STRUCTURE AND LIPID INTERACTIONS OF MEMBRANE-ASSOCIATED GLYCOSYLTRANSFERASES

Scarlett Szpyngiel
Structure and lipid interactions of membrane-associated glycosyltransferases

Cationic patches and anionic lipids regulate biomembrane binding of both GT-A and GT-B enzymes

Scarlett Szporyngiel


Printed in Sweden by Holmbergs, Malmö 2016

Distributor: Department of Biochemistry and Biophysics, Stockholm University
Till Tesla
Abstract

This thesis concerns work on structure and membrane interactions of enzymes involved in lipid synthesis, biomembrane and cell wall regulation and cell defense processes. These proteins, known as glycosyltransferases (GTs), are involved in the transfer of sugar moieties from nucleotide sugars to lipids or chitin polymers. Glycosyltransferases from three types of organisms have been investigated; one is responsible for vital lipid synthesis in Arabidopsis thaliana (atDGD2) and adjusts the lipid content in biomembranes if the plant experiences stressful growth conditions. This enzyme shares many structural features with another GT found in gram-negative bacteria (WaaG). WaaG is however continuously active and involved in synthesis of the protective lipopolysaccharide layer in the cell walls of Escherichia coli. The third type of enzymes investigated here are chitin synthases (ChS) coupled to filamentous growth in the oomycete Saprolegnia monoica. I have investigated two ChS-derived MIT domains that may be involved in membrane interactions within the endosomal pathway.

From analysis of the three-dimensional structure and the amino-acid sequence, some important regions of these very large proteins were selected for in vitro studies. By the use of an array of biophysical methods (e.g. Nuclear Magnetic Resonance, Fluorescence and Circular Dichroism spectroscopy) and directed sequence analyses it was possible to shed light on some important details regarding the structure and membrane-interacting properties of the GTs. The importance of basic amino-acid residues and hydrophobic anchoring segments, both generally and for the abovementioned proteins specifically, is discussed. Also, the topology and amino-acid sequence of GT-B enzymes of the GT4 family are analyzed with emphasis on their biomembrane association modes. The results presented herein regarding the structural and lipid-interacting properties of GTs aid in the general understanding of glycosyltransferase activity. Since GTs are involved in a high number of biochemical processes in vivo it is of utmost importance to understand the underlying processes responsible for their activity, structure and interaction events. The results are likely to be useful for many applications and future experimental design within life sciences and biomedicine.
List of publications

This thesis is based on the following papers, which will be referred to in the text by their Roman numerals. Reprints are made with permission from the publishers. An asterisk (*) indicates equal contribution from the authors.


BMRB/PDB entries associated with the papers:

I. BMRB Entry 17356, PDB accession code 2L7C
II. BMRB Entry 25699, PDB accession code 2N58
IV. BMRB Entry 19987, PDB accession code 2MPK

Additional publication not included in thesis:

Szpryngiel S., Oliveberg M. and Mäler L. Diffuse binding of Zn(2+) to the denatured ensemble of Cu/Zn superoxide dismutase 1 (2015), FEBS Open Bio 5, 56-63. BMRB entry 18968.
Contents

Abstract.................................................................................................................................................vi
List of publications .......................................................................................................................... vii
Abbreviations ...................................................................................................................................... x
1. Prologue............................................................................................................................................... 11
2. The interplay between proteins and cell membranes ................................................................. 13
   2.1. Membrane interaction mechanisms of proteins ........................................................................... 13
3. Lipids and Sugars in Biomembranes and Cell walls ................................................................. 17
   3.1 Biomembranes – function and assembly ..................................................................................... 17
   3.2 Lipids establish the structural scaffold of bio-membranes and act as ligands ...................... 18
      3.2.1 Phospholipids define the bilayer surface and act as ligands and enzyme substrates ................. 18
      3.2.2. Three glycolipids dominate in chloroplasts ........................................................................ 20
   3.3. Cell walls maintain the structural integrity of cells ..................................................................... 22
      3.3.1. Chitin is present in trace amounts in some oomycetes ......................................................... 23
      3.3.2. Lipopolysaccharides and gram-negative bacteria .................................................................. 24
4. Activity, structure and classification of glycosyltransferases ..................................................... 26
   4.1 Glycosyltransferase activity, structure and classification ...................................................... 26
   4.2. A GT-A glycosyltransferase involved in chitin synthesis may act as a target for anti-oomycete compounds .................................................................................................................. 28
      4.2.2. MIT domains and lipid recognition of chitin synthases ....................................................... 30
   4.3 GT-B glycosyltransferases regulates the glycolipid levels in A. thaliana and E. coli biomembranes ................................................................................................................................. 31
      4.3.1. Phosphate starvation triggers DGD2 activity in A. thaliana ................................................. 32
      4.3.2. The GT-B Glycosyltransferase WaaG is involved in formation of shielding carbohydrate layers in E. coli .................................................................................................................. 34
5. Monotopic GT4 glycosyltransferases share structural features coupled to their lipid interacting mechanisms ................................................................................................................................. 35
   5.1. atDGD2 and WaaG anchors to one leaflet of the bilayer with parts of the N-terminal domain .............................................................................................................................................. 35
   5.2. The C-terminal domain of atDGD2 display additional lipid interacting properties ............ 38
   5.3. Topology comparisons of GT4 enzymes reveal common features ................................................. 40
5.4. Conclusions and future prospects ................................................................. 43

6. Chitin synthases in oomycetes interact with lipids via a novel interaction mode, involving MIT domains ................................................................. 45

7. Methods for studies of glycosyltransferases ........................................... 49
   7.1. Sequence analysis and in silico methods to characterize membrane-associated GTs.................................................................................................................. 50
   7.2. Biomembrane mimetics .............................................................................. 55
   7.3. CD spectroscopy in studies of GT-derived segments structure and stability ...... 56
   7.4. NMR spectroscopy to study glycosyltransferases and lipids...................... 57
      7.4.1. Sequence assignment and 3D structure determination ......................... 58
      7.4.2. NMR Diffusion experiments may inform on peptide-bicelle interactions ...... 59
   7.5. Fluorescence spectroscopy for quantification of protein-lipid interactions .... 60

8. Socker och fett - livsviktigt för alla!............................................................... 62

10. Acknowledgements ...................................................................................... 66

11. References ...................................................................................................... 69
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>atDGD2</td>
<td>Digalactosyldiacylglycerol synthase (transferase) from <em>A. thaliana</em></td>
</tr>
<tr>
<td>CD</td>
<td>Circular Dichroism</td>
</tr>
<tr>
<td>ChS</td>
<td>Chitin Synthase</td>
</tr>
<tr>
<td>CL</td>
<td>Cardiolipin</td>
</tr>
<tr>
<td>DAG</td>
<td>Diacylglycerol</td>
</tr>
<tr>
<td>DGDG</td>
<td>Digalactosyldiacylglycerol</td>
</tr>
<tr>
<td>DHEPC</td>
<td>1,2-dihexanoyl-sn-glycero-3-phosphocholine</td>
</tr>
<tr>
<td>DLS</td>
<td>Dynamic Light Scattering</td>
</tr>
<tr>
<td>DPC</td>
<td>n-dodecyl phosphocholine</td>
</tr>
<tr>
<td>GT</td>
<td>Glycosyltransferase</td>
</tr>
<tr>
<td>HSQC</td>
<td>Heteronuclear Single Quantum Coherence</td>
</tr>
<tr>
<td>K_D</td>
<td>Dissociation constant</td>
</tr>
<tr>
<td>K_SV</td>
<td>Stern-Volmer quenching constant</td>
</tr>
<tr>
<td>LPS</td>
<td>Lipopolysaccharide</td>
</tr>
<tr>
<td>LUV</td>
<td>Large Unilamellar Vesicle</td>
</tr>
<tr>
<td>MGDG</td>
<td>Monogalactosyldiacylglycerol</td>
</tr>
<tr>
<td>MIR-WaaG</td>
<td>Membrane-Interacting Region of WaaG</td>
</tr>
<tr>
<td>MIT1/2</td>
<td>Microtubule Interacting and Trafficking domain from smChS1/2</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear Magnetic Resonance</td>
</tr>
<tr>
<td>NOE</td>
<td>Nuclear Overhauser Effect</td>
</tr>
<tr>
<td>OM</td>
<td>Outer Membrane</td>
</tr>
<tr>
<td>PA</td>
<td>Phosphatidic Acid</td>
</tr>
<tr>
<td>PC</td>
<td>Phosphatidylcholine</td>
</tr>
<tr>
<td>PE</td>
<td>Phosphatidylethanolamine</td>
</tr>
<tr>
<td>PG</td>
<td>Phosphatidylglycerol</td>
</tr>
<tr>
<td>Pi</td>
<td>Inorganic phosphate ions</td>
</tr>
<tr>
<td>PIP</td>
<td>Phosphorylated Phosphatidylinositol</td>
</tr>
<tr>
<td>PM</td>
<td>Plasmamembrane</td>
</tr>
<tr>
<td>PS</td>
<td>Phosphatidylserine</td>
</tr>
<tr>
<td>smChS</td>
<td>Chitin synthase from <em>Saprolegnia monoica</em></td>
</tr>
<tr>
<td>TM</td>
<td>Transmembrane</td>
</tr>
<tr>
<td>WaaG</td>
<td>GT involved in LPS synthesis in <em>E. coli</em></td>
</tr>
</tbody>
</table>
1. Prologue

Any student attending biochemistry courses will early on learn about the crucial role that proteins and cell membranes have in all living organisms. Both proteins and biomembranes have been delicately refined during the evolution. Cell membranes can be composed of thousands of different molecules and the expression of proteins may be fine-tuned down to such detail that it is almost difficult to comprehend how this can be possible. The removal of one component may completely strike out the cell and, eventually, the whole organism. Knowledge about these complex systems is the basis for practical applications but most importantly it is essential for all further research within natural sciences that involves biological systems.

My work regards molecular interactions involving proteins and biomembranes. Proteins associated to biomembranes are often challenging to work with \textit{in vitro} since they are large, may require association to bilayers or other interaction partners, and often have very poor solubility in water. To overcome such obstacles I have used biophysical approaches where the assemblies of interest are scaled down to smaller but rudimentary pieces of the puzzle. This approach permits us to pinpoint some of the most critical properties involved in a certain interaction, and allows us to answer questions that cannot be handled by most other means. It also leaves room for interesting speculations about the model systems and how the results may be generalized and extrapolated onto larger assemblies or organisms, including such that are not even closely related phylogenetically. It is satisfying to be able to reproduce results for one system and then predict the outcome for another, both because it strengthens the underlying physicochemical mechanisms, but also since it allows us to tackle hypotheses that have previously not been feasible due to limitations in the experimental setups or lack of sequential homology. The fact that there can also be hands-on applications based on the results is an added bonus. Since the proteins described in this work are involved in a high number of molecular events \textit{in vivo}, in many types of organisms, development of drugs (especially antibiotics), pesticides, or even novel biomaterials is of course
greatly beneficial for humanity and for the non-scientific society that we learn more about the biophysics behind the proteins that satisfies the “sweet tooth” of all living cells. What I mean about that will hopefully become clearer in the following chapters, where I also hope to visualize some of the opportunities that arise from “keeping things simple” when studying glycosyltransferases, biomembranes and related biological systems.
2. The interplay between proteins and cell membranes

Membrane proteins are involved in many processes in or close to biomembranes of plasma membranes and organelles where they act as enzymes, receptors, transporters and are directly involved in trafficking of other proteins and biomolecules. In this thesis two main types of biomembrane-interacting enzymes will be presented. Both are involved in the transfer of sugars from one molecule to another. They do, however, have differences regarding their enzymatic activity and have dissimilar membrane-interacting properties. The introduction of this thesis is divided into three parts: In Chapter 2 I will briefly describe some general lipid interaction patterns before I move on to describing the characteristics and roles of lipids in biomembranes in Chapter 3. Finally, in Chapter 4, details regarding the three GTs that were investigated in this thesis are presented.

2.1. Membrane interaction mechanisms of proteins

Proteins can attach to biomembranes in many ways, either directly or indirectly via other biomolecules or ligands. The glycosyltransferases that are investigated in this work have both integral and peripheral interaction patterns.

**Integral membrane interactions** involve more-or-less permanent anchoring into the lipid bilayer (although all protein-lipid interactions may be rapidly converted by appropriate external factors). The interactions may be distinguished as either transmembrane (TM), where the protein has regions with a hydrophobic outer surface that integrate into the hydrophobic interior composed of lipid acyl tails from both leaflets of the bilayer (Figure 1A). They may span the bilayer with one or several secondary structure elements, or domains. Integral membrane interactions may also be restricted to one part of the bilayer, so-called monotopic proteins (Figure 1B).
There are several ways this can occur, for instance via amphipathic helices that localize parallel to the membrane surface or larger hydrophobic segments (often composed of several small helices) that may penetrate into one leaflet of the bilayer. The intrinsic hydrophobic moment of amphipathic helices may favor electrostatic events involving the interface region, or amphipathic helix formation may be the sole driving force of membrane interactions (entropy driven). It appears that large conformational changes may induce monotopic interactions and based on the rapid recent progress regarding understanding these events, it is likely that conformational variability is much more common in membrane-anchoring proteins than previously expected. Conformational changes of membrane proteins and peptides may occur either upon lipid interaction or upon other ligand binding, commonly substrate molecules or subunit interactions/oligomerization. These types of conformational changes may for example induce new intrinsic secondary structure elements and/or expose
amphipathic helices; they may also expose interaction sites for other ligands that may govern lipid interaction. Both TM and monotopic proteins are enriched in cationic and aromatic amino-acid residues that may often be found in the interface region. Tryptophans and positive charges have been shown to play important roles for membrane binding in monotopic membrane proteins. Although the hydrophobic effect, discussed in Chapter 3, ultimately is what drives the interaction, such species are likely to aid in the interaction. With the help of sequence analyses of integral proteins it is often possible to predict such regions from the amino-acid sequence (see Chapter 7 for more information about protein sequence analysis of membrane-associated proteins). Sequence analysis may also aid in finding regions that from solved (x-ray) structures do not seem to include typical integral regions. The amino-acid sequences of such regions are likely to have a dynamic behavior in order to function in several structural modes (e.g. alternating between helical to β-sheet structure). This is a complex behavior since it must happen without inducement of unwanted protein aggregation (due to sticky surfaces, exposed aromatic residues, un-paired charged residues etc.).

Peripheral interactions (Figure 1B-C), in contrast, often take place without the involvement of hydrophobic amino-acid residues, and instead utilize electrostatics, ligands or other biochemical events that alter the local pKₐ. Electrostatic interactions between cationic patches or residues (Lys/Arg) and anionic lipids are common, and will result in a loose interaction approximately 3 Å from the membrane surface. The energy associated with each positive charge is roughly -1.4 kcal, independent of the type of amino-acid residue and anionic lipid. This means that differences in the interaction strength of a protein segment and two different lipid species most likely involve other processes in addition to the electrostatic component, possibly dependent on the lipid head-group (see Chapter 3 for details regarding some common lipids in vivo). Also, if a region is found to integrate further into the bilayer than what is expected for peripheral cationic helices there are likely other mechanisms involved, for example hydrophobic interactions or aromatic ring electrons as mentioned above. Interestingly, interactions between negatively charged amphipathic helices (enriched in Glu residues) and anionic lipids was also seen for a cytidylyltransferase and suggested to be involved in regulation of membrane binding. Anionic lipids induce a local pH effect with enrichment of protons, ultimately increasing the protonation degree of three Glu residues, thereby decreasing the negative
charge and increasing hydrophobicity of the helix. The authors suggest that the apparent enrichment of glutamates in interfacial regions of membrane proteins can be explained by such mechanisms. Electrostatic association can also involve other non-solvent ligands, such as metal ions or phosphate ions that either alter the local pKₐ and thereby change the protonation states of the nearby residues, or alter the hydration locally, which enhances the interaction. Electrostatic association can also involve other non-solvent ligands, such as metal ions or phosphate ions that either alter the local pKₐ and thereby change the protonation states of the nearby residues, or alter the hydration locally, which enhances the interaction. Phosphorylation or dephosphorylation of certain Ser and Tyr residues close to binding regions can disrupt or strengthen interactions that are electrostatic in nature since anionic charges of the phosphate group will be introduced or removed. This could function as a switch mechanism for other membrane interaction events. Phosphate groups may also attract protons that are clustering at the surface region of anionic membranes, hence strengthen the attraction of the protein towards the bilayer if anionic lipids are present. Metal ion binding in GT-A enzymes occurs via DXD or EXD motifs, as described in Chapter 4. The binding is accompanied by conformational changes and nucleotide-sugar binding and is stabilizing the transition state. Hence, additional roles of metal ions, except involvements in the direct activity, are possible for enzymes that show metal ion dependence for activity.

Multi-domain interactions may be present in large proteins. Both peripheral and integral proteins may have several interaction sites, and they may be separated in several domains. One specific binding site/domain may be strengthened by other interactions since weak lipid interactions, but still with important function, can be enhanced by multiple membrane-binding domains or oligomerization events. Additional interaction regions may increase membrane affinity, or the extra binding mechanism may be involved in formation of the active site by exposure of interaction sites, or chaperone ligands into a correct stereochemistry relative the active site. Both the C-terminal domain of atDGD2 investigated in paper II and the MIT domain of chitin synthases in paper IV are examples of domains of such proteins with multiple lipid interaction modes.
3. Lipids and Sugars in Biomembranes and Cell walls

In this chapter general biomembrane and cell wall properties will be introduced, and the lipids that define plasma membranes and organelles of different organisms will be described. Biomembrane components of the chloroplasts of plants will also be discussed, followed by two types of cell walls of prokaryotes and eukaryotes, with a main focus on gram-negative bacteria and oomycetes.

3.1 Biomembranes – function and assembly

It is basic knowledge that one of the most important tasks of biomembranes is **compartmentation**, i.e. they “keep things in place” and balance the concentrations of solutes in- and outside of the lipid bilayer. However, an equally important role for biomembranes is as a matrix for an immense number of macromolecular events and in regulation of enzymatic activity. Import/export mechanisms via vesicular systems guide proteins, lipids and other macromolecules to the appropriate places, and ensure that hormones and organic compounds are solvated and accessible for enzymes and other interaction partners. Membrane-associated proteins and lipids may be modified by post-translational modifications (PTMs) and thereby acquire novel functions, for instance they may act as anchors for carbohydrates (as described further in section 4) or gain charge. The biomembrane surface is an extremely dynamic milieu, and at any given moment there will be gas molecules, ions and small molecules that intervenes with lipid/protein associations and alter the local pH, as well as an ever-changing hydration layer. Water molecules at the surface of biomembranes are important ligands in that sense and the interplay between water and other molecules may sometimes drive important molecular events. Many important drugs are dependent on the intrinsic biophysical properties of biomembranes as well as the signaling and transport functions biomembranes acquire upon interactions with other biomolecules. The syntheses of lipids and membrane-
associated proteins are often a necessity for all life and are tightly controlled. Proteins are abundant and may be up-regulated if needed by specific trigger events, as exemplified in section 4.

The macromolecular assembly of biomembranes *in vivo* is of course not trivial since there are many components involved and the cell environment interferes with the lipids in many ways (some described above). It is however only one physical phenomenon that is responsible for the assembly and main morphology of all biomembranes: the hydrophobic effect.\textsuperscript{21,22} This is schematically described as a combination of low solubility of the interacting molecules and an increase in entropy upon release of the hydration layer in the interaction interfaces. Hence, the amphiphilic character of all lipids is what forces them into an (ordered) aggregated state. Whether or not this will happen under given conditions depends on the lipid type(s), the solvent, the temperature and the pressure. The overall appearance of the aggregates will also depend on head-group and acyl-chain properties (size, charges, unsaturation, PTMs etc.).\textsuperscript{20} To understand protein-lipid interactions it is important to understand that for any integral or peripheral biomembrane interaction to occur spontaneously there has to be an accompanying thermodynamic gain.

### 3.2 Lipids establish the structural scaffold of biomembranes and act as ligands

Often biomembranes of cells are composed of a few main lipid types, with specificity introduced by less abundant species and PTMs or bound proteins and addition of sterols. The bilayer will be further differentiated by flippases and phase separations that give asymmetry of the bilayers. Here a description of two types of lipids will be described, phospholipids and glycolipids.

#### 3.2.1 Phospholipids define the bilayer surface and act as ligands and enzyme substrates

Phospholipids (i.e. phosphoglycerides, based on a glycerol backbone) are the main components of most eukaryotic and bacterial biomembranes. In eukaryotes the bulk of phospholipids are synthesized and matured in the Endoplasmic Reticulum (ER), and exported to the plasma membrane (PM)
and other organelles (mitochondria, chloroplasts, vacuoles etc.). There are also enzymes in the chloroplasts and mitochondria that synthesize lipids, although the precursor molecules are imported. The thylakoid membranes of chloroplasts have no lipid synthesis but lipids instead need to be transported from other chloroplast membranes or ER via the inner envelope membrane, perhaps via vesicular transfer and/or hexagonal II platforms where MGDG accumulate. The lipid synthesis in chloroplast membranes will be further discussed in Chapter 4.

The structurally simplest phosphoglyceride lipid has its glycerol hydroxyl group esterified to **phosphatidic acid (PA)**. Although PA is not that abundant *in vivo*, it is extremely important as a precursor lipid for the synthesis of other phospholipids (Table 1). PA lipids are also involved in curvature regulation of membranes and act as regulators and substrates for many enzymes, e.g. enzymes involved in glycolipid and DAG synthesis in chloroplasts and phosphatidylylglycerophosphate (PGP) synthesis, which may then be dephosphorylated by PGPS phosphatases to **phosphatidylglycerol (PG)**. PG lipids are involved in the regulation of membrane-associated enzymes and are even the most abundant phospholipid of thylakoid membranes. In contrast to many other phospholipids, PG is synthetized in three locations in plants (ER, chloroplast or mitochondria) by three isoforms of synthases, emphasizing the importance of this lipid. PG is also the precursor molecule for synthesis of **cardiolipin (CL, or diphosphatidylglycerol)** in mitochondria and prokaryotes. Both PG and CL have anionic head-groups, and in bacterial cell membranes these two lipids may, as a response to environmental factors, substitute each other without any major effect on the membrane function. As with PA, **Phosphatidylserine (PS)** lipids are found in many cell membranes, but never in large amounts. It is likely the main precursor lipid in the synthesis of **phosphatidylethanolamine (PE)** lipids, the main component of the OM of mitochondria. PE has a small head group size compared to the major bilayer-forming lipid of most eukaryotic membranes: **phosphatidylcholine, PC**, which is likely the reason for some unique characteristics found in bacterial biomembranes and mitochondria. The trafficking of phospholipids is complex and involves vesicular pathways as well as non-vesicular lipid transfer between membranes. This interesting interplay is reviewed by Jouhet (2009) and will not be further discussed here, although a short summary of a few other phospholipids can be found in Table 1.
Table 1: Properties of some common lipids.

<table>
<thead>
<tr>
<th>Lipid</th>
<th>Charge</th>
<th>Head group</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>PA</td>
<td>-1</td>
<td>-H</td>
<td>Important for biosynthesis of other phospholipids. Trace amounts.</td>
</tr>
<tr>
<td>PC</td>
<td>0</td>
<td>-CH₂CH₂N(CH₃)₂</td>
<td>Main lipid of most eukaryotic biomembranes (excluding chloroplasts). Often a minor component of bacterial membranes.</td>
</tr>
<tr>
<td>PS</td>
<td>-1</td>
<td>-CH₂CH(NH₂)COO</td>
<td>Involved in cell cycle signaling/apoptosis of all cell types, present in low abundance.</td>
</tr>
<tr>
<td>PG</td>
<td>-1</td>
<td>-CH₂CH(OH)CH₂OH</td>
<td>Main PL in thylakoid membranes and many bacteria (sometimes modified). Involved in membrane interaction of proteins.</td>
</tr>
<tr>
<td>CL</td>
<td>-2</td>
<td>(di)phosphatidyl-glycerol</td>
<td>Unique for bacteria and mitochondrial IM. Have four acyl chains and two phospho-groups.</td>
</tr>
<tr>
<td>PI</td>
<td>-1</td>
<td>Myo-inositol</td>
<td>Phosphorylated PI lipids (PI(P)x) are present in many organelles and involved in the endosomal pathway.</td>
</tr>
<tr>
<td>PE</td>
<td>0</td>
<td>-CH₂CH₂NH₃</td>
<td>Main component of OM. Non-bilayerforming.</td>
</tr>
</tbody>
</table>

Based on “The Lipid Handbook”.

3.2.2. Three glycolipids dominate in chloroplasts

Glycolipids are mono-, di- or polyglycosylated lipids and constitute the main part of photosynthetic membranes in plants, algae and cyanobacteria, where phospholipids are often present in small amounts. In chloroplasts the biomembranes are composed of up to ~75% galactolipids, i.e. glycolipids with galactose sugars in their head-group regions (Figure 2), of three main types, described below. In plants glycolipids replace the structural and surface-charge properties of PC, PE and other lipids. They are responsible for both the structural scaffold for other biomolecules to anchor
into, and they define the boundaries of the chloroplasts but also act as ligands and substrates in enzymatic reactions. Since they carry out so many tasks \textit{in vivo} and can only be synthesized in the chloroplasts, they also need back-up systems and there are consequently several genes encoding these proteins present in most plant genomes. The expression of some of the synthases can rapidly change the lipid composition; this will be exemplified further in Chapter 4. The glycolipids are all based on diacylglycerol (DAG) from either the ER or from DAG produced in the chloroplasts (from PA lipids).

Monogalactosyldiacylglycerol (MGDG)

MGDG (Figure 2) is a non-bilayerforming lipid that may be important for stabilization and/or solubilization of proteins and protein complexes. MGDG has been shown to co-crystallize with proteins of the photosynthetic machinery and is thus likely to be involved in some way in photosynthesis.\textsuperscript{33,38} MGDG is also the precursor lipid for DGDG synthesis (i.e. DGDG is never synthetized from DAG directly or via other precursor molecules).\textsuperscript{39}

Digalactosyldiacylglycerol (DGDG)

DGDG (Figure 2) makes up a big portion of the lipid pool in chloroplasts, but in varying amounts between different tissues. DGDG easily forms planar bilayers and is likely important for the structural integrity of chloroplast membranes. DGDG may also replace phospholipids (Figure 3) in the cytoplasmic part of plant plasma membranes, mitochondria and chloroplasts.\textsuperscript{40-44} Even though MGDG and DGDG are only present in trace amount in mitochondria,\textsuperscript{30} mitochondria of \textit{A. thaliana} may alter the lipid composition when the plant is grown without P\textsubscript{i}, and the main change will be that the abundance of DGDG is strongly increased (on behalf of mainly PE),

Figure 2: Schematic chemical structures of MGDG (left) and DGDG.
from 1.5 % to 18.2 %. Some studies also point out other environmental factors that affect DGDG levels. Drought was shown to increase the DGDG:MGDG ratio in leaf plasma membranes, as well as nitrogen deficiency. DGDG levels has also been shown to play a role for the thermotolerance of *A. thaliana*. How all of this is connected is however not resolved. An enzyme involved in the synthesis of DGDG in the chloroplasts under phosphate-deprived conditions is presented in paper I-II and Chapter 4.

**Sulfoquinovosyldiaclylglycerol (SQDG)**

SQDG is a sulpholipid with anionic head group, and important for maintaining a stable anionic surface charge of chloroplast membranes. As with DGDG, SQDG may replace phospholipids (PG) upon phosphate shortage (Figure 3), hence keeping the levels of anionic lipids in chloroplast biomembranes stable. A certain level of anionic lipids has been demonstrated to be crucial for plant growth and photosynthesis, and this rescue-mechanism also illustrates the importance of anionic lipids in biomembranes.

![Figure 3: Changes in total lipid composition upon phosphate deprivation in *A. thaliana* leaf tissues. DGDG and SQDG both increase in concentration on behalf of phospholipids. Adapted from Härtel (2000).](image)

**3.3. Cell walls maintain the structural integrity of cells**

Cell walls share many of the functions that are also seen in biomembranes, such as recognition/receptor functions, permeability to solutes and
biomolecules and selective accumulation of molecules important for many physiological processes. They do however also have additional functions, most importantly they give structural support to cells and resistance to harmful external and internal influences.\textsuperscript{51} The main components of cell walls are endoplasmic carbohydrate oligo- or polymers.\textsuperscript{52} Long cellulose polymers are abundant in plants and algae, whereas shorter carbohydrate species (peptidoglycans in bacteria and pseudopeptidoglycans in Archea) are present as part of some organisms defense mechanisms (including biofilm formation).\textsuperscript{51,53} Fungi have chitin instead of cellulose, a carbohydrate that plants on the other hand are completely devoid of. It may be the main component of fungi cell walls (up to 60\% in \textit{Alomyces macrogynus}) or present in very small amounts (2\%) in the yeast \textit{Saccharomyces cerevisiae}.\textsuperscript{51}

3.3.1. Chitin is present in trace amounts in some oomycetes

Chitin is a polymer consisting of \textit{N}-acetylglucosamine moieties (derived from glucose, see figure 4) connected by ether linkages.\textsuperscript{52} Chitin is, together with cellulose, the most abundant polysaccharide in nature, nevertheless, little is known about its synthesis and interaction partners. The main research focus has instead been on its (promising) use as biomaterials. Chitin was isolated in 1811 from the cell walls of fungi. This initially gave it the name \textit{fungine}, although soon it was found to be the main component of the cuticle of beetles, from which it got its present name (from Greek chiton: “coat of mail”).\textsuperscript{54} It is not uncommon for eukaryotic organisms to carry chitin in their cell walls, but the amount certainly varies. Only 0.5 mol\% of the carbohydrates isolated from the cell walls of the oomycete \textit{Saprolegnia} species is chitin.\textsuperscript{55} This was a somewhat surprising finding as it was commonly assumed that oomycetes did not synthesize chitin. The chitin is likely involved in the structural integrity of oomycete cell walls as its tensile strength is superior to cellulose or other carbohydrate polymers.\textsuperscript{51,56} The need for increased mechanical support, hence regulation of enzymes involved in

\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{chitin_structure.png}
\caption{Schematic dimer structure of chitin.}
\end{figure}
chitin synthesis, could be distinctive depending on the present stage of the organism’s life cycle. It appears in shorter oligomeric forms that can cross-link to the much more abundant cellulose (amongst others) \textit{in vivo}.\textsuperscript{57} The formation of chitin is catalyzed by chitin synthases, as described in Chapter 4 and paper IV.

3.3.2. Lipopolysaccharides and gram-negative bacteria

Gram-negative bacteria (e.g. \textit{E. coli}) have unique cell walls composed of two types of bilayers: The Outer Membrane (OM) and the Inner Membrane (IM), separated by the peptidoglycan-rich periplasmic space (see figure 5). The IM has a similar function as the biomembranes typically seen in eukaryotic cells. The OM is important for nutrient uptake and acts as a shielding layer that protects the bacterium from destructive substances (e.g. antibiotics, detergent molecules). The OM is further unique in its lipid composition as it has two asymmetric leaflets; the inner leaflet is mainly composed of phospholipids whereas the outer leaflet is enriched in a specific type of glycolipids called \textbf{lipopolysaccharides (LPS)}. LPS is the secret behind the sturdiness often seen in gram-negative bacteria. They are composed of a hydrophobic lipid-anchoring part (Lipid A), an oligosaccharide core region and a distal and variable polysaccharide, the so-called \textit{O}-antigen repeat.\textsuperscript{31} The synthesis of LPS is conducted in the IM but the lipids are transferred to the OM by several transport mechanisms, including a periplasmic pathway and flipping of the lipid in the OM.\textsuperscript{58} Only the lipid A part and the very first sugar moieties (collectively described as \textit{kdo}) are needed for \textit{E. coli} survival \textit{in vitro}, the rest of the LPS molecule is involved in resistance against the immune responses, drugs and environmental factors. Chapter 4 includes a brief description about one of the enzymes responsible for one step of the LPS synthesis, WaaG (Figure 5), that was studied in paper III. The structure and synthesis of LPS in \textit{E. coli} and other gram-negative bacteria is reviewed in detail by Raetz and Whitfield\textsuperscript{59} and the synthesis, lipid export and virulence factors by Wang and Quinn.\textsuperscript{58}
Figure 5: The cell wall of gram-negative bacteria. There are several unique cell wall properties such as an asymmetric lipid composition (glycolipids/phospholipids) and exposed LPS lipids at the exterior of the cells, yielding a protective layer with nutrient-absorbing properties. The synthesis of LPS (here simplified for a schematic overview) is catalyzed by an array of GTs and also involves transfer from the cytosolic side of the IM to OM via large TM proteins (i.e. MsbA and Lpt enzymes, in orange). In paper III the GT WaaG (beige) is investigated, a GT that utilizes UDP-glucose from the cytosol (blue-yellow) and precursor LPS molecules in the IM to extend the LPS molecules before they are translocated to the OM.
4. Activity, structure and classification of glycosyltransferases

Synthesis of glycolipids and chitin (described in Chapter 3) involves a class of enzymes called glycosyltransferases (GTs, sometimes referred to as glycosylsynthases). In this section the general structure, classification and activity of glycosyltransferases will be described together with some specific lipid interaction details regarding GTs of relevance for this thesis. The common feature for the GTs in this thesis is that they are situated at or in the membrane.

4.1 Glycosyltransferase activity, structure and classification

In order to understand the classification and characteristics of GTs it is valuable to describe some background information. The GTs studied in this work utilize nucleotide sugars (i.e. UDP-sugars, Figure 6) as substrate donor molecules but use varying accepting molecules to generate a glycosidic bond between the donor and an acceptor molecule. If the GT utilizes hydrophobic or amphiphilic molecules (lipids/DAG) as acceptor substrates it is likely that they are localized close to biomembranes. Glycosyltransferases are mainly categorized according to their amino-acid sequence, but information from structure and the enzymatic mechanism is also useful. GTs function via either a retaining or inverting mechanism, where the main difference lies within the donor and acceptor substrate stereochemistry in the resulting molecule after the catalysis has taken place. The classification into families

Figure 6: A schematic chemical structure of UDP-galactose. GTs are involved in transfer of the sugar moiety to lipids, yielding glycosylated lipids (described in Chapter 3) such as MGDG and DGDG.
is a highly dynamic process and proteins are rearranged between families, as new information about GT structure and phylogeny is discovered. To date 99 GT families have been described in the CAZy database (visited 2016-05-24) and at least four folds have been attributed to GT proteins. The proteins investigated here are members of one of the two main folds: GT-A and GT-B. The number of genes that has been discovered (so far) that encode GTs has rapidly increased the last few years, but the number of solved (unique) structures has not followed in the same fashion. The 3D structures that have emerged lately have, however, been of huge importance for the understanding of this enzyme class and are collected in the Glyco3D database (see Chapter 7).

Enzymes involved in the synthesis of cellulose and chitin are members of the GT2 family with a **GT-A fold** (CAZy database, visited 2016-05-24). GT-A enzymes topologically consist of several domains, and can be soluble or membrane-bound. GT-A enzymes have a soluble catalytic domain, with a Rossmann fold and a vital metal ion bound by a DXD (or rarely EXD) motif, an conserved motif that positions the metal in a favorable orientation towards the nucleotide sugar (substrate donor) and is involved in the catalysis. A recent x-ray structure of the bacterial cellulose synthase (BcsA, PDB 4HG6) together with the growing cellulose chain and in complex with another protein (BcsB, located on the periplasmic side of IM) has recently been reported. Valuable structural and functional information about how carbohydrate polymers are produced and how they translocate through IM in gram-negative bacteria was obtained. The overall structure contains one TM domain and two soluble domains, one on each side of the bilayer. The catalytic domain is located on the cytosolic side where the synthesis of carbohydrate polymers takes place. The polymer is then transferred to the extracellular side through a pore in the TM domain. It is likely that chitin synthases and many other uncharacterized GT-A GTs share this type of assembly and translocation mechanism.

**GT-B enzymes** are both structurally and catalytically dissimilar to GT-A enzymes. As with GT-A enzymes they involve the typical nucleotide-binding motif (Rossmann fold), but instead of having several very dissimilar domains, they comprise two domains of similar topology. The two domains (in this work annotated as the *N-terminal* and the *C-terminal domain*) each comprise a double Rossmann fold, well established as a nucleotide-binding motif, and are connected via a flexible linker. A substrate-binding cleft is
located between the two domains, with amino-acid residues in the N-terminal domain involved in binding of the acceptor substrate, and in the C-terminal domain in binding the sugar donor.\textsuperscript{54,65} GT-B enzymes are either soluble or membrane-bound, but do not utilize TM domains but instead involve peripheral and/or monotopic interactions.\textsuperscript{60} Structurally GT-B enzymes are especially interesting since they are extremely diverse in their function (especially regarding the accepting substrate molecules) but share this very conserved fold, with only minor differences in the topology (further discussed in Chapter 6 and paper II). GT-B enzyme structure, function and membrane interaction modes were recently reviewed by Albesa-Jové et al. (2014).\textsuperscript{60}

In this thesis one GT-A (Section 4.2 and Chapter 6) and two GT-B enzymes (Section 4.3 and Chapter 5) have been investigated, belonging to the GT2 and GT4 families, respectively. To understand GTs it is important to both compare the characteristics within a protein family and fold, but also to extend the results and inspect structures and sequences of other families. The subject of GT-B enzymes in general is too broad to cover in this thesis, several review articles cover structure-function relationships of GTs.\textsuperscript{62,66,68}

### 4.2. A GT-A glycosyltransferase involved in chitin synthesis may act as a target for anti-oomycete compounds

As mentioned in Chapter 2, cellulose is the main component in the cell walls of a quite modest group of organisms called oomycetes, or water molds.\textsuperscript{55} The importance of cellulose for the formation of the cell walls (and vitality of the cells) is implied by the presence of multiple genes for the GTs responsible for cellulose synthesis in all oomycetes.\textsuperscript{69} Oomycetes are filamentous stamenopiles (see Figure 7); eukaryotic organisms related to brown algae and diatoms\textsuperscript{70} with a life cycle that puts high demands on rapid environmental adaptations of cell walls and biomembranes. Some oomycetes are severe pathogens in aqua- and agriculture, e.g. Phytophthora infestans; "late blight" on potatoes, Saprolegnia parasitica; "water mold" on fish (e.g. salmon, see figure 7), and Aphanomyces astaci; "crayfish plague" (or "kräftpest" in Swedish). Saprolegnia species are normally saprotrophic but may act as parasites on fish with compromised immune system (from pollutions/wounds). A compound called Bronopol\textsuperscript{71,72} has shown promising results as a broad-spectrum biocide, and has been able to replace the very
toxic (and banned) compound malachite green in control of oomycetes in aquaculture (against *S. parasitica*). There are however limitations with this compound and it is therefore useful to identify ways of oomycete control that are more species specific. Further ways of oomycete control in aquaculture are discussed in Brown (2015). The trace amounts of chitin discovered in some species, or the enzymes involved in its synthesis, may be a target, since the chitin quantities differ between oomycetes and the ChS present in oomycetes is different from most of those encountered in fungi. One way of selective oomycete control may therefore be to control or inhibit chitin synthases. The recent understanding that different species of oomycetes have such distinctive cell walls is also very interesting for the understanding of growth and lipid/carbohydrate adaptations and protein-lipid interplay in different stages in the life cycle of oomycetes (and other filamentous organisms).

Two chitin synthase genes (*smChS1* and *smChS2*) was recently isolated from the oomycete model organism *Saprolegnia monoica*. Chitin synthase activity was later demonstrated for the gene product of *smChS2* and it was demonstrated that the protein was mainly active in the hyphae tips of *Saprolegnia* during their vegetative phases. The fungicide nikkomycin Z was shown to inhibit the protein, which led to bursting hyphae tips, and

Figure 7: Blow fly (Calliphoridae) infected by the oomycete *Saprolegnia ferax* (“fiskmögel”). Photographed at the Nordic herbarium, Swedish Museum of Natural History (S), “no. 48 *Saprolegnia thureti*”, collected by P. Sydow (1897).
consequently death of the organism. The result from these studies, that presence of an active smChS2 enzyme is crucial for *S. monoica* and, even more so for *S. parasitica* viability, paves the way for development of less invasive pest control methods. Interestingly, ChS1 did not show synthase activity *in vitro*, which may indicate a requirement to be anchored to biomembranes or some interaction with ligands. Eventually it became clear that the presence of ChS genes is more common than previously anticipated, and that most oomycetes do have one or two genes for ChS. Not all have however been demonstrated as enzymatically active, at least not under the investigated conditions and life cycle stages.

4.2.2. MIT domains and lipid recognition of chitin synthases

The overall topologies of smChS1 and smChS2 are predicted from hydropathy analyses as typical GT-A folds, and the amino-acid sequence characteristics places them in the GT2 family, with an inverting mechanism and a large C-terminal TM domain composed of seven helices. Both smChS1 and smChS2 have a soluble N-terminal domain with a DXD-motif, but oddly there is also a domain of roughly 70 amino-acid residues that has not previously been described in GTs. Prediction tools placed this domain in the pfam 04212 group, i.e. Microtubule Interacting and Trafficking (MIT) domains (see Figure 8). MIT domains where described in 2003 by Ciccadelli

![Figure 8: The amino-acid sequences of ChS in *S. monoica*, investigated in paper IV, predicts several domains. In the C-terminus a large TM domain is present (orange), preceded by a catalytic domain (beige), and a MIT domain of unknown function. Previous work on MIT domains suggests several interaction mechanisms to other biomolecules.](image)
et al, who suggested the name of the domain based on sequence predictions and activity of several proteins (including spartin, spastin, vps4), suggesting a common function within the endosomal pathway and often found in microtubule-interacting proteins. MIT domains are commonly found in AAA ATPases, e.g. Vps4 proteins, where they interact with so-called MIMs (MIT Interacting Motifs) of their interaction partners. AAA ATPases are involved in many fundamental cell functions, where they adopt a chaperone-like function. Vps4 proteins are coupled to protein sorting and trafficking in endosome membranes where they oligomerize and bind to the endosome-bound ESCRT-III (Endosomal Sorting Complexes Required for Transport) complex via MIM motifs into functional assemblies (and binds additional ligands), acquiring ATPase activity. The assembly of the final vps4 oligomer structure and ATP leads to dissociation of parts of the ESCRT-III complex, and this in turn initiates vesicle formation. In 2011 Scott et al solved a solution NMR structure of the MIT domain of human VPS4A, showing the expected three-helix bundle with several Ala residues in helix 1 and helix 2, proposed to drive the tertiary structure formation and packing of the helices. MIT domains have so far only been demonstrated in eukaryotic cells (99.5%) and Archaea (0.5%), and were not previously demonstrated in any Stramenopile (including oomycetes) (SMART database accession code: SM00745). In paper IV and Chapter 6 the role of the MIT domains of oomycete ChS are discussed.

4.3 GT-B glycosyltransferases regulates the glycolipid levels in A. thaliana and E. coli biomembranes

Here two GT-B enzymes of the GT4 family are presented, both with important impact on lipid synthesis. Previous studies of GTs in the gram-positive bacterium Acholeplasma laidlawii have shown that the activity in vitro is regulated by anionic properties of the lipids and by biomembrane curvature. As many GT-B enzymes utilize substrates, or interact with other proteins, that are membrane bound or produce components destined for the biomembrane, it is not surprising that there is a need for tight regulation of these enzymes. Sudden changes in the environment require strategies for survival, and these need to be fine-tuned according to the organisms needs. In this section a phosphate-rescuing mechanism, using DGDG synthases, found in several plants, and the synthesis of the protective LPS by WaaG in E. coli is described. GT-B enzymes are very conserved in their three-
dimensional structure although they share low sequence identity and utilize a diverse array of substrate and donor molecules.\textsuperscript{66,68} It is important to characterize as many proteins as possible of this class, in order to understand key regions and events. The structural characteristics and lipid-interacting properties of atDGD2 (described below) and other GTs are very important for a general understanding of membrane-interacting GTs and lipid synthesis.

4.3.1. Phosphate starvation triggers DGD2 activity in \textit{A. thaliana}

Synthesis of all chloroplast envelope phospholipids occurs at several locations.\textsuperscript{33} The distribution of glycolipids in plants varies between organelles and tissues, and most interestingly also with growth conditions (Figure 3), as mentioned in Chapter 3. MGDG synthases (GT4) utilize Diacylglycerol (DAG) and UDP-galactose (Figure 6) to generate MGDG (Figure 2),\textsuperscript{83} which can be either transferred within the chloroplast membranes, or be used as substrate molecules for two DGDG synthases, together with additional UDP-Gal molecules.\textsuperscript{40,84} \textit{A. thaliana} is a small angiosperm of the Brassicaceae family, often regarded as weed but also well-known in its role as a model organism in genetic studies. The synthesis

![Diagram](image-url)

Figure 9: During normal growth conditions the bulk galactolipid synthesis in photosynthetic tissues involves two main GTs: MGD1 and DGD1. The substrate for MGDG synthesis (DAG) may be either imported from ER or synthetized by other enzymes in the inner envelope membrane. The products are transferred within the chloroplast membranes and thylakoids via largely unknown mechanisms. Adapted from Shimojima (2011).\textsuperscript{39}
of MGDG and DGDG during normal growth conditions is displayed in Figure 9 and involves the GT4 enzyme MGD1 and DAG derived from either the ER or chloroplasts, whereas the isoform MGD2/3 (Figure 10) uses DAG from the chloroplast alone. This will eventually lead to three MGDG pools, and consequently DGDG, with slightly different characteristics of the hydrocarbon tails dependent on organelle/tissue.

Environmental factors may disrupt the normal pathway of DGDG synthesis (Figure 9), leading to another scheme (Figure 10) where DGDG synthesis is catalyzed by DGD2 synthases instead of DGD1, especially in non-photosynthetic tissues (roots, stems etc.). DGDG produced by this other subset of enzymes is exported to membranes enriched in phospholipids, instead of the glycolipid-dominant thylakoid membranes. This may give the plants opportunities to recycle the phosphorous bound in phospholipids, and direct it to pathways involved in vital DNA and ATP synthesis. This is a very important survival mechanism observed in *A. thaliana* as well as many other plants. For instance, it has been demonstrated that if the plant experiences phosphate shortage, the amount of DGDG in leaves will increase from 10 to almost 30 %, whereas root-derived DGDG levels are elevated even more, relatively (from a few percent to 15 %).

![Figure 10: In roots MGDG and DGDG is synthetized by MGD2 or MGD3 and DGD2 GTs. DGD2 utilizes MGDG from those processes and the UDP comes from the ER exclusively, perhaps demonstrating a vital role for PA in chloroplasts. Adapted from Shimojima (2011).](image-url)
4.3.2. The GT-B Glycosyltransferase WaaG is involved in formation of shielding carbohydrate layers in *E. coli*

LPS on the extracellular surface may act as defense structures and a matrix for nutrients and is therefore involved in the viability and growth of many gram-negative bacteria (Chapter 3). Due to the polymeric structure of LPS molecules (illustrated schematically in Figure 5), several GT enzymes are involved in the synthesis. This work regards one of those, called WaaG (previously RfaG). WaaG is a retaining GT-B enzyme of the GT4 family, involved in an intermediate step in the LPS synthesis, where a glucose moiety is attached to the growing core region. If inhibited, the organism will be weakened, illustrating the enzyme’s importance for development of novel drugs and for the understanding of bacterial cell walls. The subcellular localization of WaaG is in the inner leaflet of the inner membrane on the cytosolic side. Still, since the emerging LPS is transported, gene deletions of WaaG in *E. coli* have been shown to destabilize the OM. In *Salmonella enterica* it has also been shown that deletion of WaaG interferes with the bacterium’s capacity to infect epithelial cells. Recently it was demonstrated that PG, CL and the detergent CHAPS could increase activity in WaaG, and in the same study the authors investigated potential drug targets against WaaG activity, since LPS increases cell viability. Hence, a detailed understanding about cell wall synthesis and glycosyltransferases involved is needed for development of novel antibiotics.
5. Monotopic GT4 glycosyltransferases share structural features coupled to their lipid interacting mechanisms

This chapter contains a summary of my work on GT-B enzymes in which some results and additional details that emerged from studies of atDGD2 and WaaG will be presented. We have studied lipid interacting properties of two monotopic enzymes of the GT4 family, derived from two very different organisms, digalactosyldiacylglycerol synthase 2 (atDGD2) from the angiosperm *A. thaliana*, and a lipopolysaccharide core biosynthesis protein (WaaG) from the gram-negative bacterium *E. coli*. By the use of biophysical methods (mainly spectroscopic), tailored membrane mimicking systems (i.e. micelles, LUVs and bicelles) and sequence analyses we elucidated the roles of the two domains of these proteins. Some of the methods used are further described in Chapter 7.

5.1. atDGD2 and WaaG anchors to one leaflet of the bilayer with parts of the N-terminal domain

As described in Chapter 4, DGD2 synthases are involved in synthesis of glycolipids mainly under phosphate-deprived conditions, and the localization of DGD1 and DGD2 differ. To investigate membrane interactions of atDGD2 short segments from the full-length protein where selected (Table 2, Figure 11) based on sequence analyses, especially results from a novel MVDA analysis, further presented in paper I and Chapter 7. Four regions of the atDGD2 N-terminal domain were identified as potentially lipid-interacting. One of the atDGD2 peptides, denoted S169-187, was early on in an *in vitro* experimental setup found to have lytic properties on vesicles and was observed also during the current project to undergo severe aggregation, probably because of its hydrophobic nature. It was concluded from these observations that S169-187 is likely to be integrated into a hydrophobic milieu (or involved in dimerization events).
Another segment close in sequence, S130-148, is also hydrophobic and was shown to transform into an α-helical structure at 30 % PG, an interesting observation since large conformational changes recently has been shown to occur in other GTs, as discussed in paper II. Two regions (S11-29 and S46-64) did not seem to interact with lipids, consistent with results from studies on other GT4 enzymes, where these regions topologically are located far from the lipid surface. It was also shown that there are unusually many Trps in the proteins sequence, and mutation studies indicated a crucial role for each one of them.94 They were also all proposed to be surface exposed, although the homology model structure used in paper II indicates that some are buried (i.e. W77, W132 and possibly W48), and are likely involved in the tertiary structure packing rather than intermolecular interaction events.
The Trps in segment S130-148 and S169-187 are however exposed, and both W139A and W177A protein variants were previously shown inactive.

In WaaG a three-helix region containing 30 amino-acid residues (Table 2) in the N-terminal domain, called MIR-WaaG (Membrane-Interacting Region of WaaG), was selected based on sequence analysis and the 3D structure (Figure 12 and 11, respectively). MIR-WaaG contains many cationic and Tyr residues, which may be involved in lipid anchoring. In paper III we determined a solution NMR structure of MIR-WaaG incorporated in micelles (PDB 2N58) and noted that it forms a three-helix structure similar to the one seen in the full-length protein, and that structure induction in vitro.

Figure 12: The region where atDGD2 and MIR-WaaG overlap in a global alignment with several membrane-associated GT-B proteins with known 3D structures (from paper I), showing the helical (dark boxes) dominance in parts of the sequences. WaaG is annotated RfαG in this figure, regions with high MPex scores are underlined in red.
was dependent on the presence of LUVs with 20-30 % anionic lipids (or detergent molecules). We investigated the charge-dependence of the lipid interactions further by several methods and also investigated differences in binding to PG vesicles and E. coli native-like vesicles with PE/PG/CL lipids. By a step-wise change in the lipid composition we obtained information about the driving force behind the interactions. We noted that, in our systems, electrostatics were important for the interaction, and that other lipid properties seemed less relevant since PE/PG/CL bicelles in general resulted in very similar results as with PG only. The strong association of MIR-WaaG to anionic lipids was detected by many complementary techniques (fluorescence quenching, anisotropy measurements) and NMR experiments (PFG diffusion). By the combined use of several NMR and fluorescence based methods we could estimate the depth of the incorporation of the individual amino-acid residues (i.e. parts of the peptide), either by information gained from side-chain atoms (CLEANEX NMR, fluorescence quenching) or backbone Cα (by the use of a novel “DPC ruler”, based on paramagnetic relaxation experiments). We noticed that the side-chain atoms of all Lys residues reached into the interface region, and the Tyr residue side-chains seemed to be exposed to the solvent. From the results we could estimate the approximate localization of MIR-WaaG in/towards the biomembrane, and consequently the orientation of the entire protein.

5.2. The C-terminal domain of atDGD2 display additional lipid interacting properties

All but one segment investigated in paper I is located in the N-terminal domain. One segment from the C-terminal domain did however also display a very strong binding to anionic vesicles, along with apparent conformational changes in CD spectra, and therefore this segment (S227-245) was characterized further by NMR and CD spectroscopy. From studies of S227-245 together with different types of anionic LUVs it appeared that the segment could alter its secondary structure upon interaction with PG lipids, an interesting property as that type of conformational changes are not trivial, typically requiring rearrangement of solvent molecules and intramolecular hydrogen bonds, as well as salt bridge dissociations. A solution NMR structure of S227-245 in DPC micelles demonstrated an α-helical structure with a clear amphipathic character, hence possibly involved in peripheral or integral biomembrane interactions. This is odd since this
region, as visible in homology models, is localized in the Rossmann motif as part of the β-sheet. This dual structural behavior may be important for the protein in recognition of substrate molecules, i.e. the large MGDG lipids that likely require conformational changes in order to fit into the active site. Large conformational changes have also been demonstrated for other proteins (see Chapter 2), as is also discussed in paper II. Information about the characteristics of such a mechanism cannot be obtained from present data, but it is clearly an important property to look for also in other GT-B enzymes, especially those of the GT4 family. Differences in the amino-acid composition of the linker elements located in between the two Rossmann motifs could also be indicative of specific substrate binding requirements and may regulate the domain movements by different sequence lengths and degrees of flexibility. This could be coupled to the conformational plasticity seen in the segment next to the linker, i.e. S227-245. Indeed, there seems to be variance in the linker lengths of different GT4 enzymes, which may then reflect on the substrate size and/or hydrophobicity.

In paper II we expanded the studies of the C-terminal domain by inclusion of the segments S240-258 and S269-287 (Table 3, Figure 11 and Figure 13). The aim was to explain the role of two seemingly surface-located and strongly charged regions, that both appeared to align close to the lipid binding interface and part of the Rossmann fold helices rather than the β-sheet. S240-258 was shown to bind to anionic lipids and upon interaction

![Figure 13: Two C-terminal domain segments in atDGD2 (front and side view) have strong ionic character that is exposed in the structure. Based on results from paper I S240-258 (yellow) orients towards the bilayer, whereas S269-287 directs many anionic residues towards the solvent. Both are connected in sequence and may undergo conformational changes in response to lipid binding.](image-url)
attain helical structure, with an obvious amphipathic character (based on sequence analysis). Due to the apparent functional diversity of the C-terminal domain of GT4 enzymes it is likely that similar lipid-interacting regions will be revealed in other monotopic GTs, along with other types of regulative function. Based on current results for the S269-287 segment it is not possible to draw any conclusions about how this part of the protein may function, however, from our results we conclude that this region is not involved in lipid binding. Since there are a few sites that are amenable for phosphorylation (Table 2) one may however speculate that perhaps this domain is involved in regulation of the enzyme. Dephosphorylation could also be involved in release of phosphate as part of the phosphate rescue mechanism, or trigger conformational changes in the C-terminal domain. Based on the model structure, the side-chain of E275 in S269-287 is a possible candidate for salt bridge formation with K246 in S240-258. Conformational changes could position these helices temporarily in such orientation towards each other that the charges could neutralize each other. This could also affect the amphipathic helix and position the hydrophobic part of it towards the lipids, either in order to attach the MGDG substrate, or to anchor the enzyme deeply into the outer leaflet of the bilayer. Also, in some inverting GT-A enzymes the acceptor binding involves stacking interactions of Trp residues between the donor and acceptor substrates, i.e. via the sugar moieties, aided by anionic residues that hydrogen bond to the pyranose rings and water molecules. The strong anionic character of S269-287 and the closeness to the functionally important W241, together with a glycosylated substrate molecule (MGD) may indicate a similar mode of action.

5.3. Topology comparisons of GT4 enzymes reveal common features

In paper I and II two different homology model structures of atDGD2 were used since it was difficult to detect the expected topology in the first structure. With the appearance of several novel three-dimensional structures, it is likely that the model structure in paper II is based on a more relevant scaffold than the one used in paper I. A comparison of the two model structures shows some regions that are not very similar, which were also less well defined in the first model. In paper II a topology analysis was done for some GTs and some details especially about the lipid-binding region in the
N-terminal domain are discussed. Many similarities between several GT4 enzymes were detected, e.g. conformational changes and membrane-binding properties. Upon closer inspection it was revealed that also the region including helices α2-3 in the N-terminal Rossmann motif (Figure 14), comprising the atDGD2 S46-64 segment (unstructured in vitro, paper I) show structural plasticity and this region has in studies of the GT4 enzyme PimA been suggested to be involved in binding to anionic lipids via cationic residues as an amphipathic helix. In WaaG this region is unstructured, and in a very recent 3D structure of the GT-B enzyme MGD1, involved in synthesis of the precursor lipids for DGD synthases as described in Chapter 4, this region could not be structurally determined.

As expected with enzymes functioning at protein-bilayer interfaces, MGD1 also displays a monotopic membrane interaction via the N-terminal domain. Another GT-B enzyme involved in LPS synthesis, WaaA, recognizes an early part of the LPS molecule (kdo) via a loop in the C-terminal domain. The membrane association is proposed to occur via cationic patches in the N-terminal domain upon conformational changes due to substrate binding. Based on the emergence of novel GT-B structures it becomes clear that both eukaryotic and gram-negative bacteria share such membrane-interacting behavior. As a comparison, cell walls and membranes of gram-positive bacteria do not contain LPS, but instead contain highly anionic molecules, the most important being teichoic acid, which can be covalently linked to a glycolipid to give lipoteichoic acids. A crystal structure was solved for a TagF polymerase from Staphylococcus epidermidis, a GT4 enzyme involved in teichoic acid polymer synthesis. The structure displays an N-terminal helical region that may interact monotonically with the lipid bilayer, and the authors note that this gives the active site an appropriate distance to the lipid substrate.
Figure 14: Topology diagrams for A) atDGD2 (based on paper I-II results) B) WaaG (based on PDB 2IW1), C) apo-PimA from *Mycobacterium smegmatis* in extended conformation (PDB 4NC9)
5.4. Conclusions and future prospects

We have demonstrated that both GT4 proteins investigated in papers I-III have similar membrane-interacting properties in the N-terminal domain, and a slightly different mechanism from some other previously studied GT-B enzymes, e.g. alMGS, a GT4 enzyme with surface-located Lys/Arg pairs in an amphipathic helix, located in the N-terminal domain and MurG (GT28 family), which anchors to biomembranes via surface-exposed Trps. The interaction of the N-terminal domains of WaaG and atDGD2 involves a hydrophobic segment with key aromatic and hydrophobic residues that anchor into one leaflet of the lipid bilayer, but also requires electrostatic components, manifested by the interplay between cationic amino-acid residues and anionic lipids. The C-terminal domain seems to govern regulation of the binding or other interaction events. More studies will be needed before functional details regarding this domain come to a consensus. The results from paper I-III have, however, made it possible to present a general hypothesis regarding monotopic bilayer interactions that is plausible also for other monotopic GT4 enzymes. A detail that we cannot resolve based on current data but that would be valuable to investigate, regards the behavior of atDGD2 when it is first presented to a lipid bilayer, and whether or not the electrostatic interaction of the C-terminal domain is preceding the integral/electrostatic binding of the N-terminal, or vice versa. The C-terminal binding may be what initially draws the protein to the anionic membrane surface, hence positioning the protein correctly towards the bilayer surface for further insertion by the hydrophobic “horn”. The anchoring of the N-terminal domain could also be the initial event that positions the active site and C-terminal lipid interacting region in a favorable manner. It seems that the two domains of GT-B enzymes have evolved for different purposes, and that the differences in substrate and membrane binding behavior amongst GT4 family members could be explained by differences in the C-terminal domain. Since most current knowledge about GT-B enzymes regards the direct enzymatic activity or the general structure/fold, it is important to also focus on such less general features in order to advance in the understanding of this class of proteins. The results from our studies have widened the current knowledge about possible structure conversions coupled to biomembrane interactions and have made it possible to propose regions that may be relevant for GT-B enzymes in lipid interactions. This novel information will also simplify the design of future studies regarding other membrane-associated GTs.
As mentioned, further studies on GT-B enzymes of the GT4 family should focus on the role of the C-terminal domain and the lipid-protein interplay. The tertiary and secondary structure conversions that seem to take place for many proteins of this family will increase general knowledge about membrane protein folding and function. Much is known about binding of the donor substrates (i.e. the nucleotide sugars), but less about the acceptor substrate binding sites and how ligand binding may alter the structure and lipid binding properties. The conserved fold of GT4 enzymes is very fascinating from a physicochemical point of view and the more understanding we obtain about the physical events that define the structure and lipid interaction events of these large and poorly characterized proteins, the easier it will be to tailor novel experiments. Every single bit of information may also be used to fine-tune the somewhat immature classification system of these proteins.

Even though our main focus in this work has been based mainly on a wish to understand the underlying physicochemical motifs of protein-membrane interactions, the practical applications that can be developed from this new knowledge are potentially important. Recycling of phosphate and other nutrients is a tricky issue especially in organic farming where the use of inorganic fertilizers is avoided. Fundamental knowledge about the role of DGDG synthases may explain what protects the plants from damage due to nutrient starvation, and possibly this information could also be useful for development of less ecologically invasive methods in agriculture. As the climate concerns are very much reality at this date there are demands for clever solutions in agriculture. The trend of overuse of antibiotics and fertilizers has led to tragic consequences that cannot be neglected. Proteins involved in the cell wall formation of gram-negative bacteria may hold the key to development of directed and less aggressive drugs.
6. Chitin synthases in oomycetes interact with lipids via a novel interaction mode, involving MIT domains

In contrast to the before mentioned papers, paper IV regards a GT-A glycosyltransferase, a chitin synthase involved in the assembly of oomycete cell walls. Although GT-A proteins have TM domains, the question examined here was not the interaction with the bilayer. As described in Chapter 4 two ChS enzymes derived from the oomycete *S. monoica* were previously predicted by sequence analysis to contain MIT domains, and in paper IV the aim was to elucidate the role of these, and to determine the three-dimensional structure of these additional domains, both in order to verify the fold and to be able to investigate properties and interaction surfaces. MIT domains from ChS1 and ChS2 (here denoted MIT1 and MIT2, respectively) from *S. monoica* were cloned, expressed and structurally characterized. A solution structure of MIT1 was determined and the results were used to also obtain a homology model structure of smChS MIT2 domain. The finding of such domains is very interesting since it is the first experimentally verified MIT domain structure from any glycosyltransferase.

We were able to demonstrate specific lipid interacting properties of both MIT domains, and additional binding of MIT2 to proteins involved in endosomal trafficking of proteins. As mentioned, GT-A enzymes are permanently anchored to biomembranes through integral domains, which is likely the case also for the proteins studied here. The results from paper IV do, however, indicate an additional lipid interaction mode that is unlikely to be purely electrostatic as interactions between different anionic lipids were not similar. Both MIT domains demonstrated strong binding to PA lipids, whereas the interaction to PS was weaker. Binding to phosphorylated PI lipids also showed variance that could not be contributed to a purely anionic character, as monophosphorylated PIP lipids were stronger binders than those that are phosphorylated in two or three positions. There must therefore
be some structural or stereochemical requirement on the interaction of the MIT domains that we have not been able to elucidate in this work. The specificity in binding to monophosphorylated PIP lipids is intriguing, since the role of MIT domains normally involves protein interactions within the endosomal pathway. Different organelles and vesicular assemblies, as well as the plasma membrane, are known to contain varying amounts of each PIP lipid type, hence acting as potential recognition partners that guides the proteins. Monophosphorylated PI(5)P showed the strongest binding, followed by PI(4)P (main localization in golgi) and PI(3)P (main occurrence in early endosomes). Some PI-related functions are mediated by direct interactions between phosphoinositides and proteins either via basic amino-acid stretches or globular structural domains. PI-binding domains may be involved in direct membrane anchoring of proteins, but PI-binding domains are also speculated to disturb cellular functions by either disrupting biomembrane assemblies or to interfere with other proteins lipid binding to PI.\textsuperscript{101}

A comparison can be made with MIT domains derived from other organisms and protein types. As described in Chapter 4, Vps4 proteins are coupled to protein sorting and trafficking in endodomes, whereas katanin is a heterodimeric protein that has a microtubule degrading mechanism, important during cell division.\textsuperscript{102} The NMR structure of an N-terminal domain from the enzyme katanin p60 has been determined,\textsuperscript{103} and although the sequence similarity was very low (20\%) it adopted a fold very similar to known MIT domains, and also displayed a cationic patch (formed by helix 2 and 3) that was proposed to be involved in tubulin binding. As described in Chapter 4, the MIT domain of AAA-ATPase VPS4 proteins interacts with MIM domains, and regulation of this interaction has very recently been attributed to a specific Lys residue in helix three. The authors suggest that PTMs of the Lysine may modulate the MIT-MIM interactions, possibly in parallel with other PTMs of close-by residues.\textsuperscript{104} A small MIT-containing protein was recently structurally and functionally characterized and shown to dimerize with a cationic patch exposed where three Arg were involved in binding to anionic lipids,\textsuperscript{105} possibly stabilizing the ESCRT complex. Clearly the role of the cationic pattern seen in MIT domains is either very diverse in function between proteins, have multiple roles or are poorly characterized to date.
It is possible that helix one and two of the MIT structures in *S. monoica* ChS are responsible for packing of the main scaffold and generation of the three-helix bundle via hydrophobic amino-acid residues in the core region, and that helix three, that also displays the largest variation in orientation, is responsible for protein-specific interactions. Charge-smoothed potential representations of both MIT1 and MIT2 (Figure 15) show that helix three has an electrostatic range from highly cationic (the region directly after the loop connecting helix two and helix three, including amino-acid residues K48, R50, K56 and K58) to highly anionic (the very C-terminal region of helix three, including amino-acid residues E62, E65, E69 and E73). The modeled structure of MIT2 shares similar characteristics but with less pronounced charge polarization. The charge-dependent lipid interactions seen for both domains in the lipid binding assays are likely to take place via the cationic patches although present work cannot reveal such detailed interactions patterns. Kuang et al. (2016) modeled MIT1 towards PC and PA membranes, and found several binding possibilities between MIT1 bilayers. One of the suggested lipid binding modes calculated for PA membranes included the cationic motif also seen here. Further experimental work on the lipid interactions of chitin synthase MIT domains will reveal more details regarding these processes.

Based on our results we conclude that the MIT domains of oomycete ChS may play a role in the intracellular Chs transport, possibly within the endosomal pathway, which leads to chitin accumulation in the hyphae tips of the growing mycelium in the asexual phase of the organism. Although the TM domain is responsible for the anchoring to biomembranes, there are lipid-binding events that may guide the protein to specific locations and perhaps regulate protein interactions, or to give stereospecific alignment to the soluble catalytic domain.

Benefits of aquaculture is easy to understand in some aspects, but the rapid increase in the demanded quantities has led to problems that is not solved in any environmentally sensible way. Present pesticides (or anti-oomycete drugs) are harsh on the surrounding waters and not optimal for continuous use. Chitin synthases involved in growth of oomycetes may be targets for novel compounds, in search for the ultimate solution for humanity, environment and animals.
Figure 15 A) The solution NMR structure of MIT1, shown in cartoon and surface charge representation (Helix 2-3 interface and helix 1 interface, respectively). Blue color indicates cationic surface and red indicates anionic surface charge. The polar character in helix three is eligible for interactions with other domains of ChS or other proteins, or with lipids. In B) an overlay is shown of the MIT1 solution structure (ice blue) and the homology model of MIT2 (green). The polar property of helix three seems less pronounced in MIT2.
7. Methods for studies of glycosyltransferases

As demonstrated earlier several methods have been used for studies on structure and protein-lipid interactions of atDGD2, WaaG and smChs. We have obtained detailed information thanks to spectroscopic methods such as **Nuclear Magnetic Resonance** (NMR), **Circular Dichroism** (CD) and **fluorescence spectroscopy**, often complemented with **light scattering** experiments. Depending on the system and choice of method several types of tailor-made **membrane-mimicking systems** (vesicles, bicelles and micelles) have been utilized to obtain information on the interplay between proteins and specific lipids. **Amino-acid sequence analyses** have been extremely useful early on in experimental setups and planning of research projects, but also to summarize and aid in interpretation of the experimental results. Here none of the methods will be described in detail; the aim is instead that this chapter will highlight strengths with the methods we have chosen and possibly clarify some details about the choice of methods and sample conditions used in the articles. The large size of membrane-associated GTs makes them difficult to structurally characterize **in vitro**, and the lipid-interacting property means that they may be very difficult to keep in solution and also that they may generate complex assemblies. It is therefore convenient to simplify the systems. As is often the case also in practice, I will begin with a description of some online tools that focuses on sequence-based predictions of protein secondary structure, fold, lipid binding and other physicochemical properties. I will then continue with a discussion about membrane-mimicking systems, and thereafter a description of the structure-based methods used in this work (CD and NMR spectroscopy) will be discussed. In the end I will describe how I have investigated protein-lipid interactions by mainly fluorescence and NMR spectroscopy.
7.1. Sequence analysis and *in silico* methods to characterize membrane-associated GTs

Studies of smaller segments (peptides or domains) in large proteins are useful since very much of the information about function and structure lies within the amino-acid sequence. The use of smaller polypeptides also simplifies data interpretation from methods where many components contribute to the signal. One example is in the commonly used far-UV CD spectroscopy (described below) where it is difficult to deduce how defined regions of a protein behave structurally based on the CD signal from the full protein. The peptides and domains investigated here have been chosen based on amino-acid sequence analyses and current knowledge about the proteins, for instance known ligands, active sites or a conserved protein fold. We have used several prediction tools and *in silico* methods to predict characteristics of atDGD2, WaaG and the MIT domains. A simplified scheme is depicted in Figure 16. The methods have two main purposes: **1) Prediction of secondary structure, fold recognition and homology modeling** in order to evaluate protein structure and/or simplify in the selection of segments for

![Diagram](image)

Figure 16: A scheme for using *in silico* tools in studies of glycosyltransferases. The route for investigations of the MIT domains (paper IV) is depicted in green, the route for investigations of GT-B enzymes (paper I-III) is shown in red.
1) Secondary structure, domain and fold prediction has been useful for us in order to rapidly gain structural information or for use where experimental methods have not been successful. Protein sequences can in principle be modeled *ab initio* into a native, intrinsically determined structure since specific amino-acid stretches have different propensities to form a particular secondary structure element.\textsuperscript{107,109} Assembly of secondary structure elements into a tertiary structure can often (but far from always) be determined. GTs have very low sequence identity which makes *ab initio* modeling difficult, but thanks to the limited number of folds that GTs adopt is that it is possible to cut corners through homology modeling, where experimentally solved structures act as structural scaffolds.\textsuperscript{57,110,111} Homology models together with sequence alignments with other similar and structurally characterized GT-B proteins aided in the selection of regions in atDGD2 that could possibly be excised without disruption of the tertiary structure. Homology models were also helpful in the topology analysis of atDGD2 in paper II. In addition, since it was experimentally challenging to produce samples suitable for NMR spectroscopy for the MIT2 domain, a homology structure was solved based on sequence homology with the MIT1 domain (paper IV). It was initially so, as mentioned in Chapter 4, that the presence of MIT domains in the ChS proteins was detected via a sequence-based domain prediction, that consequently indicated novel function and interaction partners for this type of domain. Fold and domain prediction of GTs is often successful since GTs are so structurally conserved. GT-A enzymes (e.g. ChS enzymes) are TM proteins and as such the membrane-spanning portion of the polypeptide chain may be predicted based on hydrophobic surface properties along with clustering of charged (cationic) and aromatic amino-acid residues in the lipid interface.\textsuperscript{10,109,112} For monotopic and peripheral membrane proteins the structural information is mainly achieved from comparisons along with soluble GTs (or other protein classes of the same general fold).

2) *In silico* methods to detect putative local function and membrane interaction regions

The preliminary structure information obtained from *in silico* tools makes it possible to obtain further information about the protein, as was the case of the ChS MIT domains. The domains were earlier predicted from sequences,
and we could experimentally verify their presence (paper IV). MIT domains had not been encountered in chitin synthases before, and their presence indicates protein function within the endosomal pathways where MIT domains are normally found. This was the basis for further biochemical studies on lipid and protein binding (PIP strips, Y2H). For studies of GTs a variety of uncomplicated tools and databases were useful early in the experimental setup. These are listed in Table 3, and here I will mention some methods and databases of special interest for studies of membrane interactions of proteins.

Information about GT families and 3D structures is easily obtained from GT-specific databases, which gather genomic, structural and biochemical information. Use of global sequence alignments is standard protocol as conserved residues inform about regions involved in ligand binding, enzymatic activity or association with bilayers. Some important amino-acid residues may, however, be overlooked since they are not considered to be conserved, e.g. Trp and Tyr, although both are often found in the membrane interface and probably carry a similar function in membrane-associations of integral proteins. A careful inspection of such key residues in global alignments may give valuable information, as also discussed briefly in paper II. It is also useful to use global alignment tools to search for regions that are not conserved. Such analysis may indicate regions that are in loops, linker regions or, as for Rossmann fold proteins, located in between the β sheets and possibly important for function. Globally, it is also useful to check for so-called low complexity regions (LCR), as described in paper II. LCRs are regions enriched in Arg, Lys, Glu, Pro and Ser. Identification of LCRs is useful to detect important enzyme-specific regions as proteins may have unique sequence features in regions that are quickly evolving. For GTs, with diverse substrates and interaction mechanisms, local alignments may be useful for comparison of smaller regions, since it is not uncommon with large sequence alternations such as domain swaps. Local alignments aided in analysis of different GT4 enzymes in paper II.

atDGD2 could in paper I be preliminary determined to orient towards the lipid bilayer based on results from several in silico tools and inspection of a homology model, in combination with experimental results. These parameters were based on hydrophobicity, electrostatics and the presence of seemingly surface-exposed Trps. Hydropathy plots of the full sequence give valuable information regarding membrane interacting regions of
integral (both TM and monotopic) proteins.\textsuperscript{116} In paper I a novel MVDA \textbf{algorithm} was presented for improved prediction of membrane anchored GT-B enzymes and regions likely to interact with lipids. First several GT sequences (including atDGD2) where separated as either membrane bound or soluble. Each amino-acid in the protein sequence was scored according to hydrophobicity, size and electrostatics and each parameter was compared to the respective positions in the other GTs within the same class (i.e. either soluble or membrane-associated). Regions of a defined length (22 amino-acids) were scored depending on their amino-acid properties, hence likelihood to interact with membranes. Based on a plot of the MVDA scores against the sequence of atDGD2 we could overview regions that were potentially membrane-interacting. This approach gave better results than conventionally used algorithms in predicting lipid-interacting regions of atDGD2. The \textbf{hydrophobic moment} is a measure of the amphiphatic character of a helix, defined by the vector sum of the individual side chain’s hydrophobicities.\textsuperscript{117} The larger the value, the more amphiphilic the helix potentially is, hence likelihood to be located in the bilayer-protein interface. Hydropathy scores can therefore be used to predict in-plane interacting helices (see Chapter 2). On-line tools that give information about helical content and hydrophobic moments of small peptides and generate \textbf{helical wheels} have aided in the understanding of membrane-interacting regions of atDGD2, and were used in comparison with other GT-B enzymes.
Table 3: Some tools used in this work for studies of glycosyltransferases

<table>
<thead>
<tr>
<th>Name</th>
<th>Short description</th>
<th>Webpage/references</th>
</tr>
</thead>
<tbody>
<tr>
<td>CaZy database</td>
<td>Updated information about GT families (and related enzymes).</td>
<td><a href="http://www.cazy.org">http://www.cazy.org</a>&lt;sup&gt;118-120&lt;/sup&gt;</td>
</tr>
<tr>
<td>Glyco3D</td>
<td>PDB codes and easy overview of 3D structures within a certain GT family.</td>
<td><a href="http://www.cermav.cnrs.fr/glyco3d">http://www.cermav.cnrs.fr/glyco3d</a>&lt;sup&gt;121&lt;/sup&gt;</td>
</tr>
<tr>
<td>SMART pFam</td>
<td>Domain analysis</td>
<td><a href="http://smart.embl-heidelberg.de">http://smart.embl-heidelberg.de</a>&lt;sup&gt;122&lt;/sup&gt; <a href="http://pfam.xfam.org">http://pfam.xfam.org</a>&lt;sup&gt;123&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Homology modeling</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Robetta full-chain protein structure prediction server</td>
<td><em>ab initio</em> modeling or modeling based on a protein scaffold is possible.</td>
<td><a href="http://roberta.bakerlab.org">http://roberta.bakerlab.org</a>&lt;sup&gt;124&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Partial sequence analysis</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ialign</td>
<td>Local structure alignment.</td>
<td><a href="http://bioinfo.wisc.edu/biolinux/docs/ialign.html">http://bioinfo.wisc.edu/biolinux/docs/ialign.html</a>&lt;sup&gt;125&lt;/sup&gt;</td>
</tr>
<tr>
<td>Agadir</td>
<td>Calculates the amount of helical structure for shorter segments.</td>
<td><a href="http://agadir.crg.es">http://agadir.crg.es</a>&lt;sup&gt;126&lt;/sup&gt;</td>
</tr>
<tr>
<td>HeliQuest</td>
<td>Helical wheel generation, calculation of hydrophobic moments, prediction of interactions and function of peptides.</td>
<td><a href="http://heliquest.ipmc.cnrs.fr">http://heliquest.ipmc.cnrs.fr</a>&lt;sup&gt;127&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Full sequence analysis</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mpex</td>
<td>Calculates hydropathy via a sliding window function.</td>
<td><a href="http://blanco.biomol.uci.edu/mpex">http://blanco.biomol.uci.edu/mpex</a>&lt;sup&gt;129,130&lt;/sup&gt;</td>
</tr>
<tr>
<td>AmphipaSeek</td>
<td>Predicts presence of amphipathic helices (IPM anchors).</td>
<td><a href="https://npsa-prabi.ibcp.fr">https://npsa-prabi.ibcp.fr</a>&lt;sup&gt;131&lt;/sup&gt;</td>
</tr>
</tbody>
</table>
7.2. Biomembrane mimetics

To study biomembrane interactions of membrane-associated GT-B enzymes without having to deal with redundant complexity we have utilized different membrane mimicking systems. This approach essentially starts from the simplest system applicable for the chosen instrumental setup required for the specific molecular event one wishes to investigate. Under some circumstances it is possible to use a crude “hydrophobic matrix”, i.e. micelles. The zwitterionic lipid DPC (dodecylphosphocholine) has been used in paper I and III. At the so-called Critical Micelle Concentration (CMC), roughly 1 mM for DPC at room temperature, micelles are formed. Higher concentrations are however generally used for protein studies, > 50 mM, to achieve a detectable effect in spectra. In both paper I and paper III we noticed (by use of CD spectroscopy) that a very similar secondary structure for the S227-245 peptide was obtained with simple micelles and with more complicated lipid systems. The small size of micelles is a great advantage for NMR spectroscopy studies, where larger lipid assemblies tend to broaden the signals beyond detection, and they do normally not interfere with signals in fluorescence or CD measurements of the type presented in this work. However, in studies where it is crucial to also maintain some physical characters of the biomembrane, either related to

![Micelle](image1)

![Large Unilamellar Vesicle](image2)

**Figure 17**: Micelles are small spherical aggregates without any planar surface. Bicelles are composed of long-chained lipids flanked by detergent molecules (shown in cross-section). LUVs are slightly more complex systems that can be produced in very many ways for tailored applications.
the head-group (see Chapter 3), acyl-chain properties or membrane curvature, they are not as useful as some other systems, for instance in studies of peripheral membrane interactions. In such cases we have instead used vesicles (liposomes) or bicelles (Figure 17) that better mimic bilayer-type lipid assemblies. To form stable bilayers we have used a lipid component that is bilayer-forming (i.e. PC species), and then we have altered the lipid composition to better mimic biomembranes that are relevant for our systems by introduction of either PG lipids (anionic), the glycolipid MGDG or PE/PG/CL in E. coli native-like membranes. The type of vesicles used in papers I-III are so-called large unilamellar vesicles (LUVs, see Figure 17), prepared by the extruder technique. They where found suitable for our studies since the size may be controlled (in contrast to some other vesicle types) and the lipid composition is quite easily altered without disruption of the assemblies. In NMR applications where interaction with LUVs will broaden the protein signals severely, we have found it useful to instead use isotropic bicelles, i.e. small fast-tumbling lipid assemblies composed of both long-chained and short-chained components. These types of bicelles will acquire a slightly disc-shaped structure with a lipid bilayer patch solvated by detergents. The morphology, physical characteristics and production of bicelles is widely described in the literature. Several novel glycolipid bicelle systems (MGDG/DGDG), suitable for solution NMR applications and plant GT studies, have been developed and characterized recently, as well as E. coli membrane mimicking bicelles made from native E. coli lipids. In paper II we investigated novel MGDG-containing bicelles and LUVs were anionic charge is additionally introduced, in order to evaluate the effect on atDGD2 peptides and bicelle morphology. In paper III E. coli native-like bicelles are presented with a lipid composition mimicking the IM where WaaG is located and active.

7.3. CD spectroscopy in studies of GT-derived segments structure and stability

In all papers Circular Dichroism (CD) spectroscopy have been used to study structural conversions and thermostability of the polypeptides. In the far UV region (~ 180-260 nm) one utilizes inherent chiral properties of the peptide bonds to gain information about conformation and structure. The CD spectrum will report on the type and the extent of circularly polarized light that dominates the absorption at a particular wavelength.
Differences in the CD signal between two samples may occur because of differences in the electronic transitions that take place between certain groups (mainly involving carbonyls and/or the nitrogen electrons) upon interaction or conformational changes. Spectra of model polypeptides or databases of experimentally solved structures can be used to estimate the amount of each secondary structure element. Far UV CD spectra may also indicate transitions involving side chains. They are not contributing as much as the peptide bond chromophores, but for random coil polypeptides the signal may be visible and confuse in secondary structure quantification attempts. In paper IV CD spectroscopy was used to quickly gain information about the thermal stability of the two MIT domains by investigating the temperature dependence of the signal intensity at 222 nm (Figure 18). Also, since an isosbestic point (a wavelength where all signals cross in experiments with one alternating variable, as seen for the MIT2 domain in paper IV) is expected for systems where only two main species are present, CD spectra can also be useful for information about protein folding and sample homogeneity.

![Figure 18: The stability ($T_M$) of proteins may be determined quickly from CD experiments by plotting the signal intensity at a specific wavelength against temperature.](image)

7.4. NMR spectroscopy to study glycosyltransferases and lipids

Nuclear magnetic resonance (NMR) is extremely useful to study structure and dynamics of molecules at an atomic level. NMR is based on several quantum mechanical properties of nuclei (i.e. the nuclear spin quantum number and the gyromagnetic ratio), such that some isotopes will, when subjected to an external magnetic field, generate signals whereas others will not. Due to differences in the proximate electronic environment, each type of NMR-active nucleus will also resonate with a slightly different frequency, from which so-called chemical shifts ($\delta$) are derived. Chemical shift
analysis makes it possible to obtain important information about the physicochemical environment of all individual NMR active nuclei in a sample, hence a very powerful tool with many applications in protein and lipid research. For instance, nuclei located in sequences of different secondary structure will exhibit characteristic deviations from the typical random coil chemical shifts, so-called secondary chemical shifts, \( \Delta \delta \). Especially \( H\alpha \) (paper I) or \( C\alpha \) chemical shifts (paper III, supp. data) may quickly inform on secondary structure, but a combination of many types of CS may be used, as in paper IV where SSP scores are displayed. In paper I we also analyzed the S227-245 segment by \( \Delta \delta \) patterns to detect an amphipathic character of the peptide in DPC micelles. Chemical shift-derived information (i.e. torsion angle constraints) is used in papers I, III and IV in calculation of the NMR structures, together with other structural constraints (explained below).

7.4.1. Sequence assignment and 3D structure determination

To generate information from most types of NMR experiments presented herein, each signal visible in NMR spectra needs to be assigned to a particular nucleus and amino-acid residue in the sequence. For this a sequential assignment routine is commonly used,\(^{162}\) based on information from a series of NMR experiments that generate slightly different information. The resulting data resembles a puzzle that may be very easy and quick to resolve for small molecules but could also be extremely difficult, depending on the amino-acid sequence properties, size of the molecule (large molecules cause line-broadening and spectral overlap) and sample conditions. For small peptides such as S227-245 and MIR-WaaG simple two-dimensional experiments were sufficient, and there was no need for isotope enrichment. MIT domains are, however, larger and for the NMR structure solved in paper IV isotope labeling and three-dimensional experiments were required in order to resolve the signals.\(^{163,164}\) Figure 19 demonstrates the different strategies used for each paper. The experiments used for structure determination are based either on retrieval of through-bond information via J-couplings (via TOCSY experiments), or on cross-relaxation rates arising from dipolar couplings, i.e. NOEs (via NOESY experiments). NOEs can be used to obtain through-space distance information\(^ {165}\) between atoms providing crucial information about distances between protons that are not directly connected through bonds, but still within proximity. The magnitude of the J-couplings gives information about
dihedral angles between the connected nuclei, and patterns in NOE intensities may be used to obtain secondary structure information since certain types of atoms will end up close to each other in some secondary structures, but not in others. All NMR-derived structural information is combined via a **structure calculation** routine, which generates the final ensemble of structures. The quality of the structure is controlled by data validation packages and manual inspection. Sometimes other NMR-derived constraints are used for structure refinement. Certain regions in ensemble structures may appear dynamic and/or unstructured, hence, it is valuable to also perform **NMR relaxation experiments** in order to validate the data quality and verify the structure, as in paper IV. Relaxation parameters from amide $^{15}$N-$^1$H dipolar vectors may inform on chemical exchange events due to conformational changes ($R_2$ rates, domain motions) or on internal motions ($R_1$ rates).

Figure 19: Structural characterization of GTs. The green route was followed in paper IV for investigations of the MIT1 domain. The red route schematically illustrates how atDGD2 (paper I) and WaaG (paper III) were investigated. Relaxation experiments were used to add information about rigidity and conformational changes in MIT1.

7.4.2. **NMR Diffusion experiments may inform on peptide-bicelle interactions**

Translational diffusion experiments have been used in paper I-III to investigate peptide-lipid interactions. The method is based on the slightly different position a molecule will have after a defined period of time due to Brownian motions (i.e. dependent on size, shape and viscosity of the solvent) and the NMR signal intensity can report on this via a series of experiments with slightly different pulse gradient strengths applied to the
Differences in the diffusion rates for peptides in solution and in the presence of bicelles were valuable in paper I-III in order to detect association to larger assemblies, i.e. interaction. The diffusion rates for lipids and peptides can also be compared to gain information about interaction strength, and to determine specificity regarding lipid types involved in the binding. Altered lipid diffusion upon peptide addition can indicate an altered morphology and/or size of the lipid assembly upon interaction, perhaps due to phase separations, withdrawal of lipids from the bilayer or altered protonation states.

7.5. Fluorescence spectroscopy for quantification of protein-lipid interactions

In my studies of GTs I have exploited intrinsic fluorophores of polypeptides, i.e. Tryptophan (Trp, paper I-II) and Tyrosine (Tyr, paper III), which is useful since these residues are often found in the membrane interface of both integral and peripheral proteins (as mentioned in Chapter 2). In papers I-II we have used two physical phenomena coupled to fluorescence to investigate if segments derived from atDGD2 interact with lipids, i.e. blue-

Figure 20: Trp fluorescence wavelength and emission intensity will depend on the solvent exposure (i.e. the dielectric environment). Trps in a hydrophobic milieu (acyl chains in lipids or the interior of a folded protein) will display a strong and blue-shifted signal, compared to solvent-exposed fluorophores (schematically illustrated as Trps in membrane proteins).
shifts and fluorescence quenching. In paper III we also utilized fluorescence anisotropy in studies of WaaG and attachment to many types of LUVs. The lifetime of a fluorophore is the time the electrons remains in the excited state, and will vary quite much between fluorophores with different surroundings. For instance, if a Trp is moved from a hydrophobic milieu to a polar, the electronic environment will alter the signal intensity since the quantum yield will be changed. Figure 20 illustrates how the emission of a distinct fluorophore may depend on the polarity of the solvent, referred to as **solvatochromism**. This effect is denoted either as a **blue shift** or **red shift** depending on how the frequency of the emitted light from the fluorophore is altered. In studies of WaaG and atDGD2 we observed different effects on the spectral wavelength maxima depending on which LUVs that were introduced and from this conclusions about membrane interactions could be drawn. The quantum yield is the fluorophore’s efficiency to convert absorbed light into fluorescence (i.e. the number of photons emitted compared to the number of photons absorbed). Hence, an increase in the emission signal intensity may be seen for peptides that interact with lipid bilayers, This effect was used for calculation of $K_D$ constants in paper III.

In addition, we have used fluorescence **quenching experiments** to extract details regarding the degree of solvent exposure of fluorophores in both WaaG and atDGD2. Quenchers are molecules that have a matching spectral overlap with the detected fluorophore, and may “steal” energy from it, thereby altering the efficient quantum yield and signal intensity. Energy transfer occurs between the molecules during the lifetime of the fluorescence, i.e. excitement of the quencher electrons on behalf of the fluorophores, yielding lower emission intensity since fewer fluorophores will contribute to the emitted light. Quenchers are handpicked for the specific system and can report either on solvent exposure or interaction depth into lipid bilayers. In papers I-III the common quencher **acrylamide** was used, a water-soluble small molecule with very limited capacity to permeate lipid bilayers. The quenching mechanism for acrylamide is referred to as collisional quenching where the energy transfer is dependent on close proximity. In paper III we also utilize a lipophilic quencher in order to gain dual information on the membrane interaction mechanism.
8. Socker och fett - livsviktigt för alla!


1. De två olika molekylerna (blå lipid och lila/grönt nukleotidsocker) som sockret ska flytta mellan binds in på ett plats i enzymet (beige).

2. GT-enzymet katalyserar reaktionen i det aktiva sätet (gulmarkerat) där den gamla kemiska bindningen (grön) bryts och den nya (röd) skapas.

3. Två nya molekyler har bildats. Lipiden har dragit vare reaktionen fått helt nya egenskaper som kan vara viktiga för funktioner i cellen.
Då det sällan är möjligt att studera den uppsjö av händelser och komponenter som förekommer samtidigt i ett cellmembran har jag försökt fokusa på enskilda delar åt gången genom att ”klippa ut” sådant jag tror är intressant, eller ta bort det som jag inte tror är lika relevant. På så vis har jag skapat förutsättningar för studier av individuella detaljer som annars skulle drunknat i bruset från allt annat som pågår i en cell. I mina artiklar har jag studerat förenklade system bestående av delar av proteiner (kallade peptider eller proteindomäner) och förenklade modeller av cellmembran (oftast i form av så kallade biceller eller vesiklar). Jag har använt mig av måtmetoder som gett möjlighet att se mycket små detaljer, ofta ända ned på atomnivå. Dessa metoder kompletterar varandra med information och är generellt baserade på spektroskopi, vilket innebär att man studerar effekten av att man bestrålat molekyler med ljus (elektromagnetisk strålning). En massa små processer sker med och mellan atomer i molekyler efter en sådan behandling, och eftersom reaktionerna (som synliggörs i ett spektrum) ser lite olika ut beroende på vilka egenskaper de molekyler man bestrålat har kan man dra slutsatser om både tredimensionell struktur och hur molekyler interagerar med varandra.

En stor del av mitt arbete har fokuserat på ett GT-enzym som förekommer i den lilla växten backtrav, som även pryder omslaget på denna avhandling (eller framförallt dess rötter). Backtrav heter vetenskapligt Arabidopsis thaliana, och anses av vissa vara ett litet skruttigt ”ogräs”. Men tvärtom så är det en mycket viktig och välstudierad växt inom forskningsvärlden! Den valdes nämligen ut som den första växt vars arvsmassa bestämdes i helhet och är modellväxt för många studier av växters gener och biokemiska processer. Det specifika enzym jag studerat heter atDGD2 och har visat sig ha en extra viktig roll om backtrav får brist på näringsämnen. Man tror att växter kan bryta ned mindre nödvändiga lipider i cellmembranen för att återvinna viktiga delar till mer kritiska processer. Det har tidigare inte funnits så mycket information om hur atDGD2 interagerar med lipider och cellmembran, och därför har resultaten som presenteras i mina två första artiklar varit viktiga. Jag har bland annat kommit fram till att atDGD2 har en strukturdel (proteindomän) med en ”plus-laddad” yta, som kan interagera med ”minus-laddade” lipider i cellmembranen samtidigt som den sticker in en del en bit in i cellmembranet, och på så vis håller sig kvar. Det är viktigt för enzymet att kunna göra det eftersom det är där det ska binda in sitt ena substrat som är en så kallad glykolipid. En annan proteindomän binder också till lipider genom laddningar, vilket kan vara del i en mekanism som
finreglerar enzymets inbindning till antingen membran eller de molekyler den ska interagera med, eller som medför viktiga strukturförändringar. Resultaten från mina studier är betydelsefulla för den grundläggande förståelsen av hur såväl GT-enzymer som andra membran-interagerande proteiner fungerar. Den tredje artikeln i min avhandling handlar om ett annat GT-enzym som strukturellt sett är mycket likt atDGD2, men istället förekommer i cellväggen hos den välkända mag-tarmbakterien *Escherichia coli*. *E. coli* och dess släktingar ligger bakom många problematiska sjukdomar, och den snabbväxande resistensbildningen mot antibiotika som uppenbarat sig på senare tid är mycket problematisk. Att hitta nya och mer specifika sätt att förhindra tillväxt av bakterier blir allt viktigare, och ett sätt skulle kunna vara att rikta behandlingen mot det enzym jag studerat, WaaG. Jag har studerat hur en del av WaaG-strukturen stickar in cellmembran på nästan samma sätt som atDGD2 gör, vilket understryker att det beteendet är generellt för många GT-enzymer. Att de betar sig på samma vis är intressant eftersom man nu kan anta att detta är vanligt även bland andra GT-enzymer, speciellt de som är lika varandra i strukturen. Att hitta sådana gemensamma nämnare är mycket viktigt inom forskning eftersom det tillåter genvägar i studier av nya och helt okända proteiner. Vissa resultat från artikeln bidrar dock också till en mycket detaljerad förståelse kring hur enskilda del-interaktioner är inblandade i enzymets lipid-inbindning, och det är sådan information som i slutändan potentiellt skulle kunna användas för framtagande av mer effektiva läkemedel.

Utöver WaaG och atDGD2 har jag studerat ytterligare ett glykosyltransferas i den fjärde artikeln. Strukturellt är detta GT-enzym lite annorlunda från de andra då det har en proteindomän som är helt inneslutet i membranet och förankrar det där, och en annan del som är mer skild från membranet och sköter katalys av sammanlänkning av sockermolekyler till *kitin*. *Kitin* är nog mest känt för att bygga upp insekternas exoskelett och svamparnas cellväggar, men finns också i mycket små mängder i cellmembran hos märkliga organismer som heter oomyceter. Oomyceter, eller algsvampar, har oftast en nedbrytande roll i naturen men är trots namnet inte alls svampar utan snarare släkt med brunalger. De förekommer lite varstans och hjälper till i kretsloppet av biologiskt material i såväl vatten som på land. Ibland ställer de till besvär för människor då de i sin städande roll även angriper försvarade fiskar och kräftdjur, en egenskap som inte är önskvärd i produktion av livsmedel. Enzymet jag studerat kommer från just en sådan oomycet, ur släktet *Saprolegnia*, och kallas ibland för "vattenmögel". Just
Saprolegnia-oomyceter verkar behöva kitin för att kunna tillväxa och det gör kitin-producerande enzymer, **kitinsyntaser**, till mål molekyler för utveckling av nya bekämpningsmedel som har bättre egenskaper än de som finns att tillgå idag. Jag har i två närbesläktade kitinsyntaser kunnat påvisa en för GT-enzyme helt ny slags proteindomän. Denna domän, **MIT**, har man tidigare funnit i sådana proteiner som är aktiva i ett stort nätverk av biomolekylära processer där proteiner och lipider flyttas inuti cellerna och mellan olika cellmembran, antingen för att de ska hamna på rätt plats eller för återvinning och nedbrytning. Förekomsten av MIT-domäner i kitinsyntaser var intressant på många vis, speciellt eftersom vi också kunde visa att de interagerar med lipider (som i föregående projekt). Detta är helt nya resultat som kan indikera en helt ny funktion för den sortens domän samt agera en bit i pusslet kring hur kitinsyntaser, och därigenom kitin, ansamlas där oomyceten tillväxer.
10. Acknowledgements

Dear Lena. Your perspectives about why we do research and what the PhD program is (or should be) about have always been very close to my own opinions, which is something I have really appreciated and that has made me feel both respected and qualified. I will always be thankful for your support and am very happy that I had the opportunity to work with you. The supervision I have gotten from you has been based on trust and I appreciate how you have managed to at the same time be very clear about your opinions but still keep a flexible attitude. This turned out to be exactly what I needed also during times when some things did not really go my way on the personal level.

I am of course thankful to all present and former group members for a very nice work environment and for many dear memories. Thank you Elsa Reimann, Henrik Biverståhl, Sofia Unnerståle, Weihua Ye, Jesper Lind, Andrea Bodor, Johannes Björnerås, Jobst Liebau, Pontus Pettersson, Biao Fu and Christian Brown. I am thankful for nice collaborations in the projects presented in this thesis; Weihua, Changrong Ge, Christian, Vincent Bulone, Jobst, Pontus, Irina Iakovleva, Hiroshi Matsutaka and Fang Fang Yin. I also value the teaching experiences I have had with Fatemeh Madani, Jesper, Ellinor Haglund and Peter Lachmann. Administrative tasks are perhaps not one of my main strengths, so I want to highlight the appreciation I feel toward all the administrative and technical staff. Some of you not only made things so much easier on a practical level, but also demonstrated a wonderful warmth and true understanding throughout these years. I am especially thankful towards Torbjörn and Haidi Astlind, Ann Nielsen, Britt-Marie Ohlsson, Anita Hjelm and Peter Nyberg.

The passing of Åke Wieslander was a great loss for me as he inspired me a lot in my early work on glycosyltransferases. He was not only a collaborator in the DGD2 project but was always putting the scientific hypotheses and experimental results in a wider perspective, strongly influencing my present
beliefs about glycosyltransferases and membrane interactions. I will always remember him with warmth in my heart.

Many other people have circulated through the biophysics corridor (and the rest of DBB) during all these years and I have also had valuable help by many of you. There are some people I have gotten to know at DBB that I want to thank specifically for great fun and many hours of very nice (and important!) conversations. Thank you Luminita Färnström, Maria Öhrström, Anna Broström, David Gotthold and Ana Popović-Bijelić; I’m truly delighted to have met you all. I’m also privileged to have gotten to know Astrid Gräslund, and have really valued the talks we have had about the academia in general and things not always connected to science. Stefan Nordlund’s enthusiasm and interest in all the PhD projects at DBB is impressive, as so many has already experienced. He has helped me with this thesis and my research plan but he also guided me during some difficult periods, for which I want to express appreciation. There are many others that have had impact on this work (and me): Axel Abelein, Ann-Sofi Johansson, Sebastian Wärmländer, Jüri Jarvet, Lina Leinartaite, Lisa Lang, Martin Kurnik, Jens Danielsson, Susanna Seppälä, Sylvia Eriksson, Kamila Oglecka, Nina Voevodskaya, Göran Eriksson, Christoffer Lendel, Leila Ghalebani and Ida Lindström. I must honestly say that it is very hard at this moment to remember all the people that have inspired and supported me, both scientifically and on a personal level, and I am very sorry there are people that for sure should be listed here, but are unintentionally left out.

I am amazed by the kindness and acceptance I have experienced by the people working at the Swedish Museum of Natural History, where the writing of most of the introduction chapters has taken place. In particular I wish to thank Julia Stigenberg and Mattias Forshage for comments on the text and for fun discussions about both science and bugs. I am thankful to Johannes Lundberg for helping me find interesting specimens to photograph for the front cover and oomycete picture.

In addition, one important reason that I chose to continue with PhD studies was the experiences I had and the people I met during work with my Master thesis at Linköping University. Many thanks to Per Hammarström, Ina Caesar, Karin Sörgjerd, Daniel Sjölander, Karin Almstedt, Cecilia Andrésen, Sofie Nyström and Satish Babu Moparthi, amongst others.
To my family, both old, new and extended: Jag älskar er alla, tack för allt er stöd. Babciu i Dziadku! Kocham was bardzo i brak mi was kazdego dnia! Emilie och Björn, jag är stolt och glad över att ha lärt känna er och tacksam att Tesla har så fina syskon. Tack till mina fina vänner, i synnerhet Maria-Therese Gustafsson med familj, Mikael Kinn, Hjördis Lenz, Kristin Ekström med familj samt Karin Wästlund för värdefull hjälp med Tesla och annat direkt kopplat till “livet som sådant” under senaste tiden.

Rasmus, my love, it is just amazing how patient and understanding you have been while I have been working on this thesis. The sincere love I have felt and the loyalty that I have experienced from you during this time has strengthened me tremendously.

11. References

(1) Luckey M. Membrane Structural Biology: With biochemical and biophysical foundations (2008), Cambridge University Press, United States.


(34) Shimojima M. (2011) Biosynthesis and functions of the plant sulfolipid, Prog.Lipid Res. 50, 234-239.


(51) Ruiz-Herrera J. Fungal cell wall: structure, synthesis, and assembly (2012), CRC Press, Boca Raton, USA.


(73) Brown C. Characterization of specific domains of the cellulose and chitin synthases from pathogenic oomycetes (2015), KTH Royal Institute of Technology, Stockholm.


synthase MIT domain on POPA and POPC membranes by molecular
structure of proteins from their amino acid sequence,
(108) Levitt M. (1978) Conformational preferences of amino acids in
globular proteins, Biochemistry 17, 4277-4285.
(109) von Heijne G. Sequence analysis in molecular biology : treasure trove
or trivial pursuit (1987), Academic, San Diego, Calif.
(110) Botte C., Jeanneau C., Snajdrova L., Bastien O., Imberty A., Breton C.
and Marechal E. (2005) Molecular modeling and site-directed
mutagenesis of plant chloroplast monogalactosyldiacylglycerol
synthase reveal critical residues for activity, J.Biol.Chem. 280, 34691-
34701.
modeling of glycosyltransferases, Methods Mol.Biol. 347, 145-156.
(113) Romero P., Obradovic Z., Li X., Garner E.C., Brown C.J. and Dunker
A.K. (2001) Sequence complexity of disordered protein, Proteins 42,
38-48.
amino acid sequences and sequence databases, Comput.Chem. 17, 149-
163.
biased regions in sequence databases, Meth.Enzymol. 266,
hydrophobicity scale for proteins at membrane interfaces, Nat. Struct.
Biol. 3, 842-848.
hydrophobic moment: a measure of the amphiphilicity of a helix,
a database for plant carbohydrate-active enzymes, Database (Oxford)
(119) Lombard V., Golaconda Ramulu H., Drula E., Coutinho P.M. and
Henrissat B. (2014) The carbohydrate-active enzymes database (CAZy)
in 2013, Nucleic Acids Res. 42, D490-D495.
(120) Cantarel B. L., Coutinho P.M., Rancurel C., Bernard T., Lombard V.
(CAZy): an expert resource for glycoinformatics, Nucleic Acids Res. 37,
D233-238.
Glyco3D: a portal for structural glycosciences, Methods Mol.Biol.
1273, 241-258.
77


(147) Fasman GD editor. Circular dichroism and the conformational analysis of biomolecules. ; 1996: Springer New York, USA.


