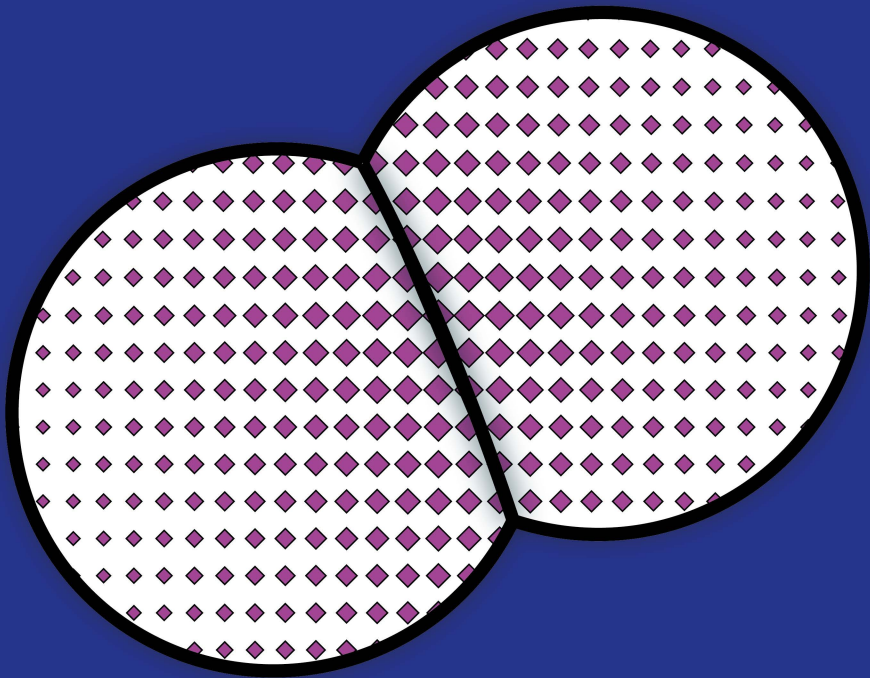# To sugarcoat bacteria

Glycoengineering of
lipooligosaccharides of bacteria
with sialic acid analogues

Hanna de Jong

# To sugarcoat bacteria

## Glycoengineering of lipooligosaccharides of bacteria with sialic acid derivatives

Hanna de Jong

# To sugarcoat bacteria

# Glycoengineering of lipooligosaccharides of bacteria with sialic acid derivatives

**Glyco-modificaties van bacteriële lipo-oligosacchariden met siaalzuuranalogen**

(met een samenvatting in het Nederlands)

**Proefschrift**

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof. dr. H.R.B.M. Kummeling,
ingevolge het besluit door college voor promoties
in het openbaar te verdedigen op

donderdag 23 juni 2022 des avonds om 6.15

door

**Hanna de Jong**

geboren op 21 maart 1993
te Utrecht

**Promotor**

Prof. dr. G.J.P.H. Boons

**Co-promotoren**

Dr. T. Wennekes
Dr. M.M.S.M. Wösten

# Table of Contents

# Chapter 1

## Introduction

## Chapter 1

Chemical biology spans the interdisciplinary field between chemistry and biology, and aims to study and manipulate biological processes at the molecular level through the application of chemistry, often with the use of small or tailor-made molecules. Two leading edge areas of research within chemical biology are represented in this thesis: the elucidation of glycobiology at the molecular-level and the use of small molecules to probe receptor ligand interactions.

## Glycans as biomolecules

Glycans are a class of biomolecules that are important for many different fields, such as cell biology, medicine and biotechnology. The significance of glycans in nature is illustrated by their great abundance on the one hand, for example the glycan cellulose being the most abundant organic biomolecule on earth, and on the other hand by its ability to alter processes through small changes in its structure, for instance only the presence or absence of a single monosaccharide determines the difference between blood group A, B and O[1–3].

Glycans are also present on the cell surface, often as conjugates with other biomolecules, like glycoproteins and glycolipids. The layer of glycans on the cell surface is referred to as the glycocalyx. The size of the glycocalyx varies per cell type, but can be up to several micrometers thick for a human endothelial glycocalyx layer[4], for instance. Given that cells are covered by glycans, it is easy to imagine that these biomolecules are key players in biology[5]. Indeed, glycans are important for interactions between cells and the immune system[6,7]. Recent examples that demonstrate the importance of glycans are the glycosylated spike protein on the coronavirus SARS-CoV-2 that shields it from antibody binding and the discovery of glycosylated RNA[8,9].

The architecture of glycans spans tremendous possibilities. Since glycans can be composed of a multitude of monosaccharides, which can be coupled to each other through multiple linkage types, and can be part of a larger conjugate, like glycoproteins or glycolipids, there is a vast number of different glycan structures (Figure 1). To make matters even more complex, the biochemical synthesis of these glycans is not templated, unlike the central dogma from DNA to proteins, and this makes it hard to predict their structure or make modifications to them, for instance[10]. In the past decades, several 'tools' from various fields have been developed to enable the study of glycans and in this thesis we will focus on tools from chemical biology.
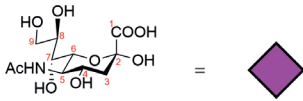
Figure 1. Structure and glycan symbol of the monosaccharide N-acetylneuraminic acid (Neu5Ac) or sialic acid. This glycan is the focus of several research chapters in this thesis. The numbers indicate the different carbon atoms and the corresponding hydroxyl groups can form different glycosidic linkages, which illustrates the diverse possibilities of glycan structures.

## Bacterial glycans

Not only eukaryotic cells are covered by glycans, but also prokaryotic cells. In case of bacteria, the cell surface is highly variant and several glycans are attached to the outer membrane as part of glycoconjugates, for instance: glycoproteins, lipopolysaccharides (LPS), lipooligosaccharides (LOS), and capsular polysaccharides (CPS) (Figure 2). LPS can be further subdivided in lipid A, a core region and an O-antigen region[11–13]. These glycoconjugates improve bacterial survival through several ways, for instance via avoiding recognition by the immune system, providing motility or cell wall integrity to the bacteria[13,14].

Also for bacterial glycosylation there are many unanswered questions. This field deals with unique challenges like the large monosaccharide diversity compared to eukaryotes[15]. Another challenge is that it was thought that protein glycosylation was restricted to eukaryotes[16], but since the discovery of bacterial protein glycosylation, the interactions of these glycoproteins and their biochemical assembly has become an quickly expanding field of study[17–20]. Another issue that has constrained research is that there is a strong focus on pathogenic bacteria because of their clinical relevance and by focusing on the glycosylation of these pathogens, the commensal bacteria are often not the first candidates to study[17,21]. Taken together the inherent challenges of glycoscience and the challenges unique to bacterial glycosylation, there is a clear need for custom made tools to meet these challenges.
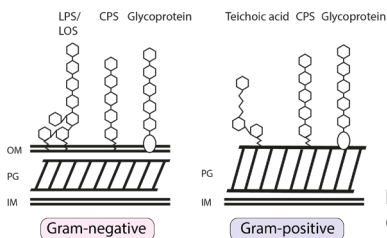


Figure 2. Schematic glycoconjugates attached to the cell surface of Gram-negative and Gram-positive bacteria. Abbreviations not in text; IM = inner membrane; PG = peptidoglycan; OM = outer membrane.

# Chapter 1

## Glycoengineering

One possible approach in glycoscience is to make use of glycoengineering (Figure 3). This chemical biology method can be applied to all glycans, both eukaryotic and prokaryotic. Glycoengineering is a term that describes the modification of glycans that occurs due to introduced external factors, for example the introduction of a bio-orthogonal group. In general, glycoengineering is the controlled editing of cell surface glycans and this method can erase or elongate the glycans, introduce unnatural groups or monosaccharides, inhibit or supplement the biosynthesis (Figure 3)[22–24]. More specifically, a set of tools was developed for glycoengineering and a selection that relates to bacterial glycans was surveyed as part of a literature review about bacterial glycan mimicry in **Chapter 2**. One such tool is selective exoenzymatic labeling (SEEL) which can modify cell surface glycans trough the introduction of a reporter group, such as a biotin or fluorophore. We have applied the SEEL technique to a bacterium, to our knowledge, for the first time, and report on this in **Chapter 3**. As an additional approach to glycoengineer bacterial glycans, we explored the power of bacterial sialyltransferases to modify their own glycans with unnatural sugars containing a reporter group in **Chapter 4**.
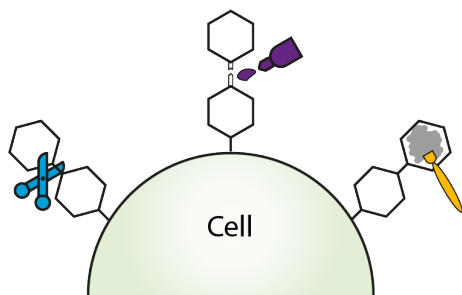


Figure 3. Schematic picture of glycoengineering tools to remove, introduce or alter glycans on the cell surface.

## A completely different world: plants and hormones

From the world of bacterial glycans we are now jumping into an altogether different world: plant hormones. Probably the most well-known hormone is auxin, or indole-3-acetic acid (IAA)(Figure 4), and this molecule acts in different biological processes such as plant growth, and plant's sensation to light[25]. Another auxin is 2,4-D (2,4-dichlorophenoxyacetic acid)(Figure 4), which is referred to as a 'synthetic' auxin because it can act in different ways than the 'natural' auxin IAA[26–28]. 2,4-D is very efficient at inducing a special kind of plant growth, namely somatic embryogenesis (SE). SE is a unique developmental process in which

differentiated somatic cells can acquire totipotency and are 'reprogrammed' to form new 'somatic' embryos[28]. This process can take place *in vitro* and is also applied in plant biotechnology, but the process occurs in nature as well for the plant named 'Mother of thousands' (*Kalanchoe daigremontiana*)[29]. 2,4-D is a small molecule, but even minor changes to the molecule can have a big impact on its biological function, as we will see for the set of 2,4-D analogues tested in a chemical biology approach to study somatic embryogenesis and microspore embryogenesis in **Chapter 5** and in the general discussion of **Chapter 6**.
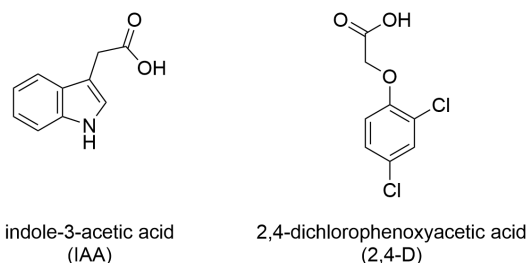


indole-3-acetic acid
(IAA)

2,4-dichlorophenoxyacetic acid
(2,4-D)

Figure 4. Structure of the auxins IAA and 2,4-D

## Aims and outline of this thesis

This thesis had two main aims: glycoengineering the sialic acid on the bacterial cell surface and establishing a structure-activity relationship (SAR) for the molecule 2,4-D in somatic embryogenesis of *Arabidopsis thaliana*. Although seemingly a world apart, as mentioned, both aims are within the field of chemical biology. The glycoengineering approach in this thesis aims to expand the current chemical biology techniques available to study bacterial glycans and the SAR approach aims to advance our understanding of the biological process somatic embryogenesis with the use of synthetic small molecules.

In **Chapter 2** we describe a selection of bacteria that are known to display glycan mimicry. These bacteria contain glycans in their glycocalyx that resemble human glycans. This form of mimicry can have many immunological effects and we speculate that one major outstanding question is why this phenomenon occurs. We finalize this chapter with an overview of chemical biology techniques we think will advance our understanding of glycan mimicry.

A chemical biology technique we have applied to bacteria is selective exoenzymatic labeling (SEEL). A lot remains unknown about bacterial glycans and we hypothesized that the SEEL technique, which has been previously used to engineer mammalian cells, would be an additional tool to glycoengineer

bacteria and further increase our understanding of bacterial glycans in the future. In **Chapter 3** we describe our efforts in applying this technique to *Neisseria gonorrhoeae* with modified nucleotide sugars of *N*-acetylneuraminic acid (Neu5Ac), and characterizing the different parameters, like the substrate scope and amount of labeling, and discuss the possible application of this technique.

Whilst performing the research for SEEL, we discovered that certain bacteria are capable of scavenging unnatural nucleotide sugars and transferring these onto their own glycocalyx. In **Chapter 4** we explore this further by glycoengineering the sialic acid on several bacteria including Nontypeable *Haemophilus influenzae*, *Neisseria gonorrhoeae*, and *Campylobacter jejuni*, and screen a selection of bacteria for their capacity to incorporate modified Neu5Ac.

In **Chapter 5** we switch to the world of plants and investigate the structure-activity relationship of the synthetic hormone 2,4-D in somatic embryogenesis of *Arabidopsis Thaliana*. We had designed and synthesized a small library of 2,4-D analogs with modifications on the aromatic ring and the carboxylic acid chain of the molecule. In collaboration with Leiden University the compounds from this library were tested for root growth inhibition, auxin response and induction of somatic embryogenesis.

Finally, **Chapter 6** describes briefly the chemical biology setting of the two approaches in this thesis and continues with a discussion of those approaches, including additional unpublished results, like the activity of 2,4-D in microspore embryogenesis.

# References

(1) Varki, A.; Kornfeld, S. Historical Background and Overview. In *Essentials of Glycobiology*; Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Darvill, A. G., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L., Seeberger, P. H., Eds.; Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press, 2017.

(2) Klemm, D.; Heublein, B.; Fink, H. P.; Bohn, A. Cellulose: Fascinating Biopolymer and Sustainable Raw Material. *Angew. Chem - Int. Ed.* **2005**, *44*, 3358–3393.

(3) Stanley, P.; Cummings, R. D. Structures Common to Different Glycans. In *Essentials of Glycobiology*; Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Darvill, A. G., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L., Seeberger, P. H., Eds.; Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press, 2017.

(4) Weinbaum, S.; Duan, Y.; Thi, M. M.; You, L. An Integrative Review of Mechanotransduction in Endothelial, Epithelial (Renal) and Dendritic Cells (Osteocytes). *Cell. Mol. Bioeng.* **2011**, *4*, 510–537.

(5) Varki, A. Biological Roles of Glycans. *Glycobiology* **2017**, *27*, 3–49.

(6) Crocker, P. R.; Paulson, J. C.; Varki, A. Siglecs and Their Roles in the Immune System. *Nat. Rev. Immunol.* **2007**, *7*, 255–266.

(7) Dube, D. H.; Bertozzi, C. R. Glycans in Cancer and Inflammation - Potential for Therapeutics and Diagnostics. *Nat. Rev. Drug Discov.* **2005**, *4*, 477–488.

(8) Walls, A. C.; Park, Y. J.; Tortorici, M. A.; Wall, A.; McGuire, A. T.; Veesler, D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **2020**, *181*, 281-292.e6.

(9) Flynn, R. A.; Pedram, K.; Malaker, S. A.; Batista, P. J.; Smith, B. A. H.; Johnson, A. G.; George, B. M.; Majzoub, K.; Villalta, P. W.; Carette, J. E.; Bertozzi, C. R. Small RNAs Are Modified with N-Glycans and Displayed on the Surface of Living Cells. *Cell* **2021**, *184*, 3109-3124.e22.

(10) Zol-Hanlon, M. I.; Schumann, B. Open Questions in Chemical Glycobiology. *Commun. Chem.* **2020**, *3*, 102.

(11) Whitfield, C.; Szymanski, C. M.; Aebi, M. Eubacteria. In *Essentials of Glycobiology*; Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Darvill, A. G., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L., Seeberger, P. H., Eds.; Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press, 2017.

(12) Erridge, C.; Bennett-Guerrero, E.; Poxton, I. R. Structure and Function of Lipopolysaccharides. *Microbes Infect.* **2002**, *4*, 837–851.

(13) Misra, S.; Sharma, V.; Srivastava, A. K. *Bacterial Polysaccharides: An Overview*; Ramawat, K. G., Mérillon, J.-M., Eds.; Springer International Publishing: Cham, 2015.

(14) Alexander, C.; Rietschel, E. T. Bacterial Lipopolysaccharides and Innate Immunity. *J. Endotoxin Res.* **2001**, *7*, 167–202.

(15) Imperiali, B. Bacterial Carbohydrate Diversity — a Brave New World. *Curr. Opin. Chem. Biol.* **2019**, *53*, 1–8.

(16) Nothaft, H.; Szymanski, C. M. New Discoveries in Bacterial *N*-Glycosylation to Expand the Synthetic Biology Toolbox. *Curr. Opin. Chem. Biol.* **2019**, *53*, 16–24.

(17) Eichler, J.; Koomey, M. Sweet New Roles for Protein Glycosylation in Prokaryotes. *Trends Microbiol.* **2017**, *25*, 662–672.

(18) Tytgat, H. L. P.; Lebeer, S. The Sweet Tooth of Bacteria: Common Themes in Bacterial Glycoconjugates. *Microbiol. Mol. Biol. Rev.* **2014**, *78*, 372–417.

(19) Tytgat, H. L. P.; de Vos, W. M. Sugar Coating the Envelope: Glycoconjugates for Microbe–Host Crosstalk. *Trends Microbiol.* **2016**, *24*, 853–861.

(20) Tan, F. Y. Y.; Tang, C. M.; Exley, R. M. Sugar Coating: Bacterial Protein Glycosylation and Host-Microbe Interactions. *Trends Biochem. Sci.* **2015**, *40*, 342–350.

(21) Comstock, L. E.; Kasper, D. L. Bacterial Glycans: Key Mediators of Diverse Host Immune Responses. *Cell* **2006**, *126*, 847–850.

(22) Griffin, M. E.; Hsieh-Wilson, L. C. Glycan Engineering for Cell and Developmental Biology. *Cell Chem. Biol.* **2016**, *23*, 108–121.

(23) Nischan, N.; Kohler, J. J. Advances in Cell Surface Glycoengineering Reveal Biological Function. *Glycobiology* **2016**, *26*, 789–796.

(24) Critcher, M.; O'Leary, T.; Huang, M. L. Glycoengineering: Scratching the Surface. *Biochem. J.* **2021**, *478*, 703–719.

(25) Whippo, C. W.; Hangarter, R. P. Phototropism: Bending towards Enlightenment. *Plant Cell* **2006**, *18*, 1110–1119.

(26) Peterson, M. A.; McMaster, S. A.; Riechers, D. E.; Skelton, J.; Stahlman, P. W. 2,4-D Past, Present, and Future: A Review. *Weed Technol.* **2016**, *30*, 303–345.

(27) Calderón Villalobos, L. I. A.; Lee, S.; De Oliveira, C.; Ivetac, A.; Brandt, W.; Armitage, L.; Sheard, L. B.; Tan,

# Chapter 1

X.; Parry, G.; Mao, H.; Zheng, N.; Napier, R.; Kepinski, S.; Estelle, M. A Combinatorial TIR1/AFB–Aux/IAA Co-Receptor System for Differential Sensing of Auxin. *Nat. Chem. Biol.* **2012**, *8*, 477–485.

(28)   Shimizu-Mitao, Y.; Kakimoto, T. Auxin Sensitivities of All Arabidopsis Aux/IAAs for Degradation in the Presence of Every TIR1/AFB. *Plant Cell Physiol.* **2014**, *55*, 1450–1459.

(29)   Garcês, H. M. P.; Champagne, C. E. M.; Townsley, B. T.; Park, S.; Malhó, R.; Pedroso, M. C.; Harada, J. J.; Sinha, N. R. Evolution of Asexual Reproduction in Leaves of the Genus *Kalanchoë*. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 15578–15583.

# Chapter 2

## Sweet impersonators: molecular mimicry of host glycans by bacteria

All bacteria display surface-exposed glycans that can play an important role in their interaction with the host and in select cases mimic the glycans found on host cells, an event called molecular or glycan mimicry. In this review, we highlight the key bacteria that display human glycan mimicry and provide an overview of the involved glycan structures. We also discuss the general trends and outstanding questions associated with human glycan mimicry by bacteria. Finally, we provide an overview of several techniques that have emerged from the discipline of chemical glycobiology, which can aid in the study of the composition, variability, interaction and functional role of these mimicking glycans.

*Hanna de Jong, Marc M. S. M. Wösten, Tom Wennekes*

# Chapter 2

## Introduction

Bacterial glycans play an important role in host-microbe interactions[1]. Glycans at the bacterial cell surface are often the first molecules to interact with the environment and can mimic glycan structures of the host. The existence of identical molecular structures on the host and a microbe is referred to as molecular or glycan mimicry[2]. Bacteria may exploit glycan mimics to hijack host biology to establish infection, but the glycan mimicry may also inadvertently induce abnormal immune responses causing serious auto-immune pathology. Bacterial surface glycans that have been found to display glycan mimicry with host glycan structures include capsular polysaccharides (CPS), and glycolipids, like lipopolysaccharide (LPS) and lipooligosaccharides (LOS). These bacterial surface glycans mimic certain classes of eukaryotic glycans for which an overview is presented in Figure 1. The bacteria that will be discussed in this review are grouped per class of eukaryotic glycans they are mimicking. A detailed overview of the bacterial glycan structures that display glycan mimicry are depicted in Table S1.

The concept of bacterial glycan mimicry and its effects on host immunity has been subject of a few excellent reviews[3,4]. Here we will focus on an updated discussion of research into the glycan structures and interacting proteins involved in bacterial host glycan mimicry and its impact on host biology. We have selected 11 bacterial species that display glycan mimicry. First, we will discuss the bacterial glycan structures involved (using a graphical notation according to the current symbol nomenclature[5], Table S1), their interaction with host receptors (see Figure 2 for an overview) and their role in bacterial pathogenesis. In a second part of the review, we will highlight the potential of recent advances in the field of chemical glycobiology and make a case for their use in future study of bacterial glycan mimicry. Of note, microbe-host mimicry is not limited to glycans and also not all bacterial glycans on microbes that interact with their host are involved in mimicry. However, discussion of other forms of molecular mimicry (e.g., protein mimicry)[6], microbial glycan interactions[1], and the nature of unique microbial glycans[7–9] is beyond the scope of the current review.

N-Acetyl-Neuraminic acid    Galactose    N-Acetyl-Galactosamine    Mannose
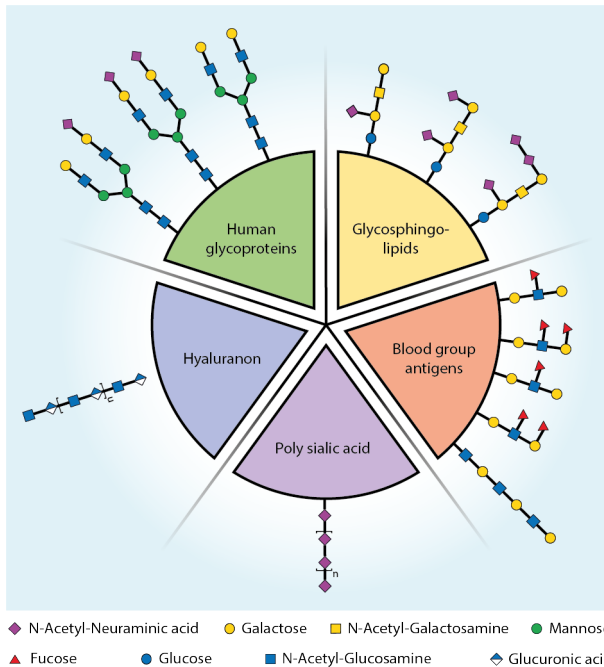Fucose            Glucose        N-Acetyl-Glucosamine        Glucuronic acid

Figure 1. Bacterial glycans can mimic the glycans of the (human) host. Here the targets of glycan mimicry are grouped for several human glycan types. Bacterial glycans that resemble these and are thus molecular mimics, can be found in SI table S1.

## Bacterial pathogens that display glycan mimicry

### Glycosphingolipids

#### *Campylobacter jejuni*

One of the most well studied forms of glycan mimicry involves the bacterial surface glycans of the food-borne pathogen *Campylobacter jejuni*. Infection by *C. jejuni* is a common cause of human gastroenteritis and in rare cases can lead to Guillain-Barré syndrome (GBS). In case of GBS, the immune system produces antibodies against the infecting microbe's glycans that cross-react with similar glycan structures on the host's nerves[10]. GBS can cause acute flaccid paralysis and is classified in different subtypes based on further clinical features[11]. Although infections from many different bacteria and viruses could potentially result in GBS[11], *C. jejuni* is the most frequent preceding infection, especially for the GBS subtypes, acute motor axonal neuropathy (AMAN) and Miller Fisher syndrome (MFS)[11,12]. *C. jejuni* infections are frequent, but only an estimated 1 per 1000 infections results in GBS, which suggest additional factors play a role in developing GBS[10,11]. For *C. jejuni* it is well documented that some forms of LOS mimic the gangliosides found on nerves[13]. Cross-reactive antibodies against different LOS

19

epitopes and gangliosides can be found in patient sera[14]. In addition, the link between the glycan mimicry of gangliosides and the development of GBS was demonstrated in rabbits[15]. The animals were administered GM1-like LOS from a clinical isolate, upon which they developed GBS[15]. *Campylobacter coli* and *Brucella melitensis* might also have GM1-like structures, because they reacted with an anti-ganglioside antibody[12], but it remains unknown whether they can cause GBS[16]. Although the evidence for a link between glycan mimicry and onset of GBS is strong, it also does not appear to be a prerequisite. For instance, Godschalk et al. used capillary electrophoresis MS to determine the LOS of *C. jejuni* strains from patients with Enteritis, GBS and MFS[17]. Although a correlation was found between *C. jejuni* strains that had mimicking LOS and the diseases GBS and MFS, there were also *C. jejuni* strains that lacked this type of LOS and did give rise to disease[17]. This suggests that other still unknown factors might play a role in developing GBS.

Variants of *C. jejuni* LOS can bind to several host receptors and although less well studied, also to host glycans (Figure 2)[18]. The Siglec 7 receptor, which occurs on monocyte-derived macrophages, Natural Killer and dendritic cells, is bound by strains associated with GBS and MFS through disialylated LOS and not monosialylated LOS[19,20]. LOS that contains a terminal GalNAc can bind to the Macrophage Galactose-type lectin (MGL, CLEC10A, CD301) on immature dendritic cells[21]. LOS structures mimicking Gd1a, GM1b and GM3 gangliosides show binding to both murine[22] and human[23] sialoadhesin (Sn, Siglec-1, CD169) that is dependent on terminal α2,3-linked *N*-acetylneuraminic acid (α2,3-Neu5Ac). The binding of LOS to Sn leads to increased uptake by the macrophages and increased IL-6 production but does not facilitate bacterial survival as lysosomal degradation still occurs. The exact role of Sn binding remains unclear, but Sn-mediated LOS ingestion by antigen presenting cells might facilitate the production of auto-antibodies[23].

The key role of the sialic acid family, mainly in the form of Neu5Ac, in glycan mimicry by *C. jejuni* LOS has also been confirmed by genetic studies. For example, mutations of either heptosyltransferase, which leads to truncated LOS[24], or Neu5Ac synthetase[25] both abolish LOS sialylation and mimicry. The sialyltransferase CstII is responsible for the sialylation of LOS and can create either α2,3- or both α2,3- and α2,8-linkages, depending on only a single point mutation[17,26]. *In vitro* studies of *C. jejuni* with sialylated LOS show it is more invasive yet less adherent[27]. Additionally, sialylated LOS facilitates translocation across Caco-2 cells, yet intracellular survival depends on the number of translocated bacteria and not on LOS structure[28]. Most studies on LOS sialylation are *in vitro*,

so the exact *in vivo* role remains to be further elucidated. However, a patient study from Finland indicated that LOS sialylation might only have a limited contribution to *in vivo* invasiveness[29]. Only 23% of the *C. jejuni* isolates had the genetic potential to sialylate the LOS[29].

Besides the genetic potential for mimicry by *C. jejuni*, the exact LOS composition in a culture is also dependent on growth conditions. For example, the model strain *C. jejuni* 11168 produces LOS with 90% of glycan structures mimicking GM1 if grown at 37 °C, which drops to 50% if grown at 42 °C[30]. Day *et al.* suggest that the heterogeneity of LOS could contribute to population fitness and hypothesized that the different LOS structures form multiple targets for auto-antibodies[30].

## *Moraxella catarrhalis*

*Moraxella catarrhalis* is a nonmotile Gram-negative pathogenic bacterium with an affinity for the human upper respiratory system. The LOS of serotype A, B and C was shown to have a common Gal-Gal-Glc motive, which can also be found on human glycolipids[31–33]. The interactions of the LOS with host biology remain to be studied. The glycan mimicry by this bacterium might aid in host colonization and invasion[33].
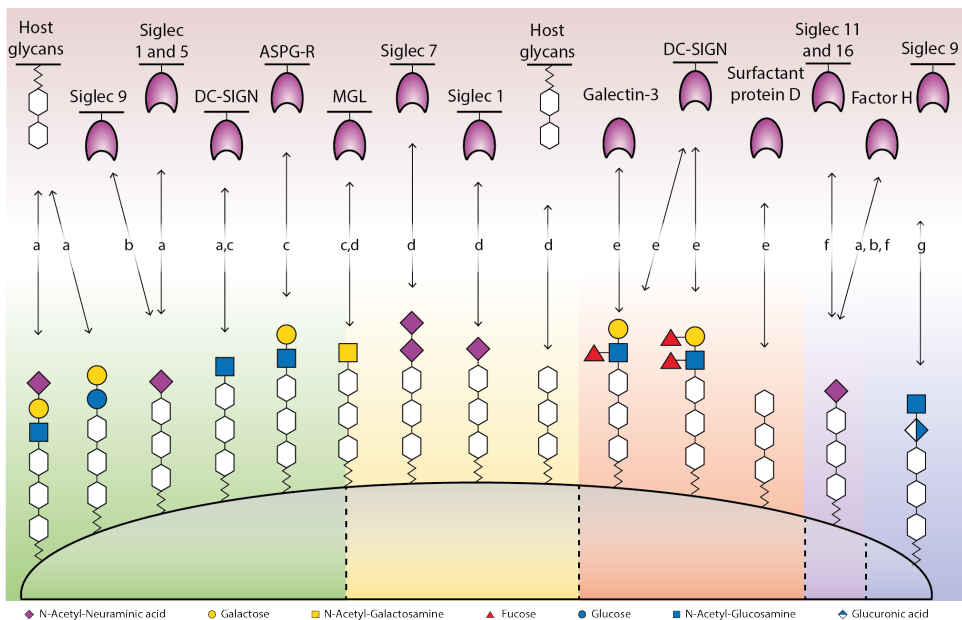


Figure 2. Interactions between the terminal surface glycan structures of mimicking bacteria and the host's glycan binding proteins or glycans. The bacteria are indicated by the following letters: (a) *N. meningitidis*, (b) Group B *streptococcus*, (c) *N. gonorrhoeae*, (d) *C. jejuni*, (e) *H. pylori*, (f) *E. coli*, (g) Group A *streptococcus*.

# Chapter 2

## Blood group antigens

### Helicobacter pylori

*Helicobacter pylori* is a Gram-negative, microaerophilic, helically shaped bacterium. It is a gastric pathogen that persists for a lifetime in most human's stomach[34] and can lead to a variety of diseases, such as gastric and duodenal ulcers, chronic gastritis and mucosa-associated lymphoid tissue[35]. This persistent colonization is associated with glycan mimicry by the *O*-antigen of *H. pylori's* LPS of Lewis blood group antigens[13,36]. A study that investigated the LPS composition of eight *Helicobacter* strains with reactivity towards antibodies directed against Lewis antigens found only *H. pylori* and *H. mustelae* displayed glycan mimicry, of which the latter is a pathogen that can be found in ferrets[37]. In an effort to understand the role of glycan mimicry in auto-immunity, Appelmelk *et al.*[38] studied seven *H. pylori* strains for mimicry with antibodies against Lewis antigens and found cross reactivity between LPS and gastric tissue.

The glycan structure of *H. pylori* LPS is important for invasiveness[39], colonization[40] and antigenicity[41], details on the biosynthesis of LPS with Lewis antigens have been previously described[42,43]. *Helicobacter pylori* with truncated LPS, which lacks the *O*-antigen, colonized less compared with wildtype[40]. Comparison of the glycan composition of two *H. pylori* strains with different virulence found they differed in the monosaccharides rhamnose and mannose, yet the significance of this for invasiveness remains unclear[39]. In relation to antigenicity, one study compared a weakly and highly antigenic form of LPS and the two LPS structures only differed in one GlcNAc residue[41], highlighting that a minor change in glycan structure can have direct impact on immunogenicity. *Helicobacter. pylori* might persist in the host by balancing the population of T-helper-1 and -2-cells through phase variation and expression of Lewis antigens[36]. This balance might be achieved via Lewis x and Lewis y that bind to DC-SIGN (Figure 2), which causes a down-regulation in T-helper-1-cells[36,42]. Besides their binding to DC-SIGN, Lewis antigens bind to several other receptors (Figure 2). Lewis x binds to Galectin-3 and is thought to contribute to adhesion[42]. *Helicobacter pylori* *O*-antigens with a lower degree of fucosylation preferentially bind to surfactant protein D and this is dependent on the phase-variable expression of the bacterial fucosyltransferase[41,42,44]. For a more detailed discussion about the function of glycan mimicry for *H. pylori*, the reader is referred to a perspective by Bergman et al.[36].

## Polysialic acid

### Escherichia coli

*Escherichia coli* is a Gram-negative, facultative anaerobic, bacterium that is commonly found in the lower intestine. Its *O*-antigen and capsule can display glycan mimicry. The *O*-antigen is composed of 10-25 repeating units of two to seven monosaccharides[45]. Serotypes O86, O90, O127 and O128, are presumed to be glycan mimicking due to their resemblance to blood group antigens[45]. For more details we refer to a comprehensive overview of *O*-antigens of *E. coli*[45]. The capsule of *E. coli* K1 and K92 consists of polysialic acid, which is either α2,8- or both α2,8- and α2,9-linked, respectively[46], and mimics the polysialic acid found on neural cell adhesion molecules[47]. Contrary to vertebrates, where polysialylation is found on the non-reducing termini of sialylated *N*- and *O*-linked glycans of certain glycoproteins, bacterial polysialic acid structures are covalently linked to an oligo-KDO that is anchored to the plasma membrane[48]. The capsule of *E. coli* K1 was shown to bind both inhibitory Siglec 11 and activating Siglec 16 (Figure 2), which give opposite signals to the immune system[49]. In addition to Siglec binding, the polysialic acid capsule possibly also recruits factor H, as described for *Neisseria meningitidis* (Figure 2)[46].

### *Neisseria meningitidis* serogroup B

*Neisseria meningitidis* is a Gram-negative bacterium that can cause meningitis and sepsis. Its serogroup B strain has a polysialic acid capsule that mimics the polysialic acid that can be found on neural cell adhesion molecules, similar to *E. coli*[47]. *N. meningitidis* serogroup C also has a polysialic acid capsule, but is α2,9-linked instead of α2,8-linked like the capsule of serogroup B. Although this serogroup C, and the serogroups W-135 and Y contain Neu5Ac, these have not been described as involved in glycan mimicry. For a more detailed discussion about the glycan structures of *Neisseria*, the reader is referred to two excellent reviews[4,50]. A possible function of the glycan mimicry by *N. meningitidis* serogroup B's polysialic acid capsule could be immune evasion through the possible recruitment of factor H (Figure 2), which suppresses the complement system, conferring serum-resistance[51]. If there is less sialylation of the capsule, there is more C3 deposition on the bacterial cell surface[51]. The binding of C3 on the cell surface is the first step in the complement pathway to form the membrane-attack complex that eventually can lead to cell death[52]. The deposited C3 can be removed by factor H, which is usually responsible for recognizing 'self', and it has been hypothesized that Neu5Ac could attract factor H[51].

# Chapter 2

# Human glycoproteins

## Neisseria meningitidis

The LOS of *N. meningitidis* contains a terminal structure, lacto-*N*-neotetraose, that is also found on human glycoproteins and glycosphingolipids[53]. Sialyltransferases from *N. meningitidis* sialylate its LOS, often with an α2,3-linkage, except for immunotype L1 in which its α2,6-linked Neu5Ac[53,54]. This sialylated LOS has been shown to bind to Siglecs Sn or 5 (Figure 2)[55]. 3-Sialyllactosamine or 3-siallylactose that is coupled to polyacrylamide probes has been shown to bind even more Siglecs. The fact that meningococcal LOS only binds two but has identical terminal structures, shows the underlying LOS glycan structure probably also play a role in this. Of note is that Siglecs Sn or 5 do have a relaxed ligand specificity and can bind sialylated galactose connected through either an α2,3- α2,6- or α2,8- linkage[55]. It has been hypothesized that binding of LOS to Siglecs could lead to release of inhibitory signals or to uptake of the bacteria with a trojan horse like effect[55]. Truncated meningococcal LOS that instead contains a terminal GlcNAc (lgtB mutant) has been found to bind DC-SIGN, another important lectin for the immune system[56]. However, because the lgtB mutant is not phase variable it probably does not have an *in vivo* contribution to infection[57]. Interestingly, a glycan array study identified 31 interactions between LOS of immunotype L3 and L8 and host glycans, including the high-affinity binding to the Thomsen-Friedenreich antigen[58]. The downstream effects of these glycan-glycan interactions on meningococcal infection remain to be investigated. In general, the glycan mimicry by meningococcal LOS is thought to be a form of immune evasion[4,50,53]. LOS sialylation could for instance prevent the binding of antibodies of the classical pathway, or inhibit parts of the complement pathway[59]. The mimicking LOS has many interactions with the host cells that might aid in immune evasion, but the detailed mechanism of immune evasion is not clear yet.

## Neisseria gonorrhoeae

*N. gonorrhoeae* is a Gram-negative, diplococci bacteria that causes the sexually transmitted gonorrhea infection. Its LOS, similar to *Neisseria meningitidis*, contains terminal lacto-*N*-neotetraose that mimics human glycoproteins and glycosphingolipids[60]. Gonococcal LOS is phase-variable and can be sialylated, typically with an α2,3-linkage except for the P[K] LOS which contains a α2,6-linkage[60,61]. Phase-variable gonococcal LOS has multiple interactions with host cells (Figure 2) and can bind to C-type lectins MGL and DC-SIGN in case of a terminal *N*-acetylglucosamine or *N*-acetylgalactosamine, respectively[62]. Binding

to DC-SIGN results in increased IL-10 production and binding to MGL skews the population to T helper 2 cells, the latter could be a form of immune evasion[62].

Sialylation of LOS plays an important factor in the interaction of *N. gonorrhoeae* with its host. Upon invasion, only bacteria with low levels of sialylated LOS invade the host epithelial cells[63]. In absence of sialylation, the exposed terminal lactosamine residues of the LOS bind to the asialoglycoprotein receptor (ASGP-R)[64] that induces clathrin-dependent receptor mediated endocytosis, thus facilitating invasion. Once inside the host cell, the bacteria can sialylate themselves by scavenging the host's CMP-Neu5Ac, resulting in serum resistance[63]. Several reasons have been postulated for why sialylation of gonococcal LOS contributes to immune evasion, including: recruitment of factor H, shielding of underlying immunogenic structures against antibodies, and engagement of human inhibitory Siglecs[65]. Finally, it is hypothesized that *N. gonorrhoeae* has evolved to specifically use Neu5Ac to commit to a human host and to avoid innate immunity, as the immune system would be triggered by LOS with *N*-glycolylneuraminic acid (Neu5Gc)[65]. Notably, when incorporating different sialic acids on the gonococcal LOS, including Neu5Ac, Neu5Gc, Pseudaminic acid (Pse) and Legionaminic acid (Leg), it was found that both Neu5Ac and NeuGc confer high serum resistance[61].

## Nontypeable *Haemophilus influenzae*

Nontypeable *Haemophilus influenzae* (NTHi) is a Gram-negative, facultatively anaerobic, opportunistic pathogen that can cause otitis media. The LOS of nontypeable *Haemophilus influenzae* is glycan mimicking when it contains a terminal *N*-acetyllactosamine, similar to *Neisseria spp*, or a lactose that is capped with a Neu5Ac[18,66–68]. The structure of NTHi375 and RdKW20 LOS have been shown to bind galactose specific lectins through microarray technologies[68,69]. There has also been a report that NTHi LOS might have a terminal KDO, illustrating LOS heterogeneity[70]. The interactions and role of these mimicking LOS in NTHi pathogenesis are not fully understood. Heise *et al.* used a probe to study the sialylation of the LOS of NTHi[71]. Combined with a sialyltransferase inhibitor they showed decreased serum resistance when the LOS was not sialylated and thus a possible function of this in glycan mimicry[71]. The sialylation of NTHi does not lead to more factor H binding, but to less deposition of C3 on the bacterial surface[72]. Sialylated LOS might also mask the underlying glycans and thus prevent the binding of antibodies of the classical complement pathway[73], a similar observation was made for Neu5Gc[74]. However, as an indication of the complexity of bacterial glycan mimicry in relation to the host immune system,

it has also been shown that the uptake by NTHi of non-human Neu5Gc from the human diet can lead to the formation of anti-Neu5Gc antibodies that were shown to bind specifically to Neu5Gc-displaying NTHi and not to nonsialylated NTHi[75].

## *Haemophilus ducreyi*

*Haemophilus ducreyi* is a human pathogen that can cause chancroid. Similar to NTHi, the LOS can contain a terminal *N*-acetyllactosamine, which can also be sialylated[76–78]. *H. ducreyi* sialylates its LOS, by scavenging Neu5Ac from the environment since it does not have the necessary biosynthetic pathway to produce Neu5Ac[79,80]. The LOS of *H. ducreyi* has been shown to contribute to binding of human foreskin fibroblast and keratinocytes, among other components[81,82], yet it remains to be investigated whether this interaction depends on glycan mimicry. Although the LOS of *H. ducreyi* can be sialylated, this does not appear to function as a form of immune evasion, because the strain is virulent regardless of the sialylation[83]. Additionally, the bacterium is resistant against serum, so sialylation is not required for the recruitment of factor H. The sialylation of the LOS might have occurred to commit to a human host[83].

## *Klebsiella pneumoniae*

*Klebsiella pneumoniae* is a Gram-negative, non-motile, facultative anaerobic, rod-shaped bacterium that colonizes the human mucosal surfaces of the oropharynx and gastrointestinal tract. The capsule is an important virulence factor and may contain terminal Neu5Ac, as suggested for *K. pneumonia* showing the hypermucoviscosity phenotype[84], but the capsule structure has not been fully elucidated yet. At first glance, the bacteria seem to engage with inhibitory Siglec 9, but further studies are needed to confirm this binding[84]. The sialylated capsule of *K. pneumonia* could prevent complement mediated phagocytosis or might attract complement factor H to deactivate C3 [52], and would thus aid in immune evasion.

## Group B *Streptococcus*

Within the genus of *Streptococcus*, a Gram-positive, spherical bacteria, the group B *streptococcus* (GBS, *Streptococcus agalactiae*) can be classified into ten serotypes that all have a terminal α2,3- Neu5Ac. GBS serotype III dampens the host immune response[85] by binding to Siglec-9 (Figure 2)[86], an inhibitory receptor found on human neutrophils and platelets[87]. In case of platelets, Neu5Ac binds

the inhibitory Siglec-9, which suppresses platelet activation, and is considered responsible for intrinsic resistance against microbial peptides[87]. Overall, the glycan mimicry of GBS could aid its survival[85,88].

## Hyaluronan

### Group A *Streptococcus*

Group A *Streptococcus* (GAS, *Streptococcus pyogenes*) is a Gram-positive, spherical bacteria that has a capsule containing hyaluranon, a glycan which can also be found in humans as part of the extracellular matrix[89]. GAS can bind to the human inhibitory receptor Siglec-9 (Figure 2)[90]. Although Siglec-9 is a sialic acid receptor, it was reported that the hyaluranon of GAS can also bind to this receptor. The interaction between Siglec-9 and GAS is specific for hyaluranon and is not observed for glycosaminoglycans such as heparan sulfate and chrondoitin sulfate[90]. As Siglec-9 binds both Group A and B *Streptococcus*[85,86], it is involved in glycan mimicry by two bacteria with structurally different glycans that also each have distinct Siglec-9 binding sites[90]. The downstream effect of this interaction with inhibitory Siglec-9 and, in general, the effect on pathology by the glycan mimicry of GAS remains to be further studied.

## General trends and outstanding questions in bacterial glycan mimicry

The previous sections provided a brief overview of the bacteria, glycans, lectins and interactions involved in glycan mimicry by bacteria. Although these encompass a very diverse set of glycan structures and bacterial species, some general trends can be observed. First, most glycan mimicking bacteria are Gram-negative. Second, mostly facultative anaerobe mucosal pathogens display glycan mimicry. Third, Neu5Ac is a recurring monosaccharide involved in glycan mimicry. We will briefly discuss the third trend, because several published hypotheses have attempted to explain this observation.

The presence of Neu5Ac in bacterial glycoconjugates stands out and the reason for its occurrence in specific bacterial species has been the subject of much discussion, but the exact reason for its involvement in glycan mimicry by bacteria is unknown. A general hypothesis is that Neu5Ac is displayed by bacteria in glycan mimicry for its prominent terminal position on mammalian glycans and its charge. An associated question is if other terminal sialic acids found on bacteria, such as Pse and Leg, also have a function in glycan mimicry. Another hypothesis for the frequent involvement of Neu5Ac in glycan mimicry is their

interaction with Siglecs. Many mimicking bacteria interact with inhibitory Siglecs and can thereby downregulate their host's immune response[91]. In turn, their interaction with inhibitory Siglecs might have led to the evolution of activating Siglecs and perhaps also to the microbial glycan activated intelectins[49,92]. Varki has postulated that these type of interactions might be tied to the Red Queen effect, an evolutionary race between host and pathogen, which might explain the diversification of glycans[93]. The engagement of Siglecs by bacterial glycans could also play a role in masking the bacteria as 'self'. For instance, human Siglecs do not recognize the non-human Neu5Gc and this glycan has also not been found on bacteria[91,94]. This suggests that bacteria have committed to using only Neu5Ac to avoid immune recognition[95,96]. These hypotheses and the possible role of sialic acids in glycan mimicry by bacteria are part of an ongoing discussion[93,94,97].

Besides the general trends mentioned, there are still many unanswered questions about glycan mimicry by bacteria, see Box 1 for the outstanding questions highlighted by us. A major outstanding question is: what is the function of glycan mimicry? For a few bacteria their interactions and influence on the host's immune system are known, with sialylated glycan mimicry playing a prominent role, but an unequivocal conclusion for the existence and function of glycan mimicry cannot yet be made. This and other outstanding questions provide ample opportunities for research. In the final part of the review, we highlight one specific discipline, chemical glycobiology, that we believe that will contribute significantly in studying glycan mimicry by bacteria.

## Box 1: outstanding questions

- What are the minimal glycan structures required by specific bacteria to achieve functional glycan mimicry?
- Besides Neu5Ac, could other sialic acids such as Pse/Leg play a role in glycan mimicry?[45,61,141]
- To what extent does sialic acid O-acetylation[142–145] occur on mimicking bacteria and does this O-acetylation affect glycan mimicry and its function?
- What is the exact location, lifetime and substrate specificity of the native sialyltransferases used by various bacteria to sialylate their outer glycans involved in glycan mimicry?
- Can human Galectins discriminate between self and non-self glycans, and thus act as a form of immunity?[128,146]
- Do commensal bacteria also use glycan mimicry to maintain themselves at the host gut microbiota interface which is continuously exposed to our immune system?

## Chemical glycobiology techniques to study glycan mimicry

Human glycan mimicry by bacteria is a complex phenomenon that involves many interacting elements and thus needs to be studied with various approaches. Frequently used approaches are biochemistry and genetics[98,99]. Techniques from these fields can, for example, demonstrate the cross-reactivity between antibodies of the host and a microorganism, or make mutations in glycosyltransferases and glycosidases involved in LPS/LOS biosynthesis to further dissect their function. Chemical glycobiology is a powerful approach that has emerged over the past two decades to study and understand glycans and their function, and carbohydrate-related proteins. The assembly of all glycans, also bacterial, is not template-driven and this poses unique challenges in studying glycobiology[100,101]. Glycans are also generally more diverse in structure and often exert their function in a multivalent fashion as heterogeneous mixtures. This structural complexity offers an opportunity for a chemistry-based strategy to unravel the complex mechanisms at play in glycobiology. Indeed, the past two decades has seen the development of many techniques for the study of microbial glycobiology, in addition to the chemical techniques to study other bacterial components like, amino acids, lipids and peptidoglycan[102,103]. In the following sections, we will highlight several of these chemical glycobiology techniques (Figure 3) and report on their application or speculate on their potential use to study bacterial mimicry of human glycans.

## Metabolic oligosaccharide engineering

Metabolic Oligosaccharide Engineering (MOE) is a technique that uses the cells' own metabolic salvage pathway to incorporate an externally administered monosaccharide-based probe into glycoconjugates. These monosaccharide-based probes typically contain an unnatural functional group that can be covalently coupled through a click reaction to a reporter group, for example biotin or a fluorophore, after incorporation into the bacterial glycans. It can serve as a technique to detect the presence of a specific monosaccharide in a complex glycoconjugates (e.g., LPS) *in vivo* or *in vitro*. Due to the possibility of attaching artificial groups to the labeled glycan, it can also be used to engineer the target glycoconjugate to investigate its structure-activity relationship or perturb an interaction with an associated lectin.

MOE has been performed with glycans that are unique for bacteria[7,8], for example bacillosamine in *N*-linked glycans of *C. jejuni*, Legionaminic acid (Leg) in *L. Pneumophilia*[104] and *C. jejuni* [unpublished data by our group] and Pseudaminic acid (Pse) in *C. jejuni, P. aeruginosa, A. baumannii, V. vulnificus*[105,106]. Two groups of

unique bacterial glycans that are especially interesting to study glycan mimicry are: 1) monosaccharides that are frequently found on the core of the LOS, like KDO and heptose, and 2) nonulosonic acids, like the previously mentioned Pse and Leg, to dissect the role of these microbial sialic acids further[8,107].

The extensive monosaccharide 'toolbox' with different chemical reporters that has been developed for MOE in mammalian cells can be used to study the role of *N*-acetylneuraminic acid and other human monosaccharides in glycan mimicry by bacteria[108,109].

MOE is often performed with monosaccharides that have a chemical reporter, which can reveal the incorporation into a potentially unknown glycoconjugate or the location of the labeled bacteria (see also imaging). A useful way to influence this process is to make use of inhibitors that can perturb the transferase enzymes of the glycan biosynthesis pathway. An illustrative example of such an approach was the study of LOS sialylation by NTHi which used a Neu5Ac probe in combination with a metabolic sialyltransferase inhibitor[71]. By itself, the development of activity-based probes and (metabolic) inhibitors to perturb, detect and identify the glycosyl hydrolase and transferase enzymes involved in the processing of bacterial glycans is an important emerging research direction towards understanding bacterial glycobiology[110–112].

MOE has two main drawbacks. One drawback is that the efficiency of MOE can vary a lot per bacterial species or strain, for instance bacteria can also degrade the probes for growth, which would require optimization per 'tool' [113]. Second, a probe is processed through the bacteria's metabolic pathway, which leads to incorporation into various types of glycoconjugates with limited control over this. An advantage of MOE is that the technique has been fairly well developed for microbiology and many of the monosaccharide probes and associated chemical reporters are commercially available. MOE provides exciting opportunities to study glycan mimicry and its function, and possibly might also provide leads for future glycotherapies[8,114].

## Enzymatic methods

Another approach to glycoengineer and thereby study the composition and function of complex glycoconjugates in a living organism is to make use of enzymes, here we describe six of these techniques. First, the technique called selective exoenzymatic labeling (SEEL), also called chemoenzymatic labeling, has been used to track *N*-glycans, identify human glycoproteins and introduce a large biomolecule[115–117]. The technique has mostly been applied to mammalian cells thus far but can be used for glycoengineering of bacteria (Chapter 3). A

detailed review of the different enzymes available and their substrate scope has been reported[118].

A second engineering technique is to remove subclasses of *N*-glycans on mammalian cells by trimming them with the appropriate glycosidases and insert whole *N*-glycans with a chemical handle, an oxazoline[119]. This technique could potentially homogenize the cell surface. This approach could also be used to homogenize the LOS of bacteria, although this would require extensive engineering of the enzymes and glycans involved.

A third technique studies the function of glycosyltransferases and their acceptors through bump-and-hole engineering, which employs a complementary set of modified enzymes and substrates that are not naturally present in a cell. Schumann et al.[120] applied this to *N*-acetylgalactosaminyl transferases to profile the targets of this family of enzymes. This technique can be used to engineer the cell surface in a targeted way with glycans or clickable groups, but also to scan for substrate specificity or labeling without the interference of other native bacterial transferases.

Most of the developed enzymatic glycoengineering techniques have only been applied to mammalian cells, but there are a few examples for bacteria. One study screened the substrate promiscuity of several enzymes involved in the biosynthesis of the heptasaccharide on the *N*-glycans of *C. jejuni* to introduce azides[121]. Attempts to label *in vivo* failed, because the labeled monosaccharide precursors were not converted to the corresponding nucleotide sugars[121]. In another study the enzymatic synthesis of a glycolipid was developed to incorporate an arabinose onto the cell surface of mycobacterium[122]. This strategy could, for example, also be applied to study the glycolipid portion of the LOS that is involved in mimicry.

The LOS of bacteria can also be glycoengineered through genetically introduced transferases[123]. This approach was applied to engineer the *O*-antigen of *E. coli* and *Salmonella enterica* serovar Typhimurium[123]. This technique can build a poly LacNAc unit, or a Lewis x motive, on a truncated lipid A core, and therefore the LOS becomes glycan mimicking.
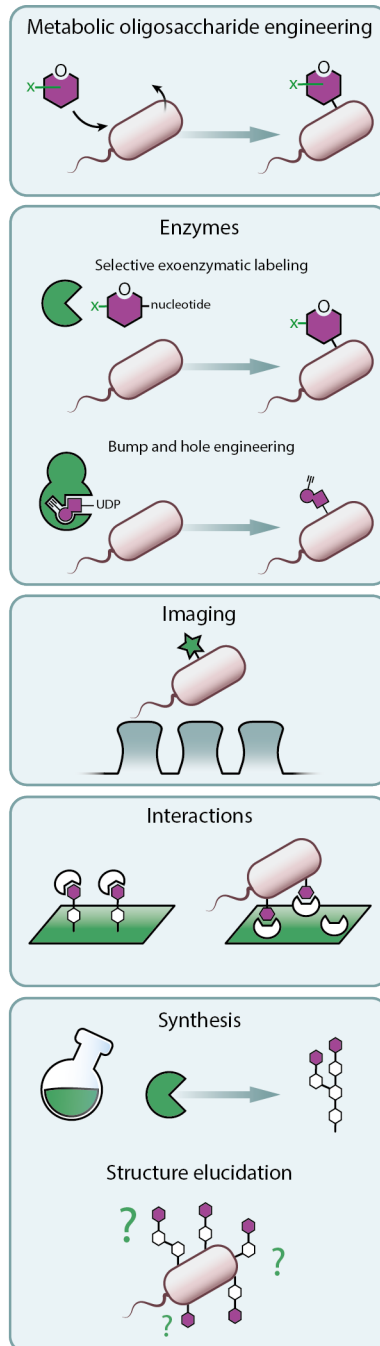
Figure 3. Chemical glycobiology techniques to study glycan mimicry on bacteria and their host.

## Imaging

In the past, labeled lectins or antibodies have often been used to indirectly detect and image certain glycans on the cell surface of bacteria. However, a powerful way to study glycan mimicry would be to directly image the actual glycans themselves. The Kasper group applied MOE to fluorescently label and track commensal bacteria in the gut via their cell surface glycans[124]. In another study, they labeled multiple glycoconjugates of the cell wall: peptidoglycan, LPS and CPS, with different fluorophores and tracked these bacteria[125]. Bacteria involved in glycan mimicry could also be labeled as such and visualized, for example on their LOS or LPS[126,127]. Imaging and spatiotemporal tracking of a specific glycan on a commensal or pathogenic bacterium in the presence of a mammalian cell model or even *in vivo* in an animal model (with for instance near-IR dyes) could provide valuable information about their location and their possible interaction partners.

## Interactions

Glycan microarrays are an important technique to study bacterial glycans and their interactions. One approach to study glycan mimicry with microarrays is to investigate the binding of mimicking glycans to lectins of the innate immune system[128]. Another approach is to study the binding of unknown bacterial glycans to lectins on a microarray in order to glycophenotype the bacterial glycan coat[129–131]. The presence of mimicking glycan structures on commensal bacteria might be elucidated with glycophenotyping. For a detailed overview of the application of microarrays to study both bacterial glycans and their interactions, we refer the reader to a recent review[132].

In case the interactions between bacterial glycans and receptors are known, it might be possible to manipulate this interaction and steer it to a receptor of choice. In a recent study, Natural Killer cells were modified via MOE with a set of Neu5Ac's that contained unnatural groups, of which some enhanced the binding to a tumor cell receptor that in turn led to increased tumor lysis[133]. When studying a specific glycan-receptor interaction multivalency effects play an important role and can be investigated with defined multimers of the bacterial glycans[134]. In case of glycan mimicry by bacteria a more fundamental application can be thought of where the binding to different receptors and the effect on the immune system can be studied, for example the binding of activating or inhibiting Siglecs by *Neisseria gonorrhoeae*[65]. A set of molecules containing bio-orthogonal click groups is available to increase binding to certain Siglecs[135].

In case the *in vivo* interactions between bacterial glycans and receptors or

binding partners are unknown, a photoreactive group, like a diazirine, might be installed on the bacterial glycans to photocrosslink nearby receptors[113] and use proteomics to identify these binding partners.

## Synthesis and structure elucidation

Progress to make glycans more accessible through automated glycan synthesis, both chemically and enzymatically, will speed up the study of glycobiology[101,136]. Some examples that are of special interest for glycan mimicry are gangliosides and poly-LacNAc repeats that were synthesized through automated synthesis in combination with enzymes[137], or the LOS of *C. jejuni* that was synthesized through chemoenzymatic synthesis[138].

Another fundamental approach in glycobiology is structure elucidation. This is typically performed through a range of techniques, such as monomer analysis, MS and NMR[139]. Mass spectrometry analysis of glycans can require a lot of expertise. Improvements can be made in purification before measuring, mass labels and annotation of the data[140].

# Summary

Bacteria are covered by glycans. In the cases described here, and possibly even more, bacteria display glycans that mimic glycans found in their human or animal host. This glycan mimicry is often found for Gram-negative and mucosal pathogens. Bacterial glycans are important for many interactions that could possibly result in infection. For bacterial glycans mimicking their host, an increasing number of interaction partners have been identified, yet their effects beyond glycan recognition by host receptors are not always clear. In general, the overall function of glycan mimicry by bacteria remains to be studied. The techniques developed for chemical glycobiology could aid in identifying bacteria that display glycan mimicry and elucidate the function of glycan mimicry.

# Acknowledgements

# References

(1) Poole, J.; Day, C. J.; von Itzstein, M.; Paton, J. C.; Jennings, M. P. Glycointeractions in Bacterial Pathogenesis. *Nat. Rev. Microbiol.* **2018**, *16*, 440–452.

(2) Damian, R. T. Molecular Mimicry: Antigen Sharing by Parasite and Host and Its Consequences. *Am. Nat.* **1964**, *XCVIII*, 129–149.

(3) Moran, A. P.; Prendergast, M. M.; Appelmelk, B. J. Molecular Mimicry of Host Structures by Bacterial Lipopolysaccharides and Its Contribution to Disease. *FEMS Immunol. Med. Microbiol.* **1996**, *16*, 105–115.

(4) Mandrell, R. E.; Apicella, M. A. Lipo-Oligosaccharides (LOS) of Mucosal Pathogens: Molecular Mimicry and Host-Modification of LOS. *Immunobiology* **1993**, *187*, 382–402.

(5) Varki, A.; Cummings, R. D.; Aebi, M.; Packer, N. H.; Seeberger, P. H.; Esko, J. D.; Stanley, P.; Hart, G.; Darvill, A.; Kinoshita, T.; Prestegard, J. J.; Schnaar, R. L.; Freeze, H. H.; Marth, J. D.; Bertozzi, C. R.; Etzler, M. E.; Frank, M.; Vliegenthart, J. F. G.; Lütteke, T.; Perez, S.; Bolton, E.; Rudd, P.; Paulson, J.; Kanehisa, M.; Toukach, P.; Aoki-Kinoshita, K. F.; Dell, A.; Narimatsu, H.; York, W.; Taniguchi, N.; Kornfeld, S. Symbol Nomenclature for Graphical Representations of Glycans. *Glycobiology* **2015**, *25*, 1323–1324.

(6) Carolin Frank, A. Molecular Host Mimicry and Manipulation in Bacterial Symbionts. *FEMS Microbiol. Lett.* **2019**, *366*, fnz038.

(7) Dube, D. H.; Champasa, K.; Wang, B. Chemical Tools to Discover and Target Bacterial Glycoproteins. *Chem. Commun.* **2011**, *47*, 87–101.

(8) Tra, V. N.; Dube, D. H. Glycans in Pathogenic Bacteria-Potential for Targeted Covalent Therapeutics and Imaging Agents. *Chem. Commun.* **2014**, *50*, 4659–4673.

(9) Imperiali, B. Bacterial Carbohydrate Diversity — a Brave New World. *Curr. Opin. Chem. Biol.* **2019**, *53*, 1–8.

(10) Van Den Berg, B.; Walgaard, C.; Drenthen, J.; Fokke, C.; Jacobs, B. C.; Van Doorn, P. A. Guillain-Barré Syndrome: Pathogenesis, Diagnosis, Treatment and Prognosis. *Nat. Rev. Neurol.* **2014**, *10*, 469–482.

(11) Jasti, A. K.; Selmi, C.; Sarmiento-Monroy, J. C.; Vega, D. A.; Anaya, J. M.; Gershwin, M. E. Guillain-Barré Syndrome: Causes, Immunopathogenic Mechanisms and Treatment. *Expert Rev. Clin. Immunol.* **2016**, *12*, 1175–1189.

(12) Yu, R. K.; Usuki, S.; Ariga, T. Ganglioside Molecular Mimicry and Its Pathological Roles in Guillain-Barré Syndrome and Related Diseases. *Infect. Immun.* **2006**, *74*, 6517–6527.

(13) Moran, A. P.; Prendergast, M. M. Molecular Mimicry in *Campylobacter Jejuni* and *Helicobacter Pylori* Lipopolysaccharides: Contribution of Gastrointestinal Infections to Autoimmunity. In *Journal of Autoimmunity*; 2001; Vol. 16, pp 241–256.

(14) Prendergast, M. M.; Moran, A. P. Lipopolysaccharides in the Development of the Guillain-Barré Syndrome and Miller Fisher Syndrome Forms of Acute Inflammatory Peripheral Neuropathies. *J. Endotoxin Res.* **2000**, *6*, 341–359.

(15) Yuki, N.; Susuki, K.; Koga, M.; Nishimoto, Y.; Odaka, M.; Hirata, K.; Taguchi, K.; Miyatake, T.; Furukawa, K.; Kobata, T.; Yamada, M. Carbohydrate Mimicry between Human Ganglioside GM1 and *Campylobacter Jejuni* Lipooligosaccharide Causes Guillain-Barré Syndrome. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 11404–11409.

(16) Van Belkum, A.; Jacobs, B.; Van Beek, E.; Louwen, R.; Van Rijs, W.; Debruyne, L.; Gilbert, M.; Li, J.; Jansz, A.; Mégraud, F.; Endtz, H. Can *Campylobacter Coli* Induce Guillain-Barré Syndrome? *Eur. J. Clin. Microbiol. Infect. Dis.* **2009**, *28*, 557–560.

(17) Godschalk, P. C. R.; Kuijf, M. L.; Li, J.; St. Michael, F.; Ang, C. W.; Jacobs, B. C.; Karwaski, M. F.; Brochu, D.; Moterassed, A.; Endtz, H. P.; Van Belkum, A.; Gilbert, M. Structural Characterization of *Campylobacter Jejuni* Lipooligosaccharide Outer Cores Associated with Guillain-Barré and Miller Fisher Syndromes. *Infect. Immun.* **2007**, *75*, 1245–1254.

(18) Day, C. J.; Tran, E. N.; Semchenko, E. A.; Tram, G.; Hartley-Tassell, L. E.; Ng, P. S. K.; King, R. M.; Ulanovsky, R.; McAtamney, S.; Apicella, M. A.; Tiralongo, J.; Morona, R.; Korolik, V.; Jennings, M. P. Glycan:Glycan Interactions: High Affinity Biomolecular Interactions That Can Mediate Binding of Pathogenic Bacteria to Host Cells. *Proc. Natl. Acad. Sci.* **2015**, *112*, E7266–E7275.

(19) Avril, T.; Wagner, E. R.; Willison, H. J.; Crocker, P. R. Sialic Acid-Binding Immunoglobulin-like Lectin 7 Mediates Selective Recognition of Sialylated Glycans Expressed on *Campylobacter Jejuni* Lipooligosaccharides. *Infect. Immun.* **2006**, *74*, 4133–4141.

(20) Heikema, A. P.; Jacobs, B. C.; Horst-Kreft, D.; Huizinga, R.; Kuijf, M. L.; Endtz, H. P.; Samsom, J. N.; van Wamel, W. J. B. Siglec-7 Specifically Recognizes *Campylobacter Jejuni* Strains Associated with Oculomotor Weakness in Guillain-Barré Syndrome and Miller Fisher Syndrome. *Clin. Microbiol. Infect.* **2013**, *19*, E106–E112.

(21) van Sorge, N. M.; Bleumink, N. M. C.; van Vliet, S. J.; Saeland, E.; van der Pol, W. L.; van Kooyk, Y.; Van Putten, J. P. M. N-Glycosylated Proteins and Distinct Lipooligosaccharide Glycoforms of *Campylobacter Jejuni*

**2**

Target the Human C-Type Lectin Receptor MGL. *Cell. Microbiol.* **2009**, *11*, 1768–1781.

(22)  Heikema, A. P.; Bergman, M. P.; Richards, H.; Crocker, P. R.; Gilbert, M.; Samsom, J. N.; Van Wamel, W. J. B.; Endtz, H. P.; Van Belkum, A. Characterization of the Specific Interaction between Sialoadhesin and Sialylated *Campylobacter Jejuni* Lipooligosaccharides. *Infect. Immun.* **2010**, *78*, 3237–3246.

(23)  Heikema, A. P.; Koning, R. I.; Rico, S. D. dos S.; Rempel, H.; Jacobs, B. C.; Endtz, H. P.; van Wamel, W. J. B.; Samsom, J. N. Enhanced, Sialoadhesin-Dependent Uptake of Guillain-Barré Syndrome-Associated *Campylobacter Jejuni* Strains by Human Macrophages. *Infect. Immun.* **2013**, *81*, 2095–2103.

(24)  Perera, V. N.; Nachamkin, I.; Ung, H.; Patterson, J. H.; McConville, M. J.; Coloe, P. J.; Fry, B. N. Molecular Mimicry in *Campylobacter Jejuni*: Role of the Lipo-Oligosaccharide Core Oligosaccharide in Inducing Anti-Ganglioside Antibodies. *FEMS Immunol. Med. Microbiol.* **2007**, *50*, 27–36.

(25)  Xiang, S. L.; Zhong, M.; Cai, F. C.; Deng, B.; Zhang, X. P. The Sialic Acid Residue Is a Crucial Component of *C. Jejuni* Lipooligosaccharide Ganglioside Mimicry in the Induction Guillain-Barré Syndrome. *J. Neuroimmunol.* **2006**, *174*, 126–132.

(26)  Gilbert, M.; Karwaski, M. F.; Bernatchez, S.; Young, N. M.; Taboada, E.; Michniewicz, J.; Cunningham, A. M.; Wakarchuk, W. W. The Genetic Bases for the Variation in the Lipo-Oligosaccharide of the Mucosal Pathogen, *Campylobacter Jejuni*. Biosynthesis of Sialylated Ganglioside Mimics in the Core Oligosaccharide. *J. Biol. Chem.* **2002**, *277*, 327–337.

(27)  Louwen, R.; Heikema, A.; Van Belkum, A.; Ott, A.; Gilbert, M.; Ang, W.; Endtz, H. P.; Bergman, M. P.; Nieuwenhuis, E. E. The Sialylated Lipooligosaccharide Outer Core in *Campylobacter Jejuni* Is an Important Determinant for Epithelial Cell Invasion. *Infect. Immun.* **2008**, *76*, 4431–4438.

(28)  Louwen, R.; Nieuwenhuis, E. E. S.; van Marrewijk, L.; Horst-Kreft, D.; de Ruiter, L.; Heikema, A. P.; van Wamel, W. J. B.; Wagenaar, J. A.; Endtz, H. P.; Samsom, J.; van Baarlen, P.; Akhmanov, A.; van Belkum, A. *Campylobacter Jejuni* Translocation across Intestinal Epithelial Cells Is Facilitated by Ganglioside-like Lipooligosaccharide Structures. *Infect. Immun.* **2012**, *80*, 3307–3318.

(29)  Ellström, P.; Feodoroff, B.; Hänninen, M. L.; Rautelin, H. Lipooligosaccharide Locus Class of *Campylobacter Jejuni*: Sialylation Is Not Needed for Invasive Infection. *Clin. Microbiol. Infect.* **2014**, *20*, 524–529.

(30)  Day, C. J.; Semchenko, E. A.; Korolik, V. Glycoconjugates Play a Key Role in *Campylobacter Jejuni* Infection: Interactions between Host and Pathogen. *Front. Cell. Infect. Microbiol.* **2012**, *2*, 9.

(31)  Masoud, H.; Perry, M. B.; Brisson, J.-R.; Uhrin, D.; Richards, J. C. Structural Elucidation of the Backbone Oligosaccharide from the Lipopolysaccharide of *Moraxella Catarrhalis* Serotype A. *Can. J. Chem.* **1994**, *72*, 1466–1477.

(32)  Schwingel, J. M.; Edwards, K. J.; Cox, A. D.; Masoud, H.; Richards, J. C.; St. Michael, F.; Tekwe, C. D.; Sethi, S.; Murphy, T. F.; Campagnari, A. A. Use of *Moraxella Catarrhalis* Lipooligosaccharide Mutants to Identify Specific Oligosaccharide Epitopes Recognized by Human Serum Antibodies. *Infect. Immun.* **2009**, *77*, 4548–4558.

(33)  Holme, T.; Rahman, M.; Jansson, P. E.; Widmalm, G. The Lipopolysaccharide of *Moraxella Catarrhalis*. Structural Relationships and Antigenic Properties. *Eur. J. Biochem.* **1999**, *265*, 524–529.

(34)  Appelmelk, B. J.; Monteiro, M. A.; Martin, S. L.; Moran, A. P.; Vandenbroucke-Grauls, C. M. J. E. Why *Helicobacter Pylori* Has Lewis Antigens. *Trends in Microbiology*. 2000, pp 565–570.

(35)  Chmiela, M.; Gonciarz, W. Molecular Mimicry in *Helicobacter Pylori* Infections. *World J. Gastroenterol.* **2017**, *23*, 3964–3977.

(36)  Bergman, M.; Del Prete, G.; van Kooyk, Y.; Appelmelk, B. *Helicobacter Pylori* Phase Variation, Immune Modulation and Gastric Autoimmunity. *Nat. Rev. Microbiol.* **2006**, *4*, 151–159.

(37)  Hynes, S. O.; Ferris, J. A.; Szponar, B.; Wadström, T.; Fox, J. G.; O'Rourke, J.; Larsson, L.; Yaquian, E.; Ljungh, Å.; Clyne, M.; Andersen, L. P.; Moran, A. P. Comparative Chemical and Biological Characterization of the Lipopolysaccharides of Gastric and Enterohepatic *Helicobacters*. *Helicobacter* **2004**, *9*, 313–323.

(38)  Appelmelk, B. J.; Simoons-Smit, I.; Negrini, R.; Moran, A. P.; Aspinall, G. O.; Forte, J. G.; De Vries, T.; Quan, H.; Verboom, T.; Maaskant, J. J.; Ghiara, P.; Kuipers, E. J.; Bloemena, E.; Tadema, T. M.; Townsend, R. R.; Tyagarajan, K.; Crothers, J. M.; Monteiro, M. A.; Savio, A.; De Graaff, J. Potential Role of Molecular Mimicry between *Helicobacter Pylori* Lipopolysaccharide and Host Lewis Blood Group Antigens in Autoimmunity. *Infect. Immun.* **1996**, *64*, 2031–2040.

(39)  Leker, K.; Lozano-Pope, I.; Bandyopadhyay, K.; Choudhury, B. P.; Obonyo, M. Comparison of Lipopolysaccharides Composition of Two Different Strains of *Helicobacter Pylori*. *BMC Microbiol.* **2017**, *17*, 226.

(40)  Logan, S. M.; Conlan, J. W.; Monteiro, M. A.; Wakarchuk, W. W.; Altman, E. Functional Genomics of *Helicobacter Pylori*: Identification of a β-1,4 Galactosyltransferase and Generation of Mutants with Altered Lipopolysaccharide. *Mol. Microbiol.* **2000**, *35*, 1156–1167.

(41)  Chmiela, M.; Miszczyk, E.; Rudnicka, K. Structural Modifications of *Helicobacter Pylori* Lipopolysaccharide: An Idea for How to Live in Peace. *World J. Gastroenterol.* **2014**, *20*, 9882–9897.

(42) Li, H.; Liao, T.; Debowski, A. W.; Tang, H.; Nilsson, H. O.; Stubbs, K. A.; Marshall, B. J.; Benghezal, M. Lipopolysaccharide Structure and Biosynthesis in *Helicobacter Pylori*. *Helicobacter* **2016**, *21*, 445–461.

(43) Li, H.; Yang, T.; Liao, T.; Debowski, A. W.; Nilsson, H. O.; Haslam, S. M.; Dell, A.; Stubbs, K. A.; Marshall, B. J.; Benghezal, M. The Redefinition of *Helicobacter Pylori* Lipopolysaccharide O-Antigen and Core-Oligosaccharide Domains. *PLoS Pathog.* **2017**, *4*, 175–178.

(44) Khamri, W.; Moran, A. P.; Worku, M. L.; Karim, Q. N.; Walker, M. M.; Annuk, H.; Ferris, J. A.; Appelmelk, B. J.; Eggleton, P.; Reid, K. B. M.; Thursz, M. R. Variations in *Helicobacter Pylori* Lipopolysaccharide to Evade the Innate Immune Component Surfactant Protein D. *Infect. Immun.* **2005**, *73*, 7677–7686.

(45) Stenutz, R.; Weintraub, A. The Structures of *Escherichia Coli* O-Polysaccharide Antigens. *FEMS Microbiol. Rev.* **2006**, *30*, 382–403.

(46) Suerbaum, S.; Friedrich, S.; Leying, H.; Opferkuch, W. Expression of Capsular Polysaccharide Determines Serum Resistance in *Escherichia Coli* K92. *Zentralblatt fur Bakteriol.* **1994**, *281*, 146–157.

(47) Steenbergen, S. M.; Wrona, T. J.; Vimr, E. R. Functional Analysis of the Sialyltransferase Complexes in *Escherichia Coli* K1 and K92. *J. Bacteriol.* **1992**, *174*, 1099–1108.

(48) Lizak, C.; Worrall, L. J.; Baumann, L.; Pfleiderer, M. M.; Volkers, G.; Sun, T.; Sim, L.; Wakarchuk, W.; Withers, S. G.; Strynadka, N. C. J. X-Ray Crystallographic Structure of a Bacterial Polysialyltransferase Provides Insight into the Biosynthesis of Capsular Polysialic Acid. *Sci. Rep.* **2017**, *7*, 5842.

(49) Schwarz, F.; Landig, C. S.; Siddiqui, S.; Secundino, I.; Olson, J.; Varki, N.; Nizet, V.; Varki, A. Paired Siglec Receptors Generate Opposite Inflammatory Responses to a Human-specific Pathogen. *EMBO J.* **2017**, *36*, 751–760.

(50) Mubaiwa, T. D.; Semchenko, E. A.; Hartley-Tassell, L. E.; Day, C. J.; Jennings, M. P.; Seib, K. L. The Sweet Side of the Pathogenic *Neisseria*: The Role of Glycan Interactions in Colonisation and Disease. *Pathogens and Disease.* 2017, p ftx063.

(51) Jarvis, G. A.; Vedros, N. A. Sialic Acid of Group B *Neisseria Meningitidis* Regulates Alternative Complement Pathway Activation. *Infect. Immun.* **1987**, *55*, 174–180.

(52) Doorduijn, D. J.; Rooijakkers, S. H. M.; Schaik, W. Van; Bardoel, B. W. Complement Resistance Mechanisms of *Klebsiella Pneumoniae*. *Immunobiology* **2016**, *221*, 1102–1109.

(53) Tsai, C. M.; Chen, W. H.; Balakonis, P. A. Characterization of Terminal NeuNAcα2-3Galβ1-4GlcNAc Sequence in Lipooligosaccharides of *Neisseria Meningitidis*. *Glycobiology* **1998**, *8*, 359–365.

(54) Tsai, C. M.; Kao, G.; Zhu, P. Influence of the Length of the Lipooligosaccharide α Chain on Its Sialylation in *Neisseria Meningitidis*. *Infect. Immun.* **2002**, *70*, 407–411.

(55) Jones, C.; Virji, M.; Crocker, P. R. Recognition of Sialylated Meningococcal Lipopolysaccharide by Siglecs Expressed on Myeloid Cells Leads to Enhanced Bacterial Uptake. *Mol. Microbiol.* **2003**, *49*, 1213–1225.

(56) Steeghs, L.; van Vliet, S. J.; Uronen-Hansson, H.; van Mourik, A.; Engering, A.; Sanchez-Hernandez, M.; Klein, N.; Callard, R.; van Putten, J. P. M.; van der Ley, P.; van Kooyk, Y.; van de Winkel, J. G. J. *Neisseria Meningitidis* Expressing IgtB Lipopolysaccharide Targets DC-SIGN and Modulates Dendritic Cell Function. *Cell. Microbiol.* **2006**, *8*, 316–325.

(57) Johswich, K. Innate Immune Recognition and Inflammation in *Neisseria Meningitidis* Infection. *Pathog. Dis.* **2017**, *75*, ftx022.

(58) Mubaiwa, T. D.; Hartley-Tassell, L. E.; Semchenko, E. A.; Jen, F. E. C.; Srikhanta, Y. N.; Day, C. J.; Jennings, M. P.; Seib, K. L. The Glycointeractome of Serogroup B *Neisseria Meningitidis* Strain MC58. *Sci. Rep.* **2017**, *7*, 5693.

(59) Lewis, L. A.; Carter, M.; Ram, S. The Relative Roles of Factor H Binding Protein, Neisserial Surface Protein A, and Lipooligosaccharide Sialylation in Regulation of the Alternative Pathway of Complement on Meningococci. *J. Immunol.* **2012**, *188*, 5063–5072.

(60) Tong, Y.; Arking, D.; Ye, S.; Reinhold, B.; Reinhold, V.; Stein, D. C. *Neisseria Gonorrhoeae* Strain PID2 Simultaneously Expresses Six Chemically Related Lipooligosaccharide Structures. *Glycobiology* **2002**, *12*, 523–533.

(61) Gulati, S.; Schoenhofen, I. C.; Whitfield, D. M.; Cox, A. D.; Li, J.; St. Michael, F.; Vinogradov, E. V.; Stupak, J.; Zheng, B.; Ohnishi, M.; Unemo, M.; Lewis, L. A.; Taylor, R. E.; Landig, C. S.; Diaz, S.; Reed, G. W.; Varki, A.; Rice, P. A.; Ram, S. Utilizing CMP-Sialic Acid Analogs to Unravel *Neisseria Gonorrhoeae* Lipooligosaccharide-Mediated Complement Resistance and Design Novel Therapeutics. *PLoS Pathog.* **2015**, *11*, e1005290.

(62) Van Vliet, S. J.; Steeghs, L.; Bruijns, S. C. M.; Vaezirad, M. M.; Blok, C. S.; Arenas Busto, J. A.; Deken, M.; Van Putten, J. P. M.; Van Kooyk, Y. Variation of *Neisseria Gonorrhoeae* Lipooligosaccharide Directs Dendritic Cell-Induced T Helper Responses. *PLoS Pathog.* **2009**, *5*, e1000625.

(63) van Putten, J. P. Phase Variation of Lipopolysaccharide Directs Interconversion of Invasive and Immuno-Resistant Phenotypes of *Neisseria Gonorrhoeae*. *EMBO J.* **1993**, *12*, 4043–4051.

(64) Harvey, H. A.; Jennings, M. P.; Campbell, C. A.; Williams, R.; Apicella, M. A. Receptor-Mediated Endocytosis

of *Neisseria Gonorrhoeae* into Primary Human Urethral Epithelial Cells: The Role of the Asialoglycoprotein Receptor. *Mol. Microbiol.* **2001**, *42*, 659–672.

(65) Landig, C. S.; Hazel, A.; Kellman, B. P.; Fong, J. J.; Schwarz, F.; Agarwal, S.; Varki, N.; Massari, P.; Lewis, N. E.; Ram, S.; Varki, A. Evolution of the Exclusively Human Pathogen *Neisseria Gonorrhoeae*: Human-Specific Engagement of Immunoregulatory Siglecs. *Evol. Appl.* **2019**, *12*, 337–349.

(66) Apicella, M. A. Nontypeable *Haemophilus Influenzae*: The Role of N-Acetyl-5-Neuraminic Acid in Biology. *Front. Cell. Infect. Microbiol.* **2012**, *2*, 19.

(67) Phillips, N. J.; Apicella, M. A.; Griffiss, J. M. L.; Gibson, B. W. Structural Characterization of the Cell Surface Lipooligosaccharides from a Nontypable Strain of *Haemophilus Influenzae*. *Biochemistry* **1992**, *31*, 4515–4526.

(68) Kalograiaki, I.; Euba, B.; Proverbio, D.; Campanero-Rhodes, M. A.; Aastrup, T.; Garmendia, J.; Solís, D. Combined Bacteria Microarray and Quartz Crystal Microbalance Approach for Exploring Glycosignatures of Nontypeable *Haemophilus Influenzae* and Recognition by Host Lectins. *Anal. Chem.* **2016**, *88*, 5950–5957.

(69) Kalograiaki, I.; Euba, B.; Fernández-Alonso, M. del C.; Proverbio, D.; St. Geme, J. W.; Aastrup, T.; Garmendia, J.; Cañada, F. J.; Solís, D. Differential Recognition of *Haemophilus Influenzae* Whole Bacterial Cells and Isolated Lipooligosaccharides by Galactose-Specific Lectins. *Sci. Rep.* **2018**, *8*, 16292.

(70) Apicella, M. A.; Coffin, J.; Ketterer, M.; Post, D. M. B.; Day, C. J.; Jen, F. E.; Jennings, M. P. Nontypeable *Haemophilus Influenzae* Lipooligosaccharide Expresses a Terminal Ketodeoxyoctanoate *in Vivo*, Which Can Be Used as a Target for Bactericidal Antibody. *Mol. Microbiol.* **2018**, *9*, e01401-18.

(71) Heise, T.; Langereis, J. D.; Rossing, E.; de Jonge, M. I.; Adema, G. J.; Büll, C.; Boltje, T. J. Selective Inhibition of Sialic Acid-Based Molecular Mimicry in *Haemophilus Influenzae* Abrogates Serum Resistance. *Cell Chem. Biol.* **2018**, *25*, 1279-1285.e8.

(72) Figueira, M. A.; Ram, S.; Goldstein, R.; Hood, D. W.; Moxon, E. R.; Pelton, S. I. Role of Complement in Defense of the Middle Ear Revealed by Restoring the Virulence of Nontypeable *Haemophilus Influenzae SiaB* Mutants. *Infect. Immun.* **2007**, *75*, 325–333.

(73) Jackson, M. D.; Wong, S. M.; Akerley, B. J. Underlying Glycans Determine the Ability of Sialylated Lipooligosaccharide to Protect Nontypeable *Haemophilus Influenzae* from Serum IgM and Complement. *Infect. Immun.* **2019**, *87*, e00456-19.

(74) Oerlemans, M. M. P.; Moons, S. J.; Heming, J. J. A.; Boltje, T. J.; De Jonge, M. I.; Langereis, J. D. Uptake of Sialic Acid by Nontypeable *Haemophilus Influenzae* Increases Complement Resistance through Decreasing IgM-Dependent Complement Activation. *Infect. Immun.* **2019**, *87*, e00077-19.

(75) Taylor, R. E.; Gregg, C. J.; Padler-Karavani, V.; Ghaderi, D.; Yu, H.; Huang, S.; Sorensen, R. U.; Chen, X.; Inostroza, J.; Nizet, V.; Varki, A. Novel Mechanism for the Generation of Human Xeno-Autoantibodies against the Nonhuman Sialic Acid *N*-Glycolylneuraminic Acid. *J. Exp. Med.* **2010**, *207*, 1637–1646.

(76) Schweda, E. K. H.; Sundström, A. C.; Eriksson, L. M.; Jonasson, J. A.; Lindberg, A. A. Structural Studies of the Cell Envelope Lipopolysaccharides from *Haemophilus Ducreyi* Strains ITM 2665 and ITM 4747. *J. Biol. Chem.* **1994**, *269*, 12040–12048.

(77) Melaugh, W.; Phillips, N. J.; Campagnari, A. A.; Tullius, M. V; Gibson, B. W. Structure of the Major Oligosaccharide from the Lipooligosaccharide of *Haemophilus Ducreyi* Strain 35000 and Evidence for Additional Glycoforms. *Biochemistry* **1994**, *33*, 13070–13078.

(78) Melaugh, W.; Campagnari, A. A.; Gibson, B. W. The Lipooligosaccharides of *Haemophilus Ducreyi* Are Highly Sialylated. *J. Bacteriol.* **1996**, *178*, 564–570.

(79) Wratil, P. R.; Horstkorte, R.; Reutter, W. Metabolic Glycoengineering with N-Acyl Side Chain Modified Mannosamines. *Angew. Chem - Int. Ed.* **2016**, *55*, 9482–9512.

(80) Goon, S.; Schilling, B.; Tullius, M. V.; Gibson, B. W.; Bertozzi, C. R. Metabolic Incorporation of Unnatural Sialic Acids into *Haemophilus Ducreyi* Lipooligosaccharides. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 3089–3094.

(81) Alfa, M. J.; Degagne, P. Attachment of *Haemophilus Ducreyi* to Human Foreskin Fibroblasts Involves LOS and Fibronectin. *Microb. Pathog.* **1997**, *22*, 39–46.

(82) Gibson, B. W.; Campagnari, A. A.; Melaugh, W.; Phillips, N. J.; Apicella, M. A.; Grass, S.; Wang, J.; Palmer, K. L.; Munson, R. S. Characterization of a Transposon Tn 916 -Generated Mutant of *Haemophilus Ducreyi* 35000 Defective in Lipooligosaccharide Biosynthesis. *J. Bacteriol.* **1997**, *179*, 5062–5071.

(83) Spinola, S. M.; Li, W.; Fortney, K. R.; Janowicz, D. M.; Zwickl, B.; Katz, B. P.; Munson, R. S. Sialylation of Lipooligosaccharides Is Dispensable for the Virulence of *Haemophilus Ducreyi* in Humans. *Infect. Immun.* **2012**, *12*, 679–687.

(84) Lee, C. H.; Chang, C. C.; Liu, J. W.; Chen, R. F.; Yang, K. D. Sialic Acid Involved in Hypermucoviscosity Phenotype of *Klebsiella Pneumoniae* and Associated with Resistance to Neutrophil Phagocytosis. *Virulence* **2014**, *5*, 673–679.

(85) Carlin, A. F.; Uchiyama, S.; Chang, Y. C.; Lewis, A. L.; Nizet, V.; Varki, A. Molecular Mimicry of Host Sialylated Glycans Allows a Bacterial Pathogen to Engage Neutrophil Siglec-9 and Dampen the Innate

Immune Response. *Blood* **2009**, *113*, 3333–3336.

(86)  Carlin, A. F.; Lewis, A. L.; Varki, A.; Nizet, V. Group B Streptococcal Capsular Sialic Acids Interact with Siglecs (Immunoglobulin-like Lectins) on Human Leukocytes. *J. Bacteriol.* **2007**, *189*, 1231–1237.

(87)  Uchiyama, S.; Sun, J.; Fukahori, K.; Ando, N.; Wu, M.; Schwarz, F.; Siddiqui, S. S.; Varki, A.; Marth, J. D.; Nizet, V. Dual Actions of Group B *Streptococcus* Capsular Sialic Acid Provide Resistance to Platelet-Mediated Antimicrobial Killing. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 7465–7470.

(88)  Yamaguchi, M.; Hirose, Y.; Nakata, M.; Uchiyama, S.; Yamaguchi, Y.; Goto, K.; Sumitomo, T.; Lewis, A. L.; Kawabata, S.; Nizet, V. Evolutionary Inactivation of a Sialidase in Group B *Streptococcus*. *Sci. Rep.* **2016**, *6*, 28852.

(89)  Mouw, J. K.; Ou, G.; Weaver, V. M. Extracellular Matrix Assembly: A Multiscale Deconstruction. *Nat. Rev. Mol. Cell Biol.* **2014**, *15*, 771–785.

(90)  Secundino, I.; Lizcano, A.; Roupé, K. M.; Wang, X.; Cole, J. N.; Olson, J.; Ali, S. R.; Dahesh, S.; Amayreh, L. K.; Henningham, A.; Varki, A.; Nizet, V. Host and Pathogen Hyaluronan Signal through Human Siglec-9 to Suppress Neutrophil Activation. *J. Mol. Med.* **2016**, *94*, 219–233.

(91)  Crocker, P. R.; Paulson, J. C.; Varki, A. Siglecs and Their Roles in the Immune System. *Nat. Rev. Immunol.* **2007**, *7*, 255–266.

(92)  Wesener, D. A.; Wangkanont, K.; McBride, R.; Song, X.; Kraft, M. B.; Hodges, H. L.; Zarling, L. C.; Splain, R. A.; Smith, D. F.; Cummings, R. D.; Paulson, J. C.; Forest, K. T.; Kiessling, L. L. Recognition of Microbial Glycans by Human Intelectin-1. *Nat. Struct. Mol. Biol.* **2015**, *22*, 603–610.

(93)  Varki, A. Nothing in Glycobiology Makes Sense, except in the Light of Evolution. *Cell* **2006**, *126*, 841–845.

(94)  Varki, A. Biological Roles of Glycans. *Glycobiology* **2017**, *27*, 3–49.

(95)  Padler-Karavani, V.; Yu, H.; Cao, H.; Chokhawala, H.; Karp, F.; Varki, N.; Chen, X.; Varki, A. Diversity in Specificity, Abundance, and Composition of Anti-Neu5Gc Antibodies in Normal Humans: Potential Implications for Disease. *Glycobiology* **2008**, *18*, 818–830.

(96)  Ng, P. S. K.; Day, C. J.; Atack, J. M.; Hartley-Tassell, L. E.; Winter, L. E.; Marshanski, T.; Padler-Karavani, V.; Varki, A.; Barenkamp, S. J.; Apicella, M. A.; Jennings, M. P. Nontypeable Haemophilus Influenzae Has Evolved Preferential Use of N-Acetylneuraminic Acid as a Host Adaptation. *MBio* **2019**, *10*, 1–11.

(97)  Varki, A.; Gagneux, P. Multifarious Roles of Sialic Acids in Immunity. *Ann. N. Y. Acad. Sci.* **2012**, *1253*, 16–36.

(98)  Ausubel, F. M. *Current Protocols in Molecular Biology*; Greene Publishing Associates: New York, 1987.

(99)  Moran, A. P.; Holst, O.; Brennan, P. J.; Itzstein, M. von. *Microbial Glycobiology: Structures, Relevance and Applications*, 1st ed.; Academic Press/Elsevier: Amsterdam, 2009.

(100)  Zol-Hanlon, M. I.; Schumann, B. Open Questions in Chemical Glycobiology. *Commun. Chem.* **2020**, *3*, 102.

(101)  Wen, L.; Edmunds, G.; Gibbons, C.; Zhang, J.; Gadi, M. R.; Zhu, H.; Fang, J.; Liu, X.; Kong, Y.; Wang, P. G. Toward Automated Enzymatic Synthesis of Oligosaccharides. *Chem. Rev.* **2018**, *118*, 8151–8187.

(102)  Zhang, Z. J.; Wang, Y.-C.; Yang, X.; Hang, H. C. Chemical Reporters for Exploring Microbiology and Microbiota Mechanisms. *ChemBioChem* **2020**, *21*, 19–32.

(103)  Griffin, M. E.; Hsieh-Wilson, L. C. Glycan Engineering for Cell and Developmental Biology. *Cell Chem. Biol.* **2016**, *23*, 108–121.

(104)  Pons, J. M.; Dumont, A.; Sautejeau, G.; Fugier, E.; Baron, A.; Dukan, S.; Vauzeilles, B. Identification of Living *Legionella Pneumophila* Using Species-Specific Metabolic Lipopolysaccharide Labeling. *Angew. Chem - Int. Ed.* **2014**, *53*, 1275–1278.

(105)  Liu, F.; Aubry, A. J.; Schoenhofen, I. C.; Logan, S. M.; Tanner, M. E. The Engineering of Bacteria Bearing Azido-Pseudaminic Acid-Modified Flagella. *ChemBioChem* **2009**, *10*, 1317–1320.

(106)  Andolina, G.; Wei, R.; Liu, H.; Zhang, Q.; Yang, X.; Cao, H.; Chen, S.; Yan, A.; Li, X. D.; Li, X. Metabolic Labeling of Pseudaminic Acid-Containing Glycans on Bacterial Surfaces. *ACS Chem. Biol.* **2018**, *13*, 3030–3037.

(107)  Clark, E. L.; Emmadi, M.; Krupp, K. L.; Podilapu, A. R.; Helble, J. D.; Kulkarni, S. S.; Dube, D. H. Development of Rare Bacterial Monosaccharide Analogs for Metabolic Glycan Labeling in Pathogenic Bacteria. *ACS Chem. Biol.* **2016**, *11*, 3365–3373.

(108)  Sminia, T. J.; Zuilhof, H.; Wennekes, T. Getting a Grip on Glycans: A Current Overview of the Metabolic Oligosaccharide Engineering Toolbox. *Carbohydr. Res.* **2016**, *435*, 121–141.

(109)  Moons, S. J.; Adema, G. J.; Derks, M. T.; Boltje, T. J.; Büll, C. Sialic Acid Glycoengineering Using N-Acetylmannosamine and Sialic Acid Analogs. *Glycobiology* **2019**, *29*, 433–445.

(110)  Luijkx, Y. M. C. A.; Jongkees, S.; Strijbis, K.; Wennekes, T. Development of a 1,2-Difluorofucoside Activity-Based Probe for Profiling GH29 Fucosidases. *Org. Biomol. Chem.* **2021**, *19*, 2968–2977.

(111)  Gloster, T. M.; Vocadlo, D. J. Developing Inhibitors of Glycan Processing Enzymes as Tools for Enabling Glycobiology. *Nat. Chem. Biol.* **2012**, *8*, 683–694.

**2**

# Chapter 2
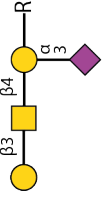
(112) Wu, L.; Armstrong, Z.; Schröder, S. P.; de Boer, C.; Artola, M.; Aerts, J. M.; Overkleeft, H. S.; Davies, G. J. An Overview of Activity-Based Probes for Glycosidases. *Curr. Opin. Chem. Biol.* **2019**, *53*, 25–36.

(113) Nischan, N.; Kohler, J. J. Advances in Cell Surface Glycoengineering Reveal Biological Function. *Glycobiology* **2016**, *26*, 789–796.

(114) Hudak, J. E.; Bertozzi, C. R. Glycotherapy: New Advances Inspire a Reemergence of Glycans in Medicine. *Chem. Biol.* **2014**, *21*, 16–37.

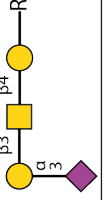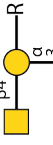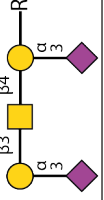(115) Mbua, N. E.; Li, X.; Flanagan-Steet, H. R.; Meng, L.; Aoki, K.; Moremen, K. W.; Wolfert, M. A.; Steet, R.; Boons, G. J. Selective Exo-Enzymatic Labeling of N-Glycans on the Surface of Living Cells by Recombinant ST6Gal I. *Angew. Chem. Int. Ed.* **2013**, *52*, 13012–13015.

(116) Sun, T.; Yu, S. H.; Zhao, P.; Meng, L.; Moremen, K. W.; Wells, L.; Steet, R.; Boons, G. J. One-Step Selective Exoenzymatic Labeling (SEEL) Strategy for the Biotinylation and Identification of Glycoproteins of Living Cells. *J. Am. Chem. Soc.* **2016**, *138*, 11575–11582.

(117) Capicciotti, C. J.; Zong, C.; Sheikh, M. O.; Sun, T.; Wells, L.; Boons, G. J. Cell-Surface Glyco-Engineering by Exogenous Enzymatic Transfer Using a Bifunctional CMP-Neu5Ac Derivative. *J. Am. Chem. Soc.* **2017**, *139*, 13342–13348.

(118) Lopez Aguilar, A.; Briard, J. G.; Yang, L.; Ovryn, B.; Macauley, M. S.; Wu, P. Tools for Studying Glycans: Recent Advances in Chemoenzymatic Glycan Labeling. *ACS Chem. Biol.* **2017**, *12*, 611–621.

(119) Tang, F.; Zhou, M.; Qin, K.; Shi, W.; Yashinov, A.; Yang, Y.; Yang, L.; Guan, D.; Zhao, L.; Tang, Y.; Chang, Y.; Zhao, L.; Yang, H.; Zhou, H.; Huang, R.; Huang, W. Selective N-Glycan Editing on Living Cell Surfaces to Probe Glycoconjugate Function. *Nat. Chem. Biol.* **2020**, *16*, 766–775.

(120) Schumann, B.; Malaker, S. Al.; Wisnovsky, S. P.; Debets, M. F.; Agbay, A. J.; Fernandez, D.; Wagner, L. J. S.; Lin, L.; Choi, J.; Fox, D. M.; Peh, J.; Gray, M. A.; Pedram, K.; Kohler, J. J.; Mrksich, M.; Bertozzi, C. R. Bump-and-Hole Engineering Identifies Specific Substrates of Glycosyltransferases in Living Cells. *Mol. Cell* **2020**, *78*, 824-834.e15.

(121) Lukose, V.; Whitworth, G.; Guan, Z.; Imperiali, B. Chemoenzymatic Assembly of Bacterial Glycoconjugates for Site-Specific Orthogonal Labeling. *J. Am. Chem. Soc.* **2015**, *137*, 12446–12449.

(122) Calabretta, P.; Hodges, H. L.; Kraft, M. B.; Marando, V.; Kiessling, L. L. Bacterial Cell Wall Modification with a Glycolipid Substrate. *J. Am. Chem. Soc.* **2019**, *141*, 9262–9272.

(123) Mally, M.; Fontana, C.; Leibundgut-Landmann, S.; Laacisse, L.; Fan, Y. Y.; Widmalm, G.; Aebi, M. Glycoengineering of Host Mimicking Type-2 LacNAc Polymers and Lewis X Antigens on Bacterial Cell Surfaces. *Mol. Microbiol.* **2013**, *87*, 112–131.

(124) Geva-Zatorsky, N.; Alvarez, D.; Hudak, J. E.; Reading, N. C.; Erturk-Hasdemir, D.; Dasgupta, S.; Von Andrian, U. H.; Kasper, D. L. *In Vivo* Imaging and Tracking of Host-Microbiota Interactions via Metabolic Labeling of Gut Anaerobic Bacteria. *Nat. Med.* **2015**, *21*, 1091–1100.

(125) Hudak, J. E.; Alvarez, D.; Skelly, A.; Von Andrian, U. H.; Kasper, D. L. Illuminating Vital Surface Molecules of Symbionts in Health and Disease. *Nat. Microbiol.* **2017**, *2*, 17099.

(126) Wang, T.-C.; Cochet, F.; Facchini, F. A.; Zaffaroni, L.; Serba, C.; Pascal, S.; Andraud, C.; Sala, A.; Di Lorenzo, F.; Maury, O.; Huser, T.; Peri, F. Synthesis of the New Cyanine-Labeled Bacterial Lipooligosaccharides for Intracellular Imaging and *in Vitro* Microscopy Studies. *Bioconjug. Chem.* **2019**, *30*, 1649–1657.

(127) Heesterbeek, D. A. C.; Muts, R. M.; van Hensbergen, V. P.; de Saint Aulaire, P.; Wennekes, T.; Bardoel, B. W.; van Sorge, N. M.; Rooijakkers, S. H. M. Outer Membrane Permeabilization by the Membrane Attack Complex Sensitizes Gram-Negative Bacteria to Antimicrobial Proteins in Serum and Phagocytes. *PLoS Pathog.* **2021**, *17*, e1009227.

(128) Stowell, S. R.; Arthur, C. M.; Mcbride, R.; Berger, O.; Razi, N.; Heimburg-Molinaro, J.; Rodrigues, L. C.; Gourdine, J. P.; Noll, A. J.; Von Gunten, S.; Smith, D. F.; Knirel, Y. A.; Paulson, J. C.; Cummings, R. D. Microbial Glycan Microarrays Define Key Features of Host-Microbial Interactions. *Nat. Chem. Biol.* **2014**, *10*, 470–476.

(129) Tytgat, H. L. P.; de Vos, W. M. Sugar Coating the Envelope: Glycoconjugates for Microbe–Host Crosstalk. *Trends Microbiol.* **2016**, *24*, 853–861.

(130) Semchenko, E.; Moutin, M.; Korolik, V.; Tiralongo, J.; Day, C. J. Lectin Array Analysis of Purified Lipooligosaccharide: A Method for the Determination of Molecular Mimicry. *J. Glycomics Lipidomics* **2011**, *1*, 105.

(131) Semchenko, E. A.; Day, C. J.; Moutin, M.; Wilson, J. C.; Tiralongo, J.; Korolik, V. Structural Heterogeneity of Terminal Glycans in *Campylobacter Jejuni* Lipooligosaccharides. *PLoS One* **2012**, *7*, e40920.

(132) Campanero-Rhodes, M. A.; Palma, A. S.; Menéndez, M.; Solís, D. Microarray Strategies for Exploring Bacterial Surface Glycans and Their Interactions with Glycan-Binding Proteins. *Front. Microbiol.* **2020**, *10*, 2909.

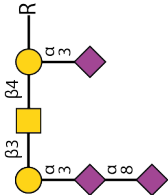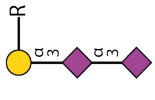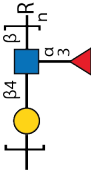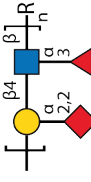(133) Wang, X.; Lang, S.; Tian, Y.; Zhang, J.; Yan, X.; Fang, Z.; Weng, J.; Lu, N.; Wu, X.; Li, T.; Cao, H.; Li,
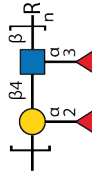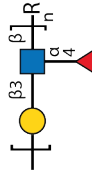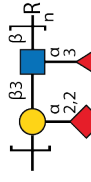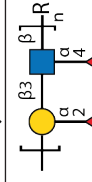
Z.; Huang, X. Glycoengineering of Natural Killer Cells with CD22 Ligands for Enhanced Anticancer Immunotherapy. *ACS Cent. Sci.* **2020**, *6*, 382–389.

(134) Bernardi, A.; Jiménez-Barbero, J.; Casnati, A.; De Castro, C.; Darbre, T.; Fieschi, F.; Finne, J.; Funken, H.; Jaeger, K. E.; Lahmann, M.; Lindhorst, T. K.; Marradi, M.; Messner, P.; Molinaro, A.; Murphy, P. V.; Nativi, C.; Oscarson, S.; Penadés, S.; Peri, F.; Pieters, R. J.; Renaudet, O.; Reymond, J. L.; Richichi, B.; Rojo, J.; Sansone, F.; Schäffer, C.; Bruce Turnbull, W.; Velasco-Torrijos, T.; Vidal, S.; Vincent, S.; Wennekes, T.; Zuilhof, H.; Imberty, A. Multivalent Glycoconjugates as Anti-Pathogenic Agents. *Chem. Soc. Rev.* **2013**, *42*, 4709–4727.

(135) Büll, C.; Heise, T.; van Hilten, N.; Pijnenborg, J. F. A.; Bloemendal, V. R. L. J.; Gerrits, L.; Kers-Rebel, E. D.; Ritschel, T.; den Brok, M. H.; Adema, G. J.; Boltje, T. J. Steering Siglec-Sialic Acid Interactions on Living Cells Using Bioorthogonal Chemistry. *Angew. Chem - Int. Ed.* **2017**, *56*, 3309–3313.

(136) Li, R.; Yu, H.; Chen, X. Recent Progress in Chemical Synthesis of Bacterial Surface Glycans. *Curr. Opin. Chem. Biol.* **2020**, *58*, 121–136.

(137) Li, T.; Liu, L.; Wei, N.; Yang, J.-Y.; Chapla, D. G.; Moremen, K. W.; Boons, G.-J. An Automated Platform for the Enzyme-Mediated Assembly of Complex Oligosaccharides. *Nat. Chem.* **2019**, *11*, 229–236.

(138) Li, T.; Wolfert, M. A.; Wei, N.; Huizinga, R.; Jacobs, B. C.; Boons, G. J. Chemoenzymatic Synthesis of *Campylobacter Jejuni* Lipo-Oligosaccharide Core Domains to Examine Guillain-Barré Syndrome Serum Antibody Specificities. *J. Am. Chem. Soc.* **2020**, *142*, 19611–19621.

(139) Reuel, N. F.; Mu, B.; Zhang, J.; Hinckley, A.; Strano, M. S. Nanoengineered Glycan Sensors Enabling Native Glycoprofiling for Medicinal Applications: Towards Profiling Glycoproteins without Labeling or Liberation Steps. *Chem. Soc. Rev.* **2012**, *41*, 5744–5779.

(140) Pilobello, K. T.; Mahal, L. K. Deciphering the Glycocode: The Complexity and Analytical Challenge of Glycomics. *Curr. Opin. Chem. Biol.* **2007**, *11*, 300–305.

(141) Lewis, A. L.; Desa, N.; Hansen, E. E.; Knirel, Y. A.; Gordon, J. I.; Gagneux, P.; Nizet, V.; Varki, A. Innovations in Host and Microbial Sialic Acid Biosynthesis Revealed by Phylogenomic Prediction of Nonulosonic Acid Structure. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 13552–13557.

(142) Lewis, A. L.; Nizet, V.; Varki, A. Discovery and Characterization of Sialic Acid O-Acetylation in Group B *Streptococcus. Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 11123–11128.

(143) Houliston, R. S.; Endtz, H. P.; Yuki, N.; Li, J.; Jarrell, H. C.; Koga, M.; Van Belkum, A.; Karwaski, M. F.; Wakarchuk, W. W.; Gilbert, M. Identification of a Sialate O-Acetyltransferase from *Campylobacter Jejuni*: Demonstration of Direct Transfer to the C-9 Position of Terminal α-2,8-Linked Sialic Acid. *J. Biol. Chem.* **2006**, *281*, 11480–11486.

(144) Lewis, A. L.; Cao, H.; Patel, S. K.; Diaz, S.; Ryan, W.; Carlin, A. F.; Thon, V.; Lewis, W. G.; Varki, A.; Chen, X.; Nizet, V. NeuA Sialic Acid O-Acetylesterase Activity Modulates O-Acetylation of Capsular Polysaccharide in Group B *Streptococcus. J. Biol. Chem.* **2007**, *282*, 27562–27571.

(145) Dzieciatkowska, M.; Brochu, D.; Van Belkum, A.; Heikema, A. P.; Yuki, N.; Houliston, R. S.; Richards, J. C.; Gilbert, M.; Li, J. Mass Spectrometric Analysis of Intact Lipooligosaccharide: Direct Evidence for O-Acetylated Sialic Acids and Discovery of O-Linked Glycine Expressed by *Campylobacter Jejuni. Biochemistry* **2007**, *46*, 14704–14714.

(146) Arthur, C. M.; Patel, S. R.; Mener, A.; Kamili, N. A.; Fasano, R. M.; Meyer, E.; Winkler, A. M.; Sola-Visner, M.; Josephson, C. D.; Stowell, S. R. Innate Immunity against Molecular Mimicry: Examining Galectin-Mediated Antimicrobial Activity. *BioEssays* **2015**, *37*, 1327–1337.

**2**
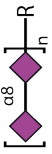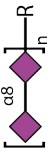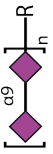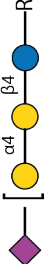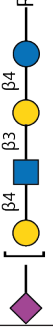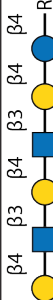
**Table S1:** this table provides an overview of the glycan structures that are involved in glycan mimicry by bacteria and are discussed in this review. The glycan structures of *Moraxella catarrhalis* and *Klebsiella pneumoniae* are not listed in the table since their glycan structures, specially the structures involved in glycan mimicry, are not fully elucidated yet. An overview of the glycan symbols used in the table can be found within. For a complete overview, the reader is kindly referred to the Symbol Nomenclature of Glycans[1].

Legend:
- ◆ N-Acetyl-Neuraminic acid
- ● Galactose
- ■ N-Acetyl-Galactosamine
- ● Glucose
- ■ N-Acetyl-Glucosamine
- ▲ Fucose
- ◆ Sialic acid
- ◆ Pseudaminic acid
- ⬡ Keto-deoxy-Mannooctanoic acid
- ◆ Glucuronic acid

| Bacteria | Part that mimics | Acronym glycan | Structure of glycan | Mimics eukaryotic glycans | Ref |
|---|---|---|---|---|---|
| *Campylobacter jejuni* | LOS | GM1 |  | Glycosphingolipids | 2,3 |
| *Campylobacter jejuni* | LOS | GM1b |  | Glycosphingolipids | 2,3 |
| *Campylobacter jejuni* | LOS | GM2 |  | Glycosphingolipids | 2,3 |
| *Campylobacter jejuni* | LOS | GM3 |  | Glycosphingolipids | 2,3 |
| *Campylobacter jejuni* | LOS | GD1a |  | Glycosphingolipids | 2,3 |

| | | | | | |
|---|---|---|---|---|---|
| *Campylobacter jejuni* | LOS | GT1a | | Glycosphingolipids | 2,3 |
| *Campylobacter jejuni* | LOS | GD3 | | Glycosphingolipids | 2,3 |
| *Helicobacter pylori* | O-antigen | Lewis x | | Blood group antigens | 2,4 |
| *Helicobacter pylori* | O-antigen | Sialyl Lewis x | | Blood group antigens | 2,4 |
| *Helicobacter pylori* | O-antigen | Lewis y | | Blood group antigens | 2,4 |
| *Helicobacter pylori* | O-antigen | Lewis a | | Blood group antigens | 2,4 |
| *Helicobacter pylori* | O-antigen | Sialyl Lewis a | | Blood group antigens | 2,4 |
| *Helicobacter pylori* | O-antigen | Lewis b | | Blood group antigens | 2,4 |

| Organism | Type | Antigen | Structure | Classification | Ref |
|---|---|---|---|---|---|
| *Helicobacter pylori* | O-antigen | H antigen 1 | | Blood group antigens | 2,4 |
| *Helicobacter pylori* | O-antigen | i-antigen | | Blood group antigens | 2,4 |
| *Escherichia coli* O86 | O-antigen | | | Blood group antigens | 5 |
| *Escherichia coli* O90 | O-antigen | | | Blood group antigens | 5 |
| *Escherichia coli* O127 | O-antigen | | | Blood group antigens | 5 |
| *Escherichia coli* O128 | O-antigen | | | Blood group antigens | 5 |
| *Escherichia coli* O24 | O-antigen | | | Blood group antigens | 5 |
| *Escherichia coli* O56 | O-antigen | | | Blood group antigens | 5 |
| *Escherichia coli* O104 | O-antigen | | | Blood group antigens | 5 |
| *Escherichia coli* O136 | O-antigen | | | Blood group antigens | 5 |
| *Escherichia coli* K92 | CPS | | | Poly sialic acid | 6 |

| Organism | Type | Structure | Source | Ref |
|---|---|---|---|---|
| *Escherichia coli* K1 | CPS | [ ◆◆ ]$_n$ α8 —R | Poly sialic acid | 6 |
| *Neisseria meningitidis* serogroup B | CPS | [ ◆◆ ]$_n$ α8 —R | Poly sialic acid | 7 |
| *Neisseria meningitidis* serogroup C | CPS | [ ◆◆ ]$_n$ α9 —R | Poly sialic acid | 7 |
| *Neisseria meningitidis* L1 | LOS | ◆ [ ●●● ] α4 β4 β4 —R | Human glycoproteins | 7 |
| *Neisseria meningitidis* L2-5 | LOS | ◆—●●■● β4 β3 β4 β4 —R | Human glycoproteins | 7 |
| *Neisseria meningitidis* L6 | LOS | ◆ [ ■●● ] β3 β4 β4 —R | Human glycoproteins | 7 |
| *Neisseria gonorrhoeae* | LOS | ◆■●●● β3 β4 β3 β4 β4 —R | Human glycoproteins | 7,8 |
| *Neisseria gonorrhoeae* | LOS | ●●■●● α3 β4 β3 β4 β4 —R | Human glycoproteins | 7,8 |
| *Neisseria gonorrhoeae* | LOS | ●■●● β4 β3 β4 β4 —R | Human glycoproteins | 7,8 |
| *Neisseria gonorrhoeae* PID2 | LOS | ●●●■●● β4 β3 β4 α3 β4 β2 —R | Human glycoproteins | 9 |
| *Nontypeable Haemophilus Influenzae* 375 | LOS | ■●●● β3 β4 β4 β2 —R | Human glycoproteins | 10,11 |
| *Nontypeable Haemophilus Influenzae* 375 | LOS | ◆●●● α2,3 β4 β4 β2 —R | Human glycoproteins | 10,11 |
| *Nontypeable Haemophilus Influenzae* 375 | LOS | ◆◆●● α2,8 α2,3 β4 β2 —R | Human glycoproteins | 10,11 |
| *Nontypeable Haemophilus Influenzae* 2019 | LOS | ◆●■ α2,3 β4 β2 —R | Human glycoproteins | 3,12 |

| Organism | | Structure | Target | Ref |
|---|---|---|---|---|
| *Nontypeable Haemophilus Influenzae* 2019 | LOS | | Human glycoproteins | 3,12 |
| *Nontypeable Haemophilus Influenzae* RdKW20 | LOS | | Human glycoproteins | 13 |
| *Haemophilus ducreyi* ITM2665 and 35000 | LOS | | Human glycoproteins | 14–16 |
| *Streptococcus agalactiae,* Group B streptococcus (GBS), Ia | CPS | | Human glycoproteins | 17,18 |
| *Streptococcus agalactiae,* Group B streptococcus (GBS), Ib | CPS | | Human glycoproteins | 17,18 |
| *Streptococcus agalactiae,* Group B streptococcus (GBS), II | CPS | | Human glycoproteins | 17,18 |
| *Streptococcus agalactiae,* Group B streptococcus (GBS), III | CPS | | Human glycoproteins | 17,18 |
| *Streptococcus agalactiae,* Group B streptococcus (GBS), V | CPS | | Human glycoproteins | 17,18 |
| *Streptococcus pyogenes,* Group A streptococcus | CPS | | Hyaluranon | 19 |

# References

(1)   Varki, A.; Cummings, R. D.; Aebi, M.; Packer, N. H.; Seeberger, P. H.; Esko, J. D.; Stanley, P.; Hart, G.; Darvill, A.; Kinoshita, T.; Prestegard, J. J.; Schnaar, R. L.; Freeze, H. H.; Marth, J. D.; Bertozzi, C. R.; Etzler, M. E.; Frank, M.; Vliegenthart, J. F. G.; Lütteke, T.; Perez, S.; Bolton, E.; Rudd, P.; Paulson, J.; Kanehisa, M.; Toukach, P.; Aoki-Kinoshita, K. F.; Dell, A.; Narimatsu, H.; York, W.; Taniguchi, N.; Kornfeld, S. Symbol Nomenclature for Graphical Representations of Glycans. *Glycobiology* **2015**, *25*, 1323–1324.

(2)   Moran, A. P.; Prendergast, M. M. Molecular Mimicry in *Campylobacter Jejuni* and *Helicobacter Pylori* Lipopolysaccharides: Contribution of Gastrointestinal Infections to Autoimmunity. In *Journal of Autoimmunity*; 2001; Vol. 16, pp 241–256.

(3)   Day, C. J.; Tran, E. N.; Semchenko, E. A.; Tram, G.; Hartley-Tassell, L. E.; Ng, P. S. K.; King, R. M.; Ulanovsky, R.; McAtamney, S.; Apicella, M. A.; Tiralongo, J.; Morona, R.; Korolik, V.; Jennings, M. P. Glycan:Glycan Interactions: High Affinity Biomolecular Interactions That Can Mediate Binding of Pathogenic Bacteria to Host Cells. *Proc. Natl. Acad. Sci.* **2015**, *112*, E7266–E7275.

(4)   Appelmelk, B. J.; Simoons-Smit, I.; Negrini, R.; Moran, A. P.; Aspinall, G. O.; Forte, J. G.; De Vries, T.; Quan, H.; Verboom, T.; Maaskant, J. J.; Ghiara, P.; Kuipers, E. J.; Bloemena, E.; Tadema, T. M.; Townsend, R. R.; Tyagarajan, K.; Crothers, J. M.; Monteiro, M. A.; Savio, A.; De Graaff, J. Potential Role of Molecular Mimicry between *Helicobacter Pylori* Lipopolysaccharide and Host Lewis Blood Group Antigens in Autoimmunity. *Infect. Immun.* **1996**, *64*, 2031–2040.

(5)   Stenutz, R.; Weintraub, A. The Structures of *Escherichia Coli* O-Polysaccharide Antigens. *FEMS Microbiol. Rev.* **2006**, *30*, 382–403.

(6)   McGowen, M. M.; Vionnet, J.; Vann, W. F. Elongation of Alternating A2,8/2,9 Polysialic Acid by the *Escherichia Coli* K92 Polysialyltransferase. *Glycobiology* **2001**, *11*, 613–620.

(7)   Mubaiwa, T. D.; Semchenko, E. A.; Hartley-Tassell, L. E.; Day, C. J.; Jennings, M. P.; Seib, K. L. The Sweet Side of the Pathogenic *Neisseria*: The Role of Glycan Interactions in Colonisation and Disease. *Pathogens and Disease*. 2017, p ftx063.

(8)   Mandrell, R. E.; Apicella, M. A. Lipo-Oligosaccharides (LOS) of Mucosal Pathogens: Molecular Mimicry and Host-Modification of LOS. *Immunobiology* **1993**, *187*, 382–402.

(9)   Tong, Y.; Arking, D.; Ye, S.; Reinhold, B.; Reinhold, V.; Stein, D. C. *Neisseria Gonorrhoeae* Strain PID2 Simultaneously Expresses Six Chemically Related Lipooligosaccharide Structures. *Glycobiology* **2002**, *12*, 523–533.

(10)  Jackson, M. D.; Wong, S. M.; Akerley, B. J. Underlying Glycans Determine the Ability of Sialylated Lipooligosaccharide to Protect Nontypeable *Haemophilus Influenzae* from Serum IgM and Complement. *Infect. Immun.* **2019**, *87*, e00456-19.

(11)  Kalograiaki, I.; Euba, B.; Fernández-Alonso, M. del C.; Proverbio, D.; St. Geme, J. W.; Aastrup, T.; Garmendia, J.; Cañada, F. J.; Solís, D. Differential Recognition of *Haemophilus Influenzae* Whole Bacterial Cells and Isolated Lipooligosaccharides by Galactose-Specific Lectins. *Sci. Rep.* **2018**, *8*, 16292.

(12)  Phillips, N. J.; Apicella, M. A.; Griffiss, J. M. L.; Gibson, B. W. Structural Characterization of the Cell Surface Lipooligosaccharides from a Nontypable Strain of *Haemophilus Influenzae*. *Biochemistry* **1992**, *31*, 4515–4526.

(13)  Apicella, M. A.; Coffin, J.; Ketterer, M.; Post, D. M. B.; Day, C. J.; Jen, F. E.; Jennings, M. P. Nontypeable *Haemophilus Influenzae* Lipooligosaccharide Expresses a Terminal Ketodeoxyoctanoate *in Vivo*, Which Can Be Used as a Target for Bactericidal Antibody. *Mol. Microbiol.* **2018**, *9*, e01401-18.

(14)  Schweda, E. K. H.; Sundström, A. C.; Eriksson, L. M.; Jonasson, J. A.; Lindberg, A. A. Structural Studies of the Cell Envelope Lipopolysaccharides from *Haemophilus Ducreyi* Strains ITM 2665 and ITM 4747. *J. Biol. Chem.* **1994**, *269*, 12040–12048.

(15)  Melaugh, W.; Phillips, N. J.; Campagnari, A. A.; Tullius, M. V.; Gibson, B. W. Structure of the Major Oligosaccharide from the Lipooligosaccharide of *Haemophilus Ducreyi* Strain 35000 and Evidence for Additional Glycoforms. *Biochemistry* **1994**, *33*, 13070–13078.

(16)  Melaugh, W.; Campagnari, A. A.; Gibson, B. W. The Lipooligosaccharides of *Haemophilus Ducreyi* Are Highly Sialylated. *J. Bacteriol.* **1996**, *178*, 564–570.

(17)  Carlin, A. F.; Lewis, A. L.; Varki, A.; Nizet, V. Group B Streptococcal Capsular Sialic Acids Interact with Siglecs (Immunoglobulin-like Lectins) on Human Leukocytes. *J. Bacteriol.* **2007**, *189*, 1231–1237.

(18)  Carlin, A. F.; Uchiyama, S.; Chang, Y. C.; Lewis, A. L.; Nizet, V.; Varki, A. Molecular Mimicry of Host Sialylated Glycans Allows a Bacterial Pathogen to Engage Neutro-

phil Siglec-9 and Dampen the Innate Immune Response. *Blood* **2009**, *113*, 3333–3336.

Secundino, I.; Lizcano, A.; Roupé, K. M.; Wang, X.; Cole, J. N.; Olson, J.; Ali, S. R.; Dahesh, S.; Amayreh, L. K.; Henningham, A.; Varki, A.; Nizet, V. Host and Pathogen Hyaluronan Signal through Human Siglec-9 to Suppress Neutrophil Activation. *J. Mol. Med.* **2016**, *94*, 219–233.

(19)

2

# Chapter 3

## Selective exoenzymatic labeling of lipooligosaccharides of *Neisseria gonorrhoeae* with sialic acid analogues

The interactions between bacteria and their host often rely on recognition processes that involve host or bacterial glycans. Glycoengineering techniques make it possible to modify and study the glycans on the host's eukaryotic cells, but only a few are available for the study of bacterial glycans. Here, we have adapted the chemical reporter strategy, selective exoenzymatic labeling (SEEL), to label the lipooligosaccharides of the bacterial pathogen *Neisseria gonorrhoeae*, using the recombinant glycosyltransferase ST6Gal1, and three synthetic CMP-sialic acid derivatives. We show that SEEL treatment does not affect cell viability and is capable of introducing an $\alpha2,6$-linked sialic acid with a reporter group on the lipooligosaccharides by Western blot, flow cytometry and fluorescent microscopy. This new bacterial glycoengineering technique allows for the precise modification, here with sialoside derivatives, and direct detection of specific surface glycans on live bacteria, which will aid in further unravelling the precise biological functions of bacterial glycans.

*Hanna de Jong, Maria J. Moure, Jet E. M. Hartman, Gerlof P. Bosman, Bart W. Bardoel, Geert-Jan Boons, Marc M. S. M. Wösten, Tom Wennekes*

## Chapter 3

## Introduction

Glycans play a crucial role in many biological processes[1]. They are particularly prevalent on the outside of the cell in the so-called glycocalyx. Efforts to manipulate and track glycans in the glycocalyx with the use of glycoengineering techniques are gaining momentum[2,3]. In glycoengineering, the glycans on a cell surface are modified by either inserting whole glycoconjugates or editing the existing glycan structures by introducing, removing or altering specific monosaccharide residues, which often entails the introduction of a chemical reporter. This approach allows studying specific glycans and the precise modification of their structure in the relevant biological context of a living cell. So far, many of the reported glycoengineering techniques have focused on mammalian cells and have contributed to a better understanding of the structure-function relationship of glycans within the context of the glycocalyx. However, in case of host-microbe interactions, comprehending cell surface glycosylation and their biological functions extends not only to mammalian cells, but also to microbial glycans, as illustrated by the many glycan interactions between a host and microbe that influence processes like bacterial pathogenesis, interactions with host receptors, and sialylation to evade immune detection by mimicry of host glycans[4,5].

Currently, few glycoengineering techniques have been reported that can modify bacterial glycans. The most widely applied approach is Metabolic Oligosaccharide Engineering (MOE)[6,7], which makes use of the cell's own metabolic pathways to incorporate monosaccharides with a chemical reporter group into sugar nucleotides that are eventually incorporated into bacterial glycans by native glycosyltransferases. MOE has been applied on microbes to engineer their cell wall,[8–12] to image them,[13,14] discover glycoproteins[15,16] or to develop new antibacterial strategies[17–19]. MOE is a powerful approach to selectively label unique bacterial monosaccharides like keto-deoxyoctanoate (KDO) in LPS[20] and *N*-acetyl muramic acid (NAM) in peptidoglycan[21,22], among other residues[6,19,23]. Several reports have described MOE with sialic acids, for instance neuraminic acid derivatives for *Haemophilus ducreyi*[24] or nontypeable *Haemophilus influenzae*[25], or legionaminic[26] and pseudaminic[12] acid for the flagella of *Campylobacter jejuni*. However, MOE is not generally applicable to bacteria if the biosynthetic machinery for the metabolic processing of the monosaccharide is absent. Other examples of glycoengineering techniques for bacteria are the chemoenzymatic synthesis of a heptasaccharide of *C. jejuni*[27] and a modification of a mycobacterial cell wall with a glycolipid derivative[28,29]. These pioneering examples of bacterial

glycoengineering show the promise of this approach, but current techniques have limited control over the type of glycoconjugate that is modified or are toxic to bacteria at higher concentrations of synthetic monosaccharide.

Our goal was to investigate whether a glycoengineering technique for mammalian cells, selective exoenzymatic labeling (SEEL),[30,31] could be adapted to bacteria and thus expand the glycoengineering toolbox for bacterial glycans to be able to manipulate and study specific glycans on the surface of live bacteria. SEEL uses an exogenous applied recombinant glycosyltransferase to selectively label the glycocalyx of a cell with tailor-made sugar nucleotide analogs. In earlier work, we reported a sialyltransferase ST6Gal1 that selectively modifies a terminal *N*-acetyllactosamine (LacNAc) on various human cell lines with α2,6-sialosides. In a two-step or one-step SEEL approach, respectively, the sialic acid either contained a bio-orthogonal azide that could be clicked to a biotin reporter group, or was already covalently coupled to a biotin[30,31]. SEEL is complementary to MOE as it has two unique features. The first one being the ability to precisely modify a particular glycan acceptor site with a specific linkage type because an enzyme of choice can be employed. Second, the monosaccharide derivative does not have to go through the multistep metabolic route of the cell, which allows for the one-step introduction of more diverse labels, such as a biotin or fluorophore. Considering these strengths of SEEL, we wanted to evaluate if SEEL can be applied to bacteria and thus function as a new strategy to modify their cell surface glycans.

We selected *Neisseria gonorrhoeae* as our target, which is a Gram-negative bacterial pathogen causing the sexually transmitted disease gonorrhea[32]. Its lipooligosaccharides (LOS) contain a terminal *N*-acetyllactosamine, which can be α2,3-sialylated by its native sialyltransferases that scavenge CMP-sialic acid from the host[32,33]. The resulting sialylated terminal glycans on the bacterium resemble mammalian *N*-glycans and glycosphingolipids, and are thus a form of glycan mimicry that conceals the bacteria from immune detection[5]. The use of terminal sialic acids to display glycans that mimic the host extends to other pathogenic bacteria and raises questions about the role of this monosaccharide in the interaction with the immune system, making sialic acid on bacteria a prime target for glycoengineering[5]. Since the established SEEL approach that we developed for mammalian cells uses the recombinant sialyltransferase, ST6Gal1, to modify terminal *N*-acetyllactosamines of *N*-glycans with sialic acid derivatives, we hypothesized that *N. gonorrhoeae* would be an ideal candidate for

adapting this glycoengineering technique to bacteria (Figure 1).

Here, we report for the first time the successful use of selective exoenzymatic labeling on bacteria. We show that SEEL labels the LOS of *N. gonorrhoeae* with a preference for LacNAc as the terminal glycan unit and that the introduced α2,6-linked sialic acid derivatives can be visualized and quantified on the bacteria.
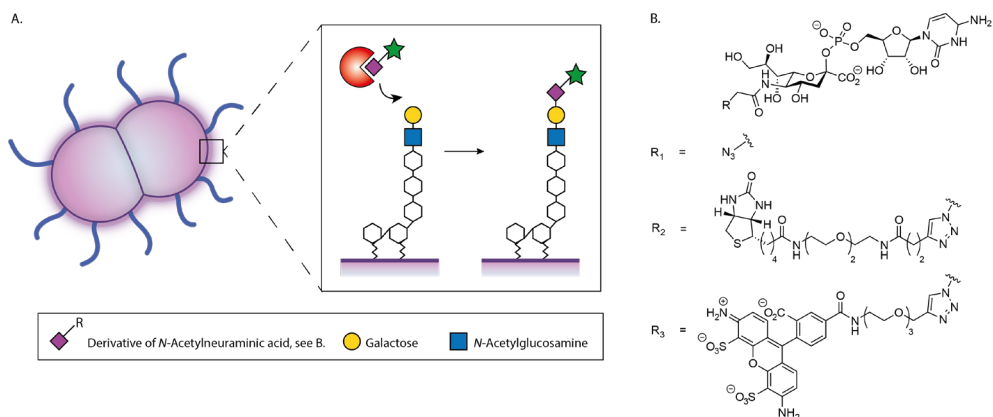


Figure 1. (A) Schematic overview of selective exoenzymatic labeling of the lipooligosaccharides on *N. gonorrhoeae* with a sialic acid that contains a reporter group. (B) Structures of azido, biotinylated and fluorescent CMP-Sia derivatives.

## Results and discussion

We set out to test the glycoengineering technique SEEL on the bacteria *Neisseria gonorrhoeae*. We started by testing one-step SEEL, using biotinylated cytidine monophosphate sialic acid (CMP-Sia-biotin), on two different strains of *N. gonorrhoeae*: one wildtype and one mutant lacking its own sialyltransferase (ST mutant)[34,35]. For both strains, we observed labeling of the LOS on Western blot and on polyacrylamide gel after silver staining through upward shifted bands of the LOS (Figure 2A). To check for any labeling of other glycoconjugates besides LOS, (glyco)protein samples of SEEL-treated bacteria were analyzed and labeling of bacterial glycoproteins was not observed (SI Figure 1A). However, the analysis of glycoproteins identified a band that corresponds to the labeled recombinant sialyltransferase ST6Gal1 (47 kDa) that can label its *N*-glycans during SEEL, as previously observed with mass spectrometry analysis by our group (unpublished data; manuscript in preparation) (SI Figure 1B). From these initial experiments, we could conclude that SEEL is indeed able to selectively label the LOS of *N. gonorrhoeae*.

## Selective exoenzymatic labeling of LOS of *N. gonorrhoeae*

Next, we wanted to evaluate if SEEL selectively labels *N*-acetyllactosamine on the terminal position of the LOS. To this end, we labeled three different isogenic *N. gonorrhoeae* strains with different terminal glycan structures, strain B contains LOS with a terminal *N*-acetyllactosamine, strain A lacks a terminal galactose thus exposing a terminal *N*-acetylglucosamine and strain C has an additional *N*-acetylgalactosamine (Figure 2B)[36]. It has been reported that *N. gonorrhoeae* can scavenge CMP-Neu5Ac from the environment[37,38] and therefore we wanted to explore if the three isogenic strains could also transfer tagged CMP-sialic acid derivatives. For both SEEL-treated bacteria and bacteria incubated with only the CMP-Sia-biotin, thus without any recombinant sialyltransferase, labeling was observed (Figure 2C). This indicates that the native sialyltransferases of *N. gonorrhoeae* can accept CMP-Sia-biotin and transfer it to the LOS. To exclude the contribution of these native enzymes in the experiments we heated the bacteria to denature these proteins and then performed SEEL (Figure 2C bottom panel, see SI Figure 2 for heat inactivated and SEEL treated WT and ST mutant). The signal for the heat-treated bacteria was strongly diminished which indicated that the native sialyltransferases contributed significantly to the previously observed labeling. Since the intensity of labeling in these experiments was diminished, it was possible to compare the labeling patterns for the different strains. As expected, weak labeling of strain A was observed and strong labeling of B, but also strong labeling of the LOS of strain C. We reasoned that strain C is labeled because the enzyme that installs the terminal *N*-acetylgalactosamine is under phase-variable expression[39,40] and it might thus not be present under the growth conditions, which exposes a terminal *N*-acetyllactosamine in strain C and makes it closely resemble strain B. Taken together, we conclude that as expected SEEL on *N. gonorrhoeae* with recombinant ST6Gal1 preferentially labels the LOS terminal *N*-acetyllactosamines.
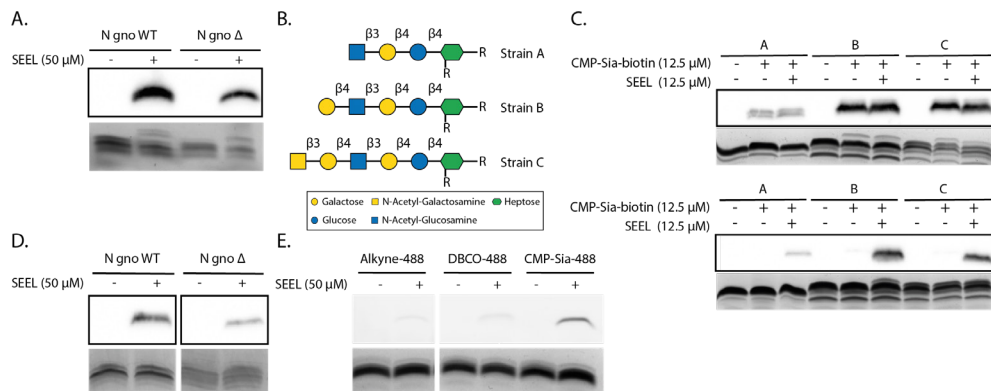
Figure 2. Western blot and silver stain analysis of *N. gonorrhoeae* LOS labeled with SEEL. (A) In a one-step reaction, the CMP-Sia-biotin is incorporated in the LOS of a wildtype and sialyltransferase mutant *N. gonorrhoeae* (N gno WT and N gno Δ, respectively), as can be seen by the signal on Western blot (upper panel) and the small increase in molecular weight visualized by silver stain (bottom panel). (B) Three different strains of LOS on *N. gonorrhoeae* to test labeling specificity. (C) The bacteria with different terminal glycans were treated with only CMP-Sia-biotin or SEEL and the LOS was analyzed with Western blot and silver stain. In the lower panel, bacteria were heated and then treated with only CMP-Sia-biotin or SEEL and the LOS was analyzed with Western blot and silver stain. (D) In a two-step reaction, the CMP-Sia-azide is incorporated in the LOS of a wildtype and sialyltransferase mutant *N. gonorrhoeae* and then clicked with an alkyne-biotin. The LOS was analyzed with Western blot and silver stain. (E) In either a two-step reaction, for alkyne-488 and DBCO-488, or a one-step reaction, with CMP-Sia-488, the mutant *N. gonorrhoeae* were labeled. The LOS was analyzed with in-gel fluorescence (upper panel) or silver stain (bottom panel). The results of this gel are combined in this figure and the raw data images are available in the SI.

After studying the targeted acceptor glycan of SEEL in *N. gonorrhoeae*, we focused on quantifying the number of bacteria that are being labeled. To achieve this, a fluorescent reporter group was introduced on the bacteria for analysis by flow cytometry. First, we tested the efficiency of two-step SEEL to introduce a reporter group via a click reaction of the sialosides with an azide to a terminal alkyne connected to a biotin reporter. The copper-catalyzed azide-alkyne cycloaddition (CuAAC) resulted in a clear signal (Figure 2D). Next, we evaluated this again, but now for both a CuAAC and strain-promoted azide-alkyne cycloaddition (SPAAC) with a terminal alkyne or DBCO group containing an Alexa Fluor 488 fluorescent dye (AF488). Both these experiments produced only a small amount of fluorescently labeled LOS on gel (Figure 2E). To compare the two-step SEEL with one-step SEEL, a CMP-sialic acid derivative with an AF488 dye, CMP-Sia-AF488, was synthesized by conjugating the fluorescent dye to CMP-NeuAz via a CuAAC reaction (see SI for full details). Compared to the click reaction in the two-step approach, this fluorescently labeled sugar nucleotide would avoid background labeling often observed with reacting a fluorescent dye onto cell surfaces or the potential cytotoxicity of CuAAC. When comparing two-step and one-step SEEL we saw a significantly stronger signal for one-step SEEL, especially for the in-gel fluorescence (Figure 2E). In previous studies, we made a similar

observation while comparing the amount of labeled mammalian glycoproteins with one-step and two-step SEEL[31]. We speculate that the two-step SEEL is less efficient because the click reaction might be sterically hindered on the surface of the bacteria. Since one-step SEEL would give a more accurate number of labeled bacteria, we continued using one-step SEEL with our newly synthesized CMP-Sia-AF488 for flow cytometry experiments. These measurements on the SEEL treated *N. gonorrhoeae* ST mutant demonstrated that there is an increase in fluorescence compared to the controls (Figure 3B) in which either only CMP-Sia-AF488 was added or SEEL treated bacteria with natural CMP-Neu5Ac. Although the flow cytometry data showed that most mutant bacteria were being labeled with SEEL (Figure 3B), it appears as a broad distribution and the level of incorporation was not as high when we compared it with treatment of WT *N. gonorrhoeae* with CMP-Sia-AF488. The native sialyltransferases of *N. gonorrhoeae* have been reported by Gulati et al.[41,42] to accept azido nucleotide sugar analogs and also microbial sialic acids. In line with this previous work, we observed that the native sialyltransferases of *N. gonorrhoeae* also accepted larger modifications such as the fluorescently labeled nucleotide sugar, which is another proof of broad substrate tolerance of this sialyltransferase. In contrast to the bacterial sialyltransferases, SEEL uniquely allows for the introduction of non-natively linked sialosides through the exogeneous enzyme, namely α2,6- instead of α2,3-Neu5Ac on the LOS of *Neisseria*[43,44]. To further optimize the SEEL protocol, we varied several parameters that showed the labeling could be increased with higher concentrations of label mix, but longer incubations times showed similar labeling as 2 hour incubation, and less enzyme or nucleotide sugar even showed a decrease in labeling (Figure 3C-F). We were also interested to gain insight into the amount of sialic acid incorporated on a bacterium's surface by SEEL. To determine the median number of fluorescently labeled sialosides per bacteria by SEEL, we used quantum beads with a known number of fluorophores to make a calibration curve (SI Figure 3)[45]. This revealed that SEEL treated bacteria on average have 19000 modifications with fluorescently modified sialic acid on their LOS (Figure 3G). In contrast, the labeling by the native bacterial enzymes of the WT was 19 times more efficient than the applied SEEL treatment. Collectively, the flow cytometry data shows that most of the SEEL treated bacteria are labeled but labeling of the LOS does not occur as efficiently as by the native sialyltransferases. This comparably lower sialylation might be caused by limited access of the recombinant enzyme to acceptor sites or a comparatively elevated local concentration of the native enzymes near the acceptor sites if they are surface-bound, yet SEEL has the unique property of being able to install a

sialoside with a linkage of choice, here a non-native α2,6-linkage. This type of glycosidic linkage cannot be obtained through the bacterial biosynthetic machinery during MOE or labeling through native sialyltransferases.
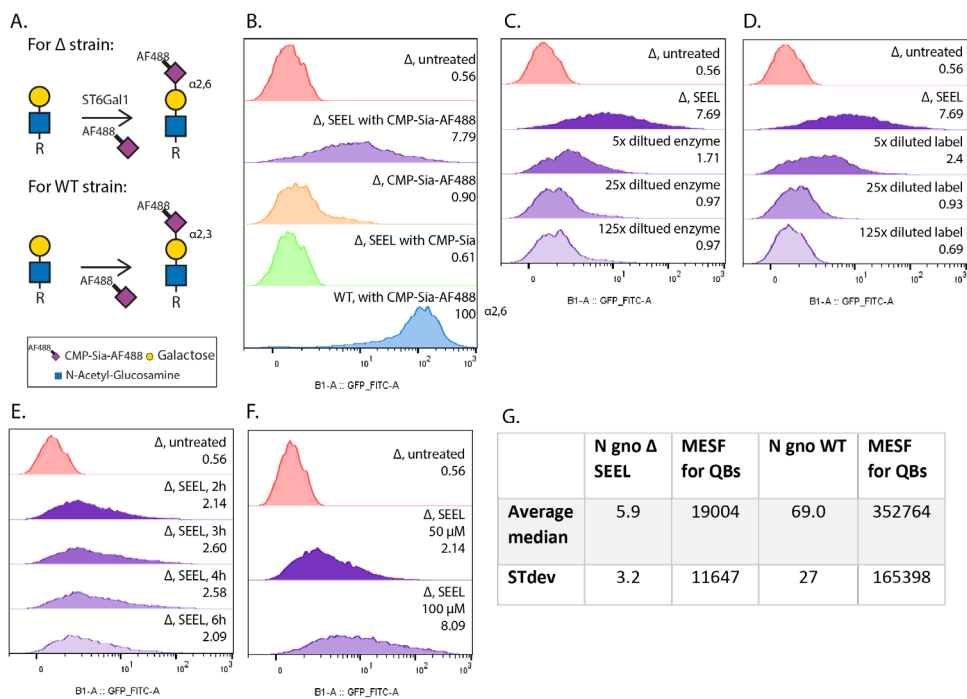


Figure 3. Flow cytometry data to quantify labeling and to test parameters of SEEL treated *N. gonorrhoeae*. (A) Schematic overview of the LOS labeling for SEEL treated mutant bacteria and WT treated with nucleotide sugar only. (B) Different conditions to confirm that the signal from the SEEL treated label originates from the SEEL treatment after 2-hour incubation. (C) Dilution series of the amount of enzyme in the SEEL label mix, 2-h incubation; decreasing amounts of ST6Gal1 (µg): 1.05; 0.21; 0.042; 0.0084. (D) Dilution series of the amount of CMP-Sia-AF488 in the SEEL label mix, 2-h incubation; decreasing concentration CMP-Sia-AF488 (µM): 50; 10; 2; 0.4. (E) SEEL labeling of *N. gonorrhoeae* after different incubation times. (F) SEEL labeling with increased concentration of SEEL label mix. (G). Quantification of the amount of fluorescence on SEEL treated bacteria or WT bacteria with CMP-Sia-AF488 from three independent measurements. The median fluorescent intensity was determined for mutant bacteria labeled by SEEL or wildtype bacteria labeled by native sialyltransferases. This number was converted to the number of modifications on the cell surface through the Molecules of Equivalent Soluble Fluorochrome (MESF) of the Quantum Beads (QBs).

Finally, we wanted to visualize the SEEL introduced sialic acid analogue. Through SEEL it is possible to label intact bacteria and since the enzyme ST6Gal1 is exogeneous, labeling takes place extracellularly[46]. The aim was to visualize the fluorescent reporter group on the bacteria that was introduced by SEEL. Fluorescence microscopy of the *N. gonorrhoeae* ST mutant that was SEEL treated with CMP-Sia-AF488 showed a bright green signal in a circular

shape around the cytoplasm in which the chromosomal DNA was stained with DAPI, and thus indicating successful labeling with CMP-Sia-AF488 (Figure 4). In agreement with the flow cytometry data, fluorescence microscopy showed a similar ratio of fluorescently labeled versus unlabeled *N. gonorrhoeae* by SEEL (SI Figure 4 and 5).

A. 488 channel          B. Merge of 488 and 405 channels   C. 488 channel, superresolution
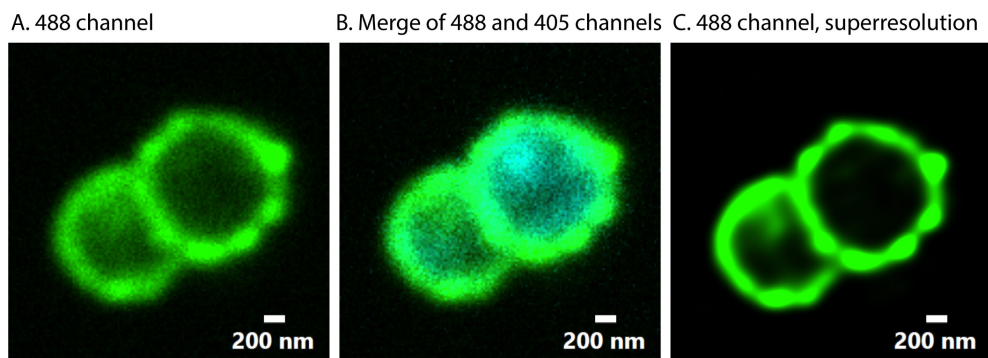


Figure 4. Fluorescence microscopy images show that *N. gonorrhoeae* are fluorescently labeled on the outside of the cell. (A) 488 channel. (B) Merge of 488 and 405 (DAPI) channels. (C) 488 channel with superresolution.

These combined results demonstrated that it is possible to adapt SEEL to bacteria. The power of SEEL, as a glycoengineering technique, is that it is targeted because it incorporates one type of monosaccharide on a specific acceptor on the cell surface, in this case Neu5Ac on the terminal *N*-acetyllactosamine of LOS. Additionally, the use of an exogenous enzyme ensures that a certain linkage type between the glycans is made because of the inherent specificity of the chosen glycosyltransferase, which can be a glycosidic linkage of choice that is non-native for the bacteria, and that the modification is presumably made extracellularly. Furthermore, as SEEL can introduce an extracellular modification in a single step, it has the potential to introduce specific cell surface modifications with large biolmolecules[47]. The single step introduction of a fluorescently labeled sialoside, as shown here, is for instance not possible via MOE[48]. Currently, MOE is the most widely applied approach to glycoengineer bacterial glycans. Although this technique has the power to hijack the metabolic process of the cell to incorporate unnatural glycans, it has disadvantages due to this requirement of metabolic processing of the externally added monosaccharide derivatives[49]. A case in point is that MOE cannot be applied to *N. gonorrhoeae* to engineer sialic acid, because the monosaccharide derivative cannot be converted to the nucleotide sugar since the bacteria lacks the required CMP synthetase in the metabolic pathway[32]. Also, MOE can be toxic to the cell as often high concentrations of these derivatives are used to ensure sufficient cellular uptake

and to achieve successful competition in the metabolic pathway with the native monosaccharide. In contrast, we evaluated SEEL and it did not show cell toxicity for *N. gonorrhoeae* (SI Figure 6). A possible limitation of SEEL includes the number of modifications made. The high glycosylation activity we observed for the native bacterial sialyltransferases with our three synthetic CMP-sialic acid derivatives could be a promising method to achieve higher modification levels. On the other hand, SEEL allows for more precise control over the introduction of a tailor-made sialic acid derivative via specific glycosidic linkage on a particular acceptor glycan. The observed activity of the native bacterial enzymes could also mean that knock outs of these are required to exclude their contribution for certain applications, like the introduction of sialic acids on specific acceptor glycans to study specific biological processes.

In conclusion, we show here for the first time that SEEL can also be applied as a glycoengineering technique for the modification of cell surface glycans on bacteria, in this case synthetic sialic acid derivatives with a reporter group on the LOS of *N. gonorrhoeae*. Bacterial SEEL represents a promising new complementary technique to engineer microbial glycans, next to MOE. In the future, the scope of SEEL to label other bacteria with a terminal LacNAc like *N. meningitidis* or *H. ducreyi*, and the use of other sugar nucleotides and corresponding glycosyltransferases can be explored, as well as the use of this technique to study the interactions between sialylated LOS of *N. gonorrhoeae* and its host.

## Acknowledgements

# References

(1) Varki, A. Biological Roles of Glycans. *Glycobiology* **2017**, *27*, 3–49.

(2) Griffin, M. E.; Hsieh-Wilson, L. C. Glycan Engineering for Cell and Developmental Biology. *Cell Chem. Biol.* **2016**, *23*, 108–121.

(3) Critcher, M.; O'Leary, T.; Huang, M. L. Glycoengineering: Scratching the Surface. *Biochem. J.* **2021**, *478*, 703–719.

(4) Poole, J.; Day, C. J.; von Itzstein, M.; Paton, J. C.; Jennings, M. P. Glycointeractions in Bacterial Pathogenesis. *Nat. Rev. Microbiol.* **2018**, *16*, 440–452.

(5) de Jong, H.; Wösten, M. M. S. M.; Wennekes, T. Sweet Impersonators: Molecular Mimicry of Host Glycans by Bacteria. *Glycobiology* **2021**.

(6) Luong, P.; Dube, D. H. Dismantling the Bacterial Glycocalyx: Chemical Tools to Probe, Perturb, and Image Bacterial Glycans. *Bioorganic Med. Chem.* **2021**, *42*, 116268.

(7) Banahene, N.; Kavunja, H. W.; Swarts, B. M. Chemical Reporters for Bacterial Glycans: Development and Applications. *Chem. Rev.* **2021**, acs.chemrev.1c00729.

(8) Sadamoto, R.; Niikura, K.; Sears, P. S.; Liu, H.; Wong, C.-H.; Suksomcheep, A.; Tomita, F.; Monde, K.; Nishimura, S.-I. Cell-Wall Engineering of Living Bacteria. *J. Am. Chem. Soc.* **2002**, *124*, 9018–9019.

(9) Sadamoto, R.; Niikura, K.; Ueda, T.; Monde, K.; Fukuhara, N.; Nishimura, S.-I. Control of Bacteria Adhesion by Cell-Wall Engineering. *J. Am. Chem. Soc.* **2004**, *126*, 3755–3761.

(10) Dumont, A.; Malleron, A.; Awwad, M.; Dukan, S.; Vauzeilles, B. Click-Mediated Labeling of Bacterial Membranes through Metabolic Modification of the Lipopolysaccharide Inner Core. *Angew. Chem. Int. Ed.* **2012**, *51*, 3143–3146.

(11) Swarts, B. M.; Holsclaw, C. M.; Jewett, J. C.; Alber, M.; Fox, D. M.; Siegrist, M. S.; Leary, J. A.; Kalscheuer, R.; Bertozzi, C. R. Probing the Mycobacterial Trehalome with Bioorthogonal Chemistry. *J. Am. Chem. Soc.* **2012**, *134*, 16123–16126.

(12) Liu, F.; Aubry, A. J.; Schoenhofen, I. C.; Logan, S. M.; Tanner, M. E. The Engineering of Bacteria Bearing Azido-Pseudaminic Acid-Modified Flagella. *ChemBioChem* **2009**, *10*, 1317–1320.

(13) Geva-Zatorsky, N.; Alvarez, D.; Hudak, J. E.; Reading, N. C.; Erturk-Hasdemir, D.; Dasgupta, S.; Von Andrian, U. H.; Kasper, D. L. *In Vivo* Imaging and Tracking of Host-Microbiota Interactions via Metabolic Labeling of Gut Anaerobic Bacteria. *Nat. Med.* **2015**, *21*, 1091–1100.

(14) Hudak, J. E.; Alvarez, D.; Skelly, A.; Von Andrian, U. H.; Kasper, D. L. Illuminating Vital Surface Molecules of Symbionts in Health and Disease. *Nat. Microbiol.* **2017**, *2*, 17099.

(15) Champasa, K.; Longwell, S. A.; Eldridge, A. M.; Stemmler, E. A.; Dube, D. H. Targeted Identification of Glycosylated Proteins in the Gastric Pathogen *Helicobacter Pylori (Hp)*. *Mol. Cell. Proteomics* **2013**, *12*, 2568–2586.

(16) Besanceney-Webler, C.; Jiang, H.; Wang, W.; Baughn, A. D.; Wu, P. Metabolic Labeling of Fucosylated Glycoproteins in *Bacteroidales* Species. *Bioorganic Med. Chem. Lett.* **2011**, *21*, 4989–4992.

(17) Kaewsapsak, P.; Esonu, O.; Dube, D. H. Recruiting the Host's Immune System to Target *Helicobacter Pylori*'s Surface Glycans. *ChemBioChem* **2013**, *14*, 721–726.

(18) Memmel, E.; Homann, A.; Oelschlaeger, T. A.; Seibel, J. Metabolic Glycoengineering of *Staphylococcus Aureus* Reduces Its Adherence to Human T24 Bladder Carcinoma Cells. *Chem. Commun.* **2013**, *49*, 7301–7303.

(19) Tra, V. N.; Dube, D. H. Glycans in Pathogenic Bacteria-Potential for Targeted Covalent Therapeutics and Imaging Agents. *Chem. Commun.* **2014**, *50*, 4659–4673.

(20) Dumont, A.; Malleron, A.; Awwad, M.; Dukan, S.; Vauzeilles, B. Click-Mediated Labeling of Bacterial Membranes through Metabolic Modification of the Lipopolysaccharide Inner Core. *Angew. Chem - Int. Ed.* **2012**, *51*, 3143–3146.

(21) Liang, H.; DeMeester, K. E.; Hou, C. W.; Parent, M. A.; Caplan, J. L.; Grimes, C. L. Metabolic Labelling of the Carbohydrate Core in Bacterial Peptidoglycan and Its Applications. *Nat. Commun.* **2017**, *8*, 1–11.

(22) Demeester, K. E.; Liang, H.; Jensen, M. R.; Jones, Z. S.; D'ambrosio, E. A.; Scinto, S. L.; Zhou, J.; Grimes, C. L. Synthesis of Functionalized N-Acetyl Muramic Acids to Probe Bacterial Cell Wall Recycling and Biosynthesis. *J. Am. Chem. Soc.* **2018**, *140*, 9458–9465.

(23) Clark, E. L.; Emmadi, M.; Krupp, K. L.; Podilapu, A. R.; Helble, J. D.; Kulkarni, S. S.; Dube, D. H. Development of Rare Bacterial Monosaccharide Analogs for Metabolic Glycan Labeling in Pathogenic Bacteria. *ACS Chem. Biol.* **2016**, *11*, 3365–3373.

(24) Goon, S.; Schilling, B.; Tullius, M. V.; Gibson, B. W.; Bertozzi, C. R. Metabolic Incorporation of Unnatural Sialic Acids into *Haemophilus Ducreyi* Lipooligosaccharides. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 3089–3094.

(25) Heise, T.; Langereis, J. D.; Rossing, E.; de Jonge, M. I.; Adema, G. J.; Büll, C.; Boltje, T. J. Selective Inhibition of Sialic Acid-Based Molecular Mimicry in *Haemophilus Influenzae* Abrogates Serum Resistance. *Cell Chem. Biol.* **2018**, *25*, 1279-1285.e8.

**3**

(26)     Meng, X.; Boons, G. J.; Wösten, M. M. S. M.; Wennekes, T. Metabolic Labeling of Legionaminic Acid in Flagellin Glycosylation of *Campylobacter Jejuni* Identifies Maf4 as a Putative Legionaminyl Transferase. *Angew. Chem - Int. Ed.* **2021**, *60*, 24811–24816.

(27)     Lukose, V.; Whitworth, G.; Guan, Z.; Imperiali, B. Chemoenzymatic Assembly of Bacterial Glycoconjugates for Site-Specific Orthogonal Labeling. *J. Am. Chem. Soc.* **2015**, *137*, 12446–12449.

(28)     Calabretta, P.; Hodges, H. L.; Kraft, M. B.; Marando, V.; Kiessling, L. L. Bacterial Cell Wall Modification with a Glycolipid Substrate. *J. Am. Chem. Soc.* **2019**, *141*, 9262–9272.

(29)     Marando, V. M.; Kim, D. E.; Calabretta, P. J.; Kraft, M. B.; Bryson, B. D.; Kiessling, L. L. Biosynthetic Glycan Labeling. *J. Am. Chem. Soc.* **2021**, *143*, 16337–16342.

(30)     Mbua, N. E.; Li, X.; Flanagan-Steet, H. R.; Meng, L.; Aoki, K.; Moremen, K. W.; Wolfert, M. A.; Steet, R.; Boons, G. J. Selective Exo-Enzymatic Labeling of N-Glycans on the Surface of Living Cells by Recombinant ST6Gal I. *Angew. Chem. Int. Ed.* **2013**, *52*, 13012–13015.

(31)     Sun, T.; Yu, S. H.; Zhao, P.; Meng, L.; Moremen, K. W.; Wells, L.; Steet, R.; Boons, G. J. One-Step Selective Exoenzymatic Labeling (SEEL) Strategy for the Biotinylation and Identification of Glycoproteins of Living Cells. *J. Am. Chem. Soc.* **2016**, *138*, 11575–11582.

(32)     Mubaiwa, T. D.; Semchenko, E. A.; Hartley-Tassell, L. E.; Day, C. J.; Jennings, M. P.; Seib, K. L. The Sweet Side of the Pathogenic *Neisseria*: The Role of Glycan Interactions in Colonisation and Disease. *Pathogens and Disease*. 2017, p ftx063.

(33)     Parsons, N. J.; Ashton, P. R.; Constantinidou, C.; Cole, J. A.; Smith, H. Identification by Mass Spectrometry of CMP-NANA in Diffusible Material Released from High Mr Blood Cell Fractions That Confers Serum Resistance on Gonococci. *Microb. Pathog.* **1993**, *14*, 329–335.

(34)     Bramley, J.; de Hormaeche, R. D.; Constantinidou, C.; Nassif, X.; Parsons, N.; Jones, P.; Smith, H.; Cole, J. A Serum-Sensitive, Sialyltransferase-Deficient Mutant of *Neisseria Gonorrhoeae* Defective in Conversion to Serum Resistance by CMP-NANA or Blood Cell Extracts. *Microb. Pathog.* **1995**, *18*, 187–195.

(35)     Gill, M. J.; Mcquillen, D. P.; Van Putten, J. P. M.; Wetzler, L. M.; Bramley, J.; Crooke, H.; Parsons, N. J.; Cole, J. A.; Smith, H. Functional Characterization of a Sialyltransferase-Deficient Mutant of *Neisseria Gonorrhoeae*. *Infect. Immun.* **1996**, *64*, 3374–3378.

(36)     Van Vliet, S. J.; Steeghs, L.; Bruijns, S. C. M.; Vaezirad, M. M.; Blok, C. S.; Arenas Busto, J. A.; Deken, M.; Van Putten, J. P. M.; Van Kooyk, Y. Variation of *Neisseria Gonorrhoeae* Lipooligosaccharide Directs Dendritic Cell-Induced T Helper Responses. *PLoS Pathog.* **2009**, *5*, e1000625.

(37)     NAIRN, C. A.; COLE, J. A.; PATEL, P. V.; PARSONS, N. J.; FOX, J. E.; SMITH, H. Cytidine 5'-Monophospho-N-Acetylneuraminic Acid or a Related Compound Is the Low Mr Factor from Human Red Blood Cells Which Induces Gonococcal Resistance to Killing by Human Serum. *J. Gen. Microbiol.* **1988**, *134*, 3295–3306.

(38)     Parsons, N. J.; Andrade, J. R. C.; Patel, P. V.; Cole, J. A.; Smith, H. Sialylation of Lipopolysaccharide and Loss of Absorption of Bactericidal Antibody during Conversion of Gonococci to Serum Resistance by Cytidine 5'-Monophospho-N-Acetyl Neuraminic Acid. *Microb. Pathog.* **1989**, *7*, 63–72.

(39)     Yang, Q. L.; Gotschlich, E. C. Variation of Gonococcal Lipooligosaccharide Structure Is Due to Alterations in Poly-G Tracts in Lgt Genes Encoding Glycosyl Transferases. *J. Exp. Med.* **1996**, *183*, 323–327.

(40)     van Putten, J. P. Phase Variation of Lipopolysaccharide Directs Interconversion of Invasive and Immuno-Resistant Phenotypes of *Neisseria Gonorrhoeae*. *EMBO J.* **1993**, *12*, 4043–4051.

(41)     Gulati, S.; Schoenhofen, I. C.; Whitfield, D. M.; Cox, A. D.; Li, J.; St. Michael, F.; Vinogradov, E. V.; Stupak, J.; Zheng, B.; Ohnishi, M.; Unemo, M.; Lewis, L. A.; Taylor, R. E.; Landig, C. S.; Diaz, S.; Reed, G. W.; Varki, A.; Rice, P. A.; Ram, S. Utilizing CMP-Sialic Acid Analogs to Unravel *Neisseria Gonorrhoeae* Lipooligosaccharide-Mediated Complement Resistance and Design Novel Therapeutics. *PLoS Pathog.* **2015**, *11*, e1005290.

(42)     Gulati, S.; Schoenhofen, I. C.; Lindhout-Djukic, T.; Schur, M. J.; Landig, C. S.; Saha, S.; Deng, L.; Lewis, L. A.; Zheng, B.; Varki, A.; Ram, S. Therapeutic CMP-Nonulosonates against Multidrug-Resistant *Neisseria Gonorrhoeae*. *J. Immunol.* **2020**, *204*, 3283–3295.

(43)     Kuhn, B.; Benz, J.; Greif, M.; Engel, A. M.; Sobek, H.; Rudolph, M. G. The Structure of Human -2,6-Sialyltransferase Reveals the Binding Mode of Complex Glycans. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2013**, *69*, 1826–1838.

(44)     Gilbert, M.; Watson, D. C.; Cunningham, A. M.; Jennings, M. P.; Young, N. M.; Wakarchuk, W. W. Cloning of the Lipooligosaccharide α-2,3-Sialyltransferase from the Bacterial Pathogens *Neisseria Meningitidis* and *Neisseria Gonorrhoeae*. *J. Biol. Chem.* **1996**, *271*, 28271–28276.

(45)     Shapiro, H. M. *Practical Flow Cytometry*, 4th ed.; Wiley-Liss: New York, 2003.

(46)     Jones, M. B.; Oswald, D. M.; Joshi, S.; Whiteheart, S. W.; Orlando, R.; Cobb, B. A. B-Cell-Independent Sialylation of IgG. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 7207–7212.

(47)     Capicciotti, C. J.; Zong, C.; Sheikh, M. O.; Sun, T.; Wells, L.; Boons, G. J. Cell-Surface Glyco-Engineering by Exogenous Enzymatic Transfer Using a Bifunctional CMP-Neu5Ac Derivative. *J. Am. Chem. Soc.* **2017**, *139*,

13342–13348.

(48)     Moons, S. J.; Adema, G. J.; Derks, M. T.; Boltje, T. J.; Büll, C. Sialic Acid Glycoengineering Using N-Acetyl-mannosamine and Sialic Acid Analogs. *Glycobiology* **2019**, *29*, 433–445.

(49)     Nischan, N.; Kohler, J. J. Advances in Cell Surface Glycoengineering Reveal Biological Function. *Glycobiology* **2016**, *26*, 789–796.

3

# Chapter 3
## Supporting information

## Material and methods

α-(2,6)-sialyltransferase (ST6Gal1), CMP-Sia-N$_3$ and CMP-Sia-biotin were prepared as reported[1]. Alkaline phosphatase (FastAP) was purchased from Thermofischer Scientific (EF0651). HRP conjugated anti-biotin antibody (200-032-211) was purchased from Jackson ImmunoResearch Laboratories. Acetylene-PEG$_4$-biotin (CLK-TA105), DBCO-PEG$_4$-biotin (CLK-A105P4), AF488-alkyne (CLK-1277), and DBCO-AF488 (CLK-1278) were purchased from Jena Bioscience.
Chocolate Columbia agar plates were purchased from BioTrading (K018P090KP). Peptone was purchased from Oxoid (LP0085).

## Bacterial strains and culture

*N. gonorrhoeae* WT, Δ STase, isogenic strains A, B and C, were gifted by Prof. Dr. Jos van Putten (Utrecht University)[2–4]. *N. gonorrhoeae* was grown on Chocolate Columbia agar plates at 37°C + 5% CO$_2$ and in HEPES medium at 37°C.

## SEEL of bacteria

One-step SEEL was performed on bacteria grown in liquid culture (5 x 10$^8$ bacteria). The bacteria were washed with buffer and were incubated with SEEL label mix at 37°C for 2 h while rotating. A typical SEEL label mix (50μL) was prepared in medium (PBS/HEPES buffer) with ST6Gal1 (1.05 uL of stock 1 mg/mL), CMP-Sialic acid derivative (50 μM), 0.34 uL BSA (2 mg/mL) and 0.34 μL alkaline phosphatase (1 U/ μL). After SEEL treatment, the bacteria were washed with buffer and prepared for application.

Two-step SEEL was performed similar to one-step SEEL, followed by a click reaction. In case of CuAAC, 100 μL reaction volume contained 100 μM Acetylene-PEG$_4$-biotin, 500 μM CuSO$_4$ and 2.5 mM sodium L-ascorbate. In case of SPAAC 100 μL reaction volume contained 100 μM DBCO-PEG$_4$-biotin. For the fluorophores AF488-alkyne and DBCO-AF488 the concentration was 1 mM.

If the bacteria were heat-inactivated, they were heated at 80°C for 15 min and treated with the described SEEL method.

## Western blotting

In case of (glyco)proteins, the samples were lysed and analyzed with a 10% SDS-PAGE gel for which the gel was run for 45-60 min at 150V. For the western blotting, the gel was electroblotted onto a PVDF membrane. The membrane was blocked (5% milk, 30 – 60 min), washed (1% milk, 5 min), stained with anti-biotin-HRP antibody (1 : 20000 in 1% milk), washed (1 % milk, followed by PBS, 5 min each), and treated with ECL western substrate for signal detection.

## LOS preparation and Tris-Tricine gel

Samples were boiled for 5 min and then treated with protease K (10 uL, 20mg/mL) overnight at 55°C. 3x Laemli buffer was added and the samples were analyzed with a 16% Tris-Tricine gel. The gel ran typically for 3 - 4 h at 20 mA and was then further analyzed by Western blotting or silver staining or in-gel fluorescence.

## Silver staining

Silver staining was performed on a 16% Tris-Tricine gel as described previously.[5] Briefly, the gel was fixed (30 min, 40% ethanol, 5% acetic acid), oxidized (5 min, 0.7% sodium periodic acid, 40% ethanol, 5% acetic acid), washed (3 x 5 min in distilled water), stained (distilled water containing 19% 0.1M NaOH, 1.3% >28% ammonium hydroxide, 3,3% 20% w/v silver nitrate), washed (3 x 5 min in distilled water), developed until bands appeared (distilled water containing 0.1% PFA 37% and 0.1% citric acid 100mg/mL), rinsed with distilled water and stopped (7% acetic acid in distilled water).

## In-gel fluorescence

In-gel fluorescence was measured on Amersham imager 600 using the Green channel (520 nm, cy3).

## Flow cytometry

Samples were fixed with 1% paraformaldehyde and 0.5% BSA. If necessary, bacteria were diluted in PBS + 0.05% BSA to not exceed 20000 counts/s in flow cytometry analysis. Flow cytometry was performed on MACSQuant flow cytometer (Miltenyi Biotech) and analysis was done with FlowJo Software (V10).

# Chapter 3

## Imaging

Bacteria (circa $3 \times 10^8$) were centrifuged, resuspended in HEPES + 1 % BSA and DAPI (4',6-diamidino-2-phenylindole) (1:50) was added. The samples were incubated for 25 min in the dark before washing (milliQ + 0.1% Tween) and then carefully resuspended in ProLong Diamond Antifade Mountant (P36961). 5 uL of sample was taken, put on a poly-L-lysine coated coverslip and mounted on a glass slide. Slides were stored at RT overnight to allow the samples to harden and subsequently stored at 4 °C.

## Growth measurements

Bacteria were treated according to conditions specified in text. Bacteria were diluted to OD=0.05 and the growth was monitored with Synergy HTX multi-mode meter in a hypoxic glove box for 24 h while shaking continuously. Data was exported and analyzed with excel/prism.

## Synthesis of CMP-Sia-AF488

### General Methods and Materials

Alexa Fluor 488 succinimidyl (NHS) ester was purchased from ThermoFisher Scientific. Other reagents were obtained from commercial sources and used as purchased. Dichloromethane (DCM) was freshly distilled using standard procedures. Other organic solvents were purchased anhydrous and used without further purification. Unless otherwise noted, all reactions were carried out at room temperature (RT) in glassware with magnetic stirring. Organic solutions were concentrated under reduced pressure with bath temperatures < 30 °C. Flash column chromatography was carried out on silica gel G60 (Silicycle, 60-200 µm, 60 Å). Thin-layer chromatography (TLC) was carried out on Silica gel 60 F254 (EMD Chemicals Inc.) with detection by UV absorption (254 nm) where applicable, by spraying with 20% sulfuric acid in ethanol followed by charring at ˷150 °C or by spraying with a solution of $(NH_4)_6Mo_7O_{24}\cdot H_2O$ (25 g/L) in 10% sulfuric acid in ethanol followed by charring at ˷150 °C. $^1$H NMR spectra were recorded on a Varian Inova 500 (500 MHz) spectrometer equipped with sun workstations. Mass spectra were recorded on an Applied Biosystems 5800 MALDI-TOF or Shimadzu LCMS-IT-TOF mass spectrometer. The matrix used was 2,5-dihydroxy-benzoic acid (DHB).
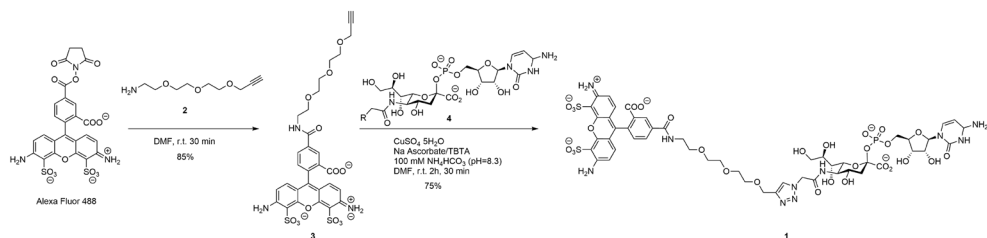
Scheme 1. Synthesis of Alexa Fluor 488 CMP-sialic acid (CMP-Sia-AF488).

## Compound **3**

To a solution of Alexa Fluor 488 NHS (1 eq, 1 mg) in DMF (680 μL), alkyne 2[6] (3 eq, 0.44 mg) and triethylamine (9 eq, 970 μL) were added. The resulting mixture was stirred at room temperature for 30 minutes in darkness until ESI-MS indicated completion of the reaction. After removing solvent under reduced pressure, the crude residue was concentrated to afford 1.1 mg of product as an orange solid in quantitative yield. ESI-MS m/z calcd for $C_{30}H_{29}N_3O_{13}S_2$, [M-1H]$^-$ : 702.1069, found, 702.1099.

## Compound **1** (CMP-Sia-AF488)

To a solution of compound **3** (1.3 eq, 1.1 mg) and CMP-NeuAz[1] (compound **4**, 1 eq, 0.79 mg) in 200 μL 0.1M $NH_4HCO_3$ was added 0.1 M sodium L-ascorbate (12.1 uL), 0.1 M $CuSO_4$ (9.61 uL), and 130 μg TBTA. The resulting mixture was stirred at room temperature for 2h 30 min in darkness. After completion of the reaction as indicated by ESI-MS, the mixture was lyophilized. The residue was purified by a C18 column using a gradient of methanol and water (from 90/10 to 30/70, v/v) to afford compound **1** (1 mg, 62%) as an orange solid. ESI-MS m/z calcd for $C_{50}H_{60}N_{10}O_{29}PS_2$, [M-1H]$^-$: 1359.2712, found, 1359.2730.

# Chapter 3

# Supplemental figures

A.
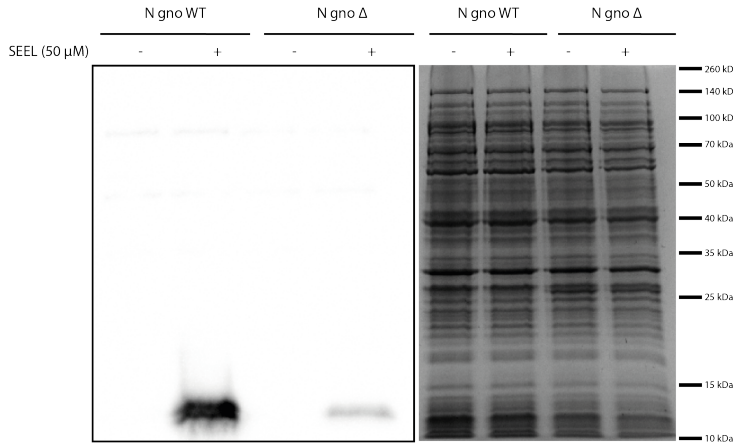


B.
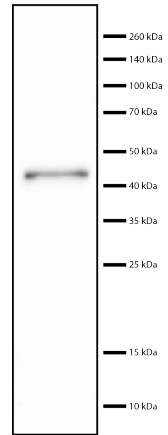


Supplementary figure 1. (A) Western blot and pageblue analysis of protein samples from SEEL treated *N. gonorrhoeae*. The bottom bands are labeled LOS. (B) SEEL label mix analyzed on Western blot. The enzyme ST6Gal1 labels itself in solution.
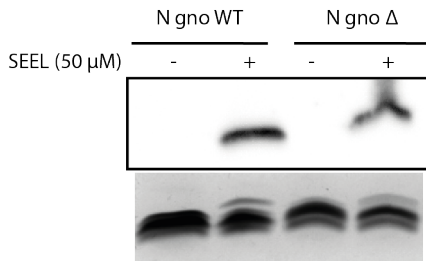


Supplementary figure 2. One-step SEEL after heat-inactivating *N. gonorrhoeae* wildtype and mutant.
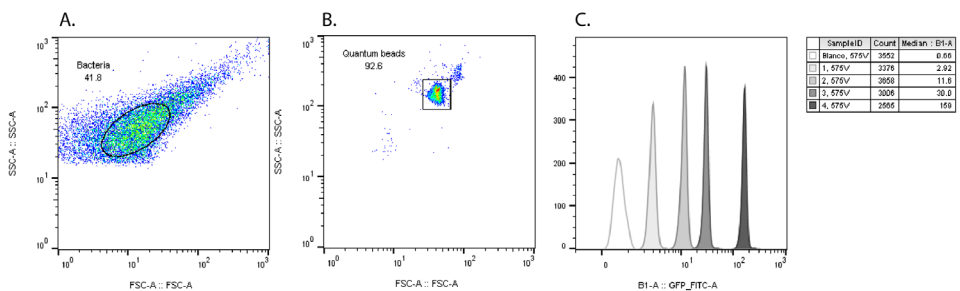


Supplementary figure 3. (A) Gating of *N. gonorrhoeae* for flow cytometry. (B) Gating of quantum beads for quantification of the number of fluorescence labels. (C) The amount of fluorescence in the 488 channel (FITC) for the different quantum beads with the mean indicated in the legend.

# Selective exoenzymatic labeling of LOS of *N. gonorrhoeae*

GFP  Merge  GFP  Merge
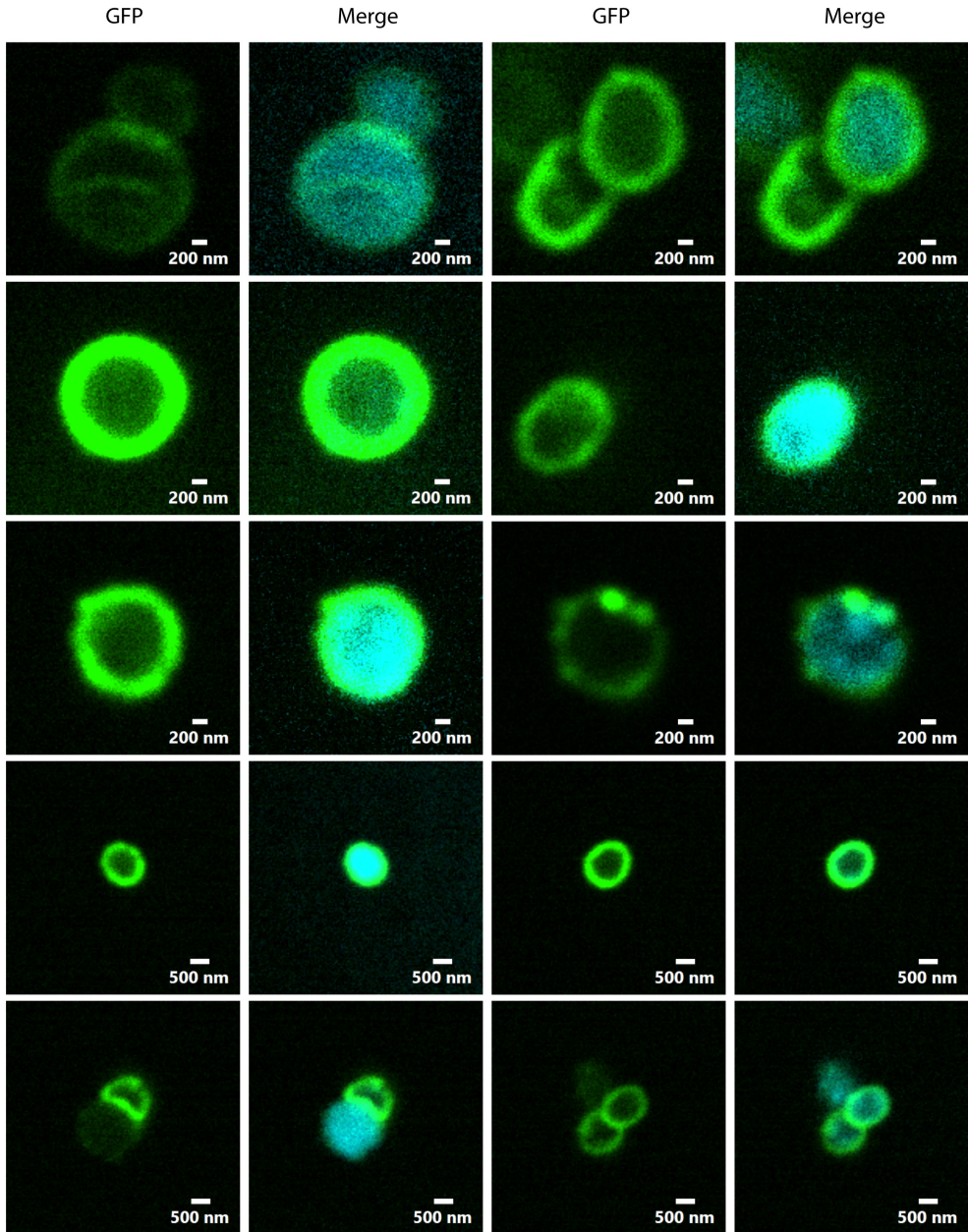
Supplementary figure 4. Fluorescence microscopy images of SEEL treated *N. gonorrhoeae* with CMP-Sia-AF488.
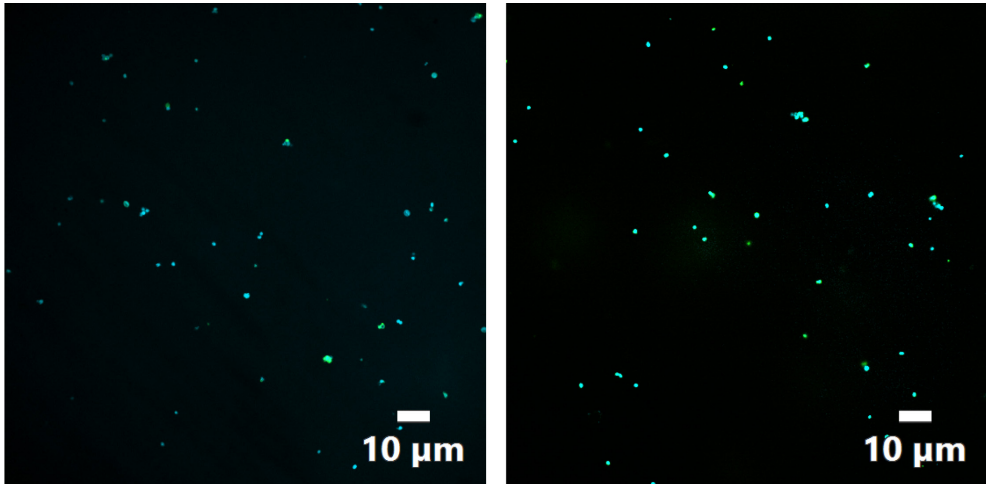
Supplementary figure 5. Fluorescence microscopy images of SEEL treated *N. gonorrhoeae* in two overviews (merge of 488 and 405 channel).
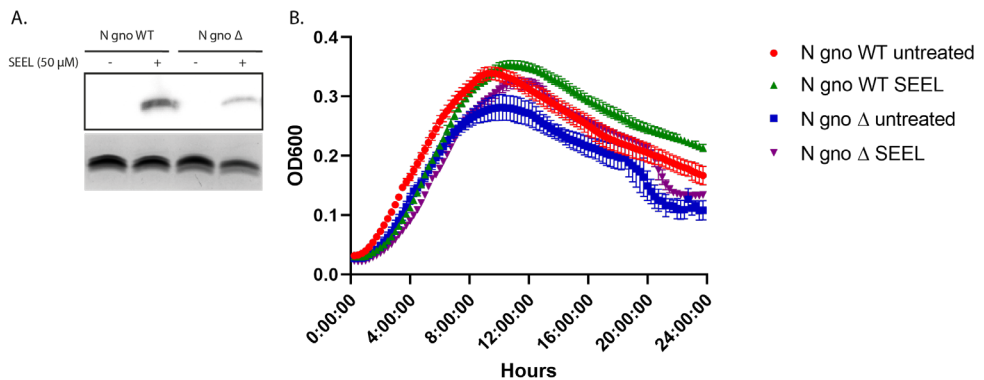


Supplementary figure 6. (A) One-step SEEL labeling in HEPES buffer, which is used for the cytotoxicity assays. (B) OD600 measured over time to test cell viability of untreated and SEEL treated bacteria for both wildtype and mutant strain of *N. gonorrhoeae*.
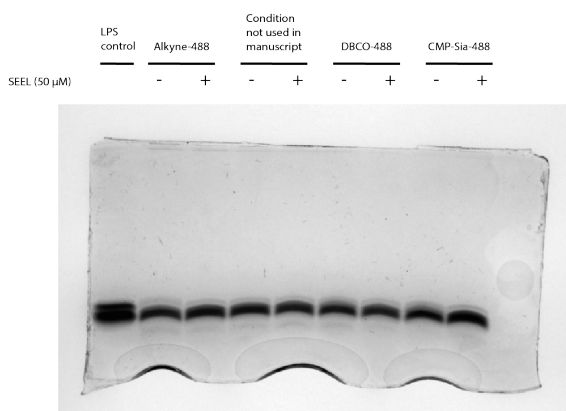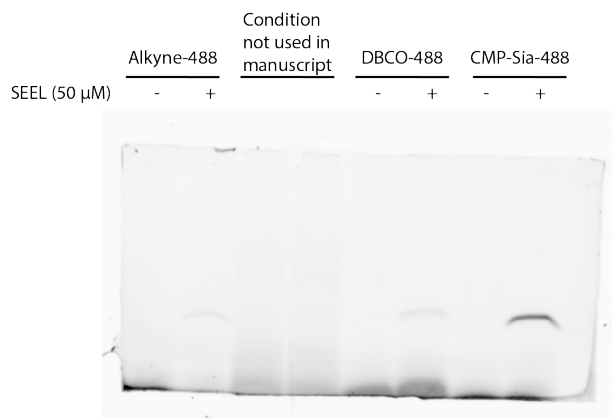
Supplementary figure 7. Raw data images of in-gel fluorescence (top) and silver stain (bottom) that were used in Figure 2E.

# Supplemental references

(1)    Sun, T.; Yu, S. H.; Zhao, P.; Meng, L.; Moremen, K. W.; Wells, L.; Steet, R.; Boons, G. J. One-Step Selective Exoenzymatic Labeling (SEEL) Strategy for the Biotinylation and Identification of Glycoproteins of Living Cells. *J. Am. Chem. Soc.* **2016**, *138*, 11575–11582.

(2)    Bramley, J.; de Hormaeche, R. D.; Constantinidou, C.; Nassif, X.; Parsons, N.; Jones, P.; Smith, H.; Cole, J. A Serum-Sensitive, Sialyltransferase-Deficient Mutant of *Neisseria Gonorrhoeae* Defective in Conversion to Serum Resistance by CMP-NANA or Blood Cell Extracts. *Microb. Pathog.* **1995**, *18*, 187–195.

(3)    Gill, M. J.; Mcquillen, D. P.; Van Putten, J. P. M.; Wetzler, L. M.; Bramley, J.; Crooke, H.; Parsons, N. J.; Cole, J. A.; Smith, H. Functional Characterization of a Sialyltransferase-Deficient Mutant of *Neisseria Gonorrhoeae*. *Infect. Immun.* **1996**, *64*, 3374–3378.

(4)    Van Vliet, S. J.; Steeghs, L.; Bruijns, S. C. M.; Vaezirad, M. M.; Blok, C. S.; Arenas Busto, J. A.; Deken, M.; Van Putten, J. P. M.; Van Kooyk, Y. Variation of *Neisseria Gonorrhoeae* Lipooligosaccharide Directs Dendritic Cell-Induced T Helper Responses. *PLoS Pathog.* **2009**, *5*, e1000625.

(5)    Tsai, C. M.; Frasch, C. E. A Sensitive Silver Stain for Detecting Lipopolysaccharides in Polyacrylamide Gels. *Anal. Biochem.* **1982**, *119*, 115–119.

(6)    Norberg, O.; Deng, L.; Yan, M.; Ramström, O. Photo-Click Immobilization of Carbohydrates on Polymeric Surfaces—A Quick Method to Functionalize Surfaces for Biomolecular Recognition Studies. *Bioconjug. Chem.* **2009**, *20*, 2364–2370.

# Chapter 4

## Glycoengineering cell surface glycoconjugates on bacterial pathogens with CMP-sialic acid analogues

Glycoconjugates on the bacterial cell surface play an important role in host-microbe interactions. Often, *N*-acetylneuraminic acid (Neu5Ac or sialic acid) is a key component of these glycoconjugates. The importance of this monosaccharide is illustrated by its overrepresentation in bacterial glycoconjugates that resemble human glycans. The functional roles of Neu5Ac on these mimicking bacterial glycans are not fully understood, but it has been shown to interact with sialic acid binding receptors of the host and shield the bacteria from recognition by the immune system. The ability to directly detect and manipulate Neu5Ac on bacterial glycoconjugates of live bacteria is a key approach to further investigate the functional roles of this modification at the molecular level. Here, we report the application of three glycoengineering techniques: Metabolic Oligosaccharide Engineering, selective exoenzymatic labeling and labeling via native sialyltransferases, to incorporate Neu5Ac with a reporter group into the lipooligosaccharides of a selection of Gram-negative bacteria. We show that these techniques are complementary based on the chosen bacterium and the goal of engineering. Furthermore, we demonstrated that the native sialyltransferases of several mucosal pathogens were able to accept extracellular nucleotide sugar analogs, thereby introducing unnatural sialosides onto their lipooligosaccharides. In addition, for a number of bacterial pathogens we show for the first time that they are able to incorporate Neu5Ac in their bacterial glycoconjugates. The choice of glycoengineering techniques to be used ultimately depends on the bacterial species and the goal of engineering, but they all can make an important contribution to unraveling glycan interactions between host and bacteria.

*Hanna de Jong, Erianna I. Alvarado Melendez, Jet E. M. Hartman, Jun Yang Ong, Maria J. Moure, Bart W. Bardoel, Astrid P. Heikema, Geert-Jan Boons, Marc M. S. M. Wösten, Tom Wennekes*

# Chapter 4

## Introduction

An important monosaccharide residue in glycans is 5-*N*-acetylneuraminic acid (Neu5Ac), or commonly referred to as sialic acid. This monosaccharide is a 9-carbon glycan that is often found at the terminal position of glycan classes like gangliosides, glycoproteins and polysialic acids[1]. In mammals, the function of this monosaccharide is diverse, and plays a key role in the interaction with the immune system[2,3]. Neu5Ac is used by bacteria as well, either as a carbon source or as a building block for their own glycoconjugates [4,5]. In some cases, the incorporation of Neu5Ac into the bacterial glycoconjugates leads to resemblance between human and bacterial glycans, a form of glycan mimicry[6]. As a result, the bacterial glycans can engage with human sialic acid binding receptors (Siglecs) or shield the bacteria from the immune system[6]. Neu5Ac is a key monosaccharide in glycan mimicry, and also in other host-pathogen interactions[7]. Despite its importance, the exact role of Neu5Ac on bacterial glycoconjugates is not fully elucidated on the molecular level. To gain further insight, there is a need to detect and manipulate Neu5Ac through glycoengineering techniques.

Several engineering approaches exist to perturb and study the roles of glycans[8–12]. Currently, a widely applied chemical approach to study glycans is the use of carbohydrates that are modified with (bioorthogonal) reporters. In this technique, named Metabolic Oligosaccharide Engineering (MOE), the carbohydrate analogs are taken up by the cell, metabolically processed into their sugar nucleotide equivalent and transferred onto the cellular glycans by native glycosyltransferases. The metabolically incorporated carbohydrate analog can subsequently be studied via its reporter by covalently attaching a tag, for example a fluorophore or biotin. MOE has also been applied on microbes to engineer their cell wall[13–17], image[18,19], discover glycoproteins[20,21] or to develop new antibacterial strategies[22–24]. Although MOE is a very useful technique, it has a few limitations. In order to modify glycans, the used carbohydrate analogs must be able to get inside the cell, metabolically processed by the native salvage pathway enzymes and finally accepted as substrates by native intracellular glycosyltransferases. Deficiencies in any of these steps often results in only limited control over both reporter incorporation level and site, e.g. glycoproteins or glycolipids[25]. The metabolically generated sugar nucleotide analogs are often substrates for multiple related intracellular glycosyltransferases. Therefore, MOE can also result in various linkage types, for example α2,3- vs α2,6-sialic acid linkages, which complicates their subsequent analysis and interpretation of their biological effects.

Another glycoengineering technique is selective exoenzymatic labeling[26,27], recently applied to the bacterium *Neisseria gonorrhoeae* by us (Chapter 3). SEEL uses an externally applied recombinant glycosyltransferase to selectively label the outside of a cell with tailor-made sugar nucleotide analogs. In case of *N. gonorrhoeae*, derivatives of CMP-Neu5Ac were introduced with reporter groups such as an azide, a biotin or fluorescent dye on the lipooligosaccharides (LOS) of this bacterium. Advantages of SEEL are the precise incorporation of a glycan onto a defined substrate on the cell surface with a known linkage type due to the inherent activity of the externally applied glycosyltransferase. The established linkage type can be non-native to the bacterium, like α2,6-Neu5Ac instead of the naturally occurring α2,3-Neu5Ac on *N. gonorrhoeae*[28], which is an unique feature of SEEL when compared to other glycoengineering techniques. Another advantage is that SEEL is non-toxic to bacteria in contrast to the often higher concentrations of monosaccharide analogs applied during MOE. A disadvantage of SEEL can be, as observed by us when applied to the bacterium *N. gonorrhoeae*, that the degree of sialylation is lower compared to the level achieved by native glycosyltransferases (Chapter 3).

During the application of SEEL to *N. gonorrhoeae* we observed that wildtype bacteria were capable of incorporating the sugar nucleotide analogs with their own sialyltransferases. For *N. gonorrhoeae* it was previously reported that it scavenges CMP-Neu5Ac from the environment and transfers this onto its lipooligosaccharides through native sialyltransferases[29]. It has been shown that *N. gonorrhoeae* can also utilize nucleotide sugar analogs with small modifications such as an azide or other glycans including legionaminic acid and keto-deoxyoctanoate[30–32]. In line with these previous observations, we found that also larger modifications on the C5 position, like a fluorescent dye, were accepted as substrates by the sialyltransferase. Intrigued by the ability to harness these enzymes to engineer their own cell surface glycoconjugates, we set out to test other bacteria for their ability to transfer sialic acid nucleotide analogs.

To test these glycoengineering techniques and the sialyltransferase activity of a set of bacterial pathogens, three species were chosen to further tune and optimize testing conditions: Nontypeable *Haemophilus influenzae* (NTHi), *Neisseria gonorrhoeae*, and *Campylobacter jejuni*. All three bacterial species contain Neu5Ac in their cell surface glycoconjugates and this has a profound effect on their interaction with the human host. NTHi is a Gram-negative bacterium, facultatively anaerobic, and an opportunistic pathogen that can cause otitis

media. NTHi can take up Neu5Ac from outside the cell and incorporate it into its LOS, leading to glycan mimicry and increased resistance against serum-mediated killing[6,33,34]. Another target was *N. gonorrhoeae* which is a Gram-negative, diplococci bacteria that causes the sexually transmitted gonorrhea infection. Although *N. gonorrhoeae* cannot synthesize sialic acid nucleotide sugars by itself since it lacks the required enzyme CMP synthetase[35], it can scavenge these nucleotide sugars from the environment[29,36], as mentioned previously. It was also demonstrated that this bacteria can scavenge other 9-carbon nucleotide sugars, like 5-glycolyl-neuraminic acid (Neu5Gc) or legionaminic acid (Leg5Ac7Ac)[30,31]. The third target was *C. jejuni*, which is a food-borne pathogen and a common cause of human gastroenteritis. In rare cases *C. jejuni* can lead to the autoimmune disease Guillain-Barré Syndrome (GBS) and the development of this disease is associated with *C. jejuni*'s LOS[37]. Neu5Ac appears to have an important effect on *C. jejuni*'s interaction with the immune system of the host[38].

Here, we first report the use of MOE, SEEL and sialyltransferase activity of three bacterial pathogens of which is known that they sialylate their cell surface glycoconjugates, namely Nontypeable *Haemophilus influenzae*, *Neisseria gonorrhoeae*, and *Campylobacter jejuni*. After having established the sialyltransferase activity of the enzymes by the three bacterial pathogens, a blast search with the bacterial glycosyltransferases was performed to select other bacteria that might sialylate their cell surface. We show that a wide range of pathogenic bacteria is able to scavenge nucleotide sugars and incorporate these into their cell surface glycoconjugates. This glycoengineering technique, in addition to the existing techniques like MOE and SEEL, is a very useful and fast method to study the function of Neu5Ac on bacterial glycoconjugates.

## Results

### Nontypeable *Haemophilus influenzae*

NTHi is an opportunistic pathogen that can utilize Neu5Ac as a carbon source or incorporate it into bacterial glycoconjugates. In a previous glycoengineering study, Heise et al.[34] demonstrated that sialic acid probes get incorporated into the LOS via the metabolic pathway and this leads to resistance against serum-mediated killing.

First, we repeated the experiment by Heise et al. and applied MOE to NTHi[34]. The sialic acid analog Neu5Az was indeed incorporated onto the LOS, but not

by the bacteria that lack the required metabolic pathway to take up and process the sialic acid analog, as demonstrated with a sialic acid transporter mutant (ΔSiaP) (Figure 1A). There was no incorporation of the Neu5Ac analog into any other cell surface glycoproteins (Figure 1B). A possible downside of MOE is the toxicity associated with the probe, which is usually an azido analog. For the highest concentration of probe (1 mM), indeed a decrease in growth is observed that indicates cell toxicity (SI Figure 1). Taken together, this data confirms the previous report of MOE for NTHi[34].

Next, we tested whether SEEL could be used to label NTHi. Therefore, a biotinylated nucleotide sugar was used in combination with a sialyltransferase from a mammalian or bacterial origin. To exclude the contribution of NTHi's sialyltransferases to the labeling, the bacteria were heated at 80 °C for 15 minutes (Chapter 3). Under these conditions we unfortunately observed no labeling of NTHi by SEEL. However, without the heat-inactivation step and without exogeneous enzyme, labeling was observed (Figure 1C). This indicated that the native sialyltransferases from NTHi are capable of modifying the LOS with an unnatural Neu5Ac analog.

Since the sialyltransferases of NTHi were capable of scavenging the nucleotide sugar, we wanted to investigate this further. First, the CMP-Neu5Ac derivative with the smallest modification, an azide in this case (CMP-Neu5Az), was applied to NTHi at two different concentrations (Figure 1D). In both cases a clear signal was observed indicating the incorporation of the Neu5Ac analog onto the bacterial LOS. Even a nucleotide sugar with a much larger modification on the C5 position, a biotin already covalently connected to the nucleotide sugar (CMP-Neu5Biotin) could be incorporated (Figure 1E). Next, we wanted to test the minimum concentration of nucleotide sugar required to observe a signal on Western blot. An experiment using a concentration range of CMP-Neu5Az showed that from the highest concentration of 200 μM to as low as 0,8 μM a signal could be observed (Figure 1F). After having established the minimal concentration of nucleotide sugar needed, a set of 10 clinical nontypeable and 2 typeable *Haemophilus influenzae* isolates with previously unidentified sialyltransferase activity were tested. The glycan structures of these nontypeable and typeable clinical isolates are unidentified. Seven out of ten NTHi incorporated the nucleotide sugar derivative and a moderate signal for both typeables was detected(Figure 1G). The use of tailor-made sugar nucleotides provided a rapid way to explore the sialyltransferase activity of multiple strains and to identify the

sialylation of previously uncharacterized cell surface glycoconjugates of clinical isolates.
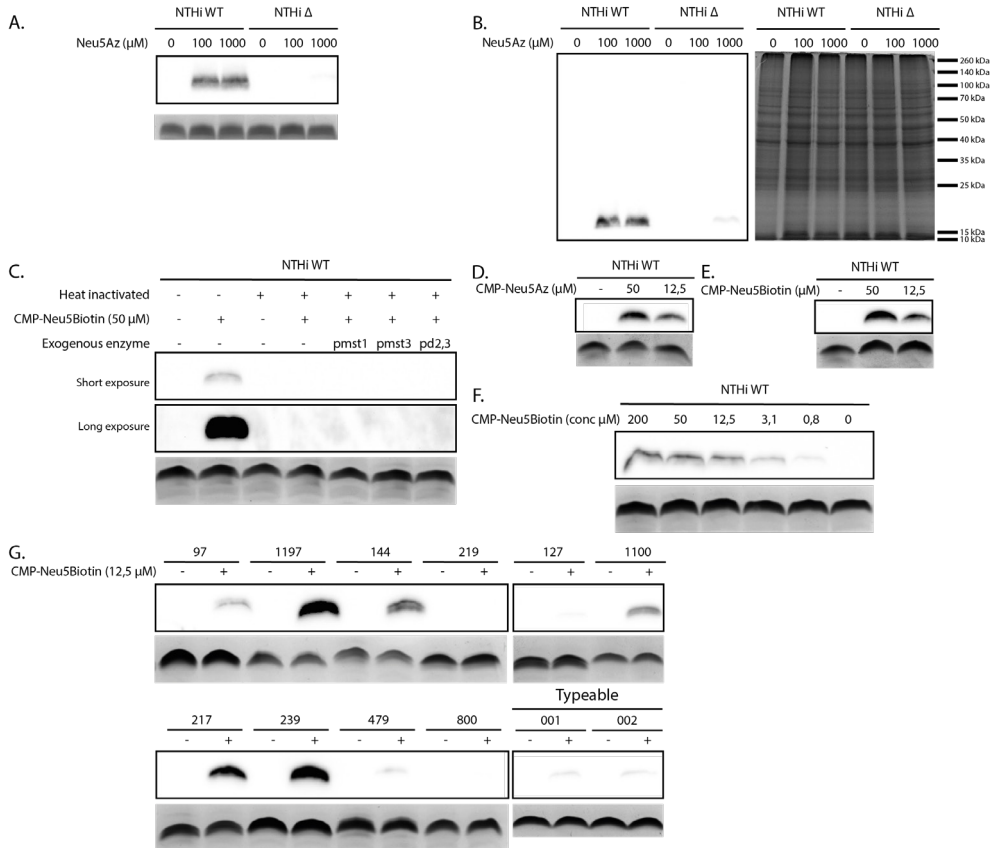


**Figure 1. Glycoengineering techniques MOE, SEEL and native sialyltransferases, applied to NTHi. (A)** MOE with Neu5Az shows incorporation into the LOS of NTHi wildtype (WT), but not for a sialic acid transporter mutant (Δ). **(B)** MOE with Neu5Az does not label any cell surface glycoconjugates of NTHi. **(C)** SEEL of NTHi with different sialyltransferases does not show labeling of the LOS of NTHi. **(D, E)** The native sialyltransferases of NTHi are capable of transferring a nucleotide sugar with a modified C5 position of sialic acid (CMP-Neu5Az and CMP-Neu5Biotin). **(F)** A concentration range shows incorporation of the nucleotide sugar (CMP-Neu5Biotin) into the LOS with different intensities on Western blot. **(G)** The native sialyltransferases of nontypeable and typeable *Haemophilus influenzae* show incorporation of CMP-Neu5Biotin to varying degrees.

## Neisseria gonorrhoeae

*N. gonorrhoeae* is a bacterial pathogen that uses Neu5Ac to mask itself from the immune system via glycan mimicry[6]. This bacterium does not have a CMP synthetase and is known to scavenge the nucleotide sugar from the extracellular environment[29,35]. CMP-Neu5Ac is transferred by the sialyltransferases (Lst) of *N. gonorrhoeae* that transfers the sialic acid with an α2,3-linkage onto the LOS for most strains[35]. There have been conflicting reports on the location

of the sialyltransferase, which could either be surface-bound or present in the cytoplasm[32,36]. As *N. gonorrhoeae* lacks the genes for the biosynthesis of its own CMP-Neu5Ac, MOE with the monosaccharide is not possible. This was verified by probing the bacteria with Neu5Az followed by a click reaction and observing no labeling of the LOS (Figure 2A). In a previous study we demonstrated that *N. gonorrhoeae* can be labeled with SEEL (Chapter 3). *N. gonorrhoeae* was included in our glycoengineering approach because the sialyltransferase Lst has been observed by us and other groups to transfer modified sugar nucleotide analogs[30–32]. To potentially harness the activity of this enzyme for glycoengineering, we continued studying its substrate scope. The azide modified nucleotide sugar (CMP-Neu5Az) was also incorporated onto the LOS, as it could be observed on Western Blot after a click reaction which was either copper- or strain-promoted, CuAAC vs SPAAC (Figure 2B). Notably, larger modifications on the C5 position were also accepted as could be seen for CMP-Neu5Biotin (Figure 2C). Similarly, the fluorescently labeled derivative was also incorporated with high efficiency as previously reported (Chapter 3).

*N. gonorrhoeae* uses Neu5Ac to become resistant against serum-mediated killing[29,39]. Gulati et al., previously showed that *N. gonorrhoeae* has altered resistance against serum-mediated killing when the bacteria were modified with other sialic acids than Neu5Ac, such as legionaminic acid[30,31]. However, the 5-azido derivative of the nucleotide sugar was not reported yet, so we investigated its effect in the presence of serum. In a serum-resistance assay with bacteria covered by Neu5Az, the bacteria were less resistant against serum compared to bacteria treated with CMP-Neu5Ac (Figure 2D). This observation fits the trend that a small modification on the glycan can perturb the interaction with factors from serum, which also has been seen for the 9-azido derivative of the nucleotide sugar[30].

4

A.



B.



C.



D.
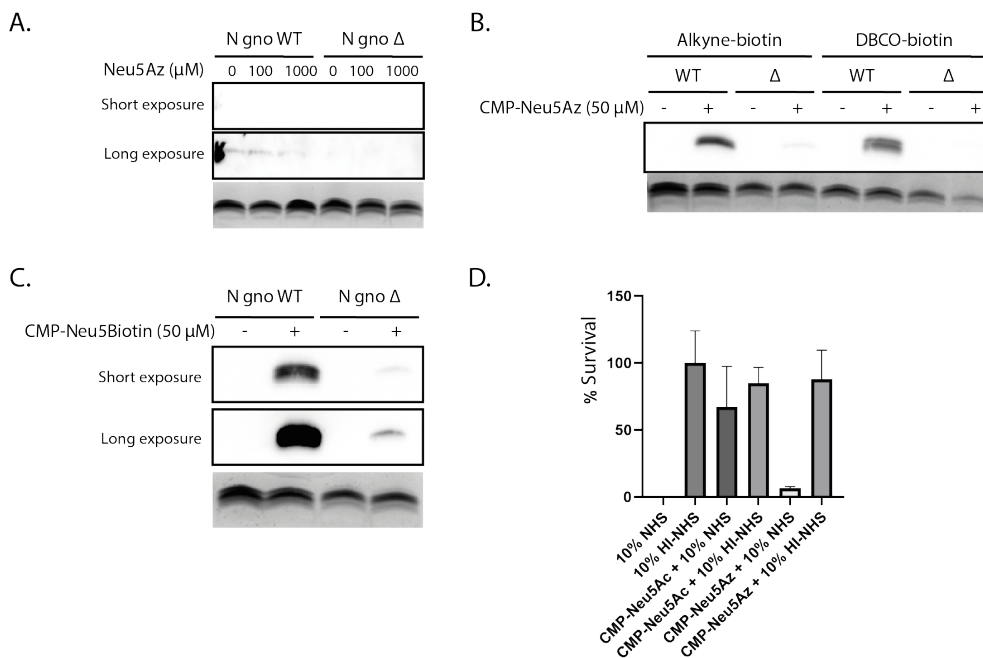


**Figure 2. Glycoengineering approaches MOE and native sialyltransferases applied to *Neisseria gonorrhoeae*. (A)** MOE with Neu5Az on *N. gonorrhoeae* does not show incorporation of the monosaccharide probe into the LOS for both wildtype and a sialyltransferase mutant (Δ). **(B)** The native sialyltransferases of *N. gonorrhoeae* incorporate an azide-modified nucleotide sugar into the LOS as observed after a click reaction with a biotin derivate through a CuAAC reaction (alkyne-biotin) and SPAAC reaction (DBCO-biotin). **(C)** The native sialyltransferases of *N. gonorrhoeae* incorporate a nucleotide sugar with a biotin derivative into the LOS. **(D)** Serum resistance of *N. gonorrhoeae* which was grown without or with a nucleotide sugar; CMP-Neu5Ac or CMP-Neu5Az. Colony forming units were counted after incubation in the presence of Normal Human Serum (NHS) or Heat-inactivated NHS (HI-NHS), and reported as % survival.

## *Campylobacter jejuni*

The enteropathogen *C. jejuni* is the leading cause of bacterial gastroenteritis in the developing world. In 1 per 1000 cases, an infection with this bacterium is associated with the autoimmune disease Guillain-Barré syndrome (GBS)[40]. The LOS of these *C. jejuni* strains mimic human glycans and this glycan mimicry is considered to be one of the causes of autoantibodies that target the host's nerves during disease[41]. *C. jejuni*'s LOS can be sialylated and bind to host receptors, such as Siglec 1 and 7[42,43]. Sialylation occurs via the sialyltransferase CstII, which can create either α2,3- or both α2,3- and α2,8-linkages depending on a point mutation[44,45]. Because of the link between the sialylated LOS of *C. jejuni* and GBS, the glycoconjugates of this bacterium form an intriguing target to manipulate and study. We focused on three strains of *C. jejuni* that are associated

with GBS: GB2, GB11 and GB19, along with the corresponding mutant strains of CstII and CMP synthetase (Figure 3A)[45–47].

First, MOE was attempted to probe *C. jejuni*'s LOS. For both the neuraminic acid probes with an azide on either the C5 or C9 position, clear labeling was observed for the wildtypes and as expected not for the sialyltransferase mutants (Figure 3B). Next, SEEL was attempted in the CstII mutants of all three strains with recombinant CstII as the external sialyltransferase, but no signal was observed that indicated labeling of the LOS by SEEL (Figure 3C). CstII has been characterized as a bifunctional sialyltransferase having both α2,3-sialyltransferase and α2,8-sialyltransferase activities. We hypothesized that CstII might not be able to initiate sialylation on the LOS acceptor sites of these specific *C. jejuni* strains, and it needed an additional sialyltransferase to create the required α2,3-Neu5Ac first. Therefore, we next attempted SEEL with either CstI or a combination of both CstI and CstII, and in both cases, labeling was observed (Figure 3D). The labeling intensity was equal for both conditions, which suggests that CstI contributes more to labeling than CstII in this case. This points to successful labeling by SEEL using an externally applied sialyltransferase from *Campylobacter* origin. Therefore, it was not surprising that the native enzymes of the WT strains could also transfer the nucleotide sugar derivative, so without any enzymes added from an external source (Figure 3C). After having established that *C. jejuni*'s own sialyltransferases could transfer the azide derivative of the nucleotide sugars, we wanted to explore this as a third glycoengineering approach. Therefore, the nucleotide sugar that was already covalently coupled to a biotin was used. This molecule was also accepted as a substrate by the bacterial sialyltransferase (Figure 3E). In an attempt to quantify the number of labeled bacteria by this approach, a fluorescent reporter group (Alexa Fluor 488 dye) was introduced on the bacteria for analysis with flow cytometry. In case of a two-step approach, labeling with an azide and then a click reaction, the signal was faint but observable for strains GB2 and GB11, but not for GB19 (SI Figure 2). Two types of click reactions were employed, one copper-catalyzed (CuAAC) and one strain-promoted (SPAAC). For SPAAC faint bands were visible for bacteria that were not treated with the azide derivative, which indicates aspecific labeling. Since the labeling of the two-step approach was not clear for in-gel fluorescence, the one-step approach was applied with a derivative that is already covalently coupled to a fluorophore, CMP-Neu5AF488. With this approach all strains could be labeled (Figure 3F) and by flow cytometry a modest shift in fluorescence was observed for all *C. jejuni* strains as shown

4

for GB2 (Figure 3G). In the absence of the sialyltransferase, GB2 ΔCstII, the fluorescence slightly increased, which indicated that there is a small amount of background fluorescence originating from the CMP-Neu5AF488 aspecifically binding to the bacteria
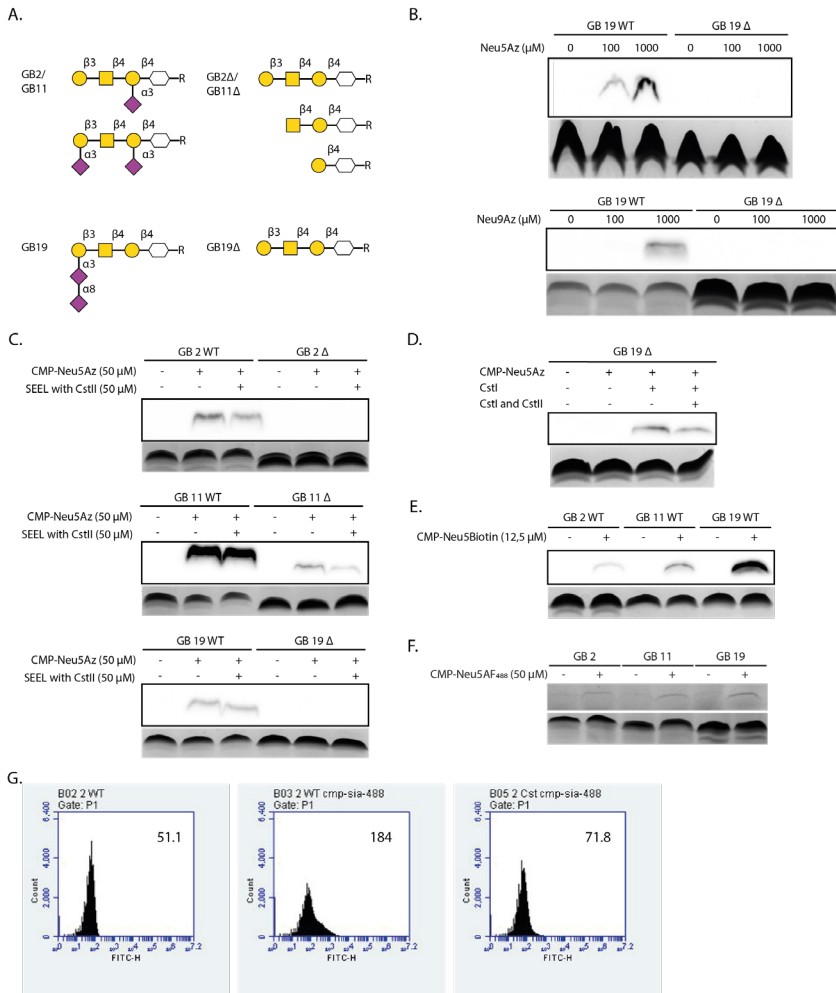


**Figure 3. Glycoengineering techniques MOE, SEEL and native sialyltransferases, applied to *Campylobacter jejuni*. (A)** LOS structures of *C. jejuni* strains associated with GBS; GB2, 11 and 19. **(B)** MOE with Neu5Az and Neu9Az on GB19 wildtype and a sialyltransferase mutant. The probe is incorporated into the LOS of the WT. **(C)** SEEL with exogenous applied CstII does not show labeling for the sialyltransferase mutants of all strains. The wildtype bacteria do show labeling, which indicates that native sialyltransferases use the azide derivative of the nucleotide sugar to label the LOS. **(D)** SEEL with exogenous applied CstI and CstII shows successful labeling for the sialyltransferase mutant. **(E)** The native sialyltransferases use the biotin derivative of the nucleotide sugar to label the LOS. **(F)** In-gel fluorescence for LOS which was modified by native sialyltransferases with CMP-Neu5AF488. **(G)** Flow cytometry data of untreated GB2 WT, GB2 WT and GB2 ΔCstII incubated with CMP-Neu5AF488 for 2 hours. Mean fluorescence of FITC indicated in the top right corner for each condition.

## Scanning for sialyltransferase activity in other pathogenic bacteria

After having established that NTHi, *N. gonorrhoeae* and *C. jejuni* could incorporate modified nucleotide sugars, we investigated whether other bacterial species were also capable of this sialyltransferase activity and would incorporate modified nucleotide sugars into either their LOS/LPS. Moreover, the screen for sialyltransferase activity of multiple strains of NTHi (Figure 1G) proofed a fast method to identify functional sialylation of LOS, which is a quick indication whether the bacteria might use Neu5Ac *in vivo*. To screen a set of bacteria for sialyltransferase activity, first a blast search was performed with the six sialyltransferases we knew or expected to be capable of this activity (Lst (α2,3-ST, *N. gonorrhoeae* strain F62), Lic3A (α-2,3-ST, NTHi R2886), SiaA (α-2,8-ST, NTHi), LsgB (NTHi), Cst-II (α-2,3-ST, *C. jejuni*) and PmST1 (α-2,3-ST, *P. multocida* pm70)). From the hits of the blast search, a selection was made based on whether the bacteria were pathogenic or commensal in humans, zoonotic or displayed glycan mimicry. Additionally, we included practical considerations such as, commercial or in-house availability, and the compatibility with the safety regulations of a biosafety level 2 laboratory. We selected 17 bacterial species to screen for sialyltransferase activity with CMP-Neu5Az. Labeling of the LOS/LPS was observed for 9 out of 17 bacterial (Figure 4).

4
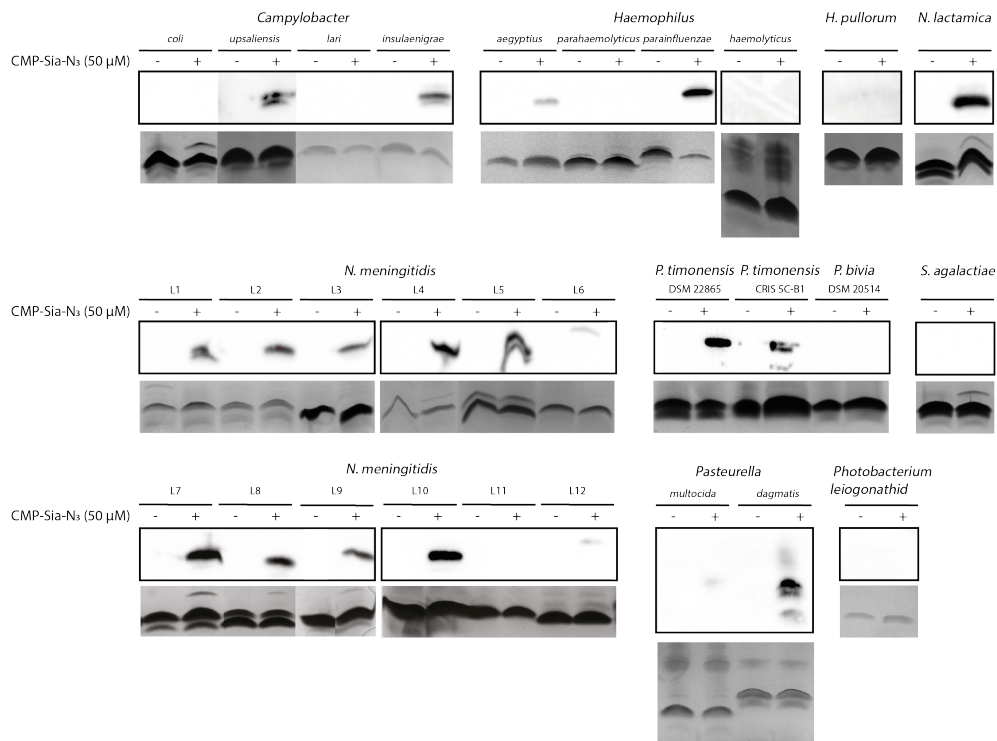
**Figure 4. Scan for the sialyltransferase activity of a set of pathogenic bacteria.** The azide derivative of the nucleotide sugar was incorporated by native sialyltransferases followed by a strain-promoted click reaction with a DBCO-biotin. The LOS/LPS was analyzed on Western blot and by silver stain. 9 out of 17 bacterial strains tested positive.

To confirm the results of the sialyltransferase assay, the genetic potential of these strains was tested by PCR. In most cases, we were able to detect the sialyltransferase gene as identified by blast result (Table 1). The lack of detection of a sialyltransferase gene in some strains might be due to the misfit of primers, as the nucleotide sequence of a number of the tested strains were not available.

Table 1. Preliminary table of PCR results. Bacterial strains that were hits from the sialyltransferase blast were tested for genomic potential through PCR. Abbreviations: - tested negative; + tested positive; sialyltransferases (lic3A, LsgB, SiaA, CstII, Pmst, Lst); TBD to be determined; ND not determined.

| | ST activity with CMP-Neu5Az | Blasted STases | PCR STases |
|---|---|---|---|
| *Campylobacter coli* | - | CstII | - |
| *Campylobacter insulaenigrae* | + | CstII | CstII |
| *Campylobacter lari* | - | CstII | - |
| *Campylobacter upsaliensis* | + | CstII | CstII |
| *Haemophilus aegyptius* | + | lic3A | lic3A, LsgB, SiaA |
| *Haemophilus haemolyticus* | - | lic3A | - |
| *Haemophilus parahaemolyticus* | - | SiaA | - |
| *Haemophilus parainfluenzae* | + | SiaA | TBD |
| *Helicobacter pullorum* | - | Pmst | TBD |
| *Neisseria lactamica* | + | Lst | Lst |
| *Neisseria meningitidis* B | + | Lst | Lst |
| *Neisseria meningitidis* C | + | Lst | Lst |
| *Neisseria meningitidis* W-135 | + | Lst | Lst |
| *Neisseria meningitidis* Y | + | Lst | Lst |
| *Neisseria meningitis* L1 | + | Lst | Lst |
| *Neisseria meningitis* L2 | + | Lst | Lst |
| *Neisseria meningitis* L3 | + | Lst | Lst |
| *Neisseria meningitis* L4 | + | Lst | Lst |
| *Neisseria meningitis* L5 | + | Lst | Lst |
| *Neisseria meningitis* L6 | - | Lst | Lst |
| *Neisseria meningitis* L7 | + | ND | Lst |
| *Neisseria meningitis* L8 | + | ND | Lst |
| *Neisseria meningitis* L9 | + | ND | Lst |
| *Neisseria meningitis* L10 | + | ND | Lst |
| *Neisseria meningitis* L11 | - | ND | Lst |
| *Neisseria meningitis* L12 | - | ND | Lst |
| *Pasteurella dagmatis* | + | Pmst | Pmst |
| *Pasteurella multocida* | + | Pmst | Pmst |
| *Photobacterium leiognathi* | - | Pmst | TBD |
| *Prevotella bivia* | + | SiaA | TBD |
| *Prevotella timonensis* | + | SiaA | TBD |
| *Streptococcus agalactiae* | - | Lst | Lst |

4

Chapter 4

# Discussion

In this work, we have evaluated three different glycoengineering techniques: MOE, SEEL and native sialyltransferases, for their capability to detect and modify terminal Neu5Ac on bacterial LOS of NTHi, *N. gonorrhoeae* and *C. jejuni*. Furthermore, we tested the scope of the native sialyltransferases on a set of bacterial pathogens.

In case of MOE, NTHi and *C. jejuni* could be labeled with Neu5Az or Neu9Az, but as expected not in *N. gonorrhoeae* since this pathogen does not have the required CMP synthetase to convert the monosaccharide into the nucleotide sugar. Our data confirm the result of NTHi[34] and the expected outcome for *N. gonorrhoeae*, and show for the first time labeling with a neuraminic acid derivative of the LOS of *C. jejuni*. MOE has been reported to also incorporate the microbial sialic acids legionaminic acid and pseudaminic acid in flagella, another glycoconjugate, of *C. jejuni*[17,48].

Previously, we had demonstrated that the glycoengineering technique SEEL could label the LOS of *N. gonorrhoeae* (Chapter 3). Similarly, we applied this technique on the LOS of NTHi by using different bacterial sialyltransferases that target a terminal galactose, but no labeling was observed. We speculate that the LOS of this NTHi strain might not be a good acceptor for the chosen sialyltransferases. In case of *C. jejuni*, SEEL of the LOS was possible with the enzyme CstI. This sialyltransferase creates an α2,3-Neu5Ac on the acceptor and is known to be active *in vitro*, but the *in vivo* activity is poorly described[49,50]. Our data show that SEEL of LOS on *C. jejuni* is possible with enzymes from *Campylobacter* origin and future experiments could test sialyltransferases from different bacterial or mammalian sources. Since we also observed labeling by the native sialyltransferases from *C. jejuni*, we focused our efforts on that glycoengineering technique and showed successful labeling with the biotinylated and fluorescent nucleotide sugar derivative.

From the scan for sialyltransferase activity the following strains showed labeling with sialic acid on their LOS/LPS: *Prevotella timonensis*, *Haemopilus aegyptius*, *Haemophilus parainfluenzae*, *Campylobacter upsaliensis*, *Campylobacter insulaenigrae*, *Pasteurella multocida*, *Pasteurella dagmatis*, *Neisseria lactamica*, and *Neisseria meningitidis*. For a more detailed discussion about the reported use of Neu5Ac by these strains, the reader is referred to the supporting information.

It is important to note that the activity of the sialyltransferases of the set of bacterial pathogens was tested in one experimental condition. Although these experiments were performed using standardized conditions, such as buffer and amount of CMP-Neu5Az, they do not exclude the possibility that under different environmental or culturing conditions, the sialyltransferase activity of these bacteria might alter. Also, it is important to note that the experiment focused on the cell surface glycoconjugates LOS/LPS, and did not study the possible incorporation into the capsule. Since we could not visualize the CPS on Western blot, the possible sialylation of the capsule might be checked with flow cytometry or confocal microscopy.

The identification of sialyltransferase activity for several bacterial pathogens could point to uptake mechanisms of nucleotide sugars by bacteria or sialyltransferases that are surface bound, similar to *N. gonorrhoeae*[36]. The enzymatic activity of these ectopic glycosyltransferases has been considered a possibility for glycosylation of mammalian cells, yet the evidence for the presence and stability of free nucleotide sugars outside the cell has been limited[51,52]. Recent studies have shown that platelets are a source for nucleotide sugars, which can be released after a stress trigger, and that sialylation via this machinery is possible *in vivo*[52–55]. The amount of free CMP-Neu5Ac was determined at 104 pmol/mg platelet or an average concentration of 75 μM in serum samples of resting mice[52,55]. Similarly, the source of neuraminic acid for *N. gonorrhoeae* originates from human red blood cells[29]. In context of bacterial glycosylation, the scan for sialyltransferase activity hints at a similar mechanism in which bacteria use nucleotide sugars available from the environment. It remains to be determined if these nucleotide sugars originate from platelets as well, or other sources such as liberation in co-culture with other bacteria. An alternative mechanism for the incorporation of nucleotide sugars could be the uptake of these nucleotide sugars from the environment and sialyltransferase activity in the cytoplasm instead of on the surface. For *N. gonorrhoeae* it was reported that the sialyltransferases are surface bound, but this finding has been recently disputed[32,36]. Instead, an alternative pathway was suggested in which the nucleotide sugars were taken up by an unknown transporter and then transferred inside the cell onto the LOS during its assembly on lipid A. From our data no definite conclusions can be drawn about the exact location of the sialyltransferase. However, if such a transporter would exist, this study suggests it would accept a broad substrate range since large modifications such as a biotin or fluorescent group on the sialic acid were tolerated. Additionally, our microscopy data for *N. gonorrhoeae* incubated with

CMP-Neu5AF488 indicated no fluorescence in the cytoplasm (Chapter 3), yet an even higher resolution might be needed to exclude the presence of the nucleotide sugar close to the inner membrane. Since the fluorescence signal was not dispersed throughout the cytoplasm and localized outside the membrane, this might indicate that the sialyltransferases are active in the periplasm or on the surface. The exact location of the sialyltransferase poses an interesting question regarding the biosynthetic pathway and it would be interesting to establish which of the mechanisms is responsible for the glycosylation, whether the nucleotide sugars are taken up and metabolized intracellularly or transferred via ectopic glycosyltransferases by bacteria.

In a comparison of the three glycoengineering techniques, MOE is an established technique to incorporate monosaccharides with reporter groups. MOE requires optimization per probe and bacteria since it relies on the metabolic processing by bacteria and these pathways differ[9]. In some cases, it is not possible to apply MOE, because the required metabolic pathway is absent, as shown for *N. gonorrhoeae*. In addition, at high concentrations of probe, the technique might be toxic to bacteria, as observed for NTHi in our experiments. Second, the technique SEEL allows the incorporation of a glycan with a reporter group on a defined substrate (Chapter 3). Additionally, SEEL is not toxic to bacteria (Chapter 3). A possible disadvantage of SEEL is that the amount of labeling is lower compared to native sialyltransferases. Third, native sialyltransferases can be harnessed to incorporate glycans with reporter groups. An advantage is that it does not require the expression and purification of exogeneous sialyltransferases and since no additional metabolic processing besides the transfer are required, only the substrate needs to be synthesized in the lab. The incorporated glycan would be processed similar to the natural substrate, which does not allow for a perturbation of the system by introducing an unknown linkage type for instance, but does allow for the study of these glycans in a 'natural' setting. Additionally, this technique reflects the activity of the enzymes *in vitro* and provides additional information next to genomic potential in rapid fashion. Overall, we consider these glycoengineering techniques complementary and based on the goal, a preferred technique will be chosen. The novel insight that a broad range of bacterial pathogens have sialyltransferase activity provide the opportunity to study the glycosylation machinery of the LOS of bacteria in more detail. In the future, the functional role of Neu5Ac on bacterial glycoconjugates could be studied through these glycoengineering techniques.

# References

(1) Varki A; RL, S.; R, S. Sialic Acids and Other Nonulosonic Acids. In *Essentials of Glycobiology*; Varki A, Cummings RD, Esko JD, et al., Ed.; Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.

(2) Varki, A.; Gagneux, P. Multifarious Roles of Sialic Acids in Immunity. *Ann. N. Y. Acad. Sci.* **2012**, *1253*, 16–36.

(3) Chang, Y. C.; Nizet, V. The Interplay between Siglecs and Sialylated Pathogens. *Glycobiology* **2014**, *24*, 818–825.

(4) Vimr, E. R. Unified Theory of Bacterial Sialometabolism: How and Why Bacteria Metabolize Host Sialic Acids. *ISRN Microbiol.* **2013**, 816713.

(5) Severi, E.; Hood, D. W.; Thomas, G. H. Sialic Acid Utilization by Bacterial Pathogens. *Microbiology* **2007**, *153*, 2817–2822.

(6) de Jong, H.; Wösten, M. M. S. M.; Wennekes, T. Sweet Impersonators: Molecular Mimicry of Host Glycans by Bacteria. *Glycobiology* **2021**.

(7) Poole, J.; Day, C. J.; von Itzstein, M.; Paton, J. C.; Jennings, M. P. Glycointeractions in Bacterial Pathogenesis. *Nat. Rev. Microbiol.* **2018**, *16*, 440–452.

(8) Dube, D. H.; Champasa, K.; Wang, B. Chemical Tools to Discover and Target Bacterial Glycoproteins. *Chem. Commun.* **2011**, *47*, 87–101.

(9) Nischan, N.; Kohler, J. J. Advances in Cell Surface Glycoengineering Reveal Biological Function. *Glycobiology* **2016**, *26*, 789–796.

(10) Griffin, M. E.; Hsieh-Wilson, L. C. Glycan Engineering for Cell and Developmental Biology. *Cell Chem. Biol.* **2016**, *23*, 108–121.

(11) Critcher, M.; O'Leary, T.; Huang, M. L. Glycoengineering: Scratching the Surface. *Biochem. J.* **2021**, *478*, 703–719.

(12) Edgar, L. J. Engineering the Sialome. *ACS Chem. Biol.* **2021**.

(13) Sadamoto, R.; Niikura, K.; Sears, P. S.; Liu, H.; Wong, C.-H.; Suksomcheep, A.; Tomita, F.; Monde, K.; Nishimura, S.-I. Cell-Wall Engineering of Living Bacteria. *J. Am. Chem. Soc.* **2002**, *124*, 9018–9019.

(14) Sadamoto, R.; Niikura, K.; Ueda, T.; Monde, K.; Fukuhara, N.; Nishimura, S.-I. Control of Bacteria Adhesion by Cell-Wall Engineering. *J. Am. Chem. Soc.* **2004**, *126*, 3755–3761.

(15) Dumont, A.; Malleron, A.; Awwad, M.; Dukan, S.; Vauzeilles, B. Click-Mediated Labeling of Bacterial Membranes through Metabolic Modification of the Lipopolysaccharide Inner Core. *Angew. Chem. Int. Ed.* **2012**, *51*, 3143–3146.

(16) Swarts, B. M.; Holsclaw, C. M.; Jewett, J. C.; Alber, M.; Fox, D. M.; Siegrist, M. S.; Leary, J. A.; Kalscheuer, R.; Bertozzi, C. R. Probing the Mycobacterial Trehalome with Bioorthogonal Chemistry. *J. Am. Chem. Soc.* **2012**, *134*, 16123–16126.

(17) Liu, F.; Aubry, A. J.; Schoenhofen, I. C.; Logan, S. M.; Tanner, M. E. The Engineering of Bacteria Bearing Azido-Pseudaminic Acid-Modified Flagella. *ChemBioChem* **2009**, *10*, 1317–1320.

(18) Geva-Zatorsky, N.; Alvarez, D.; Hudak, J. E.; Reading, N. C.; Erturk-Hasdemir, D.; Dasgupta, S.; Von Andrian, U. H.; Kasper, D. L. *In Vivo* Imaging and Tracking of Host-Microbiota Interactions via Metabolic Labeling of Gut Anaerobic Bacteria. *Nat. Med.* **2015**, *21*, 1091–1100.

(19) Hudak, J. E.; Alvarez, D.; Skelly, A.; Von Andrian, U. H.; Kasper, D. L. Illuminating Vital Surface Molecules of Symbionts in Health and Disease. *Nat. Microbiol.* **2017**, *2*, 17099.

(20) Champasa, K.; Longwell, S. A.; Eldridge, A. M.; Stemmler, E. A.; Dube, D. H. Targeted Identification of Glycosylated Proteins in the Gastric Pathogen *Helicobacter Pylori (Hp)*. *Mol. Cell. Proteomics* **2013**, *12*, 2568–2586.

(21) Besanceney-Webler, C.; Jiang, H.; Wang, W.; Baughn, A. D.; Wu, P. Metabolic Labeling of Fucosylated Glycoproteins in *Bacteroidales* Species. *Bioorganic Med. Chem. Lett.* **2011**, *21*, 4989–4992.

(22) Kaewsapsak, P.; Esonu, O.; Dube, D. H. Recruiting the Host's Immune System to Target *Helicobacter Pylori's* Surface Glycans. *ChemBioChem* **2013**, *14*, 721–726.

(23) Memmel, E.; Homann, A.; Oelschlaeger, T. A.; Seibel, J. Metabolic Glycoengineering of *Staphylococcus Aureus* Reduces Its Adherence to Human T24 Bladder Carcinoma Cells. *Chem. Commun.* **2013**, *49*, 7301–7303.

(24) Tra, V. N.; Dube, D. H. Glycans in Pathogenic Bacteria-Potential for Targeted Covalent Therapeutics and Imaging Agents. *Chem. Commun.* **2014**, *50*, 4659–4673.

(25) Bussink, A. P.; Van Swieten, P. F.; Ghauharali, K.; Scheij, S.; Van Eijk, M.; Wennekes, T.; Van Der Marel, G. A.; Boot, R. G.; Aerts, J. M. F. G.; Overkleeft, H. S. N-Azidoacetylmannosamine-Mediated Chemical Tagging of Gangliosides. *J. Lipid Res.* **2007**, *48*, 1417–1421.

(26) Mbua, N. E.; Li, X.; Flanagan-Steet, H. R.; Meng, L.; Aoki, K.; Moremen, K. W.; Wolfert, M. A.; Steet, R.; Boons, G. J. Selective Exo-Enzymatic Labeling of N-Glycans on the Surface of Living Cells by Recombinant ST6Gal I. *Angew. Chem. Int. Ed.* **2013**, *52*, 13012–13015.

(27) Sun, T.; Yu, S. H.; Zhao, P.; Meng, L.; Moremen, K. W.; Wells, L.; Steet, R.; Boons, G. J. One-Step Selective

4

Exoenzymatic Labeling (SEEL) Strategy for the Biotinylation and Identification of Glycoproteins of Living Cells. *J. Am. Chem. Soc.* **2016**, *138*, 11575–11582.

(28)   Gilbert, M.; Watson, D. C.; Cunningham, A. M.; Jennings, M. P.; Young, N. M.; Wakarchuk, W. W. Cloning of the Lipooligosaccharide α-2,3-Sialyltransferase from the Bacterial Pathogens *Neisseria Meningitidis* and *Neisseria Gonorrhoeae*. *J. Biol. Chem.* **1996**, *271*, 28271–28276.

(29)   NAIRN, C. A.; COLE, J. A.; PATEL, P. V.; PARSONS, N. J.; FOX, J. E.; SMITH, H. Cytidine 5'-Monophospho-N-Acetylneuraminic Acid or a Related Compound Is the Low Mr Factor from Human Red Blood Cells Which Induces Gonococcal Resistance to Killing by Human Serum. *J. Gen. Microbiol.* **1988**, *134*, 3295–3306.

(30)   Gulati, S.; Schoenhofen, I. C.; Whitfield, D. M.; Cox, A. D.; Li, J.; St. Michael, F.; Vinogradov, E. V.; Stupak, J.; Zheng, B.; Ohnishi, M.; Unemo, M.; Lewis, L. A.; Taylor, R. E.; Landig, C. S.; Diaz, S.; Reed, G. W.; Varki, A.; Rice, P. A.; Ram, S. Utilizing CMP-Sialic Acid Analogs to Unravel *Neisseria Gonorrhoeae* Lipooligosaccharide-Mediated Complement Resistance and Design Novel Therapeutics. *PLoS Pathog.* **2015**, *11*, e1005290.

(31)   Gulati, S.; Schoenhofen, I. C.; Lindhout-Djukic, T.; Schur, M. J.; Landig, C. S.; Saha, S.; Deng, L.; Lewis, L. A.; Zheng, B.; Varki, A.; Ram, S. Therapeutic CMP-Nonulosonates against Multidrug-Resistant *Neisseria Gonorrhoeae*. *J. Immunol.* **2020**, *204*, 3283–3295.

(32)   Jen, F. E. C.; Ketterer, M. R.; Semchenko, E. A.; Day, C. J.; Jennings, M. P.; Seib, K. L.; Apicella, M. A. The Lst Sialyltransferase of *Neisseria Gonorrhoeae* Can Transfer Keto-Deoxyoctonate as the Terminal Sugar of Lipooligosaccharide: A Glyco-Achilles Heel That Provides a New Strategy for Vaccines to Prevent Gonorrhea. *MBio* **2021**, *12*.

(33)   Apicella, M. A. Nontypeable *Haemophilus Influenzae*: The Role of N-Acetyl-5-Neuraminic Acid in Biology. *Front. Cell. Infect. Microbiol.* **2012**, *2*, 19.

(34)   Heise, T.; Langereis, J. D.; Rossing, E.; de Jonge, M. I.; Adema, G. J.; Büll, C.; Boltje, T. J. Selective Inhibition of Sialic Acid-Based Molecular Mimicry in *Haemophilus Influenzae* Abrogates Serum Resistance. *Cell Chem. Biol.* **2018**, *25*, 1279-1285.e8.

(35)   Mubaiwa, T. D.; Semchenko, E. A.; Hartley-Tassell, L. E.; Day, C. J.; Jennings, M. P.; Seib, K. L. The Sweet Side of the Pathogenic *Neisseria*: The Role of Glycan Interactions in Colonisation and Disease. *Pathogens and Disease.* 2017, p ftx063.

(36)   Shell, D. M.; Chiles, L.; Judd, R. C.; Seal, S.; Rest, R. F. The *Neisseria* Lipooligosaccharide-Specific Alpha-2,3-Sialyltransferase Is a Surface-Exposed Outer Membrane Protein. *Infect. Immun.* **2002**, *70*, 3744–3751.

(37)   Yuki, N.; Susuki, K.; Koga, M.; Nishimoto, Y.; Odaka, M.; Hirata, K.; Taguchi, K.; Miyatake, T.; Furukawa, K.; Kobata, T.; Yamada, M. Carbohydrate Mimicry between Human Ganglioside GM1 and *Campylobacter Jejuni* Lipooligosaccharide Causes Guillain-Barré Syndrome. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 11404–11409.

(38)   Heikema, A. P. Host-Pathogen Interactions in Guillain-Barré Syndrome: The Role of *Campylobacter Jejuni* Lipooligosaccharide Sialylation, Erasmus University Rotterdam, 2013.

(39)   Parsons, N. J.; Andrade, J. R. C.; Patel, P. V.; Cole, J. A.; Smith, H. Sialylation of Lipopolysaccharide and Loss of Absorption of Bactericidal Antibody during Conversion of Gonococci to Serum Resistance by Cytidine 5'-Monophospho-N-Acetyl Neuraminic Acid. *Microb. Pathog.* **1989**, *7*, 63–72.

(40)   Jasti, A. K.; Selmi, C.; Sarmiento-Monroy, J. C.; Vega, D. A.; Anaya, J. M.; Gershwin, M. E. Guillain-Barré Syndrome: Causes, Immunopathogenic Mechanisms and Treatment. *Expert Rev. Clin. Immunol.* **2016**, *12*, 1175–1189.

(41)   Van Den Berg, B.; Walgaard, C.; Drenthen, J.; Fokke, C.; Jacobs, B. C.; Van Doorn, P. A. Guillain-Barré Syndrome: Pathogenesis, Diagnosis, Treatment and Prognosis. *Nat. Rev. Neurol.* **2014**, *10*, 469–482.

(42)   Heikema, A. P.; Koning, R. I.; Rico, S. D. dos S.; Rempel, H.; Jacobs, B. C.; Endtz, H. P.; van Wamel, W. J. B.; Samsom, J. N. Enhanced, Sialoadhesin-Dependent Uptake of Guillain-Barré Syndrome-Associated *Campylobacter Jejuni* Strains by Human Macrophages. *Infect. Immun.* **2013**, *81*, 2095–2103.

(43)   Heikema, A. P.; Jacobs, B. C.; Horst-Kreft, D.; Huizinga, R.; Kuijf, M. L.; Endtz, H. P.; Samsom, J. N.; van Wamel, W. J. B. Siglec-7 Specifically Recognizes *Campylobacter Jejuni* Strains Associated with Oculomotor Weakness in Guillain-Barré Syndrome and Miller Fisher Syndrome. *Clin. Microbiol. Infect.* **2013**, *19*, E106–E112.

(44)   Gilbert, M.; Karwaski, M. F.; Bernatchez, S.; Young, N. M.; Taboada, E.; Michniewicz, J.; Cunningham, A. M.; Wakarchuk, W. W. The Genetic Bases for the Variation in the Lipo-Oligosaccharide of the Mucosal Pathogen, *Campylobacter Jejuni*. Biosynthesis of Sialylated Ganglioside Mimics in the Core Oligosaccharide. *J. Biol. Chem.* **2002**, *277*, 327–337.

(45)   Godschalk, P. C. R.; Kuijf, M. L.; Li, J.; St. Michael, F.; Ang, C. W.; Jacobs, B. C.; Karwaski, M. F.; Brochu, D.; Moterassed, A.; Endtz, H. P.; Van Belkum, A.; Gilbert, M. Structural Characterization of *Campylobacter Jejuni* Lipooligosaccharide Outer Cores Associated with Guillain-Barré and Miller Fisher Syndromes. *Infect. Immun.* **2007**, *75*, 1245–1254.

(46)  Godschalk, P. C. R.; Heikema, A. P.; Gilbert, M.; Komagamine, T.; Wim Ang, C.; Glerum, J.; Brochu, D.; Li, J.; Yuki, N.; Jacobs, B. C.; Van Belkum, A.; Endtz, H. P. The Crucial Role of *Campylobacter Jejuni* Genes in Anti-Ganglioside Antibody Induction in Guillain-Barré Syndrome. *J. Clin. Invest.* **2004**, *114*, 1659–1665.

(47)  Louwen, R.; Heikema, A.; Van Belkum, A.; Ott, A.; Gilbert, M.; Ang, W.; Endtz, H. P.; Bergman, M. P.; Nieuwenhuis, E. E. The Sialylated Lipooligosaccharide Outer Core in *Campylobacter Jejuni* Is an Important Determinant for Epithelial Cell Invasion. *Infect. Immun.* **2008**, *76*, 4431–4438.

(48)  Meng, X.; Boons, G.; Wösten, M. M. S. M.; Wennekes, T. Metabolic Labeling of Legionaminic Acid in Flagellin Glycosylation of Campylobacter Jejuni Identifies Maf4 as a Putative Legionaminyl Transferase Angewandte. *Angew. Chem - Int. Ed.* **2021**, *60*, 24811–24816.

(49)  Gilbert, M.; Brisson, J. R.; Karwaski, M. F.; Michniewicz, J.; Cunningham, A. M.; Wu, Y.; Young, N. M.; Wakarchuk, W. W. Biosynthesis of Ganglioside Mimics in Campylobacter Jejuni OH4384. Identification of the Glycosyltransferase Genes, Enzymatic Synthesis of Model Compounds, and Characterization of Nanomole Amounts by 600-MHz ^{1}H and ^{13}C NMR Analysis. *J. Biol. Chem.* **2000**, *275*, 3896–3906.

(50)  Heikema, A. P.; Strepis, N.; Horst-Kreft, D.; Huynh, S.; Zomer, A.; Kelly, D. J.; Cooper, K. K.; Parker, C. T. Biomolecule Sulphation and Novel Methylations Related to Guillain-Barré Syndrome-Associated Campylobacter Jejuni Serotype HS:19. *Microb. Genomics* **2021**, *7*.

(51)  Roth, S.; Mcguire, E. J.; Roseman, S. Evidence for Cell-Surface Glycosyltransferases. *J. Cell Biol.* **1971**, *51*, 536–547.

(52)  Jones, M. B.; Oswald, D. M.; Joshi, S.; Whiteheart, S. W.; Orlando, R.; Cobb, B. A. B-Cell-Independent Sialylation of IgG. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 7207–7212.

(53)  Manhardt, C. T.; Punch, P. R.; Dougher, C. W. L.; Lau, J. T. Y. Extrinsic Sialylation Is Dynamically Regulated by Systemic Triggers in Vivo. *J. Biol. Chem.* **2017**, *292*, 13514–13520.

(54)  Lee, M. M.; Nasirikenari, M.; Manhardt, C. T.; Ashline, D. J.; Hanneman, A. J.; Reinhold, V. N.; Lau, J. T. Y. Platelets Support Extracellular Sialylation by Supplying the Sugar Donor Substrate. *J. Biol. Chem.* **2014**, *289*, 8742–8748.

(55)  Wandall, H. H.; Rumjantseva, V.; Sørensen, A. L. T.; Patel-Hett, S.; Josefsson, E. C.; Bennett, E. P.; Italiano, J. E.; Clausen, H.; Hartwig, J. H.; Hoffmeister, K. M. The Origin and Function of Platelet Glycosyltransferases. *Blood* **2012**, *120*, 625–635.

4

## Chapter 4

## Supporting information

## Material and methods

α-(2,6)-sialyltransferase (ST6Gal1), CMP-Neu5Az and CMP-Neu5Biotin were prepared as reported[1]. The synthesis of CMP-Neu5AF488 was described in Chapter 3. The enzymes pmst1, pmst3 and pd2,3 were a kind gift from Gerlof Bosman, Utrecht University. Alkaline phosphatase (FastAP) was purchased from Thermofischer Scientific (EF0651). HRP conjugated anti-biotin antibody (200-032-211) was purchased from Jackson ImmunoResearch Laboratories. Acetylene-PEG$_4$-biotin (CLK-TA105), DBCO-PEG$_4$-biotin (CLK-A105P4), AF488-alkyne (CLK-1277), and DBCO-AF488 (CLK-1278) were purchased from Jena Bioscience. Normal Human Serum (NHS) was a kind gift from UMC Utrecht.

## Bacterial strains and culture conditions

Table 1. Bacterial strains used in this research

| Species | Strain |
| --- | --- |
| *Campylobacter coli* | |
| *Campylobacter insulaenigrae* | DSM 17739 |
| *Campylobacter jejuni* | GB 2 |
| *Campylobacter jejuni* | GB 2 ΔCstII |
| *Campylobacter jejuni* | GB 2 ΔCMP |
| *Campylobacter jejuni* | GB 11 |
| *Campylobacter jejuni* | GB 11 ΔCstII |
| *Campylobacter jejuni* | GB 11 ΔCMP |
| *Campylobacter jejuni* | GB 19 |
| *Campylobacter jejuni* | GB 19 ΔCstII |
| *Campylobacter jejuni* | GB 19 ΔCMP |
| *Campylobacter lari* | DSM 11375 |
| *Campylobacter upsaliensis* | DSM 5365 |
| *Haemophilus aegyptius* | DSM 21187 |
| *Haemophilus haemolyticus* | DSM 103601 |
| *Haemophilus parahaemolyticus* | DSM 21417 |
| *Haemophilus parainfluenzae* | DSM 8978 |
| *Helicobacter pullorum* | DSM 23160 |
| *Neisseria gonorrhoeae* | F62 |
| *Neisseria gonorrhoeae* | F62 ΔLst |
| *Neisseria lactamica* | |
| *Neisseria meningitidis* L1 | 126E |
| *Neisseria meningitidis* L2 | 35E |
| *Neisseria meningitidis* L3 | 6275 |
| *Neisseria meningitidis* L4 | 891 |
| *Neisseria meningitidis* L5 | M981 |
| *Neisseria meningitidis* L6 | M992 |
| *Neisseria meningitidis* L7 | 6155 |
| *Neisseria meningitidis* L8 | M978 |
| *Neisseria meningitidis* L9 | 120M |
| *Neisseria meningitidis* L10 | 7880 |
| *Neisseria meningitidis* L11 | 7889 |
| *Neisseria meningitidis* L12 | 7897 |
| *Neisseria meningitidis* serogroup B | |
| *Neisseria meningitidis* serogroup C | |
| *Neisseria meningitidis* serogroup W-135 | |
| *Neisseria meningitidis* serogroup Y | |
| Nontypeable *Haemophilus influenzae* | R2886 |
| Nontypeable *Haemophilus influenzae* | R2886 ΔSiaP |
| Nontypeable *Haemophilus influenzae* | 97 |
| Nontypeable *Haemophilus influenzae* | 1197 |
| Nontypeable *Haemophilus influenzae* | 144 |
| Nontypeable *Haemophilus influenzae* | 219 |
| Nontypeable *Haemophilus influenzae* | 127 |
| Nontypeable *Haemophilus influenzae* | 1100 |
| Nontypeable *Haemophilus influenzae* | 217 |
| Nontypeable *Haemophilus influenzae* | 239 |
| Nontypeable *Haemophilus influenzae* | 479 |
| Nontypeable *Haemophilus influenzae* | 800 |
| *Pasteurella dagmatis* | DSM 22969 |
| *Pasteurella multocida* | DSM 5281 |
| *Photobacterium leiognathid* | DSM 21260 |
| *Prevotella bivia* | DSM 20514 |
| *Prevotella timonensis* | DSM 22865 |
| *Prevotella timonensis* | CRIS 5C-B1 |
| *Streptococcus agalactiae* | |
| Typeable *Haemophilus influenzae* | 001 |
| Typeable *Haemophilus influenzae* | 002 |

4

Chapter 4

*Campylobacter jejuni* GB2, GB11, and GB19, wildtype and mutants were a kind gift from Astrid Heikema, ErasmusMC Rotterdam. *Neisseria meningitidis* L1-12 were a kind gift from Nina van Sorge, Amsterdam UMC and the Netherlands Reference Laboratory for Bacterial Meningitis (NRLBM). *Neisseria gonorrhoeae* and *Neisseria meningitidis* serogroup B, C, W-135 and Y were a kind gift from Jos van Putten, Utrecht University. Nontypeable *Haemophilus influenzae* R2886 and mutant were a kind gift from Jeroen Langereis, Radboudumc. Other nontypeable *Haemophilus influenzae* and typeable strains were a kind gift from Clinical Infectiology, Utrecht University.

## *Campylobacter* species

*Campylobacter coli* and *Campylobacter lari* were grown on blood agar sheep plates (BioTrading, K004P090KP) at 37 °C under microaerophilic conditions (80% $N_2$, 10% $CO_2$, 5% $O_2$, 5% $H_2$). *Campylobacter insulaenigrae* and *Campylobacter upsaliensis* were grown under the same conditions for 2 days. *Campylobacter jejuni* strains were grown under microaerophilic conditions at 42 °C with a second passage for the mutant strains to allow sufficient growth. *Campylobacter coli*, *insulaenigrae*, *lari* and *uspaliens* were grown at 37 °C and *campylobacter jejuni* at 42 °C under microaerophilic conditions in Heart Infusion broth (BioTrading, K716F100GH).

## *Haemophilus* species

*Haemophilus aegyptius*, *Haemophilus haemolyticus*, *Haemophilus parahaemolyticus*, *Haemophilus parainfluenzae* and typeable *Haemophilus influenzae* were grown on cholate agar plates with Vitox (Thermo Scientific, PO5090A 0A) at 37°C with 5% $CO_2$. Nontypeable *Haemophilus influenzae* was grown aerobically. *Haemophilus* strains were cultured in Mueller Hinton or Brain Heart infusion supplemented with *Haemophilus* test supplement (Oxoid, SR0158E), except for *Haemophilus haemolyticus* which was cultured in Tryptone soy broth.

## *Helicobacter pullorum*

*Helicobacter pullorum* was grown on blood agar sheep plates (BioTrading, K004P090KP) at 37 °C under microaerophilic conditions (80% $N_2$, 10% $CO_2$, 5% $O_2$, 5% $H_2$) and in Heart Infusion broth (BioTrading, K716F100GH).

## *Neisseria* species

*Neisseria gonorrhoeae, lactamica, and meningitidis* were cultured on chocolate columbia agar (BioTrading, K018P090KP) at 37°C with 5% $CO_2$ and in HEPES medium at 37°C.

## *Pasteurella* species

*Pasteurella multocida* and *Pasteurella dagmatis* were grown on blood agar sheep plates (BioTrading, K004P090KP) at 37 °C and in Tryptone soy broth.

## *Prevotella* species

Cultures of *Prevotella bivia* and *Prevotella timonensis* were grown in New York City (NYC) medium and Cooked Meat Medium (CMM) respectively, supplemented with vitamin K1 (final concentration 1 mg/L) and hemin (final concentration 5 mg/L). Bacteria were cultured overnight at 37 °C under anaerobic conditions in a Coy Vinyl Anaerobic Chamber. The $OD_{600}$ was measured on Biowave CO8000 Cell Density Meter.

## *Photobacterium leiognathid*

*Photobacterium leiogonathid* was cultured in marine broth (with and without agar bacteriological) at 30 °C.

## *Streptococcus agalactiae*

*Streptococcus agalactiae* was grown on blood agar sheep plates (BioTrading, K004P090KP) at 37 °C and in Tryptone soy broth.

## SEEL of bacteria

One-step SEEL was performed on bacteria grown in liquid culture (5 x $10^8$ bacteria). The bacteria were washed with buffer and were incubated with SEEL label mix at 37°C for 2 h while rotating. A typical SEEL label mix (50μL) was prepared in medium (PBS/HEPES buffer) with sialyltransferase (1.05 uL of stock 1 mg/mL, or 2 mg/mL in case of pmst3 and Pd2,3-ST), CMP-Sialic acid derivative (50 μM), 0.34 uL BSA (2 mg/mL) and 0.34 μL alkaline phosphatase (1 U/ μL). After SEEL treatment, the bacteria were washed with buffer and prepared for application.

Two-step SEEL was performed similar to one-step SEEL, followed by a click reaction. In case of CuAAC, 100 µL reaction volume contained 100 µM Acetylene-PEG$_4$-biotin, 500 µM CuSO$_4$ and 2.5 mM sodium L-ascorbate. In case of SPAAC 100 µL reaction volume contained 100 µM DBCO-PEG$_4$-biotin. For the fluorophores AF488-alkyne and DBCO-AF488 the concentration was 1 mM.

If the bacteria were heat-inactivated, they were heated at 80°C for 15 min and treated with the described SEEL method.

## MOE with Neu5Az and Neu9Az

Neu5Az and Neu9Az were synthesized according to a published procedure[2,3]. MOE was performed with 6 x 10$^8$ bacteria and these were incubated with indicated concentrations of Neu5Az and Neu9Az. Bacteria were incubated for 6h shaking at 160 rpm and for NTHi at 37 °C, and for *C. jejuni* at 42 °C under microaerophilic conditions. After incubation, the samples were washed 2 x 1 mL (buffer) and clicked via CuAAC or SPAAC, see SEEL of bacteria, washed again and further treated for (glyco)proteins or LOS.

## Western blotting

In case of (glyco)proteins, the samples were lysed and analyzed with a 10% SDS-PAGE gel for which the gel was run for 45-60 min at 150V. For the western blotting, the gel was electroblotted onto a PVDF membrane. The membrane was blocked (5% milk, 30 – 60 min), washed (1% milk, 5 min), stained with anti-biotin-HRP antibody (1 : 20000 in 1% milk), washed (1 % milk, followed by PBS, 5 min each), and treated with ECL western substrate for signal detection.

## LOS preparation and Tris-Tricine gel

Samples were boiled for 5 min and then treated with protease K (10 uL, 20mg/mL) overnight at 55°C. 3x Laemli buffer was added and the samples were analyzed with a 16% Tris-Tricine gel. The gel ran typically for 3 - 4 h at 20 mA and was then further analyzed by Western blotting or silver staining or in-gel fluorescence.

## Silver staining

Silver staining was performed on a 16% Tris-Tricine gel as described previously.[4] Briefly, the gel was fixed (30 min, 40% ethanol, 5% acetic acid), oxidized (5 min,

0.7% sodium periodic acid, 40% ethanol, 5% acetic acid), washed (3 x 5 min in distilled water), stained (distilled water containing 19% 0.1M NaOH, 1.3% >28% ammonium hydroxide, 3,3% 20% w/v silver nitrate), washed (3 x 5 min in distilled water), developed until bands appeared (distilled water containing 0.1% PFA 37% and 0.1% citric acid 100mg/mL), rinsed with distilled water and stopped (7% acetic acid in distilled water).

## In-gel fluorescence

In-gel fluorescence was measured on Amersham imager 600 using the Green channel (520 nm, cy3).

## Flow cytometry

In case of *Neisseria gonorrhoeae*, samples were fixed with 1% paraformaldehyde and 0.5% BSA. If necessary, bacteria were diluted in PBS + 0.05% BSA to not exceed 20000 counts/s in flow cytometry analysis. Flow cytometry was performed on MACSQuant flow cytometer (Miltenyi Biotech) and analysis was done with FlowJo Software (V10).

In case of *Campylobacter jejuni*, samples were fixed with 1-2% paraformaldehyde and flow cytometry was performed on FACS Accuri C6 plus (Becton Dickinson). Analysis was performed with BD Accuri C6 Plus Software or FlowJo Software (V10).

## Serum resistance assay

*N. gonorrhoeae* WT was incubated without and with 20 nmol/mL CMP-Neu5Ac or CMP-Neu5Az. After treatment, *N. gonorrhoeae* was washed and diluted to $10^6$ bacteria in HEPES buffer. Bacteria ($10^4$) were treated with 10% NHS or HI-NHS for 1 h at 37°C. Samples were 25x diluted and 50 μL was plated out on chocolate columbia agar with vitox. Plates were grown at 37°C + 5% $CO_2$ overnight, colony forming units were counted and data analysis was done with prism.

## Growth measurements

Bacteria were treated according to conditions specified in text. Bacteria were diluted to OD=0.05 and the growth was monitored with Synergy HTX multi-mode meter in a hypoxic glove box for 24 h while shaking continuously. Data

was exported and analyzed with excel or prism.

# Supporting figures

A.



B.
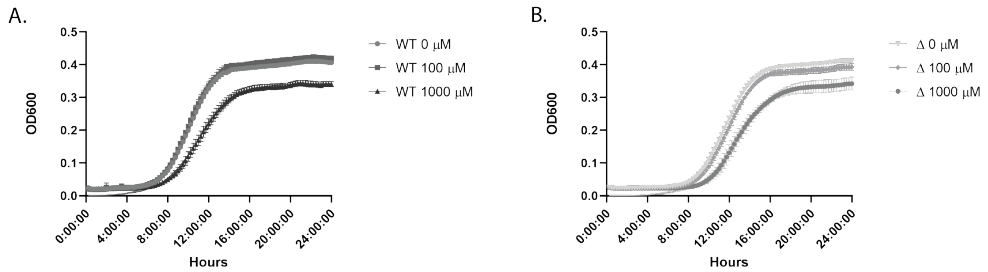


Figure 1. Growth curve of NTHi in presence of different concentrations of Neu5Az. (A) $OD_{600nm}$ of NTHi wildtype (WT). (B) $OD_{600nm}$ of NTHi transporter mutant ($\Delta$).
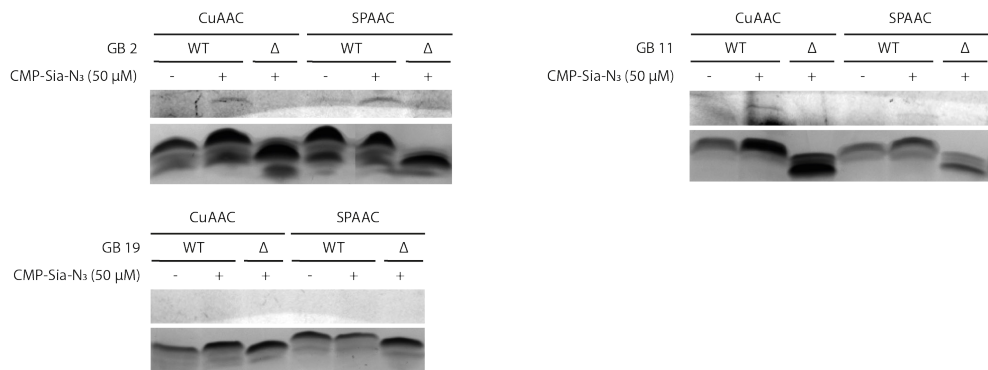


Figure 2. Two-step labeling of three strains of *C. jejuni* LOS (GB2, GB11 and GB19) with CMP-Neu5Az followed by a copper catalyzed click reaction (CuAAC) or a strain-promoted click reaction (SPAAC). The in-gel fluorescence is demonstrated in the top gel and the silver-stained LOS in the bottom gel.
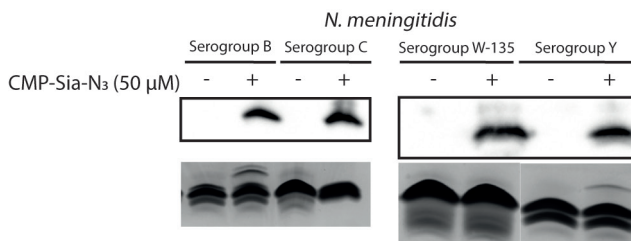


Figure 3. Labeling of the LOS via native sialyltransferases of N. meningitidis strains typed for their capsule. These strains have an unknown immunotype for their LOS.

## Supporting discussion

*Prevotella timonensis* was first isolated in 2007 from a human breast abscess[5] and is associated with bacterial vaginosis. Bacterial vaginosis is characterized by a shift in the microbiota to diverse anaerobic bacteria and a shift in the produced metabolites, such as an increase in Neu5Ac[6]. Although, to our best knowledge, little is known about the use of Neu5Ac by *P. timonensis*, the increased presence of Neu5Ac might indicate that the bacteria use this glycan as a carbon source or as a building block to use in their own glycosylation, as our data seem to indicate.

*Haemophilus aegyptius* or *Haemophilus influenzae* biogroup *aegyptius* was originally described in 1883[7]. A century later, *H. aegyptius* was again described, but as the causative agent of Brazilian purpuric fever[8]. The LOS of *Haemophilus* species can resemble human glycans[9] and this form of glycan mimicry was also demonstrated for some strains of *H. aegyptius*[10]. Sialylation of the LOS has not been described for *H. aegyptius* yet and might indicate a similar function in immune evasion as for nontypeable *Haemophilus influenzae* or *Haemophilus ducreyi*.

*Haemophilus parainfluenzae* is a Gram-negative bacterium in the human respiratory tract. In contrast to NTHi, *H. parainfluenzae* is considered a commensal bacterium. For several strains, the LOS or LPS has been characterized and Neu5Ac was not reported in theses glycan structures[11–13]. One report had scanned for the genetic potential to express an α2,3-sialyltransferase (lic3A), but could not identify this for any of the commensal bacteria[13]. Another report did identify Neu5Ac to be a part of the O-antigen of a strain[12]. Our data indicate that Neu5Ac could be a part of the LOS of this strain of *H. parainfluenzae*.

*Campylobacter upsaliensis* is a common cause of enteritis in dogs and cats, but is also considered an emerging zoonotic pathogen for humans[14–16]. Little is reported about the metabolism of Neu5Ac by this bacterium. One report described that the genes for the sialic acid synthetase or transferase were absent in a clinical isolate from a GBS patient[17]. Another study did report the genomic potential for sialic acid of a different strain of *C. upsaliensis*[18]. Our data indicate that strain DMSZ 5365 is capable of at least transferring CMP-Neu5Ac to its LOS.

*Campylobacter insulaenigrae* was first isolated in 2004 from marine mammals[19] and has been found to also infect humans in rare cases[20,21]. *C. insulaenigrae* is phylogenetically close to other *Campylobacters*, such as *C. jejuni*, *C. coli* and *C. lari*[19]. Our data indicates that *C. insulaenigrae* is capable of utilizing CMP-Neu5Ac

4

to decorate its cell surface glycoconjugates. Future studies might involve elucidating the glycan structures of this bacterium and testing if it resembles human gangliosides, similar to the LOS of *C. jejuni*[9].

*Pasteurella multocida* is a Gram-negative bacterium that mainly infects animals. The LOS/LPS of serovar 3 strains Pm70 and P1059 are described to contain sialic acid[22], which is thought to shield the bacteria from the immune system[23]. *P. multocida* has a known sialyltransferase, Pmst, that also acts as a sialidase[24]. In our experiment, *P. multocida* showed modest labeling of the LOS by native sialyltransferases. The strain tested (DSM 5281) does belong to serovar 3, but the strain number is unknown, which might explain the level of labeling. Alternatively, the experimental conditions were not optimal for sialyltransferase activity or the enzymes exhibited sialidase activity.

*Pasteurella dagmatis* is an animal pathogen that can infect humans after contact with animals[25]. *P. dagmatis* has a known sialyltransferases, PdST, that catalyze the transfer of CMP-Neu5Ac onto galactose with an α2,3-linkage[26]. In addition, this enzyme has a different mode of action and can hydrolyze CMP-Neu5Ac in absence of an acceptor substrate[27]. Moreover, this enzyme was mutated so it would create an α2,6-linkage instead of an α2,3-linkage, which can be used in chemoenzymatic synthesis for instance[27]. In agreement with the suspected sialyltransferase activity, *P. dagmatis* shows labeling of the LOS/LPS.

*Neisseria lactamica* is considered a commensal bacteria in the upper respiratory tract of humans[28]. The LOS of this Neisserial species has been described to react with antibodies against the LOS of *N. gonorrhoeae* or *N. meningitidis*, or against human glycosphingolipids[29,30]. Therefore, the LOS of *N. lactamica* could display glycan mimicry. Our data indicate that this strain is capable of utilizing CMP-Neu5Az and incorporate sialic acid in the LOS, similar to *N. meningitidis* and *N. gonorrhoeae*. It would be worthwhile to elucidate the terminal glycans structure of *N. lactamica* and identify if it is similar to other Neisserial species.

*Neisseria meningitidis* uses Neu5Ac in its LOS or capsule[31]. In some cases, the glycoconjugates of *N. meningitidis* are glycan mimicking[9]. In our screen, *N. meningitidis* strains that were typed for their LOS were scanned (L1-12). In literature, the terminal glycans of L1-L6 were described to contain sialic acid[31]. Analysis of the LOS indicated that all strains, except L6, L11 and L12, showed incorporation of CMP-Neu5Az after a strain-promoted click reaction with a

biotin reporter. This is in contrast to the previously reported glycan structures. However, for all strains the presence of the Lst gene could be confirmed (Table 1), which shows the already reported genomic potential to sialylate the LOS[32]. An alternative hypothesis is that the characterization of the strains could not identify Neu5Ac on the LOS, because it was too little during analysis or possibly cleaved off during an acidic workup that is often employed for LPS isolation. In addition to LOS typed strains, *N. meningitidis* that were typed for their CPS were analyzed, but for these the LOS types were uncharacterized. These strains did test positive for modification of their LOS which suggests that these strains have an LOS type that uses sialic acid (SI Figure 3).

# Chapter 4

# Supplemental references

(1) Sun, T.; Yu, S. H.; Zhao, P.; Meng, L.; Moremen, K. W.; Wells, L.; Steet, R.; Boons, G. J. One-Step Selective Exoenzymatic Labeling (SEEL) Strategy for the Biotinylation and Identification of Glycoproteins of Living Cells. *J. Am. Chem. Soc.* **2016**, *138*, 11575–11582.

(2) Luchansky, S. J.; Goon, S.; Bertozzi, C. R. Expanding the Diversity of Unnatural Cell-Surface Sialic Acids. *ChemBioChem* **2004**, *5*, 371–374.

(3) Han, S.; Collins, B. E.; Bengtson, P.; Paulson, J. C. Homomultimeric Complexes of CD22 in B Cells Revealed by Protein-Glycan Cross-Linking. *Nat. Chem. Biol.* **2005**, *1*, 93–97.

(4) Tsai, C. M.; Frasch, C. E. A Sensitive Silver Stain for Detecting Lipopolysaccharides in Polyacrylamide Gels. *Anal. Biochem.* **1982**, *119*, 115–119.

(5) Glazunova, O. O.; Launay, T.; Raoult, D.; Roux, V. *Prevotella Timonensis* Sp. Nov., Isolated from a Human Breast Abscess. *Int. J. Syst. Evol. Microbiol.* **2007**, *57*, 883–886.

(6) Srinivasan, S.; Morgan, M. T.; Fiedler, T. L.; Djukovic, D.; Hoffman, N. G.; Raftery, D.; Marrazzo, J. M.; Fredricks, D. N. Metabolic Signatures of Bacterial Vaginosis. *MBio* **2015**, *6*.

(7) Koch, R. *Bericht über die Thätigkeit der zur Erforschung der Cholera im Jahre 1883 nach Egypten und Indien entsandten Kommission*, NV-Onlin.; Gaffky, G., Ed.; Arbeiten aus dem Kaiserlichen Gesundheitsamte; Springer Berlin Heidelberg: Berlin, Heidelberg, 1887.

(8) Rubin, L. G.; St Geme, J. W. Role of Lipooligosaccharide in Virulence of the Brazilian Purpuric Fever Clone of *Haemophilus Influenzae* Biogroup Aegyptius for Infant Rats. *Infect. Immun.* **1993**, *61*, 650–655.

(9) de Jong, H.; Wösten, M. M. S. M.; Wennekes, T. Sweet Impersonators: Molecular Mimicry of Host Glycans by Bacteria. *Glycobiology* **2021**.

(10) Mandrell, R. E.; McLaughlin, R.; Kwaik, Y. A.; Lesse, A.; Yamasaki, R.; Gibson, B.; Spinola, S. M.; Apicella, M. A. Lipooligosaccharides (LOS) of Some *Haemophilus* Species Mimic Human Glycosphingolipids, and Some LOS Are Sialylated. *Infect. Immun.* **1992**, *60*, 1322–1328.

(11) Pollard, A.; St. Michael, F.; Connor, L.; Nichols, W.; Cox, A. Structural Characterization of *Haemophilus Parainfluenzae* Lipooligosaccharide and Elucidation of Its Role in Adherence Using an Outer Core Mutant. *Can. J. Microbiol.* **2008**, *54*, 906–917.

(12) Vitiazeva, V.; Twelkmeyer, B.; Young, R.; Hood, D. W.; Schweda, E. K. H. Structural Studies of the Lipopolysaccharide from *Haemophilus Parainfluenzae* Strain 20. *Carbohydr. Res.* **2011**, *346*, 2228–2236.

(13) Young, R. E. B.; Hood, D. W. *Haemophilus Parainfluenzae* Has a Limited Core Lipopolysaccharide Repertoire with No Phase Variation. *Glycoconj. J.* **2013**, *30*, 561–576.

(14) Acke, E. Campylobacteriosis in Dogs and Cats: A Review. *N. Z. Vet. J.* **2018**, *66*, 221–228.

(15) Bourke, B.; Chan, V. L.; Sherman, P. *Campylobacter Upsaliensis*: Waiting in the Wings. *Clin. Microbiol. Rev.* **1998**, *11*, 440–449.

(16) Man, S. M. The Clinical Importance of Emerging *Campylobacter* Species. *Nat. Rev. Gastroenterol. Hepatol.* **2011**, *8*, 669–685.

(17) Fouts, D. E.; Mongodin, E. F.; Mandrell, R. E.; Miller, W. G.; Rasko, D. A.; Ravel, J.; Brinkac, L. M.; Deboy, R. T.; Parker, C. T.; Daugherty, S. C.; Dodson, R. J.; Scott Durkin, A.; Madupu, R.; Sullivan, S. A.; Shetty, J. U.; Ayodeji, M. A.; Shvartsbeyn, A.; Schatz, M. C.; Badger, J. H.; Fraser, C. M.; Nelson, K. E. Major Structural Differences and Novel Potential Virulence Mechanisms from the Genomes of Multiple *Campylobacter* Species. *PLoS Biol.* **2005**, *3*.

(18) Richards, V. P.; Lefébure, T.; Pavinski Bitar, P. D.; Stanhope, M. J. Comparative Characterization of the Virulence Gene Clusters (Lipooligosaccharide [LOS] and Capsular Polysaccharide [CPS]) for *Campylobacter Coli*, *Campylobacter Jejuni* Subsp. *Jejuni* and Related *Campylobacter* Species. *Infect. Genet. Evol.* **2013**, *14*, 200–213.

(19) Foster, G.; Holmes, B.; Steigerwalt, A. G.; Lawson, P. A.; Thorne, P.; Byrer, D. E.; Ross, H. M.; Xerry, J.; Thompson, P. M.; Collins, M. D. *Campylobacter Insulaenigrae* Sp. Nov., Isolated from Marine Mammals. *Int. J. Syst. Evol. Microbiol.* **2004**, *54*, 2369–2373.

(20) Chua, K.; Gürtler, V.; Montgomery, J.; Fraenkel, M.; Mayall, B. C.; Grayson, M. L. *Campylobacter Insulaenigrae* Causing Septicaemia and Enteritis. *J. Med. Microbiol.* **2007**, *56*, 1565–1567.

(21) Kyotani, M.; Kenzaka, T.; Akita, H.; Arakawa, S. *Campylobacter Insulaenigrae* Bacteremia with Meningitis: A Case Report. *BMC Infect. Dis.* **2021**, *21*, 1–5.

(22) Harper, M.; Cox, A. D.; Adler, B.; Boyce, J. D. *Pasteurella Multocida* Lipopolysaccharide: The Long and the Short of It. *Vet. Microbiol.* **2011**, *153*, 109–115.

(23) Tatum, F. M.; Tabatabai, L. B.; Briggs, R. E. Sialic Acid Uptake Is Necessary for Virulence of *Pasteurella Multocida* in Turkeys. *Microb. Pathog.* **2009**, *46*, 337–344.

(24) Yu, H.; Chokhawala, H.; Karpel, R.; Yu, H.; Wu, B.; Zhang, J.; Zhang, Y.; Jia, Q.; Chen, X. A Multifunctional *Pasteurella Multocida* Sialyltransferase: A Powerful Tool for the Synthesis of Sialoside Libraries. *J. Am. Chem.*

*Soc.* **2005**, *127*, 17618–17619.

(25) Holst, E.; Rollof, J.; Larsson, L.; Nielsen, J. P. Characterization and Distribution of *Pasteurella* Species Recovered from Infected Humans. *J. Clin. Microbiol.* **1992**, *30*, 2984–2987.

(26) Schmölzer, K.; Ribitsch, D.; Czabany, T.; Luley-Goedl, C.; Kokot, D.; Lyskowski, A.; Zitzenbacher, S.; Schwab, H.; Nidetzky, B. Characterization of a Multifunctional α2,3-Sialyltransferase from *Pasteurella Dagmatis*. *Glycobiology* **2013**, *23*, 1293–1304.

(27) Schmölzer, K.; Luley-Goedl, C.; Czabany, T.; Ribitsch, D.; Schwab, H.; Weber, H.; Nidetzky, B. Mechanistic Study of CMP-Neu5Ac Hydrolysis by α2,3-Sialyltransferase from *Pasteurella Dagmatis*. *FEBS Lett.* **2014**, *588*, 2978–2984.

(28) Dorey, R. B.; Theodosiou, A. A.; Read, R. C.; Jones, C. E. The Nonpathogenic Commensal *Neisseria*: Friends and Foes in Infectious Disease. *Curr. Opin. Infect. Dis.* **2019**, *32*, 490–496.

(29) Arking, D.; Tong, Y.; Stein, D. C. Analysis of Lipooligosaccharide Biosynthesis in the *Neisseriaceae*. *J. Bacteriol.* **2001**, *183*, 934–941.

(30) Braun, J. M.; Beuth, J.; Blackwell, C. C.; Giersen, S.; Higgins, P. G.; Tzanakaki, G.; Unverhau, H.; Weir, D. M. *Neisseria Meningitidis*, *Neisseria Lactamica* and *Moraxella Catarrhalis* Share Cross-Reactive Carbohydrate Antigens. *Vaccine* **2004**, *22*, 898–908.

(31) Mubaiwa, T. D.; Semchenko, E. A.; Hartley-Tassell, L. E.; Day, C. J.; Jennings, M. P.; Seib, K. L. The Sweet Side of the Pathogenic *Neisseria*: The Role of Glycan Interactions in Colonisation and Disease. *Pathogens and Disease*. 2017, p ftx063.

(32) Tsang, R. S. W.; Law, D. K. S.; Tsai, C.-M.; Ng, L.-K. Detection of the Lst Gene in Different Serogroups and LOS Immunotypes of *Neisseria Meningitidis*. *FEMS Microbiol. Lett.* **2001**, *199*, 203–206.

4

# Chapter 5

# Structure-activity relationship of 2,4-D in somatic embryogenesis in *Arabidopsis thaliana*

2,4-dichlorophenoxyacetic acid (2,4-D) is a synthetic analogue of the natural plant hormone auxin that is widely used to kill weeds. At the same time, it is commonly used in many *in vitro* plant regeneration systems such as somatic embryogenesis (SE) and organogenesis. The effectiveness of 2,4-D in inducing SE, compared to the natural auxin indole-3-acetic acid (IAA), has been attributed to the stress triggered by this compound rather than its auxin activity. However, this hypothesis has never been thoroughly tested. Here we used a library of 40 2,4-D analogues to test the structure-activity relationship with respect to auxin activity and the capacity to induce SE in *Arabidopsis thaliana*. Based on root growth inhibition and auxin response reporter expression the 2,4-D analogues were classified into different groups, ranging from very active to not active. This showed that the type and number of substituents at different positions on the aromatic ring of the 2,4-D analogues, or different substituents at the alpha position of the carboxylate, have a significant effect on auxin activity. Molecular dynamics simulations showed that the auxin activity of the 2,4-D analogues correlated well with their TIR1-Aux/IAA coreceptor binding characteristics (enthalpy and distance). In the process, we identified two 2,4-D analogues as efficient inducers of adventitious root formation and several possible anti-auxins. Finally, we observed a strong correlation between auxin activity and SE induction efficiency, indicating that the stress-related effects by 2,4-D or its analogues are down-stream of auxin signaling.

*Hanna de Jong\*, Omid Karami\*, Victor J. Somovilla, Beatriz Villanueva Acosta, Aldo Bryan Sugiarta, Tom Wennekes, Remko Offringa*

*\* These authors contributed equally to this work*

# Chapter 5

## Introduction

The plant hormone auxin plays a central role in the development of plants. In the 1930s, the structure of the natural auxin indole-3-acetic acid (IAA) was first described[1]. A few years later, during WWII, 2,4-dichlorophenoxyacetic acid (2,4-D) was discovered as a synthetic auxin analogue that can be used as a herbicide, targeting dicots. Today, it is still broadly used as such in gardens and in agriculture[2]. Apart from the cell elongation promoting effect, which led to its discovery as auxin analogue, 2,4-D acts differently in various physiological and molecular assays compared to the natural auxin IAA[2–4].

Binding of IAA to its receptors triggers a transcriptional response. The classical nuclear IAA signaling pathway relies on the degradation of the transcriptional AUXIN/INDOLE-3-ACETIC ACID (Aux/IAA) repressors, leading to expression of auxin responsive genes[5]. The degradation of Aux/IAAs is initiated by binding of auxin to the F-Box proteins, TRANSPORT INHIBITOR RESISTANT1 (TIR1) or AUXIN SIGNALING F-BOX 1-3 (AFB1-3). Auxin acts as a molecular glue, allowing the TIR1/AFBs to recruit the N-terminal domain II of Aux/IAAs[6]. TIR1/AFBs are part of a Skp1-Cullin-F-box (SCF) E3 ubiquitin ligase complex, which after recruiting Aux/IAAs marks these proteins for degradation by the 26S proteasome[7].

The synthetic auxin 2,4-D elicits a dual response in plants, as it acts as an auxin and also induces stress[8,9]. Like IAA, 2,4-D acts through the TIR1/AFB auxin-mediated signaling pathway[3,4,8]. At high concentrations, 2,4-D acts as herbicide and selectively kills dicot weeds. The herbicidal activity of 2,4-D can be attributed to several effects, including the altering of cell wall plasticity and the increase of ethylene levels[8]. The overproduction of ethylene triggers the increased production of abscisic acid (ABA), which contributes to stomatal closure and thus eventually to plant death[8,9]. 2,4-D also induces the production of reactive oxygen species (ROS) that are very toxic to the plant[9].

Besides its use as an herbicide, 2,4-D is widely used for biological experiments to induce auxin responses and for *in vitro* regeneration systems. Our interest in 2,4-D lies in its ability to efficiently induce somatic embryogenesis (SE) at non-herbicidal concentrations. SE is a unique developmental process in which differentiated somatic cells can acquire totipotency and are 'reprogrammed' to form new 'somatic' embryos[4]. SE is a powerful tool in plant biotechnology used for clonal propagation, genetic transformation and somatic hybridization.

It prevents somaclonal variation that is often observed in plant tissue culture, and somatic embryos can be easily cryopreserved[10,11]. Although SE can be induced by IAA, other synthetic auxins, or stress[9,12], in many cases 2,4-D is most effective and it is therefore widely used for SE in many plant species. Different observations have led to the suggestion that signaling pathways activated by abiotic stress treatments and 2,4-D treatment converge to regulate a common downstream pathway[12]. 2,4-D induced SE is accompanied by upregulation of stress-related genes[13–15], and a number of stress-related transcription factors have been shown to influence the progression of SE[15–17]. This suggests that stress induced by non-herbicidal concentrations of 2,4-D[18] contributes significantly to its effectiveness as SE inducer. To further understand 2,4-D-induced SE, we established a structure-activity relationship (SAR) of 2,4-D with respect to its activity as auxin analogue and its capacity to induce SE.

Previous studies have already discovered several chemical analogues of 2,4-D with auxin-like activity. Unfortunately, however, several different bioassays have been used in these studies[19], making a direct comparison difficult. In addition, a number of possible 2,4-D analogues that based on these previous studies would be interesting candidates to assess in a structure-activity relationship (SAR) have not been tested yet. Here, we therefore formed a rationally designed library of 40 2,4-D analogues and screened them for auxin activity and SE induction in *Arabidopsis thaliana* (Arabidopsis). Based on root growth inhibition, root hair and later root induction, and auxin response reporter expression we classified compounds as very active, active, weakly active or inactive auxins, or even having anti-auxin activity. From the SAR, we concluded that an electron withdrawing group, such as a halogen, at the 4-position of the aromatic ring is important for auxin activity. Moreover, a small substitution at the carboxylate chain, a methyl or ethyl group, is tolerated. Molecular dynamics simulations indicated that the classification of the 2,4-D analogues reflected the binding characteristics (enthalpy and distance) to TIR1. SE induction capacity clearly correlated with auxin activity, indicating that the stress response induced by 2,4-D treatment and required for SE induction is downstream of auxin signaling.

## Results

### Assembly of the 2,4-D analogue library

We aimed to screen a library of 2,4-D analogues for auxin activity and SE induction in *Arabidopsis thaliana* and thereby gain insight into the SAR for this

process. To this end, we first performed a literature survey of key past auxin studies in order to identify known 2,4-D analogues for our library[8,19–30]. This provided an overview of auxin activities and the binding efficiencies of some 2,4-D-like compounds to the auxin co-receptors and gave some context on the SAR of auxin-like compounds. In addition, it offered insight into the difficulties encountered in the past in establishing a SAR, such as non-standardized assays and the use of impure compounds. However, a rational SAR of 2,4-D could not be deduced from these past studies. To achieve a more complete SAR of 2,4-D with respect to auxin activity and the induction of SE, we assembled our own library with TIR1 as the target co-receptor, 2,4-D as a lead compound, analogue hits from the literature survey and several additional structures to achieve a more complete and rational coverage of possible 2,4-D analogues. In general, analogues of the lead compound 2,4-D, either commercially available or synthesized (Supplementary Information), were selected based on modifications at two positions: the type and number of substituents on the aromatic ring or on the alpha position of the carboxylate side chain (Figure 1A). A complete overview of the library is provided in Figure 2.

## Assessing auxin activity of the 2,4-D analogues based on root growth inhibition

Exogenous auxin inhibits primary root elongation in *Arabidopsis* in a concentration-dependent manner[31]. As a first assay to establish auxin activity of the 2,4-D analogues in our library (Figure 2), we tested the inhibitory effect of low (50 nM) and high (5 μM) concentrations on the primary root growth of *Arabidopsis* seedlings (Figure 1). Similar to 2,4-D, the compounds 4-Cl, 4-Br, 2-Cl-4-F, 2,4-DB, MCPA and Mecoprop inhibited the primary root elongation efficiently, both at low and high concentrations (Figure 1B). Hence, we classified these compounds in one group called very active auxins. Interestingly, several compounds (4-F, 4-I, 2,5-D, 3,4-D, 2,4-Br, 2-F-4-Cl, 2-F-4-Br, 2,4,5-T, 2,4-DiB and 2,4-DP, MCPB) showed no inhibitory effect at 50 nM (Figure 1B), whereas they exhibited a strong inhibitory effect at 5 μM (Figure 1B). These compounds were classified as active auxins. By contrast, 2-Cl, 2-I, 3-Cl, 2,6-D, 2,4-F, $2\text{-NO}_2$-4-Cl, 2,3,4-T, 2,3,4-F, 2,4,6-T, 2,4,6-F, 2,4-DnP and 2,4-DnB had only a weak effect on root growth at 5 μM (Figure 1B) and were therefore classified as weak auxins. The remaining compounds PHAA, 2-F, 3-Me, 2,3-D, 3,5-M, 3-OMe, $4\text{-NO}_2$, 3,5-Me, 3,5-D, 2-CL-4-Formyl, 2,4,6-T, and 2,4-DnP had no obvious inhibitory effect on root length, even at 5 μM (Figure 1B), suggesting that these compounds have either very weak or no auxin activity (Figure 1B). The

above data indicate that 2,4-D analogues inhibit root growth in both dose and structure-dependent manner.
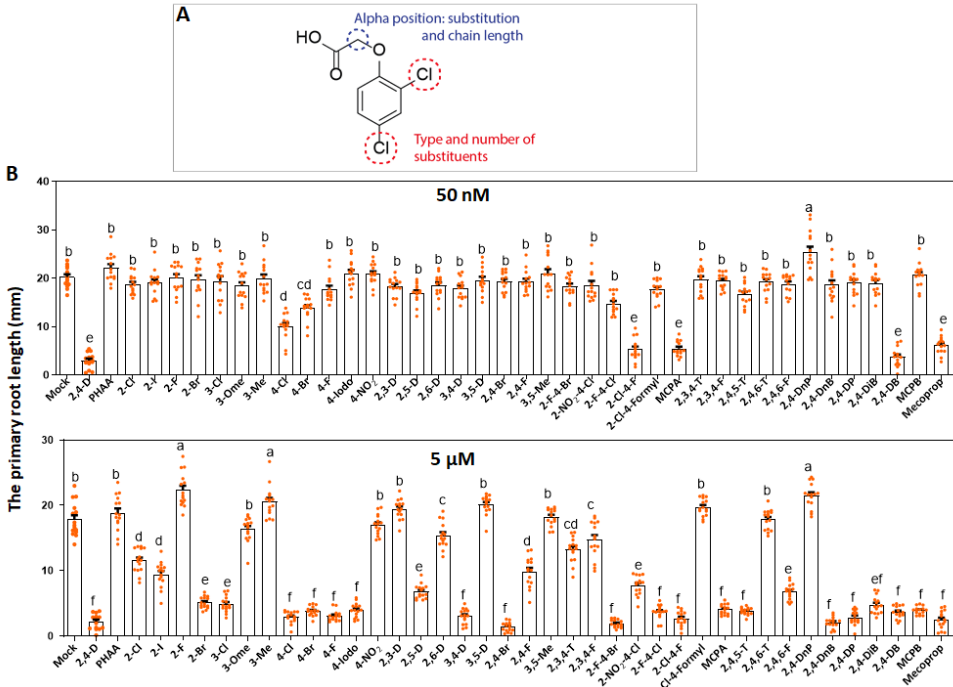


Figure 1. Effect of 2,4-D structure modifications on primary root growth inhibition in Arabidopsis. (A) Structure of 2,4-D with the different modifications indicated that compose the 2,4-D analogue library (for an overview of the analogues and their abbreviations, see *Figure 2*). (B) The primary root growth inhibition of seedlings grown in the presence of 50 nM (upper graph) and 5 µM (lower graph) of 2,4-D and 2,4-D analogues. Seedlings were first grown for 6 days on compound free-medium and subsequently transferred on medium with 2,4-D or a 2,4-D analogue for 3 days. Dots indicate the root length (mm) (n=15 biological replicates), bars indicate the mean and error bars the s.e.m. and different letters indicate statistically significant differences (P < 0.01) as determined by a one-way analysis of variance with Tukey's honest significant difference post hoc test.
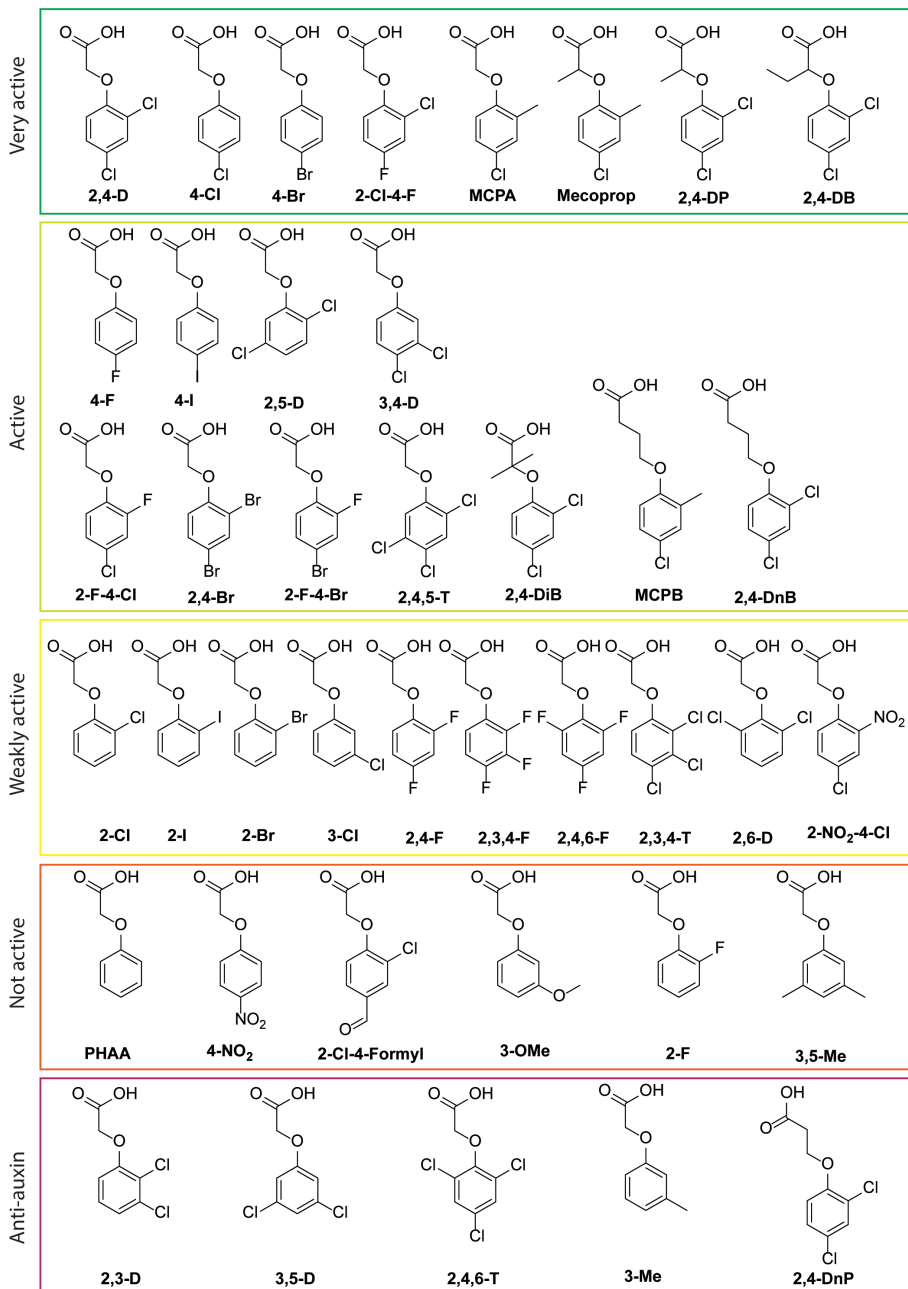
Figure 2. The library of 40 2,4-D analogues used in this study. The 2,4-D analogues were classified as very active, active or weakly active auxins, or not active, or anti-auxins based on their effect on primary root growth, lateral root development, or the expression of the auxin response reporters pDR5:GUS or R2D2.

## The auxin response induced by 2,4-D analogues corresponds to their root growth inhibition activity

The *pDR5* promoter is a synthetic, generic auxin-responsive promoter that has been extensively used to visualize the cellular auxin response in *Arabidopsis* tissues[32]. To determine whether the root growth inhibition by 2,4-D analogues correlated with a molecular auxin response, we used an *Arabidopsis* line containing *pDR5* fused to a glucuronidase (GUS) reporter (*pDR5:GUS*). 2,4-D itself induced *pDR5* efficiently, as observed by a completely blue root following histochemical staining for GUS activity (Figure 3A). Because 2,4-D induced an especially strong *pDR5* activity in the cell division area of the root tip, we next quantified the effect of the 2,4-D analogues on DR5 activity in this part of the root tip (Figure 3B).

As expected, the compounds classified as very active and active auxins (Figure 2) all strongly promoted *pDR5* activity in the cell elongation zone of *Arabidopsis* roots (Figure 3A, B). All very active auxin compounds strongly induced *pDR5* activity, in line with their root growth inhibition activity, whereas the induction of *pDR5:GUS* expression varied more for the active auxins (Figure 3A, B). The compounds that had only a slight effect on root growth and were therefore classified as weak auxins, also only slightly promoted expression of the *pDR5:GUS* reporter (Figure 3A, B). Interestingly, 3-Cl induced a stronger auxin response along the differentiation zone of the root (Figure 3A) compared to the other compounds classified as weak auxins. The remaining library members that did not lead to inhibition of root growth, also did not promote *pDR5* activity (Figure 3A, B). These findings support our conclusion that the ability of the 2,4-D analogues to inhibit root growth correlates well with their auxin response induction.

To further confirm the auxin activity of the 2,4-D analogues, we used the nuclear auxin input reporter R2D2, which acts as a proxy for the cellular sensing of auxin[33]. The R2D2 reporter consists of two parts, an auxin-degradable DII domain fused to the yellow fluorescent nVENUS that is rapidly degraded when auxin concentrations increase as auxin sensor and an auxin-nondegradable DII domain (mDII) fused to the orange fluorescent nTdTOMATO as expression control[33]. In accordance with observations based on *pDR5* activity, we detected strong down-regulation of the DII-nVENUS signal in the cell elongation zone of *Arabidopsis* roots treated with all very active and most of the active 2,4-D analogues (2,5-D, 3,4-D, 2,4-Br, 2-F-4-Br, 2,4-DP, and MCPB) and two weakly

active compounds (2-Br and 2-NO2-4-Cl) (Figure S1). The other active and weakly active compounds elicited only moderate downregulation of DII-Venus (Figure S1). The compounds that proved inactive as an auxin, as these did not inhibit root growth nor promoted *pDR5* activity, also did not lead to a detectable reduction of the DII-nVENUS signal in the cell elongation zone of roots (Figure S1). These results again supported the overall classification of the auxin activity for our library of 2,4-D analogues (Figure 2).
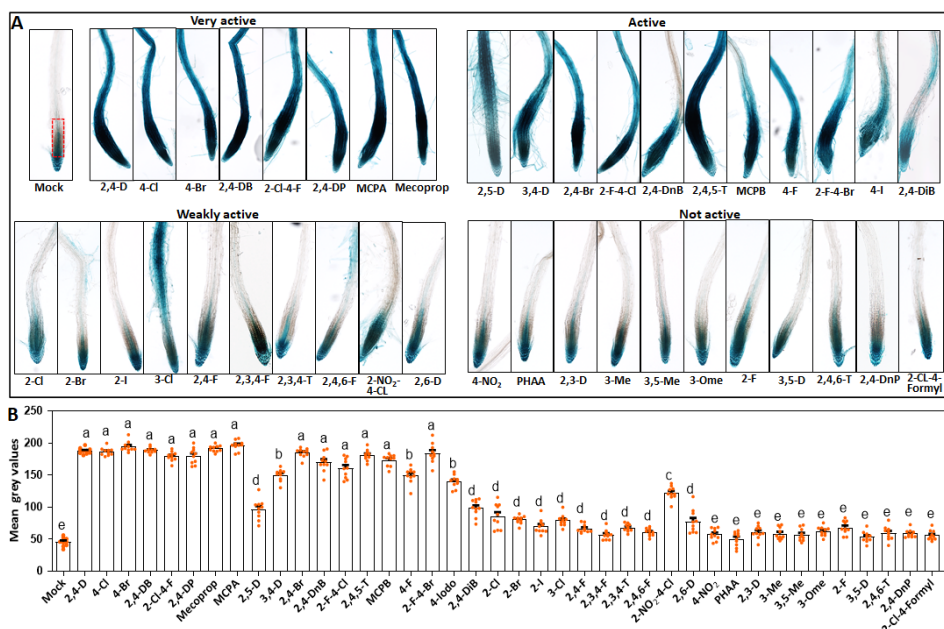


Figure 3. Auxin response induced by 2,4-D analogues (A) Expression of the auxin responsive pDR5:GUS reporter in Arabidopsis seedling roots treated for 24 hours with 5 μM of each 2,4-D analogue. For an overview of the 2,4-D analogues and their abbreviations, see Figure 2. (B) Quantification of pDR5:GUS activity in the root meristems of Arabidopsis seedlings (region indicated with dotted red line in mock). Dots indicate the individual measurement of GUS activity staining per each treatment (n= 10 biologically independent seedlings per treatment, bars indicate the mean of each treatment and error bars the s.e.m. and different letters indicate statistically significant differences (P < 0.01) as determined by a one-way analysis of variance with Tukey's honest significant difference post hoc test. In A and B, seedlings were germinated for 6 days on compound-free medium, and for treatment seedlings were transferred to medium contain 5 μM of a 2,4-D analogue and incubated for 24 hours.

## Specific 2,4-D analogues as tools to modulate root system architecture

When establishing the root growth inhibition of our 2,4-D analogues, we observed that several compounds had a unique effect on the root system architecture (RSA). The RSA of a plant describes the organization of the primary, lateral and adventitious roots. This includes root hairs that increase the surface

area and thus promote the uptake of water and nutrients[34]. Auxin treatment is known to change the RSA by inducing lateral or adventitious roots[35,36] and to positively influence the formation of root hairs in a dose-dependent manner[37]. An adventitious root (AR) refers to a plant root that forms from any non-root tissue, commonly in response to treatment with the natural auxin indole-3-butyric acid (IBA)[36,38].

All 2,4-D analogues that were designated as very active, active and weakly active positively increased the number of root hairs in Arabidopsis seedling roots (not shown). Specifically, we observed a strong effect on root hair formation on root tissues treated with 5 µM of 3-Cl, 3,4-D, 2,4-DP or Mecoprop. We found that the root hair formation was dose-dependent for all four analogues (Figure S2). Interestingly, treatment with 0,5 or 1 µM of the weakly active 3-Cl did not lead to a strong inhibition of root growth, like with the other three active or very active compounds (Figure S2), whereas it still strongly enhanced root hair development. This result reflects the weaker auxin response induced by 3-Cl in the root meristem, leading to reduced root growth inhibition, while it still induces a relatively strong auxin response in the differentiation zone, resulting in ectopic root hair formation (Figure 2A).

We also observed that 3-Cl and 2,5-D significantly promoted the number of lateral roots, whereas 2,4-D and other analogues with strong auxin activity initially induced many lateral root meristems, but these meristems quickly deteriorated into amorphous callus (Figure 4A). As lateral root and adventitious root (AR) induction are highly linked, we tested the capacity of different concentrations of 2,4-D, IBA, 3-Cl and 2,5-D in AR induction from hypocotyls of dark grown *Arabidopsis* seedlings. AR induction is a crucial process in clonal crop propagation by cuttings or shoot regeneration, and is well-known to be promoted in many plant species by the natural auxin indole-3-butyric acid (IBA). As with lateral roots and in line with previous observations[36], treatment with 2,4-D produced a low number of ARs at low µM concentrations and only undesired callus at higher µM concentrations (Figure 4B, C). Treatment with 3-Cl and 2,5-D, however, efficiently induced ARs at 5 µM (Figure 4B, C). In addition, the number of ARs induced by 3-Cl and 2,5-D was significantly higher compared to IBA treatment at a similar concentration and also compared to 57 µM IAA or 2 µM of the synthetic auxin 1-naphthaleneacetic acid (NAA) (Figure 4D), treatments that have previously been shown to efficiently induces ARs from of *Arabidopsis* hypocotyls[36]. With these results, we can conclude that 3-Cl and

5

2,5-D are excellent candidates for inducing AR formation in Arabidopsis, and that despite their structural similarity with 2,4-D, they show a unique biological activity. This probably relates to their mild and specific activity as auxin analogue, as reflected by the specific expression pattern of the *pDR5:GUS* reporter following treatment with these compounds.

Lateral root formation can be repressed by anti-auxins[35,39]. Several of the 2,4-D analogues for which we observed no clear effect on root length or *pDR5* activity, namely 3-Me, 2,3-D, 3,5-D, 2,4,6-T, and 2,4-DnP, did significantly inhibit lateral root formation, with some having a stronger effect (Figure 5A, B). In order to investigate the lateral root inhibition-responsiveness to these potential anti-auxins, we examined the effect of short term treatment (1 day) on the different developmental stages of lateral root formation, using the *pDR5:GUS* reporter activity as marker (Figure 5C)[40]. Histochemical analysis revealed that *pDR5:GUS* is expressed in the presence of anti-auxins in all lateral root stages, the lateral root initiation (stage 1), lateral root development (stage 2), lateral root emergence (stage 3) and lateral root outgrowth (stage 4) (not shown). This is in line with the observation that treatment with these compounds did not lead to a reduced expression of the *pDR5:GUS* reporter in the main root tip (Figure 2). However, these 2,4-D derived anti-auxin candidates did reduce the number of stage 1 primordia, had no effect on the number of stage 3 primordia, whereas they had a differential effect on the number of stage 2 primordia and stage 4 lateral roots (Figure 5C). The effect of these compounds on lateral root formation might indeed be caused by their activity as anti-auxins, however we cannot exclude that they affect other processes, such as auxin transport.

Based on the combined results we now classified all tested 2,4-D analogues from our library into five groups, namely very active, active, weakly active auxin analogues and not active and anti-auxin compounds (Figure 2).
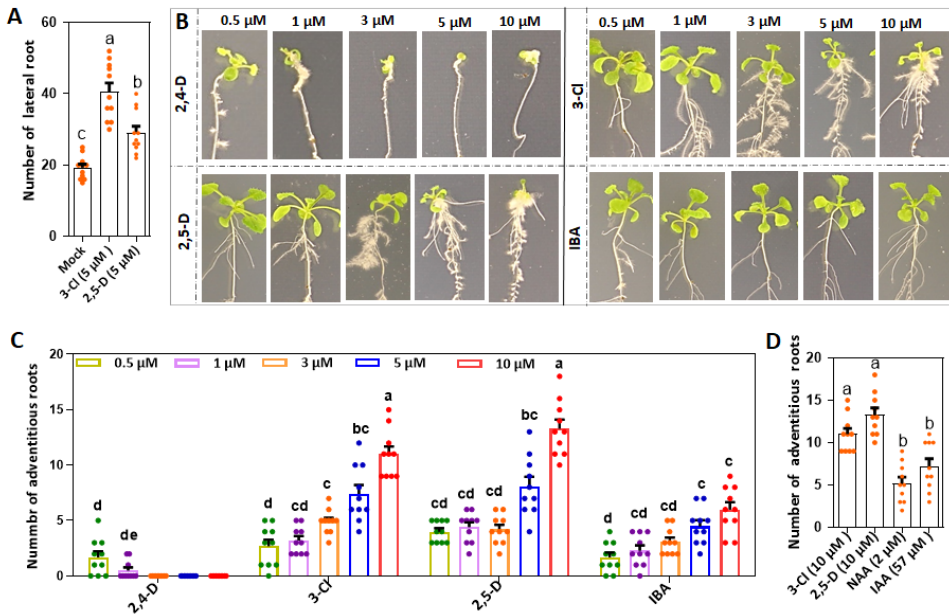
Figure 4. Efficient adventitious root induction on etiolated Arabidopsis hypocotyls by specific 2,4-D analogues. (A) Number of lateral roots formed in Arabidopsis seedlings treated with mock or 5 µM of 3-Cl and 2,5-D. Seedlings were germinated for 6 days on compound-free medium, and subsequently transferred for treatment to medium containing 5 µM of a 2,4-D analogue and incubated for 10 days. (B) The phenotype adventitious roots induced on hypocotyls of 6-day-old etiolated Arabidopsis seedlings by transferring them to medium with different concentrations of 2,4-D, 2,5-D, 3-Cl or IBA. (C) Quantification of the number of adventitious roots induced on etiolated Arabidopsis hypocotyls by 2,4-D, 2,5-D, 3-Cl or IBA (see B). (D) Comparison of the number of adventitious roots induced from etiolated Arabidopsis hypocotyls by 10 µM of 2,5-D, 10 µM of 3-Cl, 2 µM of NAA or 57 µM of IAA. In A, C and D dots indicate the number lateral roots per main root in A and the number adventitious roots per hypocotyl in C and D ( A. n=15; C. n=10; D. n=10), the bars indicate the mean value, and the error bars the s.e.m.. The letters above the bars indicate the significant difference (p < 0.01) determined by a one-way ANOVA with Tukey's honest significant difference post hoc test. In B-D, Arabidopsis seedlings were initially cultured for 6 days on hormone-free medium in the dark, and subsequently the seedlings were transferred on medium containing the indicated compound.

5
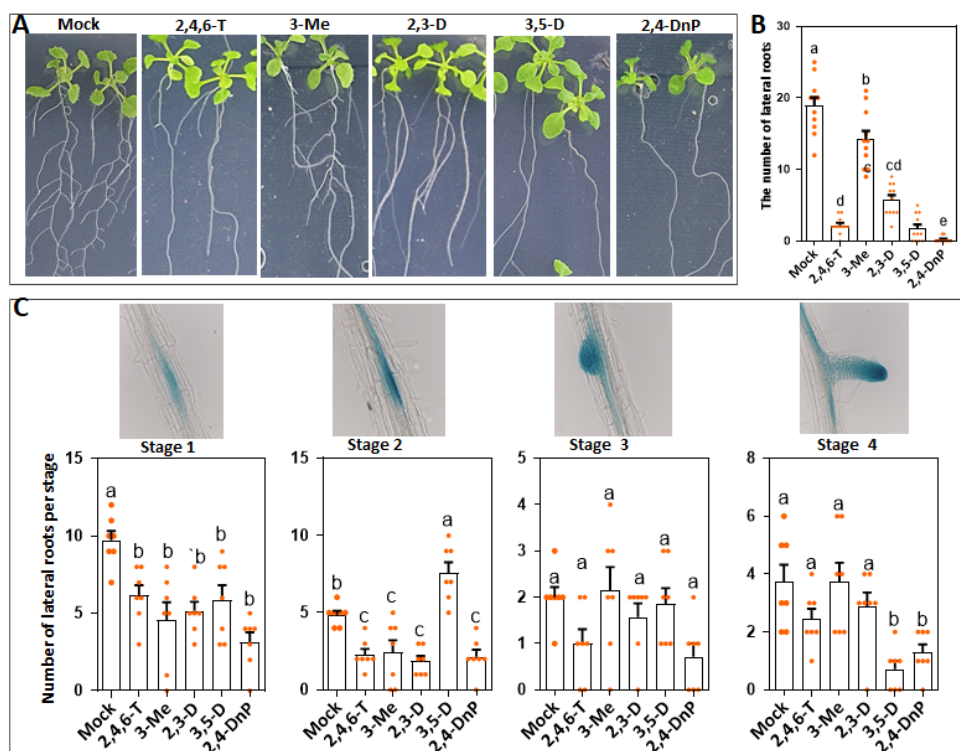
Figure 5. Several 2,4-D analogues inhibit lateral root formation in Arabidopsis. (A) Root phenotype of 16-day-old Arabidopsis seedlings grown on medium containing the indicated 2,4-D analogue. (B) The number of lateral roots formed on Arabidopsis seedling primary roots treated with mock or 5 µM of the indicated 2,4-D analogue. (C) Quantification of the effect of 1 day treatment on the different developmental stages of lateral root formation including lateral root initiation (stage 1), lateral root primordium development (stage 2), lateral root emergence (stage 3) and lateral root outgrowth (stage 4) using pDR5:GUS activity to visualize the early stages of lateral root development. In B and C dots indicate the number of lateral roots in B and number of lateral roots per developmental stage in C (B: n=10 and C: n=7), the bars indicate the mean value and the error bars the standard error. The letters above the bars indicate the significant difference (p < 0.01) determined by a one-way ANOVA with Tukey's honest significant difference post hoc test. Seedlings were first grown for 6 days on hormone-free medium, and subsequently transferred to mock medium or medium with 5 µM of the indicated compound and further incubated for 10 days in A and B and for 1 day in C.

## Molecular dynamics simulations of binding of 2,4-D analogues to the TIR1-Aux/IAA coreceptors

The auxin activity of 2,4-D occurs through its binding to the TIR1/AFB – Aux/IAA auxin co-receptors in the auxin-mediated signaling pathway[22]. We examined this binding for a subset of 18 2,4-D analogues, representing derivatives from all auxin activity classes as determined in the *in planta* experiments, by molecular dynamics (MD) simulations. Among the library entries tested *in silico* are alpha substituted compounds (methyl, ethyl or the Mecoprop derivative; 2,4-DP, 2,4-DB and Mecoprop, respectively) that due to their stereocenter can occur as the

enantiomeric R or S stereoisomers. These compounds were tested as mixture of these enantiomers in the *in planta* assays. The molecular dynamics simulation uniquely allowed us to investigate at a molecular level the potential differential binding of each enantiomer (notated as their acronym with an added R and S).

In the evaluation of the binding capacity of each auxin analogue within the system, we studied different aspects: the enthalpy of the system, the root-mean-square deviation (RMSD) of atomic positions of TIR1 along the simulation trajectory, the hydrogen bond network established by the auxin analogue and the distance between the auxin analogue and the proline (7) of the Aux/IAA peptide (Table 1). The RMSD of TIR1 reflects the system stability during the simulation. The last parameter, the distance between the auxins analogue and the proline (7) of the Aux/IAA peptide, would show us which analogue is able to establish a CH·π interaction with the degron peptide that contributes to the binding interactions and thus to the stability of the complex[41]. Our molecular dynamics simulations were based on a previously published method[42], but with a much longer trajectory of 200 ns instead of 2.5 ns.

The enthalpy was calculated using MMPBSA (Molecular Mechanics Poisson-Boltzmann Surface Area) with a lower enthalpy value indicating better binding. All the compounds classified as very active auxin analogues showed an enthalpy value below -12 Kcal/mol. Surprisingly, the weakly active 2,6-D and inactive 4-NO$_2$ were also in this range, suggesting that the binding enthalpy is not the sole parameter predictive for auxin activity. The TIR1 protein RMSD calculation was carried out over the amino acids from position 50 till the end of the protein, because the first 50 amino acids showed extra flexibility. The TIR1 N-terminal part is normally stabilized by the interaction with the SKP protein, which we did not include in our analysis. Moreover this part does not interact with the auxin analogue nor with the AUX/IAA peptide and its structure has not been completely resolved in the employed crystal structure (pdbid 2p1n)[22]. All the compounds induced similar RMSD values of the TIR1 protein ranging from 1.25 to 1.53, indicating that binding to the compounds tested do not induce significant differences in the stability of the protein backbone (Table 1). All compounds tested appeared to establish a hydrogen bond network but with different effectiveness (Figure S4). Serine (438) and arginine (403) interact through a hydrogen bond with all the derivatives, but these interactions have different prevalence. Other amino acids that are involved in most of the hydrogen bond networks established by the analogues are arginine (436) and histidine (78).

Finally, we observed the proximity between the proline (7) from the AUX/IAA peptide and the analogue, so we monitored this distance along the simulation trajectory for all the derivatives. The distance was measured specifically between the gamma carbon of proline (7) and the ring mass center of the auxin analogue, assuming that a shorter distance would predict a stronger CH·π interaction[41]. Compounds 4-NO$_2$ (5.28 Å) and 2,6-D (5 Å) showed the two largest distances at which the CH·π contribution is more or less negligible, explaining why these compounds are weakly active or inactive despite their low enthalpy. For IAA and the very active auxin analogues distance values varied between 3,7 Å (Mecoprop) and 4,5 Å (2,4-DB(R)). Several weakly active and anti-auxin compounds were also in this range. However, as the binding enthalpy was above -12 Kcal/mol, this explained why these compounds lacked auxin activity.

| Compound | Enthalpy binding (Kcal/mol) | std | Activity LU | Protein RMSD (Å) | std | Distance gamma (Å) | std |
|---|---|---|---|---|---|---|---|
| 2,4-DB(S) | -30,8 | 0,70 | very active | 1,3 | 0,16 | 4,3 | 0,33 |
| NAA | -29,6 | 1,06 | very active | 1,6 | 0,15 | 4,3 | 0,39 |
| 2,4-DB(R) | -29,3 | 0,86 | very active | 1,3 | 0,12 | 4,5 | 0,45 |
| 2,4-D | -26,9 | 0,84 | very active | 1,4 | 0,12 | 4,2 | 0,38 |
| IAA 2p1n | -26,8 | 0,80 | very active | 1,2 | 0,13 | 4,0 | 0,28 |
| Mecoprop(R) | -26,7 | 0,70 | very active | 1,5 | 0,19 | 3,7 | 0,3 |
| 2,4-DP(R) | -25,1 | 1,51 | very active | 1,5 | 0,11 | 4,0 | 0,37 |
| 2-Cl-4-F | -23,2 | 0,81 | very active | 1,3 | 0,12 | 4,0 | 0,31 |
| 2,4-DP(S) | -22,4 | 0,65 | very active | 1,2 | 0,1 | 4,2 | 0,33 |
| Mecoprop(S) | -15,1 | 0,70 | very active | 1,3 | 0,1 | 4,3 | 0,36 |
| 4-Cl | -13,0 | 0,53 | very active | 1,4 | 0,15 | 3,9 | 0,28 |
| MCPA | -12,8 | 0,69 | very active | 1,4 | 0,15 | 3,8 | 0,29 |
| 2,4-DiB | -19,5 | 0,43 | active | 1,5 | 0,14 | No CH·p | 0,53 |
| 3,4-D | -7,8 | 0,54 | active | 1,3 | 0,09 | 4,0 | 0,37 |
| 2,6-D | -36,5 | 0,76 | weakly active | 1,4 | 0,13 | No CH·p | 0,43 |
| 2,3,4-T | -8,8 | 0,57 | weakly active | 1,2 | 0,11 | 3,6 | 0,23 |
| 2,3,4-F | -8,3 | 0,67 | weakly active | 1,3 | 0,11 | 4,2 | 0,47 |
| 4-NO$_2$ | -13,6 | 0,54 | not active | 1,5 | 0,23 | No CH·p | 0,96 |
| 2,4-DnP | -8,4 | 1,06 | anti-auxin | 1,1 | 0,09 | 4,2 | 0,57 |
| 2,3-D | -6,6 | 0,69 | anti-auxin | 1,3 | 0,13 | 4,2 | 0,55 |

Table 1. Parameters obtained from the in silico enthalpy binding study of selected 2,4-D analogues and auxins to the TIR1 protein. Several parameters were calculated for the molecular dynamics simulations. The enthalpy binding of the selected compounds and the standard deviation (std) was determined. The established auxin activity of the SAR is given (activity LU). The root-mean-square deviation of the TIR1 protein (Protein RMSD) with a standard deviation was calculated as a measure of system stability during the simulation trajectory. The distance between the gamma carbon of the proline of the AUX/IAA peptide and the ring mass center of the compound was determined (Distance gamma) and the corresponding standard deviation. At shorter distances, a stronger CH π interaction is assumed.

## Somatic embryogenesis induction by 2,4-D analogues

In order to correlate auxin activity with the capacity to induce SE, we tested the ability of all 2,4-D analogues (Figure 2) to induce SE in *Arabidopsis* immature zygotic embryos (IZEs), which in prior studies has proven to be the most

competent *Arabidopsis* tissue for SE in response to 2,4-D[43].

In this assay, several very active (4-Cl, 4-Br, 2,4-DP,) and active auxins compounds (2,4-Br) were able to efficiently induce somatic embryos from IZEs to a similar extent as 2,4-D (Figure 6A).The other very active (2,4-DB 2-Cl-4-F, MCPA and Mecoprop) and active (4-F, 4-I, 2,5-D, 3,4-D, 2-F-4-Cl, 2-F-4-Br, 2,4,5-T, 2,4-DiB and MCPB) auxin analogues induced less somatic embryos on IZEs when compared with 2,4-D (Figure 6A, B). The weak auxin analogues (2-Cl, 2-I, 3-Cl, 2,6-D, 3,5-D, 2,4-F, 2,3,4-F, 2,3,4-T, 2-NO2-4-Cl, and 2,4-DnB), or the inactive (PHAA, 4-NO$_2$, 3,5-Me, 2-Cl-4-formyl, and 2,4,6-F) or anti-auxin (3-Me, 2,3-D, 3,5-D, 2,4,6-T and 2,4-DnP) compounds did not induce SE (Figure S3). Overall, these results demonstrated that the capability of compounds to induced SE correlates with their 2,4-D-like structure and their previously established auxin activity.

We also found that a higher concentration (10 μM) of 2,4-D, 4-Br, 4-I and MCPA reduced the number of embryos formed on IZEs, whereas 10 μM of 4-Cl, 4-F, or 2,4,5-T had no significant effect on the SE efficiency (Figure 6B). This different SE response to higher levels of certain 2,4-D analogues is probably associated with different metabolic properties, such as uptake and active transport, of these compounds in plant cells.

Summarizing, there is a strong correlation between auxin activity and SE induction: SE is only obtained with the very active and active auxin analogs (Figure 6A versus Figure S3) and the very active auxins are generally more efficient in inducing SE than the active auxins (Figure 6B).

5
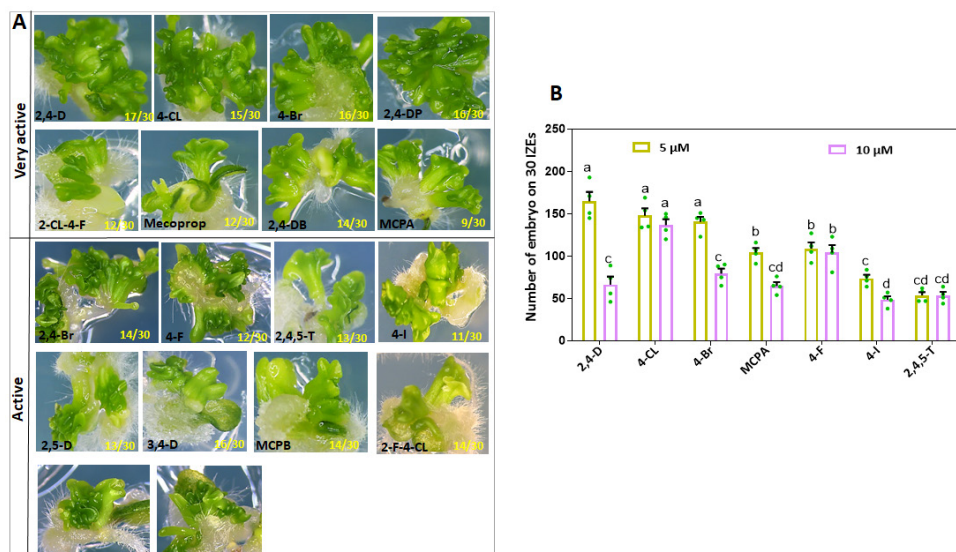
Figure 6. SE inducing capacity of 2,4-D analogues correlates with their auxin activity. (A) The phenotype of somatic embryos formed on cotyledons of Arabidopsis IZEs that were first grown for two weeks in the presence of 5 μM 2,4-D or the 2,4-D analogues classified as very active or active auxins and subsequently cultured for 1 week on medium without any supplement. (B) Quantification of the number of somatic embryos per 30 IZEs induced by 5 μM or 10 μM of 2,4-D or of the indicated 2,4-D analogue. Dots indicate the number somatic embryos produced per IZE (n=4 biological replicates, with 30 IZEs per replicate), bars indicate the mean value and error bars the s.e.m.. Different letters indicate statistically significant differences (P < 0.001) as determined by one-way analysis of variance with Tukey's honest significant difference post hoc test. The structure of each compound is provided in Figure 2.

## Discussion

Chemical biology has been instrumental in enhancing our understanding of auxin biology, as several key molecules that are currently used to specifically control auxin biosynthesis, signaling or transport have been identified in chemical genetic screens[1,44]. 2,4-D is a synthetic auxin that is widely used as a herbicide, but also as a plant growth regulator in *in vitro* regeneration and auxin research[2,45]. In this study, 2,4-D analogues with varying modifications, including the type and number of substitutions on the aromatic ring (like chlorine, bromine, fluorine, iodine, nitro, methyl group) or substitutions on or elongation of the alpha-position (Figure 2) have been evaluated for their structure-activity relationship and ability to induce SE.

## Auxin activity of 2,4-D analogues correlates to simulated binding properties to TIR1

The Arabidopsis seedling root growth inhibition assay combined with the use of the *pDR5:GUS* and *R2D2* auxin response reporters provided consistent results

with respect to classifying the 40 2,4-D orthologs as very active, active or weakly active auxins, or having no auxin activity. Importantly, this classification was in agreement with previously published results on the 2,4-D analogues PAA, 2,4 Br, 2,4,5-T, 3-Me, 2,4-DP and MCPB[44,46]. This difference in auxin activity is clearly determined by structure and chemical properties of these compounds, which relates on the one hand to their binding affinity to the TIR1/AFB[3] and AUX/IAA[46] co-receptor pair, and on the other hand to their conjugation properties or metabolic decay in plant cells. For example, 2,4-D itself it has been shown to be conjugated to Asp and Glu, and the conjugates still display residual auxin activity[47].

Molecular dynamics simulations indicated that the auxin activity of 2,4-D analogues could be correlated to their binding strength to TIR1. One important factor here was the theoretical enthalpy of the system, as all very active auxin analogues showed an enthalpy value below -12 Kcal/mol. However, the distance between Pro7 of TIR1 and the 2,4-D analogue was also important, as some of the tested analogues lacking this CH·π interaction showed no or only weak auxin activity, despite the fact that their enthalpy value was below -12 Kcal/mol.

It is important to note that, due to the expensive cost of entropy calculation combined with low reliability for such a big system, we did not calculate this parameter. However, we considered that the molecules are very similar and, although we know there will be differences in entropy among the derivatives, we assume that they will not change the results dramatically.

## Structure-activity relationship of 2,4-D analogues based on auxin activity

Based on our screen of 40 compounds, we have identified several trends for the structure auxin activity relationship of 2,4-D analogues. First, a halogen at the 4-position is more important for auxin activity than at the 2-position (4-Cl; 2-Cl-4-F; 4-I/Br/F; versus 2-Cl/I/Br/F). Second, the activity of analogues with a substitution at the 3- or 5- position remains, but such compounds are less active compared to 2,4-substitutions (2,5-D; 3,4-D; 2,4,5-T). Third, both substitutions at the alpha position (Mecoprop; 2,4-DP; 2,4-DB; 2,4-DiB) and a longer carboxylate chain with an even number of carbons are tolerated (MCPB; 2,4-DnB). This finding is in agreement with a previous study that showed that IAA analogues with a substitution on their carboxylate chain, up to n=4 carbons, can still bind to TIR1, thus allowing for a modification on this position[35]. Taken

these trends together, we conclude that an electron withdrawing group, such as a halogen, at the 4-position is important for auxin activity. In addition, we suggest that electron withdrawing or donating properties of the 2-position are less important for auxin activity and that the size of the 2-substituent could influence auxin activity, since both methyl and chloro groups are accepted at this position (like 2,4-D and MCPA). There are, however, a few outliers based on this general conclusion: 2,4-F is only weakly active, but has a strong electron withdrawing group at the 4-position, and 2,4-DnB is also weakly active, but has the same substitution as 2,4-D and a longer carboxylate chain which does not disrupt auxin activity for MCPB. Altogether, these results indicate that small modifications at different parts of 2,4-D can lead to different physiological activities, generating molecules with interesting applications in plant tissue culture, as described below.

## Several 2,4-D analogues are useful tools to modulate the root system architecture

In the assessment of 2,4-D analogues, we have found that some 2,4-D analogues (2,3-D, 3,5-D, 3-Me, 2,4,6-T and 2,4-DnP) either had no effect on or only slightly affected root growth (Figure 2), whereas they inhibited the formation of lateral roots (Figure 6). Since inhibition of lateral root formation is typical for compounds with anti-auxin activity[39], we suspected that they may act as anti-auxins. Such anti-auxin activity could be caused by high affinity binding to only one of the co-receptors, thereby preventing the formation of the F-box-Aux/IAA complex. As such, high concentrations of an anti-auxin would compete effectively with IAA for establishing the TIR1-Aux/IAA complex required for ubiquitination-mediated degradation of the Aux/IAA repressors. Interestingly, our 2,4-D analogues did not change *pDR5* activity, nor did they lead to enhanced or reduced degradation of DII-VENUS in the root tip, which would be expected for an anti-auxin. Moreover, the negative effect on lateral root formation varied per compound. For example, 3-Me only mildly affected lateral root formation, whereas 3,5-D and 2,4-DnP almost completely inhibited this process, with 3,5-D preferably blocking at stage 2 (enhanced number of stage 2 primordia) whereas 2,4-DnP resulted in a reduction of all stages. Clearly, further research on these 2,4-D analogues, such as a forward genetic screen for mutants that develop lateral roots when grown on the compound, is required to unravel the exact molecular mechanism by which they repress lateral root formation. As such, these 2,4-D analogues could uncover new components involved in lateral root formation. At the same time, they could be used to control lateral root

formation, e.g. to prevent branching during early seedling development to obtain a deeper root system.

Adventitious root (AR) formation is an organogenesis process by which new roots are produced from non-root tissues. AR induction is an important but often rate limiting step in the vegetative propagation of many horticultural and forestry plant species. Generally, AR induction is promoted by auxins, and IBA and NAA are the most common auxins used for this purpose in commercial operations[38]. Our results show that 2,5-D and 3-Cl efficiently induce AR from *Arabidopsis* hypocotyls even more efficiently than IBA or NAA. Therefore, 2,5-D and 3-Cl are recommended as new AR inducers and they may be used to resolve rooting in recalcitrant species.

Roots hairs extend from root epidermal cells in the differentiation zone of the root and support plants in nutrient absorption, anchorage, and microbe interactions[37]. Exogenous auxin treatments generally promotes root hair induction and development[37]. We showed that some 2,4-D analogues (3-Cl, 3,4-D, 2,4-DP, and Mecoprop) induced ectopic root hair formation on Arabidopsis roots, while promoting lateral root formation and having a relatively mild effect on root growth compared to 2,4-D itself. Specifically, roots of plants grown on the analogue 3-Cl produced several times more root hairs compared to other analogues. Unlike other 2,4-D analogues, 3-Cl specifically induced a strong auxin response in the differentiation zone of the root and a weaker response in the root elongation zone (Figure S2), this could explain the differential effect of 3-Cl on root biomass. We recommend 3-Cl as a promising compound for use in horticulture or agriculture to enhance root hair formation and thereby improve crop performance by enhanced ion and water uptake.

## The capacity of SE induction by 2,4-D analogues

SE is experimentally induced by application of exogenous plant growth regulators under *in vitro* conditions[48]. 2,4-D is the most well-known plant growth regulator that is widely used for SE. Our screen revealed several 2,4-D analogues (4-Cl, 4-Br, 2,4-DP, and 2,4-Br) that also efficiently induce SE from Arabidopsis IZEs, whereas other analogues (4-F, 4-I, 2,5-D, 3,4-D, 2-F-4-Cl, 2-F-4-Br, 2,4,5-T, 2,4-DiB, MCPB, MCPA, and Mecoprop) induced SE at a lower efficiency (Figure 5). Except for 3,4-D and 2,4,5-T that have been shown to efficiently induce SE in other plant species[49], SE induction by the other 2,4-D analogues has not been reported yet. We did not identify a 2,4-D analogue with enhanced capacity to

induce SE, suggesting that for our system 2,4-D is still the best compound to use. However, since 2,4,5-T efficiently induces SE in other plant species but less efficiently in Arabidopsis, it is to be expected that some of the 2,4-D analogues presented here might be used to establish an efficient SE system in other plant species. Moreover, 2,4-D treatment has been reported to lead to a significant percentage of abnormal somatic embryos in many plant species[50]. The use of the 2,4-D analogues identified here may reduce the number of abnormal embryos, but this still requires testing in the different SE systems.

We also showed that only 2,4-D analogues that classify as very active or active auxins are able to induce SE from Arabidopsis IZEs, which indicates that the capacity to induce SE is primarily linked to auxin activity. Since stress has also been identified as an important factor in SE induction[13–17], our findings suggest that at least in our SE system stress is downstream of auxin signaling, and not a parallel pathway that is additionally triggered by 2,4-D or its analogues. It is therefore unlikely that a chemical biology approach as presented here will allow to tease apart the stress and auxin pathways by identifying compounds that trigger each pathway separately. The new 2,4-D analogues identified here can be useful tools to study the importance of aspects of auxin physiology either during SE induction or through their effects on the root system architecture. For example, it would be interesting to understand why MCPA, which is classified as very active auxin based on root growth inhibition and auxin response reporter expression, shows a significantly reduced capacity to induce SE. And why does 3-Cl, classified as weakly active auxin, strongly induce the auxin response in the root differentiation zone thereby specifically promoting root hair formation? Or how can 3,5-D strongly inhibit lateral root formation without clearly affecting auxin responsive gene expression? These differences may lie in the compound specific metabolism or transport characteristics of the 2,4-D analogues, which has already been shown to differ between the natural auxin IAA and the synthetic auxins 2,4-D and 1-NAA[8,47,51].

## Materials and methods

### 2,4-D and 2,4-D-analogues

The 2,4-D analogues that were used in this study were divided into 2 main categories. The first one contained 2,4-D analogues with substituents at different positions on the aromatic ring (Figure 1). The second group consisted of 2,4-D analogues with different carbon chains at the alpha-position of 2,4-D (Figure 1). For more information about the synthesis of certain library members and the sources for commercially available compounds, the reader is kindly referred to the Supplementary Information.

The acronyms of the 2,4-D analogues were assigned based on established acronyms in literature, or based on the number and substitution type on the aromatic ring, or the type of substituent on the alpha position of the carboxylate.

### Plant materials and growth conditions

This research used *Arabidopsis thaliana* Columbia-0 (Col-0) wild-type, and the *pDR5::GUS* and *R2D2* reporter lines in the Columbia background. For *in vitro* plant culture, seeds were sterilized in 10 % (v/v) sodium hypochlorite for 10 minutes and then 4 times washed with MQ water. Sterilized seeds were plated on ½ MS medium and grown in the tissue culture room at 21°C, 16 hours photoperiod and 50 % relative humidity. After around 2 weeks, the germinated seeds were planted in soil in the climate room at 20°C, 16 hours photoperiod and 70 % relative humidity.

### *Arabidopsis* primary root growth assay adventitious root induction

The *Arabidopsis* seeds were germinated on ½ MS medium. Five-day-old seedlings were transferred to new ½ MS medium supplemented with 2,4-D analogues. The length of the primary root was quantified after 3 days, incubation with 2,4-D analogues. Primary root length and lateral root numbers were analyzed with ImageJ software.

For the adventitious root (AR) induction, the seed was first grown in complete darkness. Six-day-old seedlings were transferred to new ½ MS medium supplemented with 2,4-D analogues 16 hours photoperiod. The number of ARs induced from hypocotyls was quantified after 7 days.

### Histological GUS staining assay

In order to test auxin activity of 2,4-D analogues, the activity *pDR5:GUS* reporter

was investigated in the presence of 5 µM 2,4-D analogues. Histochemical β-glucuronidase (GUS) staining was performed as described previously[52] with some modifications. Samples were submerged in 1-2mL staining solution and incubated for 4 hours at 37°C followed by rehydration through a graded ethanol series 75% - 50% - 25% for 10 minutes each, with 5 minutes incubation between each step. *Arabidopsis* Col-0 on MS medium without any supplement was used as a control.

The tissue-specific GUS staining intensity was quantified as mean grey values by analyzing images of independent samples capturing the same region of interest (ROI) using ImageJ, as previously described by Béziat et al.[53].

## Somatic embryogenesis induction

For SE induction, 11 days old siliques of *Arabidopsis* Columbia-0 wild-type were sterilized in 10 % (v/v) sodium hypochlorite for 10 minutes and then washed with MQ water for 4 times. Sterilized siliques were dissected to acquire IZEs. IZEs were cultured on B5 medium mixed with 2,4-D analogues for 14 days in the culture room at 21°C, 16 hours photoperiod and 50 % relative humidity. The IZEs were then transferred to ½ MS medium for 7 days and the number of SEs was counted under a light microscope. The quality of the obtained SEs was then further examined by moving embryos to the MS media in square petri dishes. Some 2,4-D analogues inactive in inducing SE were combined with IAA or NAA. In this experiment, B5 medium supplemented with 2,4-D used as a control.

## Microscopy

GUS stained tissues were observed and photographed using a Nikon's Eclipse E800 microscope. Cellular and subcellular localization of DII-VENUS and TdTOMATO proteins was visualized by confocal laser scanning microscope (ZEISS-003-18533) using a 534 nm laser combined with a 488 nm LP excitation and 500-525 nm BP emission filters for DII-VENUS signals or 633 laser and 532 nm LP excitation and 580-600 nm BP emission filters for TdTOMATO signals.

## The running molecular dynamics simulations

MD simulations were run using the crystal structure of the systems comprising the TIR1, auxin responsive protein fragment, co-factor inositol hexakisphosphate, and the auxin. In particular, the structure deposited in the PDB under the ID

2P1N[22] was used as a starting structure. The auxin derivative was replaced by the different auxin candidates before running the MD simulations. Both the system preparation and the simulations were performed in the AMBER 18 suite software[54]. An updated version of the protocol described in the Hao and Yang's article was used[42]. Firstly, the system is immersed in a cubic box of 10 Å length in each direction using TIP3P (Mol. Phys, 1988, 94, 803-808) water parameters. The force fields used to obtain topography and coordinates files were ff14SB [55] and GAFF [56]. The first step to start MD simulations is a minimization only of the positions of solvent molecules keeping the solute atom positions restrained and, the second stage minimizes of all the atoms in the simulation cell. Heating the system is the third step raising gradually the temperature 0 to 300 K under a constant pressure of 1 atm and periodic boundary conditions. In addition, Harmonic restraints of 10 kcal·mol$^{-1}$ were applied to the solute, and the Berendsen temperature coupling scheme[57] was used to control and equalize the temperature. The time step was kept at 2 fs during the heating phase. Long-range electrostatic effects were modelled using the particle-mesh-Ewald method[58]. The Lennard-Jones interactions cut-off was set at 8 Å. An equilibration step for 2 ns with a 2 fs time step at a constant volume and temperature of 300 K was performed prior the production stage. The trajectory production stage kept the equilibration step conditions and was prolonged for 200 ns longer at 1 fs time step. Besides, the auxin analogs required a previous preparation step where the parameters and charges were generated by using the antechamber module of AMBER, using GAFF force field and AM1-BCC method for charges[59].

5

# Chapter 5

# References

(1) Ma, Q.; Grones, P.; Robert, S. Auxin Signaling: A Big Question to Be Addressed by Small Molecules. *J. Exp. Bot.* **2018**, *69*, 313–328.

(2) Peterson, M. A.; McMaster, S. A.; Riechers, D. E.; Skelton, J.; Stahlman, P. W. 2,4-D Past, Present, and Future: A Review. *Weed Technol.* **2016**, *30*, 303–345.

(3) Calderón Villalobos, L. I. A.; Lee, S.; De Oliveira, C.; Ivetac, A.; Brandt, W.; Armitage, L.; Sheard, L. B.; Tan, X.; Parry, G.; Mao, H.; Zheng, N.; Napier, R.; Kepinski, S.; Estelle, M. A Combinatorial TIR1/AFB–Aux/IAA Co-Receptor System for Differential Sensing of Auxin. *Nat. Chem. Biol.* **2012**, *8*, 477–485.

(4) Shimizu-Mitao, Y.; Kakimoto, T. Auxin Sensitivities of All Arabidopsis Aux/IAAs for Degradation in the Presence of Every TIR1/AFB. *Plant Cell Physiol.* **2014**, *55*, 1450–1459.

(5) Leyser, O. Auxin Signaling. *Plant Physiol.* **2018**, *176*, 465–479.

(6) Winkler, M.; Niemeyer, M.; Hellmuth, A.; Janitza, P.; Christ, G.; Samodelov, S. L.; Wilde, V.; Majovsky, P.; Trujillo, M.; Zurbriggen, M. D.; Hoehenwarter, W.; Quint, M.; Calderón Villalobos, L. I. A. Variation in Auxin Sensing Guides AUX/IAA Transcriptional Repressor Ubiquitylation and Destruction. *Nat. Commun.* **2017**, *8*, 15706.

(7) Iglesias, M. J.; Terrile, M. C.; Correa-Aragunde, N.; Colman, S. L.; Izquierdo-Álvarez, A.; Fiol, D. F.; París, R.; Sánchez-López, N.; Marina, A.; Calderón Villalobos, L. I. A.; Estelle, M.; Lamattina, L.; Martínez-Ruiz, A.; Casalongué, C. A. Regulation of SCFTIR1/AFBs E3 Ligase Assembly by S-Nitrosylation of Arabidopsis SKP1-Like1 Impacts on Auxin Signaling. *Redox Biol.* **2018**, *18*, 200–210.

(8) Song, Y. Insight into the Mode of Action of 2,4-Dichlorophenoxyacetic Acid (2,4-D) as an Herbicide. *J. Integr. Plant Biol.* **2014**, *56*, 106–113.

(9) Karami, O.; Saidi, A. The Molecular Basis for Stress-Induced Acquisition of Somatic Embryogenesis. *Mol. Biol. Rep.* **2010**, *37*, 2493–2507.

(10) Guan, Y.; Li, S. G.; Fan, X. F.; Su, Z. H. Application of Somatic Embryogenesis in Woody Plants. *Front. Plant Sci.* **2016**, *7*, 938.

(11) Bhojwani, S. S. *Plant Tissue Culture : Applications and Limitations*; Elsevier, 1990.

(12) Fehér, A. Somatic Embryogenesis - Stress-Induced Remodeling of Plant Cell Fate. *Biochim. Biophys. Acta - Gene Regul. Mech.* **2015**, *1849*, 385–402.

(13) Salvo, S. A. G. D.; Hirsch, C. N.; Buell, C. R.; Kaeppler, S. M.; Kaeppler, H. F. Whole Transcriptome Profiling of Maize during Early Somatic Embryogenesis Reveals Altered Expression of Stress Factors and Embryogenesis-Related Genes. *PLoS One* **2014**, *9*, e111407.

(14) Jin, F.; Hu, L.; Yuan, D.; Xu, J.; Gao, W.; He, L.; Yang, X.; Zhang, X. Comparative Transcriptome Analysis between Somatic Embryos (SEs) and Zygotic Embryos in Cotton: Evidence for Stress Response Functions in SE Development. *Plant Biotechnol. J.* **2014**, *12*, 161–173.

(15) Nowak, K.; Wójcikowska, B.; Gaj, M. D. *ERF022* Impacts the Induction of Somatic Embryogenesis in Arabidopsis through the Ethylene-Related Pathway. *Planta* **2015**, *241*, 967–985.

(16) Gliwicka, M.; Nowak, K.; Balazadeh, S.; Mueller-Roeber, B.; Gaj, M. D. Extensive Modulation of the Transcription Factor Transcriptome during Somatic Embryogenesis in Arabidopsis Thaliana. *PLoS One* **2013**, *8*, e69261.

(17) Mantiri, F. R.; Kurdyukov, S.; Lohar, D. P.; Sharopova, N.; Saeed, N. A.; Wang, X. D.; Vandenbosch, K. A.; Rose, R. J. The Transcription Factor MtSERF1 of the ERF Subfamily Identified by Transcriptional Profiling Is Required for Somatic Embryogenesis Induced by Auxin plus Cytokinin in *Medicago Truncatula. Plant Physiol.* **2008**, *146*, 1622–1636.

(18) Raghavan, V. Role of 2,4-Dichlorophenoxyacetic Acid (2,4-D) in Somatic Embryogenesis on Cultured Zygotic Embryos of Arabidopsis: Cell Expansion, Cell Cycling, and Morphogenesis during Continuous Exposure of Embryos to 2,4-D. *Am. J. Bot.* 2004.

(19) Ferro, N.; Bredow, T.; Jacobsen, H. J.; Reinard, T. Route to Novel Auxin: Auxin Chemical Space toward Biological Correlation Carriers. *Chem. Rev.* **2010**, *110*, 4690–4708.

(20) Simon, S.; Petrášek, J. Why Plants Need More than One Type of Auxin. *Plant Sci.* 2011, *180*, 454–460.

(21) Lee, S.; Sundaram, S.; Armitage, L.; Evans, J. P.; Hawkes, T.; Kepinski, S.; Ferro, N.; Napier, R. M. Defining Binding Efficiency and Specificity of Auxins for SCF TIR1/AFB -Aux/IAA Co-Receptor Complex Formation. *ACS Chem. Biol.* **2014**, *9*, 673–682.

(22) Tan, X.; Calderon-Villalobos, L. I. A.; Sharon, M.; Zheng, C.; Robinson, C. V.; Estelle, M.; Zheng, N. Mechanism of Auxin Perception by the TIR1 Ubiquitin Ligase. *Nature* **2007**, *446*, 640–645.

(23) van der Zaal, B. J.; Droog, F. N.; Pieterse, F. J.; Hooykaas, P. J. Auxin-Sensitive Elements from Promoters of Tobacco *GST* Genes and a Consensus as-1-like Element Differ Only in Relative Strength. *Plant Physiol.* **1996**,

*110*, 79–88.

(24) Hayashi, K. I.; Neve, J.; Hirose, M.; Kuboki, A.; Shimada, Y.; Kepinski, S.; Nozaki, H. Rational Design of an Auxin Antagonist of the SCFTIR1auxin Receptor Complex. *ACS Chem. Biol.* **2012**, *7*, 590–598.

(25) Hamilton, R. H.; Kivilaan, A.; McManus, J. M. Biological Activity of Tetrazole Analogues of Indole-3-Acetic Acid and 2,4-Dichlorophenoxyacetic Acid. *PLANT Physiol.* **1952**, *35*, 136–140.

(26) Quareshy, M.; Prusinska, J.; Kieffer, M.; Fukui, K.; Pardal, A. J.; Lehmann, S.; Schafer, P.; del Genio, C. I.; Kepinski, S.; Hayashi, K.; Marsh, A.; Napier, R. M. The Tetrazole Analogue of the Auxin Indole-3-Acetic Acid Binds Preferentially to TIR1 and Not AFB5. *ACS Chem. Biol.* **2018**, *13*, 2585–2594.

(27) Koepfli, J. B.; Thimann, K. V.; Went, F. W. Phytohormones; Structure and Physiological Activity. *J. Biol. Chem.* **1938**, *122*, 736–780.

(28) Ferro, N.; Bultinck, P.; Gallegos, A.; Jacobsen, H. J.; Carbo-Dorca, R.; Reinard, T. Unrevealed Structural Requirements for Auxin-like Molecules by Theoretical and Experimental Evidences. *Phytochemistry* **2007**, *68*, 237–250.

(29) Katekar, G. F. Auxins: On the Nature of the Receptor Site and Molecular Requirements for Auxin Activity. *Phytochemistry* **1979**, *18*, 223–233.

(30) Porter, W. L.; Thimann, K. V. Molecular Requirements for Auxin Action-I. Halogenated Indoles and Indoleacetic Acid. *Phytochemistry* **1965**, *4*, 229–243.

(31) Rahman, A.; Bannigan, A.; Sulaman, W.; Pechter, P.; Blancaflor, E. B.; Baskin, T. I. Auxin, Actin and Growth of the *Arabidopsis Thaliana* Primary Root. *Plant J.* **2007**, *50*, 514–528.

(32) Ulmasov, T.; Murfett, J.; Hagen, G.; Guilfoyle, T. J. Aux/IAA Proteins Repress Expression of Reporter Genes Containing Natural and Highly Active Synthetic Auxin Response Elements. *Plant Cell* **1997**, *9*, 1963–1971.

(33) Liao, C. Y.; Smet, W.; Brunoud, G.; Yoshida, S.; Vernoux, T.; Weijers, D. Reporters for Sensitive and Quantitative Measurement of Auxin Response. *Nat. Methods* **2015**, *12*, 207–210.

(34) Smith, S.; de Smet, I. Root System Architecture: Insights from *Arabidopsis* and Cereal Crops. *Philos. Trans. R. Soc. B Biol. Sci.* **2012**, *367*, 1441–1452.

(35) Hayashi, K.-I.; Tan, X.; Zheng, N.; Hatate, T.; Kimura, Y.; Kepinski, S.; Nozaki, H. Small-Molecule Agonists and Antagonists of F-Box Protein-Substrate Interactions in Auxin Perception and Signaling. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 5632–5637.

(36) da Costa, C. T.; Gaeta, M. L.; de Araujo Mariath, J. E.; Offringa, R.; Fett-Neto, A. G. Comparative Adventitious Root Development in Pre-Etiolated and Flooded *Arabidopsis* Hypocotyls Exposed to Different Auxins. *Plant Physiol. Biochem.* **2018**, *127*, 161–168.

(37) Lee, R. D. W.; Cho, H. T. Auxin, the Organizer of the Hormonal/Environmental Signals for Root Hair Growth. *Front. Plant Sci.* **2013**, *4*, 448.

(38) Geiss, G.; Gutierrez, L.; Bellini, C. Adventitious Root Formation: New Insights and Perspectives. In *Root Development*; Wiley-Blackwell: Oxford, UK, 2009; pp 127–156.

(39) Larsen, P. B. Anti-Auxin Compounds. U.S. Patent 0073308A1. 0073308A1, 2017.

(40) Dubrovsky, J. G.; Sauer, M.; Napsucialy-Mendivil, S.; Ivanchenko, M. G.; Friml, J.; Shishkova, S.; Celenza, J.; Benková, E. Auxin Acts as a Local Morphogenetic Trigger to Specify Lateral Root Founder Cells. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 8790–8794.

(41) Wang, J.; Yao, L. Dissecting C−H···π and N−H···π Interactions in Two Proteins Using a Combined Experimental and Computational Approach. *Sci. Rep.* **2019**, *9*, 20149.

(42) Hao, G. F.; Yang, G. F. The Role of Phe82 and Phe351 in Auxin-Induced Substrate Perception by Tir1 Ubiquitin Ligase: A Novel Insight from Molecular Dynamics Simulations. *PLoS One* **2010**, *5*, e10742.

(43) Gaj, M. D. Direct Somatic Embryogenesis as a Rapid and Efficient System for in Vitro Regeneration of *Arabidopsis Thaliana*. *Plant Cell. Tissue Organ Cult.* **2001**, *64*, 39–46.

(44) Simon, S.; Kubeš, M.; Baster, P.; Robert, S.; Dobrev, P. I.; Friml, J.; Petrášek, J.; Zažímalová, E. Defining the Selectivity of Processes along the Auxin Response Chain: A Study Using Auxin Analogues. *New Phytol.* **2013**, *200*, 1034–1048.

(45) Skůpa, P.; Opatrný, Z.; Petrášek, J. Auxin Biology: Applications and the Mechanisms Behind. In *Applied Plant Cell Biology: Cellular Tools and Approaches for Plant Biotechnology*; Nick, P., Opatrny, Z., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, **2014**; pp 69–102.

(46) Torii, K. U.; Hagihara, S.; Uchida, N.; Takahashi, K. Harnessing Synthetic Chemistry to Probe and Hijack Auxin Signaling. *New Phytol.* **2018**, *220*, 417–424.

(47) Eyer, L.; Vain, T.; Pařízková, B.; Oklestkova, J.; Barbez, E.; Kozubíková, H.; Pospíšil, T.; Wierzbicka, R.; Kleine-Vehn, J.; Fránek, M.; Strnad, M.; Robert, S.; Novak, O. 2,4-D and IAA Amino Acid Conjugates Show Distinct Metabolism in Arabidopsis. *PLoS One* **2016**, *11*, e0159269.

(48) Nic-Can, G. I.; Galaz-Ávalos, R. M.; De-la-Peña, C.; Alcazar-Magaña, A.; Wrobel, K.; Loyola-Vargas, V. M. Somatic Embryogenesis: Identified Factors That Lead to Embryogenic Repression. A Case of Species of the

5

Same Genus. *PLoS One* **2015**, *10*, e0126414.

(49)     Kouassi, M. K.; Kahia, J.; Kouame, C. N.; Tahi, M. G.; Koffi, E. K. Comparing the Effect of Plant Growth Regulators on Callus and Somatic Embryogenesis Induction in Four Elite *Theobroma Cacao* L. Genotypes. *HortScience* **2017**, *52*, 142–145.

(50)     Garcia, C.; Furtado de Almeida, A. A.; Costa, M.; Britto, D.; Valle, R.; Royaert, S.; Marelli, J. P. Abnormalities in Somatic Embryogenesis Caused by 2,4-D: An Overview. *Plant Cell. Tissue Organ Cult.* **2019**, *137*, 193–212.

(51)     Delbarre, A.; Muller, P.; Imhoff, V.; Guern, J. Comparison of Mechanisms Controlling Uptake and Accumulation of 2,4-Dichlorophenoxy Acetic Acid, Naphthalene-1-Acetic Acid, and Indole-3-Acetic Acid in Suspension-Cultured Tobacco Cells. *Planta* **1996**, *198*, 532–541.

(52)     Anandalakshmi, R.; Pruss, G. J.; Ge, X.; Marathe, R.; Mallory, A. C.; Smith, T. H.; Vance, V. B. A Viral Suppressor of Gene Silencing in Plants. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 13079–13084.

(53)     Béziat, C.; Kleine-Vehn, J.; Feraru, E. Histochemical Staining of β-Glucuronidase and Its Spatial Quantification. In *Plant Hormones: Methods and Protocols*; Kleine-Vehn, J., Sauer, M., Eds.; Springer New York: New York, NY, 2017; pp 73–80.

(54)     Case, D. A. Amber 18. *Univ. California, San Fr.* 2018.

(55)     Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**.

(56)     Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.

(57)     Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

(58)     Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N ·log( N ) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

(59)     Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: II. Parameterization and Validation. *J. Comput. Chem.* **2002**, *23*, 1623–1641.

# Supporting information
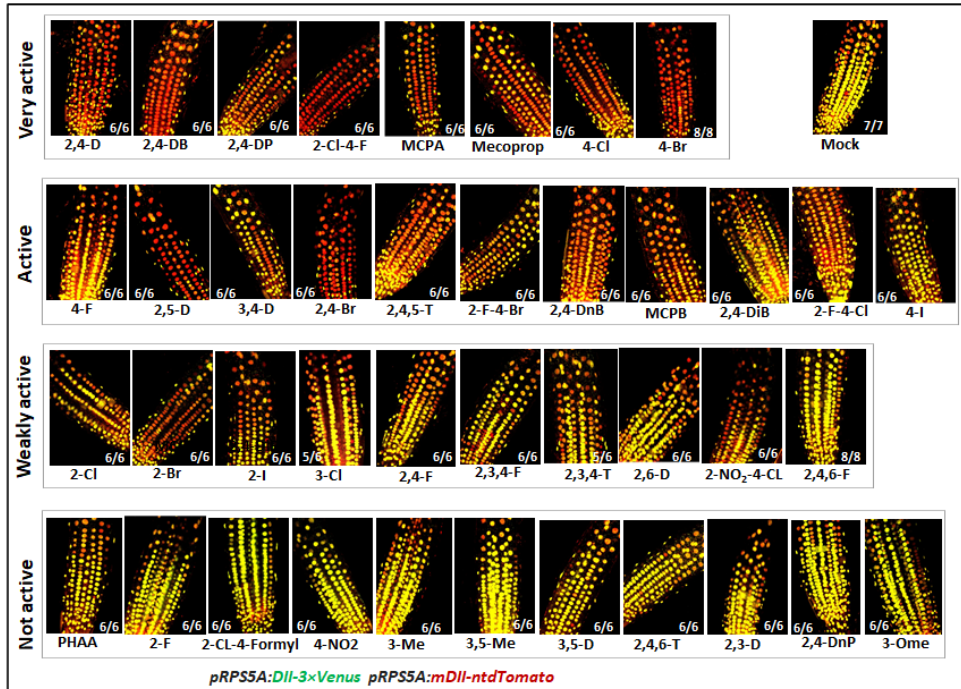


Figure S1. R2D2-reported auxin activity of the 40 2,4-D analogues. Visualization of the auxin activity of different 2,4-D analogues by using the DII-Venus/mDII-tomato (R2D2) reporter in the main root meristem of *Arabidopsis thaliana* seedlings treated with the different 2,4-D analogues. A yellow signal is indicative for no or low auxin activity, whereas an orange or red signal indicates intermediate or high auxin activity, respectively. Seeds were first grown for 6 days on control (mock) medium, and subsequently the seeds were transferred to liquid medium containing 5 µM of the 2,4-D analogues for 1 hour.

5
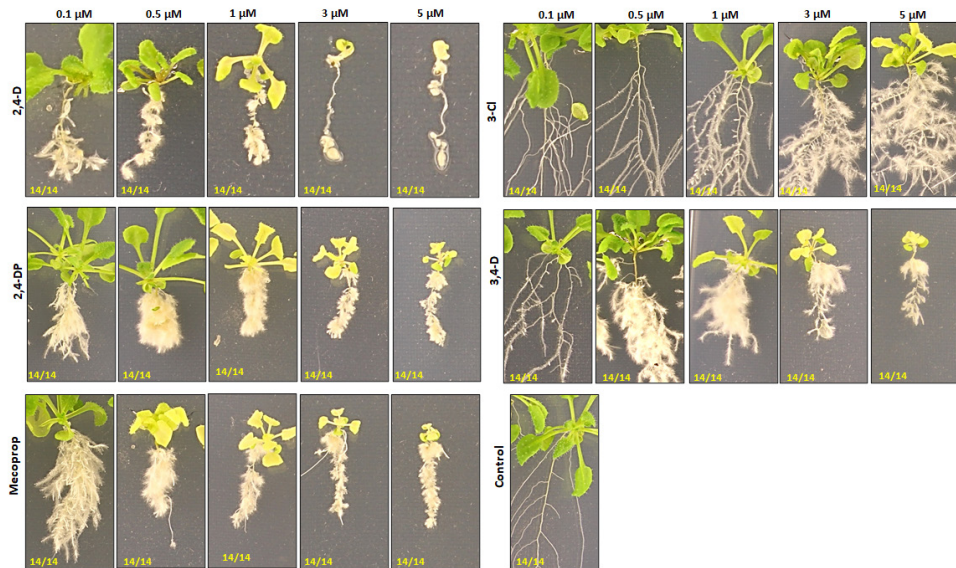
Figure S2. Certain 2,4-D analogues specifically promote root hair formation. The phenotype of Arabidopsis seedlings treated with the indicated concentration of 2,4-D, 2,4-DP, Mecoprop, 3-Cl, or 3,4-D. Seedlings were initially grown for 6 days on control medium and subsequently cultured for 15 days on medium containing the indicated concentration of the 2,4-D analogue.

Figure S3. Weakly or not active 2,4-D analogues do not induce SE. The phenotype of Arabidopsis IZEs that were first grown for two weeks in the presence of 5 µM of the indicated 2,4-D analogue and subsequently cultured for 1 week on medium without any supplement.

5

Hydrogen Bond network

**·MCPA**
·MCPA (O2) – Arginine 403 (NH$_2$) 45%
·MCPA (O2) – Serine 438 (OG) 44%
·MCPA (O2) – Arginine 403 (NE) 40%
·MCPA (O3) – Arginine 403 (NH$_2$) 36%
·MCPA (O3) – Serine 462 (OG) 36%
·MCPA (O3) – Arginine 403 (NE) 26%
·MCPA (O3) – Serine 438 (OG) 15%
·MCPA (O3) – Arginine 436 (NH$_2$) 7%

**·4-D**
- 4-D (O2) – Arginine 403 (NH$_2$) 94%
- 4-D (O3) – Serine 438 (OG) 93%
- 4-D (O3) – Arginine 403 (NE) 64%
- 4-D (O2) – Arginine 403 (NE) 7%
- 4-D (O1) – Serine 438 (OG) 2%

**2,4-D-R**
- 2,4-D-R (O3) – Arginine 403 (NH$_2$) 48%
- 2,4-D-R (O3) – Arginine 403 (NE) 45%
- 2,4-D-R (O2) – Arginine 403 (NH$_2$) 42%
- 2,4-D-R (O2) – Arginine 403 (NE) 31%
- 2,4-D-R (O3) – Arginine 436 (NH$_2$) 17%
- 2,4-D-R (O2) – Serine 438 (OG) 13%
- 2,4-D-R (O2) – Serine 462 (OG) 13%
- 2,4-D-R (O2) – Arginine 436 (NH$_2$) 7%
- 2,4-D-R (O3) – Serine 462 (OG) 3%
- 2,4-D-R (O2) – Arginine 436 (NH$_2$) 2%
- 2,4-D-R (O3) – Serine 438 (OG) 2%
- 2,4-D-R (O3) – Arginine 436 (NH$_2$) 1%

**2,4-D-S**
- 2,4-D-S (O3) – Serine 438 (OG) 54%
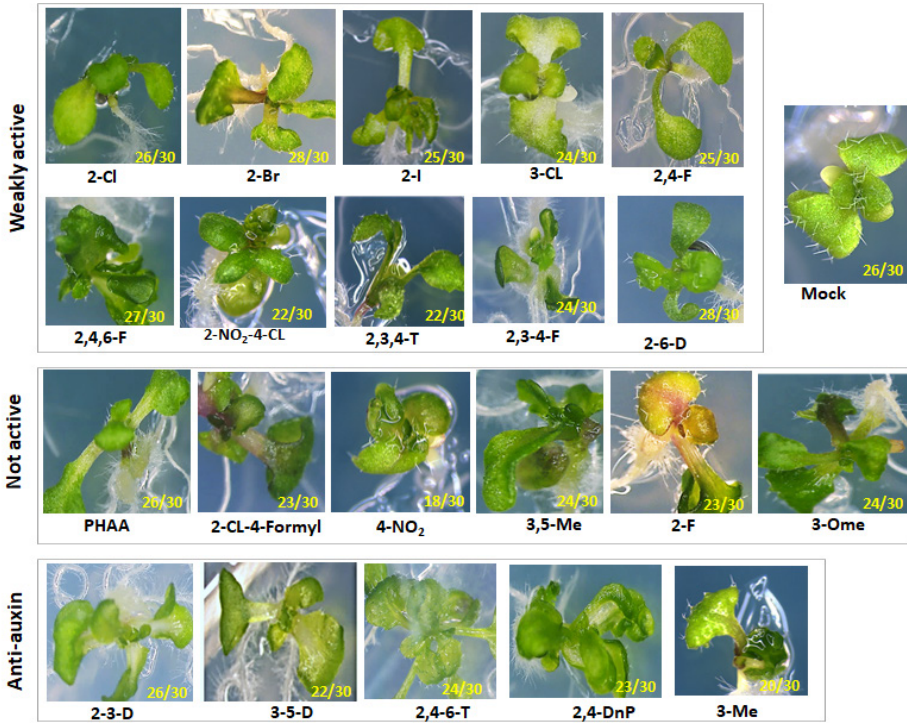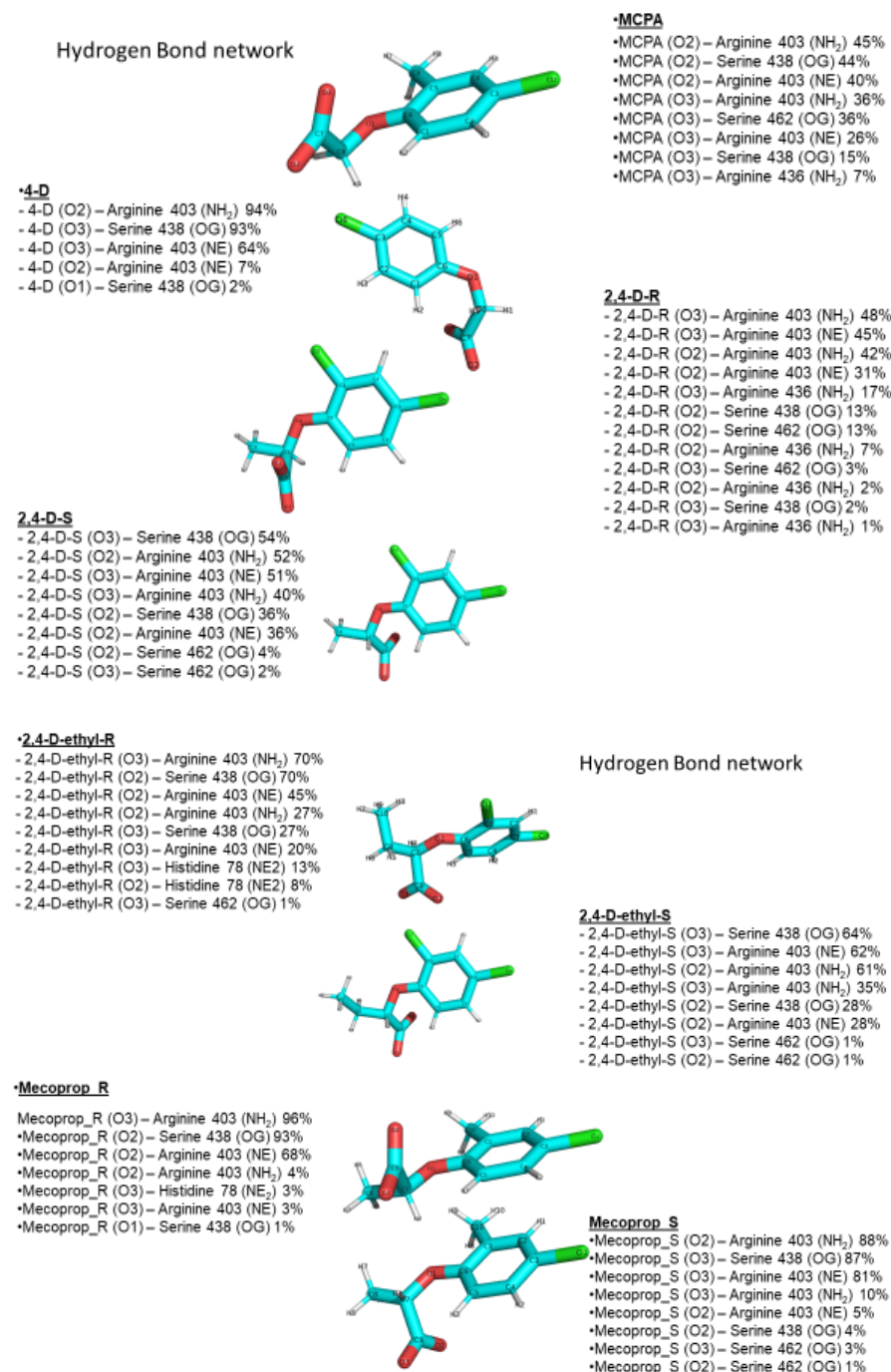- 2,4-D-S (O2) – Arginine 403 (NH$_2$) 52%
- 2,4-D-S (O3) – Arginine 403 (NE) 51%
- 2,4-D-S (O3) – Arginine 403 (NH$_2$) 40%
- 2,4-D-S (O2) – Serine 438 (OG) 36%
- 2,4-D-S (O2) – Arginine 403 (NE) 36%
- 2,4-D-S (O2) – Serine 462 (OG) 4%
- 2,4-D-S (O3) – Serine 462 (OG) 2%

**·2,4-D-ethyl-R**
- 2,4-D-ethyl-R (O3) – Arginine 403 (NH$_2$) 70%
- 2,4-D-ethyl-R (O2) – Serine 438 (OG) 70%
- 2,4-D-ethyl-R (O2) – Arginine 403 (NE) 45%
- 2,4-D-ethyl-R (O2) – Arginine 403 (NH$_2$) 27%
- 2,4-D-ethyl-R (O3) – Serine 438 (OG) 27%
- 2,4-D-ethyl-R (O3) – Arginine 403 (NE) 20%
- 2,4-D-ethyl-R (O3) – Histidine 78 (NE2) 13%
- 2,4-D-ethyl-R (O2) – Histidine 78 (NE2) 8%
- 2,4-D-ethyl-R (O3) – Serine 462 (OG) 1%

Hydrogen Bond network

**2,4-D-ethyl-S**
- 2,4-D-ethyl-S (O3) – Serine 438 (OG) 64%
- 2,4-D-ethyl-S (O3) – Arginine 403 (NE) 62%
- 2,4-D-ethyl-S (O2) – Arginine 403 (NH$_2$) 61%
- 2,4-D-ethyl-S (O3) – Arginine 403 (NH$_2$) 35%
- 2,4-D-ethyl-S (O2) – Serine 438 (OG) 28%
- 2,4-D-ethyl-S (O2) – Arginine 403 (NE) 28%
- 2,4-D-ethyl-S (O3) – Serine 462 (OG) 1%
- 2,4-D-ethyl-S (O2) – Serine 462 (OG) 1%

**·Mecoprop_R**
Mecoprop_R (O3) – Arginine 403 (NH$_2$) 96%
·Mecoprop_R (O2) – Serine 438 (OG) 93%
·Mecoprop_R (O2) – Arginine 403 (NE) 68%
·Mecoprop_R (O2) – Arginine 403 (NH$_2$) 4%
·Mecoprop_R (O3) – Histidine 78 (NE$_2$) 3%
·Mecoprop_R (O3) – Arginine 403 (NE) 3%
·Mecoprop_R (O1) – Serine 438 (OG) 1%

**Mecoprop_S**
·Mecoprop_S (O2) – Arginine 403 (NH$_2$) 88%
·Mecoprop_S (O3) – Serine 438 (OG) 87%
·Mecoprop_S (O3) – Arginine 403 (NE) 81%
·Mecoprop_S (O3) – Arginine 403 (NH$_2$) 10%
·Mecoprop_S (O2) – Arginine 403 (NE) 5%
·Mecoprop_S (O2) – Serine 438 (OG) 4%
·Mecoprop_S (O3) – Serine 462 (OG) 3%
·Mecoprop_S (O2) – Serine 462 (OG) 1%

Figure S4. The hydrogen bond network between the very active 2,4-D analogues and TIR1.

## Commercially available 2,4-D analogues

**2,4-D** (CAS number 94-75-7; Sigma-Aldrich); **3,5-Me** (CAS number 5406-14-4; Sigma-Aldrich); **PHAA** (CAS number 122-59-8; Antonides CV); **PA** (CAS number 103-82-2; Fisher Scientific); **MCPA** (CAS number 94-74-6; Sigma-Aldrich); **MCPB** (CAS number 94-81-5; Sigma Aldrich); **2,4-DnB** (CAS number 94-82-6; Sigma Aldrich); **2-Cl** (CAS number 614-61-9; Fisher Scientific); **2-F** (CAS number 348-10-7; Fisher Scientific); **2-Br** (CAS number 1879-56-7; Fisher Scientific); **3-Cl** (CAS number 588-32-9; Fisher Scientific); **4-Cl** (CAS number 122-88-3; Antonides CV); **4-F** (CAS number 405-79-8; Bio-Connect); **4-Br** (CAS number 1878-91-7; Fisher Scientific); **4-I** (CAS number 1979-94-0; Fisher Scientific); **4-NO$_2$** (CAS number 1798-11-4; Bio-Connect); **3-Me** (CAS number 1643-15-8; Bio-Connect); **3-Ome** (CAS number 2088-24-6; Fisher Scientific); **2,3-D** (CAS number 2976-74-1; Fisher Scientific); **3,4-D** (CAS number 588-22-7; Fisher Scientific); **2-Cl-4-Formyl** (CAS number 52268-20-9; Sigma-Aldrich); **2-Cl-4-F** (CAS number 399-41-7; Bio-Connect); **2,4-F** (CAS number 399-44-0; Bio-Connect); **2,4-Br** (CAS number 10129-78-9; Fisher Scientific); **2,4,5-T** (CAS number 93-76-5; Fisher Scientific); **2,4-DP** (CAS number 120-36-5; Sigma Aldrich); **Mecoprop** (CAS number 93-65-2; Sigma Aldrich); **2,4-DiB** (CAS number 1914-66-5; Bio-Connect); **2,4-DB** (CAS number 6956-86-1; Bio-Connect)

All commercially available compounds were tested for purity with $^1$H NMR and were >95 % pure, except for compound **4-I** which was circa 90% pure.

## Chemical synthesis

General procedure for the synthesis of 2,4-D analogues

The corresponding phenol of the analogue was charged to a roundbottom flask together with potassium hydroxide, bromo-or iodoacetic acid, and water. The reaction mixture was heated at 100 °C for several hours, typically five to six hours. Water was added during the reaction if the reaction mixture dried up or the reaction mixture was diluted before work-up. The reaction mixture was acidified to pH ≤ 1 and extracted with Et$_2$O. The organic layer was washed with brine, dried over Na$_2$SO$_4$, filtered and concentrated *in vacuo*. The crude was separated using flash column chromatography.
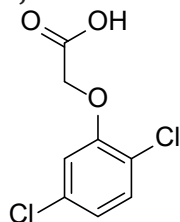
**2-I**
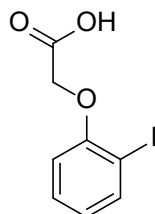


2-iodophenol (580 mg; 2.6 mmol), potassium hydroxide (1.008 g; 18.0 mmol), iodoacetic acid (1.469 g; 7.9 mmol) and water (50 mL) were charged to a roundbottom flask. The reaction mixture was heated at 100 °C for several hours. Water (20 mL) was added during the reaction, because some evaporated. The reaction mixture was diluted with water (20 mL), acidified with 6 M HCl to pH ≤ 1 and extracted with $Et_2O$ (3 x 20 mL). The organic layer was washed with brine (3 x 20 mL). The organic layer was dried over $Na_2SO_4$, filtered and concentrated *in vacuo* to yield a yellow oil. The crude was dry loaded and column chromatography (70 % PE ; 30 % EtOAc; 1 % acetic acid) yielded **2-I** (432 mg; 1.6 mmol; 59 %).

^{1}H NMR (600 MHz, CDCl3) δ 7.81 (dd, J = 7.8, 1.6 Hz, 1H), 7.33 (ddd, J = 8.2, 7.4, 1.6 Hz, 1H), 6.85 – 6.75 (m, 2H), 4.72 (s, 2H).
^{13}C NMR (600 MHz, CDCl3) δ 169.91, 155.75, 139.83, 129.73, 124.27, 112.63, 86.48, 65.90.
ESI HRMS for $C_8H_7IO_3$: m/z [M – H]$^-$ ; calcd: 276,9362; found: 276,9349.

**2,5-D**



2,5-dichlorophenol (0.296 g; 1.8 mmol), potassium hydroxide (3.57 g; 63.6 mmol), bromoacetic acid (3.68 g; 26.5 mmol) and water (50 mL) were charged to a roundbottom flask. The reaction mixture was heated at 100 °C for several hours (3) until it was concentrated. The reaction mixture was diluted with water (50 mL), acidified with 6 M HCl to pH ≤ 1 and extracted with $Et_2O$ (3 x 30 mL). The organic layer was washed with water (6 x 15 mL). The organic layer was dried over $Na_2SO_4$, filtered and concentrated in vacuo. The $Et_2O$ was washed with brine and dried over $Na_2SO_4$, filtered and concentrated *in vacuo*. Column chromatography (70 % PE ; 30 % EtOAc; 1 % acetic acid) yielded **2,5-D** (0.179
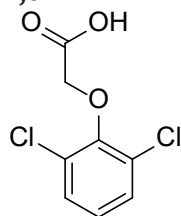
g; 0.84 mmol; 47 %).

^{1}H NMR (600 MHz, DMSO) δ 7.53 (d, J = 8.5 Hz, 1H), 7.23 (d, J = 2.3 Hz, 1H), 7.11 (dd, J = 8.5, 2.3 Hz, 1H), 4.94 (s, 2H).
^{13}C NMR (600 MHz, DMSO) δ 169.91, 154.33, 132.67, 131.50, 122.16, 120.66, 114.55, 65.68.
ESI HRMS for $C_8H_6Cl_2O_3$: m/z [M − H]⁻ ; calcd: 218,9616; found: 218,9609.
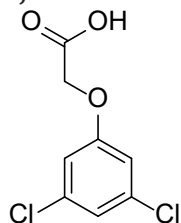
**2,6-D**



2,6-dichlorophenol (0.315 g; 1.9 mmol), potassium hydroxide (3.7 g; 65.9 mmol), bromoacetic acid (3.86 g; 27.8 mmol) and water (50 mL) were charged to a roundbottom flask. The reaction mixture was heated at 100 °C for several hours (6). The reaction mixture was diluted with water (25 mL), acidified with 6 M HCl to pH ≤ 1 and extracted with $Et_2O$ (3 x 25 mL). The organic layer was washed with brine (3 x 25 mL). The organic layer was dried over $Na_2SO_4$, filtered and concentrated *in vacuo*. Column chromatography (70 % PE ; 30 % EtOAc; 1 % acetic acid) yielded **2,6-D** (0.069 g; 0.31 mmol; 16% ).

^{1}H NMR (600 MHz, CDCl3) δ 7.34 (d, J = 8.1 Hz, 2H), 7.09 (t, J = 8.1 Hz, 1H), 4.69 (s, 2H).
^{13}C NMR (600 MHz, DMSO) δ 169.30, 150.32, 129.86, 128.71, 126.87, 69.45.
ESI HRMS for $C_8H_6Cl_2O_3$: m/z [M − H]⁻ ; calcd: 218,9616; found: 218,9603.

**3,5-D**



3,5-dichlorophenol (0.303 g; 1.9 mmol), potassium hydroxide (3.65 g; 65.1 mmol), bromoacetic acid (3.91 g; 28.1 mmol) and water (50 mL) were charged to a roundbottom flask. The reaction mixture was heated at 100 °C for several hours (5.5). The reaction mixture was diluted with water (25 mL), acidified with 6 M HCl to pH ≤ 1 and extracted with $Et_2O$ (3 x 25 mL). The organic layer was
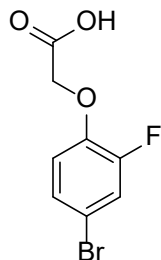
washed with brine (3 x 25 mL). The organic layer was dried over $Na_2SO_4$, filtered and concentrated *in vacuo*. The organic layer was dried over $Na_2SO_4$, filtered and concentrated *in vacuo*. Column chromatography (70 % PE ; 30 % EtOAc; 1 % acetic acid) yielded **3,5-D** (0.199 g; 0.9 mmol; 48% ).

$^1$H NMR (600 MHz, CDCl3) δ 7.04 (s, 1H), 6.83 (s, 2H), 4.68 (s, 2H).
$^{13}$C NMR (600 MHz, DMSO) δ 170.05, 159.70, 134.98, 121.23, 114.50, 65.37.
ESI HRMS for $C_8H_6Cl_2O_3$: m/z [M − H]$^-$ ; calcd: 218,9616; found: 218,9608.
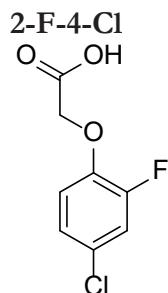
**2-F-4-Br**



2-fluoro-4-bromophenol (0.508 g; 2.7 mmol), potassium hydroxide (1.051 g; 18.7 mmol), iodoacetic acid (1.441 g; 7.7 mmol) and water (50 mL) were charged to a roundbottom flask. The reaction mixture was heated at 100 °C for several hours. Water (20 mL) was added during the reaction and before work-up (10mL). The reaction mixture was acidified with 6 M HCl to pH ≤ 1 and extracted with $Et_2O$ (3 x 25 mL). The organic layer was washed with brine (3 x 25 mL). The organic layer was dried over $Na_2SO_4$, filtered and concentrated *in vacuo*. Column chromatography (70 % PE ; 30 % EtOAc; 1 % acetic acid) yielded **2-F-4-Br** (0.441 g; 1.8 mmol; 67 %).

$^1$H NMR (600 MHz, DMSO) δ 13.14 (s, 1H), 7.55 (dd, J = 10.9, 2.4 Hz, 1H), 7.32 (ddd, J = 8.8, 2.4, 1.5 Hz, 1H), 7.07 (t, J = 9.0 Hz, 1H), 4.80 (s, 2H).
$^{13}$C NMR (600 MHz, DMSO) δ 170.05, 152.80, 151.15, 145.83, 145.76, 127.92, 127.89, 119.90, 119.75, 117.02, 117.01, 112.00, 111.95, 65.55.
$^{19}$F NMR (565 MHz, DMSO) δ -131.29 (t, J = 10.1 Hz).
ESI HRMS for $C_8H_6BrFO_3$: m/z [M − H]$^-$ ; calcd: 246,9406; found: 246,9394.
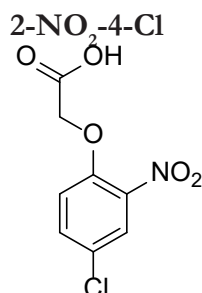
**2-F-4-Cl**



2-fluoro-4-chlorophenol (491 mg; 3.4 mmol), potassium hydroxide (1.34 g; 24.0 mmol), iodoacetic acid (1.90 g; 10.2 mmol) and water (50 mL) were charged to a roundbottom flask. The reaction mixture was heated at 100 °C for several hours. Water (20 mL) was added during the reaction, because some evaporated. The reaction mixture was acidified with 6 M HCl to pH 3 and a white solid was observed. The aqueous layer was extracted with $Et_2O$ (3 x 30 mL). The organic layer was washed with brine (3 x 30 mL) while the white solid remained in the organic layer. The organic layer was dried over $Na_2SO_4$, filtered and concentrated *in vacuo*. The crude was again dissolved in $Et_2O$ and washed with water (3 x 30 mL). Column chromatography (70 % PE ; 30 % EtOAc; 1 % acetic acid) yielded **2-F-4-Cl** (78 mg; 0.4 mmol; 20 %).

^{1}H NMR (600 MHz, CDCl3) δ 7.15 (dd, J = 10.8, 2.5 Hz, 1H), 7.06 (ddd, J = 8.8, 2.5, 1.6 Hz, 1H), 6.91 (t, J = 8.8 Hz, 1H), 4.74 (s, 2H).
^{13}C NMR (600 MHz, CDCl3) δ 172.65, 153.43, 151.77, 144.37, 127.57, 124.51, 124.48, 117.66, 117.51, 117.18, 66.38.
^{19}F NMR (400 MHz, CDCl3) δ -129.94 (t, J = 9.7 Hz).
ESI HRMS for $C_8H_6ClFO_3$: m/z [M − H]⁻ ; calcd: 202,9911; found: 202,9906.

**2-NO₂-4-Cl**



2-nitro-4-chlorophenol (0.311 g; 1.8 mmol), potassium hydroxide (0.714 g; 12.7 mmol), iodoacetic acid (0.967 g; 5.2 mmol) and water (50 mL) were charged to a roundbottom flask. The reaction mixture was heated at 100 °C for several hours. Water was added during the reaction (3 x 10 mL) and before work-up (10mL).
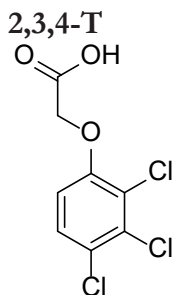
The reaction mixture was acidified with 6 M HCl to pH ≤ 1 and extracted with Et$_2$O (3 x 25 mL). The organic layer was washed with brine (4 x 20 mL). The organic layer was dried over Na$_2$SO$_4$, filtered and concentrated *in vacuo*. Column chromatography (70 % PE ; 30 % EtOAc; 1 % acetic acid; switched to 100 % EtOAc; 1 % acetic acid) yielded **2-NO$_2$-4-Cl** (0.093 g; 0.4 mmol; 22 %).

^{1}H NMR (600 MHz, DMSO) δ 8.04 (d, J = 2.7 Hz, 1H), 7.70 (dd, J = 9.1, 2.7 Hz, 1H), 7.32 (d, J = 9.1 Hz, 1H), 4.93 (s, 2H).
^{13}C NMR (600 MHz, DMSO) δ 169.61, 149.77, 140.49, 134.03, 124.90, 124.87, 117.47, 66.08.
ESI HRMS for C$_8$H$_6$ClNO$_5$: m/z [M – H]$^-$ ; calcd: 229,9856; found: 229,9851.
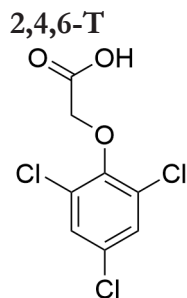
**2,3,4-T**



2,3,4-trichlorophenol (0.051 g; 0.3 mmol),  potassium hydroxide (0.099 g; 1.8 mmol), iodoacetic acid (0.174 g; 0.9 mmol) and water (10 mL) were charged to a roundbottom flask. The reaction mixture was heated at 100 °C for several hours. Water was added during the reaction (4 x 10 mL). The reaction mixture was acidified with 6 M HCl to pH ≤ 1 and extracted with Et$_2$O (3 x 20 mL). The organic layer was washed with brine (2 x 20 mL). The organic layer was dried over Na$_2$SO$_4$, filtered and concentrated *in vacuo*. Column chromatography (70 % PE ; 30 % EtOAc; 1 % acetic acid) yielded **2,3,4-T** (0.031 g; 0.1 mmol; 47 %). NMR analysis demonstrated that the compound, most likely, contained grease impurities.

^{1}H NMR (600 MHz, MeOD) δ 7.42 (d, J = 9.1 Hz, 1H), 6.98 (d, J = 9.1 Hz, 1H), 4.81 (s, 2H).
^{13}C NMR (600 MHz, MeOD) δ 176.58/170.09, 153.88, 131.69, 127.86, 125.24, 123.01, 112.15, 70.18/68.88/65.47/63.01.
ESI HRMS for C$_8$H$_5$Cl$_3$O$_3$: m/z [M – H]$^-$ ; calcd: 252,9226; found: 252,9211.

**2,4,6-T**



2,4,6-trichlorophenol (0.346 g; 1.8 mmol), potassium hydroxide (3.5 g; 62.4 mmol), bromoacetic acid (3.6 g; 25.9 mmol) and water (50 mL) were charged to a roundbottom flask. The reaction mixture was heated at 100 °C overnight. Water was added before the work-up (50mL). The reaction mixture was acidified with 6 M HCl to pH ≤ 1 and extracted with $Et_2O$ (3 x 25 mL). The organic layer was washed with brine (3 x 25 mL). The organic layer was dried over $Na_2SO_4$, filtered and concentrated *in vacuo*. Column chromatography (70 % PE ; 30 % EtOAc; 1 % acetic acid) yielded **2,4,6-T** (0.055 g; 0.2 mmol; 12 %).

$^1$H NMR (600 MHz, DMSO) δ 7.72 (s, 2H), 4.59 (s, 2H).
$^{13}$C NMR (600 MHz, DMSO) δ 169.19, 149.71, 129.67, 129.57, 129.43, 69.50.
ESI HRMS for $C_8H_5Cl_3O_3$: m/z [M − H]$^-$ ; calcd: 252,9226; found: 252,9211.

**2,3,4-F**
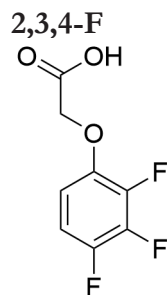


2,3,4-trifluorophenol (300 mg; 2.0 mmol), potassium hydroxide (0.928 g; 16.5 mmol), iodoacetic acid (1.139 g; 6.1 mmol) and water (40 mL) were charged to a roundbottom flask. The reaction mixture was heated at 100 °C for several hours. Water (10 mL) was added during the reaction, because some evaporated. The reaction mixture was acidified with 6 M HCl to pH ≤ 1 and extracted with $Et_2O$ (3 x 20 mL). The organic layer was washed with brine (3 x 20 mL). The organic layer was dried over $Na_2SO_4$, filtered and concentrated *in vacuo*. The crude was dry loaded and column chromatography (70 % PE ; 30 % EtOAc; 1 % acetic acid) yielded **2,3,4-F** (106 mg; 0.51 mmol; 25 %).
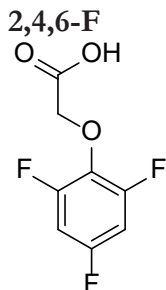
^{1}H NMR (600 MHz, MeOD) δ 6.91 (tdd, J = 10.2, 8.3, 2.5 Hz, 1H), 6.76 (tdd, J = 9.1, 4.5, 2.5 Hz, 1H), 4.80 (s, 3H).
^{13}C NMR (600 MHz, MeOD) δ 170.32, 110.24, 110.21, 110.11, 110.08, 108.97, 108.92, 65.82.
^{19}F NMR (400 MHz, CDCl3) δ -141.53 (m), -152.26 (m), -157.15 – -157.33 (m).
ESI HRMS for $C_8H_5F_3O_3$: m/z $[M – H]^-$ ; calcd: 205,0113; found: 205,0105.

## 2,4,6-F

O  OH
O
F       F

F

2,4,6-trifluorophenol (280 mg; 1.9 mmol), potassium hydroxide (1.1 g; 19.6 mmol), iodoacetic acid (1.306 g; 7.0 mmol) and water (30 mL) were charged to a roundbottom flask. The reaction mixture was heated at 100 °C for several hours. Water (10 mL) was added during the reaction, because some evaporated. The reaction mixture was acidified with 6 M HCl to pH ≤ 1 and extracted with $Et_2O$ (3 x 20 mL). The organic layer was washed with brine (3 x 20 mL). The organic layer was dried over $Na_2SO_4$, filtered and concentrated *in vacuo* to yield a yellow oil. Column chromatography (70 % PE ; 30 % EtOAc; 1 % acetic acid) yielded **2,4,6-F** (78 mg; 0.4 mmol; 20 %).

^{1}H NMR (400 MHz, CDCl3) δ 6.78 – 6.65 (m, 1H), 4.75 (d, J = 0.9 Hz, 2H).
^{13}C NMR (400 MHz, CDCl3) δ 172.43, 101.28, 101.01, 100.74, 69.48.
^{19}F NMR (400 MHz, CDCl3) δ -113.28 – -113.48 (m), -124.61 (dd, J = 8.6, 2.9 Hz).
ESI HRMS for $C_8H_5F_3O_3$: m/z $[M – H]^-$ ; calcd: 205,0113; found: 205,0101.

## 2,4-DnP

OH
O
O
Cl

Cl
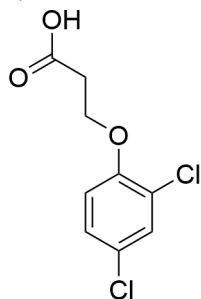
2,4-dichlorophenol (495 mg; 3.0 mmol), potassium hydroxide (1.187 g; 21.2

mmol), bromopropanoic acid (1.354 g; 8.9 mmol) and water (50 mL) were charged to a roundbottom flask. The reaction mixture was heated at 100 °C for several hours. The reaction mixture was diluted with water (20 mL), acidified with 6 M HCl to pH ≤ 1 and extracted with $Et_2O$ (3 x 20 mL). The organic layer was washed with brine (3 x 20 mL). The organic layer was dried over $Na_2SO_4$, filtered and concentrated *in vacuo* to yield a yellow oil. Column chromatography (70 % PE ; 30 % EtOAc; 1 % acetic acid) yielded **2,4-DnP** (136 mg; 0.6 mmol; 20 %).

^{1}H NMR (600 MHz, CDCl3) δ 7.36 (d, J = 2.5 Hz, 1H), 7.18 (dd, J = 8.7, 2.5 Hz, 1H), 6.88 (d, J = 8.8 Hz, 1H), 4.29 (t, J = 6.3 Hz, 2H), 2.92 (t, J = 6.3 Hz, 2H).
^{13}C NMR (600 MHz, CDCl3) δ 175.97, 152.92, 130.13, 127.63, 126.47, 124.16, 114.73, 64.79, 34.11.
ESI HRMS for $C_9H_8Cl_2O_3$: m/z [M − H]⁻ ; calcd: 232,9772; found: 292,9764.

5

# Chapter 6

General discussion and future prospects

Chapter 6

# Chemical biology

Chemical biology spans the interdisciplinary field between chemistry and biology, and aims to study and manipulate biological processes at the molecular level, often with the use of small or tailor-made molecules. This is the definition we gave at the beginning of the research described in this thesis, but this definition is not all encompassing as testified by the many different examples of the definition given by researchers with various scientific backgrounds[1]. Notwithstanding that chemical biology is a broad scientific discipline, it has seen tremendous growth through the last decades. Some examples include small molecule screens in plant biology[2–4], directed evolution of proteins[5,6] and bioorthogonal click chemistry[7,8]. The discovery and application of bioorthogonal click chemistry has set in motion a whole new field[9], with outstanding applications in other fields such as glycobiology, as exemplified by the glycans visualized during the development of a zebrafish embryo[10].

In this final chapter we discuss the two different chemical biology approaches carried out in the research that was described in **Chapter 2**, **3** and **4**, and **Chapter 5**, further discuss specific results from those studies and conclude with future prospects.

## Glycoengineering techniques to study bacterial glycans

Glycan interactions between the host and bacteria are important for cell-cell interactions, bacterial pathogenesis and interactions with surrounding molecules. Different approaches exist to study bacterial glycans. Genetic modification is an option, but this approach is limited due to the non-templated glycan biosynthesis by the cell's glycosylation machinery. An alternative is to use isotope- or radiolabeled glycans, but these often have practical limitations such as safety and cost due to the handling of radioactive material or elaborate organic synthesis to incorporate the label into the glycan probe[11]. In recent years, there has been an emergence of chemical biology approaches that make it possible to install reporter groups on bacteria and their glycans, which provide a unique opportunity to study bacterial glycoscience. In an approach dubbed glycoengineering (also called glyco-editing), glycans can be inserted, removed or modified. For which the ultimate goal would be to enable glycoengineering at high efficiency with spatiotemporal control and molecular-level precision on live bacteria.

Several glycoengineering approaches exists to specifically study bacterial glycans. A selection of these chemical biology approaches is discussed in **Chapter 2**. Besides the technique metabolic oligosaccharide engineering (MOE), which was discussed in **Chapters 2** and **4** of this thesis, other techniques include modification of the mycobacterial cell wall, the heptasaccharide of *Campylobacter jejuni* with glycans containing a reporter group or through the modification of existing glycans by a galactosidase[12–16]. These pioneering studies of bacterial glycoengineering increased our understanding of the involved bacterial glycans and how to manipulate them, yet also only represent a starting point in unraveling the diverse world of bacterial glycans[17].

There is a need to expand the tools available to glycoengineer bacteria. In **Chapter 2**, a selection of the tools from chemical biology that currently exist was presented together with a discussion on what other tools would be useful. Further on in this thesis, we have described two such approaches; the technique SEEL was adapted for bacterial glycans (**Chapter 3**) and the power of native sialyltransferases was explored to incorporate a reporter group (**Chapter 4**). These techniques are complementary based on the goal of glycoengineering and the chosen target. For instance, targeting unique bacterial monosaccharides for antibiotic strategies is an often-found goal for MOE[13]. Whereas the introduction of a large biomolecule on a precise glycan on the cell surface in a single step can thus far only be achieved through SEEL[18]. In other words, SEEL can introduce unnatural sialosides. The possible downstream effects of this modification could be studied in a biological system, for instance the binding of factor H that recognizes glycans as 'self'[19,20]. In a potential future application, SEEL can be applied to other bacteria containing a *N*-acetyllactosamine (LacNAc), which is the acceptor specificity of the exogenous enzyme ST6Gal1[21]. The presence of LacNAc is expected for certain bacteria like *Haemophilus ducreyi* or *Neisseria meningitidis*[22,23], but might also be detected through SEEL on other bacteria from the gut-microbiome for example, with unidentified glycan structures.

From a chemical point of view, the scope of both SEEL and native sialyltransferases can be expanded with additional sugar nucleotide derivatives of monosaccharides or reporter groups. An important target would be fucose, which has been described in other chemoenzymatic approaches and plays a key role in the gut at the host-microbe interface[24,25]. Several bacteria display fucose on their cell surface glycoconjugates, like the O-antigen of certain *Escherichia coli* strains or *Helicobacter pylori*, which would form interesting targets of SEEL with a

6

reported fucosyltransferase from *H. pylori* origin and GDP-fucose analogs (Figure 1)[24,26–28]. The glycoengineering techniques SEEL and native glycosyltransferase could also benefit from additional reporter groups. Reporter groups that have been successfully employed to study glycans are cyclopropene or sydnone among others[29,30]. These would provide opportunity for dual labeling, as exemplified by the study from the Kasper group for which multiple bacterial components were labeled[31]. The fluorescent labeling of bacterial glycan structures enables tracking them in the gut or to study glycan-based host-microbe interactions.
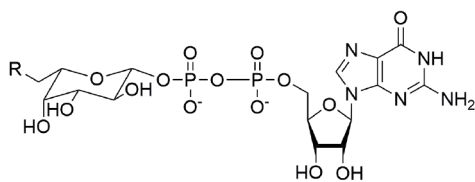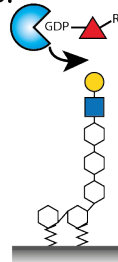


Figure 1. Possible expansion of the SEEL technique to another nucleotide sugar analog, fucose. (A) GDP-fucose analog containing an azide or fluorescent dye as a reporter group. (B) Schematic figure of the fucosylation by an exogenous enzyme on the cell surface of a bacteria with a GDP-fucose analog.

A biological question that could be addressed through a chemical biology approach is to learn more about the location of native sialyltransferases or the uptake of sugar nucleotides by bacteria. In case the glycosyltransferases are surface-bound, like for *N. gonorrhoeae*[32], this could be demonstrated with antibodies against glycosyltransferases or possibly with new probes that visualize these enzymes on the cell surface. Alternatively, the sugar nucleotides are taken up by the bacteria and then it would be interesting to identify the uptake mechanism, and the origin of nucleotide sugars[33,34].

Thus far, much remains unknown about bacterial glycans. Bacterial mono-saccharide diversity is much larger compared to eukaryotes, which provides an opportunity to target unique bacterial glycans with antibiotic strategies[13], but also raises the fundamental question if these glycans are unique to bacterial pathogens. In line with the interest to understand the pathogenicity of bacteria, much more is known about the glycan structures of pathogens than for commensals[35]. However, the vast number of bacteria of the human microbiome are considered commensal. The ascribed potential health benefits urge the scientific community to dig deeper into the glycan structures of these commensals. An interesting approach to identify the bacterial glycome, or fingerprint, is to glycophenotype bacteria.

## Small molecules in plant embryogenesis

### The synthetic auxin 2,4-D in somatic embryogenesis of *Arabidopsis thaliana*

In the second study employing a chemical biology approach, described in this thesis, a set of small molecules was tested for the induction of somatic embryogenesis (SE). In **Chapter 5** we demonstrated that the halogen on the 4-position of the aromatic ring is important for the structure-activity relationship (SAR) of the synthetic auxin 2,4-D. Although the tested set contained a logical variation of 2,4-D analogues on the aromatic ring or on the alpha position of the carboxylate, this library can be expanded, which is discussed in this chapter. One group that could be varied is the carboxylic acid of 2,4-D (Figure 2A). Instead of the carboxylic acid, other groups such as a tetrazole or sulfonamide (Figure 2B), can be synthesized and tested[36]. These groups are somewhat similar to carboxylic acid, but have slightly different properties. The tetrazole has a similar acidity (pKa ≈ 4.5-4.9 versus ≈ 3 for 2,4-D), but is bulkier compared to the carboxylic acid, and the sulfonamide is less acidic (pKa ≈ 10), but can establish the same hydrogen bonding pattern[37]. These analogues could have the capacity to distinguish which chemical properties of the carboxylic acid are important for biological activity, and thus further completing the SAR.
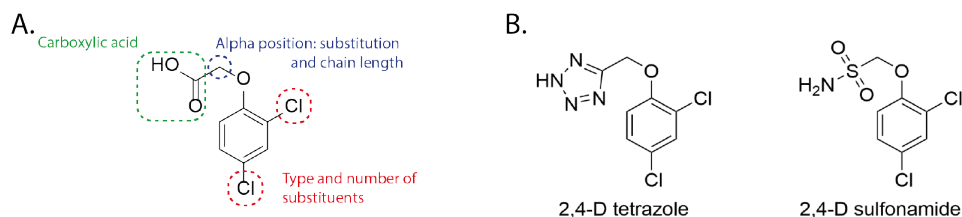


Figure 2. Analogues of 2,4-D. (A) Schematic indication of the different positions that can be varied to create analogues of 2,4-D. (B) A tetrazole analogue and sulfonamide analogue of 2,4-D which are both isosteres of the carboxylic acid.

From the SAR in **Chapter 5** it became apparent that the auxin activity can be linked to the capacity to induce SE by the 2,4-D analogues. To further understand this process at the molecular level, *in vitro* studies can be performed. In the past, several studies have shown the binding of the naturally occurring auxin indole-3-acetic acid (IAA) or 2,4-D, often to the TIR1 receptor, through in-gel binding, surface plasmon resonance or a yeast two-hybrid system[38–40]. It would be interesting to discover which combinations of F-box receptors and AUX/IAA analogues are responsible for the initiation of SE in future work.

## Chapter 6

## Small molecules that enhance microspore embryogenesis of microspores of *Brassica napus*

Embryogenesis in plants can occur via multiple developmental processes[41]. One of these processes is the previously discussed somatic embryogenesis. Another type of *in vitro* embryogenesis is microspore embryogenesis (ME), in which cultured immature male gametophytes (microspores or pollen) can be induced to form embryos. The switch from the gametophytic pathway to the embryo pathway is achieved by applying stress, typically in the form of a heat-shock[42]. Model species are used to further disseminate the developmental process of microspore embryogenesis at the molecular level and one of these models is *Brassica napus*[42].

Similar to somatic embryogenesis, we investigated if 2,4-D or 2,4-D analogues could increase the embryo yield for ME in *B. napus*. Since ME requires a stress treatment and 2,4-D is also associated with causing stress[43], the application of 2,4-D analogues in combination with a heat-shock might have the potential to increase the embryo yield.

The 2,4-D library that was used in **Chapter 5** was also tested for enhancement of ME, for which the preliminary results are described in this chapter. The following compounds were found to increase the number of embryo's: 4-I, 3,5-Me, 3-OMe, 2,3,4-T, 2,4,5-T, 2-F-4-Br, 4-Cl PA, and 2,4-D Tetrazole. An analogue was classified as active when the embryo yield increased in three or more tests for which two hits are in the same pollen stage. In comparison to somatic embryogenesis, the hits appear structurally more diverse. Some hits with modifications on the 4-position, like 4-I, are active in both systems. However, for ME, electron-donating groups on the 3- or 5- position (3,5-Me and 3-OMe) are hits, which are groups not active in SE. In addition, the acidity of the carboxylic acid seems important, because the tetrazole analogue of 2,4-D showed activity, but the sulfonamide analogue did not. Although several observations could be made, the structure-activity relationship of 2,4-D in ME is not as straightforward as for SE.

From the screen, there are no distinct structural leads for compounds that increase the yield for microspore embryogenesis. 2,4-D might have a different mode of action in ME than in SE, for which it does induce embryogenesis. A possible hypothesis could be that 2,4-D does induce the switch to totipotency in somatic, but not in microspore embryogenesis. For ME, it was reported that only

chemical induction by 2,4-D, without heat-shock, did not trigger microspore embryogenesis[44].

## Structure-activity relationship of the microspore embryogenesis enhancer C26

In an attempt to look for another chemical inducer of ME, a commercial library of compounds was screened for the enhancement of microspore embryogenesis by the group of Plant Development Systems from Wageningen University. From this library the compound 'C26' (2,4-dichloro-6-[(3-pyridinylimino)methyl] phenol) was identified as the most potent hit. The structure of C26 resembles the core of another compound that was identified as an auxin transport inhibitor (Figure 3A). To further study the mechanism of action by which C26 enhances microspore embryogenesis, we established a structure-activity relationship (SAR) study of C26 to identify the minimal molecular structure associated with enhanced microspore embryogenesis. The as of yet unpublished results of this SAR are discussed in this chapter.

According to our hypothesis, the imine present in C26 can be hydrolysed in microspore cultures, giving rise to the compounds 3-aminopyridine ('C54') and 3,5-dichlorosalicylaldehyde ('C56') (Figure 3B). As expected, nuclear magnetic resonance (NMR) analysis of C26 revealed this imine hydrolysis occurred over time in wet DMSO (Figure 3C). Furthermore, the benzoic-imine bond of C26 is sensitive to acidic conditions[45], which may further increase the hydrolysis kinetics of C26 in a cell compartment with low pH. Given that microspores were cultured in NLN-13 liquid medium at a pH of 5.8, we assumed hydrolysis would occur within certain cellular compartments.

Our SAR data revealed C26 as the most active compound with the highest embryo yield among all compounds tested in the SAR study, even compared to its predicted hydrolysis products C54 and C56 (Figure 3D). As C56 showed a significant increase in final embryo yield compared to the control, it could be suggested that C56 represents the minimal structure of the C26 molecule responsible for the enhancing activity of C26.

Based on the SAR study results, we hypothesized that the complete structure of C26 is needed to enhance microspore embryogenesis: its relatively hydrophobic character may allow for easier passage through the plasma membrane than the hydrolysis products C54 and C56. Following the cellular uptake of C26,

6

hydrolysis of C26 could occur, leading to the formation and subsequent binding of one of its hydrolysis products to its target. Alternatively, the complete C26 structure may be needed to bind to its target inside the cells. To investigate this further, we tested a compound for which the imine bond was reduced into a hydrolytically stable amine (C26-reduced). We tested this compound for embryo induction and found an embryo yield comparable to the control. This result may point to a possible different metabolic processing of C26 and C26-reduced and supports the hypothesis that C26 undergoes hydrolysis after cellular uptake.

Since the hydrolysis product C56 showed a higher embryo yield than C54, we first sought to make several modifications to the phenol ring of C26. Other dihalogens, like iodo (C59) and bromo (C60), yielded fewer embryos compared to the control. In addition, compounds with a single halogen group, a chloro (C58) or bromo (C57), also resulted in lower embryo yield when compared to C26.

Next, we wanted to test if modifications to the active hydrolysis product C56 may change the final activity. We exchanged one of the chloro groups for a fluoro group (C53). Given that the fluor atom is smaller and is a more electron-withdrawing group than the chlorine atom, we expected that this substitution would result in an active compound that is more efficient in binding to its potential target. However, our SAR study showed comparable activity to the hydrolysis product C56, suggesting that this halogen group substitution on the fourth position of the phenol ring does not change the final activity of the compound.

Besides modifications to the phenol ring of C26, we also examined if the phenol and the aldehyde groups of the hydrolysis products are important for the activity. C61, a compound containing the dichloro groups and the imine linkage but lacking a phenol group, did not enhance the final embryo yield. Furthermore, replacing the aldehyde with carboxylic acid (C55) did not result in enhanced embryogenesis. Hence, these results suggest that both the phenol and the aldehyde group are required for activity.

Taken together, we conclude that the following chemical moieties contribute to the activity of compound C26: the dichloro motive on the ring, the phenol, and the aldehyde from the hydrolysis. Furthermore, we speculate that the imine linkage is essential for the metabolic processing or cellular uptake of the compound.
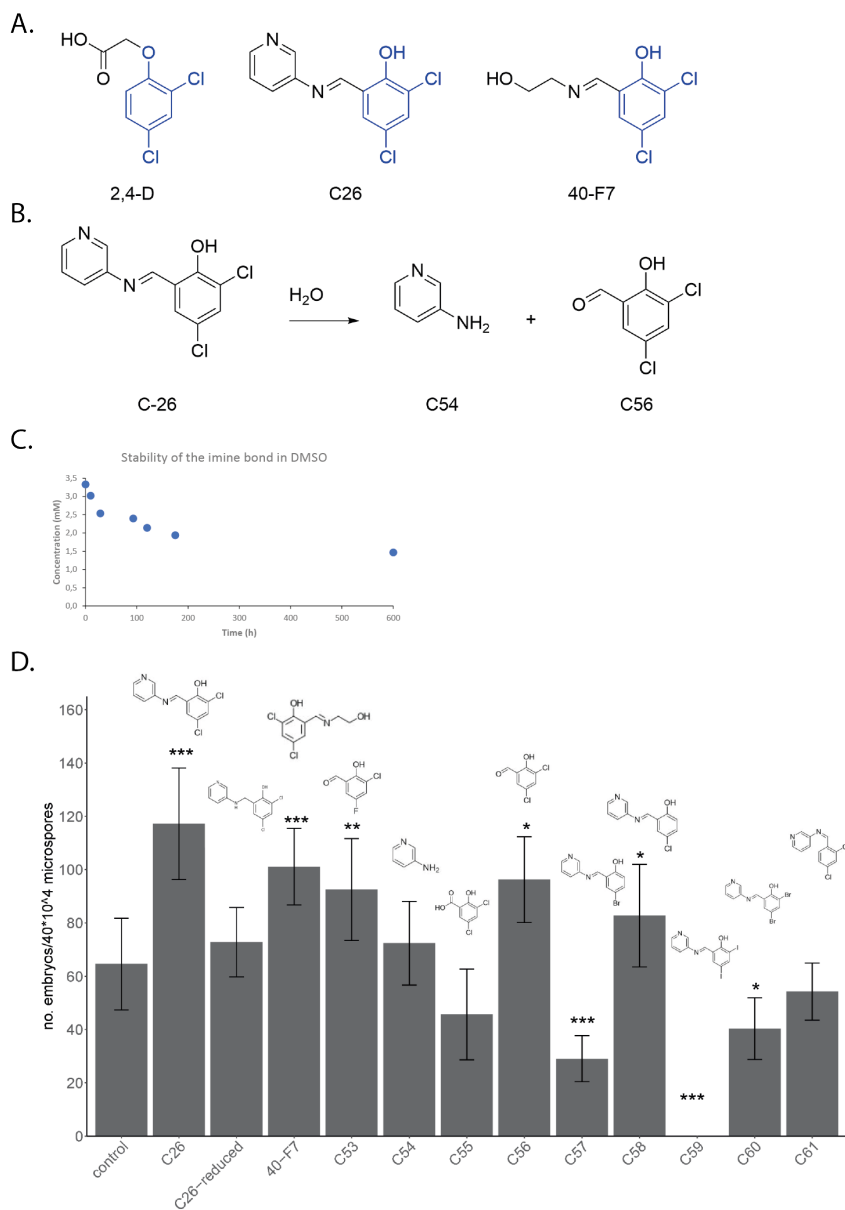
Figure 3. Structure-activity relationship (SAR) study for C26 enhancer of microspore embryogenesis. (A) Molecular structures of 2,4-D, C26, and 40-F7. Both C26 and 40-F7 share the same dichlorophenoxy core structure as 2,4-D (indicated in blue) and an imine bond. (B) Predicted hydrolysis of C26 in culture. Hydrolysis results in the formation of C54 and C56. (C) Stability of the imine bond over time measured in DMSO-d6 with NMR. (D) Embryo yield of compounds tested in the SAR study. Compound names are indicated below the bars and their molecular structures above the bars. Mean values of three biological replicates (± SD) with three technical replicates are shown. Statistical analysis was performed by comparing the control with the compound treatments using one-way ANOVA, followed by a post hoc Dunnett's test. * = $p<0.05$, ** = $p<0.01$, *** = $p< 0.001$.

# Chapter 6

## General trends of somatic and microspore embryogenesis

Both somatic and microspore embryogenesis can be initiated by different small molecules. Their effect might be linked to the intracellular level of auxin-like analogues[46]. Intracellular trafficking of auxins takes place through passive and active transport. The active transport of auxins is regulated by several classes of transport proteins: PIN-formed (PIN), P-glycoprotein (ABCB/PGP), and AUXIN RESISTANT 1/LIKE AUX1 (AUX1/LAX)[47]. The first two being efflux carriers and the latter an uptake symporter. The efflux carrier PIN can be inhibited by naphthylphthalamic acid (NPA)[48]. Inhibition of these transporters can lead to increased intracellular levels of auxins. When we tested a combination of IAA and NPA on immature zygotic embryos of *Arabidopsis thaliana,* the explant grew somatic embryos (data not shown). In an attempt to further improve this process, a set of NPA analogues was synthesized (Figure 4) for which the synthesis and results are described in this chapter. For the NPA analogues, it was hypothesized that electron withdrawing groups on the ring could alter the stability of the amide linkage, which might eventually lead to easier degradation. This would reduce the negative effects on plant development associated with the prolonged inhibition by NPA due to its stability *in planta*. These compounds were tested for inhibition of polar auxin transport[48], but unfortunately did not show a similar inhibition of auxin transport as NPA (Figure 5). This could indicate that these NPA analogues are no longer inhibitors of the PIN proteins or that these compounds are differently metabolized. Notwithstanding, the combination of IAA and NPA was an intriguing finding, because this suggests increased intracellular concentrations of auxins. In combination with the fact that 2,4-D is known to be a poor substrate for the efflux carriers[49], this could hint that auxin levels are responsible for triggering somatic embryogenesis. In case of microspore embryogenesis, C26-like compounds that contained a similar imine group, were reported to influence PIN cycling[50,51]. This could hint that auxin is an important player involved in the initiation of microspore embryogenesis and also that the auxin levels need to be precisely controlled through auxin transport during early embryogenesis. This assumption is supported by the observation that an auxin response is triggered shortly after heat-shock treatment and by the upregulated expression of *PIN1* in 3-day-old embryos when compared to pollen[52] (unpublished data by Charlotte). Taken altogether, the compounds triggering somatic and microspore embryogenesis might be responsible for increased intracellular concentrations of auxins. The possible downstream effects of these elevated concentrations of auxins on targeting F-box proteins, Aux/IAA proteins and auxin responsive genes, potentially hold the molecular

key to understanding *in vitro* embryogenesis.
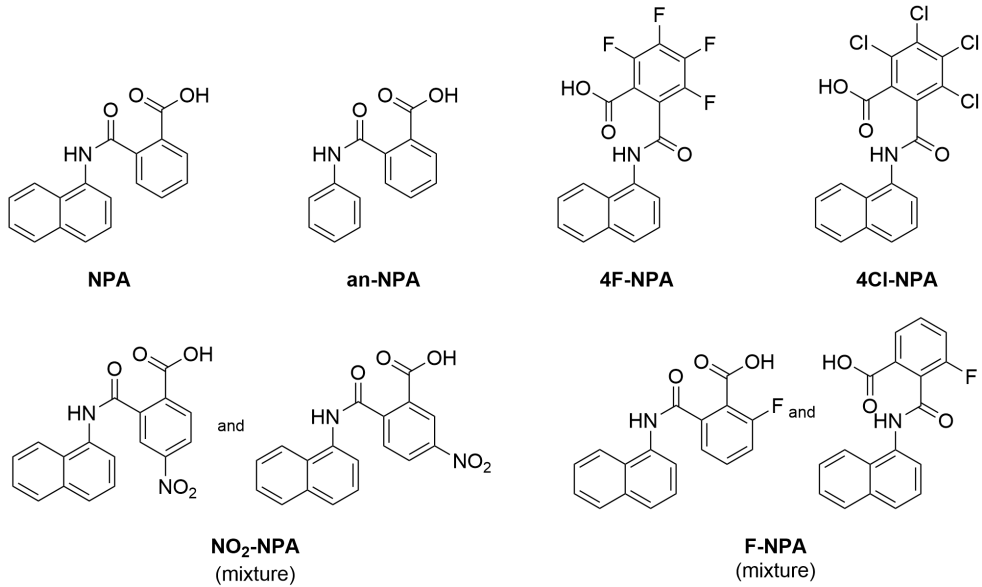


Figure 4. Structures of NPA and NPA analogues; naphthylphthalamic acid (NPA), analine derivative (an-NPA), tetrafluoro derivative (4F-NPA), tetrachloro derivative (4Cl-NPA), mixture of nitro derivatives (NO$_2$-NPA), and mixture of fluoro derivatives (F-NPA).
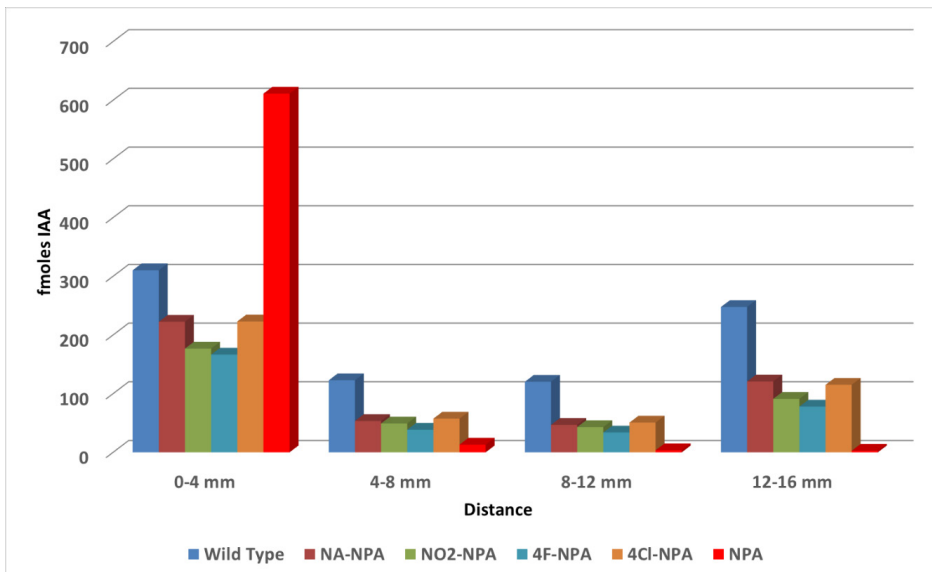


Figure 5. Auxin transport inhibition in presence of NPA and analogues, measured by Cees Boot, Leiden University. Part of the stem of *A. thaliana* were incubated with the respective NPA analogues. The stems were placed in a donor-receiver system that contained $^3$H IAA. After incubation, the stem was cut into segments (mm) and the amount of 3H IAA was measured. For NPA, auxin accumulates in the first segment which indicates that the auxin transport is inhibited. The auxin transport was not inhibited by the NPA analogues.

## Chapter 6

## Glycans in plant research

Similar to mammalian and bacterial cells, plant cells contain a glycocalyx. This glycan layer can be glycoengineered through MOE and several applications of this technique to plant cells have been published[53–57]. A previous study by our group  successfully engineered and imaged the glycans on the roots of *Arabidopsis thaliana*[58]. We hypothesized that the glycans on the pollen during microspore embryogenesis would provide an interesting target since the glycan expression patterns might change during this developmental process. This hypothesis was inspired by the revolutionary paper[10] of the Bertozzi group that visualized glycans during the development of a zebrafish embryo with a temporal resolution. In an attempt to bridge the gap between the two chemical biology approaches described in this thesis, we set out to label the glycans during microspore embryogenesis with MOE. The results of this effort are described in the following section.

For MOE, the probe $Ac_4GlcNAz$ was chosen and initial experiments showed fluorescent background labeling for pollen. To decrease aspecific fluorescent background, older pollen cultures were used and a set of fluorophores was chosen that was expected to have the least affinity for the cell wall[59]. The resulting confocal images showed auto-fluorescence of the exine of the pollen (Figure 6A). Next, suspensor embryos were used because of the thinner or absent exine in combination with more stringent washing, but confocal microscopy demonstrated background labeling for the suspensor culture (Figure 6B). In conclusion, this evaluation showed that the signal unique for the possibly incorporated glycans could not be distinguished from the aspecific fluorescent background labeling. Possibly, in another setting with different monosaccharides and fluorescent dyes, the glycans during ME might be engineered and visualized.
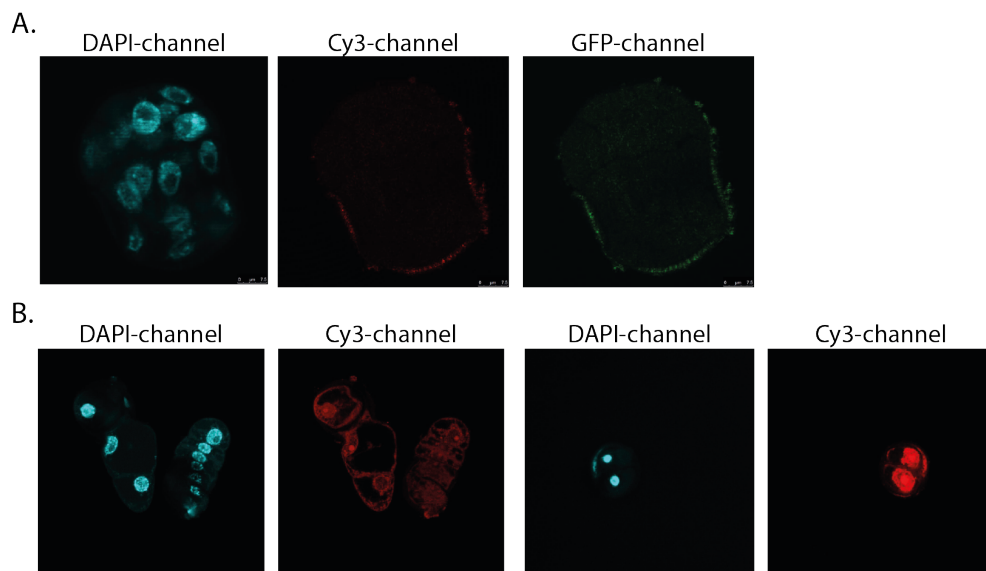
A.



B.



Figure 6. Metabolic oligosaccharide engineering attempts on the pollen of Brassica Napus during microspore embryogenesis. (A) Pollen were treated with $Ac_4GlcNAz$ and clicked with Cy3-alkyne. The exine of the pollen showed fluorescence in multiple channels, and not only in Cy3 channel of the chosen fluorophore, and thus indicating background fluorescence. (B) Suspensor embryos treated with DMSO and then Cy3-alkyne. Confocal images shown for two sets of suspensor embryos. The Cy3 channel shows fluorescent signal in the absence of the azido glycan and thus indicating background labeling.

## Concluding remarks

Chemical biology is an exciting field, which enables the study of biological processes at the molecular scale. In this thesis, all conducted studies used a chemical biology approach. In **Chapter 2**, the combination between the fields of chemical biology and glycobiology is introduced in a literature review based on the occurrence of glycan mimicry by bacteria. The possibility to study bacterial glycans is expanded through the application of the glycoengineering technique SEEL in **Chapter 3**. In another chemical biology approach, the power of bacterial sialyltransferases can be harnessed to introduce chemical reporter groups on bacterial glycans, which is described in **Chapter 4**. Finally, the power of small molecule screens in plant biology is demonstrated in **Chapter 5** by establishing a structure-activity relationship of the synthetic auxin 2,4-D in somatic embryogenesis. The described chemical biology approaches in this thesis provide or illustrate ways to use chemistry to further increases our understanding of biology. In the future, we will see many more thrilling examples of the interdisciplinary potential of chemical biology.

6

## Chapter 6

## Acknowledgement

# References

(1)     Kubicek, S.; Thiel, G. Voices of Chemical Biology. *Nat. Chem. Biol.* **2015**, *11*, 378–379.

(2)     Kaschani, F.; van der Hoorn, R. Small Molecule Approaches in Plants. *Curr. Opin. Chem. Biol.* **2007**, *11*, 88–98.

(3)     Hicks, G. R.; Raikhel, N. V. Small Molecules Present Large Opportunities in Plant Biology. *Annu. Rev. Plant Biol.* **2012**, *63*, 261–282.

(4)     Audenaert, D.; Overvoorde, P. *Plant Chemical Biology*, NV-1 onl.; Wiley: New Jersey, 2014.

(5)     Arnold, F. H. Directed Evolution: Bringing New Chemistry to Life. *Angew. Chem - Int. Ed.* **2018**, *57*, 4143–4148.

(6)     Esvelt, K. M.; Carlson, J. C.; Liu, D. R. A System for the Continuous Directed Evolution of Biomolecules. *Nature* **2011**, *472*, 499–503.

(7)     Kolb, H. C.; Finn, M. G.; Sharpless, K. B. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angew. Chem - Int. Ed.* **2001**, *40*, 2004–2021.

(8)     Devaraj, N. K. The Future of Bioorthogonal Chemistry. *ACS Cent. Sci.* **2018**, *4*, 952–959.

(9)     Thirumurugan, P.; Matosiuk, D.; Jozwiak, K. Click Chemistry for Drug Development and Diverse Chemical-Biology Applications. *Chem. Rev.* **2013**, *113*, 4905–4979.

(10)   Laughlin, S. T.; Baskin, J. M.; Amacher, S. L.; Bertozzi, C. R. In Vivo Imaging of Membrane-Associated Glycans in Developing Zebrafish. *Science* **2008**, *320*, 664–668.

(11)   Banahene, N.; Kavunja, H. W.; Swarts, B. M. Chemical Reporters for Bacterial Glycans: Development and Applications. *Chem. Rev.* **2021**, acs.chemrev.1c00729.

(12)   Dube, D. H.; Champasa, K.; Wang, B. Chemical Tools to Discover and Target Bacterial Glycoproteins. *Chem. Commun.* **2011**, *47*, 87–101.

(13)   Tra, V. N.; Dube, D. H. Glycans in Pathogenic Bacteria-Potential for Targeted Covalent Therapeutics and Imaging Agents. *Chem. Commun.* **2014**, *50*, 4659–4673.

(14)   Calabretta, P.; Hodges, H. L.; Kraft, M. B.; Marando, V.; Kiessling, L. L. Bacterial Cell Wall Modification with a Glycolipid Substrate. *J. Am. Chem. Soc.* **2019**, *141*, 9262–9272.

(15)   Lukose, V.; Whitworth, G.; Guan, Z.; Imperiali, B. Chemoenzymatic Assembly of Bacterial Glycoconjugates for Site-Specific Orthogonal Labeling. *J. Am. Chem. Soc.* **2015**, *137*, 12446–12449.

(16)   Whitworth, G. E.; Imperiali, B. Selective Biochemical Labeling of *Campylobacter Jejuni* Cell-Surface Glycoconjugates. *Glycobiology* **2015**, *25*, 756–766.

(17)   Imperiali, B. Bacterial Carbohydrate Diversity — a Brave New World. *Curr. Opin. Chem. Biol.* **2019**, *53*, 1–8.

(18)   Capicciotti, C. J.; Zong, C.; Sheikh, M. O.; Sun, T.; Wells, L.; Boons, G. J. Cell-Surface Glyco-Engineering by Exogenous Enzymatic Transfer Using a Bifunctional CMP-Neu5Ac Derivative. *J. Am. Chem. Soc.* **2017**, *139*, 13342–13348.

(19)   Gulati, S.; Schoenhofen, I. C.; Whitfield, D. M.; Cox, A. D.; Li, J.; St. Michael, F.; Vinogradov, E. V.; Stupak, J.; Zheng, B.; Ohnishi, M.; Unemo, M.; Lewis, L. A.; Taylor, R. E.; Landig, C. S.; Diaz, S.; Reed, G. W.; Varki, A.; Rice, P. A.; Ram, S. Utilizing CMP-Sialic Acid Analogs to Unravel *Neisseria Gonorrhoeae* Lipooligosaccharide-Mediated Complement Resistance and Design Novel Therapeutics. *PLoS Pathog.* **2015**, *11*, e1005290.

(20)   Blaum, B. S.; Hannan, J. P.; Herbert, A. P.; Kavanagh, D.; Uhrín, D.; Stehle, T. Structural Basis for Sialic Acid-Mediated Self-Recognition by Complement Factor H. *Nat. Chem. Biol.* **2015**, *11*, 77–82.

(21)   Kuhn, B.; Benz, J.; Greif, M.; Engel, A. M.; Sobek, H.; Rudolph, M. G. The Structure of Human -2,6-Sialyltransferase Reveals the Binding Mode of Complex Glycans. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2013**, *69*, 1826–1838.

(22)   Goon, S.; Schilling, B.; Tullius, M. V.; Gibson, B. W.; Bertozzi, C. R. Metabolic Incorporation of Unnatural Sialic Acids into *Haemophilus Ducreyi* Lipooligosaccharides. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 3089–3094.

(23)   Mubaiwa, T. D.; Semchenko, E. A.; Hartley-Tassell, L. E.; Day, C. J.; Jennings, M. P.; Seib, K. L. The Sweet Side of the Pathogenic *Neisseria*: The Role of Glycan Interactions in Colonisation and Disease. *Pathogens and Disease.* 2017, p ftx063.

(24)   Hong, S.; Shi, Y.; Wu, N. C.; Grande, G.; Douthit, L.; Wang, H.; Zhou, W.; Sharpless, K. B.; Wilson, I. A.; Xie, J.; Wu, P. Bacterial Glycosyltransferase-Mediated Cell-Surface Chemoenzymatic Glycan Modification. *Nat. Commun.* **2019**, *10*, 1799.

(25)   Pacheco, A. R.; Munera, D.; Waldor, M. K.; Sperandio, V.; Ritchie, J. M. Fucose Sensing Regulates Bacterial Intestinal Colonization. *Nature* **2012**, *492*, 113–117.

(26)   Stenutz, R.; Weintraub, A. The Structures of *Escherichia Coli* O-Polysaccharide Antigens. *FEMS Microbiol. Rev.* **2006**, *30*, 382–403.

(27)   Li, H.; Liao, T.; Debowski, A. W.; Tang, H.; Nilsson, H. O.; Stubbs, K. A.; Marshall, B. J.; Benghezal, M. Lipopolysaccharide Structure and Biosynthesis in *Helicobacter Pylori*. *Helicobacter* **2016**, *21*, 445–461.

(28)   Ma, B.; Simala-Grant, J. L.; Taylor, D. E. Fucosylation in Prokaryotes and Eukaryotes. *Glycobiology* **2006**, *16*.

**6**

(29)   Lopez Aguilar, A.; Briard, J. G.; Yang, L.; Ovryn, B.; Macauley, M. S.; Wu, P. Tools for Studying Glycans: Recent Advances in Chemoenzymatic Glycan Labeling. *ACS Chem. Biol.* **2017**, *12*, 611–621.

(30)   Chinoy, Z. S.; Montembault, E.; Moremen, K. W.; Royou, A.; Friscourt, F. Impacting Bacterial Sialidase Activity by Incorporating Bioorthogonal Chemical Reporters onto Mammalian Cell-Surface Sialosides. *ACS Chem. Biol.* **2021**, *16*, 2307–2314.

(31)   Hudak, J. E.; Alvarez, D.; Skelly, A.; Von Andrian, U. H.; Kasper, D. L. Illuminating Vital Surface Molecules of Symbionts in Health and Disease. *Nat. Microbiol.* **2017**, *2*, 17099.

(32)   Shell, D. M.; Chiles, L.; Judd, R. C.; Seal, S.; Rest, R. F. The *Neisseria* Lipooligosaccharide-Specific Alpha-2,3-Sialyltransferase Is a Surface-Exposed Outer Membrane Protein. *Infect. Immun.* **2002**, *70*, 3744–3751.

(33)   Manhardt, C. T.; Punch, P. R.; Dougher, C. W. L.; Lau, J. T. Y. Extrinsic Sialylation Is Dynamically Regulated by Systemic Triggers in Vivo. *J. Biol. Chem.* **2017**, *292*, 13514–13520.

(34)   Luijkx, Y. M. C. A.; Bleumink, N. M. C.; Jiang, J.; Overkleeft, H. S.; Wösten, M. M. S. M.; Strijbis, K.; Wennekes, T. Bacteroides Fragilis Fucosidases Facilitate Growth and Invasion of *Campylobacter Jejuni* in the Presence of Mucins. *Cell. Microbiol.* **2020**, *22*, e13252.

(35)   Tytgat, H. L. P.; de Vos, W. M. Sugar Coating the Envelope: Glycoconjugates for Microbe–Host Crosstalk. *Trends Microbiol.* **2016**, *24*, 853–861.

(36)   McMANUS, J. M.; HERBST, R. M. Tetrazole Analogs of Plant Auxins. *J. Org. Chem.* **1959**, *24*, 1464–1467.

(37)   Ballatore, C.; Huryn, D. M.; Smith, A. B. Carboxylic Acid (Bio)Isosteres in Drug Design. *ChemMedChem* **2013**, *8*, 385–395.

(38)   Kepinski, S.; Leyser, O. The *Arabidopsis* F-Box Protein TIR1 Is an Auxin Receptor. *Nature* **2005**, *435*, 446–451.

(39)   Lee, S.; Sundaram, S.; Armitage, L.; Evans, J. P.; Hawkes, T.; Kepinski, S.; Ferro, N.; Napier, R. M. Defining Binding Efficiency and Specificity of Auxins for SCF TIR1/AFB -Aux/IAA Co-Receptor Complex Formation. *ACS Chem. Biol.* **2014**, *9*, 673–682.

(40)   Calderón Villalobos, L. I. A.; Lee, S.; De Oliveira, C.; Ivetac, A.; Brandt, W.; Armitage, L.; Sheard, L. B.; Tan, X.; Parry, G.; Mao, H.; Zheng, N.; Napier, R.; Kepinski, S.; Estelle, M. A Combinatorial TIR1/AFB–Aux/IAA Co-Receptor System for Differential Sensing of Auxin. *Nat. Chem. Biol.* **2012**, *8*, 477–485.

(41)   Radoeva, T.; Weijers, D. A Roadmap to Embryo Identity in Plants. *Trends Plant Sci.* **2014**, *19*, 709–716.

(42)   Soriano, M.; Li, H.; Boutilier, K. Microspore Embryogenesis: Establishment of Embryo Identity and Pattern in Culture. *Plant Reprod.* **2013**, *26*, 181–196.

(43)   Karami, O.; Saidi, A. The Molecular Basis for Stress-Induced Acquisition of Somatic Embryogenesis. *Mol. Biol. Rep.* **2010**, *37*, 2493–2507.

(44)   Shariatpanahi, M. E.; Bal, U.; Heberle-Bors, E.; Touraev, A. Stresses Applied for the Re-Programming of Plant Microspores towards in Vitro Embryogenesis. *Physiol. Plant.* **2006**, *127*, 519–534.

(45)   Park, W.; Park, S.; Shin, H.; Na, K. Acidic Tumor PH-Responsive Nanophotomedicine for Targeted Photodynamic Cancer Therapy. *J. Nanomater.* **2016**, *2016*, 3739723.

(46)   Friml, J. Subcellular Trafficking of PIN Auxin Efflux Carriers in Auxin Transport. *Eur. J. Cell Biol.* **2010**, *89*, 231–235.

(47)   Titapiwatanakun, B.; Murphy, A. S. Post-Transcriptional Regulation of Auxin Transport Proteins: Cellular Trafficking, Protein Phosphorylation, Protein Maturation, Ubiquitination, and Membrane Composition. *J. Exp. Bot.* **2009**, *60*, 1093–1107.

(48)   Boot, K. J. M.; Hille, S. C.; Libbenga, K. R.; Peletier, L. A.; Van Spronsen, P. C.; Van Duijn, B.; Offringa, R. Modelling the Dynamics of Polar Auxin Transport in Inflorescence Stems of Arabidopsis Thaliana. *J. Exp. Bot.* **2016**, *67*, 649–666.

(49)   Hosek, P.; Kubes, M.; Lankova, M.; Dobrev, P.; Klima, P.; Kohoutova, M.; Petrasek, J.; Hoyerova, K.; Jirina, M.; Zazimalova, E. Auxin Transport at Cellular Level: New Insights Supported by Mathematical Modelling. *J. Exp. Bot.* **2012**, *63*, 3815–3828.

(50)   Nishimura, T.; Matano, N.; Morishima, T.; Kakinuma, C.; Hayashi, K. I.; Komano, T.; Kubo, M.; Hasebe, M.; Kasahara, H.; Kamiya, Y.; Koshiba, T. Identification of IAA Transport Inhibitors Including Compounds Affecting Cellular PIN Trafficking by Two Chemical Screening Approaches Using Maize Coleoptile Systems. *Plant Cell Physiol.* **2012**, *53*, 1671–1682.

(51)   Bektas, Y.; Rodriguez-Salus, M.; Schroeder, M.; Gomez, A.; Kaloshian, I.; Eulgem, T. The Synthetic Elicitor DPMP (2,4-Dichloro-6-{(E)-[(3-Methoxyphenyl)Imino]Methyl}phenol) Triggers Strong Immunity in *Arabidopsis Thaliana* and Tomato. *Sci. Rep.* **2016**, *6*, 29554.

(52)   Soriano, M.; Li, H.; Jacquard, C.; Angenent, G. C.; Krochko, J.; Offringa, R.; Boutilier, K. Plasticity in Cell Division Patterns and Auxin Transport Dependency during in Vitro Embryogenesis in *Brassica Napus*. *Plant Cell* **2014**, *26*, 2568–2581.

(53)   Anderson, C. T.; Wallace, I. S.; Somerville, C. R. Metabolic Click-Labeling with a Fucose Analog Reveals Pec-

tin Delivery, Architecture, and Dynamics in *Arabidopsis* Cell Walls. *Proc. Natl. Acad. Sci.* **2012**, *109*, 1329–1334.

(54)  Dumont, M.; Lehner, A.; Vauzeilles, B.; Malassis, J.; Marchant, A.; Smyth, K.; Linclau, B.; Baron, A.; Mas Pons, J.; Anderson, C. T.; Schapman, D.; Galas, L.; Mollet, J. C.; Lerouge, P. Plant Cell Wall Imaging by Metabolic Click-Mediated Labelling of Rhamnogalacturonan II Using Azido 3-Deoxy- d -Manno-Oct-2-Ulosonic Acid. *Plant J.* **2016**, *85*, 437–447.

(55)  McClosky, D. D.; Wang, B.; Chen, G.; Anderson, C. T. The Click-Compatible Sugar 6-Deoxy-Alkynyl Glucose Metabolically Incorporates into *Arabidopsis* Root Hair Tips and Arrests Their Growth. *Phytochemistry* **2016**, *123*, 16–24.

(56)  Zhu, Y.; Wu, J.; Chen, X. Metabolic Labeling and Imaging of N-Linked Glycans in *Arabidopsis Thaliana*. *Angew. Chem - Int. Ed.* **2016**, *55*, 9301–9305.

(57)  Simon, C.; Lion, C.; Spriet, C.; Baldacci-Cresp, F.; Hawkins, S.; Biot, C. One, Two, Three: A Bioorthogonal Triple Labelling Strategy for Studying the Dynamics of Plant Cell Wall Formation In Vivo. *Angew. Chem. Int. Ed.* **2018**, *57*, 16665–16671.

(58)  Hoogenboom, J.; Berghuis, N.; Cramer, D.; Geurts, R.; Zuilhof, H.; Wennekes, T. Direct Imaging of Glycans in Arabidopsis Roots via Click Labeling of Metabolically Incorporated Azido-Monosaccharides. *BMC Plant Biol.* **2016**, *16*, 220.

(59)  Hughes, L. D.; Rawle, R. J.; Boxer, S. G. Choose Your Label Wisely: Water-Soluble Fluorophores Often Interact with Lipid Bilayers. *PLoS One* **2014**, *9*, e87649.

6

## Chapter 6

## Experimental

## Synthesis NPA analogues
### An-NPA
Aniline (0.1 mL; 1.1 mmol) and phthalic anhydride (0.23 g; 1.6 mmol; 1.4 eq) were added to a RB with acetonitrile (5 mL). The reaction mixture was put under argon atmosphere and stirred overnight at RT. The next day, a white precipitate was formed. The organic layer was concentrated resulting in a white solid. EtOAc was added and the remaining solid was filtered. The organic layer was washed with $NH_4Cl$ (3 x 20 mL) and brine (3 x 20 mL). Column chromatography was performed (50 % PE : 50% EtOAc : 1 % acetic acid) twice to obtain An-NPA (92 mg; 0.38 mmol; 35 %).

^{1}H NMR (600 MHz, MeOD) δ 7.92 (dd, $J$ = 7.8, 1.3 Hz, 1H), 7.58 – 7.52 (m, 3H), 7.46 (m, 2H), 7.26 – 7.20 (m, 2H), 7.05 – 7.00 (m, 1H).
^{13}C NMR (600 MHz, MeOD) δ 169.75, 167.81, 139.03, 138.72, 131.91, 130.06, 129.26, 129.13, 128.37, 127.40, 124.03, 120.36.

### 4F-NPA
Naphthylamine (0.389 g; 2.72 mmol) and tetrafluorophthalic anhydride (0.61 g; 3 mmol) were added to a RB with acetonitrile (10 mL). The reaction turned into an orange suspension within minutes. The reaction was capped with inert gas and was stirred overnight. The reaction mixture was concentrated *in vacuo* and dissolved in EtOAc. The organic layer was washed with $NH_4Cl$ (1 x 60 mL) and brine (2x 80 mL) and dried over $Na_2SO_4$. The organic layer was concentrated to a light pink solid to obtain 4F-NPA (quantitative yield).

^{1}H NMR (400 MHz, DMSO) δ 10.81 (s, 1H), 8.15 – 8.07 (m, 1H), 8.01 – 7.92 (m, 1H), 7.85 (d, $J$ = 8.2 Hz, 1H), 7.67 (d, $J$ = 7.3 Hz, 1H), 7.61 – 7.51 (m, 3H).
^{13}C NMR (600 MHz, DMSO) δ 134.23, 133.07, 128.63, 128.58, 126.83, 126.76, 126.71, 126.05, 123.17, 122.78.
^{19}F NMR (400 MHz, DMSO) δ -138.23 – -138.82 (m), -139.52 – -139.78 (m), -140.33 (m), -151.30 (d, $J$ = 27.3 Hz), -153.19 (m).

### 4Cl-NPA
1-napthylamine (0.409 g; 2.9 mmol) and tetrachlorophthalic anhydride (0.948 g; 3.3 mmol) were added to a RB with acetonitrile. The reaction mixture was put under argon and left stirring for 2 days. The reaction mixture was concentrated *in vacuo*. EtOAc was added to the crude and the organic layer was washed with $NH_4Cl$ (3 x 20 mL)  and brine (3 x 20 mL). The crude was recrystallized by

refluxing in methanol for circa 1 h and the residue was washed with ice-cold DCM (2 x 15 mL) and Et$_2$O (2 x 15 mL). Only a small amount of product was obtained (0.04 g; 0.09 mmol; 3%), which showed some impurities on TLC.

$^1$H NMR (600 MHz, MeOD) δ 7.92 (dd, $J$ = 7.8, 1.3 Hz, 1H), 7.59 – 7.51 (m, 3H), 7.51 – 7.43 (m, 2H), 7.26 – 7.20 (m, 2H), 7.05 – 6.99 (m, 1H).
$^{13}$C NMR (600 MHz, DMSO) δ 165.41, 162.85, 136.27, 134.22, 133.53, 132.93, 130.03, 129.12, 128.59, 128.53, 126.76, 126.71, 126.02, 123.25, 122.46.

### NO$_2$-NPA

Naphthylamine (0.392 g; 2.7 mmol) and nitrophthalic anhydride (0.61 g; 3.2 mmol) were added to a RB with acetonitrile (10 mL). The reaction was capped with inert gas and was stirred overnight. The reaction mixture was concentrated *in vacuo* and dissolved in EtOAc. The organic layer was washed with NH$_4$Cl (3 x 30 mL) and brine (3 x 30 mL) and dried over Na$_2$SO$_4$ to obtain NO$_2$-NPA (0.71 g; 2.1 mmol; 77 %).

$^1$H NMR (600 MHz, DMSO) δ 10.65 (s, 1H), 8.37 (d, $J$ = 8.1 Hz, 1H), 8.31 (d, $J$ = 7.8 Hz, 1H), 8.22 (dd, $J$ = 6.4, 3.6 Hz, 1H), 7.97 (dd, $J$ = 6.7, 3.6 Hz, 1H), 7.89 – 7.80 (m, 3H), 7.61 – 7.51 (m, 4H).
$^{13}$C NMR (600 MHz, DMSO) δ 166.76, 164.24, 147.77, 135.50, 134.22, 133.84, 133.60, 130.69, 128.66, 128.43, 127.83, 126.53, 126.24, 126.07, 125.97, 123.58, 122.26.

### F-NPA

Naphthylamine (0.392 g; 2.7 mmol) and fluorophthalic anhydride (0.61 g; 3.2 mmol) were added to a RB with acetonitrile (10 mL). The reaction was capped with inert gas and was stirred overnight. The reaction mixture was concentrated *in vacuo* and partially dissolved in EtOAc. The organic layer was washed with NH$_4$Cl (3 x 50 mL) and brine (3 x 50 mL) and dried over Na$_2$SO$_4$. After concentration *in vacuo*, THF (15 mL) was used to dissolve residue and PE (30 mL) was added to recrystallize the product. Upon adding PE, cloudy precipitation was observed immediately. The recrystallization was left standing overnight and filtered the next morning and washed with PE to obtain F-NPA (0.601 g; 1.94 mmol; 70%).

$^1$H NMR (600 MHz, DMSO) δ 13.53 (s, 1H), 10.56 (s, 1H), 8.24 – 8.19 (m, 1H), 8.00 – 7.95 (m, 1H), 7.88 – 7.82 (m, 2H), 7.75 (d, $J$ = 7.4 Hz, 1H), 7.70 – 7.60 (m, 2H), 7.60 – 7.53 (m, 3H).
$^{13}$C NMR (600 MHz, DMSO) δ 166.71, 163.69, 160.08, 158.46, 134.23, 133.88, 131.60, 131.19, 131.13, 128.92, 128.46, 127.94, 127.80, 126.56, 126.44, 126.33, 126.26, 126.06, 123.66, 122.81, 120.30, 120.15.
$^{19}$F NMR (600 MHz, DMSO) δ -116.84.

**6**

# Chapter 6

## Synthesis isosteres 2,4-D

### 2,4-D-sulfonamide

2,4-dichlorophenol (0.21 g; 1.29 mmol) and chloromethane sulfonamide (0.32 g; 2.47 mmol) were dissolved in DMF (8 mL). $K_2CO_3$ (0.68 g; 4.92 mol) was added and the reaction was put under argon atmosphere. The reaction was heated overnight at 65 °C. The reaction mixture was diluted with EtOAc (150 mL), washed with water (3 x 50 mL), washed with brine (3 x 50 mL) and dried over $Na_2SO_4$. Crude was purified with column chromatography (80% PE : 20% EtOAc to 50% PE : 50% EtOAc) to yield product (66 mg; 0.26 mmol; 20%).

^{1}H NMR (600 MHz, MeOD) δ 7.36 (d, $J$ = 2.5 Hz, 1H), 7.23 (d, $J$ = 8.9 Hz, 1H), 7.18 (dd, $J$ = 8.9, 2.5 Hz, 1H), 5.04 (s, 2H).

^{13}C NMR (400 MHz, MeOD) δ 152.42, 129.56, 127.78, 127.58, 124.45, 117.60, 81.86.

### 2,4-D tetrazole

2,4-dichlorophenol (0.187 g; 1.15 mmol), chloroacetonitrile (0.07 mL; 1.15 mmol), anhydrous $K_2CO_3$ (0.182; 1.3 mmol) and dry acetone (5 mL) were charged to a RB. The reaction mixture was heated for 4.5 hours. The reaction mixture was quenched on ice, washed and concentrated. NMR confirmed products and a few impurities (0.19 g; 0.94 mmol). The reaction was continued with crude product. 2,4-dichlorophenoxyacetonitrile (0.116 g, crude), D-proline (0.021 g; 0.18 mmol), and $NaN_3$ (0.048 g; 0.74 mmol) and DMF (3 mL) were charged in a roundbottom flask and put under argon atmosphere. The reaction was heated at 110 °C for 2 h and cooled down to RT. The reaction mixture was quenched in ice water, acidified with HCl, and the resulting white solid was filtered and washed with water to yield product (0.085 g; 0.34 mmol; 61 %).

^{1}H NMR (400 MHz, MeOD) δ 7.45 (dd, $J$ = 2.6, 0.9 Hz, 1H), 7.30 (ddd, $J$ = 8.9, 2.5, 0.9 Hz, 1H), 7.21 (dd, $J$ = 8.8, 1.0 Hz, 1H), 5.52 (d, $J$ = 0.9 Hz, 2H).

^{13}C NMR (400 MHz, MeOD) δ 129.68, 127.68, 115.43, 60.56.

## Metabolic Oligosaccharide Engineering in Microspore Embryogenesis

Microspore cultures were prepared and incubated with $Ac_4GlcNAz$ (100 µM) or DMSO for 24 h. The samples were washed and a click reaction was performed. A typical click reaction of 1000 µL reaction volume contained 1 mM $CuSO_4$, 1 mM sodium L-ascorbate and 1 µM fluorescent dye. Samples were incubated at 4 °C while spinning in the dark. Before adding DAPI, the samples were washed 3 x and imaged on a Leica Confocal Microscope.

General discussion

6

# Appendix

Nederlandse samenvatting
About the author
List of publications
Acknowledgement

# Nederlandse samenvatting

Deze thesis bestaat uit twee verschillende onderzoeken die allebei behoren tot het vakgebied van de chemische biologie. Chemische biologie is een combinatie van chemie en biologie. Hierbij worden vaak chemische technieken toegepast om meer te leren over biologische processen op moleculair niveau. Een goed voorbeeld hiervan is het veranderen van de chemische structuur van een medicijn om meer te leren over de biologische werking daarvan. In dit promotieonderzoek zijn twee verschillende aanpakken uit de chemische biologie beschreven. In het eerste deel is de focus gericht op het ontwikkelen van chemische technieken om meer te leren over bacteriën. In het tweede deel staat het gebruik van een set moleculen centraal om meer te leren over het groeien van planten.

In het menselijk lichaam huizen heel veel bacteriën. Deze micro-organismen kunnen een positieve of een negatieve invloed uitoefenen op ons lichaam. Al deze bacteriën zijn aan de buitenkant bedekt met een laag suikers. Dit gaat niet om suiker zoals je die eet, maar om een groep van moleculen die we aanduiden met de term suikers, of glycanen. Dit laagje van suikers op bacteriën is belangrijk voor allerlei processen, zoals interacties met andere cellen, en bepaalt onder andere hoe ziekmakend een bacterie is. We komen steeds meer te weten over de samenstelling en rol van deze suikers op bacteriën, maar we weten nog heel veel niet. Om meer te weten te komen over deze suikers en hun functie kan je uitstekend gebruik maken van chemische biologie. Deze suikers worden niet volgens een vaste instructie gemaakt, zoals bijvoorbeeld wel het geval is voor eiwitten. Daarom is het handig uniek gereedschap te gebruiken, ontwikkeld vanuit de chemische biologie, om suikers te bestuderen.

In **hoofdstuk 2** is er in de wetenschappelijke literatuur gezocht naar de samenstelling van suikers op bacteriën. Op basis van deze literatuur is er een overzicht samengesteld van bacteriën die menselijke cellen nabootsen, en dus 'menselijke' suikers op hun buitenkant presenteren. Doordat bacteriën zich op deze manier presenteren, kunnen ze het immuun systeem van mensen vermijden. Daarnaast worden er in dit hoofdstuk verschillende technieken uit de chemische biologie uiteengezet die gebruikt kunnen worden om het nabootsen van suikers te bestuderen.

Om meer te weten te komen over de suikers op bacteriën kunnen we gebruik maken van technieken uit de chemische biologie. Helaas zijn er maar weinig technieken om suikers te bestuderen specifiek op bacteriën. In **hoofdstuk 3**

hebben we een techniek toegepast op bacteriën die ook wordt gebruikt om suikers op menselijke cellen te bestuderen. Dit onderzoek laat zien dat de techniek goed gebruikt kan worden om specifieke suikers op bacteriën te bestuderen. De techniek maakt gebruik van een enzym dat de suikers op de buitenkant van de bacterie zet. Het voordeel van deze techniek is dat de modificatie direct aan de buitenkant gemaakt kan worden, in tegenstelling tot bestaande technieken waarbij het suiker eerst via de binnenkant van de bacterie verwerkt moet worden. Deze techniek om suikers op de buitenkant te introduceren, is succesvol toegepast op *Neisseria gonorrhoeae*, de bacterie die de seksueel overdraagbare aandoening gonorroe veroorzaakt. Deze bacterie bootst ook menselijke suikers na. Deze techniek zou in de toekomst gebruikt kunnen worden om andere bacteriën te bestuderen die ook suikers nabootsen.

Tijdens het onderzoek van hoofdstuk 3, kwamen we erachter dat het in sommige gevallen niet nodig was om enzym toe te voegen om de suiker op de buitenkant van de bacteriën te introduceren. Dit houdt in dat de bacterie zelf het benodigde gereedschap heeft om deze reactie te doen. In **hoofdstuk 4** hebben we dit proces verder onderzocht door te kijken naar drie verschillende bacteriën: Nontypeable *Haemophilus influenzae*, *Neisseria gonorrhoeae* en *Campylobacter jejuni*. Al deze bacteriën bootsen menselijke suikers na. Er zijn verschillende technieken uit de chemische biologie toegepast om een suiker op de buitenkant van bacteriën te zetten. Deze technieken hebben ieder hun eigen voor- en nadelen. In de toekomst kan een techniek geselecteerd worden op basis van de bacterie en het doel van het onderzoek.

In **hoofdstuk 5** maken we de overgang naar het tweede deel, over het groeien van planten. In dit hoofdstuk maken we gebruik van een andere aanpak uit de chemische biologie. Hiervoor hebben we een kleine verzameling moleculen bedacht en gemaakt. Deze verzameling is vervolgens toegepast op planten. Daarbij is gekeken of deze moleculen een specifiek biologisch groeiproces konden beïnvloeden. Het ging daarbij over het bijzondere proces van het omzetten van een 'volwassen' cel naar een 'embryonale' cel. Als deze cel is omgezet naar een embryonale cel, kan deze namelijk uitgroeien tot een nieuwe volwassen plant. Dit proces vindt ook plaats in de natuur, bijvoorbeeld bij de plant *Mother of thousands* in de volksmond ook wel bommenwerper genoemd. Deze omzetting wordt ook toegepast in de biotechnologie om bijvoorbeeld planten te kweken die van zichzelf steriel zijn en zich dus niet op een andere manier kunnen voortplanten. Om enigszins inzicht te verkrijgen in dit bijzondere proces is er

## Appendix

gekeken naar de relatie tussen de activiteit en de structuur van de moleculen uit de kleine verzameling. Hieruit bleek dat moleculen die het algemene groeiproces van planten sterk konden beïnvloeden, ook goed waren in het initiëren van het omzetten van een cel. Dit onderzoek laat de kracht van kleine moleculen zien in een chemische biologisch aanpak.

*Het doel van deze samenvatting is om de boodschap van dit promotieonderzoek eenvoudig over te brengen. Om de boodschap helder te houden is ingewikkeld taalgebruik vermeden. Dit kan betekenen dat nuances verloren zijn gegaan. Deze samenvatting is dus een gegeneraliseerde, versimpelde boodschap.*

## About the author

Hanna de Jong was born on March 21st 1993 in Utrecht. After receiving her high school diploma at the Openbaar Zeister Lyceum te Zeist, she continued to study Molecular Life Sciences at Wageningen University. During her bachelor she first came into contact with microbial sialic acids whilst performing an internship under the supervision of Tjerk Sminia and Dr. Tom Wennekes at the group of Organic Chemistry. She continued with the master Molecular Life Sciences and took additional courses in chemistry at the Radboud Universiteit Nijmegen. Her first MSc internship was about the delivery of genetic material with virus-like particles under the supervision of Dr. Renko de Vries and Dr. Edwin Tijhaar at the department of Physical Chemistry and Soft Matter. For her second MSc internship, Hanna visited the laboratory of Prof. Matthew Bogyo at Stanford University (United States of America) and worked on small molecule inhibitors for the proteasome of *Plasmodium falciparum*, a parasite causing malaria. In 2017, Hanna started pursuing her PhD at the department of Chemical Biology and Drug Discovery at Utrecht University. Initially, she focused on the design and synthesis of a library of 2,4-D to establish a structure-activity relationship in somatic embryogenesis. Later, a second chemical biology approach became part of her PhD; to glycoengineer bacterial glycans with the focus on sialic acid. This project was a collaboration with the department of Infectionbiology and since then her PhD was supervised by Dr. Marc Wösten, Dr. Tom Wennekes and Prof. G. J. Boons. The research of her PhD is described in this thesis. Hanna is going to perform postdoctoral research in the group of Prof. Barbara Imperiali at the Massachusetts Institute of Technology.

## List of publications

Sweet impersonators: molecular mimicry of host glycans by bacteria
Hanna de Jong, Marc M. S. M. Wösten, Tom Wennekes
*Glycobiology*, 2022, DOI: 10.1093/glycob/cwab104

Selective exoenzymatic labeling of lipooligosaccharides of *Neisseria gonorrhoeae* with sialic acid analogues
Hanna de Jong, Maria J. Moure, Jet E. M. Hartman, Gerlof P. Bosman, Bart W. Bardoel, Geert-Jan Boons, Marc M. S. M. Wösten, Tom Wennekes
*Manuscript in preparation*

Glycoengineering cell surface glycoconjugates on bacterial pathogens with CMP-sialic acid analogues
Hanna de Jong, Erianna I. Alvarado Melendez, Jet E. M. Hartman, Jun Yang Ong, Maria J. Moure, Bart W. Bardoel, Astrid P. Heikema, Geert-Jan Boons, Marc M. S. M. Wösten, Tom Wennekes
*Manuscript in preparation*

Structure-activity relationship of 2,4-D in somatic embryogenesis in *Arabidopsis thaliana*
Hanna de Jong*, Omid Karami*, Victor J. Somovilla, Beatriz Villanueva Acosta, Aldo Bryan Sugiarta, Tom Wennekes, Remko Offringa
* These authors contributed equally to this work
*Manuscript in preparation*

## Other publications

Defining the determinants of specificity of *Plasmodium* proteasome inhibitors
Yoo, E.; Stokes, B. H.; de Jong, H.; Vanaerschot, M.; Kumar, T.; Lawrence, N.; Njoroge, M.; Garcia, A.; Van der Westhuyzen, R.; Momper, J. D.; Ng, C. L.; Fidock, D. A.; Bogyo, M.
*J. Am. Chem. Soc.* 2018, 140, 11424–11437.

## Acknowledgement

Tom, we first met during my studies in Wageningen. I still recall switching classes after hearing praise from fellow students about your organic chemistry lectures. Thank you for all your guidance during my PhD. I admire your everlasting optimism and your social skills. Those qualities, in combination with the possibility to explore research areas beyond organic chemistry, have provided the right environment for me to grow during the years of my PhD. Thank you.

Marc, I was almost two years into my PhD when we came knocking on your door to test a 'little' idea. This idea turned out very successful and so I never left. Thank you for welcoming me into your research group without hesitation. I always felt like I could ask the most basic biology questions to you, which you typically answered with a little joke in the end to reassure me. Thank you.

Geert-Jan, thank you for providing me the opportunity to do a PhD at CBDD. I have learned many amazing things about glycoscience.

Many thanks to the BBOL team. Charlotte, you were always enthusiastic about working together. I really enjoyed doing experiments together in Wageningen and also the many chats we have had after. Omid, Kim, and Remko, thank you for letting me join this team. I appreciate all your patience with explaining plant science to me and your open-mindedness towards interdisciplinary science.

During my PhD I had the opportunity to work together with many amazing scientists and visit quite a few labs. First, thank you Victor, for initiating our collaboration after seeing a poster of mine at a symposium. You proposed to do some molecular simulations, which I think are a valuable addition to our research. Mirjam, thank you for letting me use the gel imager at your department. Bart, thank you for all your help with the flow cytometer and your willingness to discuss my project. Finally, a big thanks to Astrid for sending me your precious *C. jejuni* strains and all the fun during my visits to Rotterdam.

I would also like to take the opportunity to thank my students. Jelle, thank you for all your efforts in the chemistry lab. Jet, thank you for all the hard work and fun we had during your internship at the bio lab. In general, thanks for all your efforts and allowing me to grow as a supervisor as well.

During my PhD, I was part of two departments. First, a big shoutout to all the

## Appendix

members of the Wennekes lab from CBDD (in alphabetical order): Erianna, Helena, Jun Yang, Krishna, Lemeng, Pieter, Tim, Tim, Tom, Xianke, and Yvette. Also, thanks to all the members of the Martin group who welcomed me to the East lab: Emma, Jacco, Kamal, Laurens, Matthijs, Nicola, and Tom. Thanks to all other members of CBDD: Apoorva, Arwin, Bernd, Cindy, Cyril, Dirk, Dowson, Enrico, Francesco, Frederik, Gael, George, Gerlof, Ilhan, Ingrid, Ivan, Jack, Javier, John, Justyna, Linda, Liufeng, Luca, Luuk, Margherita, Margreet, Mehman, Milka, Minglong, Nino, Nishant, Nives, Pouya, Reshmi, Roland, Roosmarijn, Rosanne, Seino, Shuela, Vito, Xuan, Yanyan, Yuji, and Zhiyong. Second, thank you to all the members of Infectionbiology: Celia, Daphne, Guus, Jinyi, Jos, Karin, Koen, Liane, Linda, Maitryee, Marc, Shaofang, Xinyue, and Xuefeng. Also, thanks to our neighbors in the Androclus: Albert, Maaike, Melanie, and Roel from Molecular Host Defense.

I have had an amazing time during my PhD and I would especially like to thank my colleagues from Leiden/Utrecht for all the fun during conferences, dinners, and parties: Charlotte, Emma, Enrico, Ingrid, Nicola, Rosanne, Tom, and Yvette.

Many thanks to Helen, Lucia, and Jeevan, for all the postcards, chats, and videocalls.

Michiel en Lisanne, we deden alle drie een PhD in een ander land en ik bewonder jullie moedige stap om in het buitenland te gaan wonen en werken. Ik vond het altijd erg leuk om jullie verhalen te horen. Lieve Lisanne, bedankt voor je geweldige vriendschap. Alhoewel we ver van elkaar wonen, weet ik dat geen afstand te groot is om ons vriendinnen te blijven noemen.

Onder het motto: dat maakt het verschil, wil ik mijn vrienden bedanken die mijn studietijd, en daarna, onvergetelijk hebben gemaakt: Iris, Jochem, Justin, Kate, Oberon en Robin. Ik ben ontzettend blij met jullie als vriendengroep.

Sherida, Bianca en Peter, Hans en Annie, ik heb me altijd welkom bij jullie gevoeld. Bedankt voor jullie warmte.

Sylvia en Leo, Merel en Koen, Karin en Mario, bedankt voor al het schaterlachen en lol tijdens diners en daarbuiten. Lieve pa en ma, bedankt voor al jullie steun. Meerdere malen heb ik vertrouwd op jullie talent om te luisteren en vragen te stellen, om daarna na te denken over wat ik wil.

Lieve Justin, 'het komt goed', was je mantra wanneer ik steun nodig had. Ik heb kunnen bouwen op jouw geloof in mij dat schuilgaat achter die woorden. Bedankt voor al je droge en relativerende humor, je onbegrensde enthousiasme over wetenschap, je drive en nieuwsgierigheid. Op naar samen nog heel veel mooie jaren!