Thesis for the degree of Doctor of Philosophy
in the Natural Sciences

Structural and Functional Studies of Membrane Proteins

For Future Development of Antimicrobial Drugs

Elin Dunevall

UNIVERSITY OF GOTHENBURG

Department of Chemistry and Molecular Biology
Gothenburg, 2019

Thesis for the degree of Doctor of Philosophy
in the Natural Sciences

Structural and Functional Studies of Membrane Proteins
For Future Development of Antimicrobial Drugs

Elin Dunevall

Cover: The membrane proteins MraY and SiaT embedded in a lipid bilayer. The arrows represent the enzymatic reaction of MraY and transport of sialic acid in SiaT.

ISBN 978-91-7833-292-2 (Print)
ISBN 978-91-7833-293-9 (PDF)
Available online at http://hdl.handle.net/2077/58373

Department of Chemistry and Molecular Biology
Division of Biochemistry and Structural Biology
University of Gothenburg
SE-405 30, Göteborg, Sweden
Printed by BrandFactory AB
Göteborg, Sweden, 2019

"Even Miracles Take a Little Time"
Fairy Godmother, Cinderella

Abstract

Antibiotic resistance is a world-wide occurring problem which threatens human health. Without development of any new and effective antibiotics, the rapid growth of antibiotic-resistant bacterial infections could put society in a situation resembling the pre-antibiotic era when a simple lung infection could kill a human being. This thesis presents two venues for targeting antibiotic resistance.

Pathogenic bacteria present in mucus rich environments are able to utilize host-derived sialic acid either as an alternative food source or by incorporating sialic acid to their surface glycoconjugates as a way to evade the host´s immune system. Hence, molecular mimicry enables bacteria to secure an ecological niche for survival. Transport of scavenged sialic acid into the cytoplasm of bacteria occurs through specific membrane bound sialic acid transporters.

The cell wall is an essential protective barrier for bacteria. The membrane bound enzyme MraY catalyzes the synthesis of lipid I, an intermediate step in the biosynthesis of peptidoglycan, the cell wall of bacteria.

This thesis presents work aimed to structurally and functionally characterize sialic acid transporters and MraY for future development of antibacterial drugs. Starting with a broad approach for expression and purification of sialic acid transporters resulted in low-resolution diffracting crystals of the *Pasteurella multocida* sialic acid TRAP transporter. In addition the X-ray crystallography structure of the sialic acid transporter SiaT from *Proteus mirabilis* was determined at 1.95 Å resolution in a substrate-bound outward-open conformation revealing a new sodium site. Furthermore, SiaT transporters have been characterized *in vivo* and the sialic acid specificity has been characterized for SiaT from *Staphylococcus aureus*. Structural comparison between MraY and the human homologue GPT have highlighted regions where to modify the natural product inhibitor tunicamycin to selectively target MraY. Further characterization of tunicamycin analogues identified potent inhibitors with reduced eukaryotic toxicity.

Sammanfattning på svenska

Vi lever i en värld där vi är beroende av antibiotika för att bland annat kunna genomföra en rad olika medicinska ingrepp. Antibiotika förhindrar tillväxten av bakteriella infektioner. Tyvärr står världen inför ett stort problem då användandet av antibiotika missköts vilket leder till att bakterier som har utvecklat resistens mot antibiotikan har möjlighet att överleva. I denna avhandling presenteras två möjliga vägar för att utveckla nya antibakteriella läkemedel.

Patogena bakterier som lever i slemrika miljöer kan ta upp sialinsyra från sin omgivning och använda det antingen som föda eller för att inkorporera sialinsyran på dess lipopolysackarider som finns på bakteriens yta. Genom att täcka cellytan med sialinsyra undviker bakterien att bli upptäckt av immunförsvaret. För att kunna utnyttja sialinsyra behöver bakterierna transportera sialinsyra in i cellen, en process som sker med specifika sialinsyratransportörer.

En bakteriecell omges av en cellvägg som fungerar som en skyddande barriär och är nödvändig för bakteriens överlevnad. Cellväggen är uppbyggd av ett peptidoglykanlager och ett av enzymen som skapar det här lagret är det membranbundna proteinet MraY.

Följande avhandling presenterar struktur- och funktionsstudier av sialinsyratransportörer samt MraY som ett steg i att i framtiden kunna utveckla antibakteriella läkemedel som inhiberar dessa proteiner.

Acknowledgements

Starting from the beginning, little did I know that this PhD journey would include so much more than I had ever imagined. Achieving a PhD includes far more than what science has to teach us. During this journey, several people have supported me both professionally and personally and I am grateful to all of you!

To my examiner Richard, thank you for being supportive during difficult times and thank you for encouraging me to complete my PhD.

My supervisor Gisela, thank you for taking me on board half way through this journey! Your knowledge, calmness and support is truly inspiring. Thank you for bringing the fun back again!

Thank you all co-authors of the papers, this work would not have been the same without all the world-wide collaborations and a special thanks to Weixiao, Aviv, Arjan and my co-supervisor Jeff.

The AZ team, I really enjoyed working together with you all. Thank you Margareta, you have an eye for details which I truly appreciate. Jenny, thank you for introducing me to the MraY project.

Rhawnie, the road was full of obstacles, but we overcame them. Thank you for being there together with me! Rosie, I wish you all the best.

Elin and Andreas, thank you for working as master students with me on different projects.

To all of you who have walked up and down the corridors in the Lundberg lab, I am glad to have had the chance to get to know so many of you. I am truly thankful for have had the opportunity to travel the world with many of you to participate in beamtimes, workshops and conferences.

Jag vill även tacka mina vänner och min familj. Mamma, Pappa och mina älskade systrar, tack för allt, ni betyder så mycket för mig!

Johan, min klippa och trygga punkt i tillvaron. Tack för att du alltid finns där för mig och för vår son. Melvin, du har verkligen gett oss en extra dimension med livet. Jag älskar er!

Publications

This thesis consists of the following research papers:

Paper I: **Elin Dunevall**, Elin Claesson, Weixiao Y. Wahlgren, Rachel A. North, Aviv Paz, Rhawnie Caing-Carlsson, Parveen Goyal, Anne Farewell, Jeff Abramson, S. Ramaswamy, Renwick C. J. Dobson and Rosmarie Friemann

"A GFP-based approach to optimize expression and purification of SiaT sialic acid transporters for structure determination"

Manuscript, (2018)

Paper II: Weixiao Y. Wahlgren*, **Elin Dunevall***, Rachel A. North*, Aviv Paz, Mariafrancesca Scalise, Paola Bisignano, Johan Bengtsson-Palme, Parveen Goyal, Elin Claesson, Rhawnie Caing-Carlsson, Rebecka Andersson, Konstantinos Beis, Ulf J. Nilsson, Anne Farewell, Lorena Pochini, Cesare Indiveri, Michael Grabe, Renwick C. J. Dobson, Jeff Abramson, S. Ramaswamy and Rosmarie Friemann

* *These authors contributed equally*

"Substrate-bound outward-open structure of a Na^+-coupled sialic acid symporter reveals a new Na^+ site"

Nature Communications (2018) 9:1753
doi: 10.1038/s41467-018-04045-7

Paper III: Rachel A. North, Weixiao Y. Wahlgren, Daniela M. Remus, Mariafrancesca Scalise, Sarah A. Kessans, **Elin Dunevall**, Elin Claesson, Tatiana P. Soares da Costa, Matthew A. Perugini, S. Ramaswamy, Jane R. Allison, Cesare Indiveri, Rosmarie Friemann and Renwick C. J. Dobson

"The sodium sialic acid symporter from *Staphylococcus aureus* has altered substrate specificity"

Frontiers in Chemistry (2018) 6:233
doi: 10.3389/fchem.2018.00233

Paper IV: Rhawnie Caing-Carlsson, Parveen Goyal, Weixiao Y. Wahlgren, **Elin Dunevall**, S. Ramaswamy and Rosmarie Friemann

"Expression, purification and crystallization of a sialic acid tripartite ATP-independent periplasmic (TRAP) transporter"

Manuscript, (2018)

Paper V: Jenny Hering, **Elin Dunevall**, Margareta Ek and Gisela Brändén

"Structural basis for selective inhibition of antibacterial target MraY, a membrane-bound enzyme involved in peptidoglycan synthesis"

Drug Discovery Today (2018) vol. 23, no. 7, pp. 1426-1435
doi: 10.1016/j.drudis.2018.05.020

Paper VI: Jenny Hering[*], **Elin Dunevall**[*], Arjan Snijder, Michael A. Jackson, Trina M. Hartman, Neil P. J. Price, Gisela Brändén and Margareta Ek
[*] *These authors contributed equally*

"Inhibition of antibacterial target MraY by modified tunicamycins"

Manuscript, (2018)

Related papers that I have co-authored but that are not included in this thesis:

Paper VII: Rebecka Andersson, Cecilia Safari, Robert Dods, Eriko Nango, Rie Tanaka, Ayumi Yamashita, Takanori Nakane, Kensuke Tono, Yasumasa Joti, Petra Båth, **Elin Dunevall**, Robert Bosman, Osamu Nureki, So Iwata, Richard Neutze and Gisela Brändén

"Serial femtosecond crystallography structure of cytochrome *c* oxidase at room temperature"

Scientific Reports (2017) 7:4518
doi:10.1038/s41598-017-04817-z

Paper VIII: Amit Sharma, Linda Johansson, **Elin Dunevall**, Weixiao Y. Wahlgren, Richard Neutze and Gergely Katona

"Asymmetry in serial femtosecond crystallography data"

Acta Crystallogr. A (2017) vol A73, pp. 93–101
doi:10.1107/S2053273316018696

Contribution

Paper I: I planned the experiments, designed the purification strategy and performed the bacterial growth assay. I performed sequence alignments and homology modeling. I was responsible for the cloning of constructs, expression and detergent screening. I purified the protein and set up crystallization trials and took part in initial screening of the crystals. I wrote the manuscript and prepared the figures.

Paper II: I was responsible for cloning, expression, purification and crystallization of the protein including initial attempts to solve the structure. I planned and performed the functional bacterial growth assay. I took part in writing the manuscript.

Paper III: I performed cloning, expression- and detergent screening, purified the protein and designed the whole-cell functional assay.

Paper IV: I took part in the cloning, expression- and detergent screening of the pWarf-TRAP constructs.

Paper V: I took part in the literature search, interpretation of the data, and writing of the manuscript along with preparing figures.

Paper VI: I was involved in purification of the protein, assay development and characterization of the compounds. I took part in the preparation of the manuscript.

Abbreviations

Here follows a list and short explanation of the different abbreviations used in this thesis.

ABC	ATP binding cassette
CMC	Critical micelle concentration
DDM	n-Dodecyl β-D-maltoside
DM	n-Decyl β-D-maltoside
FRET	Fluorescence resonance energy transfer
FSEC	Fluorescent size exclusion chromatography
GFP	Green fluorescent protein
GlcNAc	*N*-Acetylglucosamine
GPT	UDP-*N*-acetylglucosamine: dolichyl phosphate *N*-acetylglucosamine-1-phosphate transferase or GlcNAc-1-P transferase
IMAC	Immobilized metal affinity chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
KDN	Deaminated neuraminic acid
LPS	Lipopolysaccharide
MD2	Muraymycin D2
MFS	Major facilitator superfamily
MraY	UDP-*N*-acetylmuramyl-pentapeptide: undecaprenyl phosphate *N*-acetylmuramyl-pentapeptide-1-phosphate transferase or MurNAc-1-P-transferase
MST	Microscale thermophoresis
MurNAc	*N*-Acetylmuramic acid
Neu5Ac	*N*-Acetylneuraminic acid
Neu5Gc	*N*-Glycolylneuraminic acid
PDB	Protein Data Bank
PNPT	Polyprenyl-phosphate *N*-acetylhexosamine 1-phosphate-transferases
RFU	Relative fluorescence unit
SEC	Size exclusion chromatography
SiaT	Sialic acid transporter of the SSS-family
SSS	Sodium solute symporter
TRAP	Tripartite ATP-independent periplasmic
Tun	Tunicamycin

Contents

Chapter 1

Introduction

In our body as well as in any living organism, you can find the most incredible nanomachines, proteins. Proteins are recognized as important biological molecules essential to all forms of life. Characterization and structure determination of proteins allow us to understand their role and function in different processes. A chemical substance with a biological effect, often referred to as a drug, can be used to treat, prevent and cure diseases. Drugs are commonly discovered in one out of two ways; knowing the biological effect before the molecular target is identified or by knowing the molecular target first. Once a drug target has been identified and characterized, small compounds are investigated for a desired effect. Elucidating the structure of a drug target aids the development of new drugs.

Over the years antibiotics have saved millions of lives. However, the efficacy of antibiotics is reduced by the rapid emergence of resistant bacteria. Global health is threatened by a combination of overuse, misuse and lack of new developed antibiotics. Through the work within this thesis two potential drug targets, for the development of new antibiotics, have been structurally and functionally investigated.

1.1 Antibiotics and antibiotic resistance

Bacteria, one of the first life forms to appear on Earth, are diverse and their existence is widespread, from the bottom of the ocean to the inside of our intestines. Through natural selection bacteria adapt to acquire properties to survive in specific environments.

The era of antibiotics begun in the 1940s through the introduction of penicillin, first discovered by Alexander Fleming in 1928. Antibiotics have been recognized as one of the greatest advances in therapeutic medicine, saving countless of lives. With the help of antibiotics, a number of advanced clinical processes such as cancer chemotherapy and organ transplantation

are now routine procedures [1]. As antibiotics started to emerge, some bacteria also started to develop resistance towards these compounds. Antibiotic resistance is one of the most serious health threats of our time. It is a naturally occurring process, accelerated by the overuse of antibiotics in humans and animals.

Antibiotics inhibit the growth of microorganisms and can be synthetically produced as well as produced by both bacteria and fungi. Moreover, antibiotics target essential biological processes within bacteria, including cell wall biosynthesis, DNA replication, transcription and translation [2]. Resistance towards antibiotics occur when bacteria change their response towards the antibiotics and the ability of the antibiotics to effectively inhibit bacterial growth is severely reduced. Mechanisms which compromises the effectiveness of an antibiotic and therefore make the bacteria resistant, includes decreased uptake of the drug, efflux pumping of the drug and expression of enzymes which inactivate antibiotics [3,4]. Less effective antibiotics cause problems in the treatment of infections such as pneumonia, tuberculosis, blood poisoning, gonorrhoea and foodborne diseases. Some bacteria have developed the ability to resist not only one but several antibiotic drugs simultaneously, giving rise to multidrug resistant bacteria [5]. Without the development of new antibiotics, the global threat of antibiotic resistance will not be constrained.

1.2 Membrane proteins

Proteins are versatile and crucial macromolecules in living systems. They are built up by a linear polymer of amino acids, which folds into a specific three-dimensional structure determined by the amino acid sequence. The twenty amino acids, which act as building blocks of the polypeptide, all have different properties. They vary in size, shape, charge, hydrogen bonding capacity, hydrophobic character and chemical reactivity. By elucidating the structure of these important macromolecules, we will have access to a detailed map, which aids in the understanding of the function of the protein. Specific proteins, called membrane proteins, are embedded in a lipid bilayer. Membrane proteins are involved in the transport of molecules in and out of the cell, propagation of electrical impulses, control of membrane lipid composition, receiving chemical signals, regulation of intracellular vesicular transport, energy-converting processes etc [6]. Membrane proteins span the lipid bilayer and is therefore exposed to a highly amphiphilic environment. Lipid head groups occupy a layer in each end of the lipid bilayer, leaving a hydrophobic region in-between.

The architecture of the transmembrane proteins is mostly either α-helical bundles or β-barrels. The helix-bundle proteins are medically important and functionally diverse. A majority of all drug targets are membrane proteins, and these are therefore an important class of proteins for the pharmaceutical industry [7]. Hydrophobic amino acids are the dominant type of amino acids in the lipid-exposed regions of transmembrane helices. Amino acids important for function of the protein are predominantly facing the interior of the protein, embedded between hydrophobic residues. There is a strong tendency for positively charged amino acids like lysine and arginine to face the cytoplasm (inside of the cell) in contrast to the periplasm (outside of the cell); this in known as the "positive-inside rule" and it can be used for predicting the topology of membrane proteins [8, 9]. About 20-30% of all genes in both prokaryotes and eukaryotes encode membrane proteins [10].

1.3 Secondary transporters

Movement of small molecules across a lipid bilayer is facilitated by transporters. Secondary transporters can be classified as either symporters or antiporters. Symporters transport the ion and substrate in the same direction across a lipid bilayer whereas an antiporter transports the substrate and ion in opposite directions. Secondary transporters are driven by transmembrane electrochemical gradients generated by primary active pumps. The separation of charged molecules establish a membrane potential, which is located near the membrane. Secondary transporters couple the uphill transport of a substrate with the downhill transport of another (mostly ions) across the lipid bilayer. The direction of substrate transport is determined by the combined electrochemical potentials of the substrate and the driving solute. Secondary transporters are also expected to be fully reversible [11]. The stoichiometry between substrate and driving ion may vary, but it is important to note that the substrate and driving ion are always co-transported.

1.3.1 The alternating access mechanism

Alternating access to the substrate binding site from either side of the membrane is the basis of the alternating access mechanism. It is a widely used mechanism to describe the alternating conformation of the substrate binding site in the process of delivering the substrate from one side of the membrane to the other [12]. An allosteric coupling of the extracellular and intracellular gates of the transporter creates an alternating access of the two sides. The transport of substrate and ions proceeds in a number of intermediate steps that are each separated by energy barriers (Figure 1.1). Without intermediate steps which occlude the substrate from either side of the membrane, a transporter which moves its substrate against a concentration gradient become less efficient due to continuous association and dissociation of the substrate to the binding site [13]. The alternating access mechanism can be divided into three major mechanisms: the rocker switch, the rocking bundle (gated pore) and the elevator model [11].

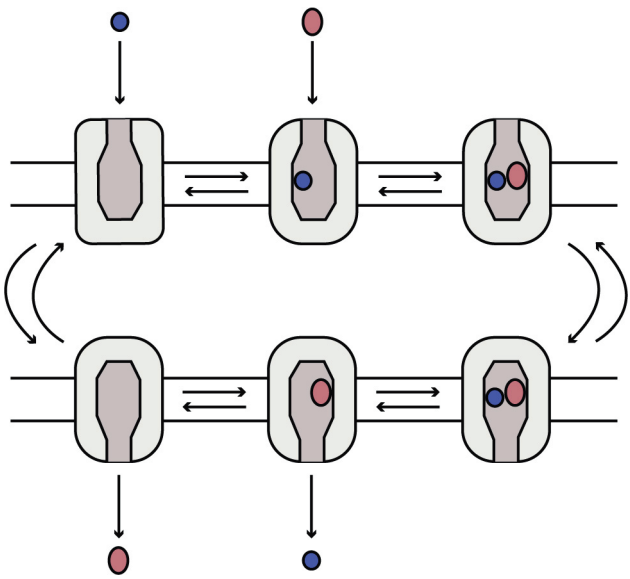

Figure 1.1: The alternating access mechanism of a symporter. A simple mechanism representation of the alternating access of the substrate binding site, through a number of interconversion steps. For a symporter both the substrate and the co-transported ion(s) needs to bind before the protein changes conformation and exposes the substrate binding site to the other side of the membrane. Figure inspired from [13].

The coupled movement between an outside and inside gate around a centrally located substrate binding site is catalyzed by substrate binding in the rocking-bundle alternating access mechanism [13]. In 2005 the crystal structure of the amino acid sodium symporter LeuT from *Aquifex aeolicus* (LeuT_{AA}) of the neurotransmitter sodium symporter (NSS) family was determined at 1.65 Å resolution using X-ray crystallography [14]. This transporter consists of twelve transmembrane helices, where the first ten build up the structural core of the protein in a fold subsequently named the LeuT-fold (Figure 1.2). The transmembrane helices TM1-TM5 and TM6-TM10 are related to each other by a pseudo-two-fold axis located in the membrane plane. TM1 and TM6 are oriented antiparallel to each other and are non-continuous helices, the breaks in these helices are located approximately halfway across the membrane bilayer. This conserved transporter architecture is characterized by three structural elements. The helix bundle (core domain) which consists of the first two helices of each repeat (TM1-TM2 and TM6-TM7), the hash domain (scaffold domain) consisting of the third and fourth helix of each repeat (TM3-TM4 and TM8-TM9) and the two peripheral arms which correspond to the fifth helix of each repeat (TM5 and TM10) [11, 15]. The two discontinues helices (TM1 and TM6) from the core domain make up a large fraction of the substrate and sodium binding sites.

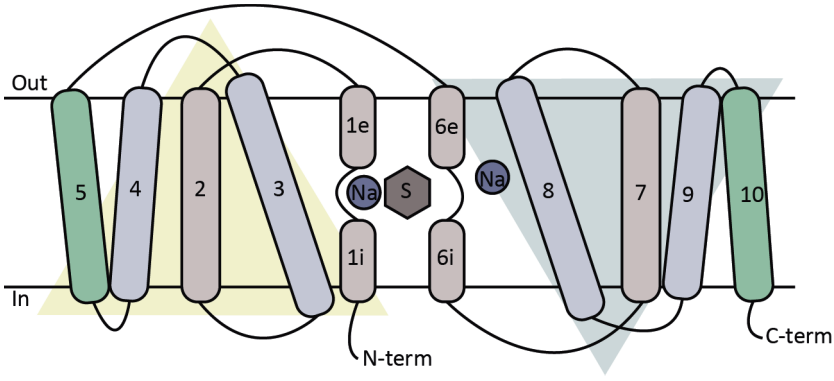

Figure 1.2: The LeuT-topology with two inverted repeats of five transmembrane helices. The three structural elements are indicated with different colors: the helix bundle (core domain) which consists of TM1-TM2 and TM6-TM7 (beige), the hash domain (scaffold domain) which consists of TM3-TM4 and TM8-TM9 (purple) and the two peripheral arms which correspond to TM5 and TM10 (green). The position of the substrate- and sodium binding sites are indicated with diamond (brown) and spheres (blue), respectively. Figure inspired from [14].

The rearrangement upon gating varies between rocking-bundle transporters with the same fold [11]. Detailed descriptions of the alternating access mechanism have been obtained from crystal structures of BetP, LeuT and Mhp1 [16–18] in different conformations.

1.3.2 Sodium binding sites

The driving force for sodium-coupled symport is the sodium ion electrochemical gradient, nevertheless there is a variation in the number of sodium binding sites and the stoichiometry between sodium: substrate which is transported across the lipid bilayer among proteins with the LeuT-fold.

In the first structure of $LeuT_{AA}$ [14], two sodium binding sites were identified. The two sodium binding sites Na1 and Na2, were located halfway across the membrane bilayer, at the unwound regions of TM1 and TM6 within 6 Å away from the substrate binding site. The Na1 sodium ion is octahedrally coordinated by amino acid residues from TM1, TM6 and TM7 as well as via the carboxyl oxygen of the substrate leucine. Na2 is coordinated by a trigonal bi-pyramidal configuration through amino acid residues from TM1 and TM8. The Na2 position between TM1 and TM8 is ∼7 Å away from the Na1 site and 5.9 Å away from the substrate binding site. Additional members of the NSS family which have also been structurally determined and which possess two sodium binding sites (Na1 and Na2) are the mulithydrophobic L-amino acid transporter MhsT from *Bacillus halodurans* [19] and the dopamine transporter from *Drosophila melanogaster* dDAT [20].

The bacterial sodium galactose transporter from *Vibrio parahaemolyticus* (vSGLT) is the first structurally determined representative of the sodium solute symporter (SSS) family and this protein possesses the conserved Na2 site [21]. vSGLT has been structurally determined in inward-facing conformations in an apo [22] and substrate-bound [21] form. The benzyl-hydantoin transporter Mhp1 belongs to the nucleobase-cation-symporter (NCS1) family, and possesses one sodium binding site (Na2) [23].

In the betaine transporter BetP, which belongs to the betaine/choline/carnitine (BCC) family, the stoichiometry between transported substrate and sodium has been determined to be 1:2 [15, 24]. The position of the second sodium binding site in BetP (Na1´) was proposed based on structural homology between repeats [25].

The Na2 site is a conserved (non-optional) sodium binding site among transporters with the LeuT-fold [15]. However, the presence of the additional non-conserved Na1 and Na1´ sites might be related to the specificity of the co-transported substrate.

1.4 Sialic acid transport

1.4.1 Sialic acid

Sialic acid, also called neuraminic acid, is a large family of nine-carbon sugar acids with among 50 structurally distinct compounds identified [26]. The most abundant and well-studied sialic acid is the 2-keto-3-deoxy-5-acetamido-D-glycero-D-galacto-nonulosonic acid (*N*-acetylneuraminic acid or Neu5Ac) (Figure 1.3) [26], an amino sugar derivative of *N*-acetylmannosamine and pyruvic acid. The neuraminic acid ring of Neu5Ac has an acetamido group at the C5 position and a glycerol tail composed of carbons 7-9, at the C6 position. The majority of sialic acid derivatives have substitutions at the C5 position and at the hydroxyl groups of the C4, C7, C8 and C9 positions [27].

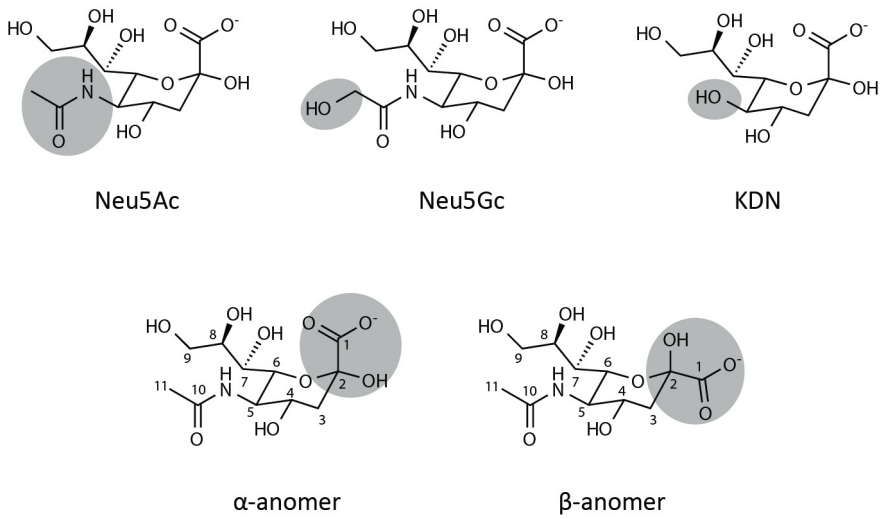

Figure 1.3: Sialic acid The structure of the three sialic acid derivatives Neu5Ac, Neu5Gc and KDN. In addition, the α- and β-anomer of Neu5Ac are displayed, carbon atoms are numbered from 1 to 11 and differences are higlighted with a grey circle.

The second most common sialic acid of most vertebrates, *N*-glycolylneuraminic acid (Neu5Gc), differs from Neu5Ac by the addition of a hydroxyl group at the C11 position (Figure 1.3). Interestingly Neu5Gc has never been detected in bacteria [26]. In humans, Neu5Gc is endogenously absent, since the gene encoding the hydroxylase which adds the *N*-glycolyl group in Neu5Gc is inactive [26, 28]. Deaminated neuraminic acid (KDN) (Figure 1.3) occurs widely among vertebrates and bacteria. Its abundant occurrence in animals is limited to lower vertebrates [29].

The carboxylate group (pK_a of 2.6) of sialic acid gives it its net negative charge at physiological pH. Sialic acid can exist in both the α- and β-conformation. In general, sialic acid is released as the α-conformation from the glycoconjugates and subsequently converted into its β-conformation before transported into the cytoplasm [27,30]. By mutarotation both forms can be interconverted as the ring opens to a straight-chain conformation followed by ring closure [26].

1.4.2 Distribution and significance of sialic acid

Sialic acid is positioned at the terminal end of glycoconjugates in both eukaryotes and prokaryotes. This position allows them to participate in interactions with the external environment, such as cell-cell communication and self-recognition. In humans the highest concentration of sialic acid is found in the brain and utilized in the central nervous system [31]. Mucus-rich environments like the respiratory and gastrointestinal tracts in humans are also rich in sialic acid with over 65% of glycans containing sialic acid residues [32, 33].

Many pathogenic bacteria have evolved the ability to use sialic acid as a shield of protection, by decorating their cell-surface. It gives them a protection against the host´s immune response [34]. The decoration of bacterial surfaces have been observed for lipopolysaccharides (LPS) in Gram negative bacteria; the flagellum and the capsular polysaccharides [31].

Sialic acid can also be used as a source of nutrition for the bacteria. Catabolism of Neu5Ac is associated with commensal and pathogenic bacteria [35, 36]. In bacteria, Neu5Ac is catabolized by five enzymes: *N*-acetylneuraminic acid lyase (NanA), *N*-acetylmannosamine kinase (NanK), *N*-acetylmannosamine 6-phosphate epimerase (NanE), *N*-acetylglucosamine-6-phosphate deacetylase (NagA), glucosamine-6-phosphate deaminase (NagB) yielding fructose-6-phosphate and ammonia [31]. The enzymes involved in degradation of Neu5Ac in *Escherichia coli* are all located within the nan operon (nanATEK-yhcH) regulated by the nanR repressor, located upstream of nanA [31]. Furthermore, the nanCMS operon is also regulated

by the NanR repressor protein. The genes encoded by the nanCMS operon includes the outer membrane porin (NanC), sialate mutarotase (NanM) and sialate *O*-acetyl esterase (NanS) [31].

Bacteria are able to scavenge free sialic acid from the environment, in addition bacteria synthesize sialidases, enzymes which cleaves terminal sialic acid from the eukaryotic host´s glycoconjugates [31,37]. The acquired sialic acid is further transported into the cytoplasm of the bacterial cell by specific sialic acid transporters [30]. In eukaryotes, the synthesis of Neu5Ac takes place in the cytoplasm. Subsequently, Neu5Ac is activated in the nucleus to CMP-Neu5Ac and transported into the Golgi apparatus by a specific CMP-Neu5Ac transporter [38] where it is recognized by sialyltransferases and attached to glycoproteins or glycolipids.

1.4.3 Sialic acid import across the bacterial cytoplasmic membrane

Bacteria transport sialic acid across their cytoplasmic membrane using specific sialic acid transporters. There are four different classes of transporters known to transport sialic acid (Figure 1.4). The type of transporter can vary between different bacterial species. The four different classes of sialic acid transporters include the ATP-binding cassette (ABC), tripartite ATP-independent periplasmic (TRAP), major facilitator superfamily (MFS) and sodium solute symporter (SSS) family [39]. It is not completely understood why some bacterial pathogens have more than one type of sialic acid transporter; possibly it can be related to the large variation of sialic acids [39]. Bacterial sialic acid transporters have low homology with the human sialic acid transporter and is therefore a valid antibacterial target.

ABC transporters

One of the largest families of transporters is the ABC superfamily [40]. The ABC-transporter family utilizes the energy generated from ATP-hydrolysis in the transport of a variety of substrates across the extracellular and intracellular membranes [41, 42]. In 2005, the first sialic acid transporter of the ABC-type was identified and characterized in *Haemophilus ducreyi* [43]. The sialic acid ABC transporter consists of four subunits: the periplasmic substrate binding subunit SatA, the integral membrane subunits SatB and SatC and the nucleotide binding domain SatD located in the cytoplasm. Furthermore, in *H. ducreyi* the SatC subunit contains both an N-terminal transmembrane domain and a C-terminal nucleotide binding domain [43].

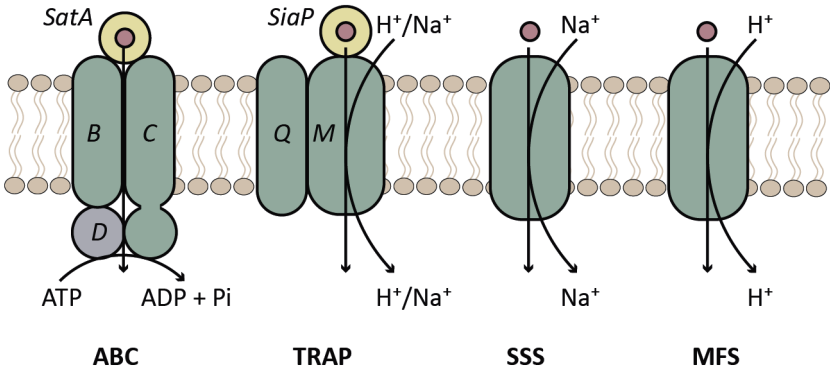

Figure 1.4: Overview of the different types of sialic acid transporters located in the cytoplasmic membrane of bacteria. ABC-transporter; generally composed of two transmembrane domains and two nucleotide binding domains which hydrolyze ATP. The Sat subunits (SatA, SatB, SatC and SatD) are labeled. TRAP-transporters are built up by three domains, the substrate binding domain (SiaP) and a large and small membrane spanning domain (SiaQ and SiaM). SSS-transporters (SiaT) and MFS-transporters (NanT) both consist of a single transmembrane spanning domain. Sialic acid is illustrated as red circles.

TRAP transporters

TRAP transporters are present in both bacteria and archaea, but not in eukaryotes [44]. The first TRAP transporter was identified in 1997 [45] and it uses an electrochemical gradient to facilitate the transport of solute over the membrane. Whether the TRAP transporter is dependent on protons or sodium ions is unclear. The first sialic acid TRAP transporter was characterized from *Haemophilus influenzae* [46, 47]. The TRAP-transporter is the predominant type in *Pasteurellaceae* and *Vibrionaceae* [48]. The sialic acid TRAP transporter is built up by three subunits. The periplasmic substrate binding subunit SiaP, which delivers sialic acid to the membrane protein subunits, SiaQ and SiaM. For some bacterial organisms these two subunits are fused together forming a SiaQM integral membrane subunit.

MFS transporters

The first sialic acid transporter to be discovered was NanT from *E. coli* [49]. This secondary transporter belongs to the major facilitator family (MFS) which uses the proton electrochemical gradient in the co-transport of sialic acid [26, 50, 51]. Interestingly, the NanT transporter is predicted to comprise of 14 transmembrane helices, this is two more than the usual 12 transmembrane helices of MFS transporters [50, 52]. The NanT transporter is mainly found in *Enterobacteriaceae* [48].

SSS transporters

In 2010 the sialic acid transporter SiaT from *Salmonella enterica* servoar Typhimurium was identified as a sialic acid transporter of the SSS family [51]. The gene encoding SiaT is widespread in both Gram positive and Gram negative bacteria [51] including the human bacterial pathogens *Staphylocccus aureus* [53], *Phobacterium profundum* [54], *Clostridium perfringens* [55] and *Clostridium difficile* [56]. Furthermore, the SiaT transporter is predominant among Firmicutes [48]. The driving force for the SSS-transporters, which co-transports a number of different substrates with sodium ions [57], is a combination of the negative membrane potential and the intracellular low concentration of sodium ions [21, 58].

1.5 Bacterial cell wall biosynthesis

1.5.1 The bacterial cell membranes

Targeting proteins involved in the bacterial cell wall biosynthesis is a valid approach to develop novel antibiotics, since mammalian cells lack a cell wall.

A prokaryotic cell protects its content with a cytoplasmic membrane. The cytoplasmic membrane consists of a phospholipid bilayer with embedded proteins which are involved in numerous of different processes within the cell, including the control of passage of molecules in and out of the cell. Most prokaryotic cells also surround themselves with a cell wall, a peptidoglycan layer, with the function to protect the cell from mechanical injury and prevent it from bursting. The two characteristically different prokaryotic cell types are the Gram positive and Gram negative (Figure 1.5).

A Gram negative bacteria is surrounded by two membrane layers, an outer membrane and an inner membrane, as opposed to the Gram positive bacteria which only possess the inner cytoplasmic membrane. The outer membrane of Gram negative bacteria is rich in lipopolysaccharides (LPS). Furthermore, the cell wall is built up by a peptidoglycan layer, which is thicker in Gram positive bacteria than in Gram negative bacteria [59]. Due to its extra outer membrane layer, the Gram negative bacteria is much more effective in blocking antibiotics from entering the cell [60, 61]. Proteins involved in bacterial cell wall biosynthesis are highly conserved and often essential [62], hence they are suitable potential antibiotic targets.

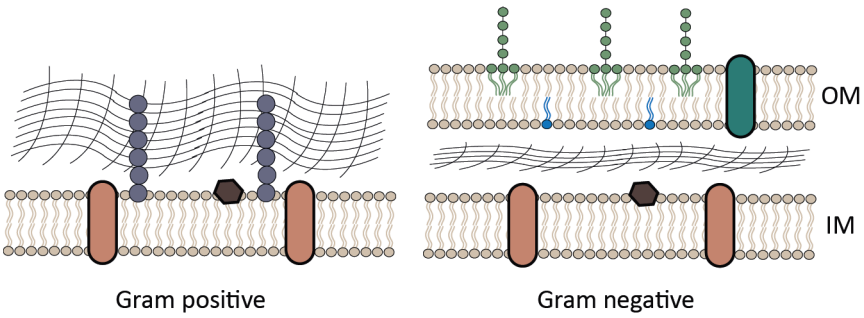

Figure 1.5: Illustration of the Gram positive and the Gram negative cell membranes. Indicated are the inner cytoplasmic membrane (IM) and the outer membrane (OM). The peptidoglycan layer is illustrated as a cross-linked network (black) and is significantly thicker in Gram positive bacteria compared to Gram negative bacteria. Shown are membrane proteins (orange), porin (green), peripheral membrane proteins (brown), lipopolysaccharide (green), lipoprotein (blue) and teichoic acids (purple). Figure inspired from [61].

1.5.2 Peptidoglycan biosynthesis

The cytoplasmic membrane of bacteria is surrounded by a peptidoglycan layer, also called murein. It is essentially a specific macromolecule which acts as a protective barrier for the cell [63]. The peptidoglycan layer gives the bacterial cell its specific shape and allows it to withstand osmotic pressure; it can also be used as a scaffold for anchoring other proteins [63, 64].

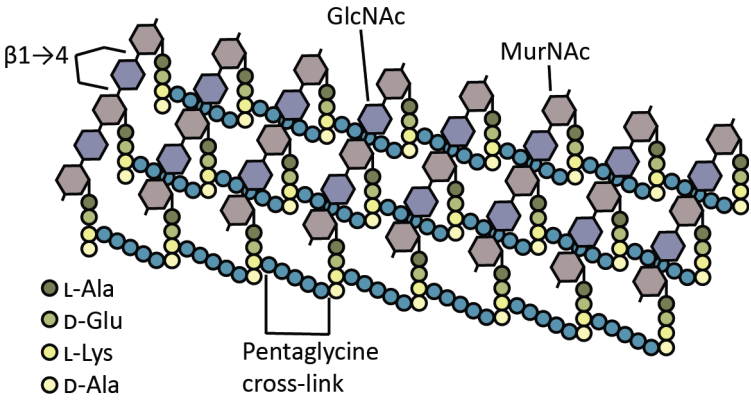

Figure 1.6: Illustration of a cross-linked peptidoglycan layer.

Peptidoglycan are linear glycan strands cross-linked by short peptides (Figure 1.6) [63, 65]. Alternating *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc) residues are linked together by β-1,4 glycosidic bonds in each glycan strand. A peptide chain is substituting the D-lactoyl group of each MurNAc residue in peptidoglycan. In the mature macromolecule the last D-Ala residue of each peptide chain is lost. Cross-linking between glycan strands often occurs between the carboxyl group of D-Ala at position 4 and the amino group of the diamino acid at position 3 by either direct cross-linking or by a short peptide bridge [63]. Variations of the glycan strands, the peptide strand and the position or composition of the interpeptide bridge occur in different bacterial species [63].

The peptidoglycan layer is assembled in three phases. In the first phase which takes place in the cytoplasm of the bacterial cell, several soluble enzymes (MurA, MurB, MurC, MurD, MurE and MurF) are invovled in the synthesis of UDP-*N*-acetylmuramyl-pentapeptide (UDP-MurNAc-pentapeptide). In the second phase, the two membrane bound enzymes MraY and MurG are involved in reactions occurring on the cytoplasmic surface of the inner membrane of the bacterial cell. MraY catalyzes the transfer of the phospho-MurNAc-pentapeptide motif of UDP-MurNAc-pentapeptide

to the lipid carrier undecaprenyl phosphate, generating lipid I in the first membrane-bound step. Secondly, MurG catalyzes the transfer of the GlcNAc motif of UDP-GlcNAc to the C4 hydroxyl MurNAc in lipid I to produce the lipid-linked β–1,4 disaccharide known as lipid II. Lastly in the third phase, lipid II is translocated to the exterior surface of the cell by an unknown mechanism [66–69]. The membrane protein MraY will be further described in chapter 1.5.5.

1.5.3 The PNPT superfamily

The polyprenyl-phosphate *N*-acetylhexosamine 1-phosphate-transferase (PNPT) superfamily includes prokaryotic enzymes responsible for bacterial cell wall biosynthesis and eukaryotic enzymes involved in *N*-linked glycosylation. This family of enzymes catalyzes the transfer of a phospho-*N*-acetylhexosamine from a cytoplasmic UDP-*N*-acetylhexosamine sugar nucleotide donor substrate to a membrane-bound polyprenyl phosphate acceptor substrate [70]. MraY, WecA, WbcO, WbpL and RgpG are all enzymes of the PNPT superfamily; they differ in their specificity for the UDP-sugar donor substrate but use the same lipid substrate, undecaprenyl-phosphate (C_{55}-P) [71]. The eukaryotic member of the PNPT family, GPT (UDP-*N*-acetyl-glucosamine: dolichyl phosphate *N*-acetylglucosamine-1-phosphate transferase or GlcNAc-1-P-transferase), is essential for *N*-linked glycosylation and located within the endoplasmic reticulum (ER) membrane. GPT uses UDP-GlcNAc and dolichol phosphate as the soluble and lipid substrate, respectively [72].

1.5.4 Natural product inhibitors

Identifying small molecules with antibacterial activity, especially for multidrug-resistant Gram negative pathogens, has proven to be challenging. The ability of an antibiotic compound to penetrate the bacterial cell wall is limited by the protective barriers, including cell membranes, complex carbohydrate networks and efflux pumps, in both Gram negative and Gram positive bacteria [73]. Pharmaceutical companies often use high-throughput screening (HTS) as a conventional drug discovery approach. However, this approach has not been as successful for the identification of novel antibacterial drugs [74]. An alternative approach is to investigate natural product inhibitors as scaffolds for the development of novel antibiotic drugs.

MraY has lately been suggested as a promising antibiotic target. An antibiotic aimed to inhibit MraY has to be able to cross the protective barriers and to reach the active site of MraY, which is located on the cytoplasmic side

of the membrane. It is also necessary to avoid toxicity issues by selectively target MraY over any other translocases. Furthermore, the antibiotic needs to possess good inhibitory activity against MraY.

Peptide inhibitors, nucleoside- and lipopeptide antibiotics are all natural product inhibitors that block the synthesis of lipid I, the product of MraY [75]. The lipopeptide amphomycin is a non-competitive inhibitor of MraY which forms a complex with the lipid substrate C_{55}-P in the presence of Ca^{2+} [76]. Protein E is a peptide produced by the bacteriophage ΦX-174 which inhibits MraY by binding at an allosteric site [75]. The group of nucleoside antibiotics can be divided into different sub-groups based on their chemical differences and/or mode of action against MraY, for example tunicamycins, ribosamino-uridines, uridylpeptides and capuramycins [75].

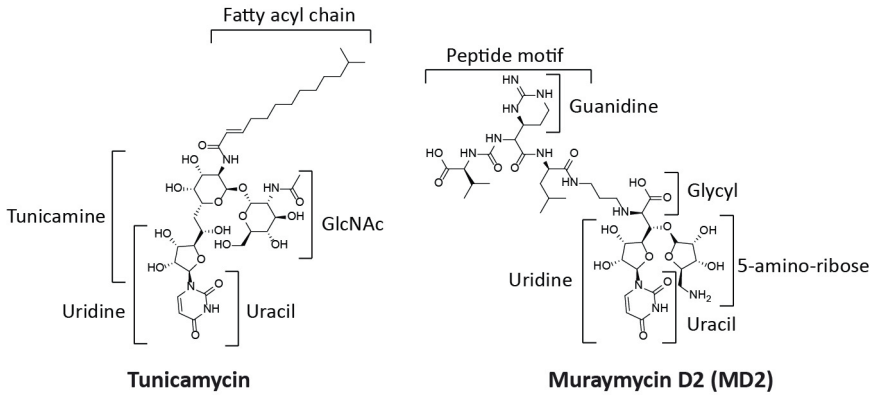

Figure 1.7: The natural product inhibitors tunicamycin and muraymycin D2 (MD2). The fatty acyl chain of tunicamycin can vary in length and branching.

Tunicamycin is a uridine-derived nucleoside antibiotic that consists of a uracil motif, a tunicamine sugar, an *N*-acetylglucosamine sugar (GlcNAc) and a fatty acyl chain (Figure 1.7). It is considered as a competitive inhibitor relative to MurNAc, the sugar nucleotide donor, and not competitive towards the undecaprenyl-phosphate substrate [77]. Tunicamycin was first discovered in 1971, as it was isolated from the fermentation broth of *Streptomyces lysosuperficus* due to its antiviral properties [78]. The synthesis of lipid I is blocked by tunicamycin in Gram positive bacteria [78,79] in contrast to Gram negative bacteria [80] where tunicamycin is unable to penetrate the outer cell membrane. It is also toxic towards eukaryotes by inhibiting GPT [75,78]. Tunicamycin is not a suitable antibiotic drug, although reduced eukaryotic toxicity has been observed for modified tunicamycins [81, 82].

Muraymycins are uridine-derived nucleoside-peptide antibiotics, first isolated from a broth of a *Streptomyces* sp. in 2002 [83]. The structure of muraymycin consist of a glycyl-uridine motif connected via an aminopropyl linker to a urea peptide moiety, and at the 5´-position the uridine is glycosylated with an aminoribose unit (Figure 1.7). Some derivatives of muraymycin have a lipophilic side chain attached at the hydroxyleucine residue of the peptide moiety [84]. In contrast to tunicamycin, most muraymycin analogues are not cytotoxic. Muraymycin D2 (MD2) has been reported to inhibit MraY, although it does not possess favorable cell permeability characteristics, hence does not show any antibacterial activity [85].

1.5.5 MraY

MraY (UDP-*N*-acetylmuramyl-pentapeptide: undecaprenyl phosphate *N*-acetylmuramyl-pentapeptide-1-phosphate transferase or MurNAc-1-P-transferase) is a membrane bound enzyme involved in the peptidoglycan biosynthesis of bacteria. It catalyzes the synthesis of C_{55}-PP-MurNAc-pentapeptide (lipid I) from the lipid substrate undecaprenyl-phosphate (C_{55}-P) and the soluble substrate UDP-MurNAc-pentapeptide (Figure 1.8). The first discovery of MraY was made in 1965 in the bacteria *S. aureus* and *Micrococcus luteus* [86, 87], and the complete amino acid sequence along with the topology have been reported for MraY from both Gram positive and Gram negative bacteria [88, 89].

Over the years, MraY from Gram positive and Gram negative bacteria have been expressed (recombinantly in *E. coli* and cell free expression) and purified [90–94]. Membrane proteins with both the N- and C-termini located on the periplasmic side are known to be difficult to overexpress, most likely related to the problems with membrane insertion and folding [95]. Cell-free expression of Gram negative MraY demonstrated the importance of incorporating negatively charged lipids for obtaining a stable and functional enzyme [62, 91].

Initial biochemical studies to measure MraY enzyme activity was performed with radiochemical assays. Radiolabeled soluble substrate was incubated with lipid substrate and enzyme, followed by extraction of the radiolabeled lipid I product [96]. The modified substrate UDP-MurNAc-L-Ala-γ-D-Glu-m-DAP(ξ-dansyl)-D-Ala-D-Ala, which contains the dansyl fluorophore, can be used as a substrate for MraY and the dansylated lipid I product exhibit enhanced fluorescence [97]. Subsequently a continuous fluorescence assay was developed [77, 80] and extended to an HTS assay [98].

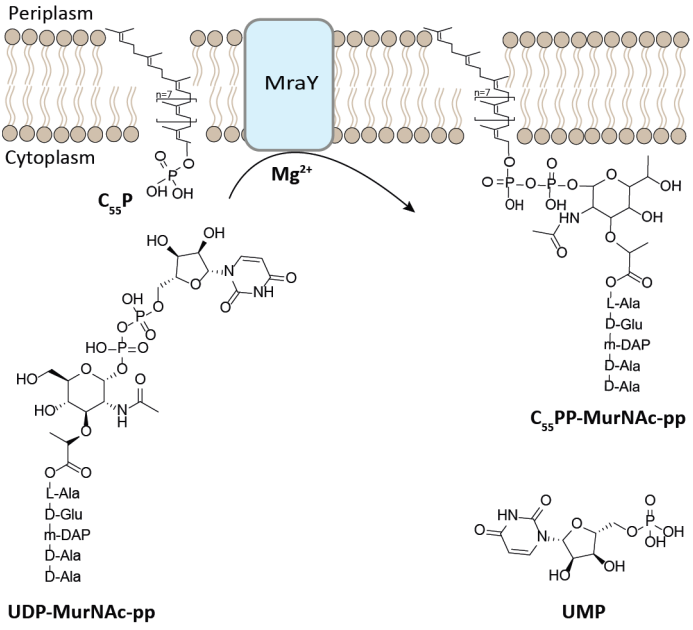

Figure 1.8: The reaction of MraY. The synthesis of lipid I occurs on the cytoplasmic side of the inner membrane in bacteria. The lipid substrate $C_{55}P$ reacts with the soluble substrate UDP-MurNAc-pentapeptide (UDP-MurNAc-pp) in a Mg^{2+} dependent reaction catalyzed by MraY, yielding $C_{55}PP$-MurNAc-pentapeptide ($C_{55}PP$-MurNAc-pp, lipid I) and UMP.

The thin layer chromatography (TLC) translocase assay has been used extensively [90, 91, 93]. Another valuable assay is the fluorescence resonance energy transfer (FRET) assay developed in 2012 [99].

The MraY enzyme activity is dependent on the divalent cation magnesium (Mg^{2+}) [90]. Amino acids involved in the coordination of the Mg^{2+} ion were predicted based on mutational studies [100], sequence similarity of the Mg^{2+}-binding motif (DDXXD/N) of prenyl transferases [101] along with the apo structure of *A. aeolicus* MraY (AaMraY) [93] (Table 1.1).

Enzyme inhibition has previously been demonstrated to be competitive with Mg^{2+} for ester-linked 5´-uridinyl dipeptides, where the inhibition of solubilized *E. coli* MraY (EcMraY) was decreased with increased concentrations of Mg^{2+} [102]. It was proposed that the primary amino terminus of the most potent inhibitor interacts with amino acids important for Mg^{2+} binding in MraY.

The binding affinity between MraY and the natural product inhibitor

Table 1.1: Amino acids proposed to be involved in Mg^{2+} coordination. Amino acids in **bold** indicate which MraY enzyme that was used as model in the assignment from the mutational studies (A), Mg^{2+}-binding motif (DDXXD/N) of prenyl transferases (B) and the structure of AaMraY (C). Amino acids are given as a reference for *Clostridium bolteae* MraY (CbMraY).

	A	**B**		**A**		**C**
CbMraY [94]	His44	Asp93	Asp94	Asp175	Asp178	Asp231
AaMraY [93]	His65	Asp117	Asp118	Asp193	Asp196	**Asp265**
BsMraY [100]	**His45**	Asp98	Asp99	**Asp174**	**Asp177**	Asp231
EcMraY [101]	His65	**Asp115**	**Asp116**	Asp195	Asp198	Asp267

tunicamycin has been determined for *Bacillus subtilis* MraY (BsMraY) (K_d = 93 ± 16 µM) using biolayer interferometry (BLI) in the presence of 10 mM $MgCl_2$ [103] and for AaMraY (K_d = 37 ± 1 nM) using isothermal titration calorimetry (ITC) in the absence of $MgCl_2$ [82]. Interestingly, the binding affinity constant is significantly different between the two methods and could be an effect of interfering Mg^{2+} [82]. For the purpose of determining the binding affinity between potential inhibitors/modified compounds and MraY, there is a need for a reliable assay which consumes small amounts of protein since the yield of purified MraY is often not very high.

Mechanism of MraY

Until 2016, there was a debate regarding the mechanism of MraY. Bouhss and co-workers concluded that MraY acts via a single displacement mechanism (one-step mechanism) where the UDP-MurNAc pentapeptide donor substrate and the lipid acceptor substrate simultaneously bind to the active site of MraY. Furthermore, they propose that a conserved aspartic acid (Asp98 in BsMraY) is involved in the deprotonation of the lipid substrate allowing for a nucleophilic attack of the pyrophosphate group of UDP-MurNAc pentapeptide [104]. Docking studies of the substrate analog UDP-GlcNAc into the active site of CbMraY [94] further strengthens this theory.

1.6 Scope of this thesis

This thesis presents the work aimed to provide structural and functional insights into sialic acid transporters which transport sialic acid across the cytoplasmic membrane of bacteria. In addition, this thesis presents structural interpretations and binding studies of the enzyme MraY involved in the synthesis of the bacterial peptidoglycan layer. Both proteins are crucial to pathogenic bacteria and as such promising candidates for the future development of antibacterial drugs.

Chapter 2 gives a brief introduction to some of the methods involved in protein production, purification and structural determination using X-ray crystallography and a few biophysical methods used for characterizing these proteins and the interaction between protein and small-molecule compounds.

Chapter 3 presents results from **Paper I** – **Paper IV**, which covers the sialic acid transporters.

Chapter 4 presents the results obtained in **Paper V** and **Paper VI**, which covers the enzyme MraY.

Chapter 5 summarizes the results and presents a future perspective.

Chapter 2

Methodology

In theory, the process from target identification to pure protein is straightforward (Figure 2.1). In reality there are a number of steps, which needs to be carefully studied and optimized, especially for membrane proteins.

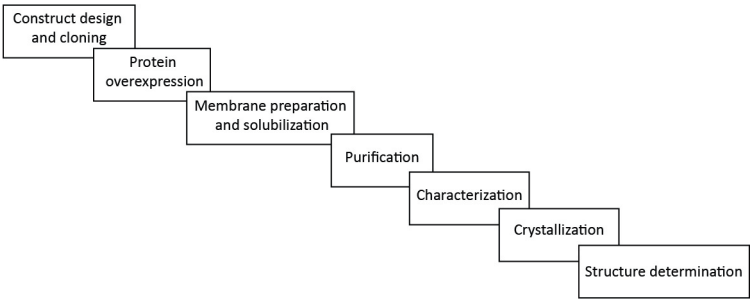

Figure 2.1: The process from identified gene to pure protein, characterization and structure determination includes many trial and error experiments.

2.1 The bacterial expression system

Protein structure determination requires high amounts of pure and stable protein. It is advantageous to use *E. coli* as the host for recombinant expression of prokaryotic proteins since it is closely related to the natural source. Furthermore, recombinant expression in *E. coli* allows for transformation of exogenous DNA, it has a short generation time and is relatively inexpensive. There are also several different modified *E. coli* strains available, suitable for different expression systems [105, 106].

A common approach for producing recombinantly expressed proteins in *E. coli* is to use the T7 system. The T7 expression host DE3 contains a

chromosomal copy of the phage T7 RNA polymerase gene. Protein expression in the *E. coli* strain BL21(DE3), developed for the overexpression of soluble proteins [107], is driven by the bacteriophage T7 RNA polymerase. T7 RNA polymerase is responsible for a much faster transcription rate compared to the *E. coli* RNA polymerase [108]. More transcribed mRNA will generate more translated protein, although if the rate is too fast, the protein is not always correctly folded [105].

Furthermore, the T7 RNA polymerase is under the control of the lacUV5 promoter which can be regulated by the lac repressor and induced with IPTG. In the T7 system, an expression vector containing the gene of interest is cloned downstream of the T7 promoter and subsequently transformed into the *E. coli* expression host. Upon induction, T7 RNA polymerase is expressed and available for transcription of the target gene. Overexpression of membrane protein is often toxic for the host, and a promoter that is not tightly regulated (*e.g* the lac promoter) often results in leaky expression. The *E. coli* Lemo21(DE3) strain has the ability to control the expression rate by using the natural inhibitor T7 lysozyme [105]. In *E. coli* Lemo21(DE3), the gene encoding T7 lysozyme is on a plasmid which is under the control of a L-rhamnose promoter (rhaP(BAD)), which is well titrable [109, 110].

2.2 Construct design

The amino acid sequence of a protein is encoded within a specific DNA sequence, a gene. As a first step, it is common to modify the regions flanking the gene of interest, *e.g.* include affinity tags, fusion proteins and protease cleavage sites which will facilitate protein expression and purification among other things.

One of the most widely used vector systems for recombinant protein expression in *E. coli* is the pET expression system [111]. The bacterial expression vector pET26b(+) has an N-terminal pelB signal sequence which directs the membrane protein to the membrane, subsequently the leader signal is removed by a signal peptidase [112].

The commercially available low-copy pWarf-vectors (Figure 2.2) [113,114] carries a C-terminal human rhinovirus 3C protease (HRV3C) cleavage site, a green fluorescent protein (GFP)-tag and an 8x histidine-tag. The pWarf(+) vector also carries a Glycophorin A (GpA) fusion (a single transmembrane spanning helix) which redirect the GFP into the cytoplasm if the C-terminal of the target protein is located on the periplasmic side. The soluble GFP protein is relatively small in size (molecular weight of 27 kDa) and functional as a monomer. Furthermore, GFP emits green light once the internal

fluorophore Ser-Tyr-Gly is excited with long wave UV-light [115]. To be able to take advantage of the GFP-tag for detection, it is important to know the location of the C-terminal of the target protein, since GFP is only functional in the cytoplasm.

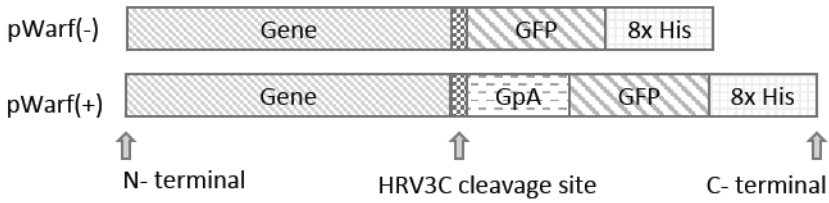

Figure 2.2: Schematics of protein constructs in the pWarf-vectors. The difference between pWarf(-) and pWarf (+) is the additional Glycophorin A (GpA) single transmembrane helix motif.

The levels of overexpressed membrane protein in whole cells can be monitored by the levels of expressed GFP, which is fused to the C-terminal of the membrane protein [113, 116, 117]. If the protein is incorrectly folded, it will not fluoresce. An additional quality control is to detect the fusion protein on an SDS-polyacrylamide gel by fluorescence (In-Gel Fluorescence) [117]. It is also possible to take advantage of the GFP-fusion protein during detergent screening, in the fluorescent size exclusion chromatography (FSEC) method.

Incorporating an affinity-tag, *e.g.* a polyhistidine-tag, can be used in immobilized metal-affinity chromatography (IMAC). This method is based on the interaction between a transition metal ion immobilized on a matrix and specific amino acid chains in the protein [118]. It is also possible to label a polyhistidine-tag with a fluorescent probe and use it in microscale thermophoresis (MST) experiments or with FSEC [92] (discussed in chapter 4.2.3).

2.3 Purification of membrane proteins

A functional stable and homogeneous membrane protein sample is a good starting point for crystallization and characterization experiments. Purification of overexpressed membrane proteins starts with breaking the cells and collecting the membrane fraction from which the proteins are extracted. Subsequently the protein is purified often using an affinity step followed by size-exclusion chromatography (SEC), as a polishing step.

2.3.1 Solubilization of membrane proteins

Membrane proteins are embedded in a lipid bilayer and needs to be solubilized into an aqueous environment where the protein is stable. Due to the intrinsic hydrophobicity of membrane proteins, they are prone to fold improperly and aggregate in solution. This is one of the major bottlenecks in membrane protein research. Solubilization of membrane proteins from their natural lipid bilayer can be done using detergent molecules. Detergent molecules are amphipathic molecules, by definition the detergents have a polar and non-polar region. Above a certain critical micelle concentration (CMC), they spontaneously form micelles (Figure 2.3). Detergents vary in charge and size, both within the head group and tail. The average number of detergent molecules per micelle is denoted as the aggregation number (N_A). Detergents with a longer tail are typically more hydrophobic, hence are less soluble, which as a consequence results in a lower CMC. Harsh detergents are often small with a polar head group [119]. Most successful detergents for α-helical membrane protein crystallization have been the alkyl-maltopyranoside and alkyl-glucopyranoside detergents [120]. A detergent with a short hydrophobic tail has a smaller micelle, which can lead to denaturation of the protein since the hydrophobic region of the membrane protein might not be completely covered. On the contrary, it is beneficial to use a detergent with a shorter carbon tail during crystallization since it will give room for forming crystal contacts between the hydrophilic regions of the protein [119].

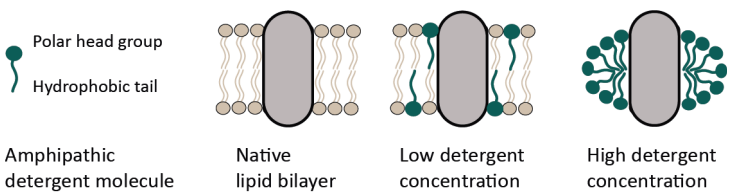

Figure 2.3: Detergent solubilization of membrane proteins. Detergent (green), membrane protein (brown) and lipid bilayer (beige).

2.3.2 Chromatography

Purification of solubilized membrane protein is most often performed in two steps, the first being an affinity purification step and the second a size exclusion chromatography step. During affinity purification the tag which is fused to the target protein is interacting with a specific resin. Polyhistidine-tags are commonly used and they have affinity for divalent metal ions like Ni^{2+} and Co^{2+}. The immobilized protein is eluted from the matrix by either changing the pH or adding free imidazole to the elution buffer to compete with the histidine-tag. By increasing the number of histidines in the polypeptide-tag, inefficient binding to the metal ion matrix might in some cases be overcome.

After an affinity purification step the protein is often reasonably pure, although it is often necessary to include a size-exclusion step as a last polishing step. During size-exclusion chromatography it is possible to check the homogeneity of the protein as well as changing the buffer composition and/or detergent if necessary.

2.4 Crystallization and structure determination

Protein molecules are small in comparison to what can be detected with the naked eye. It is therefore necessary to use a technique which is able to resolve the atomic structure, and thereby detect interactions between atoms. Chemical bond distances are in the order of 1-3 Ångström (Å) and with X-ray diffraction it is possible to detect atoms about 1 Å apart (1 Å = 10^{-10} m).

X-rays are electromagnetic radiation with a short wavelength (10^{-8}-10^{-12} m). When protein crystals are exposed to X-rays, a diffraction pattern is generated from the interference between the electrons surrounding the atoms and the X-rays. From the diffraction pattern the arrangement of atoms within the protein crystal, *i.e.* the structure of the protein can be calculated.

2.4.1 Protein crystallization

An X-ray diffraction experiment requires protein crystals. Protein crystals are built up by symmetrically arranged molecules in a three-dimensional pattern. These crystals are often grown in vapor-diffusion experiments where the protein sample is slowly dehydrated bringing the protein into nucleation, leading to crystal formation (Figure 2.4).

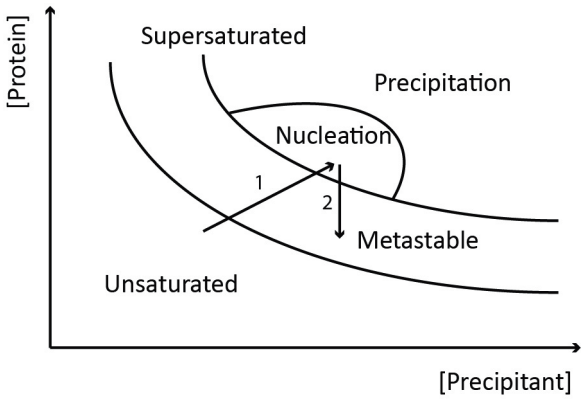

Figure 2.4: Phase diagram for protein crystallization. Illustration of the different zones involved in crystal formation. An ideal strategy for growing large crystals is indicated by the two arrows.

Protein crystallization is affected by several variables: protein concentration, ionic strength (salt concentration), pH, temperature and precipitant (*e.g.* polyethylenglycol - PEG). Protein crystals are held together with relatively weak forces compared to inorganic crystals, which are held together by electrostatic interactions of fully charged ions. The hydrogen bonds between the protein and water molecules in the solvent makes the crystal more fragile compared to an inorganic crystal.

Furthermore, membrane proteins have large hydrophobic regions covered by detergent micelles. As a consequence detergent-protein crystals are packed in a type II crystal form, where the hydrophilic ends of the protein are involved in crystal contacts and a large solvent space (containing detergents) surrounds the hydrophobic parts. There are alternative protein crystallization methods available in addition to the above mentioned vapor-diffusion method, *e.g.* membrane proteins can also be crystallized using the lipidic-cubic phase (LCP) method [121]. LCP crystals are often packed in the type I crystal form, where lipid monolayers are stabilized by hydrophobic and hydrophilic interactions. Recently, it has also been demonstrated that

it is possible to use LCP crystallization of membrane proteins which have been solubilized by styrene-maleic acid (SMA) copolymers instead of detergents [122]. SMA copolymers forms SMA nanodiscs when wrapped around planar lipid bilayer fragments containing membrane proteins and their surrounding lipids.

2.4.2 Structure determination using X-ray crystallography

The process of structure determination using X-ray crystallography goes through a number of steps:

- Collecting diffraction images.
- Identifying and integrating the reflections.
- Scale and merge the intensities.
- Solve the phase problem.
- Build and refine a model.
- Evaluate the model.

Diffraction is the interference of waves when a wave encounters an obstacle or slit comparable in size with its wavelength. The diffraction pattern which is observed when a crystal is exposed to X-rays is described by Bragg´s law (Equation 2.1), where n is a positive integer, λ is the wavelength of the incoming wave, d is the spacing between lattice planes and θ is the scattering angle.

$$n\lambda = 2dsin\theta \tag{2.1}$$

The difference in intensity between measured reflections are related to the atomic arrangement and the types of atoms in the unit cell. The unit cell is the smallest repeating unit with translational symmetry within the crystal. The structure factor F_{hkl} describes the amplitude $|F_{hkl}|$ and phase α of a wave diffracted from crystal lattice planes characterized by Miller indices h,k,l. The electron density ρ(x,y,z) at a position x,y,z in the unit cell is obtained by calculating the Fourier transform of the structure factor F_{hkl} (Equation 2.2).

$$\rho(x,y,z) = \frac{1}{V}\sum_{h}\sum_{k}\sum_{l}|F_{hkl}|e^{-2\pi i(hx+ky+lz)+i\alpha(hkl)} \tag{2.2}$$

During an experiment it is only possible to experimentally measure the intensities, which is proportional to $|F_{hkl}|^2$, whereas the information regarding the phase is not obtained; this is often referred to as the "phase problem" of crystallography.

To obtain an electron density map of the protein structure, the phase problem has to be solved. One of the most commonly used methods is molecular replacement (MR) where the phases are calculated from a similar previously solved structure. In general a sequence identity of $> 25\%$ and an r.m.s. deviation of < 2 Å between Cα atoms of the model and structure are required for a successful molecular replacement [123]. If MR does not solve the phase problem, single- or multiple-wavelength anomalous diffraction (SAD or MAD) can be used instead.

In a SAD experiment, heavy atoms have been incorporated into the protein structure at a limited number of sites, where they do not alter the actual structure of the protein in its native form. These heavy atoms will give rise to differences in the intensities of reflections that are related by inversion symmetry [124]. The anomalous differences are used to locate the position of the heavy atoms in the substructure and subsequently used for estimating the phases of the entire structure [125]. Once the phases have been obtained an electron density map can be calculated and a preliminary model can be built. The obtained model is refined until agreement between the model and the data is acceptable.

2.4.3 Validation of the data and the structure

During both data processing and model building there are a number of validation parameters calculated. Although there are guidelines of where to draw the line for a good data set and subsequently a good model it is worth to note that the final value of validation parameters might deviate a bit. The protein structural model is refined in an iterative process, evaluating the calculated parameters, the model and the electron density map.

In a first step it is necessary to decide up to which resolution limit the data should be used. The reflections observed in the higher resolution shells tend be weaker than the rest of the data, which will affect the parameters in a negative way. The signal to noise is described with $I/\sigma(I)$, which is the average ratio of reflection intensity to its estimated error.

The accuracy in the final estimation of the averaged reflection intensity is better the higher the redundancy, since every reflection is measured with a certain degree of error. R_{sym}, sometimes called R_{merge} is a measure of the spread of individual intensities of all symmetry-equivalent reflections. The correlation coefficient $CC_{1/2}$ divides the data in two randomly distributed

sets and calculates the correlation between intensities in the two half data sets. These parameters should also be monitored.

During the refinement process, R_{work} and R_{free} are calculated, which are measures of the global relative discrepancy between the experimentally obtained structure factor amplitudes and calculated structure factor amplitudes obtained from the model. The calculation of R_{free} is performed on a small fraction of the complete data which has been omitted from refinement. In this way over-interpretation of the data can be indicated. Local evaluation of the model is done by examining the local fit of the model to the electron density, evaluating the backbone torsion angles and bond lengths as well as side chain conformations of the individual amino acids.

2.5 Characterization of protein activity and the interaction with small-molecule ligands

In drug discovery, an important step is to determine the affinity between a target molecule and its interaction partner. There are several methods available for determining the affinity of a small-molecule ligand to the protein of interest. However, all methods might not work for all proteins, and/or they require a lot of optimization. In the following sub-chapter a few of these techniques will be introduced.

2.5.1 Substrate transport across a bilayer

A membrane protein transporter transports its substrate over a lipid bilayer without covalently modifying the substrate. During this process the transporter undergoes conformational changes which alternately exposes the substrate to either side of the membrane, a process which is generally more complex than many enzyme systems. Through *in vitro* experiments the flux of substrate mediated by the transporter can be characterized by the Michaelis-Menten parameters K_M and V_{max} [126], a model initially developed to describe the enzyme conversion of substrate to product. For membrane protein transporters the transport velocity reaches a plateau (steady-state) at high substrate concentrations, where above this values there is no significant increase in uptake rate. When all the substrate sites are occupied, the transporter is saturated and the rate of transport is maximal, V_{max}. From a plot of transport velocity versus substrate concentration both V_{max} and K_M can be determined. K_M is the substrate concentration at which the reaction rate is half of V_{max}.

2.5.2 Binding affinity

The strength of the interaction between a protein and its ligand/binding partner (substrate/inhibitor/drug) is determined by the binding affinity. Binding affinity is often reported as the equilibrium dissociation constant K_d. In the following reaction (Equation 2.3), P is the protein, L is the ligand and PL is the protein-ligand complex.

$$P + L \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} PL \tag{2.3}$$

The rate constant is defined as k_{on} in the forward direction and k_{off} in the reverse reaction. At equilibrium (Equation 2.4), the rate in both directions are equal hence:

$$[P][L]k_{on} = [PL]k_{off} \tag{2.4}$$

This relationship is referred to as the law of mass action and the dissociation constant K_d (Equation 2.5) describes this relationship as follows:

$$K_d = \frac{k_{off}}{k_{on}} = \frac{[P][L]}{[PL]} \tag{2.5}$$

2.5.3 Microscale thermophoresis – MST

With microscale thermophoresis (MST) it is possible to determine the equilibrium dissociation constant K_d in solution by taking advantage of the fluorescent signal of one of the binding partners.

Thermophoresis is the directed movement of molecules along a temperature gradient. This physical principal depends on changes in size, charge and solvation shell of molecules. During an MST-experiment the variation in fluorescent signal of an immobilization-free molecule is measured. The variation in fluorescent signal is a result of an induced temperature gradient obtained by an IR laser. Furthermore, the variation in the fluorescent signal correlates with the binding of a non-fluorescent ligand to the fluorescent binding partner. A local change in temperature ΔT leads to a local change in molecule concentration (depletion or enrichment) [127].

In a serial dilution, the concentration of the non-labeled ligand is increased so that the fraction of complex increases until all fluorescent molecules have formed complex with the titrated ligand. As a result, the fraction of bound molecules can be derived from the measured change in normalized fluorescence. A binding curve can be obtained, from where the K_d can be

determined, by plotting normalized fluorescence versus the concentration of non-labeled ligand.

2.5.4 Fluorescence resonance energy transfer – FRET

Fluorescence resonance energy transfer (FRET) [128] is the non-radiative energy transfer from a donor to an acceptor chromophore. FRET occurs when the emission and excitation spectra of the donor and acceptor chromophores overlap, hence they need to be in close vicinity to each other (10-100 Å).

It is common to use two different chromophores as the donor and acceptor. In the MraY FRET assay the donor fluorophore is attached to the substrate UDP-*N*-acetylmuramyl-pentapeptide (UNAM-pp). The donor fluorophore, BODIPY-FL, has an excitation wavelength of 485 nm and an emission wavelength of 520 nm. The FRET acceptor lipid, lissamine rhodamine B dipalmitotyl phosphatidylethanolamine (LRPE) has an emission wavelength of 590 nm.

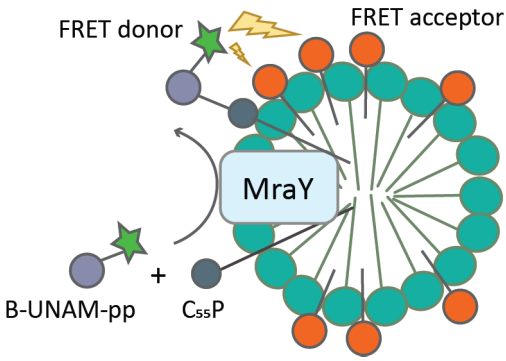

Figure 2.5: Schematic of the FRET assay used for determining MraY activity. Detergent micelles (green) contain MraY (blue), undecaprenyl phosphate ($C_{55}P$) (grey), FRET acceptor lipid (red) and UDP-*N*-acetylmuramyl-pentapeptide (UNAM-pp) (purple) labeled with FRET donor (green) (B-UNAM-pp). The FRET donor comes in close vicinity to the FRET acceptor upon transfer of the labeled UNAM-pp to the $C_{55}P$ by MraY and UMP is released (not shown). A reduction in donor fluorescence occurs during energy transfers from the FRET donor to the FRET acceptor (yellow lightning bolt). Illustration inspired from [99].

The principal of the MraY FRET assay is illustrated in Figure 2.5. MraY protein is embedded in a detergent micelle together with the lipid substrate ($C_{55}P$) and the FRET acceptor-labeled lipid (LRPE). Furthermore, this is combined with the soluble FRET donor-labeled peptide substrate (B-UNAM-pp). MraY catalyzes the transfer of the labeled soluble substrate

B-UNAM-pp to $C_{55}P$, which brings the FRET donor fluorophore in close proximity to the FRET acceptor. Upon excitation of the donor fluorophore, energy is transferred from the donor fluorophore to the acceptor fluorophore, resulting in a decrease in donor fluorescence and potentially an increase in acceptor fluorescence. The ratio change $\Delta(F_{590}/F_{520})$ increases in size with reaction time since the F_{520} intensity decreases due to the substrate being converted to product and the F_{590} intensity is almost constant [99].

By adding an inhibitor, the FRET- assay can be used to determine the IC_{50}-values of potential inhibitors. The potency of a drug can be described by the half-maximal inhibitory concentration (IC_{50}). It is a measure of how much drug is needed to inhibit a biochemical reaction by 50%.

Chapter 3

Sialic Acid Transporters

The main goal with the sialic acid transporter project was to determine the structure of a sialic acid transporter, mainly focusing on the SiaT transporter. In this process it is important to find suitable candidates for structural and functional studies. This chapter is divided into three sub-chapters describing and discussing the results from **Paper I** - **Paper IV**. Chapter 3.1 presents the work from gene to pure protein (**Paper I**, **Paper III** and **Paper IV**). In chapter 3.2 the structure of the first bacterial sialic acid transporter SiaT from *Proteus mirabilis* (*Pm*SiaT) is presented (**Paper II**). Finally, in chapter 3.3 the bacterial growth assay used to characterize SiaT is described (**Paper I** - **Paper III**).

3.1 From gene to pure protein

As with most membrane protein projects involving structure determination, it all started with a set of genes. Adopting an orthologous approach is often favorable to increase the chances of obtaining a construct suitable for structure determination. The empirical process of membrane protein production and structure determination includes a number of bottlenecks which need to be overcome.

3.1.1 Construct design and overexpression trials

For overexpressing membrane proteins, constructs are generated suitable for the expression host. The sialic acid transporters under investigation have their origin from bacterial organisms, hence the selected host for overexpression was the *E. coli* system. To increase the amount of overexpressed protein, the genes encoding the target protein can be codon optimized for

the selected expression host, to improve the protein translation efficiency affected by codon bias [129]. This approach was applied for the selected genes encoding sialic acid transporters.

Initially genes encoding SiaT were cloned into the pBAD-vector for titratable expression of the recombinant protein through regulation of carbon source [130]. Expression screening of SiaT was performed, varying L-arabinose concentration using standard protocols. With the pBAD-system, it was possible to express SiaT although in very small amounts only detectable with Western blot (data not shown). In contrast to SiaT, the pBAD expression system was more suitable for one of the TRAP constructs (*Pasteurella multocida* TRAP) (**Paper IV**).

Topology prediction [131] of genes encoding SiaT was predicted to consist of thirteen transmembrane helices with the N-terminal located on the extracellular side and the C-terminal located on the intracellular side. A set of seven SiaT genes from different bacterial organisms (*Salmonella enterica* servoar Typhimurium (*Se*), *Staphylococcus aureus* (*Sa*), *Streptococcus pneumoniae* (*Sp*), *Clostridium perfringens* (*Cp*), *Vibrio fischeri* (*Vf*), *Photobacterium profundum* (*Pp*) and *Proteus mirabilis* (*Pm*)) were therefore cloned into the pWarf(-) vector, generating SiaT-GFP fusion constructs (**Paper I**). Furthermore, additional constructs were generated for those SiaT genes which were predicted to possess a disordered region within their C-terminal [132]. The reason behind generating ΔSiaT-GFP constructs was to be proactive and foresee any potential negative impact upon crystallization since disordered regions often have a negative impact on the ability for a protein to crystallize.

SiaT-GFP and ΔSiaT-GFP constructs were screened for overexpression in the *E. coli* Lemo21(DE3) strain. The expression level was monitored by measuring the relative fluorescence (RFU) (**Paper I**). The screening aimed to find optimal temperature and an optimal concentration of L-rhamnose, where the latter acts as an inducer for expression of T7 Lysozyme. T7 Lysozyme in turn inhibits T7 RNA polymerase and reduces the levels of leaky expression from the T7 expression system [107] which could be toxic for the host cell. Of the thirteen constructs aimed for test expression, two were unsuccessfully cloned and one of these two was used as a negative control during test expression. For the remaining eleven constructs, all were successfully expressed with varying expression yield (**Paper I**).

In contrast to SiaT, the TRAP sialic acid transporter was predicted to have the C-terminal located on the extracellular side (**Paper IV**). The genes encoding TRAP sialic acid transporters from *P. multocida* (Pm), *Fusobacterium nucleatum* (Fn), *Vibrio cholera* (Vc) and *Haemophilus influenzae* (Hi) were cloned into both the pWarf(-) and pWarf(+) vector, where the latter

harbors an additional Glycophorin A transmembrane helix which directs the C-terminal of the protein into the intracellular side. The location of the C-terminal is important since GFP is not able to fluoresce if situated on the extracellular side. All pWarf(+) TRAP constructs except PmTRAP were successfully expressed. Furthermore, FnTRAP and PmTRAP constructs in pWarf(-) did not show any fluorescence, hence the topology prediction was confirmed (**Paper IV**).

In general, expression at the lowest temperature tested (25°C) gave the highest RFU values for SiaT-GFP and ΔSiaT-GFP fusion proteins, indicating a higher expression yield. Lowering the temperature slows down the protein expression process and allows the protein to be correctly folded and inserted into the membrane [133]. The optimal temperature for expression of the TRAP sialic acid transporters in the pBAD-system was determined to 18°C (**Paper IV**).

Control of expression is a key factor for recombinant membrane protein expression. Obtaining correctly folded and stable protein is important as a first step towards crystallization experiments. The pWarf-system proved to be successful for the SiaT constructs. In contrast, the pBAD-system was more suitable for the TRAP sialic acid transporters.

3.1.2 Detergent screening using FSEC

Membrane proteins are embedded in a lipid bilayer and need to be extracted upon purification. The natural lipid bilayer shields the hydrophobic regions of the membrane protein. Amphipathic detergents extract the membrane protein from their natural environment upon solubilization and covers the hydrophobic regions. An optimal detergent is able to solubilize the membrane protein efficiently and keep it in a functional conformation and prevent aggregation.

Once expression of SiaT-GFP was established, a set of promising candidates were screened for suitable detergent, using FSEC (**Paper I**). FSEC relies on the incorporation of a fluorescent tag to the protein. With FSEC, the elution profile of the SiaT-GFP fusion protein is monitored and evaluated. The method allows for screening of detergents in a relatively fast and efficient way evaluating expression levels, degree of monodispersity and approximate molecular weight of the fusion protein [134].

In **Paper I** the initial detergent screening of three SiaT-GFP fusion proteins (*Pm*SiaT, *Sa*SiaT and the third hereafter called *Xx*SiaT) using FSEC (Figure 3.1) is presented. The fluorescent elution profiles indicated monodisperse peaks for the majority of the detergents tested, although the width of the peak varied between constructs and detergents. The smaller detergent β-OG was not able to efficiently solubilize SiaT-GFP. The majority of the β-OG solubilized sample was eluted close to/in the void volume indicating aggregated sample.

The results from the FSEC profiles together with the efficiency for the detergent to solubilize SiaT-GFP (RFU-values) were combined to rank suitable detergents for solubilization and purification. The most promising detergents identified for SiaT-GFP were the gentle maltoside detergents DM and DDM, which differ from each other by the length of their carbon tail (C10 and C12, respectively). DDM was selected as the best candidate for *Pm*SiaT, and DM was selected for *Sa*SiaT and *Xx*SiaT. The choice of detergent was further examined and re-evaluated based on large-scale size-exclusion purification profiles.

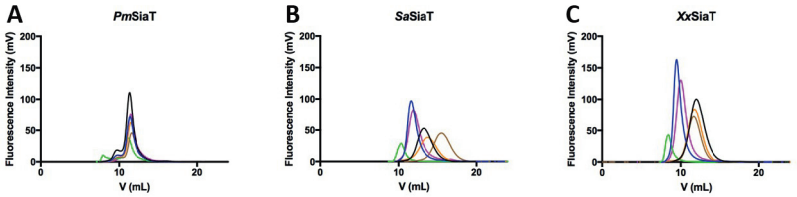

Figure 3.1: Detergent screening of SiaT from three different bacterial organisms using FSEC. SiaT-GFP was solubilized from *E. coli* membranes with different detergents and the soluble fraction was loaded onto a Superdex 200 10/300 GL column. Detergents tested: DDM (black), DM (blue), MNG12 (brown), MNG10 (orange), β-OG (green) and fc12 (pink). A) *Pm*SiaT-GFP. B) *Sa*SiaT-GFP. C) *Xx*SiaT-GFP.

In **Paper IV** the detergent screening using FSEC indicated that the new class of detergents, maltose-neopentyl glycol (MNG) [135] was the most promising for the TRAP sialic acid transporters. For VcTRAP and Pm-TRAP (not evaluated using FSEC) this result agreed with detergent screening of constructs expressed in the pBAD-system as evaluated by Western-blot.

3.1.3 Purification of SiaT

Purification of SiaT follows mainly a two-step purification scheme (**Paper I**). In the first affinity-purification step, the detergent-solubilized material was immobilized onto a HisTRAP column (IMAC purification), washed and eluted with an increased concentration of imidazole. Affinity purification is often used as a first purification step since it specifically binds an accessible affinity tag which has been introduced during cloning. Purification was monitored at 280 nm (measuring intrinsic fluorescence of proteins) and with a fluorescence detector (excitation 485 nm and emission 512 nm) measuring GFP-fluorescence.

The GFP-His-tag was removed from SiaT with HRV3C protease to increase the chances of obtaining a stable protein sample suitable for crystallization. HRV3C protease specifically cleaves the polypeptide at its recognition site, which is located after the SiaT protein and before the GFP-His-tag. To separate the GFP-His-tag from the protein after cleavage, a reverse IMAC step is often preferred since it allows for the affinity-tag to bind to the resin while the target protein can pass through. Due to complete removal of the GFP-His-tag by the HRV3C protease and the significant difference in size, this additional purification step could be excluded from the purification scheme and the tag was fully separated from the protein during size-exclusion chromatography.

Size exclusion chromatography (SEC) separates molecules based on their size, where small molecules travel a longer path through the porous stationary phase compared to large molecules which do not get delayed. As a final step the protein was purified with size-exclusion chromatography to obtain a monodisperse tag-free protein sample (Figure 3.2). Important characteristics to evaluate during size-exclusion chromatography is the homogeneity of the protein identified with a sharp Gaussian peak. Broad peaks often indicate the presence of oligomers. In order to try and capture the substrate-bound form of SiaT, the sialic acid Neu5Ac was included in the size exclusion buffer. The substrate most likely also further stabilizes the protein during purification.

Even though the FSEC detergent screen and complementing RFU-measurements of solubilized protein pointed towards using DM as detergent for the *Sa*SiaT construct (**Paper I**), large scale SEC purifications resulted in a double peak profile. Hence the detergent was exchanged to DDM which reduced the size of the first shoulder in the SEC profile (**Paper III**). Analytical ultracentrifugation experiments of *Sa*SiaT concluded that *Sa*SiaT can be purified as a monomer in the presence of ∼200 DDM detergent molecules (**Paper III**).

Establishing a purification protocol can be tedious. There are several

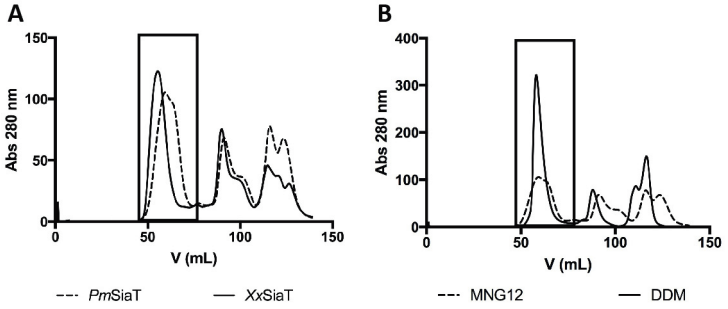

Figure 3.2: Size-exclusion profiles of SiaT. A) Comparison of profiles for *Pm*SiaT and *Xx*SiaT solubilized and purified with the detergent MNG12. *Xx*SiaT displays a narrower and more Gaussian shaped peak than *Pm*SiaT. B) Comparison of *Pm*SiaT in the detergents MNG12 and DDM. The elution profile for *Pm*SiaT in the DDM detergent is more desirable than the MNG12-profile. In both A) and B), the first peak at ∼50 mL is SiaT without any fusion-tag, peak 2 at ∼80 mL is the free GFP-His-tag, peak 3 at ∼120 mL is impurities.

variables which can affect protein stability and purity. Incorporating a GFP-fusion tag to the SiaT protein proved to be successful. The GFP-fusion tag allowed for expression screening in a much faster and reliable way. Furthermore the advantage with having a fluorescent tag allowed for detergent-screening and re-evaluation of the monodispersity and purity in subsequent purifications. *Pm*SiaT was selected as the top candidate for crystallization studies, based on expression-level and SEC profile.

3.1.4 Crystallization of SiaT

Hydrophilic interactions is an important force for the formation of protein crystals. Since membrane proteins are surrounded by a detergent micelle, the possibilities for hydrophilic interactions are limited, and the choice of detergent affects the size of the protein-detergent complex. In **Paper I** the conditions obtained from conventional crystallization screens which resulted in crystal formation of *Pm*SiaT using hanging-drop vapor diffusion experiments are described. Furthermore, optimization of crystal conditions by performing grid screens, resulted in crystals which diffracted beyond 3 Å resolution.

3.1.5 Summary

In chapter 3.1 the path from construct design and cloning of sialic acid transporters to purification and crystallization yielding SiaT crystals, which diffracted beyond 3 Å resolution has been described. Adopting an orthologous approach for identifying suitable constructs for crystallization of SiaT and TRAP sialic acid transporters and structure determination of SiaT proved to be successful. Reasonable overexpression levels of SiaT were identified in the titrable *E. coli* Lemo21(DE3) strain at low temperatures. Furthermore, the GFP-fusion tag facilitated detection of expression levels and evaluation of detergents for the SiaT constructs.

3.2 The crystal structure of *Pm*SiaT

In **Paper II** the first structure of a sialic acid transporter, SiaT from *P. mirabilis* (*Pm*SiaT) is reported. The Gram negative uropathogen *P. mirabilis* is able to catabolize host derived sialic acids as a source of nutrient. The structure of *Pm*SiaT was determined at 1.95 Å resolution, using X-ray crystallography, in its outward-open conformation with the substrate Neu5Ac and two Na^+ ions bound (Na2 and Na3) (Figure 3.3).

3.2.1 Data collection and structure determination

Optimized crystals of *Pm*SiaT (**Paper I**) allowed for data collection of a native data set of *Pm*SiaT with diffraction to 2.26 Å resolution (**Paper II**). Several attempts to obtain the phases with molecular replacement (MR) using LeuT-like structures as search models were unsuccessful. Hence the focus shifted towards the selenomethionine (SeMet)-approach, incorporating heavy atoms into the protein structure without changing its conformation. Selenomethionine replaces methionine residues in the protein when the protein is expressed in a medium containing selenomethionine as the sole source of methinonine. SeMet-incorporated crystals were obtained in similar crystallization conditions as for the native data set, with 1% OG and 20 mM Neu5Ac added to the reservoir. Several data sets of SeMet crystals were collected although none had high enough redundancy for the overall individual data sets. The phases were obtained by merging thirteen SeMet data sets (**Paper II**).

3.2.2 Overall structure of *Pm*SiaT

In agreement with topology prediction, SiaT comprises of thirteen transmembrane helices (TM0 and TM1-TM12) with the N- and C-termini facing the extracellular side and intracellular side respectively (Figure 3.3). *Pm*SiaT is captured in a substrate and sodium bound outward-open conformation [136]. SiaT belongs to the SSS-family and adopts the LeuT-fold [14, 137], where two inverted repeats (TM1-TM5 and TM6-TM10) forms the structural core of the protein (Figure 3.3B). The four-TM helix bundle (core domain) consists of the first two helices from each repeat TM1-TM2 and TM6-TM7, the hash domain (scaffold domain) consists of the third and fourth helix of each repeat TM3-TM4 and TM8-TM9 and the two peripheral arms (gating helices) correspond to the fifth helix of each repeat TM5 and TM10.

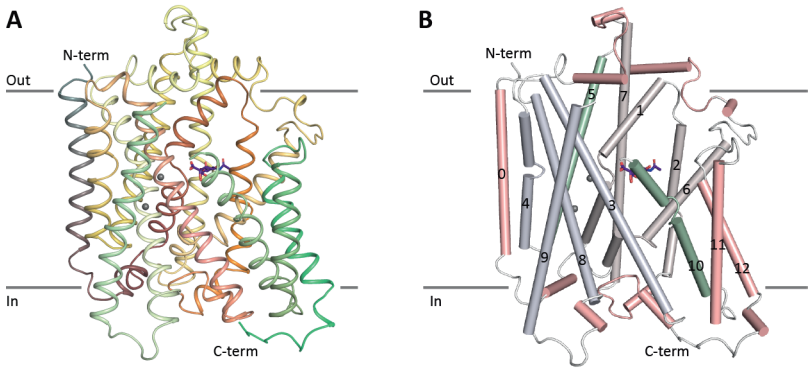

Figure 3.3: The structure of *Pm*SiaT in its outward-open sialic acid and sodium ion bound conformation. A) The *Pm*SiaT structure is colored from N- to C-terminus (blue, red, and yellow to green), Neu5Ac (purple) is displayed as sticks colored by atom and the two sodium ions are displayed as spheres (grey) (PDB ID: 5NV9). The lipid membrane is indicated with grey lines. B) Structural elements that characterize the LeuT-fold of *Pm*SiaT are displayed. The four-TM helix bundle/the core domain (TM1-TM2 and TM6-TM7 (beige)), the hash-/scaffold domain (TM3-TM4 and TM8-TM9 (purple)) and the two peripheral arms/the gating helices (TM5 and TM10 (green)). Additional helices not part of the main LeuT-fold are colored in pink. All transmembrane helices are numbered (0, 1-12).

3.2.3 Sialic acid binding site

Near the center of the protein, residues from four helices (TM1-TM3 and TM6) line the sialic acid binding site, including the discontinuous helices (TM1 and TM6) which are important for substrate binding. Discontinuous helices play an important role in the alternating access mechanism [21, 22, 137, 138].

In the sialic acid binding pocket, the β-anomer of Neu5Ac is coordinated by eight amino acid residues and seven water molecules (Figure 3.4). Three residues from TM1 (Thr58, Ser60 and Thr63) form both side- and main chain hydrogen bonds with Neu5Ac. More specifically, Ser60 and Thr63 coordinate the negatively charged carboxylate group (at C1), whereas Thr58 and Thr63 coordinate hydroxyl groups from the glycerol tail. In addition, Thr58 coordinates the hydroxyl group at the C2 position. The carboxylate group of Neu5Ac also makes a salt bridge interaction with the neighboring guanidinium group of Arg135 (TM3). Interaction between a basic amino acid residue and a sugar molecule has been observed previously in vSGLT [21].

Furthermore in the soluble sialic acid binding protein SiaP, sialic acid interacts with a conserved arginine which is important for high affinity substrate recognition by the sialic acid TRAP transporter [139]. Additional hydrogen bonds between a hydroxyl moiety of the glycerol tail is observed with Gln82 (TM2). Amino acids Phe78, Gly81 (TM2) and Phe243 (TM6) create a region with a neutral electrostatic surface which surrounds the acetyl amino moiety of the Neu5Ac methyl group.

Several water-mediated interactions are observed in the hydration layer which occupies the space between Neu5Ac and TM5-TM6. The amino acid residues Gln82, Phe78 (TM2) and Asn247, Gln250 (TM6) all participate in water-mediated interactions with Neu5Ac.

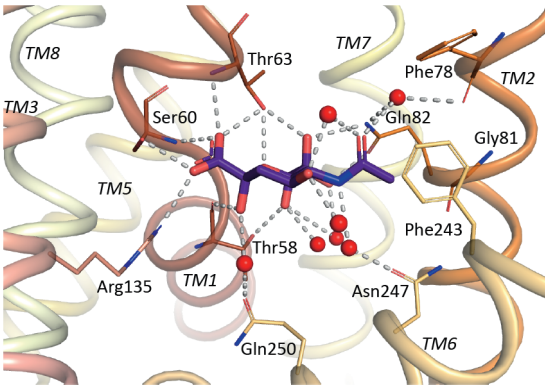

Figure 3.4: The sialic acid binding site of *Pm*SiaT. Neu5Ac (purple) is represented as sticks colored by atom and water molecules are represented as spheres (red). Neu5Ac forms interactions with amino acid residues from TM1-TM3 and TM6, including water-mediated hydrogen bonds with amino acid residues from TM2 and TM6 (PDB ID: 5NV9).

3.2.4 Characterization of the sialic acid binding site

The ability of SiaT to transport the sialic acid Neu5Ac was demonstrated with proteoliposome experiments. In these experiments the purified SiaT-transporter was inserted into an artificial bilayer. After reconstitution, a potassium ion diffusion potential was generated by adding valinomycin. Valinomycin is a lipid-soluble molecule which selectively binds potassium ions and facilitates their transfer across lipid bilayers. Upon the addition of radiolabeled substrate in the presence of sodium ions, the co-transport was initiated. Transport was terminated by removing the excess of substrate and measuring the radioactivity of the radioactive substrate which had been

transported into the proteoliposomes. Single mutations of amino acids which make side chain interactions with Neu5Ac (Thr58Ala, Ser60Ala, Thr63Ala, Gln82Asp and Arg135Glu) were designed to disrupt the transport. All except Thr58Ala, where Thr58Ala coordinates the anomeric hydroxyl at the C2 position of Neu5Ac through a main chain interaction, abolished transport.

Through proteoliposome experiments the maximal transport activity was determined to 1800 nmol/mg protein (0.4 nmol). In addition, V_{max} (187 ± 30 nmol/mg protein/min) and K_M (16 ± 4 μM) could also be determined. Furthermore, the binding affinity (K_d) between sialic acid and *Pm*SiaT was determined with isothermal titration calorimetry (ITC) and microscale thermophoresis (MST) (Table 3.1).

Table 3.1: The binding affinity constant K_d between *Pm*SiaT and different sialic acid derivatives. K_d was determined by MST and ITC (for Neu5Ac).

	K_d		
	Neu5Ac	**Neu5Gc**	**KDN**
MST	58 ± 1 μM	85 ± 2 μM	>10 mM
ITC	50 ± 4 μM	–	–

The observed difference between K_d and K_M for Neu5Ac can possibly be explained by experimental differences. In a proteoliposome experiment the protein is embedded in a lipid bilayer in contrast to the detergent micelle which surrounds the protein in the ITC and MST experiments.

3.2.5 Sodium binding sites

In an electron density map it can be difficult to distinguish sodium ions from water molecules [140]. However, there are distinguishing features of sodium ion binding sites. Sodium ions coordinate with 4-8 partners and the bond distances are less than 2.7 Å [141]. In the *Pm*SiaT structure two sodium binding sites were observed. One sodium ion was modeled in the conserved Na2 site [22] in a protruding kink in TM1 and interestingly the second sodium ion was modeled in a unique position termed Na3 between TM1 and TM8 (Figure 3.5).

In proteins with the LeuT-fold, the Na2 sodium binding site is highly conserved and is mainly formed by residues from the first helix of the first repeat and the third helix of the second repeat (TM1 and TM8) [15]. This is consistent with the Na2 site of *Pm*SiaT where the sodium ion in the Na2 site is coordinated by the carbonyl oxygen atoms of Ala56 and Leu59 (TM1), the hydroxyl groups of Ser342 and Ser343 (TM8) and the main-chain of carbonyl oxygen of Ala339 (TM8).

The sodium ion at the Na3 site is not in contact with the substrate, nor close to the Na1 or Na1´sites, which are present in some transporters with the LeuT-fold (described in chapter 1.3.2). The Na3 site is located further 6.5 Å towards the cytoplasmic side compared to the Na2 site. This is ∼14 Å away from the sialic acid binding site. The sodium ion at the Na3 site is coordinated by the main-chain carbonyl group of Ser342, the hydroxyl groups of Ser345 and Ser346 (TM8) and by the carboxyl group of Asp182 (TM5). In the electron density map there are clear density peaks for both the Na2 and Na3 sites, no negative electron density was observed when modeling a sodium ion in these sites. Both sodium ions coordinate five atoms and the bond distances are in the range of 2.2 to 2.5 Å.

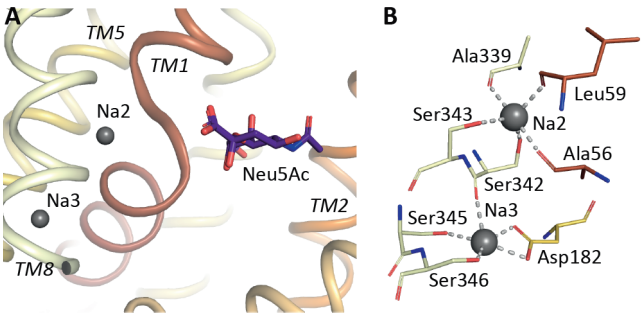

Figure 3.5: Sodium binding sites in *Pm*SiaT. A) Illustration of the location of Na2, Na3 and the substrate binding site in *Pm*SiaT. Sodium ions are illustrated as spheres (grey) and Neu5Ac (purple) as sticks colored by atom (PDB ID: 5NV9). B) The Na2 and Na3 sodium binding sites. Amino acid residues interacting with the sodium ions (grey) are represented as sticks, colored by atom.

3.2.6 Characterization of the sodium binding sites

How important are these sodium sites for the transport of sialic acid? It is known from previous studies of LeuT that when a sodium ion is bound to the Na2 site it stabilizes the structure in the substrate binding pocket [15]. Hence binding of sodium in the Na2 site is required for substrate binding.

In the proteoliposome assay (Figure 4c, **Paper II**), the single mutants Ser342Ala and Ser343Ala of the Na2 site both abolished sialic acid transport. This demonstrates the importance of the Na2 site for sialic acid transport of SiaT. Even though the sodium ion in principal would be able to bind by coordinating the remaining four amino acids, it is believed that the structural rearrangement, which occurs upon sodium binding, are important for directing/positioning the sialic acid in its binding pocket. Molecular dynamic

(MD) simulations of different states further shed light on the different roles of the sodium binding sites. The sodium sites and the substrate binding site are connected through a network of amino acids. In the crystal structure, Ser342 coordinates both the sodium ion at the Na2 and Na3 sites. In addition, the Na2 site is further connected to the substrate binding site through Leu59 which is located within the unwound segment of TM1 next to Ser60 which coordinates with the carboxyl group of Neu5Ac. In the MD simulations, Ser60 is stabilized and able to maintain a bidentate interaction with the carboxylate group of Neu5Ac due to interactions between the neighboring Leu59 and the sodium ion at the Na2 site. In the absence of sodium at the Na2 site Neu5Ac is less stable in the MD simulations, further strengthening the importance of the Na2 site.

When taking a closer look at the Na3 site, the single mutations Ser345Ala and Ser346Ala of the Na3 site did not abolish sialic acid transport in the proteoliposome experiments. However, Asp182 which has a bidentate coordination with the sodium ion at the Na3 site abolished transport when mutated to Ala. These results suggest that the binding of a sodium ion at the Na3 site still is possible if the coordination is at least four. The presence of a sodium ion at the Na3 site stabilizes the transporter in its outward-open conformation and facilitates import of sialic acid.

It has been demonstrated that a sodium ion at the Na2 site in LeuT causes an amplification of the selectivity toward sodium at the Na1 site [142]. This suggests that the order of binding would start with sodium binding to the Na2 site before the Na1 site. This is also logical since the Na1 site is located further towards the extracellular side in LeuT. However, the Na3 site in SiaT is located on the opposite side of Na2, closer to the cytoplasmic side. The location of the Na3 site would suggest that sodium binds first to the Na3 site, followed by sodium binding to the Na2 site and sialic acid binding to the substrate site. MD simulations demonstrate how a sodium ion in the Na2-only simulation results in movement of the sodium ion towards the Na3 site. In the Na2-Neu5Ac simulation, the sodium ion is stable in the Na2 position, highlighting the importance of the Na2 site upon substrate binding.

What about the stoichiometry? Cooperativity of ligand binding in a protein is described by the Hill coefficient where a cooperative relationship between ligand binding sites is indicated with a Hill coefficient > 1. In SiaT, the sodium transport is cooperative (Hill coefficient $= 1.5 \pm 0.1$). With a sodium:substrate stoichiometry of 2:1, the net transport is one positive charge.

Is the Na3 site conserved? The stoichiometry and presence of sodium binding sites in protein structure with the LeuT-fold vary, apart from the

conserved Na2 site. The newly identified Na3 site is present in 19.6% of analyzed SSS sequences, which also contains the Na2 site. However, it is not present in the closely related protein families, neurotransmitter sodium symporter (NSS), betaine/choline/carnitine (BCC) and nucleobase-cation-symport (NCS1) which all have hydrophobic residues at this site.

Interestingly there are differences between secondary transporters which adopts the LeuT-fold, concerning their stoichiometry and location of sodium sites. The *Pm*SiaT structure is the first characterized structure with a Na3 site, located on the opposite side of Na2 compared to the Na1 site. This novel Na3 site enhances the substrate-binding affinity and structurally aids in the stabilization of the outward-facing conformation of SiaT.

3.2.7 Alternating access of SiaT

Secondary transporters transport solutes across a lipid bilayer in the co-transport of the driving ion(s). The transporter alternates access to the binding sites by adopting a set of different conformations in the transport cycle. The newly determined crystal structure of *Pm*SiaT, adopts an outward-facing open conformation with both the substrate and sodium ions bound. In an attempt to explore the overall transport mechanism, an inward-facing model of *Pm*SiaT was built based on the structure of vSGLT, which adopts an inward-facing conformation [21] and morphs between the two states was generated.

The extracellular open compartment of *Pm*SiaT closes upon movement of the N-terminal of TM10 towards TM1e and TM2. The movement further directs TM9 towards TM3 and the extracellular loop helices Elh7a and Elh7b (connecting TM7 and TM8), collapses into the open compartment and interact with TM1e. A hydrophobic gate consisting of amino acids Trp404 (loop TM9-TM10), Ile67 (TM1) and Phe78 (TM2) occupy the space above the substrate binding site. In a closed conformation the outer gate is stabilized by hydrogen bonds between residues from the loop connecting TM9 and TM10 (Ala401 and Glu402) and Thr312 (Elh7) and Thr73 (TM1e) respectively.

On the inside, the inner hydrophobic gate (the intracellular loop/helix Ilh0) that occludes the substrate from the cytoplasm moves away and opens the transporter by breaking the hydrogen bonds which initially are formed with residues from the intracellular loop between TM4i and TM5. As Ilh0 moves away from the inner pore, the cytoplasmic ends of TM8 and TM9 participate in the opening and rearrange by moving away from the cavity. Salt bridges between Glu176 (loop between TM4i-TM5) and the two arginine residues Arg31 and Arg44 (Ilh0) are stabilizing the inner gate in its closed

conformation. Interestingly, Arg40 of the intracellular gating loop/helix (Ilh0) aligns with Arg5 in LeuT, which has previously been assigned to be involved in intracellular gating [17].

In addition, the conserved Arg260 (TM6) in SiaT is located below the Na3 site and interacts with Ser53 (TM1), Asp256 (TM6) and Ser346 (TM8) in the closed conformation. Based on proteoliposome experiments of the Asp260Glu mutant, which did not exhibit any uptake of Neu5Ac, it is possible that the side chain of the glutamic acid prevents the release of sodium ions and hence abolish the transport of Neu5Ac.

3.2.8 Summary

In chapter 3.2, the first determined X-ray structure of a sialic acid transporter, *Pm*SiaT, is presented. *Pm*SiaT adopts the LeuT-fold with two inverted repeats of five transmembrane helices, in an outward-open conformation bound to Neu5Ac at the substrate binding site and two sodium ions, one at the conserved Na2 site and one at the novel Na3 site.

3.3 Characterization of SiaT

3.3.1 SiaT is able to rescue the growth of *E. coli* Δ*nanT*

In order to investigate the functionality of SiaT used in the orthologous approach for identifying a suitable construct for structure determination (chapter 3.1), a whole cell bacterial growth assay [51] was performed (**Paper I**).

E. coli can use sialic acid as a carbon and nitrogen source [49]. By using an *E. coli* Δ*nanT* deletion strain, an *E. coli* strain which lacks the gene encoding the native sialic acid transporter NanT, in combination with sialic acid as the sole carbon and nitrogen source, it is possible to identify functional sialic acid transporters. If the sialic acid transporter is functional, it will transport the sialic acid into the cytoplasm where it can be further degraded and used as nutrition for the bacteria.

Seven genes predicted to encode SiaT (mentioned in chapter 3.1.1) and the gene encoding NanT from *E. coli* (*Ec*NanT) were all cloned into a low-copy vector and transformed into the *E. coli* Δ*nanT* knockout strain (*E. coli* JW3193 Δ*nanT* [143]) and grown on minimal media supplemented with Neu5Ac as the sole carbon source (**Paper I**).

The NanT transporter of *E. coli* [51] was used as a control in these experiments. All seven SiaTs were able to rescue the growth of *E. coli* Δ*nanT* when grown on minimal media supplemented with Neu5Ac as the sole carbon source (Figure 3.6) (**Paper I**). Furthermore *Se*SiaT had already been characterized [51] and was used as an additional positive control in these experiments. This bacterial growth experiment was important in the *in vivo* characterization of the predicted SiaT transporter. With this type of experiment it was not only possible to determine the functionality, it is also possible to use the setup to investigate preference for different sialic acid derivatives (discussed in **Paper I** and **Paper III**).

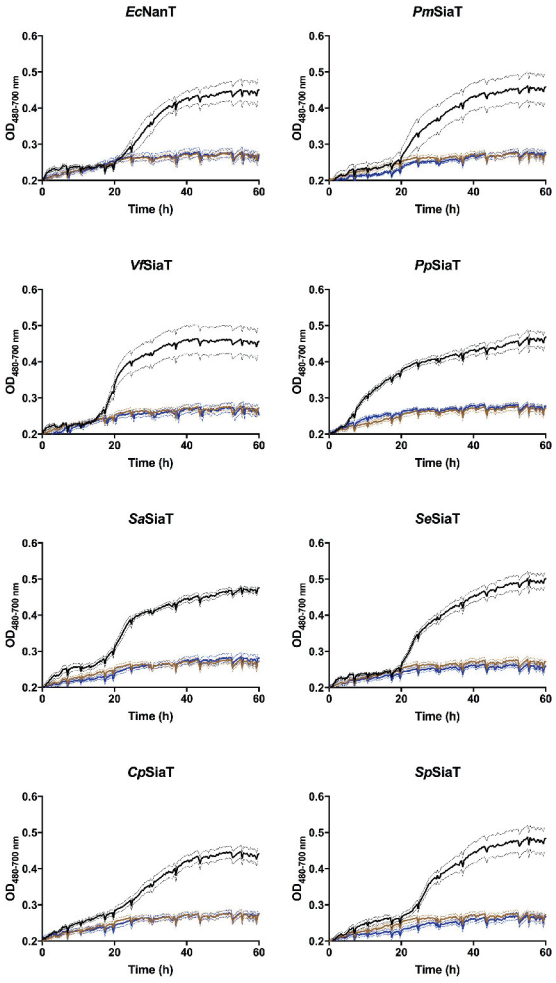

Figure 3.6: Bacterial growth experiments of predicted SiaT from seven different bacterial organisms and *E. coli* NanT. SiaT is able to rescue the growth of the *E. coli* $\Delta nanT$ strain using Neu5Ac as the sole carbon and nitrogen source. The black growth curve represents the predicted SiaT transporter grown on minimal media supplemented with Neu5Ac. The brown growth curve represents the *E. coli* JW3193 $\Delta nanT$ strain transformed with empty vector (used as a negative control) grown on minimal media supplemented with Neu5Ac and the blue growth curve represents the predicted SiaT transporter grown on minimal media without Neu5Ac. Growth curves represent the mean of six replicates ± SEM for *Ec*NanT, *Pm*SiaT, *Pp*SiaT, *Sa*SiaT, *Se*SiaT, *Sp*SiaT and five replicates ± SEM for *Cp*SiaT and *Vf*SiaT.

3.3.2 Characterization of *Pm*SiaT

In **Paper II** we demonstrate how *Pm*SiaT is able to rescue the growth of the *E. coli* Δ*nanT* strain when grown on M9 minimal media supplemented with Neu5Ac as the sole carbon source. In this experiment thiamine hydrochloride (Vitamin B1) had been added to the growth media to aid in catabolism. The optimized growth media resulted in a higher optical density compared to initial experiments. An additional positive control was included in these experiments, namely the wild type *E. coli* BW25113 strain [144]. Expression of NanT and SiaT was induced with IPTG. The growth lag observed compared to the *E. coli* wild type strain could be related to the IPTG induction. In the native *E. coli* strain, Neu5Ac is catabolized by five enzymes encoded by the nan operon (including NanT) [49–51,145,146]. The genes are induced by the presence of Neu5Ac in the growth medium through the loss of repression via the NanR protein [49, 145, 147]. In Figure 2A **Paper II**, there is a lag between when the native *E. coli* strain and the deletion strain starts their exponential growth phase. The lag could be a result of the difference in induction of the two systems and/or an increased metabolic burden caused by the overexpression of SiaT itself.

3.3.3 Sialic acid specificity between bacterial organisms?

The growth rates of SiaT from the seven different bacterial organisms, when grown on Neu5Ac, are overall relatively equal (**Paper I**). However, *Cp*SiaT displays a reduced growth rate compared to the other SiaTs investigated (Table 2, **Paper I**). The difference in growth rate might be a result of differences within the sialic acid binding site, hence allowing for different sialic acid derivatives to bind to the transporter with altered specificity. In an attempt to visualize these differences, a homology model of *Cp*SiaT was build based on the *Pm*SiaT structure (PDB ID: 5NV9) [136]. When comparing the amino acids within 5 Å of the sialic acid binding site, the majority of the differences were identified close to the acetyl amino group at the C5 position and surrounding the glycerol chain at the C6 position of Neu5Ac. It is within these regions of sialic acid where modifications often occur [148].

3.3.4 Identifying sialic acid specificity in *Sa*SiaT

Growth of *E. coli* on Neu5Ac and Neu5Gc requires the catabolic enzymes which converts sialic acid to ammonia and fructose-6-phosphate [49,145,149]. Knowing this enabled evaluation of whether *Sa*SiaT was able to transport both Neu5Ac and Neu5Gc. In **Paper III** it is demonstrated how *Sa*SiaT

is able to rescue the growth of *E. coli* $\Delta nanT$ grown on either Neu5Ac or Neu5Gc as the sole carbon source [150]. Interestingly, the growth rate is faster when the media is supplemented with Neu5Gc (1.9 ± 0.05 x 10^{-3} rate/min) compared to Neu5Ac (1.3 ± 0.3 x 10^{-3} rate/min). In **Paper III** it is speculated how this difference in growth rate can be related to the difference in amino acids within the substrate binding site of SiaT.

Characterization of *Sa*SiaT

The increased growth rate for *Sa*SiaT grown on Neu5Gc compared to Neu5Ac raised the question regarding substrate specificity. To investigate the potential difference in substrate specificity compared to *Pm*SiaT, a combination of binding studies and structural comparison based on the *Pm*SiaT structure were performed (**Paper III**).

Both MST and ITC measurements of the affinity of *Sa*SiaT for the sialic acids, demonstrated higher affinity for Neu5Gc compared to Neu5Ac. With MST the K_d of *Sa*SiaT was determined to 39 ± 4 µM and 113 ± 6 µM for Neu5Gc and Neu5Ac, respectively. Comparing with the K_d of *Pm*SiaT (85 ± 2 µM and 58 ± 1 µM for Neu5Gc and Neu5Ac, respectively) (**Paper II**), *Sa*SiaT presents a higher affinity for Neu5Gc compared to Neu5Ac, which is the opposite relationship found in *Pm*SiaT. These binding studies are in line with the *in vivo* bacterial growth assay that demonstrated increased growth rate for *Sa*SiaT when grown on Neu5Gc compared to Neu5Ac.

The hypothesis that *Sa*SiaT has altered substrate specificity compared to *Pm*SiaT was further strengthened by comparing the K_M of [^{3}H]Neu5Ac transport in *Sa*SiaT (42 ± 9 µM) with *Pm*SiaT (16 ± 4 µM) using a proteoliposome assay (Figure 5, **Paper III**).

Substrate binding site comparison between *Pm*SiaT and a model of *Sa*SiaT

Is it possible to explain the difference in substrate specificity at a molecular level? What are the differences between the *Sa*SiaT and *Pm*SiaT amino acid sequences in the substrate binding site? In **Paper II** we concluded that the overall amino acid sequence in the substrate binding site is conserved among bacterial species (Figure 3, **Paper II**). A homology model of *Sa*SiaT based on the *Pm*SiaT structure (PDB ID: 5NV9) was constructed (**Paper III**). Three of the amino acids which differ between *Sa*SiaT and *Pm*SiaT (Asn83, Asn244, Tyr79 in *Sa*SiaT aligned with Gln82, Phe243, Phe78 in *Pm*SiaT) could potentially be responsible for the substrate specificity.

The difference between Neu5Ac and Neu5Gc is a hydroxyl group at the C11 position (Figure 1.3). Most likely, the difference in affinity between

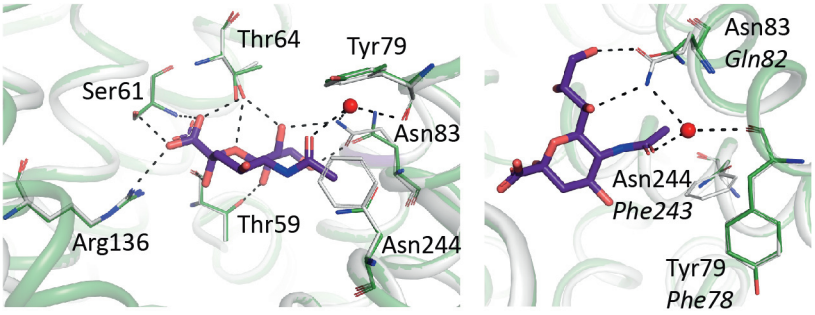

Figure 3.7: Superposition of the *Sa*SiaT homology model and the structure of *Pm*SiaT. *Sa*SiaT is colored in green, *Pm*SiaT (PDB ID: 5NV9) in grey and Neu5Ac (purple) is displayed as sticks colored by atom. A water molecule from *Pm*SiaT is displayed as a red sphere. Amino acids are labeled according to *Sa*SiaT with *Pm*SiaT labels in *italic*. Black dashed lines illustrate interactions between Neu5Ac and the corresponding amino acids. Figure adapted from [150].

*Sa*SiaT and *Pm*SiaT toward the two sialic acids is connected to the amino acid differences mentioned above. In **Paper III** we speculate about how the shorter side chain of Asn83 (*Sa*SiaT) compared to Gln82 (*Pm*SiaT) will free more space in the substrate binding site and give room for the additional hydroxyl group of Neu5Gc. Furthermore, the hydrogen bonds between Gln82 (*Pm*SiaT) and the two hydroxyl groups at C7 and C9 might still be possible, although with a longer bond length. It is also possible that the additional hydroxyl group at C11 could hydrogen bond with Asn244 (*Sa*SiaT) in a new interaction. The Tyr79 (*Sa*SiaT) creates a more hydrophilic environment towards the extracellular side, which might be important for discriminating between Neu5Ac and Neu5Gc. Previous studies have demonstrated that some bacteria can discriminate between different sialic acids [151–154]. It is also known that humans cannot synthesize Neu5Gc. However, humans can metabolically incorporate Neu5Gc from red meat and milk in their diet [151, 155–158]. In **Paper III** we demonstrate how the sialic acid transporter *Sa*SiaT has higher affinity towards the non-human synthesized Neu5Gc compared to Neu5Ac.

3.3.5 Summary

In chapter 3.3 the bacterial growth assay used in **Paper I** - **Paper III** has been described. All genes which were predicted to encode SiaT (presented in chapter 3.1) were able to rescue the growth of an *E. coli* $\Delta nanT$ strain when grown on minimal media supplemented with Neu5Ac as the sole carbon and nitrogen source. Furthermore, the assay was used in combination with sequence alignments and a homology model of *Cp*SiaT to speculate about substrate specificity between SiaTs (**Paper I**). A preference for Neu5Gc compared to Neu5Ac, in terms of growth rates of *Sa*SiaT, was further confirmed by binding studies (**Paper III**). The bacterial growth assay of SiaT is a helpful tool in exploring the sialic acid specificity in different bacterial organisms.

Chapter 4

MraY

The overall goal with the MraY project is to generate X-ray crystal structures of MraY in complex with potential inhibitors combined with biophysical characterization for the future development of antibacterial drugs. In **Paper V** the focus has been to structurally compare and analyze published MraY and GPT structures to identify key features and propose structural motives for future design of tunicamycin analogues which selectively targets MraY. In **Paper VI**, the interaction between MraY and tunicmaycin analogues has been biophysically characterized. The following chapter is divided into two sub-chapters describing the findings from **Paper V** and **Paper VI**.

4.1 MraY - a structural comparison

In 2013 the first crystal structure of the bacterial membrane-bound enzyme involved in peptidoglycan synthesis, MraY, was published [93]. MraY from the Gram negative thermophilic bacterium *A. aeolicus* (Aa) was crystallized in its apo conformation at a resolution of 3.3 Å. A few years later the apo AaMraY structure was followed by the structure of AaMraY in complex with the natural product inhibitor MD2, at a resolution of 2.95 Å [159]. The structure of MraY from the Gram positive pathogenic bacterium *Clostridium bolteae* (Cb) in complex with the natural product inhibitor tunicamycin (Tun) is the most recent structure of this membrane-bound enzyme and it was determined at 2.6 Å resolution in 2017 [94] within this project before I joined. As part of **Paper V**, structural similarities and differences between the published structures of MraY (Table 4.1) were analyzed.

Table 4.1: Structures of the bacterial membrane-bound enzyme MraY. The structures are listed in order of released to the Protein Data Bank (PDB).

Protein-Ligand	PDB ID	Resolution (Å)
AaMraY-apo [93]	4J72	3.3
AaMraY-MD2 [159]	5CKR	2.95
CbMraY-Tun [94]	5JNQ	2.6

4.1.1 Structure of MraY

MraY consists of ten transmembrane helices (TM1-TM10) with both the N- and C-terminus located on the periplasmic side (Figure 4.1A). The main differences between the overall structures of AaMraY and CbMraY are the additional periplasmic helix located between TM6 and TM7 as well as a more defined periplasmic β-hairpin in the AaMraY structure (Figure 2B, **Paper V**). In all crystal structures of MraY, a highly conserved glycine residue (Gly258 CbMraY, Gly293 AaMraY) breaks TM9 into two helical segments (TM9a and TM9b). This break results in TM9b being bent $\sim$50° relative to the membrane and protrudes $\sim$20 Å into the lipid bilayer away from the main structure. Helix 9c is a 11-residue helix located on the cytoplasmic side between TM9b and TM10 which is important for substrate selectivity due to the HHH-motif (His290, His291, His292 CbMraY) [160]. In the apo AaMraY structure helix 9c is not clearly visible.

MraY catalyzes the synthesis of C_{55}-PP-MurNAc-pentapeptide (lipid I) at the cytoplasmic side of the lipid bilayer. The active site is composed of the amphiphilic helix 9c, TM5-TM10 and loops C and D. A comparison of the active site of the three structures of MraY reveals a significant difference between the apo and the two ligand-bound complexes, where the apo AaMraY structure lacks a well-defined binding pocket.

The aspartic acid triad Asp93, Asp94 and Asp231 (CbMraY) is strictly conserved within the PNPT superfamily and important for the enzymatic activity of MraY [101]. The MraY catalysis is dependent on Mg^{2+} ions [90]. Both Asp93 and Asp94 (Asp115 and Asp116 in EcMraY) were proposed to coordinate the Mg^{2+} ion, based on sequence similarity. In the apo structure of AaMraY, Asp265 (Asp231 in CbMraY) coordinates a Mg^{2+} ion, whereas in the AaMraY-MD2 and CbMraY-Tun structures, no Mg^{2+} ion is observed. In the CbMraY-Tun structure, the ligand is interacting with Asp231, whereas in the AaMraY-MD2 structure there is no interaction between Asp265 and

MD2. Tunicamycin and Mg^{2+} are not able to interact with the enzyme simultaneously [94], furthermore MD2 does not require Mg^{2+} for binding MraY [159].

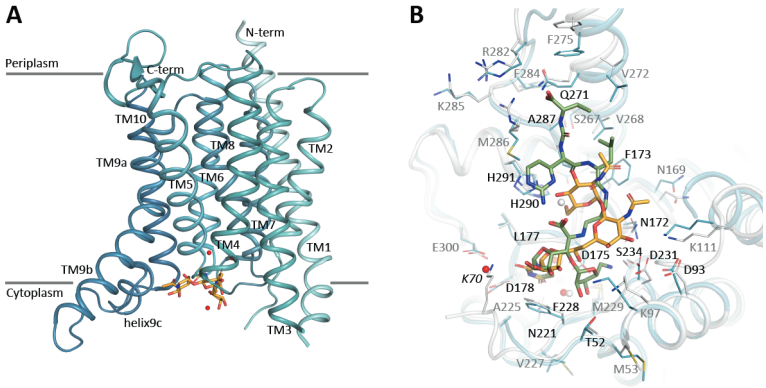

Figure 4.1: Structure of MraY. A) Overall structure of MraY (colored from N-terminal (light blue) to C-terminal (dark blue) (PDB ID: 5JNQ) in complex with tunicamycin (orange) colored by atom. The lipid membrane is indicated by grey lines and transmembrane helices (TM) are labeled, water molecules are colored in red. B) A close-up view of the active site. The CbMraY (blue)- Tun (orange) complex is superimposed with the AaMraY (white)- MD2 (green) complex (PDB ID: 5CKR). Amino acids within 5 Å of the ligands are shown as sticks and numbered according to CbMraY (except for *K70* which only is modeled in AaMraY). Figure modified from [85].

4.1.2 Interactions between MraY and the natural product inhibitors

Both the tunicamycin and MD2 ligands share a common uridine motif (Figure 1.7) with the soluble substrate UDP-MurNAc-pentapeptide (Figure 1.8). The uracil part of the two inhibitors are partially overlapping in the MraY-complex structures; amino acids Gly176, Asn221 and Phe228 in CbMraY (Gly194, Asn255 and Phe262 in AaMraY) create a small pocket for the uracil motif. Coordination between the uracil ring and the enzyme occurs via Asn221 and Asp178 as well as backbone interactions with Leu177 in CbMraY (Asn255, Asp196 and Leu195 in AaMraY). The phenylalanine Phe228 in CbMraY (Phe262 in AaMraY) is conserved throughout the PNPT superfamily and in MraY it interacts with the uracil base via a π-π stacking (Figure 4.1B).

The three conserved histidines in MraY, denoted as the HHH-motif, are located within the amphiphilic helix 9c. These histidine residues, His290, His291 and His292 in CbMraY (His324, His325, His326 in AaMraY), form interactions with both the tunicamycin and MD2 ligand. The GlcNAc ring of tunicamycin interacts with His291 in CbMraY and the guanidine motif of the peptidic moiety of MD2 interacts with His325 in AaMraY. Furthermore, His324 and His325 of AaMraY interact with a carbonyl of the peptidic moiety of MD2.

In contrast to tunicamycin, MD2 is split into two arms at the 5´-hydroxyl group: the 5´-aminoribosyl and the peptidic moiety (Figure 1.7). In AaMraY the 5´-aminoribosyl moiety of MD2 stretches toward loop D and interacts with several amino acids: Thr75, Asn190, Asp193, Gly264 and Ser268 in AaMraY (Thr52, Asn172, Asp175, Gly230 and Ser234 in CbMraY), these glycine and asparagine residues are conserved throughout the PNPT superfamily. Furthermore, the 5´-hydroxyl group of tunicamycin interacts with the conserved Asp175 in CbMraY, an important amino acid which potentially is involved in coordinating the metal ion. The backbone of Ala321 and Gln305 in AaMraY (Ala287 and Gln271 in CbMraY) both interact with the peptidic moiety of MD2 and contribute to binding affinity.

Tunicamycin has a fatty acyl chain which can vary in length. This fatty acyl chain is not observed in the electron density of the CbMraY-Tun complex structure. The Asn172 amino acid interacts with the fatty acid amide, hence indicates the direction of the fatty acyl tail up along TM5 into the membrane bilayer.

4.1.3 Structural comparison of MraY and GPT

In 2016, the first structure of human GPT was released in the PDB (PDB ID: 5LEV). The apo GPT structure was followed by three structures of GPT in complex with tunicamycin [82, 161]. Most recently, an additional apo structure and a structure of GPT in complex with the substrate UDP-GlcNAc were released in the PDB [161] (Table 4.2).

Table 4.2: Structures of the eukaryotic membrane-bound enzyme GPT. The structures are listed in order of released in the Protein Data Bank (PDB).

Protein-Ligand	PDB ID	Resolution (Å)	Mutation
GPT-apo [161]	5LEV	3.2	V264G
GPT-Tun [82]	6BW5	3.1	P129 variant
GPT-Tun [82]	6BW6	2.95	H129 variant
GPT-Tun [161]	5O5E	3.4	V264G
GPT-apo [161]	6FM9	3.6	Wild type
GPT-UDP-GlcNAc [161]	6FWZ	3.1	V264G

Both the bacterial MraY and the human GPT have ten transmembrane spanning helices and the main composition and location of the active sites are overall very similar. However, GPT has an extra cytoplasmic $\beta\alpha\beta\beta$ motif located between TM9b and helix 9c, which is not present in MraY (Figure 4.2A). The presence of this extra cytoplasmic domain in GPT positions the amphiphilic helix 9c further away from the active site compared to in MraY. Both helix 9c and the $\beta\alpha\beta\beta$ motif forms one side of the active site in GPT. The active site of GPT is narrower than in MraY, due to the loop between TM9b and the $\beta\alpha\beta\beta$ motif which protrudes into the active site of GPT.

4.1.4 Structural alignment of MraY and GPT

The activity of GPT is dependent on the presence of Mg^{2+} [162] as is the activity of MraY [90]. In the apo AaMraY structure, a Mg^{2+} ion is bound to an aspartic acid residue (Asp265 in AaMraY, Asp231 in CbMraY and Asp252 in GPT), which is conserved within the PNPT superfamily.

In the GPT-UDP-GlcNAc structure (not discussed in **Paper V**) the phosphate groups of the substrate are coordinating the Mg^{2+} ion together with Asn185 (Asn172 CbMraY) and Asp252 (Asp231 CbMraY) [161]. As Bouhss and co-workers proposed, Asp93 (located in TM3 of CbMraY, Asp115

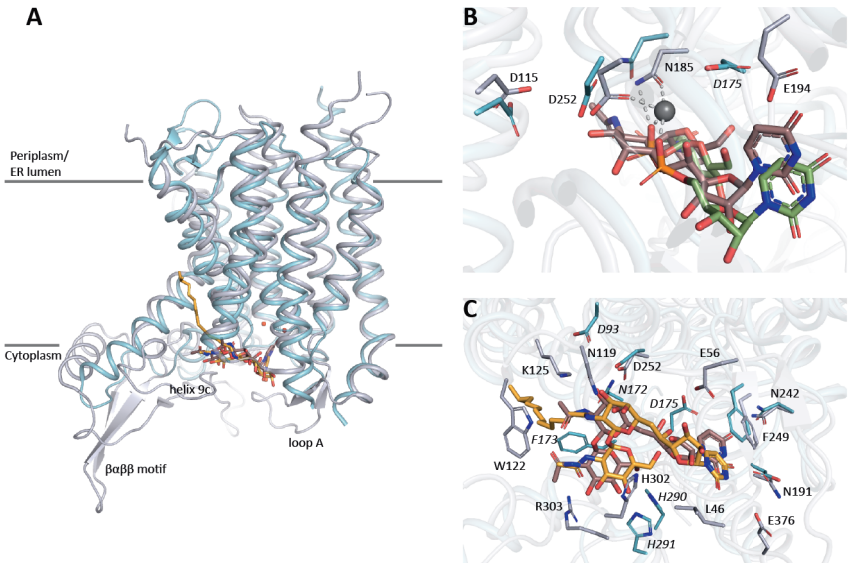

Figure 4.2: Structural comparison between GPT and MraY. A) GPT (grey)-Tun (orange) complex (PDB ID: 6BW5) superimposed with the CbMraY (blue)-Tun (violet) complex (PDB ID: 5JNQ). The lipid membrane is indicated with grey lines and water molecules are colored in red for CbMraY. B) Active site structural comparison between the GPT (grey)- GlcNAc (green) complex (PDB ID: 6FWZ) and CbMraY (blue)-Tun (violet) complex (PDB ID: 5JNQ). Amino acids important for Mg^{2+} binding are illustrated as sticks. C) Structural comparison of the active site of the GPT (grey)-Tun (orange) complex (PDB ID: 6BW5) superimposed with the CbMraY (blue)-Tun (violet) complex (PDB ID: 5JNQ). Amino acids that interact with the ligands are shown as sticks, labeled according to GPT and in *italic* for CbMraY.

in GPT) could potentially be involved in the deprotonation of the lipid substrate [104]. This aspartic acid residue is located approximately 8 Å away from the Mg^{2+} site in the GPT structure towards the proposed entry pathway of the lipid substrate, between helices TM4 and TM5 of MraY. Furthermore, Asp175 in CbMraY is positioned on the opposite side, not far away from the Mg^{2+} site ($\sim$ 4.5 Å) and could potentially also be involved in the coordination of the Mg^{2+} ion in MraY. In a sequence alignment between GPT and MraY, Asp175 aligns with Ala188 (CbMraY). However, in the structural alignment, Glu194 of GPT occupies almost the same space as Asp175 in CbMraY. Most likely Glu194 takes on the same role as Asp175 in CbMraY (Figure 4.2B).

4.1.5 Interactions with tunicamycin

Overall, the binding mode of tunicamycin is highly similar for both CbMraY (PDB ID: 5JNQ) and GPT (PDB ID: 6BW5) (Figure 4.2C). The uracil pocket is better defined in the GPT structure and the extra cytoplasmic domain creates a more enclosed active site in GPT compared to MraY. The uracil motif of tunicamycin interacts with Asn191 (Asp178 CbMraY) and the backbone of Leu46 (Lys70 AaMraY, Lys47 not modeled in CbMraY) in GPT. Furthermore, the π-π stacking between the uracil ring and the aromatic phenylalanine Phe249 of GPT (Phe228 CbMraY) is also observed in the GPT-UDP-GlcNAc structure, strengthening the importance of this interaction for the enzymatic activity. An additional interaction between the uracil motif of tunicamycin and the conserved asparagine Asn242 (Asn221 CbMraY) occurs via a water molecule in GPT (PDB ID: 5O5E), whereas in CbMraY this is a direct interaction.

Catalytic activity of GPT has previously been described to depend on a cluster of positively charged amino acids (Arg301, His302 and Arg303) in the cytoplasmic loop E [163]. More specifically, the Arg303Lys/Asn/His mutations inactivated the GPT enzyme, demonstrating the importance of this residue for stabilization of the substrate. In the GPT structures the 3´´-hydroxyl of the GlcNAc moiety of tunicamycin and UDP-GlcNAc interacts with Arg303 which is part of the $\beta\alpha\beta\beta$ motif (not present in MraY). The three conserved amino acids of GPT are not part of the amphiphilic helix 9c, in contrast to the conserved HHH-motif in MraY; but the location of these residues partially overlaps in space and presumably they share the same role in the different enzymes.

The length of the fatty acyl chain of tunicamycin varies. The direction of the fatty acyl tail is indicated by mutational studies of Asp172 in CbMraY (Asp185 in GPT) [94] and correlates well with the direction of the tail as resolved in the GPT structures where the amide of the fatty acyl chain interacts with Asp185 and positions the tail into a tunnel between TM5 and TM9 towards the lipid bilayer. Most likely the polyprenyl phosphate acceptor substrate enters the active site through the tunnel indicated by the tunicamycin fatty acid tail. Furthermore, the direction of the pentapeptide of the soluble MraY substrate is most likely located towards helix 9c and the loop connecting to TM10. The soluble substrates of MraY and GPT differs in size ($\sim$1200 Da vs $\sim$600 Da, respectively) due to the GlcNAc/MurNAc motif and a pentapeptide motif at the 3´´-hydroxy position of MurNAc in the case of the MraY substrate. It is at this position the interactions between tunicamycin and the two enzymes differ.

4.1.6 Compound selectivity based on structural differences

The natural products tunicamycin and MD2 both inhibit the enzymatic activity of MraY. Only tunicamycin inhibits GPT, whereas MD2 does not inhibit the GPT activity. The lack of inhibitory property of MD2 towards GPT is of course beneficial for avoiding potential cross-reactivity between the two enzymes when targeting MraY. Our structural alignment of the AaMraY-MD2 complex with GPT reveals a major clash between the peptidic moiety of MD2 and the additional cytosolic TM9b-TM10 domain of GPT. The soluble UDP-GlcNAc/MurNAc substrates for GPT and MraY differ in their sugar moieties. A tunicamycin analog (Tun-MurNAc) was recently synthesized which demonstrated reduced inhibition of GPT compared to tunicamycin (IC_{50} of 15 μM and 9 nM, respectively) [82]. We could rationalize this by structural alignment of the Tun-MurNAc compound with Tun as bound in the active site of GPT, which revealed a clash between the MurNAc moiety and the same TM9b-TM10 region of GPT.

4.1.7 Summary

In **Paper V** we review the current published crystal structures of MraY and GPT, two members of the PNPT superfamily of enzymes and their mode of interaction with the natural product inhibitors MD2 and tunicamycin. Furthermore, we speculate on possible routes to design a novel antibacterial agent. The structural data reviewed in **Paper V** could aid in the design of nucleoside compounds which selectively inhibit MraY.

4.2 Inhibition of MraY by modified tunicamycins

In **Paper VI**, the potency against MraY of a set of tunicamycin (Tun) analogues are evaluated by determining IC_{50} values and binding affinity. The observed trend in inhibition of MraY for the different Tun analogues is rationalized by analyzing structural motifs in context of the protein structures of MraY and GPT.

4.2.1 Modified tunicamycin analogues

Tunicamycins, streptovirudins and corynetoxins all belong to the diverse and naturally occurring tunicamycin family. The chemical structure of analogues from these groups of uridine nucleoside antibiotics differ in their length, stereochemistry and degree of unsaturation or hydroxylation of the *N*-linked acyl group [164, 165]. Streptovirudins have also been isolated with a dihydrouracil group instead of the common uridyl group [166, 167]. The end of the fatty acyl chain of tunicamycin exists in different configurations resulting in either an iso ($CH_3CH(CH_3)$-) [168], anteiso ($CH_3CH_2CH(CH_3)$-) or unbranched (CH_3CH_2-) conformation (Figure 4.3) [164]. Compared to the uridyl motif of natural products [169, 170], the 5,6-dihydrouridyl motif is a less common chemophore [171].

The natural product inhibitor tunicamycin is, as mentioned in chapter 1.5.4, toxic to eukaryotes. For a set of Tun analogues, an altered toxicity profile was reported compared to a mixture of naturally occurring Tun (fatty acyl chain lengths C14-C17) [81,171,172]. Selective hydrogenation of the *N*-acyl double bond results in the TunR1 analogue and in combination with hydrogenation of the uridyl double bond results in the TunR2 analogue [81, 171] (Figure 4.3). Mild basic treatment of the reduced 5,6-dihydrouridyl group results in amide ring opening and the Tun analogue TunR3 [171] (Figure 4.3).

In quinovosamycin (QVM), identified in *Streptomyces niger*, the GlcNAc motif at the α,β-1´´, 11´-linked position is replaced with a QuiNAc motif (6-deoxyhexosamine sugar) [172] (Figure 4.3). Both Tun and QVM are able to inhibit *N*-linked glycosylation in the yeast *Pichia pastoris* [172].

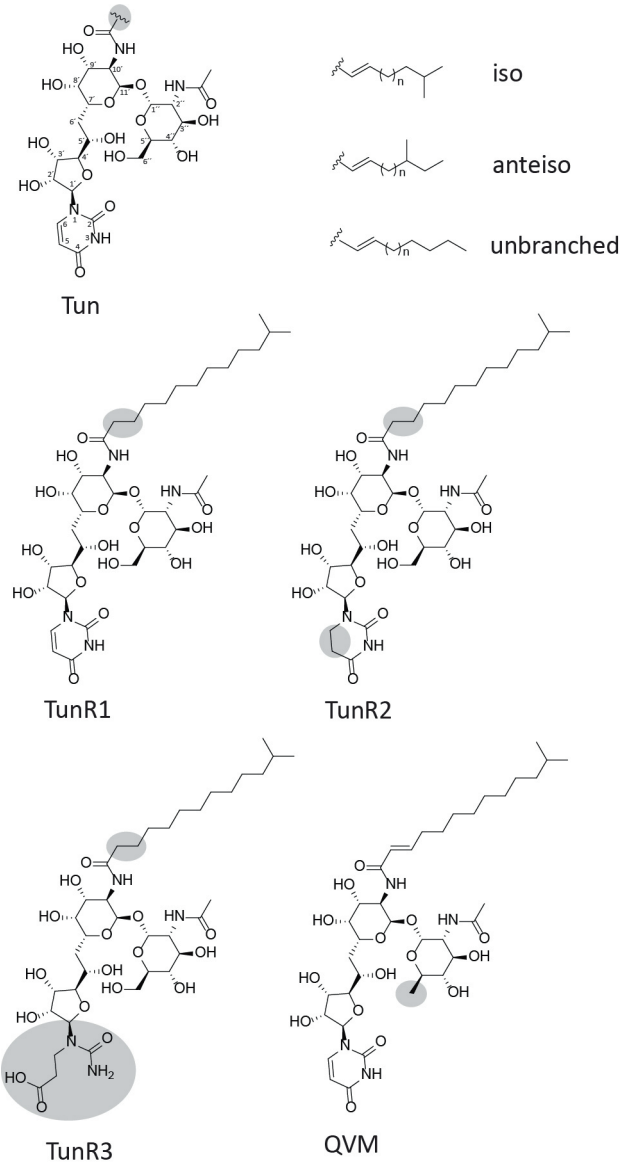

Figure 4.3: Tunicamycin analogues. The core structure of Tun and the different types of branching are displayed. The fatty acyl chain can vary in length (n = 7-10). The different Tun analogues investigated (TunR1, TunR2, TunR3 and QVM) are illustrated and the differences are highlighted with grey circles. If not defined by the name, each modified Tun analogue contains the *N*-acyl variants C-14, C-15, C-16 and C-17.

4.2.2 Rationalized inhibition of tunicamycin analogues

It has recently been demonstrated that lipid modifications of tunicamycin affect binding, where an additional introduced relatively short and unbranched fatty acyl chain inhibits MraY but not GPT [161]. In **Paper VI** we investigated the length, branching and saturation/unsaturation of the fatty acyl chain, and modifications of the uracil and GlcNAc ring of Tun. The level of inhibition (IC_{50}) was determined with a FRET-based assay [94, 99].

Tunicamycin is a mixture of homologues with variations in branching and the length of the fatty acyl chain (described in chapter 4.2.1). The inhibitory effect for different versions of Tun (15:1-17:1) was determined, demonstrating a higher potency for Tun(iso) with a shorter chain length. This is in contrast with previously observed inhibition with streptovirudins [165, 166], where a shorter chain length was more potent. Furthermore, comparing Tun(anteiso) with Tun(iso) of the same chain length, Tun(anteiso) demonstrated significantly reduced inhibition.

The hydrophobic tunnel that binds the fatty acyl chain is more enclosed in GPT compared to in MraY. Consequently, GPT is more sensitive to changes in the fatty acyl chain. Furthermore, the location of the branching also had an impact on the inhibitory effect towards MraY, reflecting the importance of maximal interactions for a potent inhibitor.

The TunR1 analogue has a saturated C2-C3 bond in the fatty acyl chain. Compered to Tun, our results demonstrate that the saturated C2-C3 bond in the fatty acyl chain of TunR1 has an unaltered inhibitory effect towards MraY. This would suggest that the hydrophobic tunnel of MraY is more adaptable for a lipid tail which resembles the lipid substrate of GPT. The lipid substrate of GPT, dolichyl phosphate, has a saturated C2-C3 bond whereas the lipid substrate of MraY, undecaprenyl phosphate, contains a C2-C3 double bond. Since the tunnel that creates the lipid binding pocket in GPT is more well defined and contains a tryptophan residue (Trp122) which closes up on the lipid [82], it is possible that alterations in the lipid chain has a larger impact on inhibition towards GPT in contrast to MraY. In a recent study [81], the effect of Tun and TunR1 inhibition on *B. subtilus* were similar. It was also discovered that TunR1 had a reduced inhibitory effect towards *Saccharomyces cerevisiae* [81].

Tunicamycin and the soluble substrates of MraY and GPT all contain a uracil group. The uracil group of tunicamycin is positioned in a well-defined pocket in both MraY and GPT (Figure 4.1B and Figure 4.2C) [82, 94, 161]. A highly conserved phenylalanine (Phe228 in CbMraY and Phe249 in GPT) interacts with the uracil ring through $\pi-\pi$ stacking. Hypothetically, reducing the double bond of the uracil ring would result in decreased inhibitory effect

since the uracil ring no longer would be planar and the $\pi - \pi$ stacking interaction would be influenced (Figure 4.4A). We investigated the effect of a reduced uracil ring by determining the IC_{50} value for the Tun analogue TunR2 (IC_{50} of TunR2 was ∼1.5 times the IC_{50} of Tun). Thus, the reduced uracil ring did not have any major effect on the inhibition of MraY, a similar observation has been noticed for streptovirudins [165]. It has previously been shown that reduced inhibition of GPT has been obtained with TunR2 [81, 171]. It is possible that the difference in inhibitory effect between MraY and GPT can be explained by the positioning of the uracil ring in the binding pocket, where the uracil ring in the CbMraY-Tun complex is positioned deeper in the pocket as opposed to the uracil ring in the GPT-Tun complex structure. As a consequence structurally aligned amino acids (Asp178 in CbMraY and Asn191 in GPT) bind to different parts of the uracil ring (the aminogroup and C4 carboxyl oxygen, respectively). A significantly reduced inhibitory effect is detected in the activity measurements for TunR3, which has an open uracil ring structure. Most likely TunR3 is not able to maintain the interactions important for binding within the uracil pocket of MraY. A previous study has indicated similar results where TunR3 was unable to inhibit *N*-glycosylation [173].

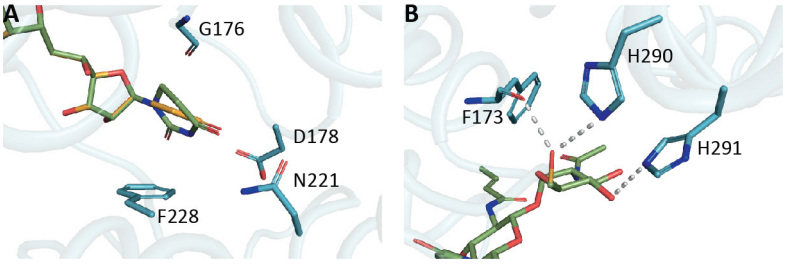

Figure 4.4: Structural comparison of tunicamycin analogues in the active site of CbMraY. A) TunR2 (green) is superposed with Tun (orange) as bound in the active site of the CbMraY (PDB ID: 5JNQ). The planar uracil ring of Tun interacts with Phe228 through $\pi - \pi$ stacking. B) QVM (green) is superposed with Tun (orange) as bound in the active site of the CbMraY (PDB ID: 5JNQ). Interactions between the 6´´-hydroxyl group of Tun and CbMraY are illustrated by dotted lines. In addition the neighbouring interaction between the 4´´-hydroxyl of the GlcNAc motif and CbMraY is also displayed. QVM lacks the 6´´-hydroxyl group.

In activity measurements, QVM16:1(iso) has demonstrated a dramatically reduced inhibitory effect, with an IC_{50} value ∼150 times lower than the IC_{50} for Tun16:1(iso). In the CbMraY structure, the 6´´-hydroxyl interacts with the backbone of Phe173 and with the side chain of His290 that is

part of the HHH-motif (Figure 4.4B). In the case of QVM, these interactions are likely disrupted leading to a decrease in inhibitory effect. QVM has previously demonstrated similar inhibitory effect as Tun towards *S. cerevisiae* GPT [172]. There is no interaction between the 6´´-hydroxyl group of Tun and GPT in the structure complex. Hence, the 6´´-hydroxyl group of Tun has a bigger impact on inhibition of MraY compared to GPT.

In **Paper V** we highlight the GlcNAc motif of tunicamycin as a plausible site for modifications to selectively target MraY. Importantly, the amino acids surrounding the GlcNAc motif of tunicamycin are different between the two enzymes, where GPT has the extra cytoplasmic $\beta\alpha\beta\beta$ motif which makes the active site narrower and where the HHH-motif of MraY, which interacts with the 6´´-hydroxyl of tunicamycin, is absent in GPT.

Lastly, we aimed to determine the binding affinity of Tun and its analogues to CbMraY. In **Paper VI** we have determined the binding affinity of commercially available Tun to CbMraY using ITC (K_d ~200 nM) in the absence of $MgCl_2$. The determined K_d for Tun is in about the same range as previously determined K_d between Tun and AaMraY (37 nM ± 1 nM) [82].

4.2.3 Alternative methods for determining binding

Identifying potent inhibitors is an important step in order to move forward in drug development. In **Paper VI** (discussed in chapter 4.2.2) we used a FRET-based activity assay to determine the inhibitory effect of Tun and Tun analogues. Furthermore, ITC was used to determine the binding affinity between CbMraY and commercially available Tun. The following section speculates about why some methods are less suitable for studying the binding between small molecules and MraY based on unpublished data.

Over the years, ITC has been used extensively to measure binding affinity. This label-free method measures the change in heat upon complex formation of ligand and macromolecule. When a macromolecule interacts with a ligand, non-covalent bonds are formed and redistributed causing heat to be either absorbed or released. An ITC experiment requires purified and stable protein in reasonably high amounts. Considering the difficulties in producing high levels of purified MraY, alternative methods that do not require as high amounts of protein were initially investigated. One of those methods is MST. The principal behind MST is described in chapter 2.5.3. Several MST experiments were performed in attempts to determine the binding affinity between CbMraY and Tun analogues. Unfortunately, it was not possible to obtain reliable binding curves with this technique (unpublished).

Why was MST not suitable for studying the binding for MraY? During MST experiments the sample is heated to induce a diffusion gradient, which

is detected by measuring the change in fluorescence. Consequently, one of the binding partners needs to be labeled with a fluorophore alternatively it is possible to take advantage of intrinsic fluorescence. To be able to adopt an approach for MST measurements to investigate many different small ligands and their binding affinity to MraY, it is preferred to label the protein. Different labeling approaches were evaluated for the binding between MraY and Tun analogues. Cysteine-labeled MraY resulted in unstable protein, an effect which could originate from the location of the cysteine residues in the transmembrane helices of MraY. In contrast to cysteine-labeled MraY, lysine-labeled MraY was stable and could be used in binding studies. However, there are lysine residues in close proximity to the active site which potentially could affect the binding, although the lysine-labeled MraY was active in the FRET assay. The binding affinity between lysine-labeled CbMraY and commercial Tun could be determined to $K_d = 86 \pm 17$ μM in the presence of 50 mM $MgCl_2$ (Figure 4.5A), a value much higher than obtained with ITC measurements (**Paper VI**).

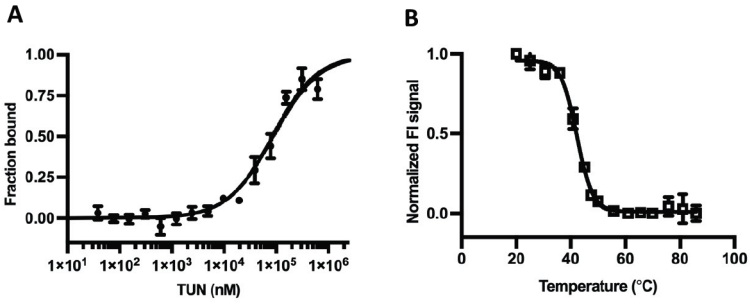

Figure 4.5: MST and TFSEC curves for CbMraY. A) Binding curve measured with MST for commercial Tun binding to CbMraY in the presence of 50 mM $MgCl_2$. The experiment was carried out in duplicates of triplicates at 24°C with an excitation power of 20% and a MST power of 80%. The K_d value was determined from the binding curve which was fitted to a quadratic solution of the law of mass action. B) TFSEC melting curve of crude-detergent solubilized CbMraY, heat incubated at different temperatures, filtered, incubated with an FSEC-probe and analyzed with FSEC. Measurements were performed in triplicates and the data was fitted to a sigmoidal dose-response curve.

In addition His-tagged labeled MraY was also tested in binding studies. The His-tagged label was thought to be advantageous since the His-tag is located far away from the cytoplasmic active site. These measurements gave similar but not as clear results as with lysine-labeled CbMraY. Lastly, label-free MST experiments were evaluated by detecting intrinsic fluorescence.

A drawback with label-free MST experiments is the potential fluorescence interference from conjugated ring-systems in the small molecules and potentially from chemicals in the buffer. Indeed the small molecules investigated did possess intrinsic fluorescence which interfered with binding experiments.

It is known from ITC measurements that Tun binds to MraY with high affinity in the absence of $MgCl_2$ (**Paper VI**) [82]. Binding between CbMraY and Tun analogues was detected in the presence of high concentrations of $MgCl_2$ at high MST laser powers. In those cases where low-affinity binding was detected, we speculate that it could be a result of the protein system being affected by the high laser power, buffer composition and/or solubility issues of Tun.

An alternative approach to investigate ligand binding to a protein is to evaluate the shift in melting temperature (T_m). The melting temperature (T_m) is a measure of the thermostability of a protein [174] and in general a ligand binding to a protein stabilizes the protein giving rise to a higher T_m. We attempted to investigate Tun binding to MraY by measuring the shift in T_m between unbound MraY and MraY bound to Tun in a thermostability fluorescence-detection size-exclusion chromatography-based assay (TFSEC). The principal behind the experiment is to solubilize the protein from the membrane and add the ligand prior heat incubation at different temperatures. The level of denatured protein was quantified through analytical FSEC where the fluorescent signal comes from a modified in-house fluorescent probe based on the P3NTA probe [92], which binds to a His-tag. A thermal transition ($T_m = 42 \pm 0.3$°C) was observed (Figure 4.5B) for CbMraY. However, we were not able to detect a significant thermal shift for CbMraY in the presence of Tun under our experimental conditions.

4.2.4 Summary

In **Paper VI** the effect of inhibition of Tun and Tun analogues towards CbMraY was investigated in a FRET-based activity assay. Furthermore, the binding affinity between Tun and CbMraY was determined with ITC. We have found that reducing the *N*-acyl double bond of the fatty acyl chain did not have a significant effect on the inhibition of MraY. In contrast, both the length of the fatty acyl chain and branching of Tun effect inhibition of MraY. In addition, the 6´´-hydroxyl group of the GlcNAc ring is an important motif for inhibition of MraY. Most likely, this motif does not contribute to binding in GPT. Modifications of the uracil ring also has an impact on inhibition and importantly our results demonstrate how the TunR1 and TunR2 analogues, which possess reduced eukaryotic toxicity profiles, are potent inhibitors of MraY.

Chapter 5

Concluding remarks

Antibiotic resistance is a naturally occurring and unavoidable process in bacteria. However, society has the responsibility to fight antibiotic resistance and slow down the process. The work presented in this thesis structurally and functionally investigate two types of membrane proteins, considered as antibacterial drug targets. The difficulties working with membrane proteins aimed for characterization and structure determination can be many. Some of these bottlenecks are brought to light in the presented papers.

The transport of sialic acid over the cytoplasmic membrane in bacteria has been studied and is presented in four papers within this thesis (**Paper I** - **Paper IV**). **Paper IV** describes the route from expression to low-resolution diffraction of the sialic acid TRAP transporter from *P. multocida*. A transporter type which has still not been structurally determined and is of high relevance for the development of inhibitors towards sialic acid transport due to its absence in eukaryotic species.

Paper I describes the process from membrane protein production to crystallization of the sialic acid transporter, SiaT. By adopting an orthologues approach combined with taking advantage of a fluorescent tag for detection, the work funneled down to crystals of *P. mirabilis* SiaT, which diffracted beyond 3 Å in resolution. In addition, the SiaT orthologues were characterized in an *in vivo* growth assay, where they were able to rescue the growth of an *E. coli* strain lacking the endogenous sialic acid transporter NanT by transporting the sialic acid, Neu5Ac, used as the sole carbon and nitrogen source in the assay.

In **Paper II** the first structure of a sialic acid transporter, SiaT from the uropathogen *P. mirabilis*, is reported at 1.95 Å resolution in an outward open conformation with both the substrate Neu5Ac and two sodium ions bound. The structure reveals a novel sodium binding site, Na3, suggested to stabilize the outward facing conformation and enhance substrate binding affinity. SiaT belongs to the SSS family of membrane protein transporters, which take

advantage of an alternating access mechanism in the co-transport of sialic acid and sodium ions over the lipid bilayer. To get a better understanding of the alternating access mechanism of sialic acid transport it would be advantageous to capture the protein in additional conformations as part of the transport cycle. With the present knowledge of the structure of SiaT, the work aimed at finding inhibitors to block sialic acid transport could be explored in more detail. Recently, with the structure of *Pm*SiaT it has been possible to model the sodium-dependent glucose transporter (SGLT) in an outward-facing conformation to understand inhibitor binding of human SGLTs [175].

In **Paper I** and **Paper III**, the substrate specificity of SiaT was explored through functional characterization and structural comparison that revealed an altered substrate specificity for *S. aureus* SiaT compared to *P. mirabilis* SiaT.

Bacterial peptidoglycan biosynthesis is an essential process for bacterial survival. In **Paper V** and **Paper VI** the membrane bound enzyme MraY which is part of this process has been structurally and functionally investigated. Differences between MraY and the eukaryotic homologue GPT have been highlighted and the effect of inhibition of MraY has been investigated for a set of tunicamycin analogues with the aim to selectively target MraY. For future studies it would be interesting to explore how these tunicamycin analogues specifically affect the activity of GPT.

With the work presented in this thesis I hope to shed some more light on one of our biggest threats towards world health, antibiotic resistance. Through the first determined structure of a sialic acid transporter it is now possible to further explore this protein as a novel antibiotic drug target. In addition, an alternative route for designing antibacterial drugs by modifying natural product inhibitors has been investigated as a possibility to selectively block the biosynthesis of peptidoglycan in bacteria.

Bibliography

[1] K. Duan and A. Kumar, "Understanding Bacterial Antibiotic Resistance and Pathogenicity Through Investigating Bacterial Membrane Proteins," *The Journal of Membrane Biology*, vol. 251, no. 1, pp. 1–3, 2018.

[2] S. B. Levy and B. Marshall, "Antibacterial resistance worldwide: causes, challenges and responses," *Nature Medicine*, vol. 10, no. 12, pp. S122–S129, 2004.

[3] J. Davies and D. Davies, "Origins and evolution of antibiotic resistance," *Microbiol Mol Biol Rev*, vol. 74, no. 3, pp. 417–433, 2010.

[4] J. M. Munita and C. A. Arias, "Mechanisms of Antibiotic Resistance," *Microbiol Spectr*, vol. 4, no. 2, pp. 1–37, 2016.

[5] H. Nikaido, "Multidrug resistance in bacteria," *Annu Rev Biochem*, vol. 78, pp. 119–146, 2009.

[6] G. von Heijne, "The membrane protein universe: what's out there and why bother?," *J Intern Med*, vol. 261, no. 6, pp. 543–557, 2007.

[7] K. Lundstrom, "Structural genomics: the ultimate approach for rational drug design," *Mol Biotechnol*, vol. 34, no. 2, pp. 205–212, 2006.

[8] G. von Heijne, "The distribution of positively charged residues in bacterial inner membrane proteins correlates with the trans-membrane topology," *Embo Journal*, vol. 5, no. 11, pp. 3021–3027, 1986.

[9] G. von Heijne, "Control of topology and mode of assembly of a polytopic membrane protein by positively charged residues," *Nature*, vol. 341, no. 6241, pp. 456–458, 1989.

[10] E. Wallin and G. von Heijne, "Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms," *Protein Science : A Publication of the Protein Society*, vol. 7, no. 4, pp. 1029–1038, 1998.

[11] D. Drew and O. Boudker, "Shared Molecular Mechanisms of Membrane Transporters," *Annu Rev Biochem*, vol. 85, pp. 543–572, 2016.

[12] O. Jardetzky, "Simple Allosteric Model for Membrane Pumps," *Nature*, vol. 211, no. 5052, pp. 969–970, 1966.

[13] L. R. Forrest and G. Rudnick, "The rocking bundle: a mechanism for ion-coupled solute flux by symmetrical transporters," *Physiology (Bethesda)*, vol. 24, pp. 377–386, 2009.

[14] A. Yamashita, S. K. Singh, T. Kawate, Y. Jin, and E. Gouaux, "Crystal structure of a bacterial homologue of Na^+/Cl^--dependent neurotransmitter transporters," *Nature*, vol. 437, no. 7056, pp. 215–223, 2005.

[15] C. Perez and C. Ziegler, "Mechanistic aspects of sodium-binding sites in LeuT-like fold symporters," *Biol Chem*, vol. 394, no. 5, pp. 641–648, 2013.

[16] S. Weyand, T. Shimamura, O. Beckstein, M. S. P. Sansom, S. Iwata, P. J. F. Henderson, and A. D. Cameron, "The alternating access mechanism of transport as observed in the sodium-hydantoin transporter Mhp1," *Journal of Synchrotron Radiation*, vol. 18, no. 1, pp. 20–23, 2011.

[17] H. Krishnamurthy and E. Gouaux, "X-ray structures of LeuT in substrate-free outward-open and apo inward-open states," *Nature*, vol. 481, no. 7382, pp. 469–474, 2012.

[18] C. Perez, C. Koshy, O. Yildiz, and C. Ziegler, "Alternating-access mechanism in conformationally asymmetric trimers of the betaine transporter BetP," *Nature*, vol. 490, no. 7418, pp. 126–130, 2012.

[19] L. Malinauskaite, M. Quick, L. Reinhard, J. A. Lyons, H. Yano, J. A. Javitch, and P. Nissen, "A mechanism for intracellular release of Na^+ by neurotransmitter/sodium symporters," *Nature structural molecular biology*, vol. 21, no. 11, pp. 1006–1012, 2014.

[20] A. Penmatsa, K. H. Wang, and E. Gouaux, "X-ray structure of dopamine transporter elucidates antidepressant mechanism," *Nature*, vol. 503, no. 7474, pp. 85–90, 2013.

[21] S. Faham, A. Watanabe, G. M. Besserer, D. Cascio, A. Specht, B. A. Hirayama, E. M. Wright, and J. Abramson, "The crystal structure of a sodium galactose transporter reveals mechanistic insights into Na^+/sugar symport," *Science*, vol. 321, no. 5890, pp. 810–814, 2008.

[22] A. Watanabe, S. Choe, V. Chaptal, J. M. Rosenberg, E. M. Wright, M. Grabe, and J. Abramson, "The mechanism of sodium and substrate release from the binding pocket of vSGLT," *Nature*, vol. 468, no. 7326, pp. 988–991, 2010.

[23] S. Weyand, T. Shimamura, S. Yajima, S. Suzuki, O. Mirza, K. Krusong, E. P. Carpenter, N. G. Rutherford, J. M. Hadden, J. O'Reilly, P. Ma, M. Saidijam, S. G. Patching, R. J. Hope, H. T. Norbertczak, P. C. Roach, S. Iwata, P. J. Henderson, and A. D. Cameron, "Structure and molecular mechanism of a nucleobase-cation-symport-1 family transporter," *Science*, vol. 322, no. 5902, pp. 709–713, 2008.

[24] M. Farwick, R. M. Siewe, and R. Kramer, "Glycine betaine uptake after hyperosmotic shift in *Corynebacterium glutamicum*," *J Bacteriol*, vol. 177, no. 16, pp. 4690–4695, 1995.

[25] K. Khafizov, C. Perez, C. Koshy, M. Quick, K. Fendler, C. Ziegler, and L. R. Forrest, "Investigation of the sodium-binding sites in the sodium-coupled betaine transporter BetP," *Proc Natl Acad Sci U S A*, vol. 109, no. 44, pp. E3035–E3044, 2012.

[26] E. R. Vimr, K. A. Kalivoda, E. L. Deszo, and S. M. Steenbergen, "Diversity of microbial sialic acid metabolism," *Microbiology and Molecular Biology Reviews*, vol. 68, no. 1, pp. 132–153, 2004.

[27] E. R. Vimr, "Unified Theory of Bacterial Sialometabolism: How and Why Bacteria Metabolize Host Sialic Acids," *ISRN Microbiology*, vol. 2013, pp. 1–26, 2013.

[28] A. Varki, "Loss of N-glycolylneuraminic acid in humans: Mechanisms, consequences, and implications for hominid evolution," *American Journal of Physical Anthropology*, vol. 116, no. S33, pp. 54–69, 2001.

[29] S. Inoue and K. Kitajima, "KDN (deaminated neuraminic acid): dreamful past and exciting future of the newest member of the sialic acid family," *Glycoconj J*, vol. 23, no. 5-6, pp. 277–290, 2006.

[30] E. Severi, D. W. Hood, and G. H. Thomas, "Sialic acid utilization by bacterial pathogens," *Microbiology*, vol. 153, no. 9, pp. 2817–2822, 2007.

[31] B. L. Haines-Menges, W. B. Whitaker, J. B. Lubin, and E. F. Boyd, "Host Sialic Acids: A Delicacy for the Pathogen with Discerning Taste," *Microbiol Spectr*, vol. 3, no. 4, pp. 1–26, 2015.

[32] C. Robbe, C. Capon, B. Coddeville, and J. C. Michalski, "Structural diversity and specific distribution of O-glycans in normal human mucins along the intestinal tract," *Biochem J*, vol. 384, no. Pt 2, pp. 307–316, 2004.

[33] A. L. Lewis and W. G. Lewis, "Host sialoglycans and bacterial sialidases: a mucosal perspective," *Cell Microbiol*, vol. 14, no. 8, pp. 1174–1182, 2012.

[34] V. Bouchet, D. W. Hood, J. Li, J. R. Brisson, G. A. Randle, A. Martin, Z. Li, R. Goldstein, E. K. Schweda, S. I. Pelton, J. C. Richards, and E. R. Moxon, "Host-derived sialic acid is incorporated into *Haemophilus influenzae* lipopolysaccharide and is a major virulence factor in experimental otitis media," *Proc Natl Acad Sci U S A*, vol. 100, no. 15, pp. 8898–8903, 2003.

[35] S. Almagro-Moreno and E. F. Boyd, "Sialic acid catabolism confers a competitive advantage to pathogenic *Vibrio cholerae* in the mouse intestine," *Infect Immun*, vol. 77, no. 9, pp. 3807–3816, 2009.

[36] S. Almagro-Moreno and E. F. Boyd, "Bacterial catabolism of nonulosonic (sialic) acid and fitness in the gut," *Gut Microbes*, vol. 1, no. 1, pp. 45–50, 2010.

[37] W. G. Lewis, L. S. Robinson, N. M. Gilbert, J. C. Perry, and A. L. Lewis, "Degradation, foraging, and depletion of mucus sialoglycans by the vagina-adapted *Actinobacterium Gardnerella vaginalis*," *J Biol Chem*, vol. 288, no. 17, pp. 12067–12079, 2013.

[38] W. Zhao, T. L. Chen, B. M. Vertel, and K. J. Colley, "The CMP-sialic acid transporter is localized in the medial-trans Golgi and possesses two specific endoplasmic reticulum export motifs in its carboxyl-terminal cytoplasmic tail," *J Biol Chem*, vol. 281, no. 41, pp. 31106–31118, 2006.

[39] R. A. North, C. R. Horne, J. S. Davies, D. M. Remus, A. C. Muscroft-Taylor, P. Goyal, W. Y. Wahlgren, S. Ramaswamy, R. Friemann, and R. C. J. Dobson, ""Just a spoonful of sugar...": import of sialic acid across bacterial cell membranes," *Biophys Rev*, vol. 10, no. 2, pp. 219–227, 2018.

[40] P. M. Jones and A. M. George, "The ABC transporter structure and mechanism: perspectives on recent research," *Cell Mol Life Sci*, vol. 61, no. 6, pp. 682–699, 2004.

[41] C. F. Higgins, "ABC transporters: from microorganisms to man," *Annu Rev Cell Biol*, vol. 8, pp. 67–113, 1992.

[42] W. Saurin, M. Hofnung, and E. Dassa, "Getting in or out: early segregation between importers and exporters in the evolution of ATP-binding cassette (ABC) transporters," *J Mol Evol*, vol. 48, no. 1, pp. 22–41, 1999.

[43] D. M. B. Post, R. Mungur, B. W. Gibson, and R. S. Munson, "Identification of a Novel Sialic Acid Transporter in *Haemophilus ducreyi*," *Infection and Immunity*, vol. 73, no. 10, pp. 6727–6735, 2005.

[44] D. J. Kelly and G. H. Thomas, "The tripartite ATP-independent periplasmic (TRAP) transporters of bacteria and archaea," *FEMS Microbiol Rev*, vol. 25, no. 4, pp. 405–424, 2001.

[45] J. A. Forward, M. C. Behrendt, N. R. Wyborn, R. Cross, and D. J. Kelly, "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," *J Bacteriol*, vol. 179, no. 17, pp. 5482–5493, 1997.

[46] S. Allen, A. Zaleski, J. W. Johnston, B. W. Gibson, and M. A. Apicella, "Novel sialic acid transporter of *Haemophilus influenzae*," *Infection and immunity*, vol. 73, no. 9, pp. 5291–5300, 2005.

[47] C. Mulligan, E. R. Geertsma, E. Severi, D. J. Kelly, B. Poolman, and G. H. Thomas, "The substrate-binding protein imposes directionality on an electrochemical sodium gradient-driven TRAP transporter," *Proc Natl Acad Sci U S A*, vol. 106, no. 6, pp. 1778–1783, 2009.

[48] N. Chowdhury, J. Norris, E. McAlister, S. Y. Lau, G. H. Thomas, and E. F. Boyd, "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*, vol. 158, no. Pt 8, pp. 2158–2167, 2012.

[49] E. R. Vimr and F. A. Troy, "Identification of an inducible catabolic system for sialic acids (nan) in *Escherichia coli*," *J Bacteriol*, vol. 164, no. 2, pp. 845–853, 1985.

[50] J. Martinez, S. Steenbergen, and E. Vimr, "Derived structure of the putative sialic acid transporter from *Escherichia coli* predicts a novel sugar permease domain," *J Bacteriol*, vol. 177, no. 20, pp. 6005–6010, 1995.

[51] E. Severi, A. H. Hosie, J. A. Hawkhead, and G. H. Thomas, "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 Microbiol Lett*, vol. 304, no. 1, pp. 47–54, 2010.

[52] Y. Huang, M. J. Lemieux, J. Song, M. Auer, and D. N. Wang, "Structure and mechanism of the glycerol-3-phosphate transporter from *Escherichia coli*," *Science*, vol. 301, no. 5633, pp. 616–620, 2003.

[53] M. E. Olson, J. M. King, T. L. Yahr, and A. R. Horswill, "Sialic Acid Catabolism in *Staphylococcus aureus*," *Journal of Bacteriology*, vol. 195, no. 8, pp. 1779–1788, 2013.

[54] S. Almagro-Moreno and E. F. Boyd, "Insights into the evolution of sialic acid catabolism among bacteria," *BMC Evolutionary Biology*, vol. 9, pp. 1–16, 2009.

[55] D. M. Walters, V. L. Stirewalt, and S. B. Melville, "Cloning, Sequence, and Transcriptional Regulation of the Operon Encoding a Putative N-Acetylmannosamine-6-Phosphate Epimerase (nanE) and Sialic Acid Lyase (nanA) in *Clostridium perfringens*," *Journal of Bacteriology*, vol. 181, no. 15, pp. 4526–4532, 1999.

[56] K. M. Ng, J. A. Ferreyra, S. K. Higginbottom, J. B. Lynch, P. C. Kashyap, S. Gopinath, N. Naidu, B. Choudhury, B. C. Weimer, D. M. Monack, and J. L. Sonnenburg, "Microbiota-liberated host sugars facilitate post-antibiotic expansion of enteric pathogens," *Nature*, vol. 502, pp. 96–99, 2013.

[57] E. M. Wright, D. D. Loo, B. A. Hirayama, and E. Turk, "Surprising versatility of Na^+-glucose cotransporters: SLC5," *Physiology (Bethesda)*, vol. 19, no. 6, pp. 370–376, 2004.

[58] S. G. Schultz and P. F. Curran, "Coupled transport of sodium and organic solutes," *Physiol Rev*, vol. 50, no. 4, pp. 637–718, 1970.

[59] D. G. Brown, "Drug discovery strategies to outer membrane targets in Gram-negative pathogens," *Bioorg Med Chem*, vol. 24, no. 24, pp. 6320–6331, 2016.

[60] K. S. McKeegan, M. I. Borges-Walmsley, and A. R. Walmsley, "Microbial and viral drug resistance mechanisms," *Trends Microbiol*, vol. 10, no. 10 Suppl, pp. S8–S14, 2002.

[61] R. M. Epand, C. Walker, R. F. Epand, and N. A. Magarvey, "Molecular mechanisms of membrane targeting antibiotics," *Biochim Biophys Acta*, vol. 1858, no. 5, pp. 980–987, 2016.

[62] E. Henrich, Y. Ma, I. Engels, D. Munch, C. Otten, T. Schneider, B. Henrichfreise, H. G. Sahl, V. Dotsch, and F. Bernhard, "Lipid Requirements for the Enzymatic Activity of MraY Translocases and *in Vitro* Reconstitution of the Lipid II Synthesis Pathway," *J Biol Chem*, vol. 291, no. 5, pp. 2535–2546, 2016.

[63] W. Vollmer, D. Blanot, and M. A. de Pedro, "Peptidoglycan structure and architecture," *FEMS Microbiol Rev*, vol. 32, no. 2, pp. 149–167, 2008.

[64] S. Dramsi, S. Magnet, S. Davison, and M. Arthur, "Covalent attachment of proteins to peptidoglycan," *FEMS Microbiol Rev*, vol. 32, no. 2, pp. 307–320, 2008.

[65] H. J. Rogers, H. R. Perkins, and J. B. Ward, *Microbial cell walls and membranes*. London: Chapman and Hall, 1980.

[66] J. V. Holtje, "Growth of the stress-bearing and shape-maintaining murein sacculus of *Escherichia coli*," *Microbiol Mol Biol Rev*, vol. 62, no. 1, pp. 181–203, 1998.

[67] N. Nanninga, "Morphogenesis of *Escherichia coli*," *Microbiol Mol Biol Rev*, vol. 62, no. 1, pp. 110–129, 1998.

[68] W. Vollmer and J. V. Holtje, "Morphogenesis of *Escherichia coli*," *Curr Opin Microbiol*, vol. 4, no. 6, pp. 625–633, 2001.

[69] T. Mohammadi, A. Karczmarek, M. Crouvoisier, A. Bouhss, D. Mengin-Lecreulx, and T. den Blaauwen, "The essential peptidoglycan glycosyltransferase MurG forms a complex with proteins involved in lateral envelope growth as well as with proteins involved in cell division in *Escherichia coli*," *Mol Microbiol*, vol. 65, no. 4, pp. 1106–1121, 2007.

[70] M. A. Lehrman, "A family of UDP-GlcNAc/MurNAc: polyisoprenol-P GlcNAc/MurNAc-1-P transferases," *Glycobiology*, vol. 4, no. 6, pp. 768–771, 1994.

[71] N. P. Price and F. A. Momany, "Modeling bacterial UDP-HexNAc: polyprenol-P HexNAc-1-P transferases," *Glycobiology*, vol. 15, no. 9, pp. 29R–42R, 2005.

[72] A. Heifetz, R. W. Keenan, and A. D. Elbein, "Mechanism of action of tunicamycin on the UDP-GlcNAc:dolichyl-phosphate Glc-NAc-1-phosphate transferase," *Biochemistry*, vol. 18, no. 11, pp. 2186–2192, 1979.

[73] R. Tommasi, D. G. Brown, G. K. Walkup, J. I. Manchester, and A. A. Miller, "ESKAPEing the labyrinth of antibacterial discovery," *Nat Rev Drug Discov*, vol. 14, no. 8, pp. 529–542, 2015.

[74] D. J. Payne, M. N. Gwynn, D. J. Holmes, and D. L. Pompliano, "Drugs for bad bugs: confronting the challenges of antibacterial discovery," *Nat Rev Drug Discov*, vol. 6, no. 1, pp. 29–40, 2007.

[75] C. Dini, "MraY Inhibitors as Novel Antibacterial Agents," *Curr Top Med Chem*, vol. 5, no. 13, pp. 1221–1236, 2005.

[76] D. K. Banerjee, "Amphomycin inhibits mannosylphosphoryldolichol synthesis by forming a complex with dolichylmonophosphate," *Journal of Biological Chemistry*, vol. 264, no. 4, pp. 2024–2028, 1989.

[77] P. E. Brandish, K. I. Kimura, M. Inukai, R. Southgate, J. T. Lonsdale, and T. D. Bugg, "Modes of action of tunicamycin, liposidomycin B, and mureidomycin A: inhibition of phospho-N-acetylmuramyl-pentapeptide translocase from *Escherichia coli*," *Antimicrobial Agents and Chemotherapy*, vol. 40, no. 7, pp. 1640–1644, 1996.

[78] A. Takatsuki, K. Arima, and G. Tamura, "Tunicamycin, a new antibiotic. I. Isolation and characterization of tunicamycin," *J Antibiot (Tokyo)*, vol. 24, no. 4, pp. 215–223, 1971.

[79] T. D. Bugg, A. J. Lloyd, and D. I. Roper, "Phospho-MurNAc-pentapeptide translocase (MraY) as a target for antibacterial agents and antibacterial proteins," *Infect Disord Drug Targets*, vol. 6, no. 2, pp. 85–106, 2006.

[80] P. E. Brandish, M. K. Burnham, J. T. Lonsdale, R. Southgate, M. Inukai, and T. D. Bugg, "Slow binding inhibition of phospho-N-acetylmuramyl-pentapeptide-translocase (*Escherichia coli*) by mureidomycin A," *J Biol Chem*, vol. 271, no. 13, pp. 7609–7614, 1996.

[81] N. P. Price, T. M. Hartman, J. Li, K. K. Velpula, T. A. Naumann, M. R. Guda, B. Yu, and K. M. Bischoff, "Modified tunicamycins with reduced eukaryotic toxicity that enhance the antibacterial activity of beta-lactams," *J Antibiot (Tokyo)*, vol. 70, no. 11, pp. 1070–1077, 2017.

[82] J. Yoo, E. H. Mashalidis, A. C. Y. Kuk, K. Yamamoto, B. Kaeser, S. Ichikawa, and S.-Y. Lee, "GlcNAc-1-P-transferase–tunicamycin complex structure reveals basis for inhibition of N-glycosylation," *Nature Structural and Molecular Biology*, vol. 25, no. 3, pp. 217–224, 2018.

[83] L. A. McDonald, L. R. Barbieri, G. T. Carter, E. Lenoy, J. Lotvin, P. J. Petersen, M. M. Siegel, G. Singh, and R. T. Williamson, "Structures of the muraymycins, novel peptidoglycan biosynthesis inhibitors," *J Am Chem Soc*, vol. 124, no. 35, pp. 10260–10261, 2002.

[84] D. Wiegmann, S. Koppermann, M. Wirth, G. Niro, K. Leyerer, and C. Ducho, "Muraymycin nucleoside-peptide antibiotics: uridine-derived natural products as lead structures for the development of novel antibacterial agents," *Beilstein Journal of Organic Chemistry*, vol. 12, pp. 769–795, 2016.

[85] J. Hering, E. Dunevall, M. Ek, and G. Branden, "Structural basis for selective inhibition of antibacterial target MraY, a membrane-bound enzyme involved in peptidoglycan synthesis," *Drug Discov Today*, vol. 23, no. 7, pp. 1426–1435, 2018.

[86] J. S. Anderson, M. Matsuhashi, M. A. Haskin, and J. L. Strominger, "Lipid-phosphoacetylmuramyl-pentapeptide and lipid-phosphodisaccharide-pentapeptide: presumed membrane transport intermediates in cell wall synthesis," *Proc Natl Acad Sci U S A*, vol. 53, pp. 881–889, 1965.

[87] W. G. Struve and F. C. Neuhaus, "Evidence for an initial acceptor of UDP-NAc-Muramyl-pentapeptide in the synthesis of bacterial mucopeptide," *Biochem Biophys Res Commun*, vol. 18, no. 1, pp. 6–12, 1965.

[88] M. Ikeda, M. Wachi, H. K. Jung, F. Ishino, and M. Matsuhashi, "The *Escherichia coli* mraY gene encoding UDP-N-acetylmuramoyl-pentapeptide: undecaprenyl-phosphate phospho-N-acetylmuramoyl-pentapeptide transferase," *J Bacteriol*, vol. 173, no. 3, pp. 1021–1026, 1991.

[89] A. Bouhss, D. Mengin-Lecreulx, D. Le Beller, and J. Van Heijenoort, "Topological analysis of the MraY protein catalysing the first membrane step of peptidoglycan synthesis," *Mol Microbiol*, vol. 34, no. 3, pp. 576–585, 1999.

[90] A. Bouhss, M. Crouvoisier, D. Blanot, and D. Mengin-Lecreulx, "Purification and characterization of the bacterial MraY translocase catalyzing the first membrane step of peptidoglycan biosynthesis," *J Biol Chem*, vol. 279, no. 29, pp. 29974–29980, 2004.

[91] Y. Ma, D. Munch, T. Schneider, H. G. Sahl, A. Bouhss, U. Ghoshdastider, J. Wang, V. Dotsch, X. Wang, and F. Bernhard, "Preparative scale cell-free production and quality optimization of MraY homologues in different expression modes," *J Biol Chem*, vol. 286, no. 45, pp. 38844–38853, 2011.

[92] A. E. Backmark, N. Olivier, A. Snijder, E. Gordon, N. Dekker, and A. D. Ferguson, "Fluorescent probe for high-throughput screening of membrane protein expression," *Protein Sci*, vol. 22, no. 8, pp. 1124–1132, 2013.

[93] B. C. Chung, J. Zhao, R. A. Gillespie, D. Y. Kwon, Z. Guan, J. Hong, P. Zhou, and S. Y. Lee, "Crystal structure of MraY, an essential membrane enzyme for bacterial cell wall synthesis," *Science*, vol. 341, no. 6149, pp. 1012–1016, 2013.

[94] J. K. Hakulinen, J. Hering, G. Branden, H. Chen, A. Snijder, M. Ek, and P. Johansson, "Mray-antibiotic complex reveals details of tunicamycin mode of action," *Nat Chem Biol*, vol. 13, no. 3, pp. 265–267, 2017.

[95] M. Saidijam, G. Psakis, J. L. Clough, J. Meuller, S. Suzuki, C. J. Hoyle, S. L. Palmer, S. M. Morrison, M. K. Pos, R. C. Essenberg, M. C. Maiden, A. Abu-bakr, S. G. Baumberg, A. A. Neyfakh, J. K. Griffith, M. J. Stark, A. Ward, J. O'Reilly, N. G. Rutherford, M. K. Phillips-Jones, and P. J. Henderson, "Collection and characterisation of bacterial membrane proteins," *FEBS Lett*, vol. 555, no. 1, pp. 170–175, 2003.

[96] W. P. Hammes and F. C. Neuhaus, "On the specificity of phospho-N-acetylmuramyl-pentapeptide translocase. The peptide subunit of uridine diphosphate-N-actylmuramyl-pentapeptide," *J Biol Chem*, vol. 249, no. 10, pp. 3140–3150, 1974.

[97] W. A. Weppner and F. C. Neuhaus, "Fluorescent substrate for nascent peptidoglycan synthesis. Uridine diphosphate-N-acetylmuramyl-(Nepsilon-5-dimethylaminonaphthalene-1-sulfonyl)penta peptide," *J Biol Chem*, vol. 252, no. 7, pp. 2296–2303, 1977.

[98] T. Stachyra, C. Dini, P. Ferrari, A. Bouhss, J. van Heijenoort, D. Mengin-Lecreulx, D. Blanot, J. Biton, and D. Le Beller, "Fluorescence detection-based functional assay for high-throughput screening for MraY," *Antimicrobial agents and chemotherapy*, vol. 48, no. 3, pp. 897–902, 2004.

[99] A. B. Shapiro, H. Jahic, N. Gao, L. Hajec, and O. Rivin, "A high-throughput, homogeneous, fluorescence resonance energy transfer-based assay for phospho-N-acetylmuramoyl-pentapeptide translocase (MraY)," *J Biomol Screen*, vol. 17, no. 5, pp. 662–672, 2012.

[100] B. Al-Dabbagh, X. Henry, M. El Ghachi, G. Auger, D. Blanot, C. Parquet, D. Mengin-Lecreulx, and A. Bouhss, "Active site mapping of MraY, a member of the polyprenyl-phosphate N-acetylhexosamine 1-phosphate transferase superfamily, catalyzing the first membrane step of peptidoglycan biosynthesis," *Biochemistry*, vol. 47, no. 34, pp. 8919–8928, 2008.

[101] A. J. Lloyd, P. E. Brandish, A. M. Gilbey, and T. D. Bugg, "Phospho-N-acetyl-muramyl-pentapeptide translocase from *Escherichia coli*: catalytic role of conserved aspartic acid residues," *J Bacteriol*, vol. 186, no. 6, pp. 1747–1757, 2004.

[102] N. I. Howard and T. D. Bugg, "Synthesis and activity of 5'-uridinyl dipeptide analogues mimicking the amino terminal peptide chain of nucleoside antibiotic mureidomycin A," *Bioorg Med Chem*, vol. 11, no. 14, pp. 3083–3099, 2003.

[103] K.-T. Chen, P.-T. Chen, C.-K. Lin, L.-Y. Huang, C.-M. Hu, Y.-F. Chang, H.-T. Hsu, T.-J. R. Cheng, Y.-T. Wu, and W.-C. Cheng, "Structural Investigation of Park's Nucleotide on Bacterial Translocase MraY: Discovery of Unexpected MraY Inhibitors," *Scientific Reports*, vol. 6, pp. 1–11, 2016.

[104] B. Al-Dabbagh, S. Olatunji, M. Crouvoisier, M. El Ghachi, D. Blanot, D. Mengin-Lecreulx, and A. Bouhss, "Catalytic mechanism of MraY and WecA, two paralogues of the polyprenyl-phosphate N-acetylhexosamine 1-phosphate transferase superfamily," *Biochimie*, vol. 127, pp. 249–257, 2016.

[105] S. Wagner, M. M. Klepsch, S. Schlegel, A. Appel, R. Draheim, M. Tarry, M. Högbom, K. J. van Wijk, D. J. Slotboom, J. O. Persson, and J.-W. de Gier, "Tuning *Escherichia coli* for membrane protein overexpression," *Proceedings of the National Academy of Sciences*, vol. 105, no. 38, pp. 14371–14376, 2008.

[106] G. L. Rosano and E. A. Ceccarelli, "Recombinant protein expression in *Escherichia coli*: advances and challenges," *Front Microbiol*, vol. 5, pp. 1–17, 2014.

[107] F. W. Studier, A. H. Rosenberg, J. J. Dunn, and J. W. Dubendorff, "Use of T7 RNA polymerase to direct expression of cloned genes," *Methods Enzymol*, vol. 185, pp. 60–89, 1990.

[108] I. Iost, J. Guillerez, and M. Dreyfus, "Bacteriophage T7 RNA polymerase travels far ahead of ribosomes *in vivo*," *Journal of Bacteriology*, vol. 174, no. 2, pp. 619–622, 1992.

[109] M. J. Giacalone, A. M. Gentile, B. T. Lovitt, N. L. Berkley, C. W. Gunderson, and M. W. Surber, "Toxic protein expression in *Escherichia coli* using a rhamnose-based tightly regulated and tunable promoter system," *Biotechniques*, vol. 40, no. 3, pp. 355–364, 2006.

[110] S. Schlegel, J. Lofblom, C. Lee, A. Hjelm, M. Klepsch, M. Strous, D. Drew, D. J. Slotboom, and J. W. de Gier, "Optimizing membrane protein overexpression in the *Escherichia coli* strain Lemo21(DE3)," *J Mol Biol*, vol. 423, no. 4, pp. 648–659, 2012.

[111] F. W. Studier and B. A. Moffatt, "Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes," *J Mol Biol*, vol. 189, no. 1, pp. 113–130, 1986.

[112] A. Deb, W. A. Johnson, A. P. Kline, B. J. Scott, L. R. Meador, D. Srinivas, J. M. Martin-Garcia, K. Dorner, C. R. Borges, R. Misra, B. G. Hogue, P. Fromme, and T. S. Mor, "Bacterial expression, correct membrane targeting and functional folding of the HIV-1 membrane protein Vpu using a periplasmic signal peptide," *PLoS One*, vol. 12, no. 2, pp. 1–23, 2017.

[113] D. Drew, D. J. Slotboom, G. Friso, T. Reda, P. Genevaux, M. Rapp, N. M. Meindl-Beinker, W. Lambert, M. Lerch, D. O. Daley, K. J. Van Wijk, J. Hirst, E. Kunji, and J. W. De Gier, "A scalable, GFP-based pipeline for membrane protein overexpression screening and purification," *Protein Sci*, vol. 14, no. 8, pp. 2011–2017, 2005.

[114] J. M. Hsieh, G. M. Besserer, M. G. Madej, H. Q. Bui, S. Kwon, and J. Abramson, "Bridging the gap: a GFP-based strategy for overexpression and purification of membrane proteins with intra and extracellular C-termini," *Protein Sci*, vol. 19, no. 4, pp. 868–880, 2010.

[115] B. J. Feilmeier, G. Iseminger, D. Schroeder, H. Webber, and G. J. Phillips, "Green fluorescent protein functions as a reporter for protein localization in *Escherichia coli*," *J Bacteriol*, vol. 182, no. 14, pp. 4068–4076, 2000.

[116] D. E. Drew, G. von Heijne, P. Nordlund, and J. W. de Gier, "Green fluorescent protein as an indicator to monitor membrane protein overexpression in *Escherichia coli*," *FEBS Lett*, vol. 507, no. 2, pp. 220–224, 2001.

[117] D. Drew, M. Lerch, E. Kunji, D. J. Slotboom, and J. W. de Gier, "Optimization of membrane protein overexpression and purification using GFP fusions," *Nat Methods*, vol. 3, no. 4, pp. 303–313, 2006.

[118] J. A. Bornhorst and J. J. Falke, "Purification of proteins using polyhistidine affinity tags," *Methods Enzymol*, vol. 326, pp. 245–254, 2000.

[119] G. G. Privé, "Detergents for the stabilization and crystallization of membrane proteins," *Methods*, vol. 41, no. 4, pp. 388–397, 2007.

[120] S. Newstead, S. Ferrandon, and S. Iwata, "Rationalizing alpha-helical membrane protein crystallization," *Protein Sci*, vol. 17, no. 3, pp. 466–472, 2008.

[121] E. M. Landau and J. P. Rosenbusch, "Lipidic cubic phases: a novel concept for the crystallization of membrane proteins," *Proc Natl Acad Sci U S A*, vol. 93, no. 25, pp. 14532–14535, 1996.

[122] J. Broecker, B. T. Eger, and O. P. Ernst, "Crystallogenesis of Membrane Proteins Mediated by Polymer-Bounded Lipid Nanodiscs," *Structure*, vol. 25, no. 2, pp. 384–392, 2017.

[123] G. L. Taylor, "Introduction to phasing," *Acta Crystallogr D Biol Crystallogr*, vol. 66, no. Pt 4, pp. 325–338, 2010.

[124] J. M. Bijvoet, "Structure of Optically Active Compounds in the Solid State," *Nature*, vol. 173, pp. 888–891, 1954.

[125] T. C. Terwilliger, G. Bunkoczi, L. W. Hung, P. H. Zwart, J. L. Smith, D. L. Akey, and P. D. Adams, "Can I solve my structure by SAD phasing? Planning an experiment, scaling data and evaluating the useful anomalous correlation and anomalous signal," *Acta Crystallogr D Struct Biol*, vol. 72, no. Pt 3, pp. 359–374, 2016.

[126] D. Vivian and J. E. Polli, "Mechanistic interpretation of conventional Michaelis-Menten parameters in a transporter system," *European Journal of Pharmaceutical Sciences*, vol. 64, pp. 44–52, 2014.

[127] C. J. Wienken, P. Baaske, U. Rothbauer, D. Braun, and S. Duhr, "Protein-binding assays in biological liquids using microscale thermophoresis," *Nat Commun*, vol. 1, pp. 1–7, 2010.

[128] T. Förster, "Zwischenmolekulare Energiewanderung und Fluoreszenz," *Annalen der Physik*, vol. 437, no. 1-2, pp. 55–75, 1948.

[129] C. Gustafsson, S. Govindarajan, and J. Minshull, "Codon bias and heterologous protein expression," *Trends Biotechnol*, vol. 22, no. 7, pp. 346–353, 2004.

[130] L. M. Guzman, D. Belin, M. J. Carson, and J. Beckwith, "Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter," *J Bacteriol*, vol. 177, no. 14, pp. 4121–4130, 1995.

[131] A. Bernsel, H. Viklund, A. Hennerdal, and A. Elofsson, "TOPCONS: consensus prediction of membrane protein topology," *Nucleic Acids Res*, vol. 37, no. Web Server issue, pp. W465–W468, 2009.

[132] Z. R. Yang, R. Thomson, P. McNeil, and R. M. Esnouf, "RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins," *Bioinformatics*, vol. 21, no. 16, pp. 3369–3376, 2005.

[133] A. Vera, N. Gonzalez-Montalban, A. Aris, and A. Villaverde, "The conformational quality of insoluble recombinant proteins is enhanced at low growth temperatures," *Biotechnol Bioeng*, vol. 96, no. 6, pp. 1101–1106, 2007.

[134] T. Kawate and E. Gouaux, "Fluorescence-detection size-exclusion chromatography for precrystallization screening of integral membrane proteins," *Structure*, vol. 14, no. 4, pp. 673–681, 2006.

[135] P. S. Chae, S. G. F. Rasmussen, R. R. Rana, K. Gotfryd, R. Chandra, M. A. Goren, A. C. Kruse, S. Nurva, C. J. Loland, Y. Pierre, D. Drew, J.-L. Popot, D. Picot, B. G. Fox, L. Guan, U. Gether, B. Byrne, B. Kobilka, and S. H. Gellman, "Maltose-neopentyl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins," *Nature methods*, vol. 7, no. 12, pp. 1003–1008, 2010.

[136] W. Y. Wahlgren, E. Dunevall, R. A. North, A. Paz, M. Scalise, P. Bisignano, J. Bengtsson-Palme, P. Goyal, E. Claesson, R. Caing-Carlsson, R. Andersson, K. Beis, U. J. Nilsson, A. Farewell, L. Pochini, C. Indiveri, M. Grabe, R. C. J. Dobson, J. Abramson, S. Ramaswamy, and R. Friemann, "Substrate-bound outward-open structure of a Na^+-coupled sialic acid symporter reveals a new Na^+ site," *Nature Communications*, vol. 9, no. 1, pp. 1–14, 2018.

[137] J. Abramson and E. M. Wright, "Structure and function of Na^+-symporters with inverted repeats," *Curr Opin Struct Biol*, vol. 19, no. 4, pp. 425–432, 2009.

[138] S. Ressl, A. C. Terwisscha van Scheltinga, C. Vonrhein, V. Ott, and C. Ziegler, "Molecular basis of transport and regulation in the Na^+/betaine symporter BetP," *Nature*, vol. 458, no. 7234, pp. 47–52, 2009.

[139] M. Fischer, A. P. Hopkins, E. Severi, J. Hawkhead, D. Bawdon, A. G. Watts, R. E. Hubbard, and G. H. Thomas, "Tripartite ATP-independent Periplasmic (TRAP) Transporters Use an Arginine-mediated Selectivity Filter for High Affinity Substrate Binding," *J Biol Chem*, vol. 290, no. 45, pp. 27113–27123, 2015.

[140] M. Nayal and E. Di Cera, "Valence screening of water in protein crystals reveals potential na^+ binding sites," *J Mol Biol*, vol. 256, no. 2, pp. 228–234, 1996.

[141] M. M. Harding, "Metal-ligand geometry relevant to proteins and in proteins: sodium and potassium," *Acta Crystallogr D Biol Crystallogr*, vol. 58, no. 5, pp. 872–874, 2002.

[142] D. A. Caplan, J. O. Subbotina, and S. Y. Noskov, "Molecular Mechanism of Ion-Ion and Ion-Substrate Coupling in the Na^+-Dependent Leucine Transporter LeuT," *Biophysical Journal*, vol. 95, no. 10, pp. 4613–4621, 2008.

[143] T. Baba, T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K. A. Datsenko, M. Tomita, B. L. Wanner, and H. Mori, "Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection," *Mol Syst Biol*, vol. 2, pp. 1–11, 2006.

[144] F. Grenier, D. Matteau, V. Baby, and S. Rodrigue, "Complete Genome Sequence of *Escherichia coli* BW25113," *Genome announcements*, vol. 2, no. 5, p. 1, 2014.

[145] A. P. Hopkins, J. A. Hawkhead, and G. H. Thomas, "Transport and catabolism of the sialic acids N-glycolylneuraminic acid and 3-keto-3-deoxy-D-glycero-D-galactonononic acid by *Escherichia coli* K-12," *FEMS Microbiol Lett*, vol. 347, no. 1, pp. 14–22, 2013.

[146] C. Mulligan, A. P. Leech, D. J. Kelly, and G. H. Thomas, "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*," *J Biol Chem*, vol. 287, no. 5, pp. 3598–3608, 2012.

[147] K. A. Kalivoda, S. M. Steenbergen, E. R. Vimr, and J. Plumbridge, "Regulation of sialic acid catabolism by the DNA binding protein NanR in *Escherichia coli*," *J Bacteriol*, vol. 185, no. 16, pp. 4806–4815, 2003.

[148] T. Angata and A. Varki, "Chemical diversity in the sialic acids and related alpha-keto acids: an evolutionary perspective," *Chem Rev*, vol. 102, no. 2, pp. 439–469, 2002.

[149] J. Plumbridge and E. Vimr, "Convergent pathways for utilization of the amino sugars N-acetylglucosamine, N-acetylmannosamine, and N-acetylneuraminic acid by *Escherichia coli*," *J Bacteriol*, vol. 181, no. 1, pp. 47–54, 1999.

[150] R. A. North, W. Y. Wahlgren, D. M. Remus, M. Scalise, S. A. Kessans, E. Dunevall, E. Claesson, T. P. Soares da Costa, M. A. Perugini, S. Ramaswamy, J. R. Allison, C. Indiveri, R. Friemann, and R. C. J. Dobson, "The Sodium Sialic Acid Symporter From *Staphylococcus aureus* Has Altered Substrate Specificity," *Frontiers in Chemistry*, vol. 6, pp. 1–11, 2018.

[151] E. Byres, A. W. Paton, J. C. Paton, J. C. Lofling, D. F. Smith, M. C. Wilce, U. M. Talbot, D. C. Chong, H. Yu, S. Huang, X. Chen, N. M. Varki, A. Varki, J. Rossjohn, and T. Beddoe, "Incorporation of a non-human glycan mediates human susceptibility to a bacterial toxin," *Nature*, vol. 456, no. 7222, pp. 648–652, 2008.

[152] X. Yu, V. T. Dang, F. E. Fleming, M. von Itzstein, B. S. Coulson, and H. Blanchard, "Structural Basis of Rotavirus Strain Preference toward N-Acetyl- or N-Glycolylneuraminic Acid-Containing Receptors," *Journal of Virology*, vol. 86, no. 24, pp. 13456–13466, 2012.

[153] Z. M. Khan, Y. Liu, U. Neu, M. Gilbert, B. Ehlers, T. Feizi, and T. Stehle, "Crystallographic and glycan microarray analysis of human polyomavirus 9 VP1 identifies N-glycolyl neuraminic acid as a receptor candidate," *J Virol*, vol. 88, no. 11, pp. 6100–6111, 2014.

[154] J. E. Stencel-Baerenwald, K. Reiss, D. M. Reiter, T. Stehle, and T. S. Dermody, "The sweet spot: defining virus-sialic acid interactions," *Nat Rev Microbiol*, vol. 12, no. 11, pp. 739–749, 2014.

[155] A. Varki, "Colloquium paper: uniquely human evolution of sialic acid genetics and biology," *Proc Natl Acad Sci U S A*, vol. 107, no. Suppl 2, pp. 8939–8946, 2010.

[156] N. M. Varki, E. Strobert, J. Dick, E. J., K. Benirschke, and A. Varki, "Biomedical differences between human and nonhuman hominids: potential roles for uniquely human aspects of sialic acid biology," *Annu Rev Pathol*, vol. 6, pp. 365–393, 2011.

[157] P. Tangvoranuntakul, P. Gagneux, S. Diaz, M. Bardor, N. Varki, A. Varki, and E. Muchmore, "Human uptake and incorporation of an immunogenic nonhuman dietary sialic acid," *Proceedings of the National Academy of Sciences*, vol. 100, no. 21, pp. 12045–12050, 2003.

[158] K. Banda, C. J. Gregg, R. Chow, N. M. Varki, and A. Varki, "Metabolism of vertebrate amino sugars with N-glycolyl groups: mechanisms underlying gastrointestinal incorporation of the non-human sialic acid xeno-autoantigen N-glycolylneuraminic acid," *J Biol Chem*, vol. 287, no. 34, pp. 28852–28864, 2012.

[159] B. C. Chung, E. H. Mashalidis, T. Tanino, M. Kim, A. Matsuda, J. Hong, S. Ichikawa, and S.-Y. Lee, "Structural insights into inhibition of lipid i production in bacterial cell wall synthesis," *Nature*, vol. 533, no. 7604, pp. 557–560, 2016.

[160] M. S. Anderson, S. S. Eveland, and N. P. Price, "Conserved cytoplasmic motifs that distinguish sub-groups of the polyprenol phosphate:N-acetylhexosamine-1-phosphate transferase family," *FEMS Microbiol Lett*, vol. 191, no. 2, pp. 169–175, 2000.

[161] Y. Y. Dong, H. Wang, A. C. W. Pike, S. A. Cochrane, S. Hamedzadeh, F. J. Wyszyński, S. R. Bushell, S. F. Royer, D. A. Widdick, A. Sajid, H. I. Boshoff, Y. Park, R. Lucas, W.-M. Liu, S. S. Lee, T. Machida, L. Minall, S. Mehmood, K. Belaya, W.-W. Liu, A. Chu, L. Shrestha, S. M. M. Mukhopadhyay, C. Strain-Damerell, R. Chalk, N. A. Burgess-Brown, M. J. Bibb, C. E. Barry Iii, C. V. Robinson, D. Beeson, B. G. Davis, and E. P. Carpenter, "Structures of DPAGT1 Explain Glycosylation Disease Mechanisms and Advance TB Antibiotic Design," *Cell*, vol. 175, no. 4, pp. 1045–1058, 2018.

[162] G. P. Kaushal and A. D. Elbein, "Purification and properties of UDP-GlcNAc:dolichylphosphate GlcNAc-1-phosphate transferase. Activation and inhibition of the enzyme," *J Biol Chem*, vol. 260, no. 30, pp. 16303–16309, 1985.

[163] N. Dan, R. B. Middleton, and M. A. Lehrman, "Hamster UDP-N-acetylglucosamine:dolichol-P N-acetylglucosamine-1-P transferase has multiple transmembrane spans and a critical cytosolic loop," *J Biol Chem*, vol. 271, no. 48, pp. 30717–30724, 1996.

[164] P. A. Cockrum and J. A. Edgar, "High-performance liquid chromatograpic comparison of the tunicaminyluracil-based antibiotics corynetoxin, tunicamycin, streptovirudin and MM 19290," *Journal of Chromatography A*, vol. 268, pp. 245–254, 1983.

[165] K. Eckardt, "Tunicamycins, Streptovirudins, and Corynetoxins, a Special Subclass of Nucleoside Antibiotics," *Journal of Natural Products*, vol. 46, no. 4, pp. 544–550, 1983.

[166] K. Eckardt, H. Thrum, G. Bradler, E. Tonew, and M. Tonew, "Streptovirudins, new antibiotics with antibacterial and antiviral activity. II. Isolation, chemical characterization and biological activity of streptovirudins A1, A2, B1, B2, C1, C2, D1, and D2," *J Antibiot*, vol. 28, no. 4, pp. 274–279, 1975.

[167] K. Eckardt, W. Ihn, D. Tresselt, and D. Krebs, "The chemical structures of streptovirudins," *J Antibiot*, vol. 34, no. 12, pp. 1631–1632, 1981.

[168] D. Duksin and W. C. Mahoney, "Relationship of the structure and biological activity of the natural homologues of tunicamycin," *J Biol Chem*, vol. 257, no. 6, pp. 3105–3109, 1982.

[169] K. Kimura and T. D. Bugg, "Recent advances in antimicrobial nucleoside antibiotics targeting cell wall biosynthesis," *Nat Prod Rep*, vol. 20, no. 2, pp. 252–273, 2003.

[170] W. Chen, J. Qi, P. Wu, D. Wan, J. Liu, X. Feng, and Z. Deng, "Natural and engineered biosynthesis of nucleoside antibiotics in Actinomycetes," *J Ind Microbiol Biotechnol*, vol. 43, no. 2-3, pp. 401–417, 2016.

[171] N. P. Price, M. A. Jackson, K. E. Vermillion, J. A. Blackburn, J. Li, and B. Yu, "Selective catalytic hydrogenation of the N-acyl and uridyl double bonds in the tunicamycin family of protein N-glycosylation inhibitors," *J Antibiot (Tokyo)*, vol. 70, no. 12, pp. 1122–1128, 2017.

[172] N. P. Price, D. P. Labeda, T. A. Naumann, K. E. Vermillion, M. J. Bowman, M. A. Berhow, W. W. Metcalf, and K. M. Bischoff, "Quinovosamycins: new tunicamycin-type antibiotics in which the alpha, beta-1",11'-linked N-acetylglucosamine residue is replaced by N-acetylquinovosamine," *J Antibiot (Tokyo)*, vol. 69, no. 8, pp. 637–646, 2016.

[173] O. H. Hashim and W. Cushley, "Minor modifications to the structure of tunicamycin lead to loss of the biological activity of the antibiotic," *Biochim Biophys Acta*, vol. 923, no. 3, pp. 362–370, 1987.

[174] R. Mancusso, N. K. Karpowich, B. K. Czyzewski, and D. N. Wang, "Simple screening method for improving membrane protein thermostability," *Methods*, vol. 55, no. 4, pp. 324–329, 2011.

[175] P. Bisignano, C. Ghezzi, H. Jo, N. F. Polizzi, T. Althoff, C. Kalyanaraman, R. Friemann, M. P. Jacobson, E. M. Wright, and M. Grabe, "Inhibitor binding mode and allosteric regulation of Na^+-glucose symporters," *Nature Communications*, vol. 9, no. 1, pp. 1–10, 2018.