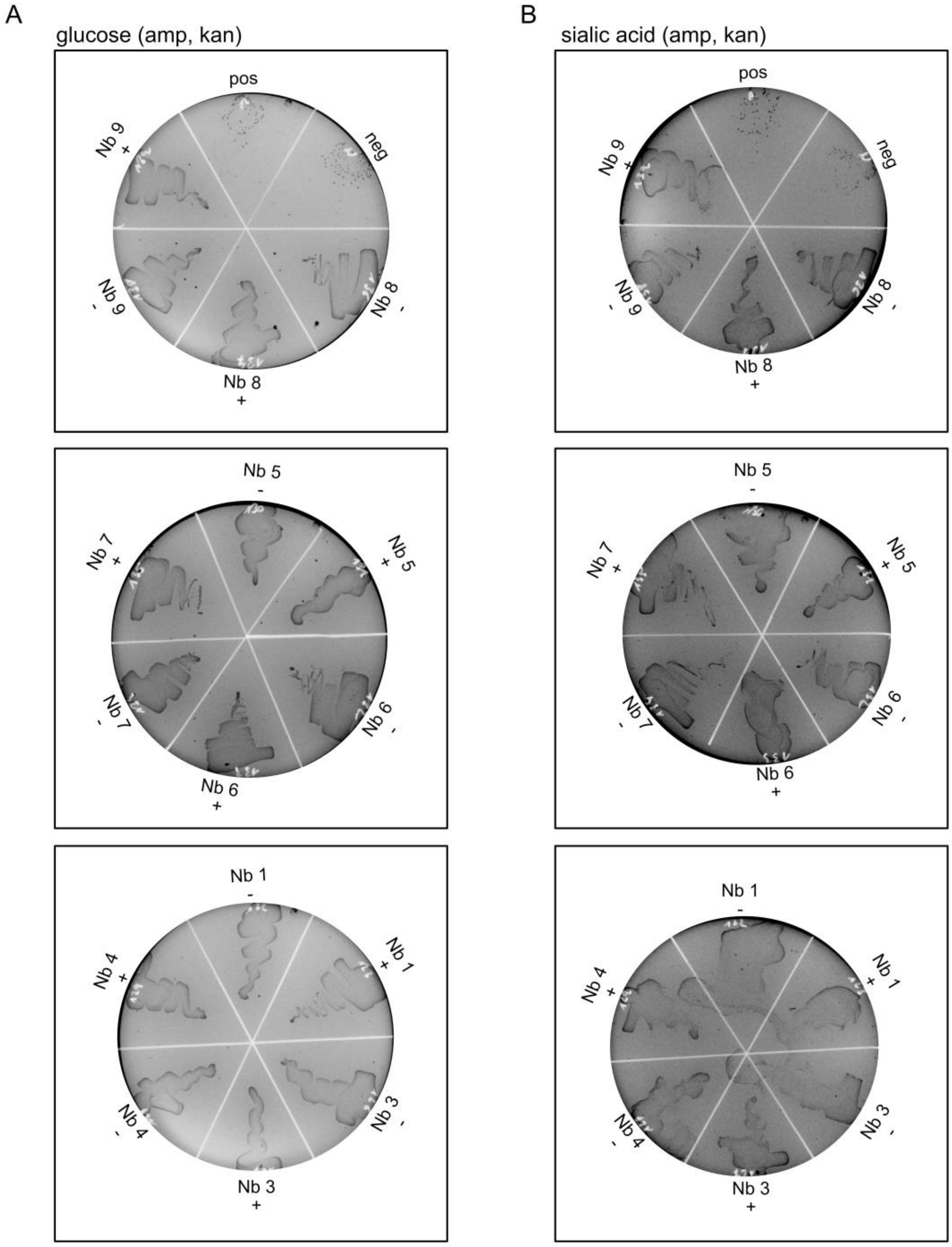


Figure SI-33: TRAP transporter growth assay with nanobodies on agar plates. A) Three plates with M9-minimal agar and glucose as nutrient, supplemented with antibiotics ampicillin and kanamycin. The positive (pos) and negative sample (neg) are illustrated and the nanobody with or without pelB signal sequence are marked with + or -, respectively. **B)** Same as A) but with sialic acid as nutrient source.

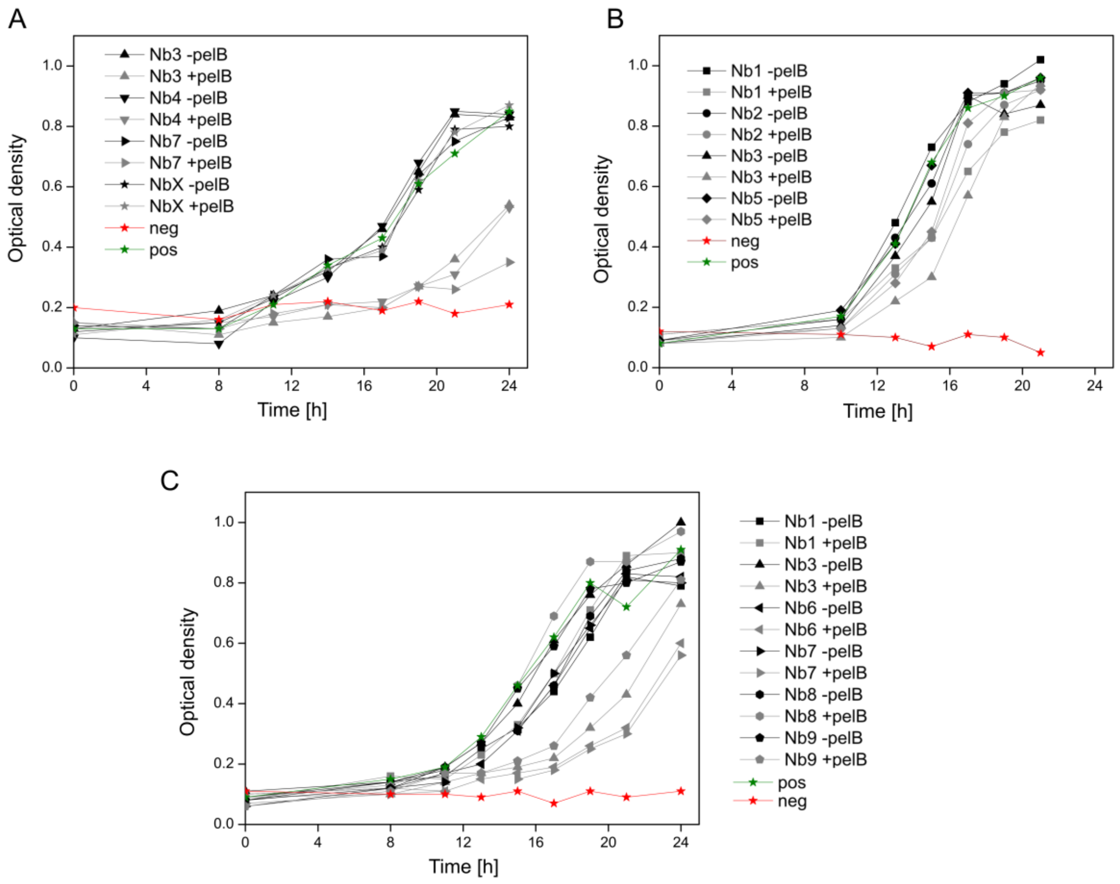


Figure SI-34: TRAP transporter growth assay with nanobodies in liquid medium. A) Growth rates (optical density) of *ΔnanT E. coli* cells in M9-minimal medium with sialic acid. Cells were supplemented with the sialic acid TRAP transporter HiSiaPQM and nanobodies with signal sequence (+pelB) or without (-pelB) (nanobody X is an unspecific control nanobody). For positive control (green), a cell culture without nanobody plasmid was used and for negative control (red), no sialic acid transporter or nanobody was transformed into the cells (just empty ampicillin vector for resistance). The data represent a repetition experiment from Figure 6-13. **B-C)** Same as A) but for cells with different supplemented nanobodies.

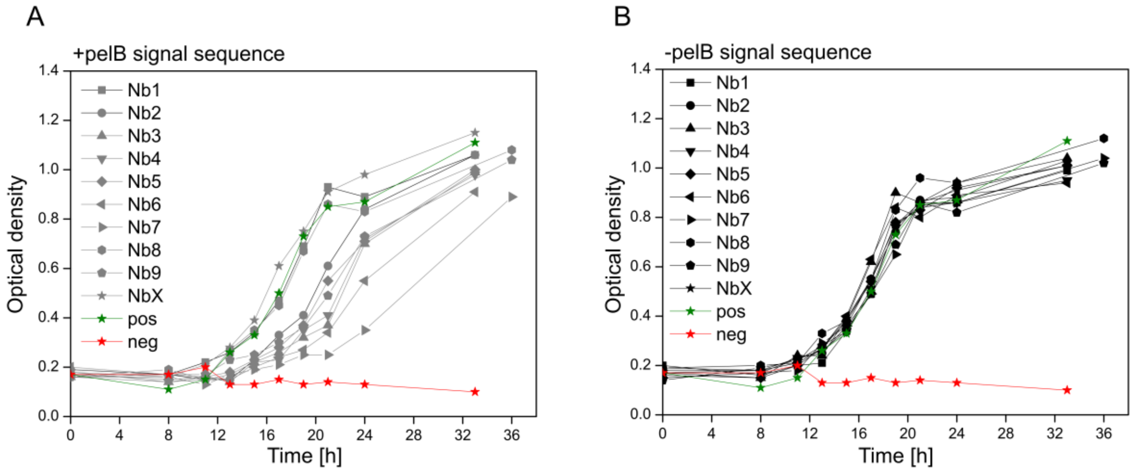


Figure SI-35: TRAP transporter growth assay with nanobodies in liquid medium. A) Growth rates (optical density) of *ΔnanT E. coli* cells in M9-minimal medium with sialic acid. Cells were supplemented with the sialic acid TRAP transporter HiSiaPQM and nanobodies with pelB signal sequence (+pelB) for periplasm transport after expression (nanobody X is an unspecific control nanobody). For positive control (green), a cell culture without nanobody plasmid was used and for negative control (red), no sialic acid transporter or nanobody was transformed into the cells (just empty ampicillin vector for resistance). The data are similar as in Figure 6-13 but for a longer time-period. **B)** Same as A) but the nanobodies were not fused to the pelB sequence (-pelB) to avoid transport into the periplasm.

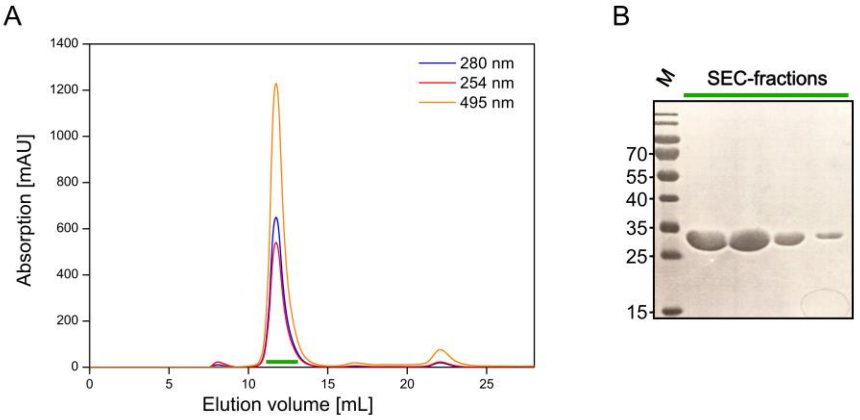


Figure SI-36: Purification and fluorophore labelling of HiSiaP K254C. A) SEC run of HiSiaP K254-fluorescein after incubation with fluorescein-5-maleimide (SD 75 10/300 with standard buffer). **B)** SDS-gel of samples from the SEC eluted fractions in A).

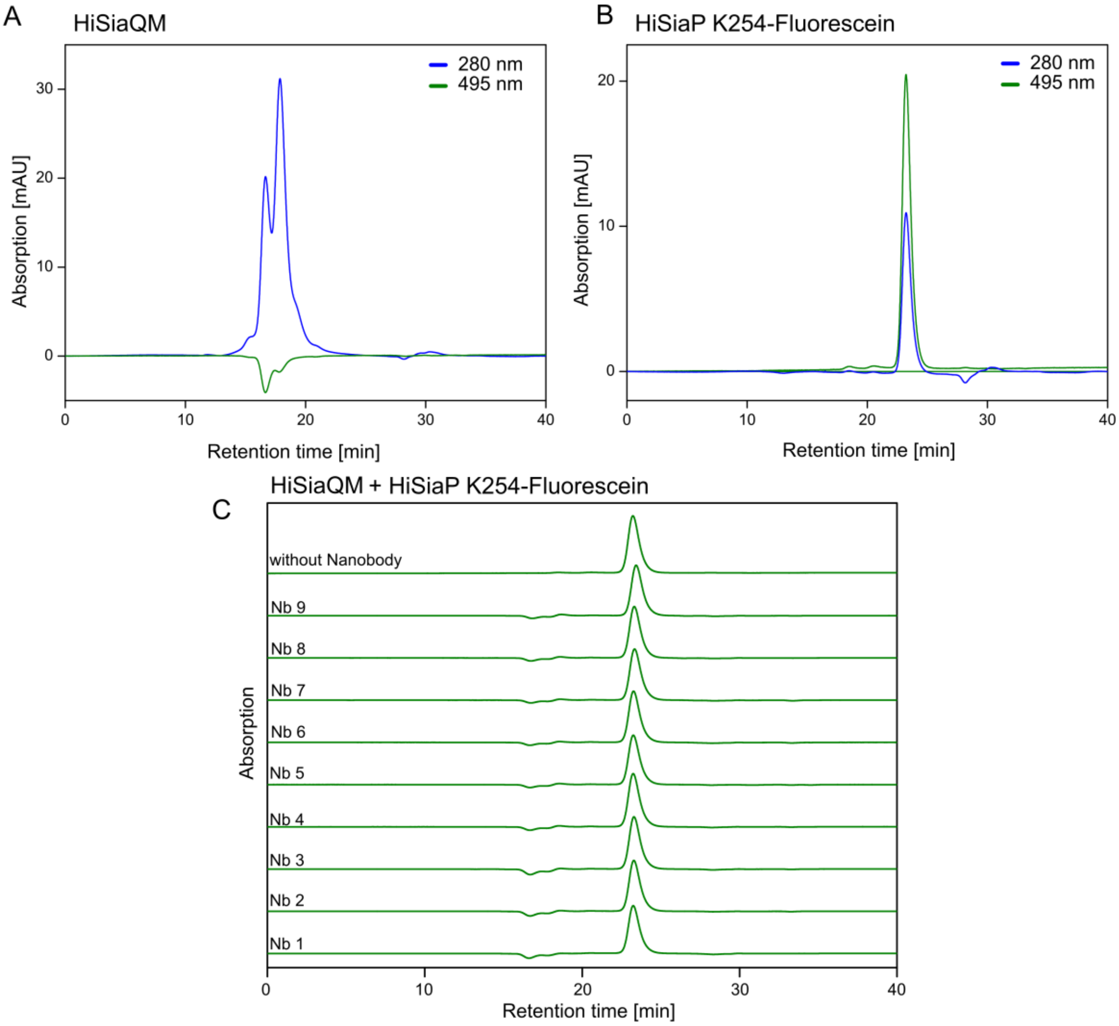


Figure SI-37: Interaction of P- and QM-domains after addition of nanobodies. A) Size-exclusion chromatogram of HiSiaQM with detection of elution at 280 nm and 495 nm (SD 200 3.2/300, Agilent HPLC). The standard buffer was supplemented with DDM and used without sialic acid (with supplementation of sialic acid in Figure 6-14). **B)** Same as A) but with HiSiaP K254C-fluorescein. **C)** Same set-up as A) but just the absorption at 495 nm is shown. The chromatograms correspond to a mixed sample of HiSiaQM and HiSiaP K254C-fluorescein and selective addition of each nanobody (1-9). The ratio between the components were chosen to 1:1:2 for HiSiaQM to nanobody to HiSiaP.

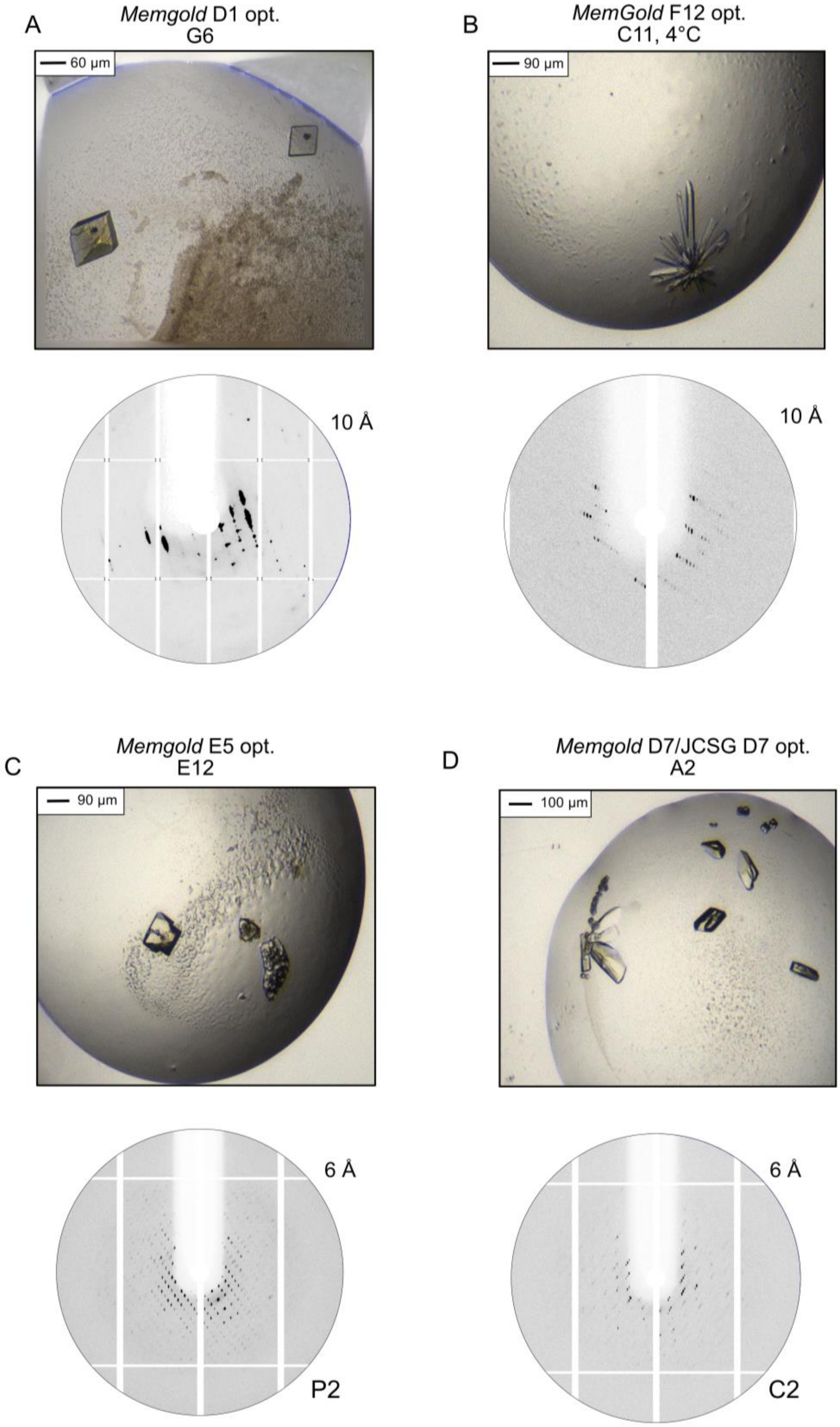


Figure SI-38: Optimized crystals and diffraction images of HiSiaQM with nanobodies 3 and 4. A-D) Top: Conditions with resulting crystals in a sitting drop, imaged shortly before collecting for diffraction. Bottom: Diffraction images, the limiting, outer circles represent the resolution of the diffraction image. In the case of a possible indexing, the resulting space group of the diffraction images are given.

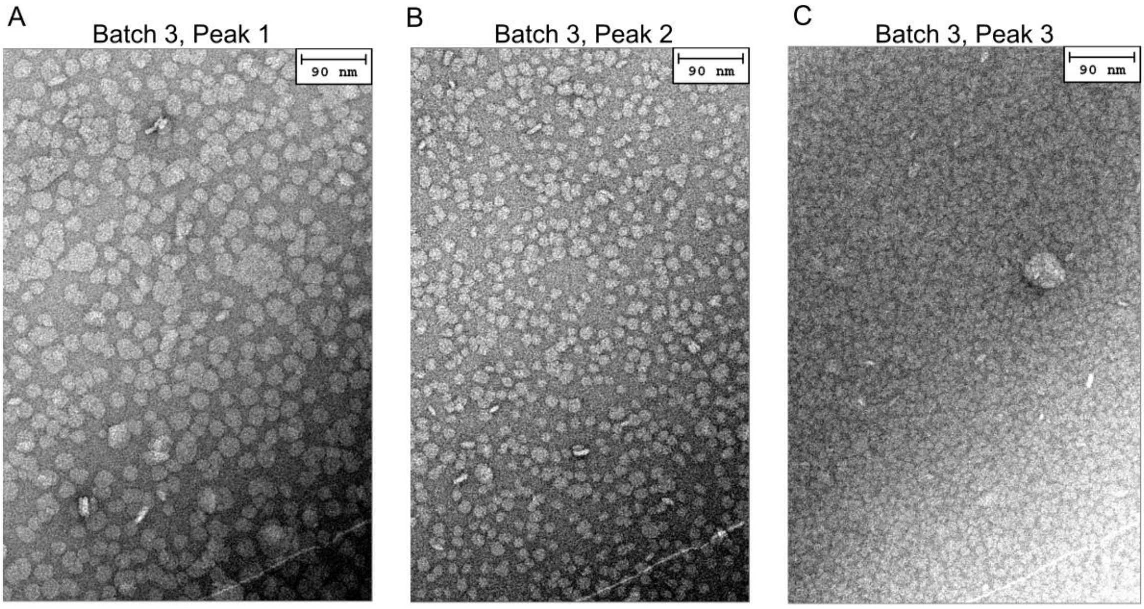
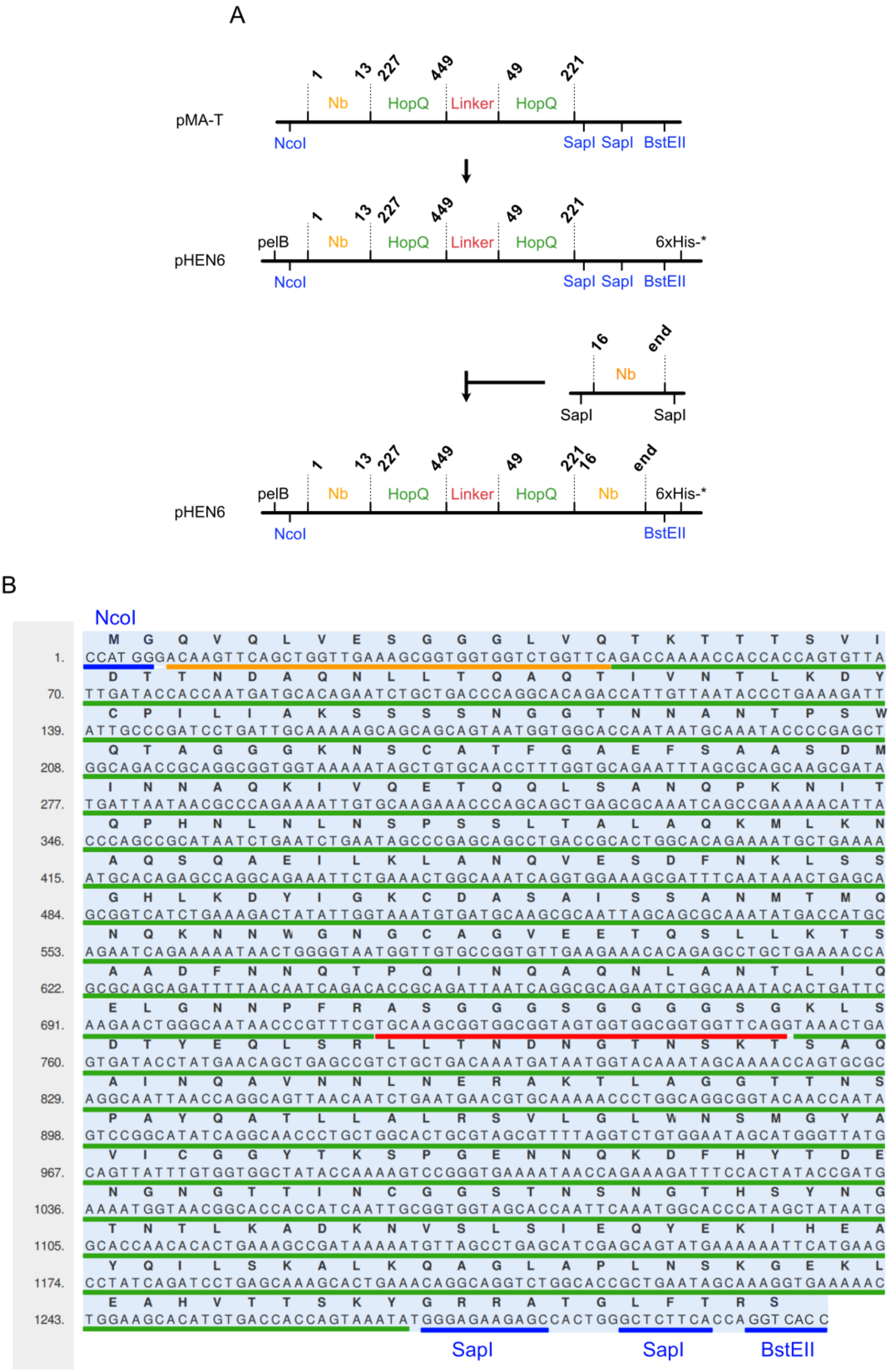


Figure SI-39: Negative stain EM of HiSiaQM in MSP-nanodiscs. A) EM images with negative stained HiSiaQM, reconstituted into MSP-nanodiscs. The sample was taken from the first SEC-fraction in Figure 7-4 D. The solution was stained with 2% uranyl acetate on glow charged grids. **B-C)** Same as A) but from different protein peak in the SEC run, as mentioned in the figure and on the chromatogram in Figure 7-4 D.



```
      M  G  Q  V  Q  L  V  E  S  G  G  G  L  V  Q  T  K  T  T  T  S  V  I
1.    CCAT GGGACAAGTTCAGCTGGTTGAAAGCGGTGGTGGTCTGGTTCAGACCAAAACCACCACCAGTGTTA
      D  T  T  N  D  A  Q  N  L  L  T  Q  A  Q  T  I  V  N  T  L  K  D  Y
70.   TTGATACCACCAATGATGCACAGAATCTGCTGACCCAGGCACAGACCATTGTTAATACCCTGAAAGATT
      C  P  I  L  I  A  K  S  S  S  S  N  G  G  T  N  N  A  N  T  P  S  W
139.  ATTGCCCGATCCTGATTGCAAAAAGCAGCAGCAGTAATGGTGGCACCAATAATGCAAATACCCCGAGCT
      Q  T  A  G  G  G  K  N  S  C  A  T  F  G  A  E  F  S  A  A  S  D  M
208.  GGCAGACCGCAGGCGGTGGTAAAAATAGCTGTGCAACCTTTGGTGCAGAATTTAGCGCAGCAAGCGATA
      I  N  N  A  Q  K  I  V  Q  E  T  Q  Q  L  S  A  N  Q  P  K  N  I  T
277.  TGATTAATAACGCCCAGAAAATTGTGCAAGAAACCCAGCAGCTGAGCGCAAATCAGCCGAAAAACATTA
      Q  P  H  N  L  N  L  N  S  P  S  S  L  T  A  L  A  Q  K  M  L  K  N
346.  CCCAGCCGCATAATCTGAATCTGAATAGCCCGAGCAGCCTGACCGCACTGGCACAGAAAATGCTGAAAA
      A  Q  S  Q  A  E  I  L  K  L  A  N  Q  V  E  S  D  F  N  K  L  S  S
415.  ATGCACAGAGCCAGGCAGAAATTCTGAAACTGGCAAATCAGGTGGAAAGCGATTTCAATAAACTGAGCA
      G  H  L  K  D  Y  I  G  K  C  D  A  S  A  I  S  S  A  N  M  T  M  Q
484.  GCGGTCATCTGAAAGACTATATTGGTAAATGTGATGCAAGCGCAATTAGCAGCGCAAATATGACCATGC
      N  Q  K  N  N  W  G  N  G  C  A  G  V  E  E  T  Q  S  L  L  K  T  S
553.  AGAATCAGAAAAATAACTGGGGTAATGGTTGTGCCGGTGTTGAAGAAACACAGAGCCTGCTGAAAACCA
      A  A  D  F  N  N  Q  T  P  Q  I  N  Q  A  Q  N  L  A  N  T  L  I  Q
622.  GCGCAGCAGATTTTAACAATCAGACACCGCAGATTAATCAGGCGCAGAATCTGGCAAATACACTGATTC
      E  L  G  N  N  P  F  R  A  S  G  G  G  S  G  G  G  G  S  G  K  L  S
691.  AAGAACTGGGCAATAACCCGTTTCGTGCAAGCGGTGGCGGTAGTGGTGGCGGTGGTTCAGGTAAACTGA
      D  T  Y  E  Q  L  S  R  L  L  T  N  D  N  G  T  N  S  K  T  S  A  Q
760.  GTGATACCTATGAACAGCTGAGCCGTCTGCTGACAAATGATAATGGTACAAATAGCAAAACCAGTGCGC
      A  I  N  Q  A  V  N  N  L  N  E  R  A  K  T  L  A  G  G  T  T  N  S
829.  AGGCAATTAACCAGGCAGTTAACAATCTGAATGAACGTGCAAAAACCCTGGCAGGCGGTACAACCAATA
      P  A  Y  Q  A  T  L  L  A  L  R  S  V  L  G  L  W  N  S  M  G  Y  A
898.  GTCCGGCATATCAGGCAACCCTGCTGGCACTGCGTAGCGTTTTAGGTCTGTGGAATAGCATGGGTTATG
      V  I  C  G  G  Y  T  K  S  P  G  E  N  N  Q  K  D  F  H  Y  T  D  E
967.  CAGTTATTTGTGGTGGCTATACCAAAAGTCCGGGTGAAAATAACCAGAAAGATTTCCACTATACCGATG
      N  G  N  G  T  T  I  N  C  G  G  S  T  N  S  N  G  T  H  S  Y  N  G
1036. AAAATGGTAACGGCACCACCATCAATTGCGGTGGTAGCACCAATTCAAATGGCACCCATAGCTATAATG
      T  N  T  L  K  A  D  K  N  V  S  L  S  I  E  Q  Y  E  K  I  H  E  A
1105. GCACCAACACACTGAAAGCCGATAAAAATGTTAGCCTGAGCATCGAGCAGTATGAAAAAATTCATGAAG
      Y  Q  I  L  S  K  A  L  K  Q  A  G  L  A  P  L  N  S  K  G  E  K  L
1174. CCTATCAGATCCTGAGCAAAGCACTGAAACAGGCAGGTCTGGCACCGCTGAATAGCAAAGGTGAAAAAC
      E  A  H  V  T  T  S  K  Y  G  R  R  A  T  G  L  F  T  R  S
1243. TGGAAGCACATGTGACCACCAGTAAATATGGGAGAAGAGCCACTGGGCTCTTCACCA GGT CAC C
```

Figure SI-40: Construction of HiSiaQM nanobody-derived megabody. A) Cloning strategy of nanobody-derived megabody after (Uchański et al., 2019, Uchański et al., 2021). The colour-code corresponds to Figure 7-9. **B)** Designed and external synthesized gene, same colours as in A) (synthesized at GeneArt).

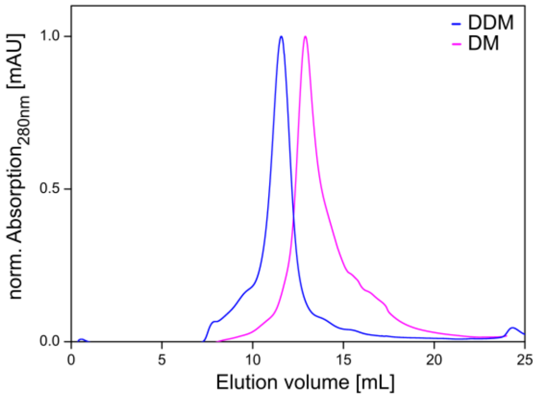


Figure SI-41: Purification of HiSiaQM in different detergents. SEC runs of HiSiaQM wildtype purification after extraction of the protein with DDM detergent and purification with DDM (blue) or with change to DM during Ni^{2+}-AC (magenta) (SD 200 10/300 column, standard buffer with corresponding detergent).