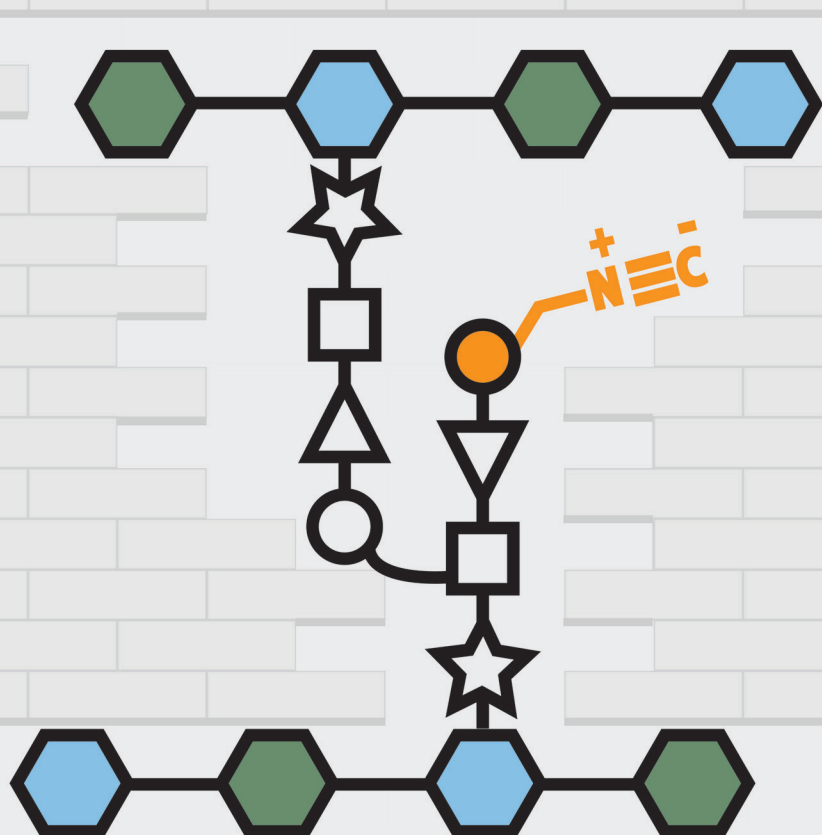


Probing the Bacterial Cell Wall with Chemical Biology Tools



Tjerk J. Sminia

Propositions

1. Reviews without page limit restrictions should not be tagged as “mini-review” by the publisher. (Carbohydrate Research author guidelines)
(this thesis)
2. The reactivity of certain Garner aldehyde derivatives is dramatically influenced by the stereochemistry of a β -positioned methyl-group. (Schmölzer et al., *Carbohydrate Res*, **2013**, 367, 1-4)
(this thesis)
3. Lowering the price of cancer medicines by increasing drug development efforts in academia, as proposed by Workman et al., is not possible in the current academic system. (Workman et al., *Cell*, **2017** 168, 579-583 and Kneller, *Nat. Rev. Drug Discovery*, **2010**, 9, 867-882)
4. The use of an electronic labjournal accelerates the writing of experimental procedures for publications.
5. Improper use of an automated sample changer for NMR spectrometers results in a decrease of effective measurement time, defeating its purpose.
6. Reading scientific articles can be stimulated by having regular literature clubs.
7. Chess lessons should be obligatory in elementary school.
8. The lack of sufficient places to park your bike at public buildings contradicts the governments bike to work initiative. (<http://www.nationalefietsprojecten.nl>)

Propositions belonging to the thesis entitled:

“Probing the bacterial cell wall with chemical biology tools”

Tjerk J. Sminia

Wageningen, 4 December 2017

*Probing the Bacterial Cell Wall
with Chemical Biology Tools*

Tjerk J. Sminia

Thesis committee

Promotors

Prof. Dr H. Zuilhof
Professor of Organic Chemistry
Wageningen University & Research

Prof. Dr W.M. de Vos
Professor of Microbiology
Wageningen University & Research

Co-promotor

Dr T. Wennekes
Assistant professor, Chemical Glycobiology
Utrecht University

Other members

Prof. Dr M.H. Zwietering, Wageningen University & Research
Prof. Dr L. Brunsveld, Eindhoven University of Technology
Prof. Dr J.H. van Maarseveen, University of Amsterdam
Dr K.M. Bonger, Radboud University Nijmegen

This research was conducted under the auspices of the Graduate School VLAG (Advanced studies in Food Technology, Agrobiotechnology, Nutrition and Health Sciences).

Probing the Bacterial Cell Wall with Chemical Biology Tools

Tjerk J. Sminia

Thesis

submitted in fulfilment of the requirements of the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus,
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Monday 4 December 2017
at 1.30 p.m. in the Aula.

Tjerk J. Sminia

Probing the Bacterial Cell Wall with Chemical Biology Tools,
196 pages.

PhD thesis, Wageningen University, Wageningen, The Netherlands (2017)
With references, with summary in English

ISBN 978-94-6343-708-0

DOI 10.18174/425378

To my parents

Table of Contents

| | | |
|---|--|-----|
| List of Abbreviations | | 8 |
| Chapter 1 | General Introduction | 11 |
| Chapter 2 | Getting a Grip on Glycans: a Current Overview of the Metabolic Oligosaccharide Engineering Toolbox | 27 |
| Chapter 3 | Probing Peptidoglycan Synthesis and L-fucose Salvage in the Gut Commensal <i>Akkermansia muciniphila</i> with Bioorthogonal Chemical Reporters | 69 |
| Chapter 4 | Metabolic Glycan Labelling with Azido-Monosaccharides in Gut Microbiome Members | 83 |
| Chapter 5 | Total Synthesis towards Pseudaminic Acid Probes | 109 |
| Chapter 6 | General Discussion | 145 |
| Appendix A | Supporting information for Chapter 3 | 167 |
| Summary | | 179 |
| Samenvatting | | 183 |
| Acknowledgement | | 187 |
| About the Author | | 191 |
| List of Publications | | 193 |
| Overview of completed training activities and education | | 195 |

List of Abbreviations

| | | | |
|-----------------------|---|----------|--|
| Ac | acetyl | ESI | electron spray ionisation |
| aq. | aqueous | Et | ethyl |
| Ar | aromatic | et al. | <i>et alii</i> (and others) |
| Az | azidoacetyl | EtOAc | ethyl acetate |
| Ac ₄ FucAz | 1,2,3,4-tetra- <i>O</i> -acetyl-6-azido-L-fucose | EtOH | ethanol |
| Ac ₄ GalAz | 1,3,4,4-tetra- <i>O</i> -acetyl-6-azido-D-galactosamine | Fmoc | 9H-fluoren-9-ylmethoxycarbonyl |
| Ac ₄ GlcAz | 1,3,4,4-tetra- <i>O</i> -acetyl-6-azido-D-glucosamine | FT-IR | fourier transform infrared spectroscopy |
| Ac ₄ ManAz | 1,3,4,4-tetra- <i>O</i> -acetyl-6-azido-D-mannosamine | Fuc | fucose |
| BCN | bicyclo[6.1.0]nonyne | g | gram |
| Bn | benzyl | Gal | galactose |
| Boc | <i>tert</i> -butoxycarbonyl | GalNAc | <i>N</i> -acetyl galactosamine |
| br | broad | GalNAz | <i>N</i> -azidoacetyl galactosamine |
| Bu | butyl | GC | gas chromatography |
| Bz | benzoyl | GDP | guanosine diphosphate |
| CMP | cytidine monophosphate | Glc | glucose |
| COSY | correlation spectroscopy | GlcNAc | <i>N</i> -acetyl glucosamine |
| C _q | quaternary carbon atom | GlcNAz | <i>N</i> -azidoacetyl glucosamine |
| CuAAC | copper(I)-catalysed alkyne-azide cycloaddition | GlcNCyc | 1,3,4,4-tetra- <i>O</i> -acetyl- <i>N</i> -methylcyclopentane- α , β -D-glucosamine |
| CSA | camphor sulfonic acid | h | hours |
| d | doublet | HOBt | 1-hydroxybenzotriazole |
| DAP | diaminopropionic acid | HPLC | high performance liquid chromatography |
| DBCO | dibenzocyclooctyne | HRMS | high resolution mass spectroscopy |
| DBU | 1,8-diazabicyclo[5.4.0]undec-7-ene | HSQC | heteronuclear single quantum coherence spectroscopy |
| DCC | <i>N,N'</i> -dicyclohexylcarbodiimide | Hz | hertz |
| dd | doublet of doublet | invDA | inverse electron demand Diels-Alder |
| DIG | digoxigenin | IR | infrared |
| DiPEA | <i>N,N</i> ,diisopropyl- <i>N</i> -ethylamine | <i>J</i> | coupling constant |
| DMAP | 4-(<i>N,N</i> -dimethylamino)pyridine | KDN | keto-deoxynonulsonic acid |
| DMF | dimethylformamide | KDO | 3-deoxy-D-manno-oct-2-ulosonic acid |
| DMS | dimethylsulfide | l | liter |
| dt | doublet of triplets | LB | lysogeny broth |
| e.g. | <i>exempli gratia</i> (for example) | LCMS | liquid chromatography mass spectrometry |
| eqv. | equivalents | Leg | legionaminic acid |

| | | | |
|----------------|---|------------------|---|
| LPS | lipopolysaccharide | SDS-PAGE | sodium dodecyl sulfate polyacrylamide gel electrophoresis |
| <i>m</i> | <i>meta</i> | sias | sialic acid |
| M | molar(s) | SPAAC | strain-promoted alkyne-azide cycloaddition |
| m | multiplet | SPANC | strain-promoted alkyne-nitrone cycloaddition |
| m/z | mass over charge ratio | SPS | solvent purification system |
| Man | mannose | <i>tert</i> | tertiary |
| ManNAc | <i>N</i> -acetylmannosamine | t | triplet |
| mg | milligram | TBAF | tetra- <i>n</i> -butylammonium fluoride |
| min | minute(s) | TBS | <i>tert</i> -butyldimethylsilyl Ethers |
| ml | millilitre | <i>t</i> Bu | <i>tert</i> -butyl |
| mmol | millimol(s) | <i>t</i> -BuOK | Potassium <i>tert</i> -butoxide |
| MOE | metabolic oligosaccharide engineering | TEMPO | 2,2,6,6-tetramethyl-1-piperdinyloxy |
| MS | mass spectrometry | Tf | trifluoromethanesulfonyl (triflate) |
| MSA | microbial sialic acid | TFA | trifluoroacetic acid |
| MTBE | methyl- <i>tert</i> -butyl ether | THF | tetrahydrofuran |
| NADHP | nicotinamide adenine dinucleotide phosphate | TLC | thin layer chromatography |
| NaH | sodium hydride | TPAP | tetrapropylammonium perruthenate |
| NAP | 2-naphthylmethyl | Ts | <i>para</i> -toluenemethylene (tosyl) |
| NBS | <i>N</i> -bromosuccinimide | UDP | uridine diphosphate |
| Neu5Ac | <i>N</i> -acetylneuraminic acid | Xyl | xylose |
| NMR | nuclear magnetic resonance | δ | chemical shift |
| <i>o</i> | <i>ortho</i> | | |
| <i>p</i> | <i>para</i> | PG probes | |
| Pd/C | palladium on activated charcoal | DA-DA | D-ala-D-ala |
| PE | petroleum ether | EDA | ethynyl-D-ala |
| Ph | phenyl | EDA-DA | ethynyl-D-ala-D-ala |
| PEG | polyEthylene Glycol | DA-EDA | D-ala-ethynyl-D-ala |
| PG | peptidoglycan | ADA | azido-D-ala |
| ppm | part per million | ADA-DA | azido-D-ala-D-ala |
| Pse | pseudaminic acid | DA-ADA | D-ala-azido-D-ala |
| PTM | post-translational modification | NADA | NBD-amino-D-alanine |
| q | quartet | HADA | HCC-amino-D-alanine |
| ref | reference | | |
| R _f | retardation factor | | |
| rt | room temperature | | |
| s | singlet | | |



Chapter 1

General Introduction

"Cluster One"

The Division Bell, Pink Floyd

Carbohydrates

Carbohydrates are all around us. Besides buying them in the candy shop, carbohydrates, also called sugars or glycans, are also present everywhere in Nature. Carbohydrates are the most abundant class of naturally occurring organic compounds, making up approximately 75% of Earth's biomass.¹ Carbohydrates are structurally very diverse and can vary a lot in their size and complexity, e.g. from 5 to literally thousands of carbon atoms. The name carbohydrate is derived from the empirical formula first postulated for this class of compounds: $C \cdot H_2O$, a hydrate of carbon.¹ This original empirical formula actually no longer suffices for the whole field of carbohydrates. For example, nitrogen-containing carbohydrates (e.g. sialic acid) cannot be described using this empirical formula. The complexity and many rules related to their names and structure are, however, out of scope of this thesis, but are discussed in several excellent books and reviews.¹⁻² Carbohydrates are built from smaller chemical units called monosaccharides. The multitude of different monosaccharides and the endless combinations thereof in complex carbohydrates result in the huge functional diversity of carbohydrates. Single differences in a monosaccharide of a complex carbohydrate on red blood cells, for example, determine our blood type. On the other hand, a polymeric glucose carbohydrate, called cellulose, is important for the structure of all plants on earth.³ In the human body we find a great number of carbohydrates with varying complexity. A simple but abundant one is glycogen a glucose polymer, which is our body's reserve energy supply. The monosaccharide *N*-acetylglucosamine is, however, a major component of the cell wall. The monosaccharide, sialic acid,⁴ is found as a terminal unit on many complex carbohydrates on the outer surface of human cells and plays important roles in cell recognition and communication. In general, complex carbohydrates can be found on the outer surface of almost all living cells. This dense sugar coating, unique to each type of cell or organism, is essential for cell recognition and communication processes between our own cells, but also for example between our cells and bacterial cells in the gut. In short, carbohydrates are essential molecules of life and can be viewed as the social messengers of a cell (Figure 1).

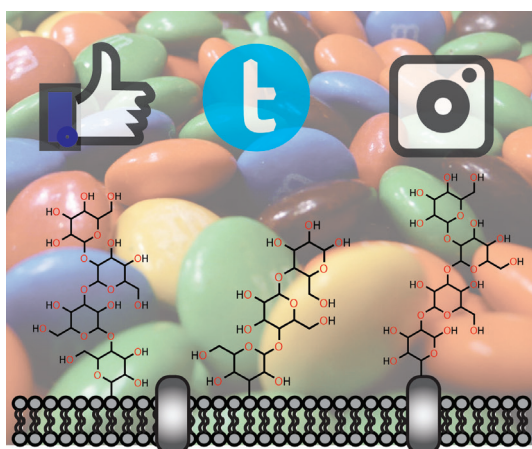


Figure 1. Carbohydrates are the social messengers of a cell.

The molecular mechanism behind the biosynthesis of complex carbohydrates and how their structure directs biological processes in health and disease is, however, poorly understood. Using (organic) chemistry to design and synthesise these carbohydrates is a powerful approach to investigate this biological role at the molecular-level.

Carbohydrate chemistry, the sub-discipline in chemistry mainly focussing on carbohydrates, has always been of high interest since the pioneering work by Nobel laureate Emil Fischer (Figure 2) in the early 20th century.^{1,5} Nowadays, carbohydrates are used in a wide variety of applications ranging from pharmaceuticals,⁶ food,⁷ vaccines,⁸ and bio-sensors.⁹ Carbohydrates also have a major influence on our health and are often implicated in diseases.¹⁰⁻¹¹ Every cell – for example, a mammalian cell or a bacterial cell – has a unique carbohydrate layer, which is typically related to their (dys)functioning. For example, surface-bound carbohydrates, such as sialic acids, on cancer cells are often over-expressed or have unusual terminal glycans.¹² The unique cell-surface carbohydrate layer is not encoded in the genes of an organism and the biosynthesis of their complex structures is not template-driven. The tools to study DNA or proteins at the molecular level can therefore not be applied on carbohydrates. The design and chemical synthesis of carbohydrates and enhanced unnatural versions thereof as “molecular tools” can be used to unravel carbohydrate-related biological processes. This chemistry-based approach using molecular tools in combination with analytical techniques such as mass spectrometry and fluorescence microscopy to study biological phenomena is called “chemical biology” and has accelerated the study of carbohydrates.¹³⁻¹⁴ This thesis presents a study that uses such a chemistry-based approach to investigate the role of bacterial carbohydrates in the human gut. The subsequent sections will provide a general introduction on the various scientific disciplines and research subjects involved in this study.

Chemical microbiology

Although an all-encompassing definition of chemical biology is challenging,¹⁵ in the context of this study it is defined as a research discipline where chemists develop tailor-made molecular tools to study and modify biological processes at the molecular level, often in close collaboration with biologists. It is closely related to cell biology and biochemistry, but in chemical biology molecules and their tuning and monitoring abilities are central rather than biochemical pathways. Sub-disciplines within this field are chemical glycobiology, focussing on the carbohydrates involved in biological pathways, or proteomics using molecular tools, mass spectrometry and advanced computer programs to unravel the proteome. Chemical biology, initially, was mainly applied to numerous mammalian cell lines. In recent years, however, chemists started to also use this approach to look in more detail at other organisms such as plants and microbes.¹⁶⁻¹⁸ Chemical biology is a nice example of how our modern interdisciplinary approach to science can bring new insight and techniques to the study of biology. This is, however, far from a new phenomenon. Actually, the first chemist working in the field of microbiology was Louis Pasteur (Figure 2). His early work on isomers of D-tartrate illustrated that microorganisms can catalyse specific chemical reactions.¹⁹ Another pioneer was the Dutch scientist Martinus Beijerinck, whom was originally trained as a botanical expert but changed his research interest towards microbiology

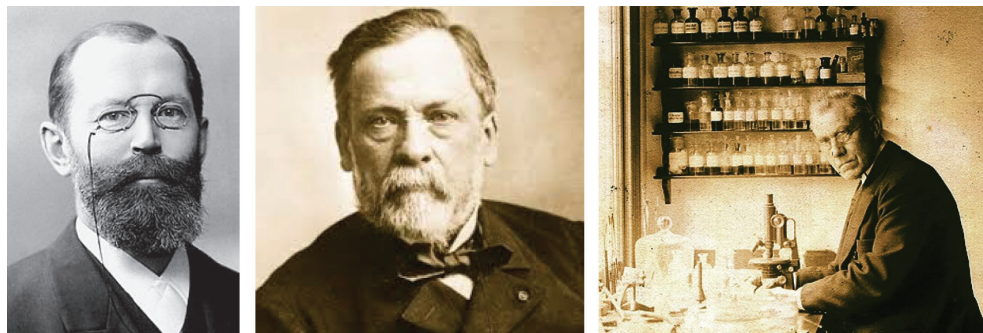


Figure 2. Emil Fischer, 1852 –1919 (left), Louis Pasteur, 1822 –1895 (center) and Martinus Beijerinck, 1851 –1931 (right).

(Figure 2).¹⁹ He formulated the enrichment culture technique. In this technique microorganisms are isolated from their natural habitat and enriched with optimal nutrients and incubation conditions to favour one microorganism over the other.

The overall aim of the study presented in this thesis was to gain new insights in the role of bacterial carbohydrates in the human gut. Chemical biology tools are well suited for this aim. The study of microbes at the molecular level with chemical tools can therefore be seen as a forgotten sub-discipline in chemical biology that only recently received new attention: chemical microbiology.

Gut microbiome

All the various microbes found in and on our body, collectively called the human microbiome, play an essential role in keeping us alive, especially the gut microbiome is important for our well-being. The complexity of the gut microbiome, however, is high, with millions of genes that vastly exceed the coding capacity of our own genome.²⁰ In other words, many diseases can be related to an imbalance in the bacterial composition within the gut. Specific diseases such as Crohn disease, are often directly linked to bacteria in our gut.²¹ Such a direct link is, however, only in specific cases such as *C. difficile* infections applicable.²³ In order to understand the role of our microbiome in even more detail important questions such as “do individuals share a core human microbiome”; “is there a correlation between microbial population structure and host genotype”; “Do differences in the human microbiome correlate with differences in human health”; and “are differences in the relative abundance of bacteria important” have to be answered.²⁴ Currently, the gut microbiome is being mapped in great detail to fully understand how bacteria interact among themselves and with our cells in the human gut.²⁵ In the future a better understanding of the function of our gut microbiome and its systemic impact will result in better diagnostics or therapies to improve the quality of life. The presence of carbohydrates on both the outer surface of gut bacteria and our own cells suggest they play a role on the effect of the microbiome on human health. Besides the common eukaryotic carbohydrate monomers, bacteria

have an extended additional set of unique bacterial monosaccharides.¹⁷ The role of these bacterial monosaccharides is often not fully understood. Especially those human gut bacterial sugars present on their outer surface which are in close contact with humans are of special interest. Our aim is to use chemical microbiology to study carbohydrates in the cell wall of human-gut bacteria at the molecular level.

Peptidoglycan and the bacterial cell envelope

A variety of carbohydrates is present in the bacterial cell envelope and may provide rigidity and shape to the cell, serve as a barrier for nutrients and endogenous molecules, and interact with other bacteria and host cells. Bacteria can be divided into two major classes, the Gram-positive and Gram-negative. The difference between these two classes can be seen in Figure 3 and is caused by a carbohydrate-containing cell wall component, called peptidoglycan, and the presence of an outer membrane in Gram-negative bacteria that is decorated with a variety of carbohydrate-containing molecules.

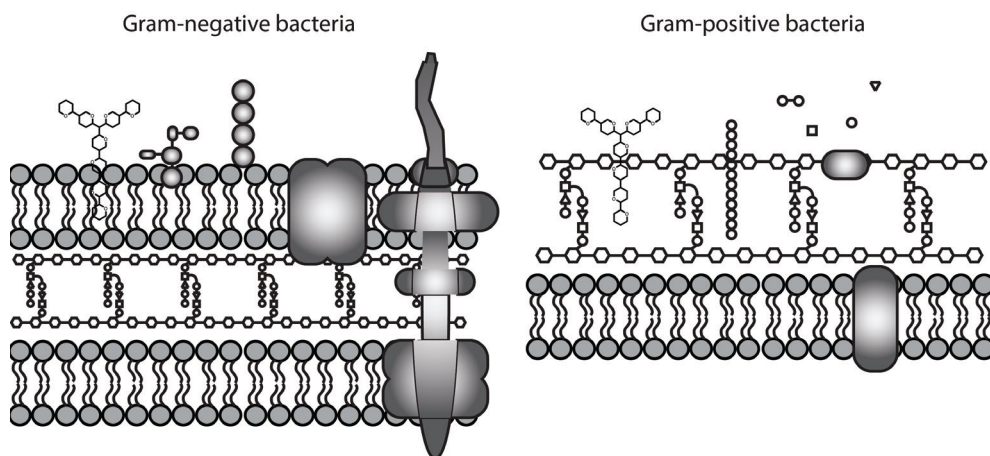


Figure 3. Schematic view of the architecture of the cell envelope in Gram-positive and Gram-negative bacteria.

The Gram-positive cell-envelope (figure 3, left panel) consists of an inner membrane and a large cell wall made up of peptidoglycan that is often covered by glycosylated molecules, such as fimbriae (also termed pili), S-layers or other glycoproteins, and capsular polysaccharides (CPS).

The Gram-negative cell-envelope (figure 3, right panel) includes two membrane, outer and inner membrane that includes the relatively small peptidoglycan layer. The outer membrane is covered with a dense layer of (lipo)proteins, glycans and lipids. These include a lipopolysaccharide layer or simply LPS (Figure 4). The LPS layer can contain a variety of glycans, such as ketodeoxyoctonate (KDO), glucose (Glc), galactose (Gal), and *N*-acetylglucosamine (GlcNAc).

Peptidoglycans are part of both Gram-positive and -negative bacteria and consist of a mesh-like structure of polymeric chains of a repeating disaccharide-penta-peptide unit. The disaccharide

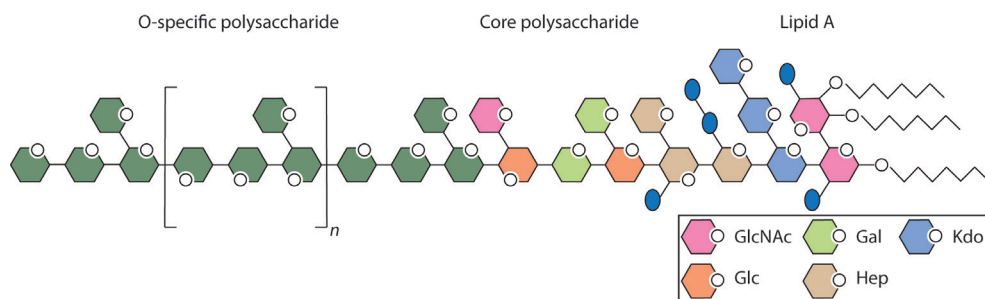


Figure 4. Schematic view of LPS layer of bacteria. GlcNAc: *N*-acetylglucosamine, Glc: glucose, Gal: galactose, Hep: heptose, Kdo: ketodeoxyoctonate.

consist of the two monosaccharides, *N*-acetylglucosamine and *N*-acetylmuramic acid (MurNAc), that are β -(1,4) linked. The penta-peptide consist in most cases of L-alanine (L-Ala), D-glutamic acid (D-Glu), *meso*-diaminopimelic acid (DAP), and two D-alanines (D-Ala). Crosslinking is often seen between the *meso*-diaminopimelic acid residue and D-alanine, and gives extra rigidity and strength to the cell wall. The full penta-peptide chain is synthesised sequentially by ligases inside the bacteria and subsequently exported over the inner membrane and further polymerised. The biosynthesis of peptidoglycan is a complex multistep process where a variety of genes are involved. Recent reviews discuss the biosynthesis in full detail.²⁶⁻²⁷ The presence of D-amino acids in PG is of specific interest since it is absent in mammals. This resulted over the years in antibiotics that specifically target the biosynthesis of these amino acids such as the β -lactams and D-cycloserine. To improve the current antibiotics and develop new, more information is needed to understand PG dynamics in full detail. Furthermore, in a variety of bacteria it is still not known whether PGs are present. Some are completely devoid of PG, such as *Mycoplasma* spp., while for others belonging to the superphylum Planctomycetes-Verrucomicrobia-Chlamydia cluster, the PG presence has been enigmatic. One of these, belonging to the Verrucomicrobia, is *Akkermansia muciniphila*. In **chapter 3** and **4** we describe studies performed on various bacteria, including *A. muciniphila*, in order to investigate their surface glycan structures and for whether they can produce PG. One of the methods to study PG is metabolic labelling, the incorporation of chemical probes in biological systems using the metabolic pathways provided by the organism [see **chapter 2** for a review on metabolic oligosaccharide engineering, **chapter 3** for studies on peptidoglycan labelling, and **chapter 4** for experimental work on metabolic oligosaccharide engineering in human-gut microbes]. Metabolic labelling has proven to be a powerful technique to study PG synthesis and dynamics in more detail. In 2014, Maurelli and co-workers were among the first to use metabolic labelling to replace the natural D-Ala–D-Ala dipeptide for a D-Ala–D-Ala alkyne probe.³² This probe was incorporated directly by the endogenous enzymes in Chlamydiae, demonstrating the presence of PG in bacteria belonging to this phylum. Side-chain modifications on the α -position were already known to not interfere with the biosynthesis of PG. In the biosynthesis of PG a dipeptide intermediate is involved, and for that reason most probes are also based on this dipeptide intermediate. Over the last years several PG dipeptide probes were synthesised (Figure 5).

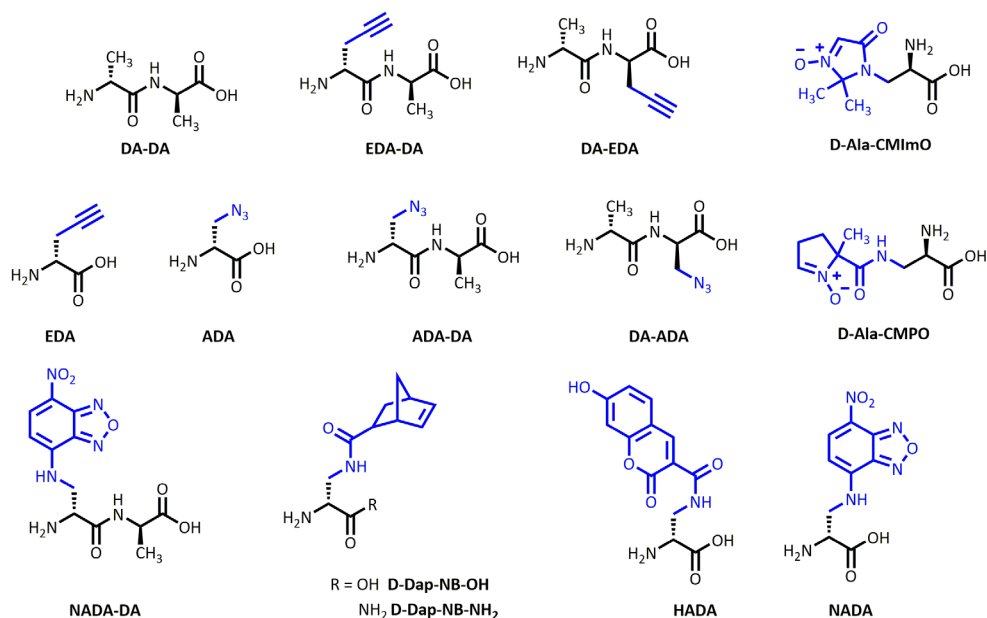


Figure 5. D-Ala and D-Ala-D-Ala peptidoglycan probes.

The molecular mechanism behind the visualisation of most of these PG dipeptide probes relies on a two-step bioconjugation reaction using the copper-catalysed azide alkyne cycloaddition (CuAAC) or Strain-Promoted Azide Alkyne Cycloaddition (SPAAC).³³⁻³⁵ In the first step the unnatural PG dipeptide is taken up by the cell and incorporated into the peptidoglycan layer. In the second step a bioconjugation, for example click chemistry, is performed to visualise the PG dipeptide probe with a fluorescent dye or antibody. A variety of combinations are known nowadays, with azide-alkyne click chemistry to be the most popular choice. The current method developed by Maurelli and co-workers has a few drawbacks.³² The alkyne dipeptide is not commercially available and the click chemistry uses copper, a known toxic in many biological pathways. Other bioconjugation reactions such as SPAAC, tetrazine ligation or Strain-Promoted Alkyne-Nitrone Cycloaddition (SPANC) do not need copper.³⁶⁻³⁹ Chemical PG probes using these bioconjugation reactions have also been developed (Figure 5). Bioconjugation without the use of copper has found applications in many organisms such as mammalian cell lines, plants, and bacteria.¹⁴ New bioconjugation techniques such as SPANC or tetrazine ligation are compatible with the CuAAC or SPAAC and thereby enable dual labelling using both techniques.⁴⁰ This dual labelling has successfully been applied in bacteria to study PG dynamics (*vide infra*).

O-glycans in the Human Gut

Although peptidoglycan forms the largest glycan-polymer in bacteria a variety of other glycans can be found on the human cell membrane. Two distinct types of glycans are: *O*-linked glycans and *N*-linked glycans. A glycan connected to an oxygen of an amino acid such as serine, threonine, or tyrosine via a glycosidic linkage is called a *O*-linked glycan. When glycans are linked to an amine in for example asparagine it is called *N*-linked. The most abundant type of *O*-linked glycans in eukaryotes and bacteria is the mucin-type which is characterised by the presence of *N*-acetyl- α -D-galactosamine.²⁵ Mucin-type *O*-linked glycosylation is an important post-translational modification and involved in a variety of physiological responses and diseases.⁴¹ Mucin-type *O*-glycans dominate the inner surface of the intestinal tract and thereby are suspected to have a significant role in regulating the microbial composition of the gut microbiome. Some bacteria can also degrade the mucin type *O*-glycans themselves. One of this mucin-degrading bacteria is *Akkermansia muciniphila*. *A. muciniphila*, a species from the phylum Verrucomicrobia, was recently discovered in the human intestinal tract.²⁸ *A. muciniphila* is a mucin-degrading human-gut bacterium, and has been implicated in the regulation of body weight and metabolism, is present since early childhood in our gut, and can easily rise to 3% of the total bacterial count in human feces during adulthood.^{28-31, 42-44}

A more detailed study of this species, among others, is important to understand our human microbiome. Metabolic labelling also provides a tool to study these glycans in more detail. A combination of PG and glycan labelling called dual labelling would even make it possible to target and track selectively different glycans under the same physiological conditions. The labelling of other cell glycoconjugates is, of course, not possible with the PG dipeptide probes. For the labelling of these species other probes have, however, been developed. Among the well-studied probes are azido variants of *N*-acetylglucosamine (GlcNAz), *N*-acetylgalactosamine (GalNAz), azido-fucose, and 5-acetylneuraminic acid (Neu5Az).^{14, 45} Almost all monosaccharide within the mucin-type *O*-glycans can be studied with these azido-probes. More recent probes with cyclopropene and isonitrile bioconjugation handles have also been developed and successfully used for metabolic labelling.^{38, 46-48} Other glycans such as glucose or *N*-acetylmannosamine (ManNAc) are, in general, not suitable for metabolic labelling in bacteria, since their biosynthetic pathways are absent in most bacteria. A few examples where Ac₄ManNAz showed metabolic labelling are, however, known.⁴⁹ The role of carbohydrates that are present in gut bacteria is of high interest. Unfortunately selectivity in labelling of these carbohydrates is often the limiting factor. For example, *Bacteroides thetaiotaomicron*, a human gut symbiont is known to degrade more than a dozen types of glycans while species like *Anaerostipes rhamnosivorans* or *Faecalibacterium prausnitzii* (both well-studied butyrate-producing gut bacteria) can only degrade a limited amount of glycans.⁵⁰⁻⁵² Additionally, as mentioned, bacteria also use many monosaccharides and complex carbohydrates that are unique to their domain. If this bacterial carbohydrate composition can be mapped in more detail, a better understanding of these glycoconjugates and the bacteria themselves will be obtained.

Microbial sialic acids

One class of carbohydrates that only occurs in bacteria that interact with humans are microbial sialic acids (MSA). Sialic acids are the most abundant natural family of nonulosonic acids.⁵³ Nonulosonic acids are nine-carbon α -ketoacid sugars typically found on the surfaces of glycoconjugates in mammalian cells. Sialic acids were first identified in vertebrates over 80 years ago by Ernst Klenk and Gunnar Blix.⁵⁴ More than 50 natural occurring sialic acid derivatives are known nowadays in a variety of cellular organisms such as insects and fish eggs.^{4,55} All sialic acids are characterised by their 3-deoxy-D-glycero-D-galacto-2-nonulosonic acid skeleton. At C1 they have a carboxylic acid group which at physiological pH is deprotonated. Three major classes of sialic acid are known: *N*-acetylneuraminic acid (Neu5Ac), *N*-glycolylneuraminic acid (Neu5Gc) and 3-deoxy-D-glycero-D-galacto-2-nonulosonic acid (KDN). Especially Neu5Ac gained a lot of interest over the years due to the connection in mammalian tumours.⁵⁶ Sialic acids are always found as α -linked carbohydrates on glycoconjugates.

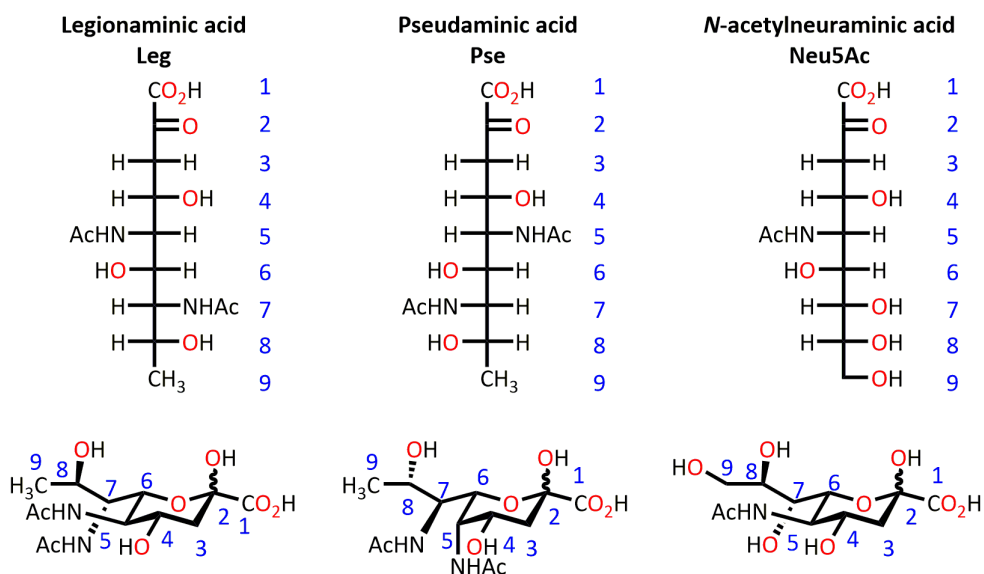


Figure 6. Fischer projection and chemical structure of the microbial sialic acids, legionaminic acid and pseudaminic acid, and the most abundant human sialic acid, *N*-acetylneuraminic acid.

Like other sialic acids, MSAs can be found on *N*- and *O*-linked glycans, glycosphingolipids, glycosaminoglycans, glycosphingolipid anchors, and lipo-oligo/polysaccharides. Sialic acids have been implicated in vital neurological and immunological processes and various diseases such as cancer, the common flu, and cholera (see for excellent reviews Varki and Crockers⁵⁶⁻⁵⁸). As a matter of fact, many pathogenic bacteria have decorated their cell (wall) surfaces with sialic acid residues or derivatives thereof in order to mask their presence for the human immune response.⁵⁹

General Introduction

Some bacterial species can express sialic acid themselves, while other bacteria have developed a unique class of sialic acids which are only present in those microbes of the species that are in close contact with humans. The best studied MSA are legionaminic (Leg) and pseudaminic acid (Pse) (Figure 6).⁶⁰ In the O-chain polysaccharide of the LPS of *Pseudomonas aeruginosa* O7 and O9 and *Shigella boydii* Pse was isolated. In the O-chain of LPS in *Legionella pneumophila* serotype 1 Leg was isolated in 1994.⁶¹ Legionaminic acid greatly resembles N-acetylneuraminic acid. At the C9-position legionaminic acid has a methyl group compared to the C9-hydroxy group in N-acetylneuraminic acid and the 7-OH in Neu5Ac is changed to 7-NHAc in Leg (Figure 6). Pseudaminic acid has a different stereochemistry at the C5, C7, and C8 position compared to legionaminic acid. Thus, legionaminic acid has a D-Gal D-Gly configuration and pseudaminic acid a L-Man L-Gly configuration. Two natural epimers of legionaminic acid are known 4-*epi*-legionaminic acid and 8-*epi*-legionaminic acid with D-glycero-D-talo and L-glycero-D-galacto configuration, respectively. The discovery of pseudaminic acid was relevant, since it was long thought that all nonulosonic acids were direct regio- and stereochemical derivatives of neuraminic acid.⁶² Pseudaminic acid are found both as α - and β -linked carbohydrates, on the terminal end of glycans or internally linked. Most of these linkages are nitrogen-based on either the C-5 or C-7 position.⁶³ *Campylobacter*, *Helicobacter*, *Escherichia*, and *Legionella pneumophila* are among the most interesting species to study pseudaminic and legionaminic acid in more detail, because all these species are major players in diseases related to the human gut. An interesting detail is that all these bacterial species are Gram-negative bacteria. This is mainly because their flagella are heavily glycosylated with pseudaminic acid residues.⁶⁴ It does not, however, mean that other glycoconjugates are absent in Pse, or that Pse is only present in Gram-negative bacteria.

In recent years, microbial sialic acids have been discovered in both non-pathogenic and pathogenic bacteria. Many bacterial species have Pse and/or Leg or epimers thereof in their glycans. A complete overview of species with Pse or Leg was given in 2014 by Zunk and Kiefel.⁶⁰ Derivatives and epimers of these two classes are known and found in a variety of bacterial species such as acinetaminic acid in the capsule of *Acinetobacter baumannii*.⁶⁵ An overview of known natural epimers and derivatives of pseudaminic acid is shown in Figure 7.

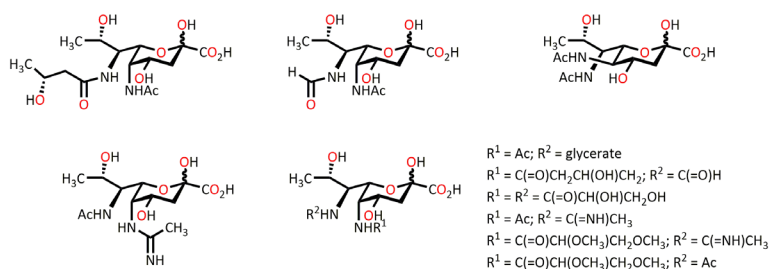


Figure 7. Natural occurring pseudaminic acid derivatives.

Biological significance of pseudaminic acid

In order to unravel the biological role of MSAs we have to identify the classes of proteins involved in their lifecycle. Three distinct classes of proteins are involved in the lifecycle of a MSAs. The first class are the glycosyltransferase enzymes, known as sialyltransferases. Glycosyltransferases can place (hence transfer) the (microbial) sialic acid onto proteins or present glycoconjugates. The second class of proteins, the so-called lectins, has interaction with the MSA on the outer surface. Lectins, carbohydrate-binding proteins, and sometimes also antibodies interact with the sialic acid to trigger a biological response through binding. The last class are the glycosylhydrolases that can hydrolyse the glycosidic bond of the sialic acid with its aglycone, also called sialidases. After this hydrolysis the lifecycle of (microbial) sialic acids ends, it can be recycled or degraded. Nearly nothing is known about these three classes of proteins for MSAs. Improved knowledge about the enzymes involved with MSAs and how the interaction takes place could lead to further understanding of the role of MSAs in bacteria and perhaps also of interactions with their host.⁶⁶

Aim of this thesis

Bacterial carbohydrates form a large group of very diverse carbohydrates which play a major but as of yet poorly understood role in the human body, especially in the gut. Synthetic organic chemistry enables the synthesis of these (complex) bacterial carbohydrates, and the design of tailor-made enhanced versions that can function as molecular tools to study the biology. This thesis presents a study in which we aim to investigate the role of bacterial carbohydrates in the human gut. We aim to contribute to unravelling the role of these bacterial carbohydrates in the human gut using a combination of synthetic organic chemistry and microbiology: chemical microbiology.

Outline of this thesis

In this thesis we have made a start to explore the role of bacterial sugars in the human gut. For the study of peptidoglycan we designed and synthesised two new peptidoglycan probes. This work is described in **chapter 3**. Labelling was performed in both *E. coli* and *A. muciniphila* with these novel probes. The newly developed PG probes were non-toxic as showed by viable count analysis. Other glycoconjugates in *E. coli* and *A. muciniphila* were also targeted with azido-variants of Ac₄GlcNAc, Ac₄GalNAc, and Ac₄FucAc. This technique is called metabolic oligosaccharide engineering and is described in a review in **chapter 2**. For visualisation of the labelling we performed confocal microscopy with all samples. To quantify this labelling flow cytometry was used. The orthogonal isonitrile PG and azido glycan probes were also used in a dual labelling experiment to label PG and fucose-glycans simultaneously.

Experimental work with MOE in human gut related bacteria is described in **chapter 4**. A small collection of azido-glycans, Ac₄GlcNAz, Ac₄GalNAz, Ac₄FucAz, Ac₄RhamAz, and Neu5Az were used in relevant human-gut bacteria such as *A. muciniphila*, *A. rhamnosivorans*, and other butyrate-producing bacteria. As a proof of principle MOE was tested in a co-culture of a combination of these gut bacteria. Confocal microscopy and flow cytometry were used to visualise and quantify the labelling.

In order to label more complex carbohydrates in the human gut or man-related bacteria we aimed for the total synthesis of pseudaminic acid. This microbial sialic acid is found on the outer surface of bacteria. Chemical synthesis towards Pse was known at the start of this work, but no extensive chemical biology experiments had been performed. First, a literature overview of known syntheses towards Pse and Leg is discussed. Next, we describe our efforts towards these chemical probes of Pse in **chapter 5**. Key reactions to obtain Pse chemical probes, such as the Henry reaction and tethered aminohydroxylation, were performed and optimised where possible. The Garner aldehyde derivative, derived from L-threonine, proved to be a versatile important building block towards Pse.

Finally, **chapter 6** provides a general discussion about all previous chapters. Furthermore it discusses unpublished work and looks into the future of this exciting and only new-born field, chemical microbiology.

References

1. Stick, R. V., *Carbohydrates: The Sweet Molecules of Life*. **2001**.
2. Lindhorst, T., *Essentials of Carbohydrate Chemistry and Biochemistry*. **2007**.
3. Varki, A., *Glycobiology* **2017**, 27 (1), 3-49.
4. Varki, A., *Nature* **2007**, 446 (7139), 1023-9.
5. *The Nobel Prize in Chemistry*. **1902**.
6. Ernst, B.; Magnani, J. L., *Nat. Rev. Drug Discov.* **2009**, 8 (8), 661-77.
7. Cui, S. W., *Food Carbohydrates: Chemistry, Physical Properties, and Applications*. **2005**.
8. Hecht, M. L.; Stallforth, P.; Silva, D. V.; Adibekian, A.; Seeberger, P. H., *Curr. Opin. Chem. Biol.* **2009**, 13 (3), 354-9.
9. Geissner, A.; Anish, C.; Seeberger, P. H., *Curr. Opin. Chem. Biol.* **2014**, 18, 38-45.
10. Guarner, F.; Malagelada, J. R., *Lancet* **2003**, 361 (9356), 512-9.
11. Koropatkin, N. M.; Cameron, E. A.; Martens, E. C., *Nat. Rev. Microbiol.* **2012**, 10 (5), 323-335.
12. Pifferi, C.; Berthet, N.; Renaudet, O., *Biomater. Sci.* **2017**.
13. Aebersold, R.; Mann, M., *Nature* **2003**, 422 (6928), 198-207.
14. Sminia, T. J.; Zuilhof, H.; Wennekes, T., *Carbohydr. Res.* **2016**, 435, 121-141.
15. Bucci, M.; Goodman, C.; Sheppard, T. L., *Nat. Chem. Biol.* **2010**, 6 (12), 847-54.
16. Hoogenboom, J.; Berghuis, N.; Cramer, D.; Geurts, R.; Zuilhof, H.; Wennekes, T., *BMC Plant. Biol.* **2016**, 16 (1), 220.
17. Tra, V. N.; Dube, D. H., *Chem. Commun.* **2014**, 50 (36), 4659-73.
18. Zhu, Y.; Wu, J.; Chen, X., *Angew. Chem.* **2016**, 55 (32), 9301-5.
19. Madigan, M. T.; Martinko, J. M.; Bender, K. S.; Buckley, D. H.; Stahl, D. A., *Brook biology of microorganisms*, **2015**.
20. Qin, J.; Li, R.; Raes, J.; Arumugam, M.; Burgdorf, K. S.; Manichanh, C.; Nielsen, T.; Pons, N.; Levenez, F.; Yamada, T.; Mende, D. R.; Li, J.; Xu, J.; Li, S.; Li, D.; Cao, J.; Wang, B.; Liang, H.; Zheng, H.; Xie, Y.; Tap, J.; Lepage, P.; Bertalan, M.; Batto, J. M.; Hansen, T.; Le Paslier, D.; Linneberg, A.; Nielsen, H. B.; Pelletier, E.; Renault, P.; Sicheritz-Ponten, T.; Turner, K.; Zhu, H.; Yu, C.; Li, S.; Jian, M.; Zhou, Y.; Li, Y.; Zhang, X.; Li, S.; Qin, N.; Yang, H.; Wang, J.; Brunak, S.; Dore, J.; Guarner, F.; Kristiansen, K.; Pedersen, O.; Parkhill, J.; Weissenbach, J.; Meta, H. I. T. C.; Bork, P.; Ehrlich, S. D.; Wang, J., *Nature* **2010**, 464 (7285), 59-65.
21. Hill-Burns, E. M.; Debelius, J. W.; Morton, J. T.; Wissemann, W. T.; Lewis, M. R.; Wallen, Z. D.; Peddada, S. D.; Factor, S. A.; Molho, E.; Zabetian, C. P.; Knight, R.; Payami, H., *Mov. Disord.* **2017**.
22. Sokol, H.; Pigneur, B.; Watterlot, L.; Lakhdari, O.; Bermudez-Humaran, L. G.; Gratadoux, J. J.; Blugeon, S.; Bridonneau, C.; Furet, J. P.; Corthier, G.; Grangette, C.; Vasquez, N.; Pochart, P.; Trugnan, G.; Thomas, G.; Blottiere, H. M.; Dore, J.; Marteau, P.; Seksik, P.; Langella, P., *Proc. Natl. Acad. Sci. USA* **2008**, 105 (43), 16731-6.
23. de Vos, W. M.; de Vos, E. A., *Nutr. Rev.* **2012**, 70 Suppl 1, S45-56.
24. Shetty, S. A.; Hugenholtz, F.; Lahti, L.; Smidt, H.; de Vos, W. M., *FEMS Microbiol. Rev.* **2017**, 41 (2), 182-199.
25. Tytgat, H. L.; de Vos, W. M., *Trends Microbiol.* **2016**, 24 (11), 853-861.
26. Gautam, A.; Vyas, R.; Tewari, R., *Crit. Rev. Biotechnol.* **2011**, 31 (4), 295-336.
27. Lovering, A. L.; Safadi, S. S.; Strynadka, N. C., *Annu. Rev. Biochem.* **2012**, 81, 451-78.
28. Derrien, M.; Vaughan, E. E.; Plugge, C. M.; de Vos, W. M., *Int. J. Syst. Evol. Microbiol.* **2004**, 54 (Pt 5), 1469-76.
29. Derrien, M.; Van Baaren, P.; Hooiveld, G.; Norin, E.; Muller, M.; de Vos, W. M., *Front. Microbiol.* **2011**, 2, 166.
30. Derrien, M.; Belzer, C.; de Vos, W. M., *Microb. Pathog.* **2016**.
31. Plovier, H.; Everard, A.; Druart, C.; Depommier, C.; Van Hul, M.; Geurts, L.; Chilloux, J.; Ottman, N.; Duparc, T.; Lichten

- stein, L.; Myridakis, A.; Delzenne, N. M.; Klievink, J.; Bhattacharjee, A.; van der Ark, K. C.; Aalvink, S.; Martinez, L. O.; Dumas, M. E.; Maiter, D.; Loumaye, A.; Hermans, M. P.; Thissen, J. P.; Belzer, C.; de Vos, W. M.; Cani, P. D., *Nat. Med.* **2016**.
32. Liechti, G. W.; Kuru, E.; Hall, E.; Kalinda, A.; Brun, Y. V.; VanNieuwenhze, M.; Maurelli, A. T., *Nature* **2014**, 506 (7489), 507-10.
33. Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B., *Angew. Chem.* **2002**, 41 (14), 2596-9.
34. Tornøe, C. W.; Christensen, C.; Meldal, M., *J. Org. Chem.* **2002**, 67 (9), 3057-64.
35. Agard, N. J.; Prescher, J. A.; Bertozzi, C. R., *J. Am. Chem. Soc.* **2004**, 126 (46), 15046-7.
36. Stockmann, H.; Neves, A. A.; Stairs, S.; Brindle, K. M.; Leeper, F. J., *Org. Biomol. Chem.* **2011**, 9 (21), 7303-5.
37. McKay, C. S.; Chigrinova, M.; Blake, J. A.; Pezacki, J. P., *Org. Biomol. Chem.* **2012**, 10 (15), 3066-70.
38. Wainman, Y. A.; Neves, A. A.; Stairs, S.; Stockmann, H.; Ireland-Zecchini, H.; Brindle, K. M.; Leeper, F. J., *Org. Biomol. Chem.* **2013**, 11 (42), 7297-7300.
39. Sherratt, A. R.; Chigrinova, M.; McKay, C. S.; Beaulieu, L. P. B.; Rouleau, Y.; Pezacki, J. P., *RSC Adv.* **2014**, 4 (87), 46966-46969.
40. Sherratt, A. R.; Chigrinova, M.; MacKenzie, D. A.; Rastogi, N. K.; Ouattara, M. T.; Pezacki, A. T.; Pezacki, J. P., *Bioconjug. Chem.* **2016**, 27 (5), 1222-6.
41. Hang, H. C.; Bertozzi, C. R., *Bioorg. Med. Chem.* **2005**, 13 (17), 5021-34.
42. Collado, M. C.; Derrien, M.; Isolauri, E.; de Vos, W. M.; Salminen, S., *Appl. Environ. Microbiol.* **2007**, 73 (23), 7767-70.
43. Derrien, M.; van Passel, M. W.; van de Bovenkamp, J. H.; Schipper, R. G.; de Vos, W. M.; Dekker, J., *Gut microbes* **2010**, 1 (4), 254-268.
44. Lievin-Le Moal, V.; Servin, A. L., *Clin. Microbiol. Rev.* **2006**, 19 (2), 315-37.
45. Laughlin, S. T.; Bertozzi, C. R., *Nat. Protoc.* **2007**, 2 (11), 2930-44.
46. Patterson, D. M.; Nazarova, L. A.; Xie, B.; Kamber, D. N.; Prescher, J. A., *J. Am. Chem. Soc.* **2012**, 134 (45), 18638-43.
47. Yang, J.; Seckute, J.; Cole, C. M.; Devaraj, N. K., *Angew. Chem.* **2012**, 51 (30), 7476-9.
48. Stairs, S.; Neves, A. A.; Stockmann, H.; Wainman, Y. A.; Ireland-Zecchini, H.; Brindle, K. M.; Leeper, F. J., *ChemBioChem* **2013**, 14 (9), 1063-7.
49. Geva-Zatorsky, N.; Alvarez, D.; Hudak, J. E.; Reading, N. C.; Erturk-Hasdemir, D.; Dasgupta, S.; von Andrian, U. H.; Kasper, D. L., *Nat. Med.* **2015**, 21 (9), 1091-100.
50. Bui, T. P.; de Vos, W. M.; Plugge, C. M., *Int. J. Syst. Evol. Microbiol.* **2014**, 64 (Pt 3), 787-93.
51. Marcobal, A.; Barboza, M.; Sonnenburg, E. D.; Pudlo, N.; Martens, E. C.; Desai, P.; Lebrilla, C. B.; Weimer, B. C.; Mills, D. A.; German, J. B.; Sonnenburg, J. L., *Cell host & microbe* **2011**, 10 (5), 507-14.
52. Salyers, A. A.; Vercellotti, J. R.; West, S. E.; Wilkins, T. D., *Appl. Environ. Microbiol.* **1977**, 33 (2), 319-22.
53. Angata, T.; Varki, A., *Chem. Rev.* **2002**, 102 (2), 439-69.
54. Faillard, H., *Trends Biochem. Sci.* **1989**, 14 (6), 237-241.
55. Varki, N. M.; Varki, A., *Lab. Invest.* **2007**, 87 (9), 851-7.
56. Varki, A., *Trends Mol. Med.* **2008**, 14 (8), 351-60.
57. Crocker, P. R., Siglecs: *Curr. Opin. Struct. Biol.* **2002**, 12 (5), 609-15.
58. Chen, X.; Varki, A., *ACS Chem. Biol.* **2010**, 5 (2), 163-76.
59. Severi, E.; Hood, D. W.; Thomas, G. H., *Microbiology* **2007**, 153 (Pt 9), 2817-22.
60. Zunk, M.; Kiefel, M. J., *RSC Adv.* **2014**, 4 (7), 3413-3421.

61. Knirel, Y. A.; Rietschel, E. T.; Marre, R.; Zahringer, U., *Eur. J. Biochem.* **1994**, 221 (1), 239-45.
62. Knirel, Y. A.; Vinogradov, E. V.; L'Vov V, L.; Kocharova, N. A.; Shashkov, A. S.; Dmitriev, B. A.; Kochetkov, N. K., *Carbohydr. Res.* **1984**, 133 (2), C5-8.
63. Knirel, Y. A.; Shevelev, S. D.; Perepelov, A. V., *Mendeleev Commun.* **2011**, 21 (4), 173-182.
64. Wilhelms, M.; Fulton, K. M.; Twine, S. M.; Tomas, J. M.; Merino, S., *J. Biol. Chem.* **2012**, 287 (33), 27851-62.
65. Kenyon, J. J.; Marzaioli, A. M.; De Castro, C.; Hall, R. M., *Glycobiology* **2015**, 25 (6), 644-654.
66. Baker, J. L.; Celik, E.; DeLisa, M. P., *Trends Biotechnol.* **2013**, 31 (5), 313-23.



Chapter 2

Getting a Grip on Glycans:

a Current Overview of the Metabolic Oligosaccharide Engineering Toolbox

"Wat's... Uh The Deal"

Obscured by Clouds, Pink Floyd

Abstract

This review discusses the advances in metabolic oligosaccharide engineering (MOE) from 2010-2016 with a focus on the structure, preparation, and reactivity of its chemical probes. A brief historical overview of MOE is followed by a comprehensive overview of the chemical probes currently available in the MOE molecular toolbox and the bioconjugation techniques they enable. The final part of the review focusses on the synthesis of a selection of probes and finishes with an outlook on recent and potential upcoming advances in the field of MOE.

This work was published in

T. J. Sminia, H. Zuilhof, T. Wennekes, *Carbohydrate Research*, **2016**, 435, 121

1. Introduction

The field that studies glycobiology has been in existence for well over a century and was first properly defined over 25 years ago.¹ Since then the intimate and complex involvement of carbohydrates, also called glycans, in a vast number of health- and disease-associated biological communication processes has been further substantiated and many new connections have been discovered. One can even consider them a companion to nucleic acids and proteins as key players in the storage and transfer of information in virtually all organisms. In stark contrast to these established 'central dogma' biomolecules, however, stands our understanding of the relationship between glycan molecular structure and their biology activity. The reason we are still almost literally in the dark is that studying glycans poses unique challenges when compared with proteins and nucleotides, related to their highly complex, often branched structures. An additional challenge is that the assembly of glycans – their structure – is not template-driven (encoded in genes), and instead occurs via complex and dynamic processing and trimming of the glycan chain under the action of a series of competitive glycosyltransferase and glycosidase enzymes. The study of how glycans are assembled by glycosyltransferases, catabolised by glycosidases, recognised by lectins, and how or when taken together this enables the dynamic storage and transfer of information via their molecular structure is therefore at the frontier of science. The field that uses a chemistry-based approach to answers these largely unanswered questions about glycobiology has been coined 'Chemical Glycobiology'. The 15 years that have passed since it was first defined as such by Bertozzi have seen it blossoming and many excellent reviews cover all these advances.²⁻⁴

The current review focusses on a chemical glycobiology strategy that uses the glycan salvage pathways that most organisms possess. These pathways allow a cell to convert monosaccharides, liberated from complex glycans or originating from the extracellular space, into their nucleotide-activated counterparts that are substrates for glycosyltransferases located in the endoplasmic reticulum and Golgi apparatus. This eventually results in the monosaccharide becoming part of complex cellular

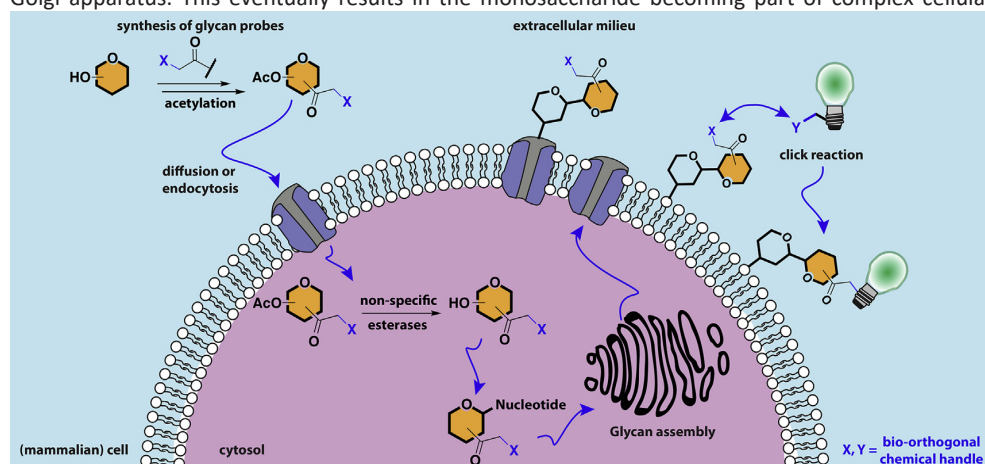


Figure 1. Schematic overview of metabolic oligosaccharide engineering.

glycan structures, e.g. *N*-glycans and *O*-glycans on proteins and glycolipids (Figure 1). In the 1990's it was discovered that the enzymes that participate in these glycan salvage pathways sometimes allow for unnatural modifications in the structure of their monosaccharide-based substrate.^{5,6} This enabled the design of tailor-made monosaccharide-based chemical probes that are labelled with a small unnatural chemical handle that can be addressed after incorporation under specific conditions at a predefined moment. Processing of these probes via the glycan salvage pathways allows for direct profiling of cellular glycans. This technique is nowadays often called metabolic oligosaccharide engineering (MOE) and it complements other approaches to profile cellular glycans, like tracing incorporation of isotope (radio)labeled glycans, indirect profiling using (chemical) lectins or direct chemical tagging of native glycans, e.g. the use of sodium periodate-mediated generation of taggable aldehydes from neuraminic acid. For a comprehensive overview of all the applications developed thus far with MOE we direct the reader to various recent reviews.⁷⁻¹⁸ This review focusses on another integral part of MOE, namely the fact that all chemical probes evaluated for MOE have to be designed and synthesised. We here provide an up-to-date comprehensive overview of the current toolbox of chemical probe structures implemented in MOE, i.e. the chemical handles employed to enable (bioorthogonal) bioconjugation reactions for profiling labeled glycans. We illustrate their use with selected recent examples of applications in various organisms. Additionally we provide a selected overview of various synthetic routes that are nowadays used to obtain these probes. The review finishes with an outlook that highlights some recent and prospective advances in the application of MOE.

2. Historical overview of MOE and bioconjugation reactions

Key to the success of any metabolic oligosaccharide engineering approach is the ability of the labelled glycans to be selectively tagged after incorporation via a bioconjugation reaction. The requirements for bioconjugate reactions suitable for use with MOE are largely known, such as compatibility with complex biological mixtures, stability of the bioconjugate, high reaction rates, and ideally no interference with native biochemical processes (known as bio-orthogonality).¹⁷ Many of the currently available bioconjugation reactions have already been applied in MOE (see Table 1), and many efforts to further extend and improve the application and capabilities of bioconjugation reactions are ongoing. In this section we limit ourselves to a brief summary; the development, reaction kinetics, and general handling of these reactions are highlighted in several recent reviews.^{13, 16, 17, 19-21}

Among the first bioconjugation reactions to be applied for MOE were aldehyde/ketone condensation reactions with hydrazides and aminoxy groups to give hydrazones and oximes, respectively (Figure 2).²² With the discovery of azide and alkyne conjugations, the field of metabolic labelling exploded.²³ Azides are reduced under Staudinger conditions and trapped intramolecularly with a biocompatible phosphine reagent as discovered by the Bertozzi group.²³ In addition, azides are easily used in the Copper-catalysed Azide Alkyne Cycloaddition (CuAAC) reaction.²⁴ In the presence of copper(I) ions, azide and alkyne substrates react rapidly to form stable triazoles. This hallmark reaction in click chemistry is nowadays still the most used technique. In order to exclude toxic copper from biological systems two

Overview of the Metabolic Oligosaccharide Engineering Toolbox

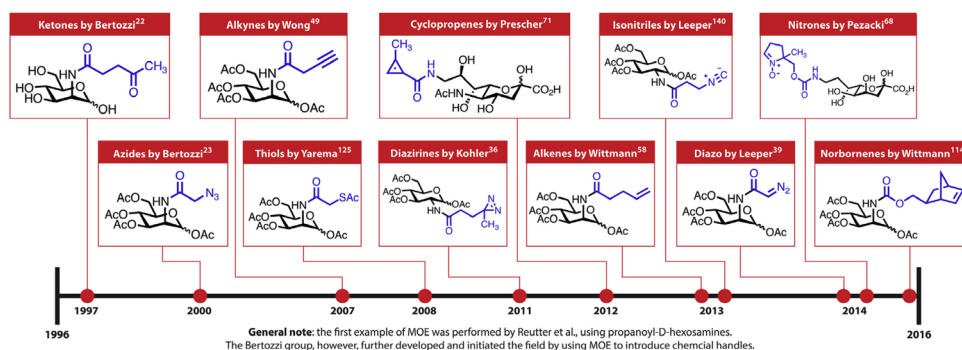


Figure 2. Timeline of the history of metabolic oligosaccharide engineering probes.

methods were developed. The first method uses amine-triazole based ligands to accelerate the CuAAC reaction and limit the required amount of copper, while the other method uses strained alkynes to improve the reaction kinetics in a copper-free manner.⁷ The use of strained alkynes in biological systems is known as the Strain-Promoted Azide Alkyne Cycloaddition (SPAAC) reaction.²⁵ Despite the exclusion of copper, the SPAAC reaction is not always superior over the CuAAC reaction, since some strained alkyne reagents are known to have a higher background labelling due to side reactions with cellular thiols.²⁶ Several recent advances in the field have introduced photoactivation to provide spatial and temporal control in bioconjugation reactions. The Boons laboratory introduced a photochemically triggered SPAAC click reaction with a dibenzocyclooctyne derivative.²⁷ A photocontrolled version of the Staudinger-Bertozzi ligation has also been reported and successfully applied in living organisms such as zebrafish.²⁸ A recent reinvestigation of the photo-induced 1,3-dipolar cycloaddition between an alkene and tetrazole has questioned its bio-orthogonality and alkene selectivity of the reactive intermediate nitrile imine.²⁹ This bioconjugation reaction has to the best of our knowledge not yet been applied in MOE.

Table 1. Overview of bioconjugation reactions.

| Name | Reactant 1 | Reactant 2 | Remarks | Average reaction rate ³⁰ | REF |
|-----------------------------------|----------------------------|-----------------|--|---|--------|
| Staudinger-Bertozzi ligation | azide | phosphine | phosphazene trapped intramolecularly as amide | $k \sim 10^{-3} \text{ M}^{-1} \text{ s}^{-1}$ | 23 |
| CuAAC | alkyne | azide | Cu-catalysed with ascorbic acid Cu(I)-regeneration | $k \sim 10\text{-}200 \text{ M}^{-1} \text{ s}^{-1}$ | 31, 32 |
| SPAAC | strained alkyne | azide | Ring strain is driving force of the reaction | $k \sim 10^2\text{-}1 \text{ M}^{-1} \text{ s}^{-1}$ | 25 |
| Hydrazide ketone ligation | ketone | hydrazide | product is hydrazone | $k \sim 10^3 \text{ M}^{-1} \text{ s}^{-1}$ | 5, 6 |
| Oxime ligation | ketone | amino-oxy | product is oxime | unknown | 33 |
| invDA | (strained/terminal) alkene | tetrazine | dihydropyrazine product | $k \sim 1\text{-}10^4 \text{ M}^{-1} \text{ s}^{-1}$ | 34 |
| 1,3-dipolar cycloaddition | (strained/terminal) alkene | nitrile imine | nitrile imines are formed <i>in situ</i> after UV activation of tetrazoles | $k \sim 890 \text{ M}^{-1} \text{ s}^{-1}$ | 35 |
| Photo-activated 1,3-cycloaddition | terminal alkene | tetrazole | product is pyrazole | $k \sim 4.41 \pm 0.34 \times 10^{-2} \text{ M}^{-1} \text{ s}^{-1}$ | 27 |
| Photo-activated diazirines | diazirine | | UV activation needed | unknown | 36 |
| Isonitrile ligation | isonitrile | tetrazine | stable product for tertiary isonitrile | $k \sim 57.5 \pm 1.5 \times 10^{-2} \text{ M}^{-1} \text{ s}^{-1}$ | 37 |
| Hetero Diels Alder | vinyl sulfide | diol | product is thioacetal | $k \sim 1.5 \pm 0.1 \times 10^{-3} \text{ M}^{-1} \text{ s}^{-1}$ | 38 |
| Diazo ligation | diazo | strained alkyne | product is pyrazole | $k \sim 0.04 \text{ M}^{-1} \text{ s}^{-1}$ | 39 |
| SPOQC | strained alkyne | quinone | among fastest known bioconjugation reactions | $k \sim 496 \pm 70 \text{ M}^{-1} \text{ s}^{-1}$ | 40 |
| SPANC | alkyne | nitron | | $k \sim 39 \text{ M}^{-1} \text{ s}^{-1}$ | 41 |

Other photochemical activation reactions have also been used in combination with MOE. Photoactivation by UV irradiation of aryl azides, diazirines, and benzophenone results in the formation of reactive nitrene, carbene and radical intermediates, respectively, that react with almost any structure in close proximity, even aliphatic C-H bonds.⁴² These photo-active probes can be taken up by the cell without significant toxic effects; however, UV irradiation can of course damage the cells. The generated reactive intermediate will react with proteins in the near vicinity of the labelled glycan, and thereby provide fruitful insights in protein-glycan interactions in biological pathways.

A recently discovered bioconjugation reaction is the inverse electron demand Diels Alder reaction (invDA).^{34, 43} The reaction is fast and irreversible, and most tetrazines are easy to make and stable in aqueous media. For MOE, the dienophile can be a terminal alkene, cyclopropene, diazo, or isonitrile moiety. The invDA reactions are superior over the CuAAC or SPAAC reaction in terms of reaction kinetics and biocompatibility. A photoactivatable tetrazine for use in invDA was recently reported by Fox and co-workers.⁴⁴ Dihydropyrazines were oxidised to tetrazines upon irradiation by red light (660 nm) in

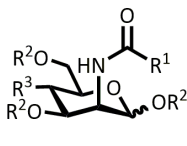
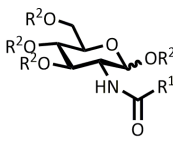
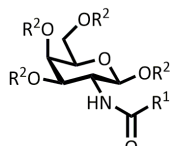
the presence of methylene blue. Upon addition of strained alkenes (e.g. *trans*-cyclooctene (TCO)), the generated tetrazine reacts rapidly with these. The same dihydrotetrazine precursors could also be chemo-enzymatically activated using a peroxidase. Beside the three main classes of bioconjugation, azide-alkyne, tetrazine invDA ligation and photo-chemical activation, other methods are also being explored for use in MOE, but not widely applied. Several recent developments are discussed in the outlook.

3. Overview of chemical probes their applications and tagging-mechanism

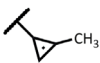
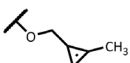
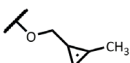

The tagging of labelled glycans after a MOE experiment can be performed with a wide variety of different chemical handles. Azides, alkynes, and cyclopropenes are among the most popular nowadays. Other frequently used chemical handles are diazirines, diazo, isonitriles, thiols, and nitrones. The chemical synthesis of these chemical handles varies from basic amide couplings to more complex synthetic strategies. In most routes the endpoint of the chemical synthesis is a fully acetylated monosaccharide-based chemical probe. The overall *modus operandi* of MOE is that acetylated monosaccharide-based chemical probes are fed to target cells or an organism, as the acetyl esters facilitate the passive uptake of the probe.^{5, 6} An additional beneficial result of acetylated probes is the easier purification thereof by standard procedures such as flash column chromatography. Once inside the cell, the acetyl groups are hydrolysed by nonspecific cytosolic esterases and the probe is processed by the glycan salvage pathway and transformed into nucleotide-activated donor substrates for glycosyltransferase enzymes in the glycosylation pathways. These glycosylation pathways take place in the endoplasmic reticulum and Golgi apparatus of the cell. There, the labelled-donors are incorporated into the many complex cellular glycan structures, e.g. *N*-glycans and *O*-glycans on proteins and glycolipids. Finally, the small chemical handle can be addressed to directly profile the targeted monosaccharide in the cells complex glycans.

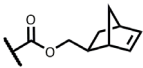
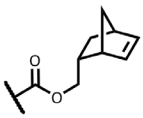
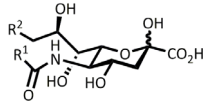
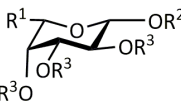
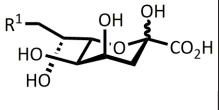
Since its inception in 1997 an extensive toolbox of chemical probes has been developed for use in MOE. Table 2 provides a comprehensive overview of the glycan-based chemical probes that have thus far been used in MOE and tested in a variety of organisms from mammalian cell lines to zebrafish and bacteria, and more recently also in plants. The table is organised per glycan and functional group. In general, the ManNAc, GlcNAc, and GalNAc salvage pathways are most tolerant towards unnatural glycans with small or sometimes even substantial chemical handles, as can be seen by their overwhelming presence in Table 2. Especially acyl substituents on the 2-position of these 2-deoxy amino glycans are well tolerated. In early work by the Bertozzi group, 2-deoxy 2-azido glycans were found to be either toxic or not converted to cell-surface glycans, indicating that the *N*-acyl group itself is essential for enzymatic recognition and conversion of ManNAc analogues.⁴⁵ The overview of Table 2 is followed by sections, subdivided on the basis of the chemical handles in these probes, that discuss their development, mode of action and selected applications.

Table 2. Overview of glycan-based chemical probes for metabolic oligosaccharide engineering.

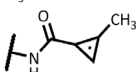
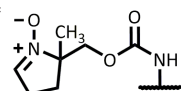
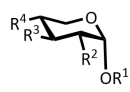
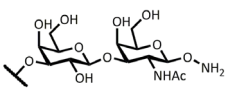
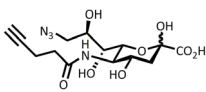
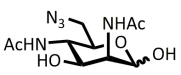
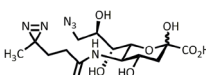
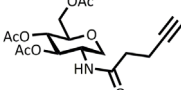
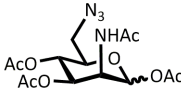
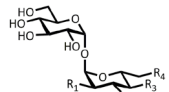
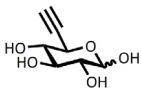
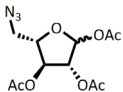
| Label | <i>N</i> -acetylmannosamine | REF | <i>N</i> -acetylglucosamine | REF | <i>N</i> -acetylgalactosamine | REF |
|---------------------|--|--------|--|-----|--|-----|
| |  | |  | |  | |
| Natural glycans | ManNAc R ¹ = CH ₃ R ² = O R ³ = H | 6 | GlcNAc R ¹ = CH ₃ R ² = H | 4 | GalNAc R ¹ = CH ₃ R ² = H | 6 |
| Aliphatic analogues | ManNPro R ¹ = CH ₂ CH ₃ R ² = O R ³ = H | 6 | GlcNPro R ¹ = CH ₂ CH ₃ R ² = H | 6 | GalNPro R ¹ = CH ₂ CH ₃ R ² = H | 6 |
| Ketones | ManNLev R ¹ = (CH ₂) ₂ COCH ₃ R ² = O R ³ = H | 22 | GlcNLev R ¹ = (CH ₂) ₂ COCH ₃ R ² = H | 22 | GalNLev R ¹ = (CH ₂) ₂ COCH ₃ R ² = H | 22 |
| Azides | Ac₄ManNAz R ¹ = CH ₂ N ₃ R ² = O R ³ = Ac | 23 | Ac₄GlcNAz R ¹ = CH ₂ N ₃ R ² = Ac | 23 | Ac₄GalNAz R ¹ = CH ₂ N ₃ R ² = Ac | 23 |
| | ManNAz R ¹ = CH ₂ N ₃ R ² = O R ³ = H | 23 | GlcNAz R ¹ = CH ₂ N ₃ R ² = H | 23 | GalNAz R ¹ = CH ₂ N ₃ R ² = H | 23 |
| | 1,3,4-O-Bu₃ManNAz R ¹ = CH ₂ N ₃ R ² = R ³ = Bu | 46 | | | | |
| | Ac₄ManN2Azaryl R ¹ = CH ₂ PhN ₃ R ² = O R ³ = Ac | 47 | | | | |
| | Ac₃-4-Az-ManNAc R ¹ = OAc R ² = Ac R ³ = N ₃ | 48 | | | | |
| Alkynes | Ac₄ManNAIk R ¹ = (CH ₂) ₂ C≡CH R ² = O R ³ = Ac | 49, 50 | Ac₄GlcNAIk R ¹ = (CH ₂) ₂ C≡CH R ² = Ac | 51 | Ac₄GalNAIk R ¹ = (CH ₂) ₂ C≡CH R ² = Ac | 52 |
| | Ac₄ManPoc R ¹ = OCH ₂ C≡CH R ² = O R ³ = Ac | 53 | Ac₄GlcPoc R ¹ = OCH ₂ C≡CH R ² = Ac | 53 | Ac₄GalPoc R ¹ = OCH ₂ C≡CH R ² = Ac | 53 |

Overview of the Metabolic Oligosaccharide Engineering Toolbox

| | | | | | | |
|------------------|--|--------|--|--------|---|----|
| Thiols | Ac₄ManNTGc R ¹ = CH ₂ SAc R ² = O R ³ = Ac | 54 | Ac₄GlcNTGc R ¹ = CH ₂ SAc R ² = Ac | 54 | Ac₄GalNTGc R ¹ = CH ₂ SAc R ² = Ac | 54 |
| Cyclopropene | Ac₄ManNCyc R ¹ =  R ² = O R ³ = Ac | 55 | | | | |
| | Ac₄ManNCyoc R ¹ =  R ² = O R ³ = Ac | 56 | Ac₄GlcNCyoc R ¹ =  R ² = Ac | 57 | Ac₄GalNCyoc R ¹ =  R ² = Ac | 57 |
| Terminal alkenes | Ac₄ManNPTl R ¹ = (CH ₂) ₂ -CH=CH ₂ R ² = O R ³ = Ac | 58 | | | | |
| | Ac₄ManNHxl R ¹ = (CH ₂) ₃ -CH=CH ₂ R ² = O R ³ = Ac | 58 | | | | |
| | Ac₄ManNALoc R ₁ = OCH ₂ -CH=CH ₂ R ₂ = O R ₃ = Ac | 59 | Ac₄GlcNALoc R ¹ = OCH ₂ -CH=CH ₂ R ² = Ac | 59 | | |
| | Ac₄ManNBeoc R ¹ = O(CH ₂) ₂ -CH=CH ₂ R ² = O R ³ = Ac | 59 | Ac₄GlcNBeoc R ¹ = O(CH ₂) ₂ -CH=CH ₂ R ² = Ac | 59 | | |
| | Ac₄ManNPeoc R ¹ = O(CH ₂) ₃ -CH=CH ₂ R ² = O R ³ = Ac | 59 | Ac₄GlcNPeoc R ¹ = O(CH ₂) ₃ -CH=CH ₂ R ² = Ac | 59 | | |
| | Ac₄ManNHeoc R ¹ = O(CH ₂) ₄ -CH=CH ₂ R ² = O R ³ = Ac | 59, 60 | Ac₄GlcNHeoc R ¹ = O(CH ₂) ₄ -CH=CH ₂ R ² = Ac | 58, 59 | | |
| Diazirines | Ac₄ManNDAz(3Me) R ¹ = (CH ₂) ₃ -1-methyl diaziriny R ² = O R ³ = Ac | 61 | Ac₄GlcNDAz(3Me) R ¹ = (CH ₂) ₃ -1-methyldiaziriny R ² = Ac | 61 | Ac₄GalNDAz(2Me) R ¹ = (CH ₂) ₂ -1-methyl diaziriny R ² = Ac | 62 |

| | | | | | | |
|-------------------|--|-----|--|----|--|----|
| | Ac₄ManNDaz(4Me) $R^1 = (\text{CH}_2)_4\text{-1-methyl}$ diaziriny $R^2 = \text{O}$ $R^3 = \text{Ac}$ | 61 | Ac₄GlcNDaz(4Me) $R^1 = (\text{CH}_2)_4\text{-1-methyldiaziriny}$ $R^2 = \text{Ac}$ | 61 | | |
| Isonitrile | Ac₄ManN-<i>n</i>-Iso $R^1 = (\text{CH}_2)_2\text{-N}\equiv\text{C}$ $R^2 = \text{O}$ $R^3 = \text{Ac}$ | 63 | Ac₄GlcN-<i>n</i>-Iso $R^1 = (\text{CH}_2)_2\text{-N}\equiv\text{C}$ $R^2 = \text{Ac}$ | 63 | Ac₄GalN-<i>n</i>-Iso $R^1 = (\text{CH}_2)_2\text{-N}\equiv\text{C}$ $R^2 = \text{Ac}$ | 63 |
| | Ac₄ManN-<i>t</i>-Iso $R^1 = \text{C}(\text{CH}_3)_2\text{CH}_2\text{-N}\equiv\text{C}$ $R^2 = \text{O}$ $R^3 = \text{Ac}$ | 63 | Ac₄GlcN-<i>t</i>-Iso $R^1 = \text{C}(\text{CH}_3)_2\text{CH}_2\text{-N}\equiv\text{C}$ $R^2 = \text{Ac}$ | 63 | Ac₄GalN-<i>t</i>-Iso $R^1 = \text{C}(\text{CH}_3)_2\text{CH}_2\text{-N}\equiv\text{C}$ $R^2 = \text{Ac}$ | 63 |
| Diazo- ketones | Ac₄ManDiaz $R^1 = \text{CHN}_2$ $R^2 = \text{O}$ $R^3 = \text{Ac}$ | 52 | | | Ac₄GalDiaz $R^1 = \text{CHN}_2$ $R^2 = \text{Ac}$ | 39 |
| Nor bornene | Ac₄ManNNorboc_{exo} $R^1 =$  $R^2 = \text{Ac}$ Ac₄ManNNorboc_{endo} $R^1 =$  $R^2 = \text{Ac}$ | 114 | | | | |
| | Sialic acid | | Fucose | | KDO | |
| |  | |  | |  | |
| | Neu5Hex $R^1 = (\text{CH}_2)_3\text{C}\equiv\text{CH}$ $R^2 = \text{OH}$ | 64 | FucAl or 6-alkynyl Ac₄fucose $R^1 = \text{C}\equiv\text{CH}$ $R^2 = \text{Ac}$ $R^3 = \text{Ac}$ | 65 | 9Al-KDO $R^1 = \text{C}\equiv\text{CH}$ | 66 |
| | SiaNAI $R^1 = (\text{CH}_2)_2\text{C}\equiv\text{CH}$ $R^2 = \text{OH}$ | 67 | | | KDO-alkyne $R^1 = \text{O}(\text{CH}_2)_2\text{C}\equiv\text{CH}$ | 68 |
| | SiaNProc $R^1 = \text{OCH}_2\text{C}\equiv\text{CH}$ $R^2 = \text{OH}$ | 69 | Ac₄FucAz $R^1 = \text{CH}_2\text{N}_3$ $R^2 = R^3 = \text{OAc}$ | 70 | KDO-N₃ $R^1 = \text{N}_3$ | 66 |

Overview of the Metabolic Oligosaccharide Engineering Toolbox

| | | | | | | |
|--|--|----|---|--------|--|-----|
| | 9-Cp-NeuAc R ¹ = CH ₃ R ² =  | 67 | FucAz R ¹ = CH ₂ N ₃ R ² = GDP R ³ = H | 70, 72 | KDO-HMMPO R ¹ =  | 68 |
| | 9BA-Neu5Ac R ¹ = OH R ² = NHCH ₂ COC ₆ H ₄ COH | 73 | | | | |
| | Xylose | | Miscellaneous | | Two functionalities | |
| |  | |  Hydrazone | 74 |  9AI-KDO | 75 |
| | UDP-2-XylAz R ¹ = UDP R ² = N ₃ R ³ = OAc R ⁴ = OAc | 76 |  6Az-2,4-diacetamido2,4,6-trideoxy-D-mannopyranose | 77 |  9AzSiaDAz | 75 |
| | UDP-3-XylAz R ¹ = UDP R ² = OH R ³ = N ₃ R ⁴ = OH | 76 |  1-deoxy-Ac₃GlcNAIk | 51 | | |
| | UDP-3-XylAz R ¹ = UDP R ² = OH R ³ = OH R ⁴ = N ₃ | 76 |  6Az-Ac₃ManNAc | 45 | | |
| | Trehalose | | | | Glucose | |
| |  | | | |  6dAG | 78 |
| | 2-TreAz: R ¹ = N ₃ ; R ² = OH; R ³ = OH; R ⁴ = OH | 79 | | | Arabinose | |
| | 3-TreAz: R ¹ = OH; R ² = N ₃ ; R ³ = OH; R ⁴ = OH | 79 | | |  Ac₃ArabAz | 156 |
| | 4-TreAz: R ¹ = OH; R ² = OH; R ³ = N ₃ ; R ⁴ = OH | 79 | | | | |
| | 6-TreAz: R ¹ = OH; R ² = OH; R ³ = OH; R ⁴ = N ₃ | 79 | | | | |

3.1 Azides

MOE with azido-labelled glycans is still the most widely adopted variant, and is used in a broad range of applications. Azido-containing MOE probes have been successfully used for metabolic labelling in various organisms and tissues, such as pancreatic cells^{46, 80}, mice⁸¹, and primary hippocampus neuron cells.⁸² Several recent reviews can be found that specifically cover this topic.^{17, 83, 84} Perhaps their most famous application, the labelling of neuraminic acid in complex glycans, is covered in these reviews. This section focusses on their recent application in glycan labelling in *C. elegans*, zebrafish, bacteria, and plants.

Currently, all azido probes are *O*-acetylated when used in MOE. Despite this being the current standard method, investigating other esters or groups to mask the probe and facilitate uptake by the target organism or cell might be a way to further optimise MOE. Yarema et al. have shown that *O*-butyrate probes are also successfully taken up and recognised by the cellular metabolic machinery.⁸⁶ Since the early discovery that azido-acyl groups on amino sugars are well tolerated and recognised by the enzymatic machinery in organisms, no further changes have been made to this golden standard chemical handle. The tolerance for bigger chemical handles such as cyclopropenes and diazirine suggests that the design of the azido acetyl groups might still be further explored.

***C. elegans*:** The efficiency of metabolic incorporation of azidosugars in *C. elegans* was initially unknown. Therefore, Bertozzi and co-workers first determined the ability for metabolic incorporation of Ac₄ManNAz, Ac₄GalNAz, and Ac₄GlcNAz into *C. elegans* glycans and glycoconjugates.⁸⁷ This revealed that *C. elegans* is unable to incorporate Ac₄ManNAz and Ac₄GlcNAz, but labelling with Ac₄GalNAz worked. The azido-glycans were first incubated in *E. coli* and the bacteria were afterwards consumed by *C. elegans*. It is therefore unknown whether *C. elegans* can actually take up these azido-glycans itself or if *E. coli* metabolises them and *C. elegans* obtains them by metabolism of *E. coli* glycans. It should also be noted that MOE in *E. coli* with Ac₄GalNAz or Ac₄GlcNAz has never been reported, while it has been reported for other Gram-negative bacteria.^{88, 89} Recently, labelling of glyconjugates using Ac₄GlcNAz was also observed in primary embryonic *C. elegans* cells.⁹⁰

Zebrafish: The zebrafish is an excellent organism to study *O*-glycosylation and *N*-glycosylation with MOE and the fact that it is a transparent organism makes fluorescent imaging easier. The research of MOE in zebrafish started in the group of Bertozzi in 2008.⁹¹ Ac₄GalNAz was the only studied substrate and showed incorporation in *O*-glycans at concentrations higher than 25 μ M, as was visualised with DIFO-488. Sialylated glycans were studied in developing zebrafish using Ac₄ManNAz. Here MOE clearly proved superior over the previous performed work on sialylated glycans using aldehyde ligation techniques by Baskin et al.⁹²

In later studies fucosylation was studied in developing zebrafish. The fucose salvage pathway is active in zebrafish embryos, but efficient processing of azide-modified fucose was poor. Azido-fucose labelling strategies in zebrafish are therefore limited to fucosylated glycoproteins active during the development. In addition, it has been reported that in certain cases azido-fucose can be toxic for

Overview of the Metabolic Oligosaccharide Engineering Toolbox

mammalian cells.^{49, 93} For this reason, the use of alkynyl-fucose analogues in mammalian cell lines has since been the preferred approach.

Bacteria: Glycans are important for the viability of bacteria and almost all bacteria require glycans as a carbon source for growth. A wide variety of glycans is exclusively found in microbes, and the uniqueness of these bacterial glycans makes them ideal targets for the development of new antibiotics that interfere with the bacterial metabolism.⁹⁴ For this reason, metabolic labelling of bacterial glycans is a rapidly growing field of research. Both pathogenic bacteria - like *Haemophilus ducreyi*,⁹⁵ *Campylobacter jejuni*,⁹⁶ and *Helicobacter pylori* - and non-pathogenic bacteria - like *Escherichia coli* MG1655,⁶⁶ and *Bacteroides fragilis*,^{89, 97} have been successfully labelled in this way.

In bacteria, probe incorporation and competing probe degradation via catabolic processes can also play an important role and is highly dependent on the type of bacteria. Rather high concentrations of azido-glycans are usually added in MOE experiments with bacteria, compared to other organisms or cell lines, to compensate for this factor. These rather high concentrations have, however, not shown significant toxicity in bacteria. Even the copper used for the CuAAC reaction, which is toxic for most organisms, is better tolerated by bacteria. If needed, a biocompatible CuAAC reagent, e.g. the triazole ligand BTAA, can be used to further reduce copper toxicity as shown by Wu and co-workers.⁹⁷ Since many SPAAC reagents have become commercially available (e.g. BCN-rhodamine), both SPAAC and CuAAC have been successfully applied in MOE on bacteria. The wide variety of non-common glycans present in bacteria requires extensive and complex synthetic routes to systematically generate the probes needed to label specific rare microbial glycans. Several synthetic strategies to these unique bacterial glycans have, however, been reported, and a selection thereof is detailed in the next section.^{96, 98-103} One interesting example is the metabolic incorporation of azido-Pseudaminic acid (Az-Pse), a bacterial glycan belonging to the class of sialic acids.⁹⁶ Mutant *C. jejuni* cells deficient in Pse pathway were fed with the azido-Pse precursor, 6-deoxy-AltdiNAz. This resulted in the decoration of the bacterial flagella with azido-Pse. The presence of azido-Pse on the flagella restored the motility of *C. jejuni*. Subsequently, the azido-glycan was visualised using biotin-Staudinger reagent and later visualised using streptavidin which was separated on SDS-PAGE. The single labelling in whole-cell flagellin lysates, as shown by SDS-PAGE, indicates that flagellin proteins are the only cell-surface-associated proteins that are glycosylated with Pse. The study of these bacterial glycans is therefore of interest as a potential new antibiotic target.

Another Gram-negative bacterium which has been evaluated with MOE, using Ac₄GlcNAz and GlcNAz, is *H. pylori*. Both azido-glycans were tested to check the need for *O*-acetylation, to investigate the presently still unknown mechanism of uptake by bacteria, and to probe for the presence of esterase activity inside the bacteria. The per-acetylated GlcNAz was easily incorporated into *N*-glycans of *H. pylori*¹⁰⁴, but incorporation in *O*-linked glycans was not seen. Furthermore, the non-acetylated GlcNAz was not metabolically consumed and installed on cell surface glycoconjugates. Recently, an extensive study with a panel of azido-glycans was recently performed by Kasper and co-workers on human-gut related pathogenic and non-pathogenic bacteria in Gram-negative and Gram-positive bacteria.⁸⁹ Prior

to this work, Seibel and co-workers looked into MOE of *S. aureus* with GlcNAz.¹⁰⁵

As mentioned, many bacteria use microbe-specific glycans to construct their cell wall, and the disaccharide trehalose is such a glycan. Bertozzi and co-workers introduced an azide on this disaccharide for metabolic labelling in mycobacteria, a class of bacteria that includes the pathogen *Mycobacterium tuberculosis*.⁷⁹ Four different trehalose-azide analogues were synthesised to study a recycling process that might be conserved across mycobacterial species. All four trehalose-azide analogues showed labelling in a variety of mycobacteria. In a next step this method can therefore be used to study glycolipid distribution, trafficking, and dynamics of trehalose biosynthesis, which in turn might provide clues on how to target these bacteria with new classes of antibiotics.

Another relevant microbe-specific glycan is 3-deoxy-D-manno-octulosonic acid (KDO), an essential and characteristic part of lipopolysaccharides in Gram-negative bacteria.¹⁰⁶ To characterise Gram-negative bacteria, metabolic labelling with KDO-based probes is thus highly relevant. Vauzeilles and co-workers synthesised a 9Az-KDO glycan (see table 2), and incubated this together with several classes of bacteria.⁶⁶ Successful labelling was achieved in *E. coli* K12, *E. coli* O86, *S. Typhimurium*, and *L. pneumophila*.

Azido-glycans in combination with MOE are proving to be a valuable tool to label glycans in many types of bacteria. Increasing our insight into bacterial glycobiology through among other MOE and finding ways to perturb bacterial glycans in pathogenic bacteria will hopefully provide new leads for the development of much needed next-generation antibiotics.

Plants: Metabolic oligosaccharide engineering is relatively new to plant sciences. The first implementation of MOE in plants was reported in 2012 by Somerville and co-workers.⁶⁵ They successfully incorporated FucAl into *Arabidopsis thaliana* and subsequent CuAAC labelling allowed fluorescence imaging of fucosylated glycans in the cell wall. Imaging indicated that the fucose-alkyne containing glycans were initially incorporated to locations on the cell surface that are likely involved in exocytic vesicle fusion. Recently, azido-KDO was metabolically incorporated into the RG-II glycan of the pectin-rich cell wall of *A. thaliana* and *Nicotiana tabacum*.¹⁰⁷ Successful labelling via CuAAC ligation with Alexa Fluor® 488-alkyne was observed. Dual labelling using Fuc-Al and KDO-N₃ was performed by taking advantage of the two bio-orthogonal reactions. This allowed simultaneous imaging of Fuc-Al in RG-I and KDO-N₃ in RG-II. Chen and coworkers have reported on the metabolic incorporation and imaging of *N*-linked glycans in *A. thaliana* with Ac₄GlcNAz.¹⁰⁸ Through MOE and direct fluorescent imaging of *N*-glycans they showed that protein *N*-glycosylation in *A. thaliana* root tissue possess distinct patterns of distribution in the various developmental zones.¹⁰⁸ We recently also reported on imaging of plant glycans in *A. thaliana* with both Ac₄GlcNAz and GlcNAz, the latter probably via active uptake. In our study we observed that *A. thaliana* can interestingly also metabolically incorporate Ac₄GalNAz, probably via an as of yet unidentified enzymatic GalNAc to GlcNAc epimerase activity. We also observed successful MOE in *A. thaliana* with Ac₄FucNAz and the novel L-arabinofuranose probe, Ac₃ArabAz. Finally we also demonstrated the successful application of strain-promoted azide-alkyne

Overview of the Metabolic Oligosaccharide Engineering Toolbox

cycloaddition (SPAAC) and inverse electron demand Diels-Alder (invDA) click reactions to allow for improved imaging of MOE in live plants.¹⁵⁶

3.2 Alkynes

Alkynes were among the first handles used in metabolic oligosaccharide engineering. The virtual absence of alkynes in nature, their relative small size, their fast reactivity with azides under CuAAC conditions, and their stability makes them very suitable for metabolic labelling. An alkyne derivative of *N*-acetylmannosamine (25 μ M) and 6-alkynyl Ac₄fucose (200 μ M) were tested in Jurkat cells back in 2007 by Chi-Huey Wong and co-workers.⁴⁹ In mammals, *N*-acetylmannosamine is the precursor for the sialic acid, *N*-acetylneuraminic acid (Neu5Ac), and the alkynylated probe versions therefore result in alkynylated sialic acids on glycoconjugates. The fucose alkyne was able to label fucosylated glycans *in vivo*.

A more advanced study that compared Ac₄ManNAI with Ac₄ManNAz in six different cell lines at a concentration of 50 μ M was reported by Bertozzi et al. in 2009.¹⁰⁹ In all six cell lines metabolic labelling of Neu5Ac in glycans with the alkynyl-glycans was more efficient than with the azido-glycans. In a next step, *in vivo* analysis of the alkynyl glycan incorporation in mice showed that after a daily injection of Ac₄ManNAI for seven days, organs were harvested to show metabolic labelling in all tissues.

MOE in bacteria with alkyne glycans has been performed in the Gram-negative class bacteroides. MOE was achieved with alkyne-fucose analogues in *Bacteroides fragilis* and *Parabacteroides distasonis*. A wide variety of glycoproteins proved enriched with alkyne-fucose analogues.⁹⁷ However, the rising popularity of the SPAAC reaction has made the alkynyl glycans that require the generally toxic copper for CuAAC less popular for metabolic labelling. This trend is despite the discovery of bio-compatible copper reagents. The advantage of these reagents is the reduced amount of free Cu(I) ions which is beneficial for labelling in living organisms.¹¹⁰

McClosky et al. have developed a 6-deoxy-alkynyl glucose analogue for metabolic glucose labelling experiments in *A. thaliana* roots.⁷⁸ They based their choice for the alkyne handle on the unreported assumption that azido analogues of fucose are less well incorporated in the cell wall of plants, compared to their alkyne analogues. Acetylated 6-deoxy-alkynyl glucose successfully labelled glycans in *A. thaliana* roots. Other relevant alkynyl-functionalised analogues of glucose, rhamnose, mannose, and sucrose were also tested, but did not show any incorporation. Whether 6-deoxy-alkynyl glucose was selectively incorporated in the cell wall or also in other cellular components was not yet investigated, but would of course be highly interesting. These initial examples of the application of MOE in plants highlights the potential for the further development of this technique to study glycobiology in plants.

3.3 Cyclopropenes, norbornenes, and alkenes

Alkene-functionalised glycans have also been successfully used for MOE. The loss in reaction speed for alkene-tetrazine conjugation compared to CuAAC conjugation is compensated by the stability and easy synthesis of alkene carbohydrates. Wittmann and co-workers synthesised a variety of alkene-equipped *N*-acetylmannosamine analogues.⁵⁸ The length of the linker did not influence the successful labelling in HEK293T and Hela cells.⁵⁸ In later studies, the length of the linker going from ManNAloc to ManHeoc showed an increase in second-order rate constant for the invDA with tetrazines.⁵⁹ New possibilities in MOE emerged with the advent of strained alkenes (or alkynes) that react extremely fast under physiological conditions via an inverse electron demand Diels-Alder reaction (invDA). Early investigations with trans-cyclooctene (TCO) and bicyclononyne (BCN) showed the potential of this new ligation technique in cellular environments.¹¹¹

Cyclopropenes are strained alkenes that are ideal for MOE with respect to their bio-orthogonal reactivity, and small size (similar to diazirines). Obtaining sufficient stability is essential for the reactive cyclopropenes probes, as it is known that they are prone to polymerisation, and can be reactive towards thiols and other nucleophiles present in cells.¹¹² Substituents on the cyclopropene ring, such as a methyl, showed a big improvement in stability, while retaining their reactivity.^{71, 111} A few years later a carbamate linker instead of an amide linker to the glycan further improved the stability and reactivity.^{57, 113} The groups of Devaraj and Prescher worked independently on cyclopropene-linked glycan analogues for MOE.^{55, 71, 111} After Devaraj proved the biocompatibility of cyclopropenes in phospholipid labelling, the Prescher group showed the first application of MOE with a cyclopropene-linked sialic acid analogue (9-Cp-Neu5Ac).^{71, 111} Flow cytometry analysis with different concentrations of the probe showed efficient metabolic incorporation, although the optimum was found at the relative high concentration of 1 mM. In a second step, the compatibility during dual labeling with an azido sialic acid probe was tested. Both cyclopropene and azido sialic acids were added to Jurkat cells and incubated for 24 hours. Successful labelling with corresponding fluorescent dyes showed the orthogonality of tetrazine click with cyclopropene glycans and the SPAAC click with azido glycans. In subsequent years, the groups of Devaraj, Prescher, and Wittmann have reported even more successful examples of cyclopropenes glycans.^{55, 57, 113} The utilised concentrations nowadays vary between 10-100 μ M. The independent investigations of cyclopropene glycans resulted in different names for the same glycan analogues. Two distinct classes, the amide-linked cyclopropenes and the carbamate-linked cyclopropenes have been developed. The amide-linked cyclopropenes are known as NMCp or NCyc for the 3-methylated cyclopropene glycan; the carbamate-linked cyclopropenes are named NCycoc or CCp.

In recent years, the effect of either an amide or carbamate linker between the glycan and cyclopropene moiety has been investigated. Different probes based on glucosamine, galactosamine, and mannosamine with the different linkers have been synthesised and tested in various cell lines. Prescher et al. took a closer look at the effect of the carbamate linker in comparison to the amide linker for mannose-based chemical probes.⁵⁷ The invDA-reactions are approx. 100 times faster with carbamates at C-3 compared to their amide derivative. A clear 130-fold increase in fluorescence signal

was obtained by flow cytometry when the two linkers were directly compared. The disappearance of fluorescence output after addition of the natural *N*-acetylmannosamine substrate suggested that the cyclopropene-ManNAc used the sialic acid biosynthetic pathway.

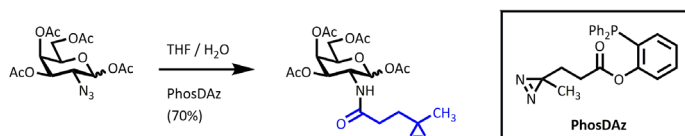
Cyclopropene-labelled glycans are attractive probes for use in MOE. They display fast enough reaction kinetics, short syntheses for their preparation and different variants (amide and carbamate scaffolds) are available. Over the last years, the relative high concentration of tetrazine counterpart needed for post incorporation tagging has also been reduced and is now in the range of 50-100 μ M, further enhancing the potential of this method.⁵⁶ Another alkene variant was recently reported by Wittmann and co-workers.¹¹⁴ Norbornene-modified mannosamine derivatives were used for MOE and successfully applied in HEK 293T cells. *Exo* and *endo* norbornene-modified mannosamine derivatives were made. The reaction kinetics of *exo* norbornene mannosamine proved faster than those of the *endo* norbornene mannosamine when reacted in an invDA reaction with a tetrazine.

3.4 Diazirine

Diazirines are small highly strained heterocycles. The metabolic labelling step is followed by UV activated crosslinking of the incorporated diazirine with nearby biomolecules. When the cell is irradiated at a 254-400 nm wavelength N_2 will be expelled leaving a highly reactive carbene on the chemical handle attached to the glycan, which will react very rapidly with a wide variety of different functional groups inside or on the outside of the cell. The advantage of this photo-crosslinking is that it can be initiated in intact cells. Furthermore, various processes surface-glycans participate in can be studied (e.g. sialidases activity or lectin interactions). The use of photo-crosslinking is not limited to cell-surface glycans but has of course also successfully applied to other biomolecules.⁴² A disadvantage of diazirines as chemical handles in probes is their lack of selectivity. Once the carbene is generated it will react with everything in close proximity.

Before diazirines were explored as metabolic labelling functionalities Paulson et al. reported in 2005 the use of aryl azides on Neu5Ac (known as 9-AAz-NeuAc) as a labelling agent for cell-surface glycoproteins or glycolipids.¹¹⁵ A benefit of aryl azides is the possibility of using the Staudinger-Bertozzi ligation as an extra option to address the labelled glycans. Both tagging methods proved successful in labelling cell surface glycans. The use of diazirines in MOE was developed further by the group of Kohler.⁴² Different types of glucosamine, mannosamine, and galactosamine *N*-acyl-diazirine derivatives were synthesised and tested in a wide variety of cell types. The effect of *N*-acyl chain length of diazirines on MOE was investigated by Kohler and co-workers in 2011.⁶¹ Diazirine-modified *N*-acetylmannosamine analogues with 2, 3 or 4 methylene groups in between the *N*-acyl chain and diazirine were synthesised and tested in BJA-B, Daudi, and Jurkat cells. All probes were incorporated on sialylated glycoprotein or sialylated glycolipids to some extent. It was, however, clearly seen that the short linker (2CH₃) was incorporated most efficiently, and this resulted in effective cross-linking after photo-irradiation.

A combination of the Staudinger-Bertozzi ligation and UV irradiation was performed by Jewett and co-workers.⁶² They developed a reagent that uses the traceless Staudinger-Bertozzi ligation to convert azides on glycans into an acyl-diazirines handle (Scheme 1). The method has not yet been tested on azido-glycans inside cells yet, but it looks promising.

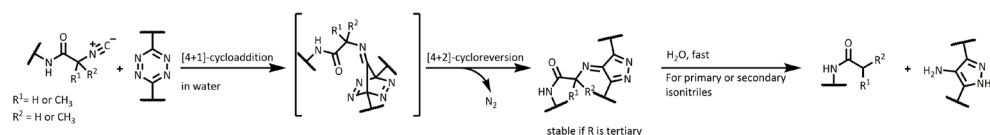


Scheme 1. A traceless Staudinger-Bertozzi reagent to install diazirines onto azido-glycans.

3.5 Isonitrile group

Isonitriles, also known as isocyanides, are overall neutral molecules that can undergo a [4+1] cycloaddition with tetrazines or strained alkenes. Although isonitriles have been found in some natural products, they are only weak nucleophiles and show excellent stability in biological systems.¹¹⁶ Hence, isonitriles do not react with common electrophiles like ketones, aldehydes or imines. Although isonitrile-containing probes themselves are stable, their imine-containing products after a [4+1] cycloaddition can decompose over time (Scheme 2). Inside cells, cycloaddition products of primary or secondary isonitriles are easily hydrolysed into amines. Products from tertiary isonitriles are more stable but bigger and therefore harder to incorporate into glycoconjugates. Primary isonitriles with an additional methylene group tautomerise after the click to a more stable α,β -unsaturated imine. Primary isonitrile-containing glycans were used at high concentration (200 μ M) to test for toxicity, but none was observed.³⁷

After initial experiments with isonitrile labelling on proteins proved successful, the Leeper group recently showed the potential of isonitriles as new metabolic oligosaccharide probes for the *in vitro* labelling of mammalian cells.^{37, 63} Three different probes based on *N*-acylated mannosamine, galactosamine, and glucosamine were synthesised and successfully tested on Lewis lung carcinoma cells. A two-step labelling method with tetrazine–biotin followed by neutravidin–DyLight680 provided the best signal to noise ratio and flow cytometry was used to quantify incorporation. This study showed that these isonitrile glycan probes were successfully incorporated into glycans and are stable over the time period of labelling. Interestingly, when comparing incorporation of the glucosamine- and mannosamine-based probes they show a reversed trend in comparison with azido-glycans. It is well known that compared to ManNAz, GlcNAz is poorly incorporated into glycoconjugates, but Ac₄GlcN-



Scheme 2. Tetrazine ligation with primary, secondary, or tertiary isonitriles.

n-Iso was efficiently incorporated, whereas Ac₄ManN-*n*-Iso showed poor incorporation.¹¹⁷ However, overall azido glycans appear to be incorporated slightly better than isonitrile glycans.

In conclusion, isonitrile-containing glycans have been successfully applied in MOE, but for not yet fully understood reasons they need relatively high concentration (200 μ M) compared to azido glycans (50 μ M). Similar to click reactions with alkene/cyclopropene versus azides, isonitriles show orthogonal reactivity to azides, which enables future use in dual-labelling MOE experiments.

3.6 Diazo group

In 2014 the Leeper group reported a new bioconjugation reaction with the diazoacetyl group.³⁹ This group is smaller than its azide counterpart and reacts under physiological conditions with strained alkenes or alkynes.¹¹⁸ Furthermore, the diazoacetyl group can be used in parallel with an alkyne-based chemical handle to enable dual bio-orthogonal labelling strategies.

The reaction of diazo compounds with strained alkenes or alkynes has been known for many years. For instance, their reactivity with cycloalkynes like DIBO and DIFO has been reported by Raines et al. and indicated their potential as chemical handles in metabolic probes.¹¹⁸ The Leeper and Raines groups independently investigated the reaction kinetics of various diazo compounds with cycloalkynes and alkenes (for both $\pm 0.04 \text{ M}^{-1} \text{ s}^{-1}$), solvent and stability.^{39, 118} The Leeper group reported the first *in vivo* labelling with diazo glycans.³⁹ A two-step [3+2] cycloaddition with simple diazo building blocks was performed. In the first step an invDA reaction was performed with the highly strained (*E,E*-1,5-cyclooctadiene (COD) (Figure 3). In the second step a [3+2] cycloaddition isomerisation results in stable adducts. Imaging the incorporation of a diazoacetyl-GalNAc derivative (Ac₄GalNDiaz) was achieved in Lewis lung carcinoma (LL2) cells via a two-step method using an initial click reaction with biotin-TMDIBO, a stable dibenzocyclo-octyne, and subsequent visualisation with avidin or neutravidin-appended fluorescent dyes. This two-step labelling procedure provided a better signal-to-background ratio compared to a one-step labelling method. Although successful, the level of labelling achieved with Ac₄GalNDiaz, as analysed by flow cytometry, was significantly lower when compared with the Ac₄GalNAz. The exact reason(s) for this low level of incorporation or subsequent labelling is still not known, but one reason might be that the diazo-glycan is a poor substrate for the enzymes involved in the galactosamine salvage pathway or for the associated glycosyltransferases.

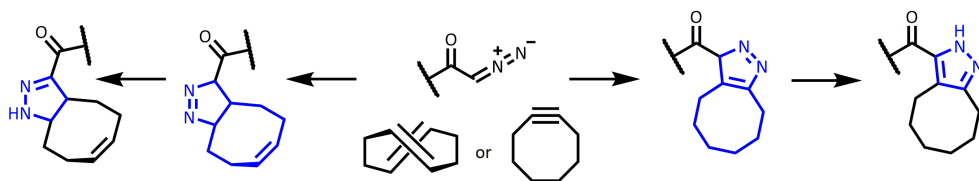


Figure 3. Mechanisms [3+2] cycloaddition of diazo substrates with strained alkenes or alkynes.

A peracetylated diazo-mannose glycan (Ac₄ManDiaz) has been successfully incorporated in Jurkat, CHO K1, HEK293T and HeLa cell lines by the Raines group.⁵² This matches the fact that the enzymes involved in the conversion of *N*-acetylmannosamine into Neu5Ac have a high tolerance for alternatively *N*-acylated derivatives.⁸⁵ However, also here Ac₄ManDiaz was incorporated less efficiently compared to ManNAz as shown by flow cytometry. It is known that diazo compounds are prone to degradation at low pH via C-protonation and subsequent hydrolysis. So if endosomes are involved in the uptake of diazo-glycans, this could also explain the general lower incorporation level.

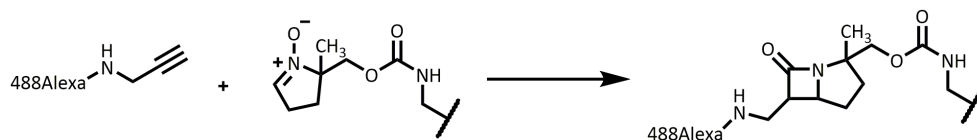
Toxicity of diazo glycans was only observed after prolonged labelling times of 60 minutes at 10 μ M concentration. Diazo compounds are not affected by the CuAAC reaction and a successful dual-labelling strategy with azido-glycans has been explored.⁵² Analysis by flow cytometry and microscopy showed that both click reactions could be performed without cross-reactivity and are independent of the labelling order. No cross-linking was observed for diazo- and azido-glycans, which is an undesired known phenomenon for alkyne and azido glycan combinations. This clearly illustrates the potential of dual-labelling techniques with diazo-glycans.

3.7 Nitron-alkyne cycloaddition followed by rearrangement

A new copper-catalysed bioconjugation reaction was discovered by Pezacki and co-workers in 2014, involving coupling a terminal alkyne and nitron to form a β -lactam ring (Scheme 3).¹¹⁹ This reaction is fast and both starting materials and product are stable in aqueous media. The copper-catalysed nitron-alkyne cycloaddition followed by rearrangement (CuANCR) has been successfully performed in living cells, like Huh-7 cells, and *E. coli*. For the labelling in *E. coli* a nitron-bearing KDO derivative (HMMPO) was synthesised. *E. coli* was incubated overnight in M9 minimal medium containing 4 mM KDO with or without an alkyne or nitron chemical handle. KDO-HMMPO and KDO-alkyne were successfully incorporated into the bacterial LPS as shown by SDS-PAGE analysis. The alkyne nitron cycloaddition can also be performed with strained alkynes (SPANC), and is compatible with cyclic or acyclic nitrones.¹²⁰

4. Synthetic strategies for selected members from the MOE toolbox

The MOE toolbox is filled with unnatural glycan derivatives and most of these are not yet commercially available. Therefore most studies that are undertaken using MOE still start with the synthesis of the required probes using a combination of carbohydrate chemistry and organic chemistry. In this section we provide an overview of synthesis strategies used for a diverse selection of glycan-based chemical probes used in MOE. In general, the same synthetic strategies apply for the construction of all three of the much used *N*-acylated mannosamine, galactosamine, and glucosamine-based chemical probes. For that reason we here only discuss several representative examples for the synthesis of the mannosamine-based chemical probes. A comprehensive overview of chemical probes used for MOE can be found in the previous section in table 2.



Scheme 3. Nitron-alkyne cycloaddition.

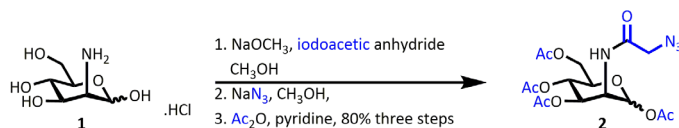
4.1 Azido-labelled glycans

The first azido-glycans to be reported were the mannosamine, galactosamine, and glucosamine-based *N*-azidoacetyl glycans. These per-acetylated glycans are easily prepared via standard carbodiimide-mediated amide bond formation with azido acetic acid and the respective amino sugar (Scheme 4).^{23, 45} Alternatively, iodoacetic anhydride or iodoacetic acid can be coupled to mannosamine, followed by nucleophilic displacement with sodium azide to provide the *N*-azidoacetyl glycan. The final step includes acetylation of all free hydroxyl groups under standard conditions. Aldolases inside mammalian cells are able to convert Ac₄ManNAz **2**, once deacetylated, towards a sialic acid, to yield azido-labelled neuraminic acid (Neu5Az).²³ Neuraminic acid is the most abundant sialic acid, and known to be important for many cellular communication and processes. Additional neuraminic acid-based chemical probes for MOE are known and will be discussed later.

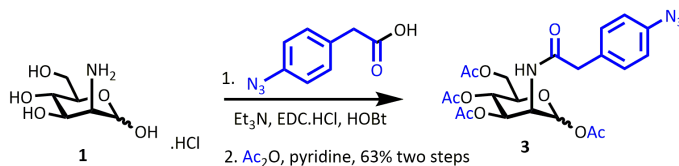
Photoactivatable phenyl azide-based chemical probes have also been made via a similar route. 2-(4-Azidophenyl)acetic acid was coupled to mannosamine hydrochloride **1**, and subsequent acetylation provided the Ac₄Man-2-arylAz **3**, which is a precursor for Sia-5-arylAz (Scheme 5).⁴⁷

Other azido-substituted mannosamine derivatives have been synthesised as well. For example, 4-azido *N*-acetylmannosamine derivative **9** shown in Scheme 6 is made from a 1,6-anhydrosugar intermediate **6** that itself is made in three steps from *N*-acetylmannosamine **4**.¹²¹ A Mitsunobu reaction on the 1,6-anhydro intermediate **6** was performed to give a 3,4-epoxide **7**. Selective axial opening of the epoxide **7** with lithium azide provided, after opening of the 1,6-anhydrosugar **8** a di-amino glycan **9**. This di-amino mannosamine glycan **9** has been used for metabolic labelling in HEK293 cells lacking the key enzyme for sialic acid biosynthesis to improve incorporation rate into glycans.⁴⁸ For membrane permeability glycan **9** was peracetylated before it was used in MOE.

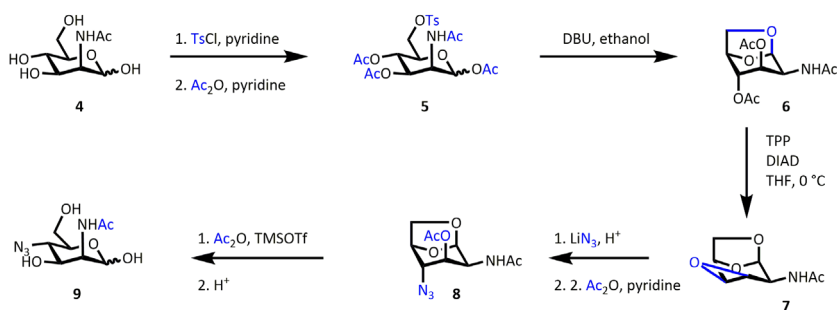
6-azido-*N*-acetylmannosamine **11** is converted inside the cell to 9-azido-*N*-acetylneuraminic acid. The 6-azido glycan **11** can be synthesised from *N*-acetylmannosamine **4** by selectively tosylating the 6-OH **10** (Scheme 7).⁴⁵ Nucleophilic substitution with sodium azide at 50 °C for 3 days gave, after acetylation, the 6-azido Ac₃-*N*-acetylmannosamine **11**.



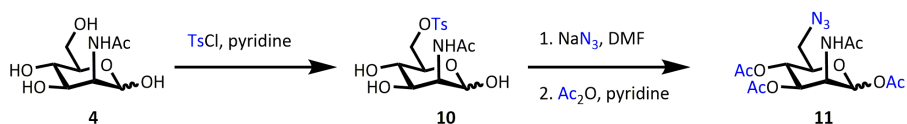
Scheme 4. Synthesis of Ac₄ManNAz **2**.

Scheme 5. Synthesis of $\text{Ac}_4\text{Man-2-arylAz}$ **3**.

3-Deoxy-D-manno-octulosonic Acid (KDO) is a subclass of sialic acids having an eight-carbon skeleton. It is highly abundant and essential component in the outer-cell membrane of lipopolysaccharide Gram negative bacteria and also found in the RG-II pectin of plant cell walls.¹⁰⁷ Human gut bacteria like *H. pylori* are known to have KDO in their LPS, and Dumont and co-workers imaged this via metabolic labelling with KDO probes.⁶⁶ The 8-azido KDO **14** is readily synthesised from D-arabinose **12** (Scheme 8). After selective tosylation of the primary alcohol, the remaining free alcohols are acetylated, and the resulting mixture is then treated with NaN_3 in DMF. In the final step sodium methoxide gives the 5-Az-5-deoxy-D-arabinofuranose **13** in an overall yield of 15%. 8-Azido KDO **14** is made by the reaction between 5-Az-5-deoxy-D-arabinofuranose **13** and sodium oxaloacetate under acidic conditions, followed by a decarboxylation.

Scheme 6. Synthesis of 4-azido ManNAc **9**.

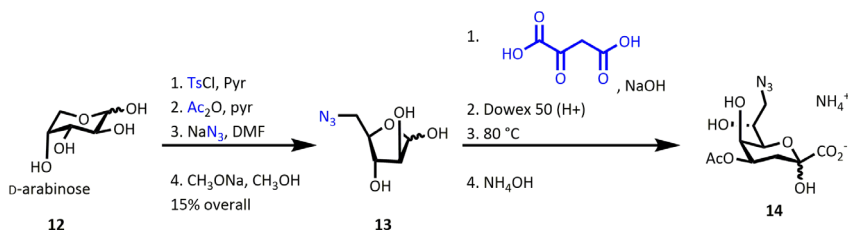
Metabolic oligosaccharide engineering can also be a highly valuable tool to study the action or inaction of glycosyltransferases as recently shown by Bertozzi et al. with the use of UDP-azido glycans as metabolic chain inhibitors.⁷⁶ We here discuss the synthesis of three different UDP azido-glycans and the outlook discusses their application in MOE. UDP-2-XylAz **22** was synthesised from D-xylose **15** as shown in Scheme 9. After protection of the anomeric center **16**, selective protection of the

Scheme 7. Synthesis of 6-azido- Ac_3 -N-acetylglucosamine **11**.

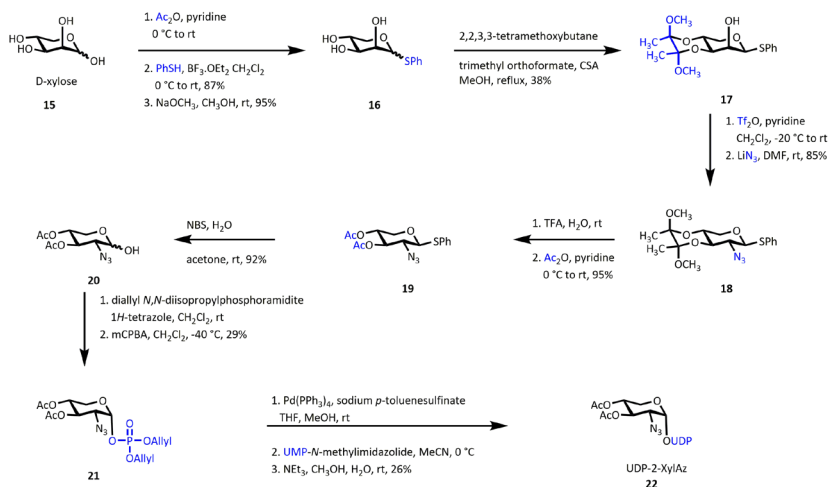
Overview of the Metabolic Oligosaccharide Engineering Toolbox

3- and 4-OH with a cyclohexane-1,2-diacetal 'Ley' protecting group made functionalisation of the 2-position possible **17**.⁷⁶ After activation of the 2-OH with Ti_2O in pyridine, nucleophilic displacement was performed with LiN_3 . The UDP-donor was prepared by first selectively deprotecting the anomeric centre using *N*-bromosuccinimide. The phosphate analogue was made by treatment with diallyl-*N,N*-diisopropylphosphoramidite. UMP-*N*-methylimidazolidine was added, after Pd-catalysed isomerisation and cleavage of the allyl groups, to finally provide UDP-2-XylAz **22**.

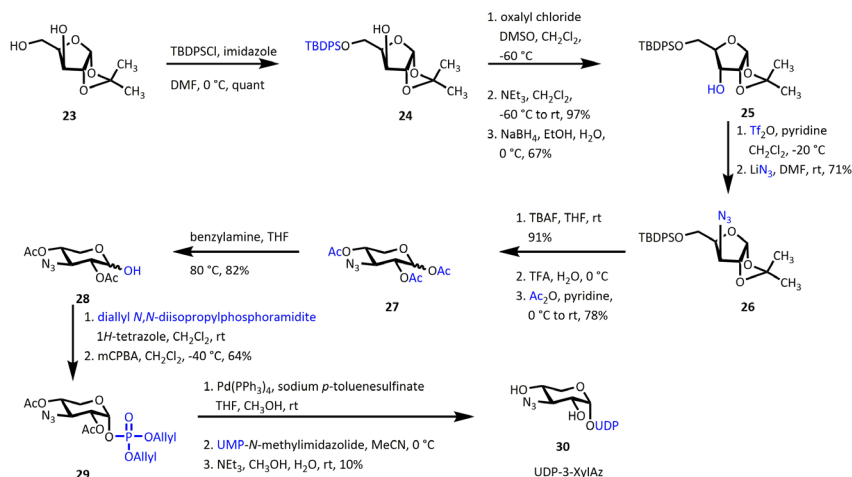
The synthesis of UDP-3-XylAz **30** was performed in 15 steps starting from 1,2-*O*-isopropylidene- α -D-xylofuranose **23** (Scheme 10).⁷⁶ To install the azide functionality on the 3-position, the hydroxyl on this position was first converted from an axial to an equatorial hydroxyl group by oxidation and subsequent stereoselective reduction **25**. Thereafter, conversion to the 3-OTf and its substitution with LiN_3 provided the 3-azido glycan **26**. The UDP group was installed in the same manner as for the 2-XylAz glycan **22**.



Scheme 8. Synthesis of 8-Az-KDO **14**.

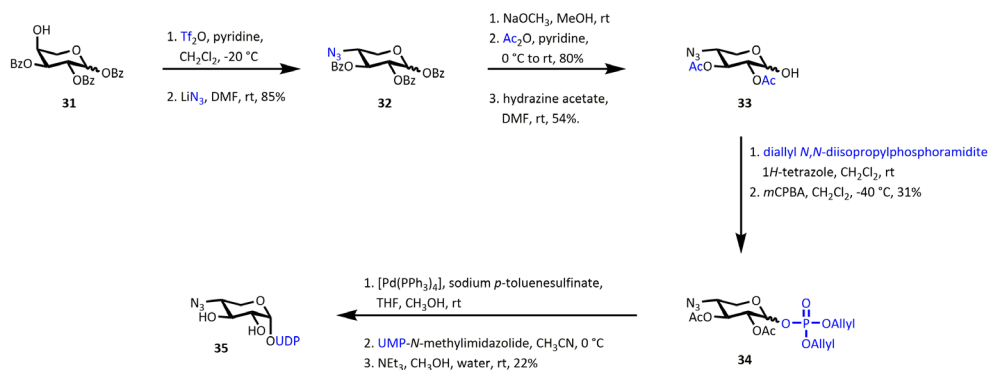


Scheme 9. Synthesis of UDP-2-XylAz **22**.

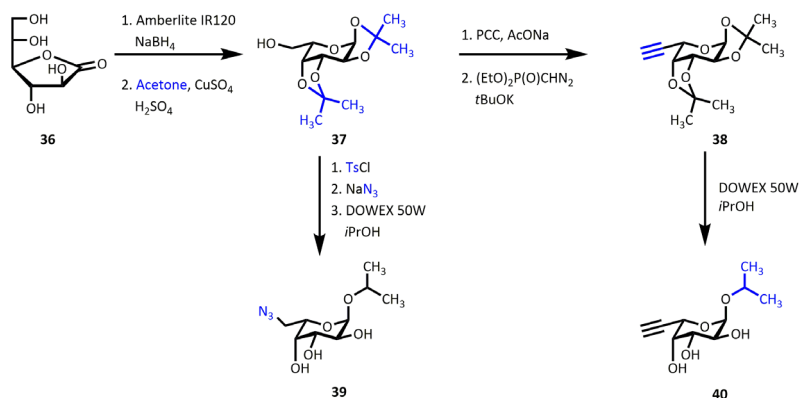


Scheme 10. Synthesis of UDP-3-XylAz 30.

4-OH Benzoylated L-arabinose **31** proved a perfect starting point for the synthesis of UDP-4-XylAz **35** (Scheme 11).^{76, 122} The 4-OH was reacted with triflic anhydride followed by substitution of the triflate with LiN₃ starting product is easily synthesised from L-arabinose. Deprotection of all benzoyl groups followed by acetylation, and selective anomeric deprotection with hydrazine acetate gave acetylated 1-OH 4-XylAz **33**. Installation of the allyl phosphodiester analogue proceeded in moderate to low yield. The UDP group was finally installed by removal of the allyl esters followed by coupling with UMP-*N*-methylimidazole. Purification of synthesised UDP-glycans is a challenging step. The described UDP-glycans were for example purified in a HPLC C-18 column with tributylammonium bicarbonate and methanol as mobile phases. Once the crude product eluted from the column it was further purified by size exclusion chromatography and sodium ion exchange resin.



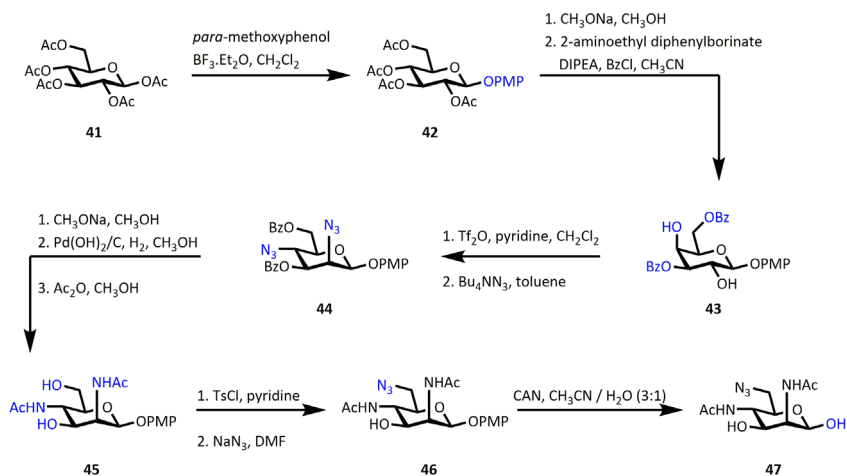
Scheme 11. Synthesis of UDP-4-XylAz 35.



Scheme 12. Synthesis of 6-Az fucose **39** and 5-alkyne fucose **40**.

L-Galactonic acid γ -lactone **36** was selected as a starting point for the synthesis of azido fucose **39** (FucAz) and alkyne fucose **40** (FucAl) analogues (Scheme 12).⁷⁰ A reduction of the lactone with sodium borohydride followed by acetonide protection gave the 6-OH glycan **37**. The 6-OH was activated and nucleophilic displacement using NaN₃ followed by acidic deprotection gave the 6-azido fucose **39** as the isopropyl protected anomeric center. FucAl **40** was made as follows; The 6-OH intermediate **37** was oxidised to the aldehyde with PCC followed by treatment with the Seyferth/Gilbert reagent.¹²³ Acid deprotection provided FucAl **40** in excellent yield. For MOE the isopropyl acetal was cleaved and the hydroxyl groups were acetylated. These azide and alkyne fucose analogues were also used to make their GDP-donor counterparts for use in MOE.

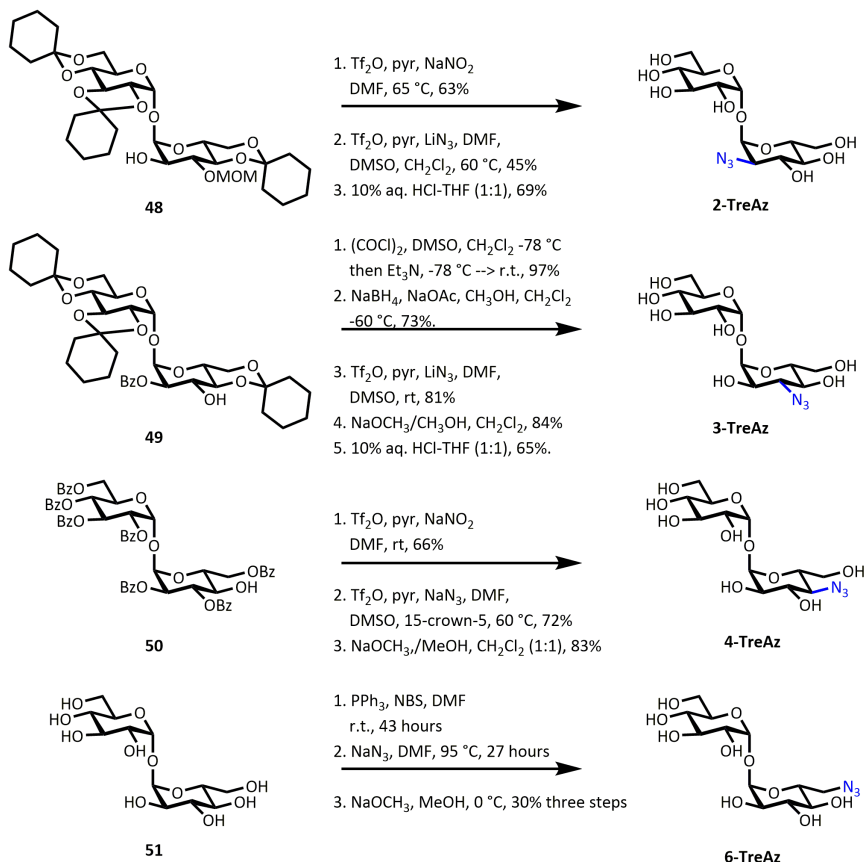
As mentioned, the microbial sialic acid, legionaminic acid, is a relevant target for metabolic labelling to study its glycobiology in human associated pathogenic bacteria. An 6-azido hexose-precursor **47** was synthesised starting from β -D-glucose pentaacetate **41** (Scheme 13).⁷⁷ A glycosylation reaction



Scheme 13. Synthesis of 6-azido Leg Hexose **47**.

with *para*-methoxyphenol followed by Zemplén deacetylation and subsequent selective benzylation gave the 2,4-dihydroxyl glucose derivative **43**. This intermediate was treated with triflic anhydride and reacted with tetrabutylammonium azide to install the two azides **44**. Removal of the benzoyl groups and reduction of the azides followed by controlled acetylation gave the 3,6-dihydroxy intermediate **45**. Selective tosylation of the primary alcohol, followed by nucleophilic displacement with sodium azide and oxidative deprotection of the anomeric PMP group provided the probe **47**.

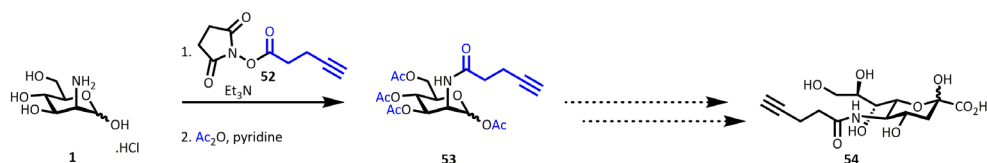
The earlier discussed trehalose azido sugars were synthesised from a series of known differentially protected trehalose building blocks **48-51** (Scheme 14).⁷⁹ Their synthesis involved alcohol activation, followed by nucleophilic substitution with an azide source and complete deprotection to provide azido-trehalose library (2-TreAz, 3-TreAz, 4-TreAz, and 6-TreAz).



Scheme 14. Synthesis of 2-TreAz, 3-TreAz, 4-TreAz, and 6-TreAz.

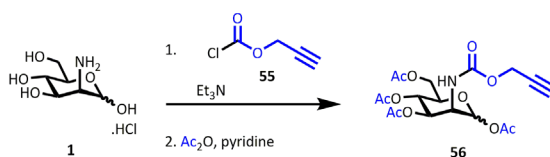
4.2 Alkyne-labelled glycans

Terminal alkynes on carbohydrates were among the first probes used for metabolic oligosaccharide engineering. Their easy synthesis and stability made it possible to synthesise a variety of different alkyne-glycans (see table 1). Alkyne-labelled *N*-acetylmannosamine glycans are made from *N*-mannosamine hydrochloride **1** in a two-step procedure (Scheme 15). Under basic conditions the amine group is reacted with a NHS activated alkyne **52**.⁴⁹ Subsequent acetylation with acetic anhydride and pyridine provided the probe **53**. The alkyne-labelled *N*-acetylmannosamine **53** can either be directly used for metabolic labelling or first converted to alkyne-labelled *N*-acetylneuraminic acid **54** using an enzymatic reaction or the chemical equivalent using a Barbier alkylation and ozonolysis.

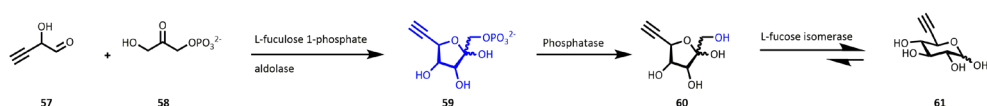


Scheme 15. Synthesis of mannosamine alkyne glycans **53**.

For the synthesis of the carbamate-linked alkyne probes, reported by Pratt et al., mannosamine, galactosamine, or glucosamine hydrochloride were reacted with propargylchloroformate **55** (Scheme 16).⁵³ Subsequent acetylation under standard conditions gave the peracetylated propargyl alkyne glycans **56**.



Scheme 16. Synthesis of Ac₄ManIPoc **56**.

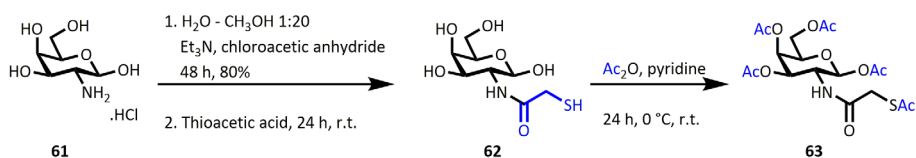


Scheme 17. Synthesis of D-glucose alkynes **60**.

4.3 Thiol-labelled glycans

Thiol-containing *N*-acetyl amino sugars **63** are made from commercially available glucosamine, galactosamine, or mannosamine derivatives. Thioacetic acid is coupled to provide, after full acetylation, the thio acetate probe **63** that can be used for MOE (Scheme 18).^{54, 125}

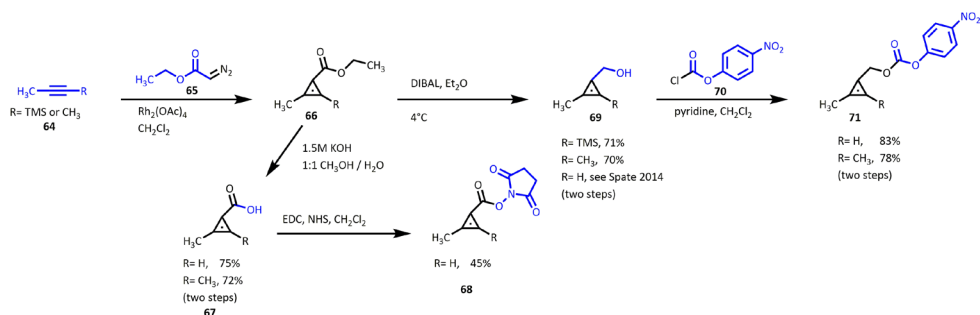
The thioacetyl handle has not been reported yet for the traditional metabolic labelling via click chemistry, but the thiols are biocompatible and have been successfully incorporated onto surface glycans.¹²⁶



Scheme 18. Synthesis of thio-*N*-acetylgalactosamine probe **63**.

4.4 Cyclopropene-labelled glycans

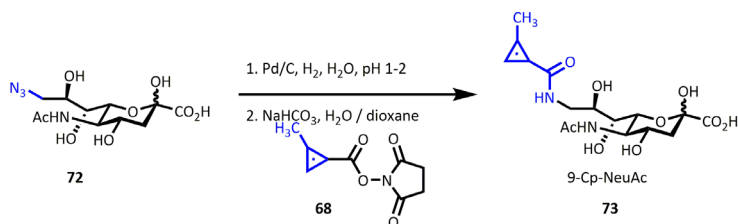
Highly strained cyclopropenes react very fast with tetrazines making them excellent candidates for MOE.⁵⁵ The synthesis of cyclopropenes is not trivial, however, and yields below 50% for the activated cyclopropenes are the standard (Scheme 19). Also, the amide formation to attach them to amino sugars is a low yielding step.⁷¹ The synthesis of 2-methyl cyclopropenes, in which the methyl substituent is important for stability, starts with a rhodium-catalysed cyclopropenation of the appropriate alkyne **64** with ethyl diazoacetate **65**, followed by a reduction of the ethyl ester.⁵⁷ The resulting cyclopropene-alcohol products **69** are known to be volatile and prone to polymerisation. Installation of the cyclopropene handle onto the amino sugar is achieved via the NHS **68** or *p*-nitrophenyl carbonate activated intermediate **71**.



Scheme 19. Synthesis of activated cyclopropenes handles **68** and **71**.

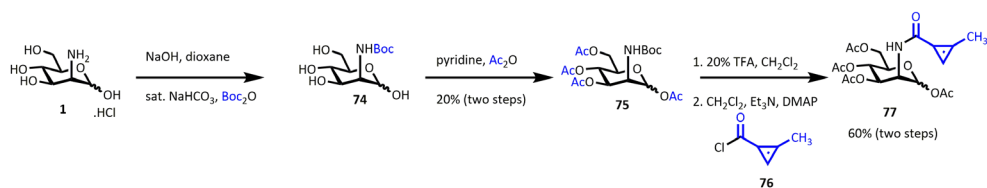
Overview of the Metabolic Oligosaccharide Engineering Toolbox

To improve the stability and reaction yields another route to the activated cyclopropene handles was developed. A TMS protected alkyne is reacted with ethyl diazoacetate under catalytic amounts of rhodium tetraacetate (Scheme 19).⁵⁶ Reduction of the ester to the alcohol with DIBAL was performed. Next, desilylation and subsequent activation was performed in a one-pot reaction using Bu₄NF to deprotect and *p*-nitrophenyl chloroformate to activate the alcohol **69**. This crystalline activated cyclopropene (71; R = CH₃) can easily be stored and used for the installation of the cyclopropene group onto amino sugars. 9-Azido *N*-acetylneuraminic acid **72** is an ideal starting point for the synthesis of 9-cyclopropenyl neuraminic acid **73** (Scheme 20). After Pd-catalysed hydrogenation, NHS-activated cyclopropene **68** is added to give the final product.⁷¹



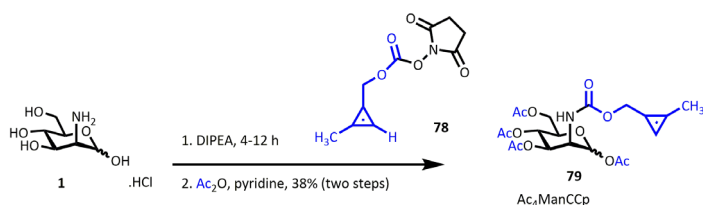
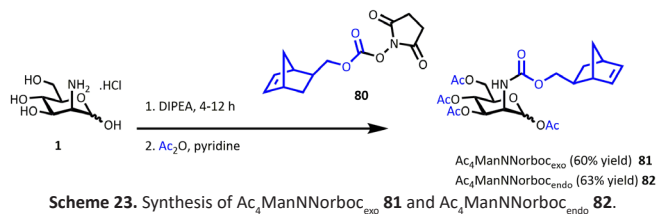
Scheme 20. Synthesis of 9-Cp-Neu5Ac **73**.

The mannosamine cyclopropenes (Ac₄ManNCyc and Ac₄ManCCP) are made from D-mannosamine*HCl **1**. When D-mannosamine*HCl **1** was directly reacted with the acid chloride of a cyclopropene handle **76**, it yielded several products making the synthesis low yielding and the purification difficult (Scheme 21). A three-step procedure with a Boc protection, followed by acetylation, and Boc deprotection gave the acetylated mannosamine **75**. This O-protected amino sugar could then be converted into the cyclopropene probe **77** in good yield (60% over two steps).^{55, 127}



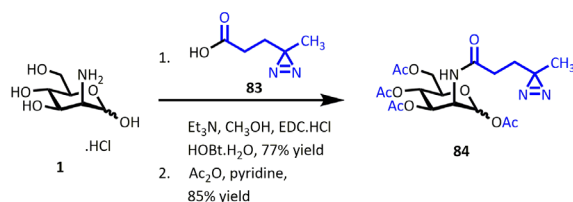
Scheme 21. Synthesis of Ac₄ManNCyc **77**.

The less reactive NHS-ester of a cyclopropene handle **78** was directly coupled to D-mannosamine*HCl **1**, and subsequent acetylation produced Ac₄ManCCP **79** in 38% yield (Scheme 22).^{56, 57} The same route was followed to install norbornene-derived mannosamine probes **81** and **82**. For that a succinimide activated norbornene derivative **80** was synthesised in three steps (Scheme 23).

Scheme 22. Synthesis of Ac_4ManCCp **79**.Scheme 23. Synthesis of $\text{Ac}_4\text{ManNNorboC}_{\text{exo}}$ **81** and $\text{Ac}_4\text{ManNNorboC}_{\text{endo}}$ **82**.

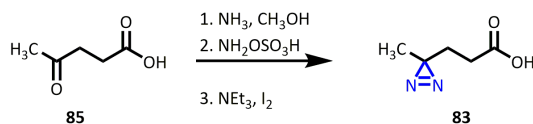
4.5 Diazirine-labelled glycans

Diazirine-based chemical probes **84** are synthesised from mannosamine, glucosamine, or galactosamine via an amide coupling with the diazirine building block **83** under standard amide-coupling techniques (Scheme 24). Similar synthesis with longer linkers and other amino sugars are known.⁶¹

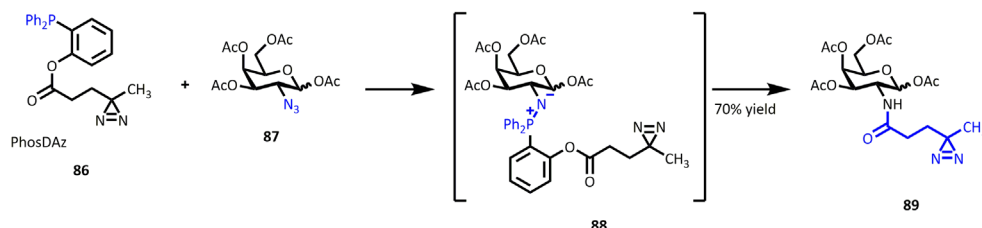
Scheme 24. Synthesis of mannosamine-diazirines derivatives **84**.

The diazirine building block **83** is commercially available or can be made from levulinic acid **85** (Scheme 25). Levulinic acid **85** is converted into the diazirine handle **83** in a two-step reaction with hydroxylamine-*O*-sulfonic acid and ammonia, followed by triethylamine and iodine in methanol (Scheme 25).³⁶ With this building block in hand, Kohler and co-workers made the $\text{Ac}_4\text{ManNDaz}(3\text{Me})$ **84** and $\text{Ac}_4\text{GlcNDaz}(3\text{Me})$.⁶¹

Jewett and co-workers recently reported an elegant alternative way to install diazirines on azide-containing carbohydrate derivatives using a traceless Staudinger reaction.⁶² They successfully reduced a C-2 azide, for example glucosamine and galactosamine **87**, with the phosphine reagent PhosDAz **86**, trapping the formed aza-ylide by an intramolecular reaction **88** with the ester-appended diazirine **89** (Scheme 26).



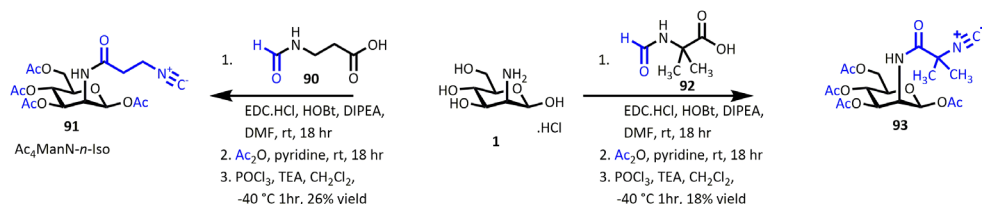
Scheme 25. Synthesis of diazirine building block **83**.



Scheme 26. Synthesis of galactosamine-diazirine **89** via Staudinger reduction.

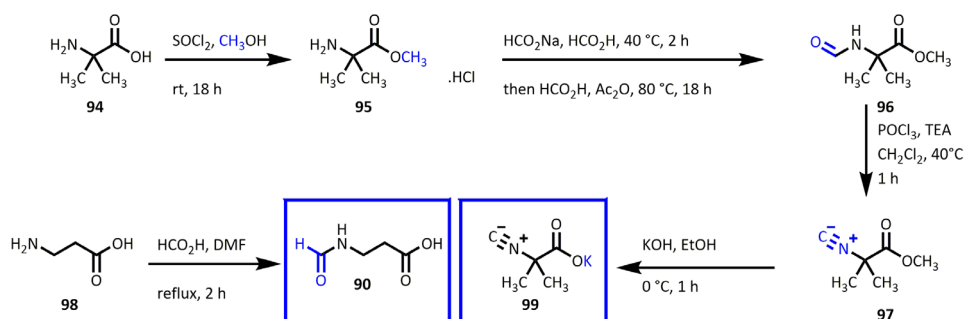
4.6 Isonitrile-labelled glycans

In 2013 Leeper and co-workers developed isonitrile-labelled probes for MOE (Scheme 27).⁶³ The amide coupling of 2-deoxy 2-amino glycans (e.g. D-mannosamine*HCl **1**) with a β -alanine formamide **90** under standard conditions gives the formamide-functionalised carbohydrate. Subsequent acetylation followed by formamide dehydration with POCl_3 gives the isonitrile-containing probe **91**. Isonitriles are stable under neutral conditions and can be purified by column chromatography or HPLC. The amide couplings can also be performed with the isonitrile β -alanine carboxylic acid in an one step procedure. Amide couplings with formamide or isonitriles are challenging and generally result in low yields (<30%). Dehydration of formamides can be performed via different procedures. The most common procedure uses POCl_3 , but this rather harsh condition can lead to decomposition of the product. Other methods for dehydration exist, for example, Burgess reagent is able to perform the same dehydration step under much milder conditions.¹²⁸



Scheme 27. Synthesis of primary **91** and tertiary isonitrile **93** mannosamine glycans.

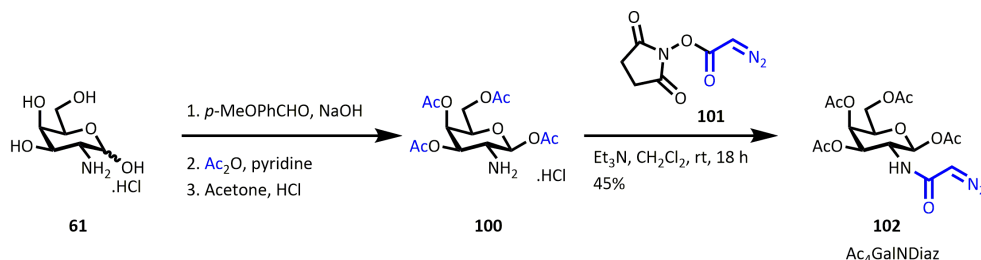
Tertiary isonitriles can also be synthesised in a direct coupling between the amino sugar and the isonitrile using potassium 2-isocyano-2-methylpropanoate **99** (Scheme 28). 2-Amino-2-methylpropanoic methyl ester **95**, made from the corresponding acid **94**, is exposed to *in situ* formed acetic formic anhydride **96** and dehydration of the formamide intermediate with POCl₃ gives the methyl-protected tertiary isonitrile **97**. Saponification of the methyl ester provides the isonitrile **99** ready for coupling to an amino sugar.



Scheme 28. Synthesis of formamide building block **90** and tertiary isonitrile building block **99**.

4.7 Diazo-labelled glycans

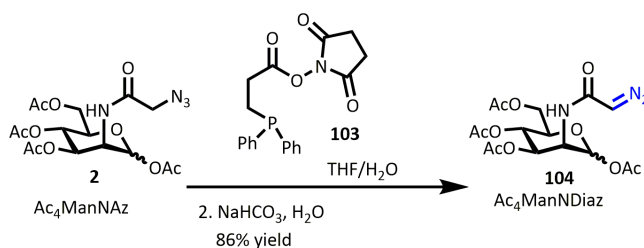
The diazo derivative of *N*-acetylmannosamine or *N*-acetylgalactosamine **61**, ManDiaz or GalDiaz, respectively, are easily synthesised via the route published by the Leeper group (Scheme 29).³⁹ Starting from galactosamine **61**, Ac₄GalDiaz **102** can be synthesised anomerically pure in a three-step process involving transient protection of the amine as its imine with anisaldehyde, acetylation of the hydroxyls, imine hydrolysis, followed by installing the diazoacetyl group via its NHS ester derivative



Scheme 29. Synthesis of Ac₄GalNDiaz **102** by Leeper et al.³⁹

Overview of the Metabolic Oligosaccharide Engineering Toolbox

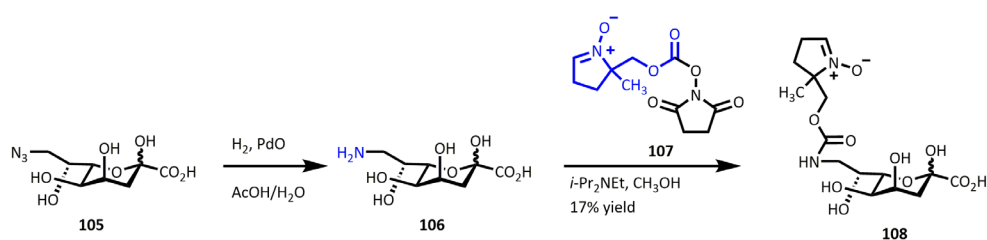
Another method to synthesise Ac₄ManDiaz **104** was reported by the Raines group (Scheme 30).⁵² A deimidogenation of the parent azido-glycan, via an acyl triazenophosphonium salt intermediate, gave the diazo-glycan in high yields.¹²⁹ Both diazo-glycans proved to be bio-orthogonal as incubation at pH 7.4 at 37 °C showed no reaction of diazo acetyl glycans with biologically relevant functional groups such as thiols, carboxylic acids, and amines. Diazo-acetyl glycans are stable compounds and did not show any degradation during synthesis or purification.



Scheme 30. Synthesis of Ac₄ManNDiaz **104** by Raines et al.⁵³

4.8 Nitrone-labelled glycans

Very recently, the use of nitrone-labelled glycans in MOE was reported for the first time (Scheme 31).¹¹⁹ Nitrone-labelled KDO **108** was made from the earlier discussed 8-azido KDO precursor **105**. The azido glycan was reduced using PdO, H₂ under slightly acidic conditions followed by amide coupling with an activated nitrone-succinimide **107**. The resulting product was purified by preparative HPLC before it was used for metabolic labelling. Metabolic labelling with these nitrone glycans were performed in *E. coli* and Huh-7 cells. The CuANCR reaction was also successfully applied to synthesise unnatural amino acids which were incorporated into *E. coli*, *L. innocua*, and *L. lactis*.¹²²



Scheme 31. Synthesis of nitrone-KDO **108** by Pezacki et al.¹¹⁹

5. Outlook

In this final section some exciting recent advances are discussed that might impact MOE in the near future, as well as several MOE-related developments that might advance the field of chemical glycobiology in general.

5.1 Selective exo-enzymatic labelling

A promising complementary method to MOE for the incorporation of labelled glycans is Selective Exo-Enzymatic Labelling (SEEL), introduced in 2013 by Boons and co-workers.¹³⁰ In this technique, intact cells are exposed to extracellularly added recombinant glycosyltransferases, for example sialyltransferases, and incubated together with labelled donor derivatives, for instance C-5 or C-9 azido-modified CMP-sialic acids. The recombinant sialyltransferase enzymes still recognise and transfer sialic acid donors with an azide introduced on these positions. As a result, an azide-labelled sialic acid was introduced exclusively on the cell surface glycans that displayed the appropriate terminal acceptor glycans. SEEL thus has several advantages over MOE, for example the possibility of selectively labelling glycans on the cell surface and the benefit of not interfering with the intracellular metabolic pathways. The main disadvantage is the need to synthesise the valuable labelled-glycan donors and prepare and/or add the required glycosyltransferases.

5.2 Metabolically generated glycosyltransferases inhibitors

Glycosyltransferases have, with the exception of glucosylceramide synthase that can be reliably targeted by lipophilic iminosugars (like Zavesca and AMP-DNM), remained elusive towards a general small molecule design principle for their inhibition. We clearly lack structural insight in glycosyltransferases when compared to their metabolic counterpart, glycosidases, and hence hardly any effective, bioavailable glycosyltransferase inhibitor currently exists. Many different approaches to glycosyltransferases inhibitors are thus known and discussed in a recent review by Gloster and Vocadlo.¹³¹ However, it has recently been shown that the principles that drive MOE can also be used to perturb the activity of glycosyltransferases. The two MOE-related strategies that make use of the cell's glycan salvage pathways are trojan horse donor inhibitors and metabolic chain inhibitors.

The Paulson group developed a new technique to tune intracellular glycosyltransferase activity through the metabolic generation of transferase inhibitors from unnatural glycans derivatives – the trojan horse – by hijacking the salvage pathway (Figure 4).^{115, 131-137} Fluorinated analogues of fucose and sialic acid were shown to be taken up and metabolised by cells into fluorinated donors that still tightly bind their corresponding glycosyltransferase, but are no longer transferred to an acceptor (Figure 5).¹³⁸ This is caused by the fact that the required active site oxocarbenium ion transition state is destabilised by the neighbouring fluorine atom and is thus no longer formed. The hijacking of the salvage pathway via this trojan horse strategy also causes a reduced production of the natural

Overview of the Metabolic Oligosaccharide Engineering Toolbox

nucleotide sugar donors. The Vocadlo group published a similar strategy that instead uses glycan derivative in which the endocyclic oxygen is replaced by a thio-group.¹³⁸ These are similarly metabolised into glycosyltransferase substrates that still bind the active site, but are no longer transferred.

An elegant example of metabolic chain inhibitors is the work on azido-xylose analogues by Bertozzi and co-workers.⁷⁶ The azido-xylose glycans were used as chain-terminating metabolic inhibitors of glycosaminoglycans (GAG) synthesis in zebrafish. The lack of a salvage pathway that could convert azido-xylose glycans to UDP-activated forms gave an extra synthetic challenge. The synthesis of the 2-, 3- and 4-azido UDP-xylose was discussed in section 4.1. The UDP-azido-xylose analogues were added to the zebrafish embryos via microinjection for further study. After development, a SPAAC reaction with difluorocyclooctyne-AlexaFluor 488 was performed on the embryos. Successful labelling was only observed with UDP-4-XylAz as confirmed by confocal microscopy and flow cytometry. The mechanism of action, hence the reason why only UDP-4-XylAz showed successful labelling, is not known yet.

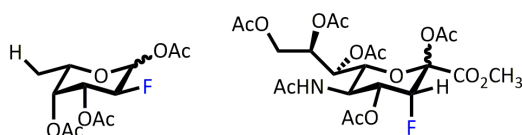


Figure 4. Two glycosyltransferases inhibitors developed by Paulson et al.¹³⁹

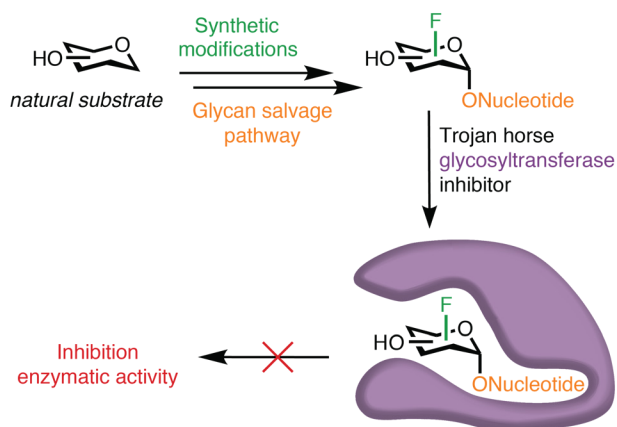


Figure 5. Overview of principle behind MOE-generated trojan horse glycosyltransferase inhibitors.

5.3 Dual labelling

Dual labelling is an area of MOE that is gaining increasing attention. Two different types of dual-labelling strategies are currently being investigated. The first strategy uses two metabolically incorporated glycan-based chemical probes with different bio-orthogonal reactive groups, while the second strategy uses two bio-orthogonal reactive groups within the same molecule. A recent example of the first approach is the successful simultaneous incorporation of Ac₄GalNAz and Ac₄GlcN-*n*-Iso and the simultaneously labelling of both these glycans in a single step using SPAAC and invDA. Other reported examples used nitrone- and azido-labelled glycans for bioconjugation via SPANC and SPAAC, and norbornene or cyclopropene-labelled glycans in combination with azido-labelled glycans for dual invDA and SPAAC labelling in HEK 293T cells.^{53, 114, 139, 140} An example of the second strategy is the successfully metabolic incorporation of a sialic acid derivative containing both an azide and a photocleavable diazirine group.⁷⁵

5.4 What's next for MOE?

Besides the already impressive toolbox of probes and associated bioconjugation methods that are currently known for use in MOE, several new bioconjugation reactions have been reported recently that might also end up as part of MOE toolbox. Four highly interesting cases are: a) the reaction of *N*-oxides with boron reagents, as developed by Bertozzi; b) the reaction between cyclopropenones and functionalised phosphines; c) 2H-azirines that form after ring opening nitrile ylides, and react spontaneously with alkene dipolarophiles to produce stable $\delta 1$ -pyrrolines; and d) the reaction between strained cycloalkynes and sulfenic acids.^{68, 141-144} On the other hand, the 'old' established oxime ligation reaction,³³ between an aldehyde and oxy-amine, is also still a powerful ally for bioconjugation reactions in MOE, especially in view of the recent finding that oximes are formed faster in water at neutral pH at -20 °C.¹⁴⁵ One of the fastest known click reactions at the moment is the Strain-Promoted Oxidation-Controlled Cyclooctyne-1,2-Quinone Cycloaddition (SPOCQ) reaction, and its speed might offer chances with respect to competing biological processes and allow specific labelling.⁴⁰ The compatibility of these reactions in polar media suggest future possibilities for their use in metabolic oligosaccharide engineering.

The CuAAC and SPAAC are still by far the most popular click reaction currently used for MOE. Many studies have focussed on the improvement of the reactivity of the agents by changing the characters of the alkyne moiety. Such changes involve the ring strain in the cyclic molecule (SPAAC reagents), the lipophilicity of SPAAC reagents (introduction of polar groups), CuAAC-accelerating ligands, and the use of iodo-alkynes.²⁴ Interestingly, changing the electronic character of the azide was not investigated until recently. One method uses sulfonyl azides as reactive precursor for a variety of applications.²⁴ In another method, van Delft and co-workers have shown the potential of difluoro azido glycans to improve the speed of SPAAC. They synthesised a difluoro azido acetic acid precursor, which can be installed via peptide chemistry to glucosamine, galactosamine, or mannosamine glycans.¹⁴⁶ This reactive probe has not been used for MOE yet.

Overview of the Metabolic Oligosaccharide Engineering Toolbox

Most of the methods described for the labelling of glycans can also be applied to other biomolecules such as lipids and proteins. A complete overview of the possibilities and achievements in this field is highlighted in other reviews.^{40, 147, 148}

In the past 15 years many different mammalian cell lines have been successfully used in metabolic oligosaccharide engineering studies. Furthermore, different organisms like zebrafish, *C. elegans*, and bacteria have also been studied. A promising new target for the study of glycans with MOE are plants. At this point, only a few studies have been published that report the use of MOE in plants.^{65, 78, 107, 108, 149, 156} These studies alone have already provided valuable insights in the spatiotemporal dynamics of plant glycans, but many questions in plant glycobiology are still unanswered.

Metabolic oligosaccharide engineering uses the glycan salvage pathway of cells. As a consequence labelled glycans end up in all biomolecules that contain these glycans, e.g. *N*- and *O*-glycans on proteins and glycolipids.¹⁵⁰ This can be a drawback, as it is difficult to distinguish exactly where the incorporated glycans end up and what the ratio of incorporation is across the various glycoconjugates. To further advance the application of MOE and the information that can be obtained from its use, methods should be developed that either allow for more selective MOE or accurate identification and quantification of the various labelled glycan populations. One method that show promise in this area is the combination of fluorescence resonance energy transfer (FRET) and MOE which allows for intracellular detection of glycosylation of a protein of interest.¹⁵¹⁻¹⁵⁵

6. Acknowledgements

Space limitations necessitated severe restrictions in the selection of MOE literature to review, so we apologise to researchers whose work was not cited here. We acknowledge the reviewers for their usefull comments on an earlier version of this manuscript. The authors thank the Netherlands Foundation for Scientific Research (NWO) for funding via a VENI (722.011.006) and VIDI (723.014.005) grant to TW.

7. References

1. T.W. Rademacher, R.B. Parekh, R.A. Dwek, *Annu. Rev. Biochem.*, **57** **1988** 785-838.
2. C.R. Bertozzi, L.L. Kiessling, *Science*, **291** **2001** 2357-2364.
3. L.L. Kiessling, R.A. Splain, *Annu. Rev. Biochem.*, **79** **2010** 619-653.
4. L.X. Wang, B.G. Davis, *Chem. Sci.*, **4** **2013** 3381-3394.
5. H. Kayser, C. Geilen, C. Paul, R. Zeitler, W. Reutter, *FEBS*, **301** **1992** 137-140.
6. H. Kayser, R. Zeitler, C. Kannicht, D. Grunow, R. Nuck, W. Reutter, *J. Biol. Chem.*, **267**, **1992** 16934-16938.
7. J.C. Jewett, C.R. Bertozzi, *Chem. Soc. Rev.*, **39** **2010** 1272-1279.
8. S.H. Yu, M.R. Bond, C.M. Whitman, J.J. Kohler, *Methods Enzymol.*, **478** **2010** 541-562.
9. C.Y. Wu, C.H. Wong, *Chem. commun.*, **47** **2011** 6201-6207.
10. S. Gautam, T.J. Gniadek, T. Kim, D.A. Spiegel, *Trends Biotechnol.*, **31** **2013** 258-267.
11. J.E. Hudak, C.R. Bertozzi, *Chem. & Biol.*, **21** **2014** 16-37.
12. M. King, A. Wagner, *Bioconjug. Chem.*, **25** **2014** 825-839.
13. D.M. Patterson, L.A. Nazarova, J.A. Prescher, *ACS Chem. Biol.*, **9** **2014** 592-605.
14. V.N. Tra, D.H. Dube, *Chem. Commun.*, **50** **2014** 4659-4673.
15. K.N. Chuh, M.R. Pratt, *Glycoconj. J.*, **2015**.
16. C.T. Saeui, E. Urias, L. Liu, M.P. Mathew, K.J. Yarema, *Glycoconj. J.*, **32** **2015** 425-441.
17. B. Cheng, R. Xie, L. Dong, X. Chen, *ChemBioChem.*, **17** **2016** 11-27.
18. S.H. Rouhanifard, L.U. Nordstrom, T. Zheng, P. Wu, *Chem. Soc. Rev.*, **42** **2013** 4284-4296.
19. S.T. Laughlin, C.R. Bertozzi, *Nat. Protoc.*, **2** **2007** 2930-2944.
20. C.I. Schilling, N. Jung, M. Biskup, U. Schepers, S. Brase, *Chem. Soc. Rev.*, **40** **2011** 4840-4871.
21. N. Nischan, J.J. Kohler, *Glycobiology*, **26** **2016** 789-796.
22. L.K. Mahal, K.J. Yarema, C.R. Bertozzi, *Science*, **276** **1997** 1125-1128.
23. E. Saxon, C.R. Bertozzi, *Science*, **287** **2000** 2007-2010.
24. J.E. Hein, V.V. Fokin, *Chem. Soc. Rev.*, **39** **2010** 1302-1315.
25. N.J. Agard, J.A. Prescher, C.R. Bertozzi, *J. Am. Chem. Soc.*, **126** **2004** 15046-15047.
26. R. van Geel, G.J. Puijn, F.L. van Delft, W.C. Boelens, *Bioconjug. Chem.*, **23** **2012** 392-398.
27. A.A. Poloukhine, N.E. Mbua, M.A. Wolfert, G.J. Boons, V.V. Popik, *J. Am. Chem. Soc.*, **131** **2009** 15769-15776.
28. L. Shah, S.T. Laughlin, I.S. Carrico, *J. Am. Chem. Soc.*, **138** **2016** 5186-5189.
29. Z. Li, L. Qian, L. Li, J.C. Bernhammer, H.V. Huynh, J.S. Lee, S.Q. Yao, *Angew. Chem.*, **55** **2016** 2002-2006.
30. K. Lang, J.W. Chin, *ACS Chem Bio.*, **9** **2014** 16-20.
31. V.V. Rostovtsev, L.G. Green, V.V. Fokin, K.B. Sharpless, *Angew. Chem.*, **41** **2002** 2596-2599.
32. C.W. Tornoe, C. Christensen, M. Meldal, *J. Org. Chem.*, **67** **2002** 3057-3064.
33. S. Ulrich, D. Boturyn, A. Marra, O. Renaudet, P. Dumy, *Chemistry*, **20** **2014** 34-41.
34. M.L. Blackman, M. Royzen, J.M. Fox, *J. Am. Chem. Soc.*, **130** **2008** 13518-13519.
35. Z. Yu, Y. Pan, Z. Wang, J. Wang, Q. Lin, *Angew. Chem.*, **51** **2012** 10600-10604.
36. Y. Tanaka, J.J. Kohler, *J. Am. Chem. Soc.*, **130** **2008** 3278-3279.
37. H. Stockmann, A.A. Neves, S. Stairs, K.M. Brindle, F.J. Leeper, *Org. Biomol Chem.*, **9** **2011**, 7303-7305.
38. Q. Li, T. Dong, X. Liu, X. Lei, *J. Am. Chem. Soc.*, **135** **2013** 4996-4999.
39. L. Josa-Cullere, Y.A. Wainman, K.M. Brindle, F.J. Leeper, *RSC Adv.*, **4** **2014** 52241-52244.

Overview of the Metabolic Oligosaccharide Engineering Toolbox

40. A. Borrmann, O. Fatunsin, J. Dommerholt, A.M. Jonker, D.W. Lowik, J.C. van Hest, F.L. van Delft, *Bioconjug. Chem.*, **26** **2015** 257-261.
41. C.S. McKay, J. Moran, J.P. Pezacki, *Chem. Comm.*, **46** **2010** 931-933.
42. N.D. Pham, R.B. Parker, J.J. Kohler, *Curr. Opin. Chem. Biol.*, **17** **2013** 90-101.
43. N.K. Devaraj, R. Weissleder, S.A. Hilderbrand, *Bioconjug. Chem.*, **19** **2008** 2297-2299.
44. H. Zhang, W.S. Trout, S. Liu, G.A. Andrade, D.A. Hudson, S.L. Scinto, K.T. Dicker, Y. Li, N. Lazouski, J. Rosenthal, C. Thorpe, X. Jia, J.M. Fox, *J. Am. Chem. Soc.*, **138** **2016**, 5978-5983.
45. E. Saxon, S.J. Luchansky, H.C. Hang, C. Yu, S.C. Lee, C.R. Bertozzi, *J. Am. Chem. Soc.*, **124** **2002** 14893-14902.
46. Y. Tian, R.T. Almaraz, C.H. Choi, Q.K. Li, C. Saeui, D. Li, P. Shah, R. Bhattacharya, K.J. Yarema, H. Zhang, *Clin. Proteom.*, **12** **2015** 11.
47. S.J. Luchansky, S. Goon, C.R. Bertozzi, *ChemBioChem.*, **5** **2004** 371-374.
48. H. Moller, V. Bohrsch, J. Bentrop, J. Bender, S. Hinderlich, C.P. Hackenberger, *Angew. Chem.*, **51** **2012** 5986-5990.
49. T.L. Hsu, S.R. Hanson, K. Kishikawa, S.K. Wang, M. Sawa, C.H. Wong, *Proc. Nat. Acad. Sci. U.S.A.*, **104** **2007** 2614-2619.
50. S.R. Hanson, T.L. Hsu, E. Weerapana, K. Kishikawa, G.M. Simon, B.F. Cravatt, C.H. Wong, *J. Am. Chem. Soc.* **129** **2007** 7266-7267.
51. B.W. Zaro, K.N. Chuh, M.R. Pratt, *ACS Chem. Bio.*, **9** **2014** 1991-1996.
52. K.A. Andersen, M.R. Aronoff, N.A. McGrath, R.T. Raines, *J. Am. Chem. Soc.*, **137** **2015** 2412-2415.
53. L.A. Bateman, B.W. Zaro, K.N. Chuh, M.R. Pratt, *Chem. Comm.*, **49** **2013** 4328-4330.
54. J. Du, P.L. Che, U. Aich, E. Tan, H.J. Kim, S.G. Sampathkumar, K.J. Yarema, *Bioorg. Med. Chem. Lett.*, **21** **2011** 4980-4984.
55. C.M. Cole, J. Yang, J. Seckute, N.K. Devaraj, *ChemBioChem.*, **14** **2013** 205-208.
56. A.K. Spate, H. Busskamp, A. Niederwieser, V.F. Schart, A. Marx, V. Wittmann, *Bioconjug. Chem.*, **25** **2014** 147-154.
57. D.M. Patterson, K.A. Jones, J.A. Prescher, *Mol. BioSys.*, **10** **2014** 1693-1697.
58. A. Niederwieser, A.K. Spate, L.D. Nguyen, C. Jungst, W. Reutter, V. Wittmann, *Angew. Chem.*, **52** **2013** 4265-4268.
59. A.K. Spate, V.F. Schart, S. Schollkopf, A. Niederwieser, V. Wittmann, *Chem. Eur. J.*, **20** **2014** 16502-16508.
60. H.S. Beckmann, A. Niederwieser, M. Wiessler, V. Wittmann, *Chemistry*, **18** **2012** 6548-6554.
61. M.R. Bond, H.C. Zhang, J. Kim, S.H. Yu, F. Yang, S.M. Patrie, J.J. Kohler, *Bioconjug. Chem.*, **22** **2011** 1811-1823.
62. A.M. Ahad, S.M. Jensen, J.C. Jewett, *Org. Lett.*, **15** **2013** 5060-5063.
63. S. Stairs, A.A. Neves, H. Stockmann, Y.A. Wainman, H. Ireland-Zecchini, K.M. Brindle, F.J. Leeper, *ChemBioChem.*, **14** **2013** 1063-1067.
64. A. Homann, R.U. Qamar, S. Serim, P. Dersch, J. Seibel, *Beilstein J. Org. Chem.*, **6** **2010** 24.
65. C.T. Anderson, I.S. Wallace, C.R. Somerville, *Proc. Nat. Acad. Sci. U. S. A.*, **109** **2012** 1329- 1334.
66. A. Dumont, A. Malleron, M. Awwad, S. Dukan, B. Vauzeilles, *Angew. Chem.*, **51** **2012** 3143-3146.
67. J. Wang, B. Cheng, J. Li, Z. Zhang, W. Hong, X. Chen, P.R. Chen, *Angew. Chem.*, **54** **2015** 5364-5368.
68. A.R. Sherratt, M. Chigrinova, C.S. McKay, L.P.B. Beaulieu, Y. Rouleau, J.P. Pezacki, *RSC Adv.*, **4** **2014** 46966-46969.
69. J. Vanbeselaere, D. Vicogne, G. Matthijs, C. Biot, F. Foulquier, Y. Guerardel, *Chem. Comm.*, **49** **2013** 11293-11295.
70. M. Sawa, T.L. Hsu, T. Itoh, M. Sugiyama, S.R. Hanson, P.K. Vogt, C.H. Wong, *Proc. Nat. Acad. Sci. U. S. A.*, **103** **2006** 12371-12376.
71. D.M. Patterson, L.A. Nazarova, B. Xie, D.N. Kamber, J.A. Prescher, *J. Am. Chem. Soc.*, **134** **2012** 18638-18643.

72. K.W. Dehnert, B.J. Beahm, T.T. Huynh, J.M. Baskin, S.T. Laughlin, W. Wang, P. Wu, S.L. Amacher, C.R. Bertozzi, *ACS Chem. Bio.*, **6** **2011** 547-552.
73. Y. Zeng, T.N. Ramya, A. Dirksen, P.E. Dawson, J.C. Paulson, *Nat. Methods.*, **6** **2009**, 207-209.
74. A. Pulsipher, M.E. Griffin, S.E. Stone, J.M. Brown, L.C. Hsieh-Wilson, *J. Am. Chem. Soc.*, **136**, **2014** 6794-6797.
75. L. Feng, S. Hong, J. Rong, Q. You, P. Dai, R. Huang, Y. Tan, W. Hong, C. Xie, J. Zhao, X. Chen, *J. Am. Chem. Soc.*, **135** **2013** 9244-9247.
76. B.J. Beahm, K.W. Dehnert, N.L. Derr, J. Kuhn, J.K. Eberhart, D. Spillmann, S.L. Amacher, C.R. Bertozzi, *Angew. Chem.*, **53** **2014** 3347-3352.
77. J. Mas Pons, A. Dumont, G. Sautejeau, E. Fugier, A. Baron, S. Dukan, B. Vauzeilles, *Angew. Chem.*, **53** **2014** 1275-1278.
78. D.D. McClosky, B. Wang, G. Chen, C.T. Anderson, *Phytochemistry*, **2016**.
79. B.M. Swarts, C.M. Holsclaw, J.C. Jewett, M. Alber, D.M. Fox, M.S. Siegrist, J.A. Leary, R. Kalscheuer, C.R. Bertozzi, *J. Am. Chem. Soc.*, **134** **2012** 16123-16126.
80. M.M. Rochefort, M.D. Girgis, J.S. Ankeny, J.S. Tomlinson, *Glycobiology*, **24** (**2014**) 62-69.
81. J. Rong, J. Han, L. Dong, Y. Tan, H. Yang, L. Feng, Q.W. Wang, R. Meng, J. Zhao, S.Q. Wang, X. Chen, *J. Am. Chem. Soc.*, **136** **2014** 17468-17476.
82. K. Kang, S. Joo, J.Y. Choi, S. Geum, S.P. Hong, S.Y. Lee, Y.H. Kim, S.M. Kim, M.H. Yoon, Y. Nam, K.B. Lee, H.Y. Lee, I.S. Choi, *Proc. Nat. Acad. Sci. U. S. A.*, **112** **2015** E241-248.
83. E.M. Sletten, C.R. Bertozzi, *Acc. Chem. Res.*, **44** **2011** 666-676.
84. B. Belardi, C.R. Bertozzi, *Chem Biol*, **22** **2015** 983-993.
85. J. Du, M.A. Meledeo, Z. Wang, H.S. Khanna, V.D. Parachuri, K.J. Yarema, *Glycobiology*, **19** **2009**, 1382-1401
86. Z. Wang, J. Du, P.L. Che, M.A. Meledeo, K.J. Yarema, *Curr. Opin. Chem. Biol.*, **13** **2009** 565-572.
87. S.T. Laughlin, C.R. Bertozzi, *ACS Chem. Biol.*, **4** **2009** 1068-1072.
88. S.T. Laughlin, *Thesis*, **2008**.
89. N. Geva-Zatorsky, D. Alvarez, J.E. Hudak, N.C. Reading, D. Erturk-Hasdemir, S. Dasgupta, U.H. von Andrian, D.L. Kasper, *Nat. Med.*, **21** **2015** 1091-+.
90. A.R. Burnham-Marusch, C.J. Snodgrass, A.M. Johnson, C.M. Kiyoshi, S.E. Buzby, M.R. Gruner, P.M. Berninsone, *PLoS One*, **7** **2012** e49020.
91. S.T. Laughlin, J.M. Baskin, S.L. Amacher, C.R. Bertozzi, *Science*, **320** **2008** 664-667.
92. J.M. Baskin, K.W. Dehnert, S.T. Laughlin, S.L. Amacher, C.R. Bertozzi, *Proc. Nat. Acad. Sci. U. S. A.*, **107** **2010** 10360-10365.
93. D. Rabuka, S.C. Hubbard, S.T. Laughlin, S.P. Argade, C.R. Bertozzi, *J. Am. Chem. Soc.*, **128** **2006** 12078-12079.
94. M. Zunk, M.J. Kiefel, *RSC Adv.*, **4** **2014** 3413-3421.
95. S. Goon, B. Schilling, M.V. Tullius, B.W. Gibson, C.R. Bertozzi, *Proc. Nat. Acad. Sci. U. S. A.*, **100** **2003** 3089-3094.
96. F. Liu, A.J. Aubry, I.C. Schoenhofen, S.M. Logan, M.E. Tanner, *ChemBioChem*, **10** **2009**, 1317-1320.
97. C. Besanceney-Webler, H. Jiang, W. Wang, A.D. Baughn, P. Wu, *Bioorg. Med. Chem. Lett.*, **21** **2011** 4989-4992.
98. M. Emmadi, S.S. Kulkarni, *Nat. Prod. Rep.*, **31** **2014** 870-879.
99. M.N. Amin, A. Ishiwata, Y. Ito, *Carbohydr. Res.*, **341** **2006** 1922-1929.
100. Y.J. Lee, A. Kubota, A. Ishiwata, Y. Ito, *Tet. Lett.*, **52** **2011** 418-421.
101. M. Zunk, J. Williams, J. Carter, M.J. Kiefel, *Org. Biomol. Chem.*, **12** **2014** 2918-2925.
102. J.T. Williams, L. Corcilius, M.J. Kiefel, R.J. Payne, *J. Org. Chem.*, **81** **2016** 2607-2611.

Overview of the Metabolic Oligosaccharide Engineering Toolbox

103. S.R. Sanapala, S.S. Kulkarni, *J. Am. Chem. Soc.*, **138** **2016** 4938-4947.
104. D.H. Dube, M.B. Koenigs, E.A. Richardson, *Abstr. Pap. Am. Chem. S.*, **239** **2010**.
105. E. Memmel, A. Homann, T.A. Oelschlaeger, J. Seibel, *Chem. Comm.*, **49** **2013** 7301-7303.
106. L. Cipolla, L. Gabrielli, D. Bini, L. Russo, N. Shaikh, *Nat. Prod. Rep.*, **27** **2010** 1618-1629.
107. M. Dumont, A. Lehner, B. Vauzeilles, J. Malassis, A. Marchant, K. Smyth, B. Linclau, A. Baron, J. Mas Pons, C.T. Anderson, D. Schapman, L. Galas, J.C. Mollet, P. Lerouge, *Plant J.*, **85** **2016** 437-447.
108. Y. Zhu, J. Wu, X. Chen, *Angew. Chem.*, **55** **2016** 9301-9305.
109. P.V. Chang, X. Chen, C. Smyrniotis, A. Xenakis, T. Hu, C.R. Bertozzi, P. Wu, *Angew. Chem.*, **48** **2009** 4030-4033.
110. D. Soriano Del Amo, W. Wang, H. Jiang, C. Besanceney, A.C. Yan, M. Levy, Y. Liu, F.L. Marlow, P. Wu, *J. Am. Chem. Soc.*, **132** **2010** 16893-16899.
111. J. Yang, J. Seckute, C.M. Cole, N.K. Devaraj, *Angew. Chem.*, **51** **2012** 7476-7479.
112. M. Matsuura, Y. Saikawa, K. Inui, K. Nakae, M. Igarashi, K. Hashimoto, M. Nakata, *Nat. Chem. Biol.*, **5** **2009** 465-467.
113. A.K. Spate, V.F. Schart, J. Hafner, A. Niederwieser, T.U. Mayer, V. Wittmann, *Beilstein J. Org. Chem.*, **10** **2014** 2235-2242.
114. A.K. Spate, J.E. Dold, E. Batroff, V.F. Schart, D.E. Wieland, O.R. Baudendistel, V. Wittmann, *ChemBioChem.*, **17** **2016** 1374-1383.
115. S. Han, B.E. Collins, P. Bengtson, J.C. Paulson, *Nat. Chem. Biol.*, **1** **2005** 93-97.
116. M.J. Garson, J.S. Simpson, *Nat. Prod. Rep.*, **21** **2004** 164-179.
117. H.C. Hang, C. Yu, D.L. Kato, C.R. Bertozzi, *Proc. Nat. Acad. Sci. U. S. A.*, **100** **2003** 14846-14851.
118. N.A. McGrath, R.T. Raines, *Chem. Sci.*, **3** **2012** 3237-3240.
119. A.R. Sherratt, N. Nasheri, C.S. McKay, S. O'Hara, A. Hunt, Z. Ning, D. Figeys, N.K. Goto, J.P. Pezacki, *ChemBioChem.*, **15** **2014** 1253-1256.
120. C.S. McKay, M. Chigrinova, J.A. Blake, J.P. Pezacki, *Org. Biomol. Chem.*, **10** **2012** 3066-3070.
121. R. Thomson, M. Von Itzstein, *Carbohydr. Res.*, **274** **1995** 29-44.
122. D.A. MacKenzie, A.R. Sherratt, M. Chigrinova, A.J. Kell, J.P. Pezacki, *Chem. Comm.*, **51** **2015** 12501-12504.
123. D. Seyferth, R.S. Marmor, P. Hilbert, *J. Org. Chem.*, **36** **1971** 1379-&.
124. W.D. Fessner, C. Gosse, G. Jaeschke, O. Eyrisch, *Eur. J. Org. Chem.*, **2000** 125-132.
125. S.G. Sampathkumar, A.V. Li, M.B. Jones, Z. Sun, K.J. Yarema, *Nat. Chem. Biol.*, **2** **2006** 149-152.
126. S.G. Sampathkumar, A.V. Li, K.J. Yarema, *Nat. Prot.*, **1** **2006** 2377-2385.
127. D.C. Xiong, J. Zhu, M.J. Han, H.X. Luo, C. Wang, Y. Yu, Y. Ye, G. Tai, X.S. Ye, *Org. Biomol. Chem.*, **13** **2015** 3911-3917.
128. S.M. Creedon, H.K. Crowley, D.G. McCarthy, *J. Chem. Soc. Perk. T. 1*, **1998** 1015-1017.
129. H.H. Chou, R.T. Raines, *J. Am. Chem. Soc.*, **135** **2013** 14936-14939.
130. N.E. Mbua, X. Li, H.R. Flanagan-Steet, L. Meng, K. Aoki, K.W. Moremen, M.A. Wolfert, R. Steet, G.J. Boons, *Angew. Chem.*, **52** **2013** 13012-13015.
131. T.M. Gloster, D.J. Vocadlo, *Nat. Chem. Biol.*, **8** **2012** 683-694.
132. T.N. Ramya, E. Weerapana, L. Liao, Y. Zeng, H. Tateno, L. Liao, J.R. Yates, 3rd, B.F. Cravatt, J.C. Paulson, *Mol. Cell Proteom.*, **9** **2010** 1339-1351.
133. D.J. Vocadlo, H.C. Hang, E.J. Kim, J.A. Hanover, C.R. Bertozzi, *Proc. Nat. Acad. Sci. U. S. A.*, **100** **2003** 9116-9121.
134. D.J. Vocadlo, *Curr. Opin. Chem. Biol.*, **16** **2012** 488-497.
135. S. Cecioni, D.J. Vocadlo, *Curr. Opin. Chem. Biol.*, **17** **2013** 719-728.
136. S. Wang, D.L. Shen, D. Lafont, A.S. Vercoutter-Edouart, M. Mortuaire, Y. Shi, O. Maniti, A.

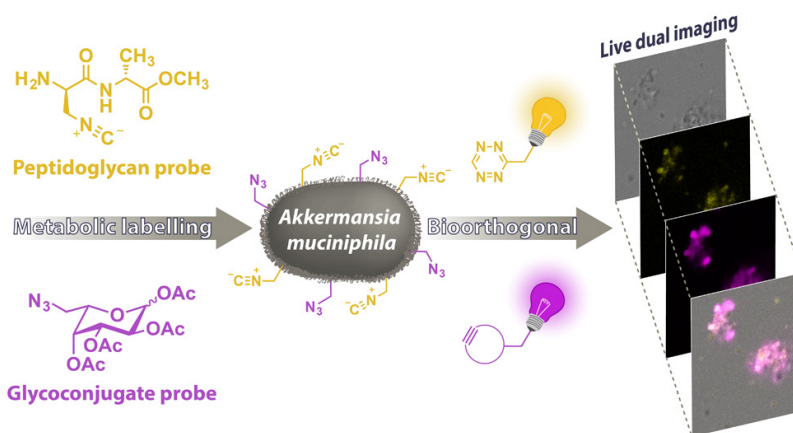
- Girard-Egrot, T. Lefebvre, B.M. Pinto, D. Vocadlo, S. Vidal, *MedChemComm*, **5** **2014** 1172-1178.
137. C.D. Rillahan, A. Antonopoulos, C.T. Lefort, R. Sonon, P. Azadi, K. Ley, A. Dell, S.M. Haslam, J.C. Paulson, *Nat. Chem. Biol.*, **8** **2012** 661-668.
138. T.M. Gloster, W.F. Zandberg, J.E. Heinonen, D.L. Shen, L. Deng, D.J. Vocadlo, *Nat. Chem. Biol.*, **7** **2011** 174-181.
139. A.R. Sherratt, M. Chigrinova, D.A. MacKenzie, N.K. Rastogi, M.T. Ouattara, A.T. Pezacki, J.P. Pezacki, *Bioconjug. Chem.*, **27** **2016** 1222-1226.
140. Y.A. Wainman, A.A. Neves, S. Stairs, H. Stockmann, H. Ireland-Zecchini, K.M. Brindle, F.J. Leeper, *Org. Biomol. Chem.*, **11** **2013** 7297-7300.
141. J. Kim, C.R. Bertozzi, *Angew. Chem.*, **54** **2015** 15777-15781.
142. H.W. Shih, J.A. Prescher, *J. Am. Chem. Soc.*, **137** **2015** 10036-10039.
143. R.K. Lim, Q. Lin, *Chem. Commun.*, **46** **2010** 7993-7995.
144. T.H. Poole, J.A. Reisz, W.L. Zhao, L.B. Poole, C.M. Furdul, S.B. King, *J. Am. Chem. Soc.*, **136** **2014** 6167-6170.
145. S.M. Agten, D.P. Suylen, T.M. Hackeng, *Bioconjug. Chem.*, **27** (**2016**) 42-46.
146. R. van Geel, M.A. Wijdeven, R. Heesbeen, J.M. Verkade, A.A. Wasiel, S.S. van Berkel, F.L. van Delft, *Bioconjug. Chem.*, **26** **2015** 2233-2242.
147. H.C. Hang, J.P. Wilson, G. Charron, *Acc. Chem. Res.*, **44** **2011** 699-708.
148. K. Lang, J.W. Chin, *Chem. Rev.*, **114** **2014** 4764-4806.
149. I.S. Wallace, C.T. Anderson, *Front. Plant Sci.*, **3** **2012** 89.
150. A.P. Bussink, P.F. van Swieten, K. Ghauharali, S. Scheij, M. van Eijk, T. Wennekes, G.A. van der Marel, R.G. Boot, J.M. Aerts, H.S. Overkleeft, *J. Lipid Res.* **48** **2007** 1417-1421.
151. B. Belardi, A. de la Zerda, D.R. Spiciarich, S.L. Maund, D.M. Peehl, C.R. Bertozzi, *Angew. Chem.*, **52** **2013** 14045-14049.
152. Y. Haga, K. Ishii, K. Hibino, Y. Sako, Y. Ito, N. Taniguchi, T. Suzuki, *Nat. Commun.*, **3** **2012** 907.
153. W. Lin, Y. Du, Y. Zhu, X. Chen, *J. Am. Chem. Soc.*, **136** **2014** 679-687.
154. W. Lin, L. Gao, X. Chen, *ChemBioChem*, **16** **2015** 2571-2575.
155. F. Doll, A. Buntz, A.K. Spate, V.F. Scharf, A. Timper, W. Schrimpf, C.R. Hauck, A. Zumbusch, V. Wittmann, *Angew. Chem.*, **55** **2016** 2262-2266.
156. J. Hoogenboom, N. Berghuis, D. Cramer, R. Geurts, H. Zuilhof, T. Wennekes, *BMC Plant Biol.*, **16** **2016** 220.

Chapter 3

Probing Peptidoglycan Synthesis and L-fucose Salvage in the Gut Commensal *Akkermansia muciniphila* with Bioorthogonal Chemical Reporters

"Another brick in the wall, Part 1"

The Wall, Pink Floyd



Published soon:

Tjerk J. Sminia, Hanne L.P. Tytgat, Steven Aalvink, Hanna de Jong, Marcel Tempelaars, Han Zuilhof, Tjakko Abée, Willem M. de Vos, and Tom Wennekes

Abstract

Our gut microbiota directly influences human physiology in health and disease. The diverse glycoconjugates in the bacterial cell envelope and our gut mucus dominate the microbiota-host interface and have been implicated in these complex interactions. We however currently lack a sufficient molecular understanding of the composition of these glycoconjugates for most individual species in our gut to study their functional activities. Here we report the chemically guided study of glycoconjugates in a key mucin-degrading gut microbiota member of the Verrucomicrobia phylum, *Akkermansia muciniphila*. Two novel non-toxic tetrazine click-compatible peptidoglycan probes allowed for the first time the demonstration of peptidoglycan synthesis in this intestinal species. Furthermore, metabolic labelling with azido-monosaccharides revealed that *A. muciniphila* can recycle components of mucin O-glycans, notably via an L-fucose salvage pathway. Finally, combining these bioorthogonal chemical probes allowed for the simultaneous detection of both peptidoglycan synthesis and fucosylation of cell surface glycoconjugates in live anaerobic gut bacteria. These findings enhance our molecular understanding of *Akkermansia muciniphila*'s basic physiology and cell envelope and offer a general approach to directly study different cell envelope glycoconjugates in other gut microbiota members.

Introduction

Research of the past decade has uncovered the great diversity and abundance of the human gut microbiota and how these mainly anaerobic bacteria directly influence our physiology in health and disease. It protects against invading microbes, aids in degrading non-digestible dietary components, produces key nutrients and vitamins, and primes our immune system.¹⁻³ Mapping the microbial molecules that mediate microbiota-host interactions is crucial to understand how this bacterial community controls human wellbeing by preserving mucosal homeostasis. We currently lack knowledge on the key molecules and basic physiology of its individual species to understand these complex interactions. Glycans and their glycoconjugates, both human and microbial, dominate the host-microbiota interface. Especially the diverse glycoconjugates in the cell envelope of individual microbiota members are ideally positioned to modulate these interactions. It is therefore crucial that we acquire a molecular understanding of their composition and activities. As the assembly and composition of glycans is not genetically encoded, a chemically guided approach is ideally suited to identify and modulate their structures in their native setting to allow subsequent study of their functional activities. To achieve this for gut microbiota there is however a critical need for validated bioorthogonal chemical reporters and techniques that enable these studies in living anaerobic bacteria.

We set out to develop such a method for the investigation of the cell envelope glycoconjugates of a key member of a healthy human gut microbiota: *Akkermansia muciniphila*, a Gram-negative, oval-shaped and non-motile bacterium that belongs to the phylum *Verrucomicrobia*.⁴⁻⁵ It colonises

the mucus layer of the gastrointestinal tract and specialises in the degradation of complex mucosal *O*-glycans. As a mucin-degrader, *A. muciniphila* is less dependent on the influx of dietary components, being able to use mucus as its sole carbon and nitrogen source,⁵ and supports other microbiota members by providing access to mucus monosaccharides. This places it in a key position to shape the overall microbial composition. Due to its close proximity within the mucus layer to its human host *A. muciniphila* may also exert a relatively large effect on our immune system that it has indeed been found to stimulate. *A. muciniphila* also stimulates host mucin production, increases mucus layer thickness and strengthens intestinal barrier function.⁶⁻¹⁰ Its lowered intestinal colonisation has been correlated with inflammatory bowel conditions in humans. An intervention study in mice showed its potential to protect against high-fat-diet induced obesity and more recently, our laboratory showed the potential of *A. muciniphila* as a therapeutic for human obesity and associated disorders.¹¹⁻¹²

To understand and modulate the interaction of *A. muciniphila* with its host and the other gut microbiota members it is essential to first comprehend its own basic physiology at the molecular level and especially its understudied cell envelope glycoconjugates. An important outstanding question in this respect is if *A. muciniphila*, as a member of the PVC clade⁵, has peptidoglycan (PG) in its cell envelope. If present, this glycoconjugate plays a key role in cell division, cell envelope maintenance and through human peptidoglycan recognition proteins also in microbiota-host interaction.¹³ Another key question is whether, besides degrading, it can salvage the monosaccharides that it obtains from mucin *O*-glycans and use them to modulate the composition of its own cell envelope glycoconjugates. This would allow mimicry of host glycan composition that, especially for L-fucose, has been shown to confer a gut colonization advantage.¹⁴

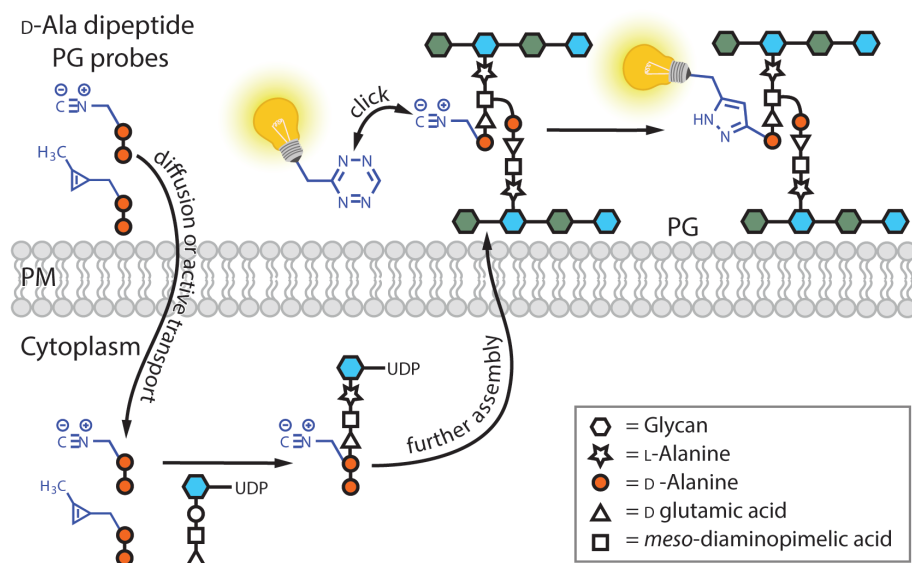


Figure 1. Overview of metabolic labeling with PG probes. PM: plasma membrane, PG: peptidoglycan, UDP: uridine diphosphate.

Here, we report the development of two non-toxic novel D-Ala-dipeptide based PG probes, equipped with an isonitrile or cyclopropene chemical reporter, that were used to prove, for the first time, that *A. muciniphila* indeed synthesises peptidoglycan in its cell envelope periplasm (see Figure 1). Both probes and their labelling with a tetrazine click reaction proved non-toxic under anaerobic culturing conditions and allowed for PG detection via flow cytometry and confocal microscopy in live bacteria. Metabolic labelling of *A. muciniphila* with azido-labelled monosaccharide analogs of key O-glycan components, uncovered its ability to salvage L-fucose. Finally, we show that a combination of our novel isonitrile-based PG probe and 6-azido-L-fucose enables the simultaneous imaging of PG and fucosylated glycoconjugates in the cell envelope of live *A. muciniphila*.

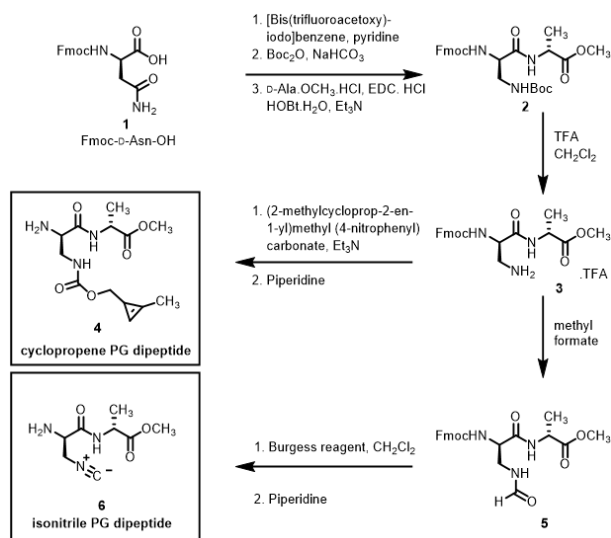
Results and Discussion

Design and Synthesis of New Cyclopropene and Isonitrile-based Peptidoglycan Probes. Our first goal was to develop probes that would be able to detect and image peptidoglycan synthesis in live anaerobic bacteria. As a constituent of bacterial cell envelopes that preserves cell rigidity and shape, PG forms an anchoring point for many cell wall-associated molecules and plays a key role in cell growth and division.¹⁵⁻¹⁷ If present it forms a unique opportunity to get insight in the cell envelope dynamics of bacterial cells and as mentioned it is currently unknown if PG is produced by *A. muciniphila*. PG consists of an alternating polymeric chain of β -1,4-linked *N*-acetylglucosamine and *N*-acetylmuramic acid with a pentapeptide attached to the latter that enables crosslinking between chains. The so-called PG stem peptide contains two terminal D-Alanine residues that are of special interest, as these are unique to bacterial PG. In recent years, this fact has been exploited to study PG via metabolic incorporation of unnatural D-amino acid-based probes (Figure 1). Although some of these probes have a prepended bulky fluorophore, most probes rely on a secondary click reaction step between a chemical reporter group on the metabolically incorporated D-Ala-based probe and a complementary group attached to a fluorophore.¹⁸⁻²⁸ This smaller size of the initial probe also better permits PG labeling of Gram-negative bacteria whose outer membrane make them less accessible to fluorophore-prepended bulky PG probes. Liechti and co-workers used this approach to prove the presence of PG in another PVC clade member, *Chlamydia trachomatis*, with an alkyne-containing D-Ala dipeptide probe.²² Their probe however relies on a copper-catalysed azide-alkyne cycloaddition (CuAAC) in which the used copper is toxic to bacteria. In general, there is a lack of bioorthogonal PG probes that closely mimic D-alanine and even more crucially that have been validated for their efficacy in non-toxic labelling and detection of bacterial PG *in vivo*.

The [4+1] cycloaddition of isonitriles and tetrazines is a recent addition²⁹⁻³⁰ to the bioorthogonal click chemistry toolbox and we envisioned that an isonitrile group would allow for a sterically very close D-alanine mimic. This close mimicry should result in only a small perturbation of the PG macromolecular structure once incorporated and thereby limit an effect on the dynamics of PG production and degradation. Additionally, the inverse electron demand Diels Alder (invDA) reaction between strained alkenes and tetrazines is among the fastest bioorthogonal reactions currently known and allows for the use of relatively compact cyclopropene chemical reporters.³¹ Both these copper independent click

reactions have been proven to be bioorthogonal and were already applied successfully in *in vivo* glycan imaging.^{30, 32-33} We therefore designed a D-Ala-based probe with either an isonitrile or cyclopropene chemical reporter group. As modified single D-Ala amino acids are sometimes poorly incorporated into PG and are not selective in labelling PG formation and maturation dynamics^{21, 28}, we choose to adapt the design reported by Liechti and co-workers based on a D-Ala dipeptide motif.²² By installing the bio-orthogonal isonitrile or cyclopropene chemical reporter group on the *N*-terminus of the D-Ala dipeptide we could target labelling to the crucial fourth position of the stem peptide²² that is actively involved in PG formation and maturation (Figure 1).

We started our synthesis route with the commercially available Fmoc-D-asparagine-OH **1** for both the cyclopropene and isonitrile probe (see Scheme 1, detailed synthetic procedures in **Appendix A**). A hypervalent iodine mediated Hofmann rearrangement under basic conditions efficiently produced an amine intermediate in 99% yield.³⁴ After temporary Boc-protection of the obtained amine, a peptide coupling under standard carbodiimide coupling conditions with D-Ala-OCH₃-HCl produced dipeptide precursor **2**. The Boc-group was removed with TFA, yielding the free amine dipeptide as a TFA salt intermediate **3**. From this point onwards the synthesis of the cyclopropene **4** and isonitrile **6** dipeptide PG probe diverged. In the case of the cyclopropene-containing probe, the free amine was reacted with activated *para*-nitrophenylcarbonate cyclopropene resulting in a cyclopropene handle on the D-Ala dipeptide.³² The synthesis of this cyclopropene handle was reported by Prescher and co-workers and later adapted by us and others.^{32-33, 35-36} From the diverse family of cyclopropenes we chose a methyl-cyclopropene carbamate handle as the most suitable in terms of stability during probe synthesis and reactivity during the labelling.



Scheme 1. Synthesis of novel cyclopropene- and isonitrile-containing D-Ala dipeptide PG probes **4** and **6**.

In case of the isonitrile PG probe, the free amine was reacted with methyl formate. Different methods for the formation of a formamide are known,³⁷⁻³⁸ but the procedure using methyl formate proved the most reproducible in providing **5**. Other procedures such as formic acid in presence of dicyclohexylcarbodiimide (DCC) resulted in byproduct formation while other methods with catalytic boric acid gave no product formation. With formamide **5** in hand we next assessed the standard method for formamide dehydration with POCl_3 . Although complete consumption of **5** was observed, this method was not successful in producing the isonitrile. A milder dehydration method using Burgess reagent did however succeed in producing the isonitrile.³⁹ Next, piperidine treatment removed the Fmoc group in both PG dipeptide probe precursors and provided our target PG probes **4** and **6** (Scheme 1). We expected that the remaining methyl ester on the C-terminal D-Ala amino acid would be hydrolysed by nonspecific bacterial esterases, as these dipeptides are believed to selectively label PG via a cytoplasmic route.²² The ester might also facilitate the initial passive uptake of the dipeptide probe **4** and **6**.

Novel PG probes are non-toxic to bacteria. The aim of our bioorthogonal PG probes was to image PG synthesis in live bacteria, but the incorporation of our unnatural D-Ala dipeptides **4** and **6** might have detrimental effects on bacterial growth, viability and physiology. Therefore, we first evaluated their toxicity. As a benchmark, we first tested toxicity of the probes in *Escherichia coli*, which did not reveal any detrimental effects of the probes on specific growth rate (Figure 2A). Also in *A. muciniphila* no significant growth defect was observed after addition of our isonitrile and cyclopropene probes. Whilst addition of the copper-dependent alkyne probe of Liechti et al., which was used as a benchmark, resulted in a significant growth defect in *A. muciniphila* cultures (Figure 2B).²²

Next, we evaluated the combined toxicity of metabolic PG probe incorporation and the subsequent click reaction with fluorescent reporter groups for our probes in comparison to the alkyne PG probe with associated CuAAC click reaction. The viability of the bacteria was assessed after the click reaction by determining colony-forming units on CP-mucin agar for *A. muciniphila* and LB agar for *E. coli* (Appendix A, Figure S7C and S7D). This clearly revealed the expected cytotoxicity of the copper-dependent alkyne PG dipeptide probe and subsequent CuAAC reaction. In stark contrast with the alkyne probe, both *E. coli* and *A. muciniphila* showed sustained viability when subjected to our novel cyclopropene **4** and isonitrile **6** PG probes in combination with the bioorthogonal invDA or [4+1] cycloaddition reaction.

Labelling of PG in *E. coli* and *Akkermansia muciniphila*. The efficacy of both our novel, non-toxic dipeptide probes (**4** and **6**) to detect and image PG synthesis was first successfully confirmed in *E. coli* using flow cytometry (FCM) and confocal imaging experiments (Figure S1 & S2 in Appendix A) and favorably compared to the established alkyne-based PG probe of Liechti and coworkers.²² This initial result showcases the potential wider application of our probes for PG imaging.

We next assessed PG presence in the cell envelope periplasm of *A. muciniphila*. Anaerobic cultures were incubated in the presence of either probe **4** or **6** for 1 hour prior to sampling. The invDA or [4+1]

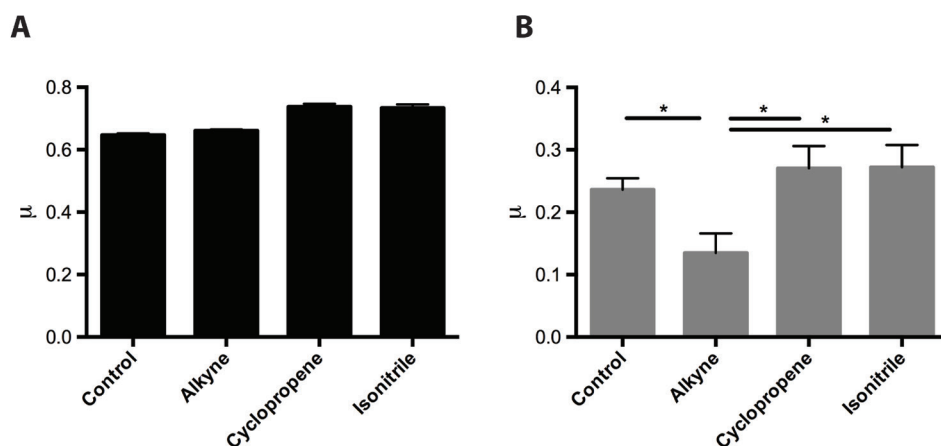


Figure 2. Specific growth rate (μ) of *E. coli* (A) and *A. muciniphila* (B) in the presence of PG dipeptide probes (data represents mean of triplicates, error bars depict SEM, T = 37 °C).

cycloaddition with the fluorophore-appended click reagents, such as Cy3 tetrazine dyes, proceeded fast and did not affect the morphology of *A. muciniphila*. Confocal microscopy images of *A. muciniphila* probed (Figure 3 & Figure S4 in Appendix A) with the non-toxic PG dipeptide probes showed clear fluorescence and for the first time provide compelling evidence for PG synthesis in this intestinal species. The efficacy of PG labelling with isonitrile probe **6** in *A. muciniphila* was also evaluated by FCM analysis (Figure 5B), which was in line with the obtained confocal images. Overall, using both probes, the presence of PG in the human gut microbiota member *A. muciniphila* could be shown. The compatibility of these non-toxic probes and subsequent click reactions with bacterial growth opens perspectives towards the further use of these probes to monitor PG dynamics *in vivo* for this and other gut microbiota members.

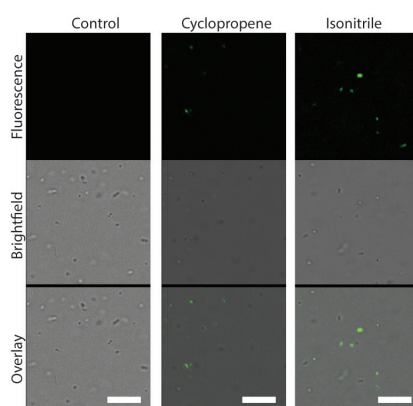


Figure 3. Confocal microscopy images of *A. muciniphila* with and without addition of the novel non-toxic PG dipeptide probes. Scale bar = 5 μ m.

Probing *Akkermansia muciniphila* with Chemical Reporters

Another advantage of these new PG probes is their orthogonality with other types of bioorthogonal chemistries such as the strain-promoted azide-alkyne cycloaddition (SPAAC).

Investigation of glycan salvage in *A. muciniphila*. In addition to the role of peptidoglycan in the cell envelope maintenance and structure, other bacterial cell envelope glycoconjugates also play a crucial but not yet fully understood role in microbiota-host interactions.⁴⁰ In this second part of the study we aimed to assess the ability of *A. muciniphila* to salvage glycans, such as mucin O-glycan derived monosaccharides, from its environment to incorporate in its own cell surface glycoconjugates, like glycoproteins and LPS. In recent years, L-fucose has especially been identified as a key glycan in a human gut microbiome homeostasis and dysbiosis.^{14, 41} Besides an engineered pathway in *E. coli*,⁴² a native glycan salvage pathway for L-fucose has up till now only been reported for *Bacteroides fragilis*, another important inhabitant of the human gut.¹⁴ *B. fragilis* can salvage mucin glycan derived L-fucose monomers, and use these to decorate its surface polysaccharides. Both enzyme functions necessary to salvage fucose are present in a bifunctional enzyme Fkp, capable of converting L-fucose to L-fucose 1-phosphate and generation of GDP-L-fucose.¹⁴

To assess if the mucin-foraging *A. muciniphila* can use mucin-derived monosaccharides to incorporate in its surface glycoconjugates, we set up a metabolic oligosaccharide engineering (MOE) experiment (Figure 4A).³¹ MOE relies on the presence of glycan salvage pathways to incorporate monosaccharides equipped with a reporter group (e.g. click handle) into glycoconjugates. For this reason, MOE is becoming an increasingly popular technique for the detection and tracking of bacterial glycans, also in human gut microbiota members, thus conferring evidence on the presence of glycan salvage pathways in these species.⁴³⁻⁴⁷

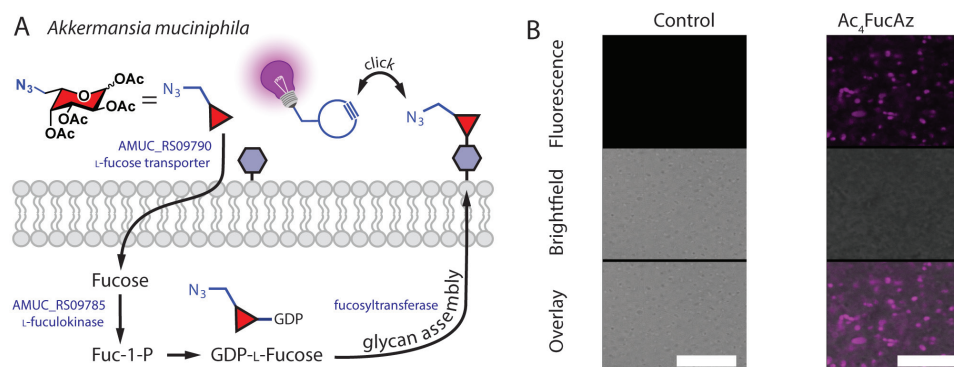


Figure 4. (A) MOE labelling with Ac₄FucAz in *A. muciniphila* and fucose salvage pathway (B) *A. muciniphila* with control and Ac₄FucAz, labeled with Cy-5. Scale bar = 5 μ m.

We evaluated incorporation of the 6-azido analogue of L-fucose, Ac₄FucAz, (Figure 4B). Ac₄FucAz was reported to be toxic in mammalian cell lines⁴⁸ and therefore the alkyne variant was used in various organisms including bacteria.⁴⁹ Here we applied Ac₄FucAz for the first time in bacteria, as this allowed us to use the SPAAC reaction for imaging, excluding the use of cytotoxic copper, and allowing for a follow up dual labelling experiment with our PG probes, thus enabling in-depth studies of the cell envelope surface glycoconjugates. Results showed strong incorporation of Ac₄FucAz in the surface glycan layer of *A. muciniphila* (Figure 5B). Confocal imaging results were confirmed by FCM analysis (Figure 5C).

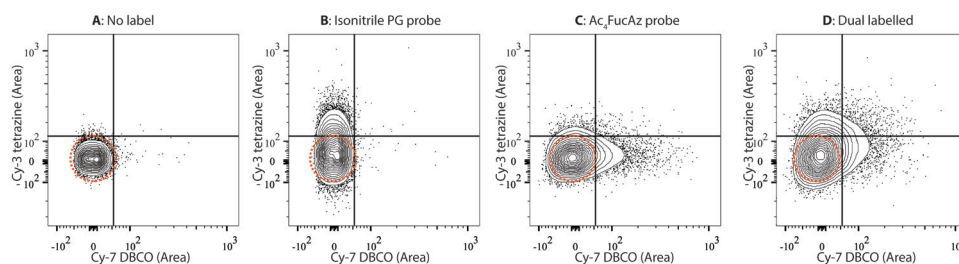


Figure 5. FCM analysis of Cy-3/Cy7 fluorescently labelled *Akkermansia muciniphila* cells with; A: no probe added, B: with PG isonitrile probe (4), C: with Ac₄FucAz probe, D: after dual labelling with both probes.

The presence of highly fucosylated glycoconjugates on the cell surface and thus a fucose salvage pathway in *A. muciniphila* is striking but corroborated by a recently established metabolic model of *A. muciniphila*.⁹ This model showed the genomic potential of *A. muciniphila* to salvage L-fucose, as it indicates the presence of a L-fucose transporter in *A. muciniphila* (AMUC_RS09790) and a L-fuculokinase (AMUC_RS09775) that can generate the crucial L-fucose-1-phosphate needed for salvage (Figure 6A). The genome also harbors several pyrophosphorylases, that can play a role in producing the GDP-L-fucose substrate for fucosyltransferases, but it remains to be investigated which of these pyrophosphorylases in its genome are involved in the production of GDP-L-fucose.⁵⁰ Finally, *A. muciniphila* has also been found to harbor fucosyltransferases from the GT11 CAZy family.⁵¹

We further determined whether two other azide-modified probe versions of monosaccharide O-glycan components, *N*-acetylglucosamine (Ac₄GlcNAz) and *N*-acetylgalactosamine (Ac₄GalNAz), could be salvaged by *A. muciniphila* and incorporated in its surface glycoconjugates. A MOE experiment indeed indicated significant incorporation of GlcNAz and to a lesser extent of GalNAz, thus indicating salvage of GlcNAc and GalNAc by *A. muciniphila* (Figure S5). Overall this successful application of MOE for detection and imaging of bacterial glycan salvage and cell surface glycoconjugates in *A. muciniphila* will permit in the future the molecular level study of this and other bacteria in their glycan-dominated anaerobic intestinal environment.

Proof of principle. Live dual imaging of PG synthesis and surface fucosylation of the bacterial cell envelope. Dual labeling applications, in which both PG and the dynamics of other key glycoconjugates in the cell envelope can be tracked simultaneously or longitudinally in live anaerobic bacteria would open interesting perspectives to study their role at the microbiota-host interface. The orthogonality of our novel PG dipeptide probes **4** and **6** with other types of bioorthogonal chemical reporters should permit this. We therefore performed a dual labelling experiment of PG and surface fucosylation in *A. muciniphila*. We used Ac₄FucAz for labelling cell surface glycans in *A. muciniphila* and the isonitrile dipeptide probe **6** to label PG. For confocal analysis, the incorporated isonitrile PG dipeptide probe was clicked with Cy-3 tetrazine while incorporated Ac₄FucAz was clicked with Cy-5 DBCO (Figure 6). Single labelling control images for both probes are available in Figure S3. Overlay images of live bacteria show labelling with both probes in *A. muciniphila* thus illustrating the orthogonality and bioorthogonality of these probes and the attachment of their respective reporter fluorophores (Figure 6). Dual labelling was also evaluated using FCM. To ensure compatibility with available FCM filters, Ac₄FucAz was clicked with Cy-7 DBCO, whilst the isonitrile probe was clicked with a Cy-3 tetrazine. Dual labelling of both peptidoglycan and fucosylated cell surface glycans in live *A. muciniphila* cells was confirmed by the FCM data (Figure 5D).

This successful proof of principle experiment illustrates the potential for *in vivo* follow-up studies of cell envelope glycoconjugates in *A. muciniphila* and other anaerobic gut microbiota members. It allows the simultaneous spatial and temporal tracking and modulation of these microbial molecules. One of them being peptidoglycan synthesis in the bacterial cell envelope. It functions as a marker for viability and active cell envelope maintenance, and is involved in microbiota-host interaction via human peptidoglycan binding proteins. The other being the study of surface cell envelope glycoconjugates directly, like fucosylated glycoconjugates, that dominate the microbiota-host interface and are suspected to influence it, among others via mimicry of host glycan composition.

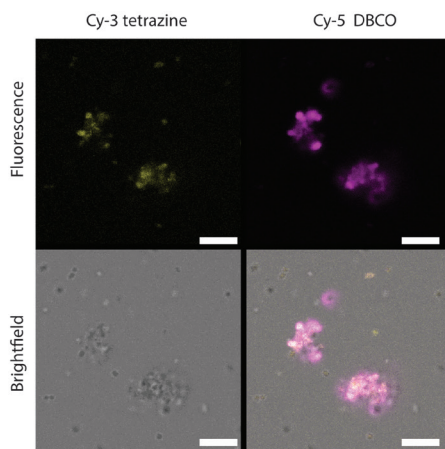


Figure 6. Dual labelling in *A. muciniphila* with isonitrile PG dipeptide probe **6** labelled with Cy-3 tetrazine (top left), Ac₄FucAz with Cy-5 DBCO (top right), brightfield (bottom left), and overlay (bottom right). Scale bar = 5 μ m.

Conclusions

The newly developed non-toxic cyclopropene **4** and isonitrile **6** D-Ala dipeptide PG probes, together with the MOE study, allowed us to successfully perform the first cell envelope engineering experiments in a live human gut microbiota member, namely the mucin degrader, *A. muciniphila*. This provided for the first time direct evidence for the presence of PG, cell surface glycoconjugates and an active L-fucose glycan salvage pathway in this bacterium. Both these classes of cell envelope glycoconjugates are positioned at the microbiota-host interface and are implicated but not yet well understood mediators of the complex interactions tied to human wellbeing.

Here we unambiguously show the presence of surface fucosylated glycoconjugates in *A. muciniphila* thus substantiating their predicted presence based on its genomic potential. Elucidating these basic aspects of *A. muciniphila* molecular physiology will aid in further developing therapeutic opportunities for this key human gut microbiota member. Furthermore, the orthogonality of these new PG probes with the strain promoted azide alkyne cycloadditions (SPAAC) allowed for live dual labelling and imaging of both PG and fucosylated glycoconjugates in their native cell envelope setting.

The here described novel approach, using both D-Ala dipeptide and monosaccharide analogues – simultaneously – with bioorthogonal chemical reporters, adds to the bacterial metabolic labelling molecular toolbox. It offers an interesting and powerful new chemically guided approach to functionally analyse the role of cell envelope glycoconjugates in key microbiota members. Their application in both an aerobic and anaerobic bacterium illustrate their potential for the *in-situ* investigation of the glycoconjugate repertoire of both aerobic and anaerobic microbiota members.

Associated content

The Supporting Information, including probe synthesis, can be found in **Appendix A** (page 167).

Acknowledgement

The authors thank the Netherlands Foundation for Scientific Research (NWO) for funding via a VENI (722.011.006) and VIDI (723.014.005) grant of TW. HT holds a Marie Skłodowska-Curie fellowship of the European Union's Horizon 2020 research and innovation program under agreement No 703577 (Glycoli) to support her work at ETH Zurich. Work at the Laboratory of Microbiology was supported by the Soehngen Institute of Anaerobic Microbiology (SIAM Gravity Grant of NWO, 024.002.002) and NWO Spinoza award of WMdV. We gratefully acknowledge Jan Willem Borst for confocal microscopy support (Laboratory of Biochemistry, Wageningen University). Yvette Luijkx and Jona Merx are acknowledged for synthesis support.

References

1. Arumugam, M.; Raes, J.; Pelletier, E.; Le Paslier, D.; Yamada, T.; Mende, D. R.; Fernandes, G. R.; Tap, J.; Bruls, T.; Batto, J. M.; Bertalan, M.; Borruel, N.; Casellas, F.; Fernandez, L.; Gautier, L.; Hansen, T.; Hattori, M.; Hayashi, T.; Kleere bezem, M.; Kurokawa, K.; Leclerc, M.; Levenez, F.; Manichanh, C.; Nielsen, H. B.; Nielsen, T.; Pons, N.; Poulain, J.; Qin, J.; Sicheritz-Ponten, T.; Tims, S.; Torrents, D.; Ugarte, E.; Zoetendal, E. G.; Wang, J.; Guarner, F.; Pedersen, O.; de Vos, W. M.; Brunak, S.; Dore, J.; Antolin, M.; Artiguenave, F.; Blottiere, H. M.; Almeida, M.; Brechot, C.; Cara, C.; Chervaux, C.; Cultrone, A.; Delorme, C.; Denariac, G.; Dervyn, R.; Foerstner, K. U.; Friss, C.; van de Guchte, M.; Guedon, E.; Haimet, F.; Huber, W.; van Hylckama-Vlieg, J.; Jamet, A.; Juste, C.; Kaci, G.; Knol, J.; Lakhdari, O.; Layec, S.; Le Roux, K.; Maguin, E.; Merieux, A.; Melo Minardi, R.; M'Rini, C.; Muller, J.; Oozeer, R.; Parkhill, J.; Renault, P.; Rescigno, M.; Sanchez, N.; Sunagawa, S.; Torrejon, A.; Turner, K.; Vandemeulebrouck, G.; Varela, E.; Winogradsky, Y.; Zeller, G.; Weissenbach, J.; Ehrlich, S. D.; Bork, P., *Nature* **2011**, *473* (7346), 174-80.
2. Qin, J.; Li, R.; Raes, J.; Arumugam, M.; Burgdorf, K. S.; Manichanh, C.; Nielsen, T.; Pons, N.; Levenez, F.; Yamada, T.; Mende, D. R.; Li, J.; Xu, J.; Li, S.; Li, D.; Cao, J.; Wang, B.; Liang, H.; Zheng, H.; Xie, Y.; Tap, J.; Lepage, P.; Bertalan, M.; Batto, J. M.; Hansen, T.; Le Paslier, D.; Linneberg, A.; Nielsen, H. B.; Pelletier, E.; Renault, P.; Sicheritz-Ponten, T.; Turner, K.; Zhu, H.; Yu, C.; Li, S.; Jian, M.; Zhou, Y.; Li, Y.; Zhang, X.; Li, S.; Qin, N.; Yang, H.; Wang, J.; Brunak, S.; Dore, J.; Guarner, F.; Kristiansen, K.; Pedersen, O.; Parkhill, J.; Weissenbach, J.; Bork, P.; Ehrlich, S. D.; Wang, J., *Nature* **2010**, *464* (7285), 59-65.
3. The Human Microbiome Project Consortium, Structure, function and diversity of the healthy human microbiome. *Nature* **2012**, *486* (7402), 207-14.
4. Belzer, C.; de Vos, W. M., *ISME J.* **2012**, *6* (8), 1449-58.
5. Derrien, M.; Vaughan, E. E.; Plugge, C. M.; de Vos, W. M., *Int. J. Syst. Evol. Microbiol.* **2004**, *54* (Pt 5), 1469-76.
6. Derrien, M.; Belzer, C.; de Vos, W. M., *Microb. Pathog.* **2016**.
7. Derrien, M.; Van Baaren, P.; Hooiveld, G.; Norin, E.; Muller, M.; de Vos, W. M., *Front. microbiol.* **2011**, *2*, 166.
8. Png, C. W.; Linden, S. K.; Gilshenan, K. S.; Zoetendal, E. G.; McSweeney, C. S.; Sly, L. I.; McGuckin, M. A.; Florin, T. H., *Am. J. Gastroenterol.* **2010**, *105* (11), 2420-8.
9. Reunanen, J.; Kainulainen, V.; Huuskonen, L.; Ottman, N.; Belzer, C.; Huhtinen, H.; de Vos, W. M.; Satokari, R., *Appl. Environ. Microbiol.* **2015**, *81* (11), 3655-62.
10. Shin, N. R.; Lee, J. C.; Lee, H. Y.; Kim, M. S.; Whon, T. W.; Lee, M. S.; Bae, J. W., *Gut* **2014**, *63* (5), 727-35.
11. Everard, A.; Belzer, C.; Geurts, L.; Ouwerkerk, J. P.; Druart, C.; Bindels, L. B.; Guiot, Y.; Derrien, M.; Muccioli, G. G.; Delzenne, N. M.; de Vos, W. M.; Cani, P. D., *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110* (22), 9066-71.
12. Plovier, H.; Everard, A.; Druart, C.; Depommier, C.; Van Hul, M.; Geurts, L.; Chilloux, J.; Ottman, N.; Duparc, T.; Lichtenstein, L.; Myridakis, A.; Delzenne, N. M.; Klievink, J.; Bhattacharjee, A.; van der Ark, K. C.; Aalvink, S.; Martinez, L. O.; Dumas, M. E.; Maiter, D.; Loumaye, A.; Hermans, M. P.; Thissen, J. P.; Belzer, C.; de Vos, W. M.; Cani, P. D., *Nat. Med.* **2016**.
13. Royet, J.; Gupta, D.; Dziarski, R., *Nat. Rev. Immunol.* **2011**, *11* (12), 837-51.
14. Coyne, M. J.; Reinap, B.; Lee, M. M.; Comstock, L. E., *Science* **2005**, *307* (5716), 1778-1781.
15. Egan, A. J.; Cleverley, R. M.; Peters, K.; Lewis, R. J.; Vollmer, W., *FEBS J* **2016**.
16. Typas, A.; Banzhaf, M.; Gross, C. A.; Vollmer, W., *Nature Rev. Microbiol.* **2012**, *10* (2), 123-136.
17. Vollmer, W.; Seligman, S. J., *Trends Microbiol.* **2010**, *18* (2), 59-66.

18. Gale, R. T.; Brown, E. D., *Curr. Opin. Microbiol.* **2015**, *27*, 69-77.
19. Kocaoglu, O.; Carlson, E. E., *Nat. Chem. Biol.* **2016**, *12* (7), 472-8.
20. Kuru, E.; Hughes, H. V.; Brown, P. J.; Hall, E.; Tekkam, S.; Cava, F.; de Pedro, M. A.; Brun, Y. V.; VanNieuwenhze, M. S., *Angew. Chem.* **2012**, *51* (50), 12519-23.
21. Kuru, E.; Tekkam, S.; Hall, E.; Brun, Y. V.; Van Nieuwenhze, M. S., *Nat. Protoc.* **2015**, *10* (1), 33-52.
22. Liechti, G. W.; Kuru, E.; Hall, E.; Kalinda, A.; Brun, Y. V.; VanNieuwenhze, M.; Maurelli, A. T., *Nature* **2014**, *506* (7489), 507-10.
23. Pidgeon, S. E.; Fura, J. M.; Leon, W.; Birabakaran, M.; Vezenov, D.; Pires, M. M., *Angew. Chem.* **2015**, *54* (21), 6158-62.
24. Pidgeon, S. E.; Pires, M. M., *Chem. Commun.* **2015**, *51* (51), 10330-3.
25. Sherratt, A. R.; Chigrinova, M.; MacKenzie, D. A.; Rastogi, N. K.; Ouattara, M. T.; Pezacki, A. T.; Pezacki, J. P., *Dual Bioconjug. Chem.* **2016**, *27* (5), 1222-6.
26. Shieh, P.; Siegrist, M. S.; Cullen, A. J.; Bertozzi, C. R., *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111* (15), 5456-61.
27. Siegrist, M. S.; Swarts, B. M.; Fox, D. M.; Lim, S. A.; Bertozzi, C. R., *FEMS Microbiol. Rev.* **2015**, *39* (2), 184-202.
28. Siegrist, M. S.; Whiteside, S.; Jewett, J. C.; Aditham, A.; Cava, F.; Bertozzi, C. R., *ACS Chem. Biol.* **2013**, *8* (3), 500-5.
29. Stockmann, H.; Neves, A. A.; Stairs, S.; Brindle, K. M.; Leeper, F. J., *Org. Biomol. Chem.* **2011**, *9* (21), 7303-5.
30. Stairs, S.; Neves, A. A.; Stockmann, H.; Wainman, Y. A.; Ireland-Zecchini, H.; Brindle, K. M.; Leeper, F. J., *Chembiochem* **2013**, *14* (9), 1063-7.
31. Sminia, T. J.; Zuilhof, H.; Wennekes, T., *Carbohydr Res.* **2016**, *435*, 121-141.
32. Patterson, D. M.; Nazarova, L. A.; Xie, B.; Kamber, D. N.; Prescher, J. A., *J. Am. Chem. Soc.* **2012**, *134* (45), 18638-18643.
33. Yang, J.; Seckute, J.; Cole, C. M.; Devaraj, N. K., *Angew. Chem.* **2012**, *51* (30), 7476-9.
34. Lau, Y. H.; Spring, D. R., *Synlett* **2011**, (13), 1917-1919.
35. Patterson, D. M.; Jones, K. A.; Prescher, J. A., *Mol. Biosyst.* **2014**, *10* (7), 1693-7.
36. Spate, A. K.; Schart, V. F.; Hafner, J.; Niederwieser, A.; Mayer, T. U.; Wittmann, V., *Beilstein J. Org. Chem.* **2014**, *10*, 2235-42.
37. Aguiam, N. R.; Castro, V. I.; Ribeiro, A. I. F.; Fernandes, R. D. V.; Carvalho, C. M.; Costa, S. P. G.; Pereira-Lima, S. M. M. A., *Tetrahedron* **2013**, *69* (43), 9161-9165.
38. Nguyen, T. B.; Sorres, J.; Tran, M. Q.; Ermolenko, L.; Al-Mourabit, A., *Org. Lett.* **2012**, *14* (12), 3202-5.
39. Creedon, S. M.; Crowley, H. K.; McCarthy, D. G., *J. Chem. Soc. Perkin Trans. I.* **1998**, (6), 1015-1017.
40. Tytgat, H. L.; de Vos, W. M., *Trends Microbiol.* **2016**, *24* (11), 853-861.
41. Comstock, L. E.; Kasper, D. L., *Cell* **2006**, *126* (5), 847-50.
42. Yi, W.; Liu, X.; Li, Y.; Li, J.; Xia, C.; Zhou, G.; Zhang, W.; Zhao, W.; Chen, X.; Wang, P. G., *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106* (11), 4207-12.
43. Dumont, A.; Malleron, A.; Awwad, M.; Dukan, S.; Vauzeilles, B., *Angew. Chem.* **2012**, *51* (13), 3143-6.
44. Geva-Zatorsky, N.; Alvarez, D.; Hudak, J. E.; Reading, N. C.; Erturk-Hasdemir, D.; Dasgupta, S.; von Andrian, U. H.; Kasper, D. L., *Nat. Med.* **2015**, *21* (9), 1091-100.
45. Liu, F.; Aubry, A. J.; Schoenhofen, I. C.; Logan, S. M.; Tanner, M. E., *Chembiochem* **2009**, *10* (8), 1317-20.
46. 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., *J. Am. Chem. Soc.* **2012**, *134* (39), 16123-6.
47. Tra, V. N.; Dube, D. H., *Chem. Commun.* **2014**, *50* (36), 4659-73.
48. Hsu, T. L.; Hanson, S. R.; Kishikawa, K.; Wang, S. K.; Sawa, M.; Wong, C. H., *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104* (8),

- 2614-9.
49. Besanceney-Webler, C.; Jiang, H.; Wang, W.; Baughn, A. D.; Wu, P., *Bioorg. Med. Chem. Lett.* **2011**, 21 (17), 4989-92.
50. Ottman, N.; Huuskonen, L.; Reunanen, J.; Boeren, S.; Klievink, J.; Smidt, H.; Belzer, C.; de Vos, W. M., *Front. Microbiol* **2016**, 7, 1157.
51. Sun, Z.; Harris, H. M.; McCann, A.; Guo, C.; Argimon, S.; Zhang, W.; Yang, X.; Jeffery, I. B.; Cooney, J. C.; Kagawa, T. F.; Liu, W.; Song, Y.; Salvetti, E.; Wrobel, A.; Rasinkangas, P.; Parkhill, J.; Rea, M. C.; O'Sullivan, O.; Ritari, J.; Douillard, F. P.; Paul Ross, R.; Yang, R.; Briner, A. E.; Felis, G. E.; de Vos, W. M.; Barrangou, R.; Klaenhammer, T. R.; Caufield, P. W.; Cui, Y.; Zhang, H.; O'Toole, P. W., *Nat. Comm.* **2015**, 6, 8322.



Chapter 4

Metabolic Glycan Labelling with Azido-Monosaccharides in Gut Microbiome Members

"Any colour you like"

The Dark Side of the Moon, Pink Floyd

Abstract

Glycans and glycan-based interactions shape the human gut microbiome. Tools to study these key ligands at a molecular-level are, however, currently insufficient albeit crucial to understanding their role in the human gut. Metabolic oligosaccharide engineering (MOE) is a technique that, if a bacterial salvage pathway for a monosaccharide is present, can explore the variety of glycans containing this monosaccharide on the cell surface of human microbiome members. To this end, we selected a set of monosaccharides that are found in bacterial mucin *O*-glycans and equipped these with an azido chemical reporter to enable bioorthogonal labelling with a fluorophore after MOE. Using these glycan-based probes, we hypothesised that, if present, the salvage pathways of key members of the human gut microbiome such as *Akkermansia muciniphila*, *Anaerostipes rhamnosivorans*, *Intestimonas butyriciproducens*, and *Eubacterium hallii* can salvage these glycan-based probes and built it into their glycans. Via MOE we could show the complexity of cell surface glycans in *A. muciniphila*, *A. rhamnosivorans*, *I. butyriciproducens*, and *E. hallii*. Furthermore, the application of MOE in *Escherichia coli* was studied extensively and this resulted in a protocol that enables the labelling of glycans in this model organism. As a proof of principle, MOE was also performed in co-cultures of key microbiota members to investigate the effect on efficiency of glycan salvage and labelling of surface glycans when co-culturing these species. Overall, this study provides proof-of-principle results on new MOE-based techniques and labels to explore the role of glycans in several key human microbiome members.

Manuscript in preparation: Tjerk J. Sminia, Hanne L.P. Tytgat, Nam T. P. Bui, Han Zuilhof, Willem M. de Vos, and Tom Wennekes

Introduction

The human gut microbiome is home to a variety of bacterial species and is vital to our well-being. Although we normally benefit from their presence, a microbial imbalance in our gut is often linked to diseases, but how these are linked is far from understood.¹ Since all cells are covered with glycans, they also play an important role in the functioning of these bacterial species. However, the fact that glycan synthesis is not template-driven brings an extra challenge to explore the complex biological role of these important biomolecules.² Many outstanding questions remain such as what role glycans play in the inter- and intraspecies interactions between bacteria and human cells.³ Nor is it fully understood why bacteria modify their proteins or lipids with glycans. The intestine is home to a heavily *O*-glycosylated mucus layer, covering the intestinal epithelial cell layer and thereby providing a protective barrier. Mucus consists of a variety of high-molecular-mass glycoproteins called mucins. The main glycans found in these mucins are *O*-glycans that contain *N*-acetylglucosamine, *N*-acetylgalactosamine, fucose and sialic acid.⁴ Mucins offer protection and are increasingly being implicated in host-gut microbiome communication processes, but also serve as a carbon and energy source to our gut bacteria.⁵ A chemistry-based approach to study these glycans in more detail is to label them with chemical reporters using metabolic oligosaccharide engineering (MOE). MOE has been developed in the last 15 years, and is a tool to discover and understand glycan processes in cells in greater detail (Chapter 2 and Figure 1).⁶ MOE operates via the principle that certain unnatural monosaccharide derivatives are often still recognised by the enzymatic machinery of cells and allows for them to be taken up into the cellular glycan salvage pathways and eventual incorporation into cellular glycans. Via this approach, monosaccharides bearing unnatural azido-acetyl groups, which do not interfere with biological processes, so called bio-orthogonal chemical reporters, can be successfully built into complex glycans.⁶

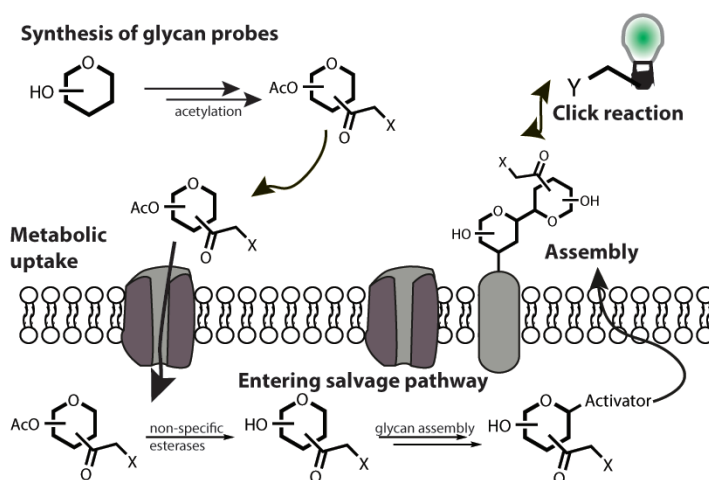


Figure 1. Schematic overview of metabolic oligosaccharide engineering.

A variety of bio-orthogonal groups are known nowadays, among which the azides, strained alkynes, and cyclopropenes are the most popular.⁶ When these bio-orthogonal groups are present on the glycans a click reaction can attach an additional molecule, often a fluorescent dye, to the glycans. Over the years, MOE has proven to be a powerful tool to directly image glycans in a variety of mammalian cell lines, bacteria, and plants.⁶⁻⁹

After successful synthesis of a chemical probe, MOE starts when the glycan-based probe is taken up by the human or bacterial cell. This can either be via active or passive transport. In the cytosol, non-specific esterases can hydrolyse any present masking group (often acetyl esters). Next, the glycan-based probe enters the salvage pathway and is activated by endogenous enzymes for use as a substrate of glycosyltransferases in further glycan assembly. After the assembly and localisation on the cell surface, the conjugation reaction (e.g. click reactions such as copper-catalysed azide alkyne cycloaddition (CuAAC) or strain-promoted azide alkyne cycloaddition (SPAAC) can happen. A more detailed overview of MOE, including examples and limitations of MOE, can be found in **chapter 2**.

The study of gut bacteria using MOE is not new. In 2015, Kasper and co-workers reported an extensive study of MOE with Ac₄GlcNAz, Ac₄GalNAz, and Ac₄ManNAz in classes of gut commensals such as *Bacteroides* and *Lactobacillus*.¹⁰ The probes used in that study include azido or alkyne derivatives of fucose, *N*-acetylglucosamine, or *N*-acetylgalactosamine.^{10-16,17} Kasper and co-workers expanded their approach in 2017 by specifically tagging three surface macromolecules, peptidoglycan, lipopolysaccharide and capsular polysaccharide, either simultaneously or individually in living anaerobic commensal bacteria.¹⁸ For PG labelling a fluorescent dye (HADA, see **chapter 1**) was installed on D-Ala, while for LPS and capsular polysaccharides, azide and cyclopropene chemical reporters were installed on *N*-acetylgalactosamine or KDO.

The glycan-based probes used in this MOE study were selected based on their presence as monosaccharides in mucin *O*-glycans or other carbohydrate-based carbon sources present in food (such as L-rhamnose in plants). In recent years, L-fucose has especially been identified as a glycan that exerts influence on the human gut and its microbiome. The overall goal was to assess if members of the human gut microbiota could salvage these monosaccharides from the environment and incorporate them in their own glycans, as was earlier shown in *Bacteroides fragilis*.¹⁹ The azido-glycan-based probes Ac₄GlcNAz **1**, Ac₄GalNAz **2**, Ac₄FucAz **3**, and Neu5Az **5** were synthesised following a protocol based on the work of Bertozzi and co-workers.⁵⁵ The synthesis of Ac₄RhamAz **4**, was adapted from Chen and co-workers.²⁰ To promote passive cellular uptake of the probes, acetylated azido-monosaccharides were used.

Metabolic Glycan Labelling in Gut Microbiome Members

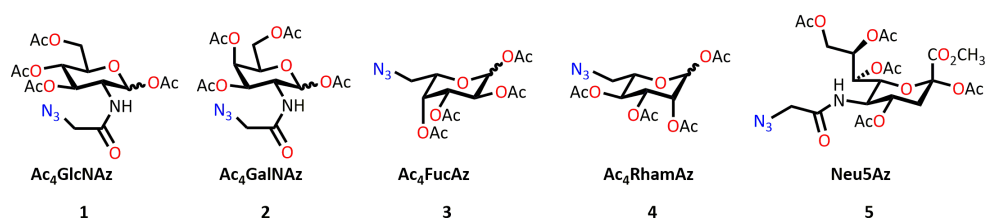


Figure 2. Overview of tested azido-monosaccharides.

As targets for these labelled monosaccharides, eight species present in the human gut core microbiota were selected in this study; *Bacteroides fragilis*, *Escherichia coli*, *Akkermansia muciniphila*, *Anaerostipes rhamnosivorans*, *Intestimonas butyriciproducens*,²¹ *Faecalibacterium prausnitzii*,²² *Ruminococcus bromii*²⁷ and *Eubacterium hallii*.²³ Below an overview of these bacteria is presented along with a brief explanation about their characteristics and specific reasons for their selection (Table 1).

Table 1. Overview of tested microbes in this study.

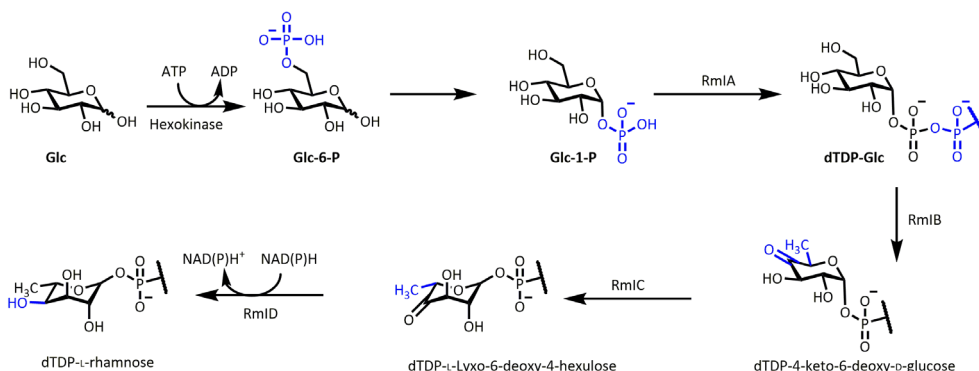
| Overview tested bacterial species | Metabolic property | REF |
|---|--|--------|
| <i>Akkermansia muciniphila</i> | Mucus degrader and propionate producer | 5 |
| <i>Eubacterium hallii</i> L2-7 | Glycan degrader and butyrate/vitB12 producer | 24 |
| <i>Faecalibacterium prausnitzii</i> | Glycan degrader and butyrate producer | 25 |
| <i>Intestimonas butyriciproducens</i> AF211 | Lysine/glucose degrader and butyrate/vitB12 producer | 21, 26 |
| <i>Ruminococcus bromii</i> | starch-degrading species | 27 |
| <i>Bacteroides fragilis</i> | β-lactamase producer | 28 |
| <i>Escherichia coli</i> | vitamin K2 producer | 29 |
| <i>Anaerostipes rhamnosivorans</i> 1Y-2 | rhamnose degrader and butyrate producer | 30 |

As a model system for the initial optimisation of the MOE protocol and imaging techniques, *Bacteroides fragilis* was chosen since MOE using Ac₄FucAl and Ac₄FucAz **3** is well established in this species.^{10, 12} *B. fragilis* is a Gram-negative obligate anaerobe that produces eight capsular polysaccharides (A-H) and many glycoproteins.³¹ The abundantly fucosylated polysaccharide A on the surface of *B. fragilis* is involved in immunomodulation and a mutant of *B. fragilis* ΔPSA does not express this polysaccharide.^{12, 19} As the latter is a nice control, both strains were selected as targets for MOE using Ac₄FucAz **3**.

The application of MOE has already been investigated extensively in *B. fragilis*, but it has never been reported in the widely used *E. coli* lab strain MG1655. This is quite striking given the ubiquitous use of *E. coli* MG1655 in the lab and the abundant availability of all kinds of mutants for this strain, also for glycan-related metabolic pathways. Aware of these possibilities and in part also curious about the reason for the lack of examples with MOE in this model organism, we also selected it for investigation. Furthermore, a variety of strictly anaerobic gut bacteria was selected. For example, the mucus

degrader *A. muciniphila* that was recently discovered to possess a fucose salvage pathway as shown by Ottman et al.,³² and was already extensively studied by us as reported in **chapter 3**. A recent study also showed the therapeutic potential of *A. muciniphila* towards obesity and diabetes.³³

Butyrate-producing Gram-positive bacteria in the gut are important for maintaining human health, are involved in key pathways for energy metabolism, and are key members of the intestinal food chain including cross feeding with many other gut members.³⁴ Glycans are considered as one of the main sources for the production of butyrate by these bacteria.³⁵ It is therefore crucial to identify all butyrate-producing species in the human gut and to study this process in more detail to understand the mechanisms underlying this production. We therefore selected *A. rhamnosivorans*, *I. butyriciproducens*, *F. prausnitzii*, and *E. hallii* according to their unique characteristics as model organisms in this process to further investigate (table 1).³⁶ *F. prausnitzii* is among the most abundant bacteria in the human gut of healthy adults and can amount up to more than 5% of the total bacterial population.²² The strictly anaerobic species *I. butyriciproducens* and *A. rhamnosivorans*³⁰ have been isolated only recently and cultivated from the human gut, and can therefore be studied in more detail with MOE and the set of glycan-based probes.



Scheme 1. Suggested biosynthesis of GlcNAc-6-P and dTDP-L-rhamnose from Glc.

A prerequisite to perform MOE in bacteria is the presence of a glycan salvage pathway.³⁷⁻³⁸ Salvage pathways of GlcNAc, GalNAc, fucose, and Neu5Ac have been studied using MOE and reported in various organisms, including bacteria.³⁹⁻⁴² Nevertheless, it is never proven that our selection of gut bacteria do have such a salvage pathway. Besides that, a rhamnose salvage pathway has never been identified in bacteria and the current belief is that all bacterial L-rhamnosides are derived from *de novo* biosynthesised L-rhamnose. To investigate if a salvage pathway might be present in anaerobic gut bacteria, the azido-rhamnose probe **4** was developed and included in our study. L-Rhamnoside biosynthesis occurs via a glycosyltransferase enzyme that adds L-rhamnose onto other bacterial

Metabolic Glycan Labelling in Gut Microbiome Members

glycoconjugates.⁴³ The substrate for these L-rhamnosyltransferase is dTDP-L-rhamnose that is biosynthesised in three steps from dTDP-glucose (Scheme 1). *A. rhamnosivorans* can make use of this biosynthesis pathway to synthesise dTDP-L-rhamnose and GlcNAc-6-P. The dTDP-L-rhamnose synthesis is unique for *A. rhamnosivorans* compared to other human gut bacteria. The azido-rhamnose probe **4** can therefore, if the corresponding dTDP donor can be biosynthesised from it in the cell, selectively label *A. rhamnosivorans* in the presence of other gut bacteria.

First, we aimed at direct imaging of specific monosaccharides in complex cell surface glycans in core human gut microbiota members using MOE as the labelling technique. Secondly, the capacity of the different core members of the human gut to salvage glycan-based probes based on common monosaccharide glycans was evaluated. Ultimately, as a proof of principle, we also investigated MOE as a tool to label specific gut microbiota bacteria in a complex mixture based on their glycome.

Results and Discussion

Metabolic Oligosaccharide Engineering in *Bacteroides fragilis*

MOE relies on salvage pathways to incorporate the probes but does not enable studying every step in the biosynthesis in great detail. Another potential drawback of MOE is that certain azido probes might be converted into other azido-glycans before ending up on the cell-surface. Last but not least, the efficiency of labelling is often low. This is mainly caused by the competition for incorporation into the salvage pathway between the azido-monosaccharides and naturally present monosaccharides. The latter are often also more easily accepted by the involved salvage enzymes as they represent their native substrate. Another reason for the low efficiency might be that intracellular glycans will also be labelled with the azide probes, as MOE is not selective for surface glycans (**chapter 2**). Despite these challenges MOE enabled new insights into the world of glycans of core members of the human gut.

MOE with an alkyne or azido fucose has already been successfully applied in *Bacteroides fragilis*.^{10, 12} Inspired by this, we used the azido-L-fucose probe **3** as a positive control.⁴⁴ Azido chemical reporters are more flexible in the subsequent bioconjugation.⁶ For example, the copper-free click chemistry can be used.⁴⁵ Although, previous experiments showed that the fucose-salvage pathway in zebrafish is not able to process this azido-L-fucose probe **3**, MOE experiments in plants and bacteria did show salvage and incorporation.^{10, 44}

For live-cell imaging the use of SPAAC labelling is preferred over CuAAC labelling since no cytotoxic copper is needed.⁴⁵ The commercially available DBCO dyes linked to Cy-3 or Cy-5 were used for the SPAAC labelling. The DBCO Cy-3 or DBCO Cy-5 SPAAC reactions were chosen as the bioconjugation reaction of choice since these reagents are commercially available. To compare our data with the existing protocols, CuAAC labelling was initially performed in *B. fragilis*. The azido-fucose probe **3** was added to a freshly inoculated culture and incubated overnight. Labelling was evaluated using

confocal microscopy after a 30 minute CuAAC reaction with an Alexa Fluor 488 alkyne conjugate and subsequent additives (Figure 3). An optimum was found after 2 hours incubation, while after 3.5 hours the fluorescent signal was no longer observed, probably due to excessive metabolism. For further MOE experiments in *B. fragilis* we therefore chose this 2 hour incubation time as the optimum. Other acetylated glycan-based probes (Figure 2) such as *N*-azidoacetylglucosamine (Ac_4GlcNAz , **1**), *N*-azidoacetylgalactosamine (Ac_4GalNAz , **2**), or 5-azido neuraminic acid (Neu5Az , **5**) have been successfully tested in *B. fragilis* by others.¹⁰

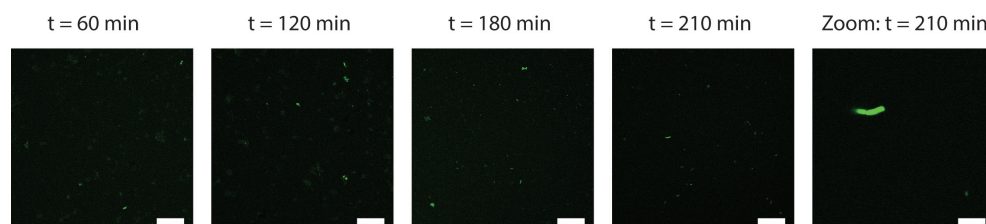


Figure 3. Metabolic oligosaccharide engineering in *Bacteroides fragilis* with azido-L-fucose **3**. Labelling was performed with CuAAC using Alexa Fluor 488 alkyne. Scale bar = 25 μm and for zoom 5 μm .

B. fragilis can express a variety of capsular polysaccharides among which the fucosylated polysaccharide A (PSA) is the most abundant³¹ and as mentioned a ΔPSA mutant was available which lacks this polysaccharide.³¹ Metabolic labelling with the azido-fucose probe **3** in this mutant should therefore be lower compared to the wild type *B. fragilis*. This was indeed observed in our experiments when both species were treated with Ac_4FucAz **3** and provides evidence that indeed fucosylated *B. fragilis* cell surface glycans are being labelled through salvage of **3** (Figure 4).

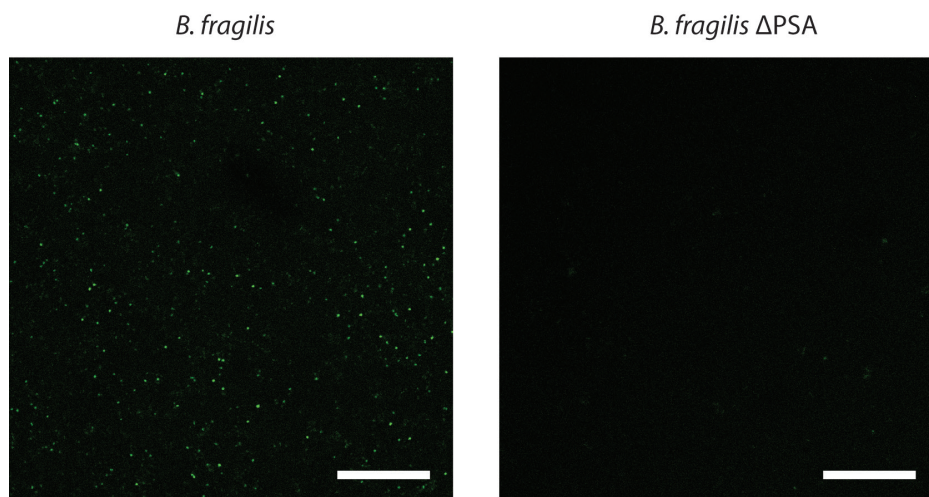


Figure 4. Metabolic oligosaccharide engineering in *Bacteroides fragilis* and *Bacteroides fragilis* Δ PSA with Ac_4FucAz **3**. Labelling was performed with CuAAC using Alexa Fluor 488 alkyne. Scale bar = 25 μm .

Next, we evaluated SPAAC labelling with a DBCO Cy-3 dye on a dense overnight *B. fragilis* culture containing Ac_4FucAz probe **3**. This resulted in fluorescent bacteria (Figure 5) and in the control culture without **3** SPAAC labelling with DBCO Cy-3 dye did not result in any fluorescent signal after washing, and the amount of a-specific labelling was thus comparable to that observed while using the CuAAC control reaction.

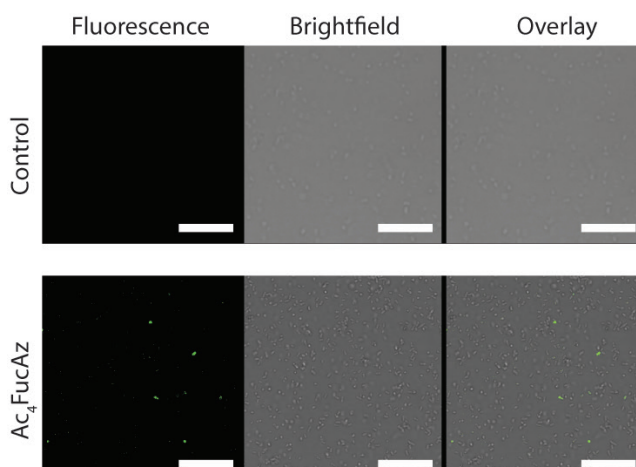


Figure 5. *Bacteroides fragilis* with Ac_4FucAz **3** and SPAAC labelling with DBCO Cy-3 dye. Scale bar = 5 μm .

Metabolic Oligosaccharide Engineering in *Escherichia coli*

The positive outcome of this initial study in *B. fragilis* prompted us to investigate the application of MOE in *Escherichia coli* as the starting point, the wild-type *E. coli* MG1655, the common *E. coli* lab strain, was selected together with the *N*-acetylglucosamine-based chemical probe, Ac₄GlcNAz **1**. To evaluate if the LB medium, which is known to cause autofluorescence, might interfere with the fluorescent signal, the M9 minimal medium was also tested. No effects were observed between the two media, and therefore we decided to use LB in all our experiments since *E. coli* grows faster in LB compared to in M9. The standard protocol for MOE involves overnight incubation with the chemical probe. However, application of this protocol on *E. coli* did not result in observable labelling after a CuAAC reaction with an Alexa Fluor alkyne dye. We speculated that one of the factors that could cause this negative result is *E. coli*'s short generation time of roughly 20 minutes. This could lead to successfully labelled glycans being "diluted" over the ever-expanding *E. coli* population in the sample and subsequent lack of fluorescent signal. An additional process that might contribute to this negative result is that the relatively high metabolic rate of wild-type *E. coli* MG1655 probably resulted in increased chemical probe catabolism instead of salvage. Furthermore, a comparatively low level of non-specific esterase activity has been reported for *E. coli*.⁴⁶ This might result in incomplete intracellular unmasking of the acetyl esters on the probe, which are there to facilitate passive probe uptake. To investigate the latter we conducted an experiment in which non-acetylated GlcNAz was incubated with *E. coli*, but still no labelling was observed. This could be caused by either insufficient uptake of the probe or because of the high turnover rate and metabolic rates. To circumvent the fast metabolism in *E. coli* we decided to carry out an experiment where the probe was dosed into the culture medium at multiple times at 30 minute intervals during the incubation and the culture was also assessed for labelling after every 30 minute interval (Figure 6). This resulted in labelled glycans on *E. coli* indicating that uptake of the probe is fine. After 180 minutes incubation with the probe and three cycles of dosing additional probe the fluorescence signal did start to significantly diminish (data after 180 min not shown).

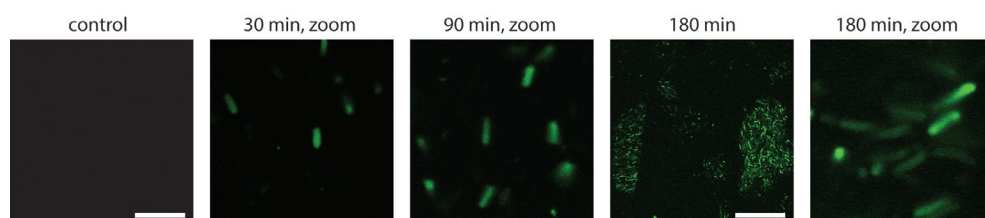
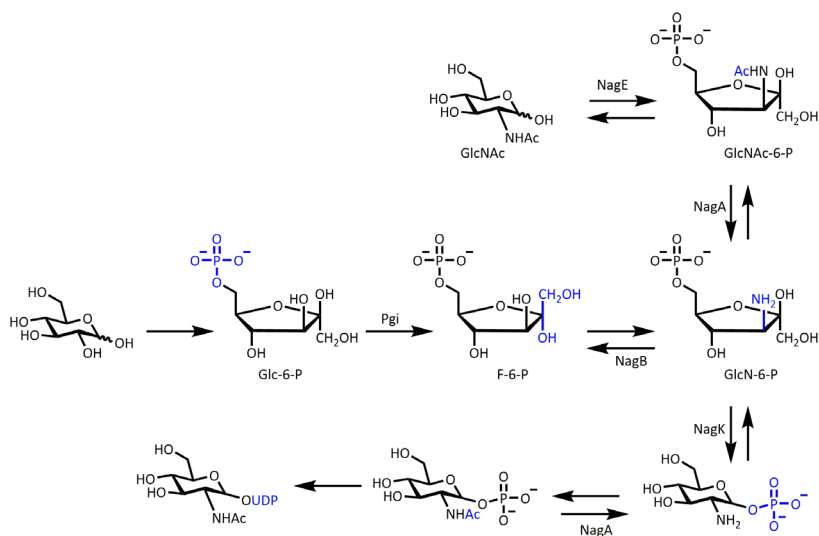


Figure 6. MOE in *Escherichia coli* with Ac₄GlcNAz **1** (100 μ M final concentration). Every 30 minutes additional Ac₄GlcNAz was added. Alkyne Alexa Fluor 488 (5 μ l of 1 mM stock) was used for CuAAC reaction. Scale bar = 25 μ m and for zoom 5 μ m.

Metabolic Glycan Labelling in Gut Microbiome Members

Although the circumstances that achieved this labelling were artificial they did for the first time reveal that MOE is, in principle, possible in *E. coli* MG1655. Next, we studied the individual chemical pathways in *E. coli* MG1655 that are involved in GlcNAc metabolism and salvage. The Keio collection, a single-gene deletion collection of all nonessential genes in *E. coli* K-12 was selected to further investigate these individual chemical pathways that might influence probe metabolism.⁴⁷ From this collection we selected the following mutants in which enzymes involved in GlcNAc-UDP biosynthesis are affected (Scheme 2): NagK (*N*-acetyl-D-glucosamine kinase), NagE (*N*-acetyl D-glucosamine PTS permease), NagA (*N*-acetyl glucosamine 6 P deacetylase), NagB (D-glucosamine 6 P deaminase) and Pgi (phosphoglucose isomerase). Next to these mutants we also re-evaluated MOE in *E. coli* MG1655 itself, but with NaN₃ present during the incubation to arrest its overall metabolic processes.⁴⁸



Scheme 2. Biosynthesis of UDP-*N*-acetylglucosamine.

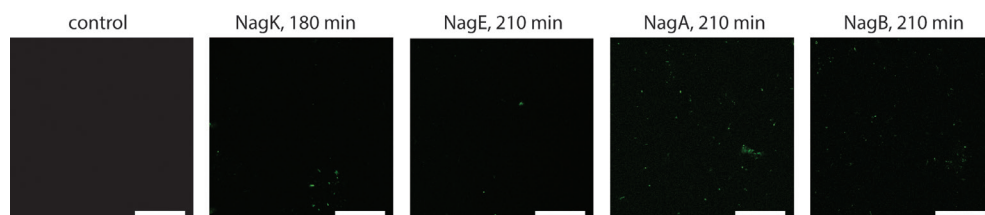


Figure 7. Metabolic oligosaccharide engineering in selected *Escherichia coli* Keio collection mutants with Ac₄GlcNAz 1 (100 μM final concentration). Alexa Fluor 488 Alkyne (5 μl of 1 mM stock) was used for CuAAC reaction. Scale bar = 25 μm.

The incubation time for the Ac₄GlcNAz **1** probe in this experiment had an optimum around 210 minutes for the NagA *E. coli* mutant (Figure 7). Longer incubation times showed less labelling or no labelling at all. For the NagK mutant the optimum was found around 180 minutes. Both NagE and NagB mutants did not result in any labelling at all. NagA catalyses the hydrolysis of the *N*-acetyl group, since labelling was observed, the *N*-azidoacetyl group was not affected by this hydrolysis. For NagK, which catalyses the phosphorylation of *N*-acetyl-D-glucosamine derived from cell-wall degradation, the product of this reaction is a substrate for NagA (GlcNAc-6-P). The other single-gene deletions are involved in other cellular processes such as transport inside the cell (NagE) or reversible isomerisation-deamination towards fructose (Pgi and NagB). NagE is not necessary for *N*-acetylated glycans while the latter gene NagB might not be active with the azido group present on *N*-acetylglucosamine. The last single-gene deletion we tested was Pgi (phosphoglucose isomerase). This gene is a major player in the isomerisation of 6-phosphate glucose. We did observe labelling but less compared to the above described single-gene deletion mutants.

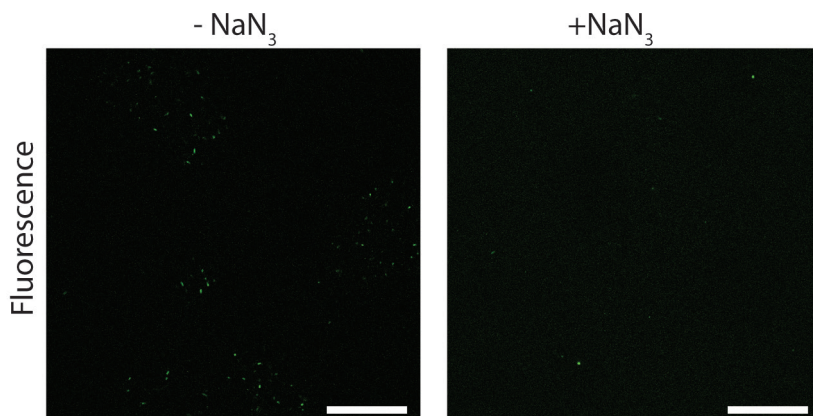


Figure 8. MOE in *Escherichia coli* with (right) or without (left) addition of NaN₃ with Ac₄GlcNAz **1** (100 μM final concentration). Alexa Fluor 488 Alkyne (5 μl of 1 mM stock) was used for CuAAC reaction. Scale bar = 25 μm.

The addition of sodium azide resulted in no labelling at all (Figure 8). The cytotoxic effect of sodium azide on biochemical processes, e.g. the inhibition of cytochrome oxidase in Gram-negative bacteria, is responsible for this.⁴⁹ The above-described study illustrates the challenges for MOE in *E. coli*. Others have also looked into MOE in *E. coli* but were never clear in their findings nor reported detailed findings to illustrate their hypothesis.^{9, 50} This study shows, for the first time, that MOE in *E. coli* can be achieved, albeit in an artificial way. In order to further elucidate the metabolic pathways in *E. coli* a more detailed study with the Keio collection might be needed but with the current toolbox not possible. First of all, other chemical reporters (e.g. cyclopropenes or isonitriles) can have a different effect on the single-gene deletion mutants but are synthetically difficult targets. Second, the study where we dosed additional azido-glycan probe showed labelling but is time consuming and needs

the availability of large amounts of chemical probe. This is therefore not suitable for further studies. Moreover, this system was highly artificial and therefore not directly related to the biological system. ECOM4 (*Escherichia coli* cytochrome oxidase mutant 4), a mutant of *E. coli* which lacks the cytochrome, might be interesting to further investigate.⁵¹⁻⁵² The metabolism of this species is hampered and thus could positively influence MOE in *E. coli*. Of course, this system remains artificial and therefore cannot resemble the real *in vivo* situation.

Preliminary evidence for rhamnose salvage in *A. rhamnosivorans*

After successfully achieving MOE in *B. fragilis* and *E. coli*, our investigation was continued in anaerobic human gut bacteria with our set of glycan-based probes. We first explored MOE in mono-cultures of *A. rhamnosivorans* and *A. muciniphila* with rhamnose probe **4** and fucose probe **3** (Figure 9). Control experiment showed no α -specific labelling with DBCO Cy-3 or Cy-5 dye in both cultures. When the azido rhamnose probe **4** was incubated with either of these bacteria, confocal microscopy showed fluorescence in *A. rhamnosivorans* but not in *A. muciniphila*. The azido fucose probe **3** only labelled glycoconjugates in *A. muciniphila*. This result provides preliminary evidence for the presence of a rhamnose salvage pathway in *A. rhamnosivorans* and a fucose salvage pathway in *A. muciniphila*. For *A. rhamnosivorans* this information sums up the tolerance towards rhamnose since, previous reported data showed that this strain is capable to ferment rhamnose which is unique in *Anaerostipes* genus.³⁰ The fucose salvage pathway in *A. muciniphila* was reported earlier by us when we performed peptidoglycan and glycoconjugates dual labelling in *A. muciniphila*, and was backed up by a metabolic model describing *A. muciniphila*. (**chapter 3**).

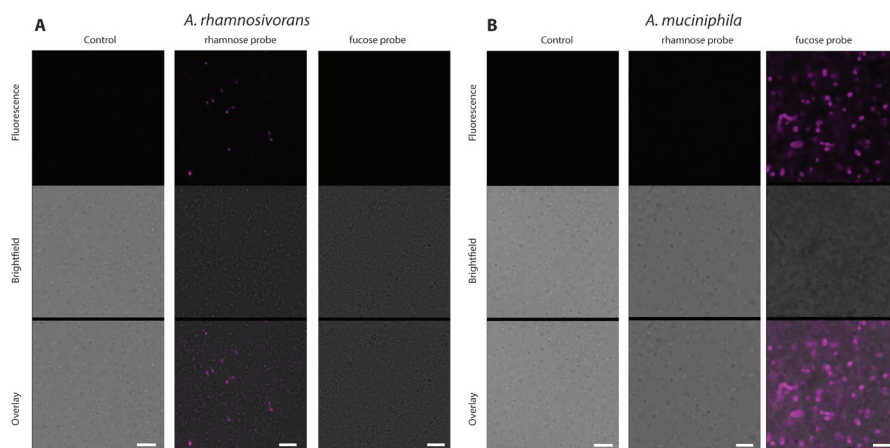


Figure 9. Metabolic oligosaccharide engineering in *A. rhamnosivorans* (panel A) and *Akkermansia muciniphila* (panel B) with control (left), Ac_4RhamAz probe **4** (middle), and Ac_4FucAz probe **3** (right) clicked with DBCO Cy-5 dye. Scale bar = 5 μm .

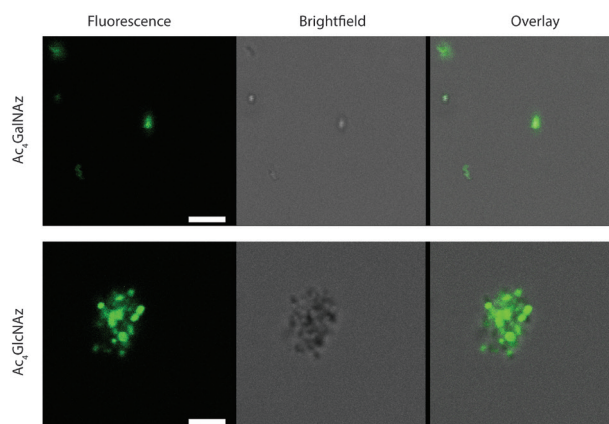


Figure 10. Metabolic oligosaccharide engineering in *Akkermansia muciniphila* with Ac_4GalNAz **2** (top) and Ac_4GlcNAz **1** (bottom) clicked with DBCO cy5. Scale bar = 5 μm .

Labelling of other glycans in the cell envelope of *A. muciniphila* was also investigated using our other probes. *N*-azidoacetylglucosamine **1** and *N*-azidoacetylgalactosamine **2** both showed successful labelling in *A. muciniphila*. According to confocal microscopy images, Ac_4FucAz **3** was more efficiently salvaged compared to Ac_4GlcNAz **1** and Ac_4GalNAz **2**. These experiments were in line with our results obtained in **chapter 3**. An exciting direction for future studies is for these *in vitro* experiments to be translated to *in vivo* experiments. In line with the work of Kasper and co-workers *A. muciniphila* and *A. rhamnosivorans* labelled bacteria can be administered to mice via oral or direct injection into the gut for further elucidation of labelled glycans *in vivo*.¹⁸

Quantification of fluorescent-labelled cells in *E. hallii*, *I. butyriciproducens*, *A. rhamnosivorans*, and *A. muciniphila*

To quantify the labelling efficiency of MOE in members of the gut microbiota, we performed flow cytometry on MOE labelled bacteria. All bacterial cells were stained with Syto-9 before flow cytometry analysis. By using this method any aspecific dye aggregates and non-DNA-containing vesicles (e.g. lysed bacteria) could be excluded from the analysis. The Syto-9 labelled cell population in flow cytometry could then be further analysed. The APC channel detector displays the Cy-5 fluorescence (horizontal axis), while the side scatter is correlated to the particle size (vertical axis). Cells that are stained with Syto-9 but are not SPAAC labelled with a dye via MOE are visible in quadrant 1 (Q1). In Q3 and Q4 the cells or cellular fragments smaller than the control cells can be observed. In general no cell population was found in these quadrants. Upon SPAAC labelling with a Cy-5 dye after MOE the cell population shifts to Q2 and thus showed increased fluorescence while obtaining the standard bacterial size. Cells that are stained with Syto-9 but are not SPAAC labelled with a dye via MOE are visible in quadrant 1 (Q1). In Q3 and Q4 the cells or cellular fragments smaller than the control cells can be observed. In general no cell population was found in these quadrants. Upon SPAAC labelling with a Cy-5 dye after MOE the cell population shifts to Q2 and thus showed increased fluorescence while obtaining the standard bacterial size.

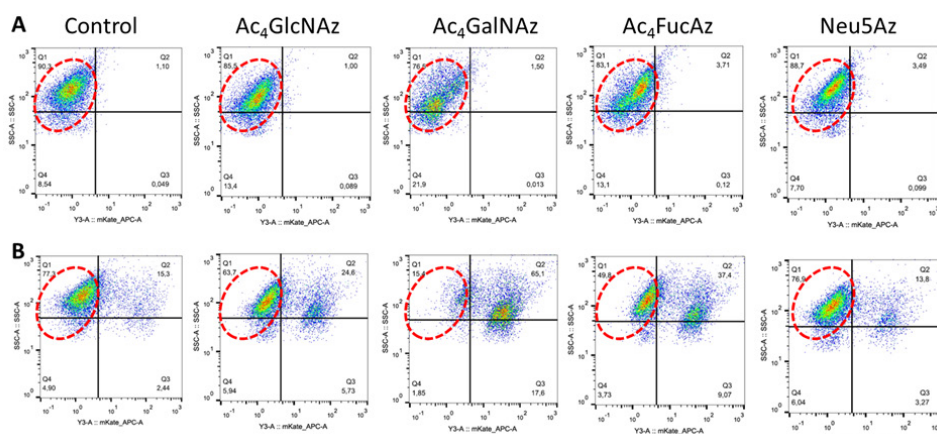


Figure 11. Flow cytometry results for MOE in *Eubacterium hallii* with glycan-based probes. SPAAC with DBCO Cy-5 after 24 hours incubation with probe. A = without DBCO Cy-5 dye. B = with DBCO Cy-5 dye. The major cell population in the control (A) is circled with a red dashed line to assist visual assessment of changes.

The generation time in *E. hallii* is rather short (20 min) compared to other human-gut members included in this study. The metabolic activity in *E. coli*, which resulted in degradation of the probe, was however not seen in *E. hallii* (Figure 11). When the DBCO fluorescent dye was added to the control sample a small population was labelled. This background labelling amounted to ~15% of the total population. The sample incubated with Ac₄GlcNAz **1** (25% labelled cells) or Ac₄FucAz **3** (37% labelled cells) probe showed a small increase in fluorescently labelled cells compared to the control sample. In contrast, a clear shift in population was seen for Ac₄GalNAz **2**. In total 65% of the population was fluorescently labelled. The sample with Neu5Az probe **5** did not show any labelling above the background signal. An explanation for this might be the absence of neuraminic acid in *E. hallii*.

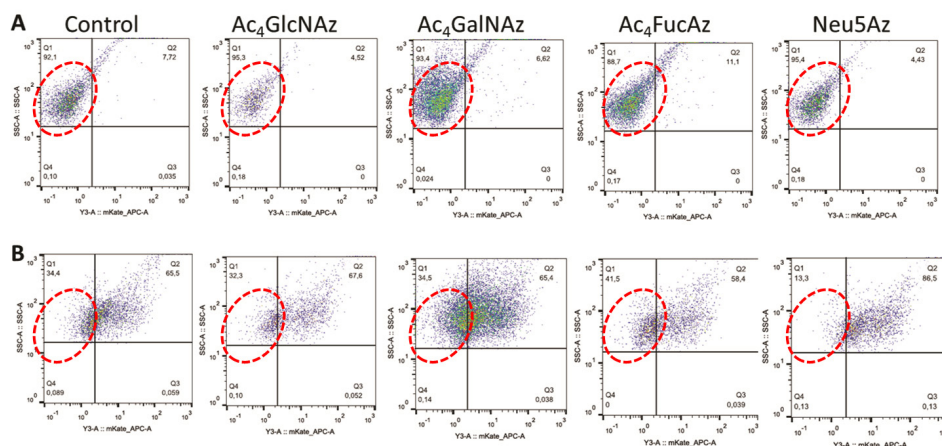


Figure 12. Flow cytometry result for MOE in *Intestimonas butyriciproducens* with glycan-based probes. Click reaction with DBCO Cy-5 performed after 24 hours incubation. A = without DBCO Cy-5 dye. B = with DBCO Cy-5 dye. The major cell population in the control (A) is circled with a red dashed line to assist visual assessment of changes.

Next, *Intestimonas butyriciproducens* was incubated with the chemical probes. A diminished growth of *I. butyriciproducens* was observed in all samples except Ac₄GalNAz (Figure 12). When the click reaction with the fluorescent dye was performed a high background labelling was observed in the control sample. The shift in the observed positive samples was comparable to the shift observed in the control, so no significant incorporation of the probes was observed. Therefore, it seems that *I. butyriciproducens* cannot salvage the Ac₄GlcNAz **1** (68% labelled cells), Ac₄GalNAz **2** (65% labelled cells) and, Ac₄FucAz **3** (58% labelled cells) probes. After the addition of Neu5Az **5** probe 86% of the cells were fluorescent labelled, which is significantly higher than the background level of the control (66% labelled cells). From this set of experiments the high background labelling, even in the absence of azido-glycans needs to be explored further. For example, it is not clear if the fluorescent dye reacts with *I. butyriciproducens*. For future studies a richer medium or a higher sample loading would provide a way to increase the cell population. The drawback might, however, be the presence of now competing sugars in the medium, which may result in a diminished uptake of the probe. Another possibility can be to add an azido- or alkyne-functionalised D-glucose derivative to the medium, with the hypothesis that it would be metabolised, since D-glucose is an important nutrient for *I. butyriciproducens*.

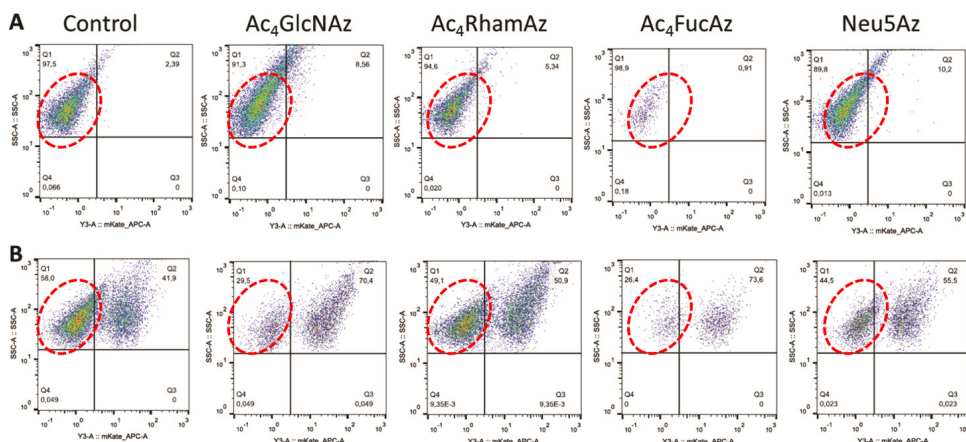


Figure 13. Flow cytometry result for MOE in *Anaerostipes rhamnosivorans* with glycan-based probes. Click reaction with DBCO Cy-5 performed after 24 hours incubation. A = without DBCO Cy-5 dye. B = with DBCO Cy-5 dye. The major cell population in the control (A) is circled with a red dashed line to assist visual assessment of changes.

The azido rhamnose probe **4** was successfully tested in *A. rhamnosivorans* as showed by confocal microscopy (*vide supra*). Next, we tested Ac₄GlcNAz **1**, Ac₄FucAz **3**, and Neu5Az **5**. A high background was observed for all experiments in *A. rhamnosivorans* (Figure 13). The dense cultures, except Ac₄FucAz **3**, showed up top 40% α-specific fluorescent labelling in the control sample. The culture containing Ac₄FucAz **3** showed a diminished growth since flow cytometry did not even count 10.000 events. Toxic effects of the probe, poor growth, and/or the absence of important carbon sources might be the cause. Despite the poor growth, labelling was observed rising up to 74% for Ac₄FucAz **3**. Also Ac₄GlcNAz **1** (70% labelled cells), Ac₄RhamAz **4** (51% labelled cells), and Neu5Az **5** (56% labelled cells) showed labelling with the fluorophore above the background signal. These results provide preliminary evidence for the presence of a rhamnose salvage pathway in *A. rhamnosivorans*. Future studies are still needed, in order to show that the rhamnose probe **4** is salvaged and not catabolised into another metabolite. For example, isolation of the rhamnose-labelled glycans and detailed mass spectrometry may reveal the presence of 6-azidorhamnose on the cell surface glycans.

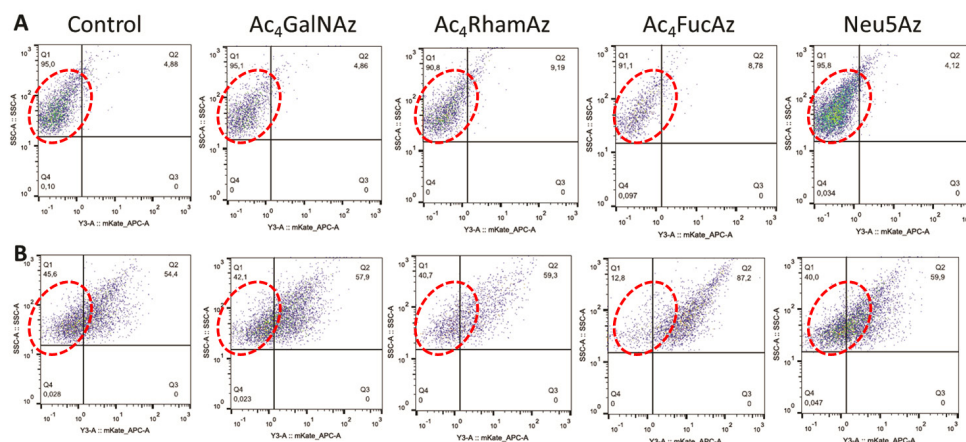


Figure 14. Flow cytometry result for MOE in *Akkermansia muciniphila* with glycan-based probes. Click reaction with DBCO Cy-5 performed after 24 hours incubation. A = without DBCO Cy-5 dye. B = with DBCO Cy-5 dye. The major cell population in the control (A) is circled with a red dashed line to assist visual assessment of changes.

In *A. muciniphila* the growth was also poor when compared to previous experiments (Figure 14). Despite the high background labelling the sample containing Ac₄FucAz probe **3** showed a clear shift in population. Interestingly, fucose labelling in *A. muciniphila* showed aggregation of bacteria as seen by confocal microscopy (Figure 9). Due to time restrictions, the corresponding flow cytometry analysis were not used to further investigate this phenomenon. Aggregates might be the result of fucose lectin present on the surface of *A. muciniphila* since the formation of aggregates was solely seen in *A. muciniphila* with the Ac₄FucAz probe **3**.⁵³ Alternative washing methods such as with added Tween®, or use of more hydrophilic fluorescent dyes such as sulfo-DBCO-Cy-5 should be evaluated in future experiments to see the effect on aggregate formation. Many reported MOE protocols use different washing procedures without clear statements or evidence for their effectiveness. For the other azide probes no clear effect was seen when it was compared to the control sample. Previous attempts showed metabolic labelling with an Ac₄GalNAz probe **2** (Figure 10), but this was not reproduced in this series of experiments. Metabolic labelling remains a challenging technique, and further optimisation in *A. muciniphila* is therefore needed to make conclusions about labelling with Ac₄GalNAz **2** or Ac₄GlcNAz **1**.

Other members of the human-gut microbiome such as *F. prausnitzii* and *R. bromii* were also of high interest for our metabolic labelling studies. *F. prausnitzii* and *R. bromii* are, however, known to be oxygen-sensitive and therefore could not handle the amount of oxygen introduced upon addition of the azide probe, resulting in diminished or no growth, respectively.^{27, 54} To prevent oxygen from entering, further experiments should add the probe in an oxygen-free solution by flushing it with an inert gas. Another factor worth investigating is the solvent (DMSO) in which the probe is dissolved which influences growth negatively. Further optimisation of both growing conditions and medium are also necessary to study glycoconjugates in *F. prausnitzii* and *R. bromii*.

Quantification of fluorescent-labelled cells in co-cultures of *I. butyriciproducens* and *A. muciniphila*, and *A. muciniphila* and *A. rhamnosivorans*

The human gut microbiome consists of a wide variety of bacteria. MOE can, in principle, enable selective labelling in bacteria if a salvage pathway is absent in one species or if the probe is not metabolised / taken up by a specific species in the presence of another that does. In **chapter 2** techniques are discussed which might explore the salvage pathway in more detail in the near future. In this series of experiments we tested MOE in co-cultures using different glycan-based probes. Our aim was to label selective one species over the other in these co-cultures. As a start, we performed MOE in co-cultures of *A. muciniphila* in combination with either *I. butyriciproducens* or *A. rhamnosivorans*. This selection was based on the distinct morphology of the bacteria (Figure 15), which enabled visual confirmation if we indeed could label different bacterial species using MOE.

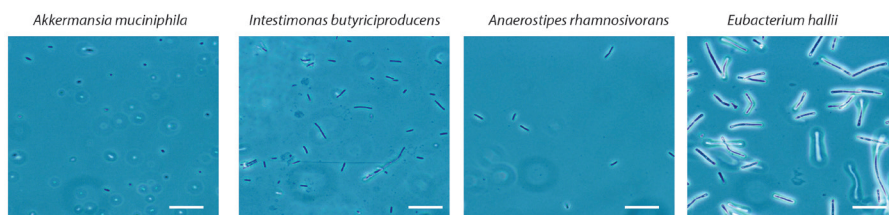


Figure 15. Microscope images of bacterial species investigated in this study. Scale bar = 10 μ m.

For the co-culture studies the different generation times of the species were taken into account. Despite this consideration the distribution of *A. muciniphila* and *I. butyriciproducens* was not equal (Figure 16D). Future experiments should focus on how to achieve an equal distribution of the co-culture. In the co-culture of *I. butyriciproducens* and *A. muciniphila* the Ac₄FucAz probe **3** was added. With the negative labelling result in the mono-culture for *I. butyriciproducens* and the positive result in the mono-culture of *A. muciniphila* we expected to selectively label *A. muciniphila* and not *I. butyriciproducens*. In the control samples prior to fluorescent labelling, the cell population (encircled with red dashed line) is seen in Q1 (Cy-5 channel plotted against side scatter; Figure 16A). When the DBCO dye was added a ~30% shift of the population to Q2 was observed (Figure 16B). In the overlay plot it is clearly seen that the whole cell population is shifted after fluorescent labelling (Figure 16C), thereby indicating that successful labelling is achieved. However, the high amount of a-specific labelling, observed in the mono-culture of *I. butyriciproducens* (*vide supra*), prevents a clear interpretation of these co-culture results. For that reason, further experiments are needed. As a start, the mono-culture of *I. butyriciproducens* needs to be repeated with different glycan-based probes and fluorescent dyes to decrease the a-specific labelling.

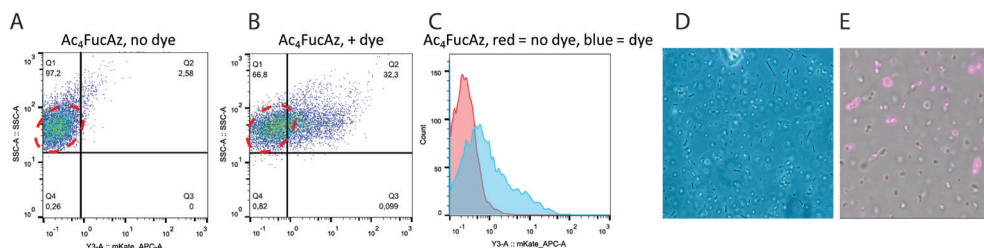


Figure 16. Flow cytometry results (A, B, and C) of co-culture 1 of *Akkermansia muciniphila* and *Intestimonas butyriciproducens* and microscopy images of culture (D) and confocal microscopy of fluorescent labelled *Akkermansia muciniphila* cells (E). The major cell population in the control (A) is circled with a red dashed line to assist visual assessment of changes. Scale bar is 5 μ m.

The second co-culture we investigated consisted of *A. muciniphila* and *A. rhamnosivorans* with Ac₄RhamAz **4** and DBCO Cy-5 dye (Figure 17). In both mono-cultures a shift in the population with Ac₄RhamAz **4** was observed. The initial results showed that rhamnose probe **4** is only incorporated in *A. rhamnosivorans*, and thereby suggest the presence of a rhamnose salvage pathway in *A. rhamnosivorans*.

The population shift in *A. muciniphila* can be considered as a-specific labelling since *A. muciniphila* cannot metabolise L-rhamnose. In the co-culture study the cell population before fluorescent labelling but with the rhamnose probe **4** fitted in Q1 (Cy-5 channel plotted against side scatter) (Figure 17A). When the fluorescent DBCO dye was added 20% of the cell population shifted towards Q2 (Figure 17B) and can thus be considered fluorescent labelled with the Cy-5 dye. Since only Ac₄RhamAz **4** can react with the fluorescent dye a first conclusion is that the cells found in Q2 are Ac₄RhamAz **4** labelled *A. rhamnosivorans*. Confocal microscopy of mono-cultures *A. rhamnosivorans* and *A. muciniphila* proved that only in *A. rhamnosivorans* fluorescent bacteria were obtained (Figure 17E). Future studies are however necessary to further investigate and validate this.

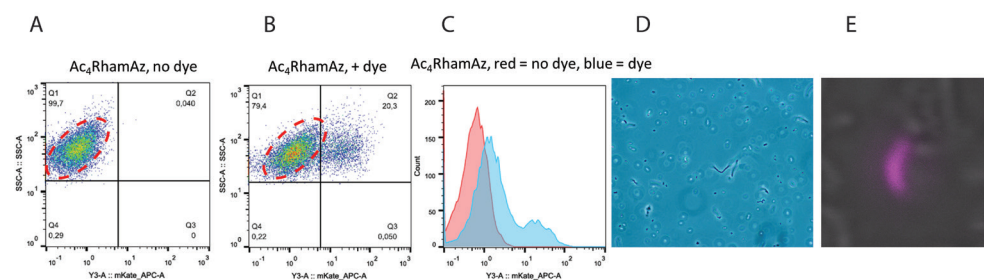


Figure 17. Flow cytometry results (A, B, and C) and microscopy images (D and E) of co-culture 2 of *Akkermansia muciniphila* and *Anaerostipes rhamnosivorans* and microscopy images of culture (D) and confocal microscopy of fluorescent labelled *Anaerostipes rhamnosivorans* cells (E). The major cell population in the control (A) is circled with a red dashed line to assist visual assessment of changes. Scale bar is 5 μ m.

Conclusion and Outlook

We here present preliminary results that give strong evidence that metabolic oligosaccharide engineering (MOE) might become a highly useful tool to label and directly detect glycans in human gut bacteria. MOE has the potential to further explore members of the complex microbiota. The results shown in this chapter have added new insights into how to apply MOE on human gut microbiota. The azido-fucose probe was used to successfully label fucosylated cell surface glycans in *B. fragilis*. This azido-fucose probe enables SPAAC labelling besides CuAAC labelling. SPAAC labelling is important for live cell imaging, something which was not possible with the alkyne-fucose probe.

After this successful labelling we focussed on MOE in *E. coli* MG1655, which has never been reported on this strain with glycan-based probes. Labelling in *E. coli* was possible when the culture was dosed every 30 minutes with additional glycan probe to cope with the fast metabolism and dilution of probe due to the short generation time in *E. coli*. With the current MOE toolbox this artificial system cannot be studied in more detail.

Our next target were human gut-related species. A variety of azido-monosaccharides were tested in mono-cultures of *E. hallii*, *I. butyriciproducens*, *A. rhamnosivorans*, and *A. muciniphila*. A-specific labelling, and thus a high background signal, was observed in almost all cultures. Future experiments should focus on diminishing this high background labelling. Other washing procedures with for example Tween® can be performed or the usage of other more hydrophilic dyes. The tested azido-probes were not salvaged by *I. butyriciproducens*. Optimisation of the medium, incubation time, and time of addition of the azido-probes can be performed in future studies to further improve the obtained labelling in *E. hallii* and *A. muciniphila*. All glycan probes, except Neu5Az **5**, showed fluorescent labelling, as backed up by flow cytometry. It seems however logical that Neu5Az **5** gave no labelling in these and other species. In mammalian cell lines, which are known to have more neuraminic acid compared to bacteria, labelling with Ac₄ManNAz results in azido-labelled neuraminic acid and is a well-established method nowadays.

To facilitate a faster metabolic uptake and easier synthesis the glycan probes were acetylated. The acetyl groups need to be removed inside the cytosol by non-specific esterases in order to be effective. One should take into consideration however, the enzyme activity of these esterases which might be diminished or not active at the time point the probes are added. If so, other enzymes can interfere in the process and thereby metabolise the glycan probe as a pure carbon source. Future studies should investigate the enzyme activity by using a set of non-acetylated and acetylated probes.

The nutrients inside the medium are important for a good growth of the bacteria. For example, *A. rhamnosivorans* can either grow on rhamnose or glucose in bicarbonate buffered medium. Future studies can investigate if *A. rhamnosivorans* can grow on pure rhamnose probe **4**. The same idea can be performed for *A. muciniphila* with azide-labelled mucin or Ac₄GlcNAz probe **1**. For the study in co-cultures additional experiments are needed. More control experiments (e.g. more counts in flow cytometry of both bacteria) and parameters (e.g. use of other dyes) need to be optimised to make decisive conclusions.

Metabolic oligosaccharide engineering provides a fast method to label and track glycans in human-gut bacteria. A remaining challenge is to characterise the glycans in more detail. One method to study glycans or glycoproteins in more detail is enrichment. After probe incubation, the azido-glycans can be conjugated with biotin or digoxigenin (DIG). In the next step, selective capturing of this label with streptavidin or anti-DIG results in a conjugated glycoprotein which can be separated on gel or blot. In a final stage glycoprotein staining, such as PAS staining, can be performed to visualise glycoproteins and subsequent downstream analysis using immunoblots and mass spectrometry. In **chapter 6** further recommendations about these techniques are given. In the near future more optimisation is, however, needed to make the labelling more selective and efficient. If this can be optimised, a full labelling and tracking of human-gut bacteria can be explored and mapped in full detail.

Experimental

1. Chemical procedures

General considerations and materials. All moisture-sensitive reactions were carried out under an argon atmosphere, using oven-dried glassware overnight, which was allowed to cool in a desiccator over anhydrous KOH, unless otherwise stated. Solvents were either purified over aluminium oxide under argon using a Pure Solv 400 solvent purification system (Innovative Technology, Amesbury, USA) (CH_2Cl_2 , Et_2O , toluene, and THF) or dried over 3Å molecular sieves (CH_3CN , CH_3OH , Pyridine). Commercially acquired chemicals were used without further purification unless stated otherwise. Alexa Fluor 488 alkyne, DBCO-Cy3, and DBCO-Cy5 were purchased from Sigma-Aldrich. All chemicals used for biological experiments were dissolved in DMSO (1 mM stock solution) and filter-sterilised prior to use.

Construction of positive control. Azide magnetic beads (Clickchemistrytools) with Cy-3/Cy-5 DBCO dye were used as a positive control. The beads were washed 2 times with PBS on a magnetic stand. Click reagents were added (CuAAC: dye, CuSO_4 and ascorbic acid in 500 μl PBS, SPAAC: dye in 500 μl PBS) and incubated for 1 hour at room temperature. The beads were washed 3 times with PBS. The samples were mounted on microscope slides and analysed by confocal microscopy.

Initially we tested the CuAAC reaction using commercially available CuAAC reagents. The same procedures as in **chapter 3** were applied in the gut microbiota species. Later, like in **chapter 3** we switched to SPAAC as our labelling method of choice. The CuAAC and SPAAC labelling reagents were evaluated using commercially available azido-magnetic beads. The samples were analysed by confocal microscopy to reveal fluorescent beads (Figure 18).

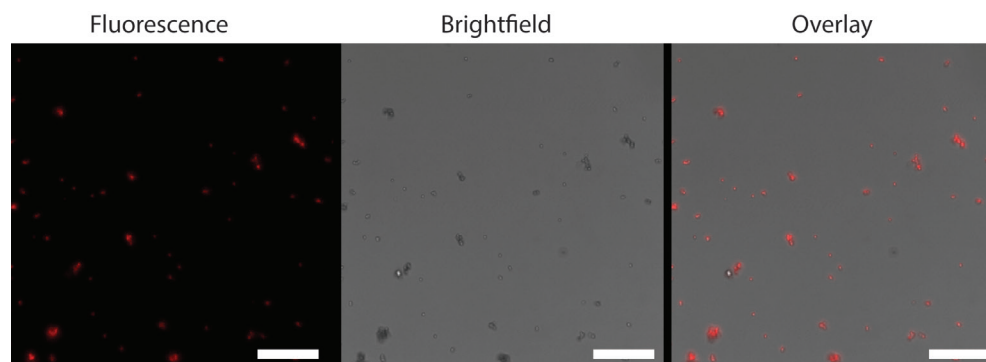


Figure 18. Positive control with azide magnetic beads and DBCO Cy-5. Scale bar = 20 μm .

2. Synthesis of chemical probes

Azido-monosaccharides Ac₄GlcNAz **1**, Ac₄GalNAz **2**, Neu5Az **5**, Ac₄RhamAz **4**, and Ac₄FucAz **3** were synthesised following literature procedures.^{12, 20, 55-56}

3. Biological procedures

Bacterial strains and growth conditions. Bacteria used in this study are *Bacteroides fragilis*, *Anaerostipes rhamnosivorans*, *Intestimonas butyriciproducens*,²¹ *Eubacterium hallii*, and *Akkermansia muciniphila* MucT (ATCC BAA-835). *A. muciniphila* was grown in CP medium as described earlier or in a synthetic medium where mucin was replaced by 6 g/l threonine, 6 g/l proline, CaCl₂ (11 g/l) and *N*-acetylglucosamine (50 mM).⁵ Mucin (5% w/v) was purified essentially as described by Derrien et al.⁵ CP medium was used for *I. butyriciproducens* and *A. rhamnosivorans* with the addition of either glucose or rhamnose (50 mM), YCFA medium was used for *E. hallii*. Co-cultures of *A. muciniphila* with *I. butyriciproducens* or *A. rhamnosivorans* were grown in CP medium with the above described additives. *B. fragilis* was grown on basal media. For the co-culture of *A. muciniphila* and *I. butyriciproducens*, *I. butyriciproducens* was inoculated first and *A. muciniphila* was added 5 hours later.

Application of bioorthogonal probes. All probes were dissolved in DMSO and stored in the -20 °C freezer and filter-sterilised before use. All media were autoclaved before use. Phosphate buffer solution (PBS) was made as follows: 800 ml H₂O containing potassium phosphate monobasic (1.12 g), NaCl (72 g), and sodiumphosphate dibasic (6.36 g).

Bacterial strains were grown till stationary phase when experiments were performed. Strains were inoculated and grown till stationary phase. From this culture, 1000x was taken and inoculated into fresh medium. For *B. fragilis* add glycan-probes (50 µM, final concentration) and incubate the cell culture overnight (16 hours) at anaerobic conditions at 37 °C. For *A. rhamnosivorans*, *I. butyriciproducens*, and *E. hallii* add glycan-probes (50 µM, final concentration) and incubate the cell culture overnight (10-16 hours) at anaerobic conditions at 37 °C. For *A. muciniphila*: add glycan-probes (50 µM, final concentration) and incubate the cell culture for maximum 10 hours at anaerobic conditions at 37 °C.

From this culture, take 0.5-1.0 ml cell culture and put in centrifuge for 10 min at 4 °C, 14,000 rpm. Wash cells twice with PBS. Pour off supernatant and suspend cells in click buffer (For CuAAC: 5 µl dye, 10 µl CuSO₄ (100mM) and 40 µl ascorbic acid (20 mM) in 500 µl PBS, For SPAAC: 500 µl PBS and 5 µl dye.) Incubate for 60 minutes at room temperature in the dark. Put samples in centrifuge (10 min at 4 °C, 14,000 rpm) and pour off supernatant. Wash the cells 3 times with 0.5 ml PBS, centrifuge in between. For flow cytometry or confocal analyses the supernatant was suspended in 100 µl PBS.

4. Biological analysis

Confocal microscopy. Ibidi μ -slide 8 well glass bottom slides were used for confocal microscopy. Live imaging was performed on a confocal laser scanning microscope (Leica TCS SP8X system). DBCO Cy-3 and Cy-5 were excited using a white light continuum laser. Images were captured using a 633 water-immersion objective with a numeric aperture of 1.2 and a pinhole set to 1 Airy unit. Images were captured and processed by ImageJ or LAS AF Lite.

Flow cytometry. Analysis of glycan labelled cells was performed using a MACSQuant® Analyzer 10 Flow Cytometer. Single cell populations were localised using FCS and SSC (488-nm 20 mW laser), gating pulse Area over pulse Width. Furthermore, Syto-9 (488-nm, 502 LP, 530/30 nm) was used to counterstain and distinguish cells from background. Data analysis was performed using FlowJo X (Treestar, USA).

Acknowledgement

We are grateful to B. Ø. Palsson for providing the ECOM4 strain. Yvette Luijkx, Tim Leenders, and Lemeng Chao are acknowledged for the synthesis of Ac₄FucAz, Neu5Az, Ac₄RhamAz. Jan-Willem Borst is thanked for his guidance and help for confocal microscopy. Bart Bardoel and Suzan Rooijackers (both UMC Utrecht) and Marcel Tempelaars (FHM) are thanked for their help in flow cytometry analysis. Steven Aalvink, Kees van der Ark, and Nico Claassens (all MIB) are thanked for their guidance during microbiology experiments.

References

1. Carding, S.; Verbeke, K.; Vipond, D. T.; Corfe, B. M.; Owen, L. J., *Microb. Ecol. Health Dis.* **2015**, *26*, 26191.
2. Zaia, J., *Nature Methods* **2011**, *8* (1), 55-57.
3. Tytgat, H. L.; de Vos, W. M., *Trends Microbiol.* **2016**, *24* (11), 853-861.
4. Tailford, L. E.; Crost, E. H.; Kavanaugh, D.; Juge, N., *Front. Genet.*, **2015**, *6*, 81.
5. Derrien, M.; Vaughan, E. E.; Plugge, C. M.; de Vos, W. M., *Int. J. Syst. Evol. Microbiol.* **2004**, *54*, 1469-76.
6. Sminia, T. J.; Zuillhof, H.; Wennekes, T., *Carbohydr. Res.* **2016**, *435*, 121-141.
7. Patterson, D. M.; Nazarova, L. A.; Prescher, J. A., *ACS Chem. Biol.* **2014**, *9* (3), 592-605.
8. Prescher, J. A.; Bertozzi, C. R., *Nat. Chem. Bio.* **2005**, *1* (1), 13-21.
9. Siegrist, M. S.; Swarts, B. M.; Fox, D. M.; Lim, S. A.; Bertozzi, C. R., *FEMS Microbiol. Rev.* **2015**, *39* (2), 184-202.
10. Geva-Zatorsky, N.; Alvarez, D.; Hudak, J. E.; Reading, N. C.; Erturk-Hasdemir, D.; Dasgupta, S.; von Andrian, U. H.; Kasper, D. L., *Nature Med.* **2015**, *21* (9), 1091-100.
11. Liu, F.; Aubry, A. J.; Schoenhofen, I. C.; Logan, S. M.; Tanner, M. E., *ChemBioChem* **2009**, *10* (8), 1317-20.
12. Besanceney-Webler, C.; Jiang, H.; Wang, W.; Baughn, A. D.; Wu, P., *Bioorg. Med. Chem. Lett.* **2011**, *21* (17), 4989-92.
13. Dumont, A.; Malleron, A.; Awwad, M.; Dukan, S.; Vauzeilles, B., *Angew. Chem.* **2012**, *51* (13), 3143-6.
14. 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., *J. Am. Chem. Soc.* **2012**, *134* (39), 16123-6.
15. Memmel, E.; Homann, A.; Oelschlaeger, T. A.; Seibel, J., *Chem. Commun.* **2013**, *49* (66), 7301-3.
16. Fugier, E.; Dumont, A.; Malleron, A.; Poquet, E.; Mas Pons, J.; Baron, A.; Vauzeilles, B.; Dukan, S., *PLoS One* **2015**, *10* (6), e0127700.
17. Sherratt, A. R.; Chigrinova, M.; McKay, C. S.; Beaulieu, L. P. B.; Rouleau, Y.; Pezacki, J. P., *RSC Adv.* **2014**, *4* (87), 46966-46969.
18. Hudak, J. E.; Alvarez, D.; Skelly, A.; von Andrian, U. H.; Kasper, D. L., *Nat. Microbiol.* **2017**, *2*, 17099.
19. Coyne, M. J.; Reinap, B.; Lee, M. M.; Comstock, L. E., *Science* **2005**, *307* (5716), 1778-1781.
20. Zhu, Y.; Chen, X., *ChemBioChem* **2017**.
21. Bui, T. P.; Shetty, S. A.; Lagkouvardos, I.; Ritari, J.; Chamlagain, B.; Douillard, F. P.; Paulin, L.; Piironen, V.; Clavel, T.; Plugge, C. M.; de Vos, W. M., *Environ. Microbiol. Rep.* **2016**, *8* (6), 1024-1037.
22. Miquel, S.; Martin, R.; Rossi, O.; Bermudez-Humaran, L. G.; Chatel, J. M.; Sokol, H.; Thomas, M.; Wells, J. M.; Langella, P., *Curr. Opin. Microbiol.* **2013**, *16* (3), 255-61.
23. Collins, M. D.; Lawson, P. A.; Willems, A.; Cordoba, J. J.; Fernandez-Garayzabal, J.; Garcia, P.; Cai, J.; Hippe, H.; Farrow, J. A., *Int. J. Syst. Bacteriol.* **1994**, *44* (4), 812-26.
24. Duncan, S. H.; Louis, P.; Flint, H. J., *Appl. Environ. Microbiol.* **2004**, *70* (10), 5810-7.
25. Duncan, S. H.; Hold, G. L.; Harmsen, H. J.; Stewart, C. S.; Flint, H. J., *Int. J. Syst. Evol. Microbiol.* **2002**, *52* (Pt 6), 2141-6.
26. Bui, T. P.; Ritari, J.; Boeren, S.; de Waard, P.; Plugge, C. M.; de Vos, W. M., *Nat. Commun.* **2015**, *6*, 10062.
27. Moore, W. E. C.; Cato, E. P.; Holdeman, L. V., *Int. J. Syst. bacteriology* **1972**, *22* (2), 78-80.
28. Onderdonk, A. B.; Kasper, D. L.; Cisneros, R. L.; Bartlett, J. G., *J. Infect. Dis.* **1977**, *136* (1), 82-9.
29. Gerdes, S. Y.; Scholle, M. D.; Campbell, J. W.; Balazsi, G.; Ravasz, E.; Daugherty, M. D.; Somera, A. L.; Kyrpides, N. C.; Anderson, I.; Gelfand, M. S.; Bhattacharya, A.; Kapatral, V.; D'Souza, M.; Baev, M. V.; Grechkin, Y.; Mseeh, F.; Fonstein, M. Y.; Overbeek, R.; Barabasi, A. L.; Oltvai, Z. N.; Osterman, A. L., *J. Bacteriol.* **2003**, *185* (19), 5673-84.
30. Bui, T. P.; de Vos, W. M.; Plugge, C. M., *Int. J. Syst. Evol. Microbiol.* **2014**, *64* (Pt 3), 787-93.
31. Surana, N. K.; Kasper, D. L., *Immunol. Rev.* **2012**, *245* (1), 13-26.

32. Ottman, N.; Davids, M.; Suarez-Diez, M.; Boeren, S.; Schaap, P. J.; Martins Dos Santos, V. A. P.; Smidt, H.; Belzer, C.; de Vos, W. M., *Appl. Environ. Microbiol.* **2017**.
33. Plovier, H.; Everard, A.; Druart, C.; Depommier, C.; Van Hul, M.; Geurts, L.; Chilloux, J.; Ottman, N.; Duparc, T.; Lichtenstein, L.; Myridakis, A.; Delzenne, N. M.; Klievink, J.; Bhattacharjee, A.; van der Ark, K. C.; Aalvink, S.; Martinez, L. O.; Dumas, M. E.; Maiter, D.; Loumaye, A.; Hermans, M. P.; Thissen, J. P.; Belzer, C.; de Vos, W. M.; Cani, P. D., *Nat. Med.* **2016**.
34. Russell, W. R.; Hoyles, L.; Flint, H. J.; Dumas, M. E., *Curr. Opin. Microbiol.* **2013**, *16* (3), 246-54.
35. Koropatkin, N. M.; Cameron, E. A.; Martens, E. C., *Nat. Rev. Microbiol.* **2012**, *10* (5), 323-335.
36. Louis, P.; Flint, H. J., Diversity, *FEMS Microbiol. Lett.* **2009**, *294* (1), 1-8.
37. Kayser, H.; Geilen, C.; Paul, C.; Zeitler, R.; Reutter, W., *FEBS* **1992**, *301* (2), 137-140.
38. Kayser, H.; Zeitler, R.; Kannicht, C.; Grunow, D.; Nuck, R.; Reutter, W., *J. Biol. Chem.* **1992**, *267* (24), 16934-16938
39. Boyce, M.; Carrico, I. S.; Ganguli, A. S.; Yu, S. H.; Hangauer, M. J.; Hubbard, S. C.; Kohler, J. J.; Bertozzi, C. R., *Proc. Natl. Acad. Sci. USA* **2011**, *108* (8), 3141-6.
40. Hollister, J.; Conrad, H.; Jarvis, D. L., *Glycobiology* **2003**, *13* (6), 487-95.
41. Laughlin, S. T.; Baskin, J. M.; Amacher, S. L.; Bertozzi, C. R., *Science* **2008**, *320* (5876), 664-7.
42. Sawa, M.; Hsu, T. L.; Itoh, T.; Sugiyama, M.; Hanson, S. R.; Vogt, P. K.; Wong, C. H., *Proc. Natl. Acad. Sci. USA* **2006**, *103* (33), 12371-6.
43. Giraud, M. F.; Naismith, J. H., *Curr. Opin. Struct. Biol.* **2000**, *10* (6), 687-96.
44. Hoogenboom, J.; Berghuis, N.; Cramer, D.; Geurts, R.; Zuilhof, H.; Wennekes, T., *BMC Plant. Biol.* **2016**, *16* (1), 220.
45. Agard, N. J.; Prescher, J. A.; Bertozzi, C. R., *J. Am. Chem. Soc.* **2004**, *126* (46), 15046-7.
46. Antonczak, A. K.; Simova, Z.; Tippmann, E. M., *J. Biol. Chem.* **2009**, *284* (42), 28795-800.
47. Baba, T.; Ara, T.; Hasegawa, M.; Takai, Y.; Okumura, Y.; Baba, M.; Datsenko, K. A.; Tomita, M.; Wanner, B. L.; Mori, H., *Mol. Syst. Biol.* **2006**, *2*, 2006 0008.
48. Snyder, M. L.; Lichstein, H. C., *J. Infect. Dis.* **1940**, *67* (2), 113-115.
49. Lichstein, H. C.; Soule, M. H., *J. Bacteriol.* **1944**, *47* (3), 253-7.
50. Laughlin, S. T., *Chemical Approaches to Imaging Glycans in Living Animals.* **2008**.
51. Ouwerkerk, J. P.; van der Ark, K. C.; Davids, M.; Claassens, N. J.; Robert Finestra, T.; de Vos, W. M.; Belzer, C., *Appl. Environ. Microbiol.* **2016**.
52. Portnoy, V. A.; Scott, D. A.; Lewis, N. E.; Tarasova, Y.; Osterman, A. L.; Palsson, B. O., *Appl. Environ. Microbiol.* **2010**, *76* (19), 6529-40.
53. Johansson, E. M.; Cruz, S. A.; Kolomiets, E.; Buts, L.; Kadam, R. U.; Cacciarini, M.; Bartels, K. M.; Diggle, S. P.; Camara, M.; Williams, P.; Loris, R.; Nativi, C.; Rosenau, F.; Jaeger, K. E.; Darbre, T.; Reymond, J. L., *Chem. Biol.* **2008**, *15* (12),
54. Sokol, H.; Pigneur, B.; Watterlot, L.; Lakhdari, O.; Bermudez-Humaran, L. G.; Gratadoux, J. J.; Blugeon, S.; Bridonneau, C.; Furet, J. P.; Corthier, G.; Grangette, C.; Vasquez, N.; Pochart, P.; Trugnan, G.; Thomas, G.; Blottiere, H. M.; Dore, J.; Marteau, P.; Seksik, P.; Langella, P., *Proc. Natl. Acad. Sci. USA* **2008**, *105* (43), 16731-6.
55. Laughlin, S. T.; Bertozzi, C. R., *Nat. Protoc.* **2007**, *2* (11), 2930-44.
56. Saxon, E.; Luchansky, S. J.; Hang, H. C.; Yu, C.; Lee, S. C.; Bertozzi, C. R., *J. Am. Chem. Soc.* **2002**, *124* (50), 14893-902.



Chapter 5

Total Synthesis towards Pseudaminic Acid Probes

"What Shall We Do Now?"

Is There Anybody Out There, Pink Floyd

Work in progress: Tjerk J. Sminia, Han Zuilhof, and Tom Wennekes

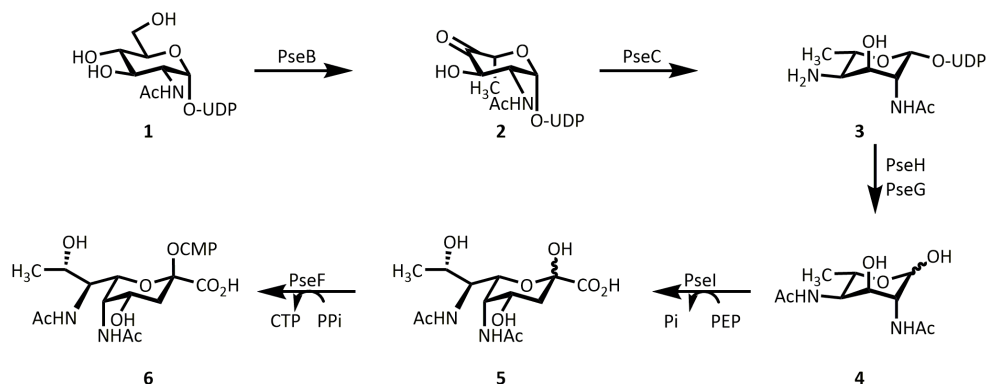
A comparative overview:

(Bio)-Synthesis of Pseudaminic Acid and Legionaminic Acid

In recent years, several synthesis of pseudaminic (Pse) and legionaminic acid (Leg) have been published by different research groups across the world. Our work on the synthesis of pseudaminic acid was partly inspired by these routes. The following section provides a comparative overview of the Pse and Leg biosynthesis pathways and the synthetic approaches published up to this point towards Pse and Leg or variants thereof.

Biosynthesis Pseudaminic Acid and Legionaminic Acid

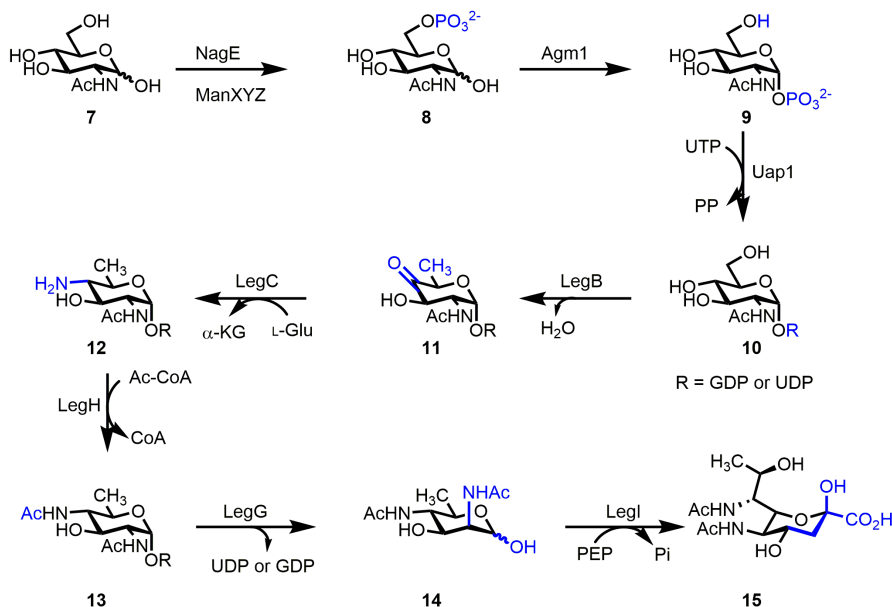
The pseudaminic acid biosynthesis starts with an activated *N*-acetylglucosamine glycan **1** (Scheme 1). Although a different biosynthesis might be present in other bacteria the biosynthesis as characterised in *Campylobacter jejuni* and *Helicobacter pylori* will be discussed. A short effective five step biosynthesis using a variety of different enzymes is used by these bacteria. In the first step the PseB enzyme, a NAD(P)-dependant C-6 dehydratase and C-5 epimerase, produces ketone **2** via the *N*-acetylglucosamine UDP glycoside **1**. The second amine is installed on the C-4 position with PseC. *N*-acetylation and removal of the UDP activator is achieved by PseH and PseG. The aldol reaction is performed by PseI using phosphoenolpyruvate (PEP). In natural glycoconjugates, sialic acids are exclusively present as the α -sialoside. Although glycosylation mechanism are not exactly known, PseF makes a β -CMP pseudaminic acid precursor **6** selectively which is ready for glycosylation reactions.



Scheme 1. Biosynthesis of pseudaminic acid.

Almost all enzymes involved in the biosynthesis of pseudaminic acid have been isolated and characterised. For example PseG, the fourth enzyme in the biosynthesis, was characterised via X-ray crystallography with either the *apo*-form or the UDP-product bound.¹ The biosynthesis of legionaminic acid was unravelled by Logan and co-workers in 2009 (Scheme 2).² The most striking point was

the discovery of GDP-GlcNAc **10** as the key building block rather than UDP-GlcNAc which is more common for this class of nonulosonic acid pathways. The pathway unravelled at that time started with dehydratase LegB to eliminate water from the C4,6 of the GDP glycan **10**. The 4-keto intermediate **11** was transferred to an amine using LegC, a PLP-dependent aminotransferase. The resulting di-amino glycan **12** was acetylated using acetyltransferase LegH. LegG, a hydrolysing 2-epimerase transposed the equatorial acetylated amine into the axial acetylated amine. In the final transformation LegI condenses the hydrolysed glycan **14** with pyruvate to produce Leg **15**. The unique GDP-linked intermediates suggested that production on large scale would be very challenging.



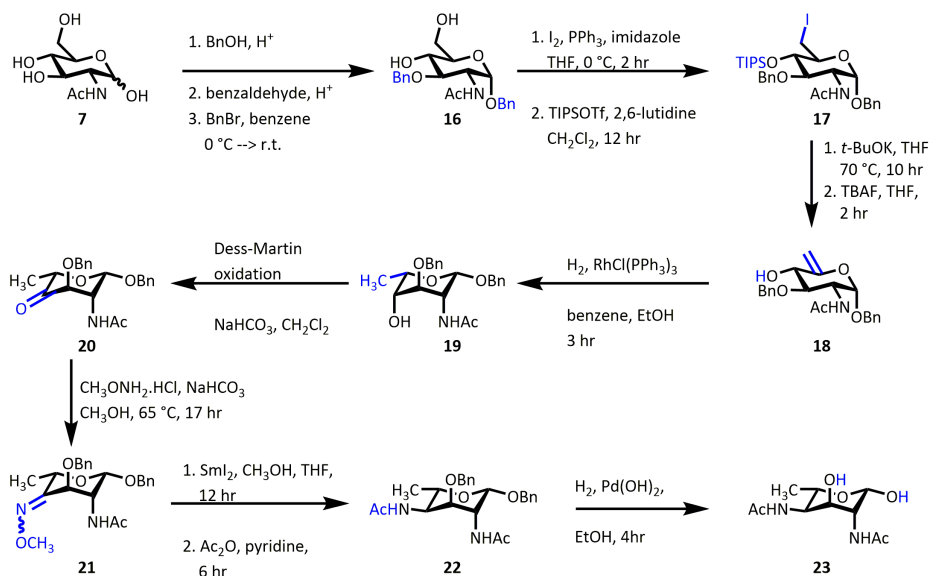
Scheme 2. *De novo* biosynthetic pathway by Schoenhofen and co-workers.

A few years later, in 2016, an updated biosynthesis was reported by Boddy et al.³ The expertise of producing other nonulosonic acids for example Neu5Ac in *E. coli* resulted in the hypothesis that legionaminic acid could also be produced in *E. coli*. For this system UDP-diNAcBac served as the most important glycan intermediate. The resulting optimised biosynthetic pathway produced 120 mg/l of Leg in *E. coli*. Chemical glycosylation studies with legionaminic acid are important. For that purpose an α -phenylthioglycoside was generated from pure Leg. First the *O*-acetyl groups were de-acetylated under standard Zemplén conditions followed by ester formation with CH_3OH in TFA. The protected carboxylic acid was activated with acetyl chloride and subsequently reacted with thio-phenol to give the α -phenylthioglycoside. The α -anomer is important since the β -anomer is the natural anomer in solution.

Synthesis of Pseudaminic Acid by Ito and co-workers

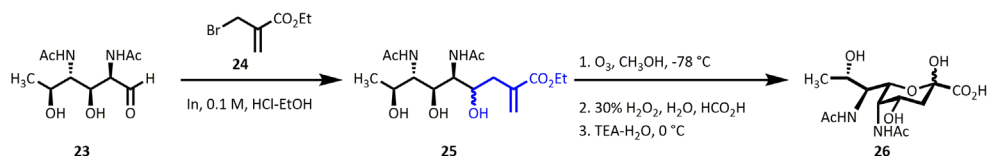
In 2010, Ito and co-workers reported the synthesis of Pse from 6-hexose precursor 6-deoxy-AltdiNAc **23** starting from *N*-acetyl-D-glucosamine (D-GlcNAc) (Scheme 3).⁴ The synthesis was inspired by the biosynthesis of Pse and was achieved in 17 steps.

2-acetamido-2-deoxy-D-glucose (D-GlcNAc, **7**) was protected at the anomeric centre with a benzyl group. The product was treated with benzaldehyde to give the 4,6-benzylidene benzyl glycoside. The 3-OH position was protected with another benzyl group to give intermediate **16**.⁵⁻⁶ Primary iodination with I_2 , imidazole, and triphenylphosphine gave 4-OH benzyl glycoside. To optimise the formation of olefin **18** a two-step procedure was followed, first the 4-OH was protected with a triisopropylsilyl (TIPS) group followed by β -elimination of the 6-iodine with *t*-BuOK and finally deprotection with *tert*-butylammonium fluoride (TBAF) to give olefin **18**. Wilkinson's catalyst was surprisingly selective in reducing the exocyclic olefin towards 2,6-dideoxy-*N*-acetyl-L-idosamine **19**. The most challenging step in the synthesis of 6-deoxy-AltdiNAc **23** was the instalment of the 4-deoxy amino functionality. A straightforward S_N2 -type reaction by activation and subsequent displacement with sodium azide was not successful. Therefore, a reductive amination from an oxime intermediate **21** was achieved. 4-OH glycoside **19** was oxidised to the ketone **6** and treated with *O*-methylhydroxylamine hydrochloride to give a mixture of *E/Z* isomers. The reduction was performed with one-electron reducing agent samariumdiodide to give 4-acetamidoaltrose derivative **22** as a single isomer after acetylation of the crude mixture. The other isomer was not formed probably due to unfavourable 1,3-diaxial interaction. Hydrogenation gave the biosynthetic intermediate of Pse **23**.



Scheme 3. Synthesis of intermediate 6-deoxy-AltdiNAc by Ito and co-workers.

In nature chain elongation of the hexose precursor is achieved with pyruvic acid by an aldol condensation, catalysed by aldolases.⁷ A chemical synthesis equivalent of this step is the Barbier allylation.⁸ This process is well known in organic chemistry literature and has been performed numerous times on other sialic acid precursors.⁹⁻¹⁰ In the Ito route, the Barbier allylation with indium in the presence of commercially available bromomethylacrylate ester **24** yielded the desired *erythro* product **25** in a 5:4 ratio to the *threo* isomers (Scheme 4). Several attempts to improve the selectivity by adding Lewis acids and chiral ligands did not improve the selectivity. Ozonolysis of the olefin and saponification of the ester gave pseudaminic acid **26**. Glycosylation with a Pse-based glycosylation donor was also investigated. A dibenzyl phosphite Pse donor was reacted with 6-OH protected glucose derivative to give the corresponding disaccharide in an excellent $\alpha:\beta$ ratio of 10:1. The significant by-product in the glycosylation reaction was an eliminated 2,3-dehydro Pse derivative, a side reaction that is often observed in reactions with sialic acids.



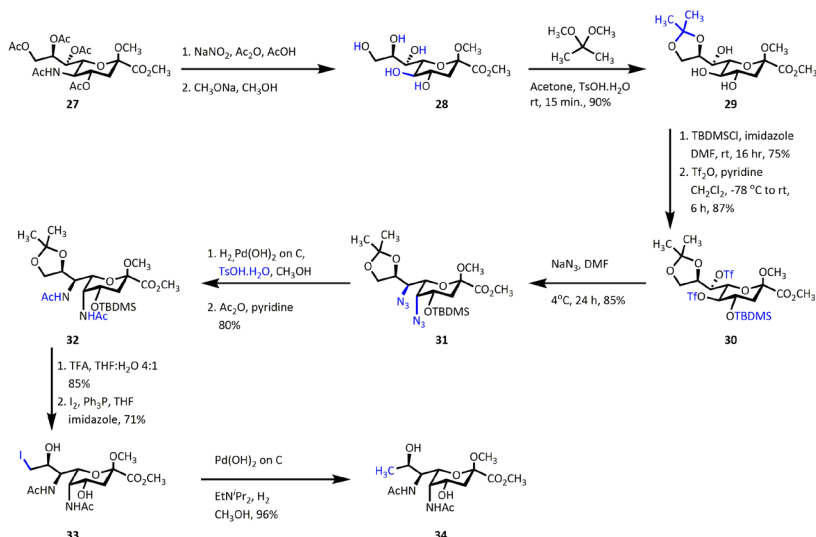
Scheme 4. Final steps in the synthesis of Pse by Ito and co-workers.

A drawback of the synthesis performed by Ito and co-workers is that the protecting group pattern in intermediate 6-deoxy-AltdiNac **23** and pseudaminic acid **26** does not discriminate between the two amine functionalities. For chemical biology studies, which is the aim of our studies, selective chemical tags on either the 5-NH₂ or the 7-NH₂ need to be installed. The two amine functionalities can of course be discriminate early in the synthesis but this might influence the selectivity of key steps such as the reduction with Wilkinson's catalyst or the chain elongation under Barbier conditions. An advantage of the Ito route is that the synthesis of intermediate 6-deoxy-AltdiNac **23** is relative fast. The continuation towards Pse **26** can be achieved in an overall step count of only 17 steps and provided enough material to perform glycosylation studies.

Synthesis by Kiefel and co-workers

The synthesis route towards 8-*epi*-pseudaminic acid **34** as published by Kiefel and co-workers started from keto-deoxynonulosonic acid **28** (KDN), a member of the class of nonulosonic acid which also include *N*-acetylneuraminic acid (Scheme 5).¹¹ The starting material, KDN, is commercially available, but quite expensive and was therefore synthesised from Neu5Ac **27** via the method of Schreiner and Zbiral, and later optimised by Glatzhofer et al.¹²⁻¹³ The key steps in this synthesis are the introduction of an amine on C-7 with inverted stereochemistry, the inversion of the amine functionality on the C-5 position, and the deoxygenation at the C-9 position of Neu5Ac. The methyl ester β -methyl glycoside of Neu5Ac **27** was treated with sodium nitrite in acetic anhydride and acetic acid. The initial conditions described by Glatzhofer were further optimised to a procedure that uses a large excess of sodium

Total Synthesis towards Pseudaminic Acid Probes



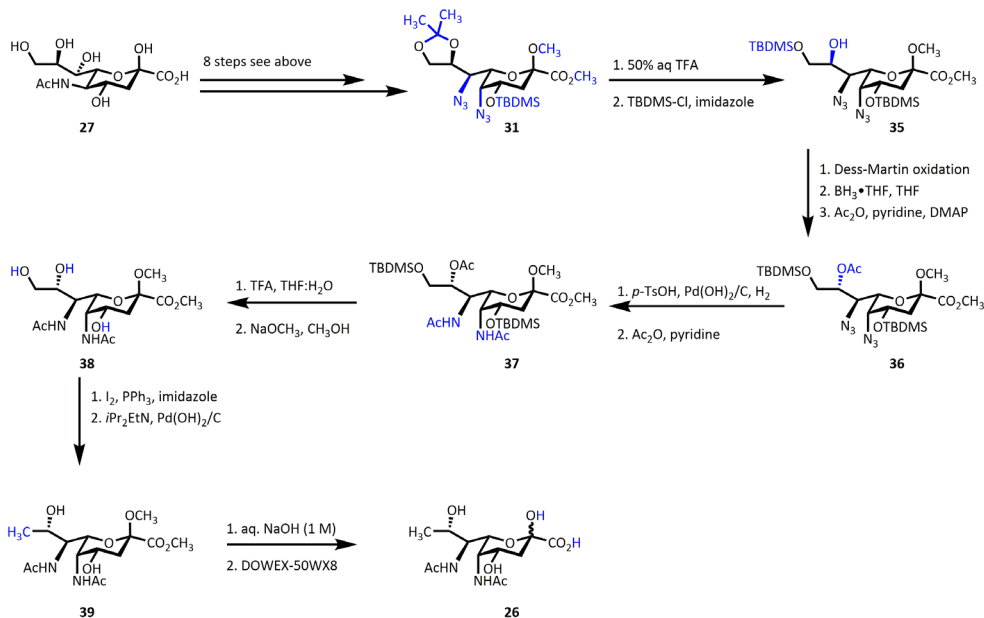
Scheme 5. Synthesis of 8-*epi*-pseudaminic acid by Kiefel and co-workers.

nitrite at elevated temperatures for several hours. This optimised process resulted in an improved yield of 57% of the protected KDN intermediate. Zémlen deacetylation gave the β -methyl glycoside methyl ester of KDN **28** as the starting point for the synthesis of 8-*epi* pseudaminic acid **34**.

An acetoneid was installed on the 8,9-position. A short reaction time proved crucial to reduce the formation of by-products. The C-4 hydroxyl **29** was selectively silylated with *tert*-butyldimethylsilyl chloride in the presence of imidazole and DMF. With only the C-5 and C-7 hydroxyls unprotected, the instalment of the amine functionalities on these positions could commence. The two alcohols at C-5 and C-7 were activated by triflic anhydride and exposed to an excess of sodium azide to provide the *bis*-azide **31** in excellent yield. A Staudinger reduction in the presence of acetic anhydride of the *bis*-azide **31** was performed to give selectively the 5-azido 7-acetamide intermediate **32**. Simple hydrogenation of the *bis*-azide resulted in the formation of a 1,5-lactam by-product while formation of the *bis*-acetamide was never seen. The discrimination of the two azides in this intermediate is of high interest for the synthesis of pseudaminic acid analogues including chemical probes of pseudaminic acid for metabolic labelling. Finally, *bis*-acetamide **32** was synthesised after hydrogenation with $\text{Pd(OH)}_2/\text{C}$ in methanol with trace amounts of *p*-toluenesulfonic acid. Trifluoroacetic acid treatment gave the free hydroxyl precursor. In the last step iodination of the primary alcohol at C-9 and subsequent treatment with $\text{Pd(OH)}_2/\text{C}$ with DIPEA and H_2 gave 8-*epi*-pseudaminic acid **34**.

The synthesis of 8-*epi*-pseudaminic acid was achieved in 9 steps starting from Neu5Ac **27**. The authors do not comment if 8-*epi*-pseudaminic acid **34** is a good precursor for the natural occurring pseudaminic acid. Given this fact, this route is probably not likely to be of interest for the synthesis of pseudaminic acid on large scale. Although 8-*epi*-pseudaminic acid **34** is not known to be present in bacteria it might be of interest as a false positive in biological essays.

Recently, Kiefel and co-workers published a revised version of this route that is able to produce natural pseudaminic acid **26** (Scheme 6).¹⁴ From intermediate **31** the synthesis was continued with TFA treatment followed by TBS protection of the primary alcohol. The inversion on the 8-position, the key step in this sequence, was possible by oxidation with Dess-Martin reagent to the ketone followed by selective reduction with $\text{BH}_3 \cdot \text{THF}$ complex. This procedure was conducted in an one-pot experiment with a 74% isolated yield of the desired epimer **36**. Reductive acetylation of the di-azide **36** was unsuccessful probably due to the C-8 alcohol. The C-8 alcohol was therefore first acetylated followed by hydrogenation with Pearlman's catalyst and *p*-TsOH as catalyst to give the di-amine which was immediately converted into the di-acetamide intermediate **37**.

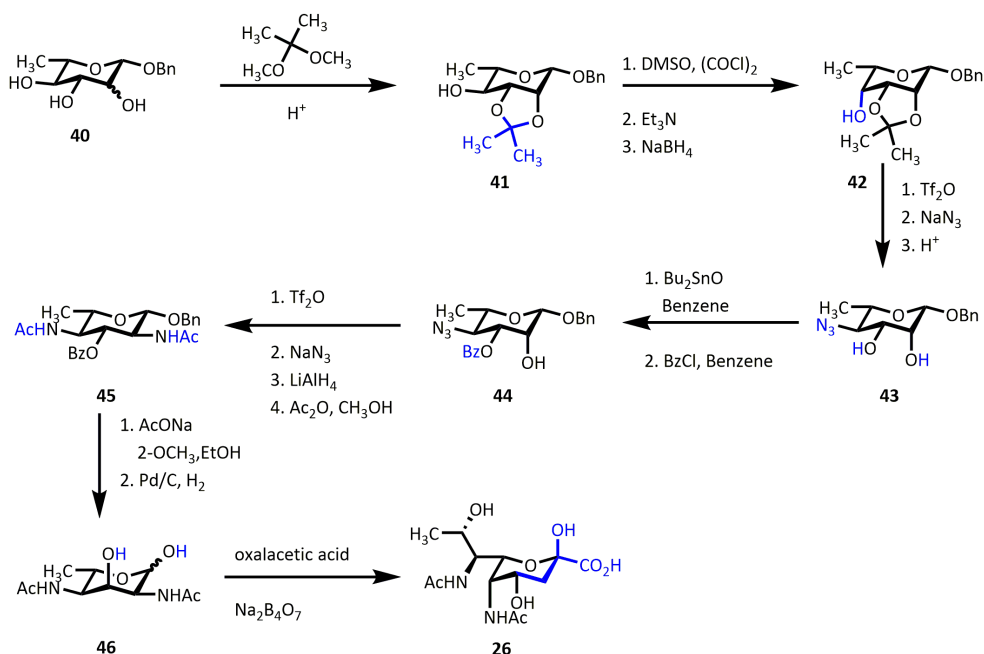


Scheme 6. Synthesis of pseudaminic acid by Kiefel and co-workers.

In the hydrogenation step, a catalytic amount of *p*-TsOH proved necessary to prevent the formation of a lactam byproduct which was seen after the intramolecular attack of the C-5 amine onto the carbonyl of the C-1 methyl ester. Deoxygenation of the 9-position was achieved via the same method with iodination followed by addition of base and Pd/C. Finally, saponification of the methyl ester and treatment with DOWEX H^+ resin provided pseudaminic acid **26**.

Synthesis Pseudaminic Acid by Zahringer

Around 2000, the group of Zahringer performed a tremendous effort in synthesising all natural occurring microbial sialic acids.¹⁵⁻¹⁸ For the synthesis of pseudaminic acid the starting point was L-rhamnose, the least expensive 6-deoxyhexose. L-Rhamnose was first converted in the methyl glycoside. Subsequently, installation of an 2,3-acetonide, benzylation on the C-3 position, and removal of the acetonide gave the 4-OBz L-rhamnose intermediate.¹⁹ Acetylation gave the 1,2,3-tri-*O*-acetyl L-rhamnose intermediate. A glycosyl bromide was made and subsequently reacted with NaCN. Selective de-acetylation of the C-3 position yielded the building block **40** which was used as the starting point for the synthesis of pseudaminic acid. Bu₂SnO-mediated benzylation of 1,2-diol was performed as previous described.¹⁹ The mixture of benzyl by-products was separated from the desired benzyl β-rhamnopyranoside benzoates.



Scheme 7. Synthesis of pseudaminic acid by Zahringer and co-workers.

Treatment of the 2,3,4-triol **40** with 2,2-dimethoxypropane gave the 2,3-acetonide **41**. A Swern oxidation gave the C-4 ketone which after reduction selectively produced the axial C-4 alcohol **42**.²⁰ The L-taloside derivative was treated with triflic anhydride and reacted with sodium azide to give 4-azido 4-deoxy L-taloside **43**. The acetonide was cleaved and the same procedure was followed to install the 2-azido 2-deoxy moiety. For this, a tin ketal was installed. This enabled the selective installation of a benzoyl moiety on C-3 **44**. Now, activation of the C-2 alcohol followed by reaction with sodium azide gave 2-azido 2-deoxy L-rhamnose precursor. The diazide was reduced with LiAlH₄ and protected with acetyl groups. Finally, the 3-OH position was inverted by activation with MsCl and

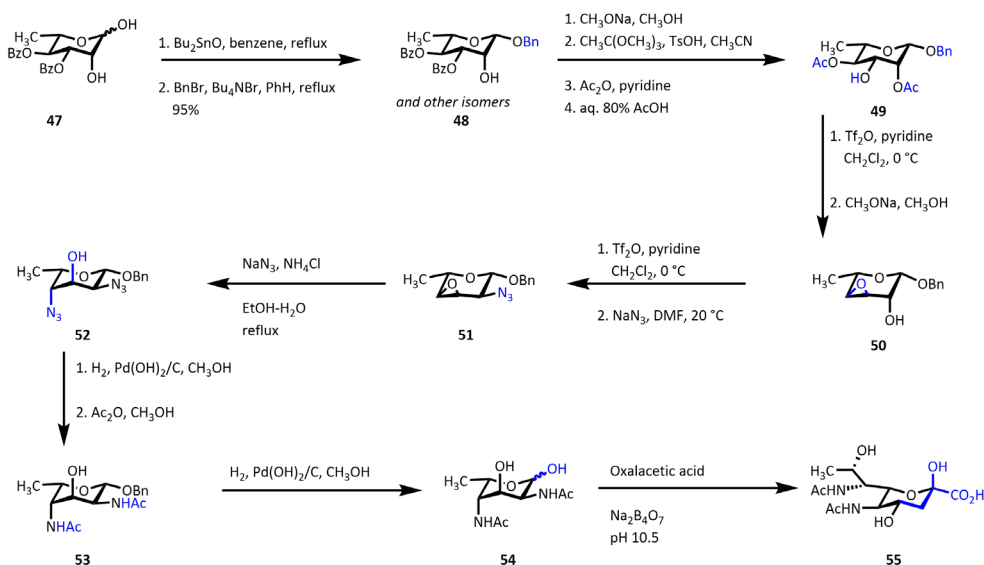
subsequent opening with aqueous 2-methoxyethanol. The microbial sialic acid precursor was hereby synthesised in 18 steps. The benzyl glycoside **45** was deprotected and the resulting anomeric mixture was condensed with oxalacetic acid and $\text{Na}_2\text{B}_4\text{O}_7$. After epimerisation of the C-2 position in the starting monosaccharide one of the products resulted in pseudaminic acid as well other related epimers. The desired pseudaminic acid **26** was obtained in only 3% yield. The above described synthetic route towards pseudaminic acid by Zahringer and co-workers is not suitable for further studies. The last step is low yielding and not selective, resulting in a mixture of epimers, without any further possibility to continue. The epimers however could be separated and therefore a detailed characterisation of all epimers by NMR was provided.

Legionaminic acid

Zahringer route to 8-*epi*-Legionaminic Acid from L-rhamnose

Zahringer and co-workers were the first to report on the synthetic route of 8-*epi*-legionaminic acid.^{16, 18} In a few years two synthetic procedures were developed (Scheme 8 and Scheme 9).

Zahringer and co-workers started the synthesis of 8-*epi*-legionaminic acid from L-rhamnose. Several modifications achieved the desirable configuration of the C-2 and C-4 amines. The 3,4-OBz L-rhamnose **47** was converted to a tin ketal with Bu_2SnO in refluxing benzene before regioselective benzylation was performed.¹⁹ This reaction gave the desired product as well as isomers where benzoyl migration

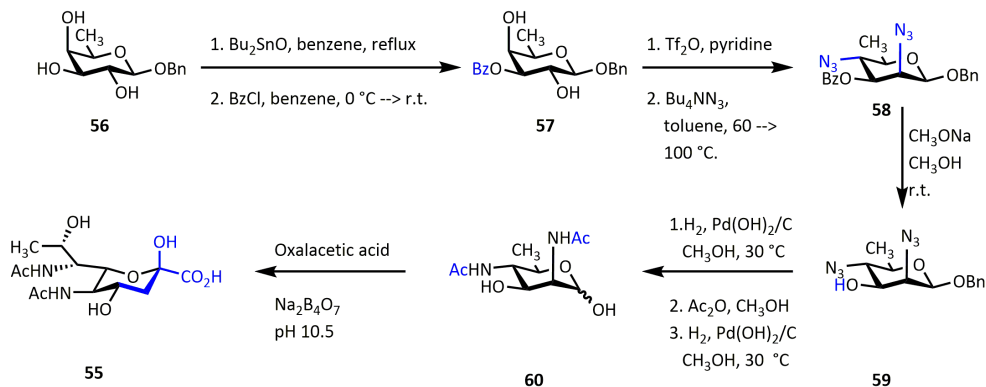


Scheme 8. Zahringer synthesis route towards 8-*epi*-legionaminic acid from L-rhamnose.

occurred. Debenzoylation with sodium methoxide gave the 2,3,4-triol L-rhamnose in large quantities. The 3-OH intermediate **49** was made via a short sequence of reactions. The benzoyl groups were removed followed by installation of an orthoester. This orthoester was opened selective to give the 3-OH intermediate **49**. The free hydroxyl was reacted with triflic anhydride to activate the alcohol. Zemplén de-acetylation and intramolecular attack of the triflate by the liberated alkoxide, resulted in formation of the epoxide **50**. The 2-OH was activated with triflic anhydride at 0 °C followed by nucleophilic displacement with sodium azide to give 2-deoxy 2-azido glycan **51**. At elevated temperatures the epoxide was selectively opened from the axial position to give the di-azide **52**. A sequence of hydrogenation with $\text{Pd}(\text{OH})_2$, *N*-acetylation, and again hydrogenation to remove the benzyl group resulted in the 8-*epi*-legionaminic acid precursor **54**. Oxalacetic acid in the presence of $\text{Na}_2\text{B}_4\text{O}_7$ at basic pH gave a condensation reaction resulting in a mixture of nonulosonic isomers. The isomers were separated by anion-exchange chromatography to give 8-*epi*-legionaminic acid **52** in 18% yield while the 4-*epi*-legionaminic acid was obtained in 20% yield.

Zahringer route to 8-*epi*-Legionaminic acid from D-fucose

A few years later Zahringer and co-workers also developed a synthesis towards 8-*epi*-legionaminic acid from D-fucose **56** (Scheme 9).¹⁷ After selective protection of the 3-OH with a benzoyl group the two remaining alcohols were activated with triflic anhydride. Nucleophilic substitution with Bu_4NN_3 gave the di-azide product **58**. The di-azide was reduced with $\text{Pd}(\text{OH})_2/\text{C}$ and immediately acetylated. The di-amino mannose derivative **60** can be chemically or enzymatically elongated by condensation with oxalacetic acid under basic conditions or natural enzymes isolated from bacterial species having 8-*epi*-legionaminic acid.

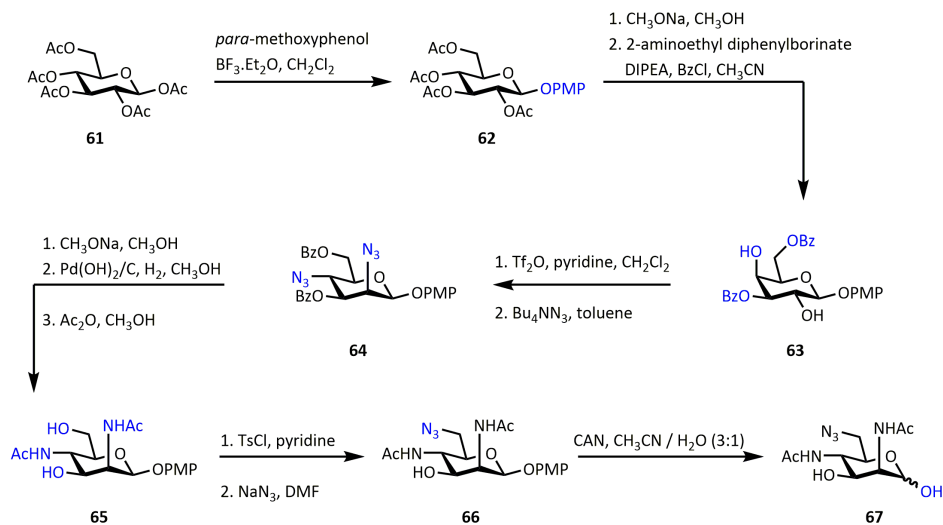


Scheme 9. Zahringer synthesis route towards 8-*epi*-legionaminic acid from D-fucose.

This synthesis of 8-*epi*-legionaminic acid by Zahringer and co-workers is relatively short. A main disadvantage is the lack of orthogonal protection of both amines. For metabolic labelling this would be a requisite. In order to make metabolic probes based on this route the yields have to be optimised. For example the first steps and the final step occurred in poor yields.

Vauzeilles route for an 9-azido Legionaminic Acid Derivative

Metabolic oligosaccharide engineering is a powerful tool to study surface glycans (**chapter 2**). Vauzeilles and co-workers reported in 2014 a metabolic labelling probe for legionaminic acid.²¹ An 6-azido hexose, 6-azido-2,4-diacetamido-2,4,6-trideoxy-D-mannopyranose **67** was synthesised in 11 steps (Scheme 10). The non-acetylated variant was also synthesised in order to study the differences in metabolic uptake into the cell. The synthesis was mainly inspired by the Zahringer synthesis with modification or improvements where necessary.

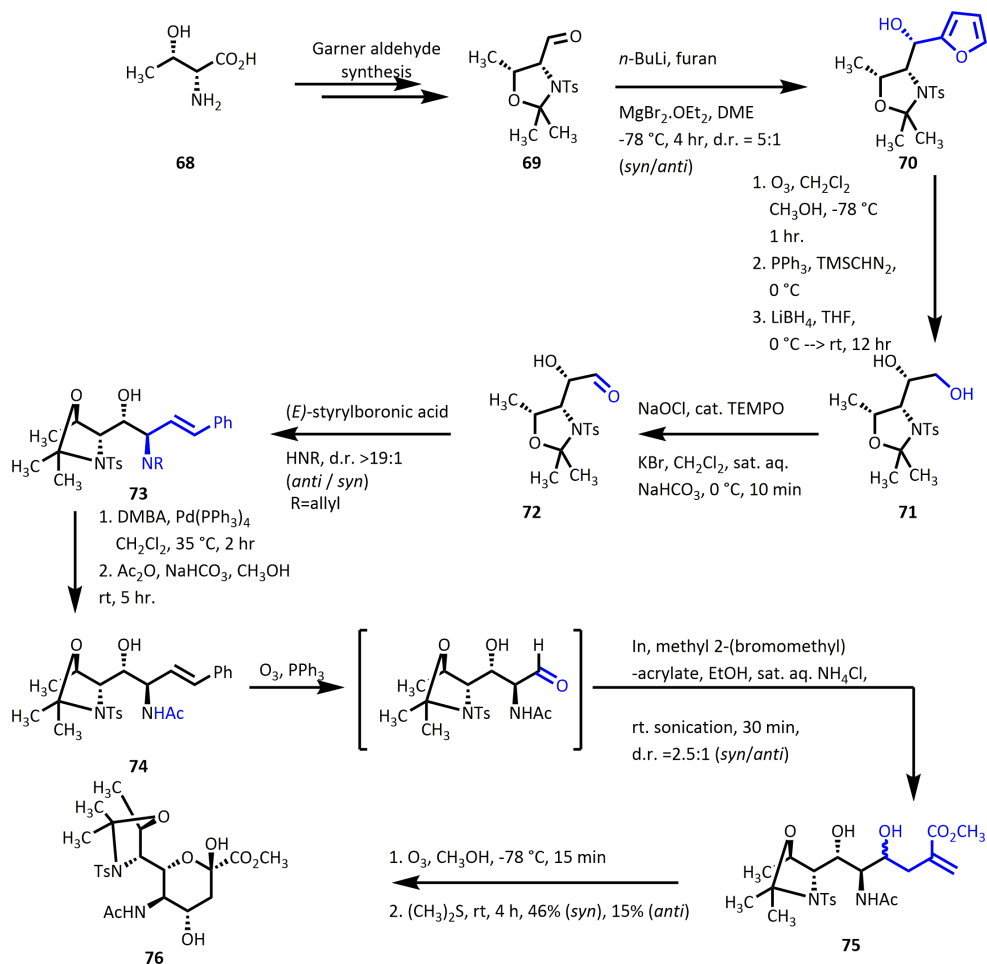


Scheme 10. Synthesis of legionaminic acid azido-precursor by Vauzeilles and co-workers.

β-D-galactose penta-acetate **61** was first protected at the anomeric centre as *para*-methoxyphenyl glycoside **62**. Zemplén de-acetylation followed by selective installation of two benzoyl groups on the 3 and 6-position was achieved in 70% overall yield (two steps). The two remaining alcohols **63** were activated with triflic anhydride and reacted with Bu_4NN_3 to give **64**. The di-azide was debenzoylated under Zemplén conditions followed by hydrogenation with $\text{Pd}(\text{OH})_2/\text{C}$. The two amines were acetylated in one-pot to give **65**. The alcohol on the 6-position was activated with TsCl and immediately reacted with sodium azide. The 6-azido legionaminic acid precursor **66** was deprotected at the anomeric centre to achieve an overall yield of 17%. For metabolic labelling purposes 6-azido legionaminic acid precursor **67** is a valuable substrate. The performed route by Vauzeilles and co-workers is well suited for this purpose.²¹

Seeberger Route Legionaminic Acid

The structural complexity of legionaminic acid is one of the reasons why only a few synthesis towards legionaminic acid are known. In 2015, Seeberger et al. published an elegant *de novo* synthesis of orthogonally protected legionaminic acid in 12 steps from aldehyde intermediate **69** (Scheme 11).²²



Scheme 11. Formal synthesis of legionaminic acid by Seeberger and co-workers.

D-Threonine **68** was chosen as a starting point to make, according to a slightly modified literature procedure, Garner aldehyde derivative **69**.²³ The intermediate D-*allo*-threonine is commercially available but notoriously expensive. Therefore, D-*allo*-threonine was made via methyl ester protection and subsequently *N*-benzoylation of D-threonine. Followed by the necessary C3-epimerisation which was performed by thionyl chloride induced cyclisation and acidic hydrolysis. Next, D-*allo*-threonine was treated with dimethoxypropane under catalytic acidic conditions to install the acetonide. Reduction of the ester with DIBAL-H gave the desired aldehyde **69**. To install the amine functionalities on aldehyde **69** a Petasis borono-Mannich reaction was chosen.²⁴ For this, aldehyde **69** was first reacted in a Cram-chelated organometallic addition reaction with 2-lithiofuran to yield the desired *syn*-alcohol **70**.²⁵ Ozonolysis in the presence of methanol gave a methyl ester which was reduced with lithium

borohydride to give diol **71**. X-ray crystallography confirmed the correct stereo-chemical assignment. Oxidation of the primary alcohol was rather challenging. Dess-Martin oxidation or TEMPO-mediated oxidation gave low conversion and no conversion at all, respectively. A modified TEMPO-oxidation following Anelli's optimisation gave the desired aldehyde **72** in quantitative yield without loss of the stereochemical purity.²⁶⁻²⁷ With the aldehyde **68** in hand the Petasis borono Mannich reaction was tested and extensively optimised. Monoallyl amine in ethanol in the presence of (*E*)-styrylboronic acid gave exclusive formation of the *anti*-diastereoisomers with d.r. > 19:1 in 76% yield. The allyl groups on the installed amine functionality were isomerized with Pd(0), hydrolysed and the liberated primary amine was acetylated for further use. Oxidative cleavage with ozone was problematic in the work-up. Resin bounded triphenylphosphine finally resulted in complete conversion to the aldehyde without loss of the stereochemistry outcome of the reaction. The mild indium-mediated Barbier allylation was performed in aqueous media while sonicated to install a masked pyruvate unit **75**. A mixture of alcohols was obtained which was used in the next step for the oxidative cleavage. A reductive workup and separation of the diastereoisomers gave protected 4-*epi*-legionaminic acid (not shown) and protected legionaminic acid **76** in good yield.

The synthesis of legionaminic acid as performed by Seeberger and co-workers is full of excellent synthetic manipulations which requires skilled organic chemists. Although a large scale synthesis via this route is probably not suitable to produce enough material to do extensive glycosylation and subsequent studies. The use of Garner aldehyde derivative ones again proves how valuable this intermediate is for synthesis of microbial sialic acids. The use of an amino acid as starting point is, given the complex structure, a smart starting point for both pseudaminic and legionaminic acid.

Conclusion comparative overview

Over the last years pseudaminic and legionaminic acid achieved a lot of interest. It is therefore not surprisingly that new synthetic routes have been developed towards Pse or Leg, precursors or derivatives thereof. New routes will probably be published soon after which the biological studies can have greater impact, and from which new insights in the relevance of Pse and Leg will likely follow soon.

Results and Discussion Synthesis Towards Pseudaminic acid

The L-mannose and L-glycine configurations in Pse (pseudaminic acid, **26**) provide a formidable synthetic challenge for which currently only a few chemical synthesis strategies exist (*vide supra*). Inspired by the biosynthesis of Pse in Nature and the known synthesis for Pse, we envisioned a two-phase strategy.^{4, 11} In the first phase a Pse hexose derivative **76**, the natural precursor towards Pse, would be synthesised. From this intermediate we should then be able to prepare Pse **26** using the well-developed chemistry for this final C-3 elongation in sialic acid chemical synthesis.^{4, 10, 28} In Figure 1 the retrosynthetic analysis for Pse is shown.

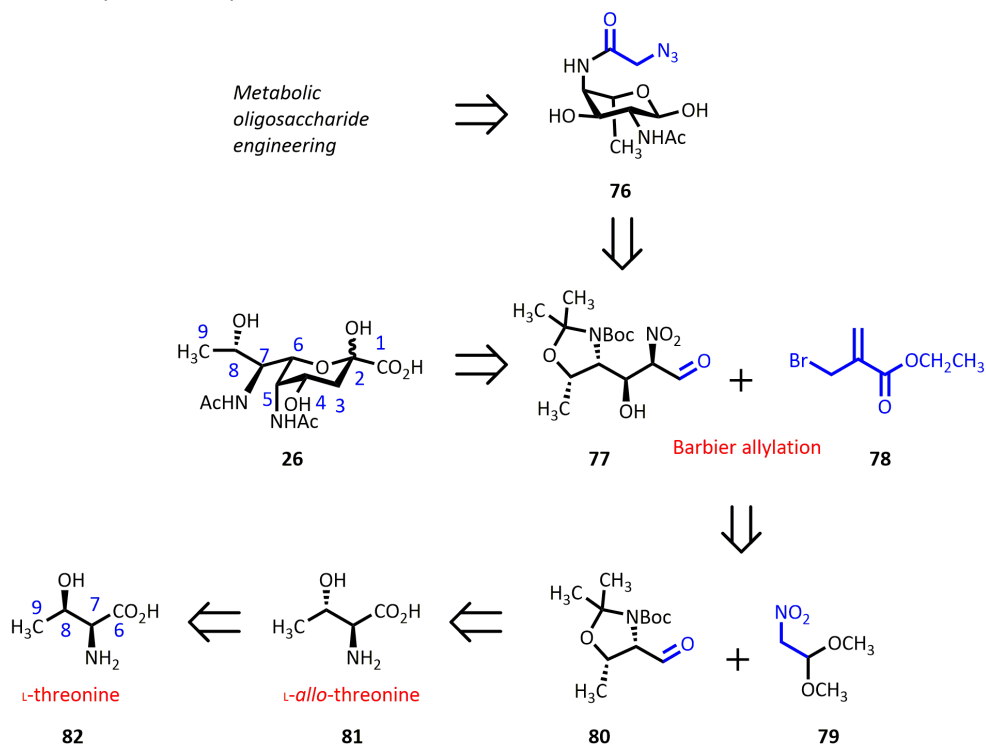
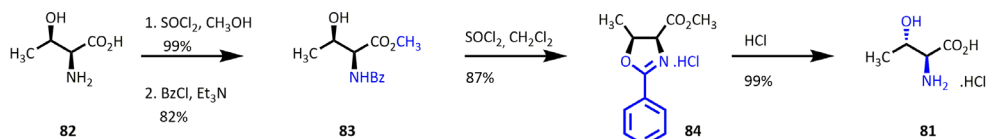


Figure 1. Retrosynthetic analysis of pseudaminic acid **26** and Pse azido precursor **76**.

The key step in this proposed route would be the Henry reaction between a nitro **79** and α -aminoaldehyde moiety **80**. The corresponding nitro-alcohol **77** can then be further functionalised, in the second phase, to Pse with a Barbier allylation and ozonolysis. The α -aminoaldehyde **80**, a Garner aldehyde derivative, needed for the Henry reaction can be synthesised from L-threonine **82** via the L-allo-threonine **81** intermediate. The nitro counterpart for the Henry reaction can be synthesised in a two-step procedure. The initial goal of the project, to synthesise chemical probes for metabolic Pse labelling is suitable via this route. The orthogonality of the amine functionalities (NBoc protection and a NO_2/N_3 masked amine) should result in controlled selective instalment of the azido-acetamide groups. As explained in **chapter 2**, other bioorthogonal groups such as cyclopropene or isonitrile also

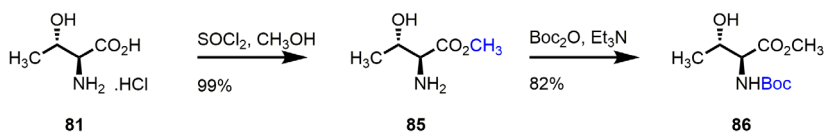
are worth considering. Orthogonal protection groups during the synthesis allow for easy selective modification of our targets. The orthogonal protection groups allow us to synthesise a variety of chemical probes from the same intermediate. This might be of interest since some chemical probes might interfere with the biological pathways.

L-*allo*-Thr **81** would be a very expensive starting material to employ for a large scale synthesis of Pse **26** at €400,- per gram. L-Thr **82** on the other hand is a cheap starting material at €13,- per gram. The first step in the route is therefore to invert the stereochemistry of the alcohol moiety in L-Thr **82** (Scheme 12).



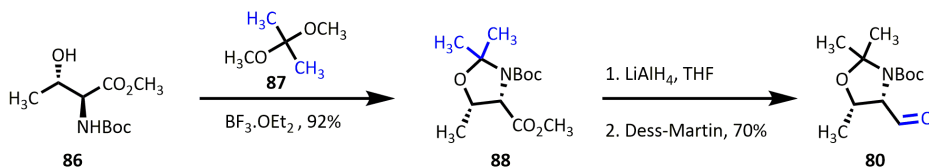
Scheme 12. Synthesis of L-*allyl*-threonine.HCl **82** from L-threonine.HCl **81**.

For the conversion of L-Thr **82** into L-*allyl*-Thr **81** the carboxylic acid was protected as a methyl ester and the amine with a benzoyl group according to standard procedures.^{29,30} Alcohol **83** was inverted via a two-step method.³¹ First thionyl chloride was added to give a cyclic oxazole intermediate **84** with inversion of stereochemistry at C-3. When the oxazole intermediate **84** was poured into acid, L-*allyl*-Thr **81** was obtained in quantitative yield. Overall L-*allyl*-Thr **81** was obtained in gram-quantities over 4 steps in excellent yields and employing only one purification method, either crystallisation or column chromatography. The total synthesis continued with two more protection steps (Scheme 13).



Scheme 13. Synthesis of BocHN-L-*allyl*-Thr-OCH₃.

Carboxylic acid **81** was reprotected as a methyl ester and amine **85** was Boc-protected using Boc anhydride.^{29,32} An isopropylidene ring was installed on the alcohol **86** (scheme 14). The rigid ring system in combination with the Boc-carbamate resulted in rotamers as seen by NMR.



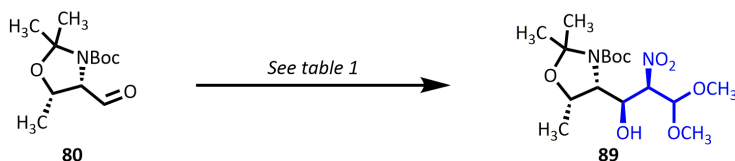
Scheme 14. Synthesis of Garner aldehyde derivative **80**.

Methyl protected carboxylic acid **88** was converted into an α -aminoaldehydes **80** via a two-step procedure involving ester reduction and reoxidation of the alcohol intermediate to the aldehyde. To

Total Synthesis towards Pseudaminic Acid Probes

preempt stability issues with the α -chiral aldehyde **80**, the alcohol was stored in larger quantities while the α -aminoaldehyde was always directly used after synthesis. Overall the Garner aldehyde derivative **80** was obtained in 36% overall yield in 9 steps.

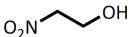
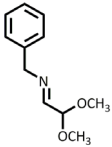
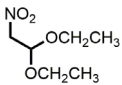



With **80** in hand, the key nitro-aldol reaction could be explored. For our target, Pse, the stereoselectivity of this reaction needs to be tuned to yield a *syn* alcohol-amino/nitro motif. This motif corresponds to the C5-amino and C5-ring oxygen in Pse. The nitro-aldol reaction, also known as the Henry reaction, is a powerful reaction discovered in the late 19th century by L. Henry (Scheme 15).³³⁻³⁴ Over the last decades numerous synthetic transformations have been employed using the Henry reaction as a key step.³⁵ Also α -aminoaldehydes have been successfully used in the nitro-aldol reaction (*vide infra*).



Scheme 15. Generic reaction scheme for the nitro-aldol reaction with Garner aldehyde derivative.

Table 1. Overview of performed attempts in the nitro-aldol reaction with Garner aldehyde derivative.

| Reagents | Entry | Catalyst or additive | solvent | reaction time | yield (mixture) | T (°C) | REF |
|---------------|------------|----------------------------|--------------------------------|---------------|-----------------|------------|---------------|
| 79 | 89a | 93 3M in THF | THF | 1-4 days | 17 % | -40 - 0 °C | ³⁶ |
| | 89b | LiOH | THF-H ₂ O (3:2) | 1-4 days | 30% | 0 °C | |
| | 89b | NaOH | THF-H ₂ O (3:2) | 1-4 days | 30% | 0 °C | |
| | 89c | LiOH | dioxane-H ₂ O (3:2) | 16 h | 9% | 0 °C | |
| | 89d | wt-HbHNL | TBME / Phosphate buffer | | - | rt | ³⁷ |
| 90 | 90a | LiOH | THF-H ₂ O (3:2) | 1-4 days | 13% | 0 °C | ³⁸ |

| | | | | | | | |
|--|------------|--|-------------------------------|--------|--------|------|---------------|
|  91 | 91a | LiOH | THF-H ₂ O (3:2) | 3 days | traces | rt | |
| | 91b | cetyltrimethyl ammonium chloride | NaOH (0.25 M) | 3 days | - | rt | ³⁹ |
| | 91c | LiOH | THF-H ₂ O (3:2) | 16 h | - | 0 °C | |
| | 91d | LiOH | THF-H ₂ O (3:2) | 16 h | - | 0 °C | |
|  92 | 92a | NbCl ₃ (DME) | THF | | - | rt | ⁴⁰ |
|  90 | 90b | hydroxynitrile lyase from <i>He- vea brasiliensis</i> | TBDME | 4 days | - | rt | ⁴¹ |
|  79 | 89e | cetyltrimethyl ammonium chloride | NaOH (0.25 M) THF | | | | ³⁹ |
|  79 | 89f | Al ₂ O ₃ | THF | | | | ⁴² |
|  79 | 89g | <i>tert</i> -butyl ammonium fluoride | THF | | - | | ⁴³ |

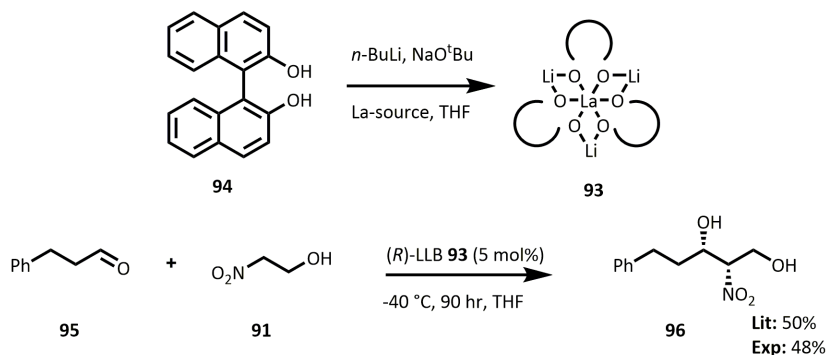
*conversion is consumption of starting material as evaluated by TLC and NMR. Yields are reported when isolated products were purified and characterised. - = not determined/analysed. rt = room temperature.

We started, inspired by Shibasaki et al, with a catalytic Henry reaction using a lanthanum-lithium binol complex^{93,36-38} A set of different rare earth metal catalyst have been designed and tested by Shibasaki and co-workers for the nitro-aldol reaction.³⁹ The lanthanum-lithium binol complex **93** (Table 1, entry 1) was for our substrate the best option since this catalyst showed good *syn*-selectivity in high yield

Total Synthesis towards Pseudaminic Acid Probes

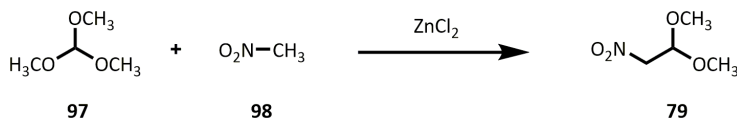
on a similar substrate in the work of Shibasaki.^{37, 40} We prepared catalyst **93** ourselves and tested it in a simple nitro-aldol reaction (Scheme 16). For the preparation of the catalyst two lanthanum sources can be used, $\text{LaCl}_3 \cdot 7\text{H}_2\text{O}$ or $\text{La}(\text{O}-i\text{-Pr})_3$. Both substrates were equally selective but the former lanthanum source is cheaper.

The synthesis of this lanthanum-lithium complex **93** was successfully performed as showed by a test reaction between an aldehyde **95** and nitroethanol **91** (see experimental section for detailed synthesis). Other rare earth metal catalyst were not investigated since catalyst **93** clearly showed the best results in the desired *syn*-stereochemistry.⁴¹



Scheme 16. Show-case example of lanthanum-catalysed nitro-aldol reaction.

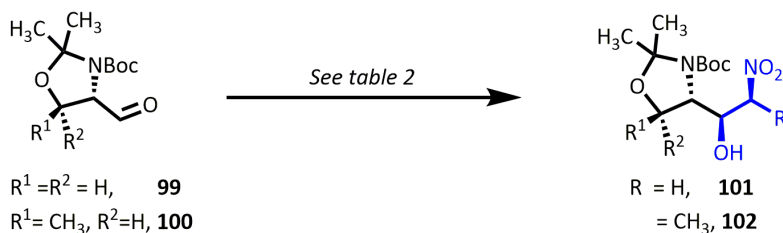
The nitro-reagent **79** was made in a one-step procedure from trimethylformate **97** and nitromethane **98** and distilled before use as described by Royer et al., (Scheme 17).⁴² The distillation reaction was a low yielding step but gave enough material to perform the Henry reactions. Highly pure Zn-catalyst (dried after refluxed in thionyl chloride⁴³) or higher catalyst loading did however not improve the reaction yield. For the synthesis of diethoxynitroethane the same procedure was followed. Initial screening of the Henry reaction was performed with dimethoxynitromethane **79** but often the procedure was compared with diethoxynitroethane. The Henry reaction proved to be rather challenging with our aldehyde. The Shibasaki procedure using the lanthanum catalyst **93** provided, after extensive optimisation a poor 17% yield and was obtained as a mixture of diastereoisomers (entry 89a).⁴⁰ Separation on a chiral column (Chiralpak IA column with mobile phase: hexane / 2-propanol) was possible but we decided to look for alternative, better yielding reactions. Furthermore, the reaction was not well suited for the large scale we aimed to synthesise this intermediate at. Another example of the Henry reaction, employing a simple inorganic base, was reported by Schmölzer et al. in 2013.⁴⁴



Scheme 17. Synthesis of nitro-acetaldehyde **79**.

They showed that inorganic bases such as LiOH in THF–water mixtures were able to perform the nitro-aldol reaction in excellent yield and scale on a diastereomer of the Garner aldehyde derivative (Scheme 18). Based on this procedure, we reproduced their experiment with our aldehyde **80** and 1.2–7 equivalents of the nitro reagent **79**. With cooling (0 °C) and long reaction times (varying between 1–4 days) we found an optimum after 4 days (entry 89b). An isolated yield of only 30% was however obtained of the desired product as an inseparable mixture of diastereomer. According to NMR, a 1:1 ratio of diastereomers was obtained of two products. These two products corresponds to *anti-syn* **89** and *anti-anti* **89**. Changing the ratio between THF and water did not improve the yield compared to the standard ratio of THF – water 3:2. THF-water froze if the reaction was performed below -5 °C. Dioxane, which does not freeze at -5 °C like THF, also did not improve the reaction yields (entry 89d). The addition of more base (LiOH) 1.0 or 2.0 equivalents also did not improve the reaction outcome. Changing the bases from LiOH to NaOH or KOH did not improve the reaction conversion or yield (entry 89c). Organic bases such as Et₃N or DBU gave no conversion at all. The Henry reaction with inorganic bases was also performed in deuterated solvents inside a NMR tube. We clearly observed fast product formation in the first hours but it clearly reached an equilibrium at a later stage. When we changed the nitro-reagent **79** to the diethoxynitroethane **90** we were able to isolate product in poor yield (entry 90a). Other nitro-reagents such as nitroethanol **91** or THP-protected nitroethanol hardly gave conversion to the desired products (entry 91a–d). An old literature procedure with a benzylimine acetal **92** and stoichiometric amounts of niobium trichloride was also investigated.⁴⁵ The proposed product an amino-alcohol was however never formed (entry 92a).

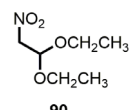
Nitro-aldol condensation reactions have also been catalysed using cetyl trimethylammonium chloride, aluminium or *tert*-butylammonium fluoride, but this did not work in our hands (entry 89e–g).^{46–48} Inspired by nature, we also performed an enzymatic reaction which is able to perform nitro-aldol reaction.^{49–50} For this enzymatic transformation we used hydroxynitrile lyase from *Hevea brasiliensis*. This class is generally used for the formation and cleavage of cyanohydrins. At pH 7 in phosphate buffer we performed first a test reaction with benzaldehyde and diethoxynitroethane but only obtained 1% yield after extensive column chromatography. With aldehyde **80** we never observed product formation (entry 90b).



Scheme 18. Generic reaction scheme for the nitro-aldol reaction with Garner aldehyde **99** and an isomer of Garner aldehyde derivative **100**.

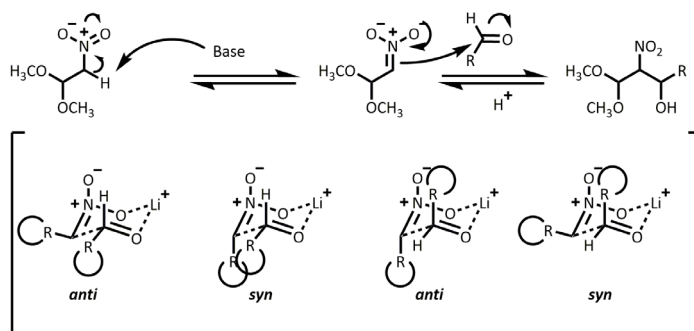
Total Synthesis towards Pseudaminic Acid Probes

Table 2. Overview of performed attempts in the nitro-aldol reaction with Garner aldehyde and isomers thereof.

| Reagents | Catalyst or additive | Solvent | Reaction time | conversion | yield (mixture) | T (°C) | REF |
|--|----------------------|---------------------------------|---------------|------------|--|--------|---------------|
|  90 | LiOH | THF - H ₂ O (3:2) | 16 h | - | 101 : 19 % 102 : 37 % | - 0 °C | ³⁸ |

*conversion is consumption of starting material as evaluated by TLC. Yields are reported when isolated products were purified and characterised. - = not determined/analysed.

To test the reproducibility of the literature procedure of Schmölzer et al. we synthesised the isomer of the Garner aldehyde derivative and the Garner aldehyde (Scheme 18).⁵¹ For compound **102** we could isolate the desired product in 37% yield (Table 2). The Henry reaction with Garner aldehyde **101** yielded 19% of the desired product. These results clearly indicated the potential of the Henry reaction on Garner aldehyde isomers but also clearly stated the importance of the correct stereochemistry of these isomers.



Scheme 19. Proposed reaction mechanism nitro-aldol reaction.

All steps in the nitro-aldol reaction are completely reversible. After deprotonation with a base the resonance stabilised anion easily performs the aldol reaction. If lithium ions are present a chair-like conformation is obtained (Scheme 19). Different conformations are possible depending a lot on the sterics of the R-groups. If the Garner aldehyde derivative **80** derived from L-*allo*-Thr **81** was used we obtained the product in poor yield. The corresponding Garner aldehyde derivative **100** derived from L-Thr **82** gave however satisfying outcome.⁴⁴

The installation of an amino-alcohol functionality onto α -aminoaldehyde **80** is notoriously difficult. A Cram-chelate controlled organometallic addition on α -aminoaldehyde **80** is for example not possible.²⁵ The desired *anti*-addition cannot be controlled due to steric effects and therefore will result in mainly *syn*-addition (Figure 2). *Anti*-addition is achieved when nucleophilic attack happens from the *si* face in conformation **B**. The *N,O*-isopropylidene moiety, C-4 methyl, and the carbonyl group have influence on the conformation, resulting in rotamer A as the preferred conformation.

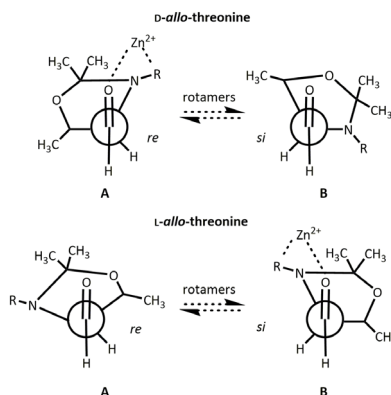


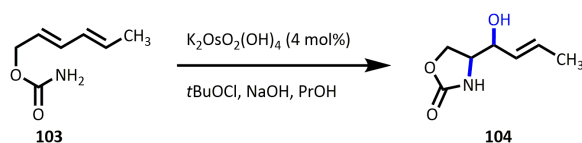
Figure 2. Schematic scheme for the Cram-chelate organometallic addition on Garner aldehyde isomers.

On the enantiomer aldehyde **100** such a Cram-chelate controlled organometallic addition is possible as showed by Seeberger and co-workers.²² This method was used for the *de novo* synthesis of legionaminic acid in 2015.

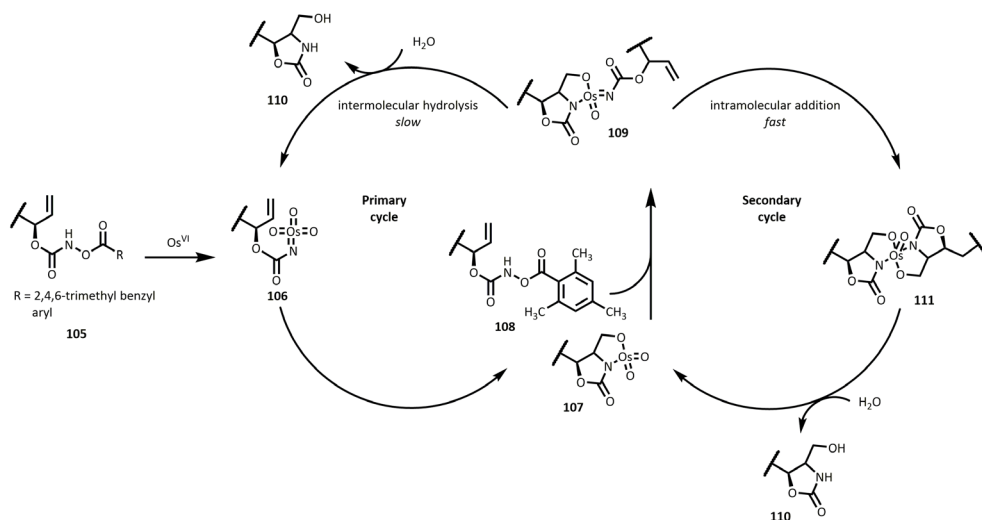
Tethered aminohydroxylation

After all the above described efforts to perform a successful nitro-aldol reaction we started looking for other synthetic reactions to install an amine-alcohol motive in our molecule. A possible synthetic manipulation to install an amine and alcohol moiety would be the tethered aminohydroxylation.⁵² The tethered aminohydroxylation is a variant of the aminohydroxylation.⁵³ In the aminohydroxylation, first described in 1996 by Sharpless and co-workers, alkenes are turned into vicinal amino alcohols. Since it also occurred in an enantioselective fashion this transformation is known as the Sharpless asymmetric aminohydroxylation. One of the limitations of this reaction is the lack of regiochemistry on more complicated unsymmetrical alkenes. To circumvent this potential problem Donohoe et al., tethered the source of nitrogen to an achiral allylic alcohol **103**.⁵⁴ This method achieved regiochemistry of the amino alcohol motif in the product **104** (Scheme 20).

The mechanism for the tethered aminohydroxylation was proposed by Donohoe and co-workers in 2012 (Scheme 21).⁵⁵ It is in line with the proposed mechanism for the Sharpless asymmetric dihydroxylation.⁵⁶ In the first step an trioxoimido osmium complex **106** is formed between the osmium salt and the *N*-oxy-carbamate **105**. An intramolecular [3+2] cycloaddition follows between the osmium complex and the alkene. The formed cycloadduct **107** can undergo another oxidation with new *N*-oxy-carbamate **108**. If hydrolysis is fast, product **110** is formed immediately. It is, however, also possible that a second cycloaddition happens with a second alkene. The secondary cycle strongly indicates that the addition of asymmetric ligands to improve the enantioselectivity will not help since the osmium metal remains hindered during the catalytic cycle.



Scheme 20. Tethered aminohydroxylation of allylic carbamates.

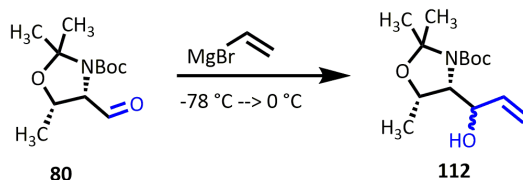


Scheme 21. Proposed reaction mechanism of the tethered aminohydroxylation.

For our system the tethered aminohydroxylation would be helpful to install the amino-alcohol moiety and circumvent the need for the troublesome Henry reaction. To build the tethered amino source a Grignard reaction on aldehyde **80** was performed. Unfortunately, no selectivity was obtained in this step. Moreover, the diastereoisomers of **112** could not be separated at this stage.

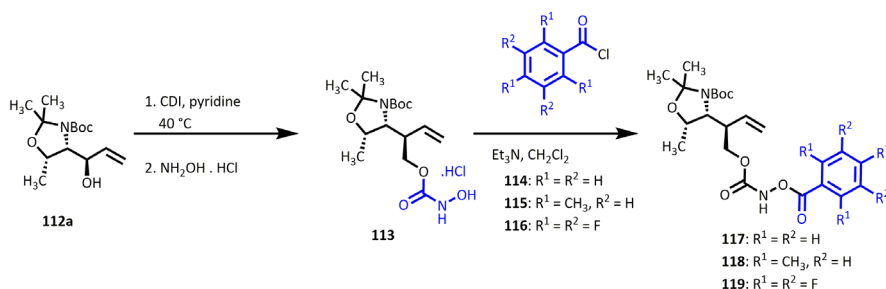
The vinyl alcohol **112** was further functionalised towards the tethered amine source. A *N*-oxy-carbamate **113** was installed using hydroxylamine hydrochloride after CDI activation. This one-pot reaction was performed on gram-scale with variable success (56% yield). One of the reasons was the high polarity of the product, which caused problems especially in column chromatography.

A few aryl chlorides **114-116** were coupled to the *N*-oxy-carbamate linker **113** to explore the effect on the efficiency in the tethered aminohydroxylation. In 2007 Donohoe and co-workers reported that the penta-fluor aryl moiety **116** was the best substrate for the tethered aminohydroxylation.⁵⁷ The yield with this group improved significantly and the osmium catalyst loading could be lowered to 1 mol% without the loss of diastereoselectivity. In later work by Donohoe et al, 2,4,6-trimethylbenzoyl chloride **115** showed the best reaction yields on the most complex structures.⁵⁸ A clear explanation



Scheme 22. Grignard reaction on Garner aldehyde derivative **80**.

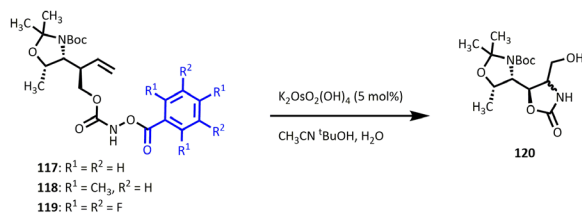
for this observed behaviour was however not provided. Benzoyl chloride **114** will result in an aryl substituent without extra electronic potential. The 2,4,6-trimethylbenzoyl chloride **115** is even more bulky than the benzoyl variant **114**. Last, penta-fluor benzoyl chloride **116** was coupled to explore the effect of its electronic withdrawing character. In our hands the synthesis of both benzoyl **114** and 2,4,6-trimethylbenzoyl *N*-hydroxy carbamates **115** went fine, yielding foamy white products. The installation of 2,4,6-trimethylbenzoyl **118** was performed in 84% yield compared to 77% yield for the benzoyl moiety **117**. The penta-fluor moiety **116** caused more problems. We therefore decided to focus on the two other substrates.



Scheme 23. Synthesis towards the tethered amine source **117-119**.

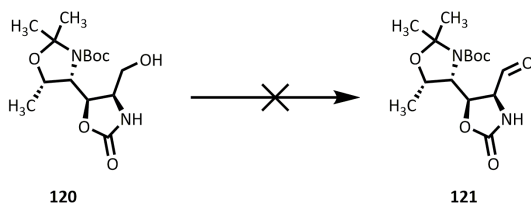
The tethered aminohydroxylation was performed on both substrates **117** and **118** (Scheme 24). In both cases addition was performed from the same side giving *syn*-products. Since both diastereoisomers were present in the reaction mixture we obtained again a mixture of diastereoisomers. At this point, however, partial separation was possible.

Next, we aimed for the Barbier allylation.²⁸ The Barbier allylation is widely explored in the synthesis of several naturally occurring higher carbon sugars.^{10, 59-60} One of the main advantages is that the Barbier allylation can be performed on unprotected carbohydrates. As a consequence, this reaction can be performed in aqueous solvents and is therefore relative safe and simple. The original procedure as developed by Barbier used magnesium, nowadays other metals such as Sn, In, or Zn are mostly used. The Ito route also used this chain-elongation reaction in their synthesis towards pseudaminic acid.⁴ A commercially available bromomethylacrylate ester was used in combination with indium. Two isomers, *erythro* and *threo* were obtained with a slight preference for *erythro*. Attempts to improve the selectivity with Lewis acids and chiral ligands did not help. We envisioned a same lack of selectivity for



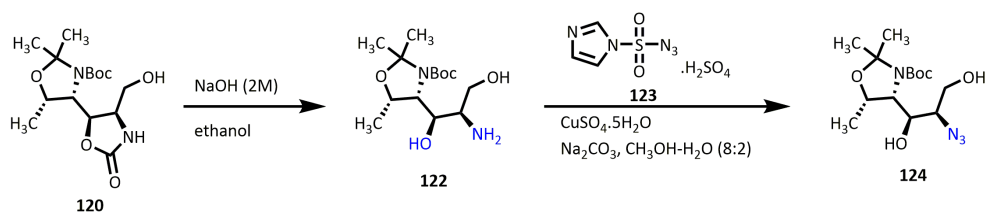
Scheme 24. Synthesis of cyclic carbamate using the tethered aminohydroxylation.

our building block in the Barbier allylation. To perform the Barbier allylation, the primary alcohol **120** needed to be oxidised to aldehyde **121** (Scheme 25). A variety of oxidation reactions were evaluated to achieve this. Dess-Martin periodane reagent only gave trace amounts of the product, leaving the majority of the starting material untouched.⁶¹⁻⁶² Swern oxidation with DMSO and oxalyl chloride gave degradation during the work-up. Oxidation with TEMPO, IBX, or tetrapropylammonium perruthenate (TPAP) resulted in no conversion of the starting material. Acetylation of the primary alcohol also led to *N*-acetylation. This was not seen as a potential problem, however, the attempted selective removal of the *O*-acetyl groups after the *N*-acetylation gave back the initial starting material. In literature only one example was found where this motif of a primary alcohol next to a cyclic carbamate was oxidised.⁶³ In our hands this procedure also gave product formation but never reached full consumption of starting material. Since the cyclic carbamate need to be removed at a later stage anyway we decided to switch the order of reactions. First the cyclic carbamate was removed and then we focussed on the oxidation of the primary alcohol. The synthesis continued with the hydrolysis of the cyclic carbamate under basic conditions. The conversion was smoothly completed in two hours and the crude mixture **122** was submitted to the next step without further purification.



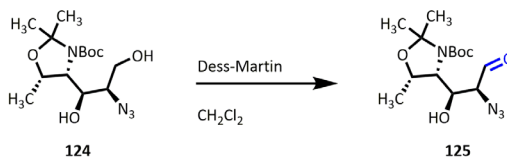
Scheme 25. Proposed oxidation of the primary alcohol.

To discriminate between the two amine functionalities we wanted to protect the amine. The protection group should be acid stable and small, in order not to interfere with later modifications. We therefore decided to install an azide on this position. Deprotection is easily performed under hydrogenation yielding an primary amine which can later be modified towards an azido-acetyl group which can be used for metabolic labelling. The recently developed diazotransfer reagent **123** by Stick et al. was used to install the *syn*-azido alcohol **124** (scheme 24).⁶⁴⁻⁶⁶ Different imidazole sulfonyl diazotransfer reagents were developed in recent years but the imidazole sulfonyl diazotransfer sulfate salt **123** proved to be efficient, shelf-stable, and easy to prepare. After the removal of the cyclic carbamate and conversion of the amine into the azide we turned back to the oxidation of the primary alcohol. Dess-Martin periodane was now, indeed able to oxidise alcohol **124** selectively towards the aldehyde in the presence of the secondary alcohol (scheme 25).



Scheme 26. Synthesis of *syn*-azido alcohol **124**.

Time limitations and the availability of material prohibited us to continue at this stage. Further investigations are needed to accomplish a successful total synthesis of Pse or its metabolic hexose precursor. For metabolic labelling studies a smart design of the probes is desired. If a variety of chemical probes will be evaluated one common synthetic intermediate would be ideal. Furthermore orthogonal protection groups are preferred in order to selectively modify functional groups and to synthesise the chemical probes.



Scheme 27. Oxidation of the primary alcohol **124** with Dess-Martin periodane.

Experimental section

General

For reactions conducted under anhydrous conditions glassware was dried overnight in an oven at 150 °C and was allowed to cool in a desiccators over anhydrous KOH. Anhydrous reactions were carried out under nitrogen. Reagents/solvents for anhydrous reactions were obtained from a Pure Process Technology Glass Contour solvent purification system (SPS) (CH_2Cl_2 , THF, diethyl ether, toluene), with 3 Å molecular sieves (DMF, CH_3CN , CH_3OH , toluene, pyridine) or Na_2SO_4 (acetone). Commercially acquired chemicals were used without further purification unless stated otherwise.

Flash Column Chromatography

Flash column chromatography was carried out according to standard procedures using silica gel (40-63 μm mesh) from Screening Devices, Amersfoort.⁶⁰ Prior to loading an appropriate eluent was chosen by running a TLC of the crude mixture which was to be purified. Loading was performed with celite or small volumes of the eluent.⁶⁷

Thin Layer Chromatography (TLC)

Thin layer chromatography was carried out on and visualised using UV (254 nm), Hanessian stain, sulfuric acid, KMnO_4 , I_2 or ninhydrin stain. The quoted R_f values are rounded to the nearest 0.05.

Nuclear Magnetic Resonance (NMR)

^1H - and ^{13}C -NMR were run on a Bruker Avance 400 NMR spectrometer (400 MHz). Chemical shifts are reported in parts per million (ppm) with reference to tetramethylsilane (TMS) as internal standard. Coupling constants (J) are reported in Hertz (Hz), rounded to the nearest 0.5 Hz. Signal assignment was made from unambiguous chemical shifts and COSY, HSQC, HMBC, and/or DEPTQ.

High Resolution Mass Spectroscopy (HRMS)

High-resolution ESI mass analyses were recorded on a Orbitrap high-resolution mass spectrometer. The orbitrap high-resolution mass spectrometer system consisted of a ESI ion source (Ion-Sense, Saugus, USA) coupled to an Exactive high resolution MS system (Thermo Fisher Scientific, San Jose, CA, USA). The MS was calibrated daily using ProteoMass™ LTQ/FT-Hybrid ESI (positive and negative mode) Cal Mix (Sigma Aldrich), which is applicable for the m/z range 100–2000. XCalibur software (version 2.1) was used for instrument control, data acquisition and data processing.

Infrared Spectroscopy (IR)

Infrared spectra were recorded on a FT-IR spectrometer apparatus (Bruker) with a universal ATR accessory. Both solid and oily compounds were recorded. FT-IR signals are reported in wavenumbers (cm^{-1}).

Optical Rotation Measurements

Optical rotation was recorded on a Perkin Elmer 241 polarimeter. The polarimeter temperature was allowed to equilibrate at 23 °C for 3 hours prior to the first measurements. Products (50.0 mg) were dissolved in HPLC grade CH_3OH using a 5 ml graduated flask to give a final volume of 5 ml. Optical rotation was recorded using the sodium D line (589 nm) at 23 °C and are given in units of $10^{-1} \text{ deg dm}^2 \text{ g}^{-1}$.

Cooled reactions

All reactions below room temperature were cooled using a cryostat apparatus using acetone as cooling agent.

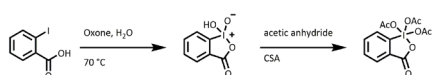
Diazotransfer reagent



1H-Imidazole-1-sulfonyl azide.H₂SO₄. The imidazole sulfonyl diazotransfer reagent was prepared according to literature procedures.⁶⁴⁻⁶⁶

Synthesis of Dess Martin periodane

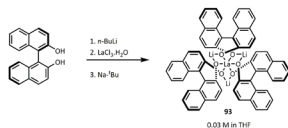
Two literature procedures were followed for the synthesis of Dess Martin periodane.⁶¹⁻⁶²



Oxone (37.20 g, 60 mmol, 1.3 eqv) was dissolved in H₂O. To the solution iodobenzoic acid (12.90 g, 20 mmol, 1 eqv) was added and the resulting mixture was stirred at

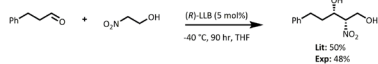
70 °C for 3 h. The reaction mixture was cooled to 5 °C and the resulting precipitate was washed with water and acetone. The white solid was concentrated *in vacuo* and dried at high vacuum overnight.

The white solid was dissolved in acetic anhydride (100 ml) and equipped with a calcium chloride tube. Camphor sulfonic acid (100 mg) was added and the suspension was stirred at 80 °C. The reaction mixture was completely dissolved after 1.5 hour, stirred for 30 min, and cooled to 5 °C. The white precipitate was filtered and washed with anhydrous ether to obtain the Dess Martin reagent which was stored at -20 °C under argon. MP: 131 °C.



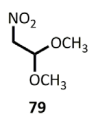
Lanthanum lithium catalyst.⁴⁰ *R*-BINOL (500 mg, 1.75 mmol) was dissolved in anhydrous THF (9 ml) and charged to an oven-dried roundbottomflask. The resulting mixture was cooled to 0 °C before *n*-Buli (2.19 ml, 3.51 mmol) was added. Meanwhile, LaCl₃·H₂O (240 mg, 0.65 mmol) was dissolved/suspended in THF and sonicated for 30 min. The

binol-lithium mixture was slowly added to the lanthanum suspension over 3 min and was stirred at room temperature for 1 hour. Na^tBu in THF (0.50 M) was added over 3 min and stirring was continued 18 h before it was heated to 50 °C and stirred for two days. The orange mixture was allowed to cool to room temperature without stirring and stored in the fridge. The supernatant of the lanthanum-lithium catalyst **93** was used as a 0.03 M solution in THF.



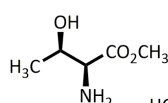
Test reaction. 2-Nitroethanol and 3-phenylpropanal were distilled before use. A roundbottom flask was charged with (*R*)-LLB (5 mol%, 0.03 M in THF) and THF (5 ml) and stirred at

-40 °C. 2-nitroethanol (0.06 ml, 0.84 mmol, 1.1 eqv) was added after 30 min. After 30 min, 3-phenylpropanal (0.01 ml, 0.76 mmol, 1.0 eqv) was added and the reaction mixture was stirred for 90 h at -40 °C. The reaction mixture was quenched with HCl (2 ml) and slowly warmed to room temperature. The aqueous phase was extracted three times with EtOAc (3 × 20 ml) and the combined organic phases were washed with brine (3 × 20 ml), dried over Na₂SO₄, and concentrated *in vacuo*. The product was recrystallised from hexane:ether (1:1, 20 ml) to give 90 mg of the product (47%).

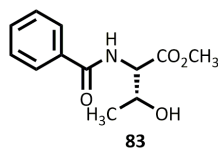


1,1-Dimethoxy-2-nitroethane⁴². To a 500 ml roundbottom flask, equipped with a magnetic stirring bar, 5-cm Vigreux column, column head, Claisen distilling head, and thermometer, were charged trimethylorthoformate (**97**, 100 ml, 0.6 mol), nitromethane (**98**, 180 gram, 2.95 mol) and anhydrous ZnCl₂ (5 gram). The reaction mixture was heated at 90 °C and stirred overnight. Methanol (5 ml) was collected whereafter the remaining mixture was filtered at room temperature. The brown liquid was distilled using the same distillation setup. At reduced pressure, nitromethane (bp ca 30 °C at 35 mmHg) and 1,1-dimethoxy-2-nitroethane (**79**, bp ca 130 °C at 1 mmHg, 3.63 gram, 6% yield) were obtained as colourless liquids. ¹H NMR (400 MHz, CDCl₃) δ 4.92 (1H, t, *J* = 5.6, CH), 4.40 (2H, d, *J* = 6.0, CH₂), 3.35 (6H, s, 2 × OCH₃). ¹³C NMR (100 MHz, CDCl₃) δ 100.4 (CH), 76.0 (CH₂NO₂), 54.5 (OCH₃). The same procedure was followed to make **1,1-diethoxy-2-nitroethane (90)**, 18 ml, 18% yield). ¹H NMR (400 MHz, CDCl₃) δ; 4.99 (1H, t, *J* = 5.8, CH), 4.35 (2H, d, *J* = 6.0, CH₂), 3.60-3.45 (4H, m, 2 × OCH₂), 1.06 (6H, t, *J* = 8.0, 2 × CH₃). Anhydrous ZnCl₂ was prepared from refluxing ZnCl₂ in thionyl chloride.⁴³

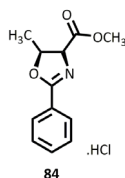
Garner's Aldehyde Derivative Synthesis



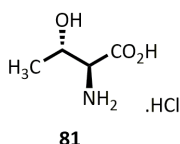
(2S,3R)-Methyl 2-amino-3-hydroxybutanoate hydrochloride (H₂N-L-Thr-OCH₃.HCl). To a roundbottom flask with anhydrous methanol (50 ml), thionyl chloride (4.83 ml, 67.0 mmol) was added dropwise over 5 min. The reaction mixture was cooled to 0 °C after which L-threonine (**82**, 8 g, 67 mmol) was added over 5 min in several smaller portions. The reaction mixture was stirred for 1 hour at 0 °C, followed by refluxing for 4 h and stirring overnight at room temperature. A water aspirator was used to concentrate the reaction mixture under reduced pressure. The resulting viscous oil (11.36 g, 99%) was used without any further purifications. *R*_f = 0.23 (3:1:2; CH₂Cl₂:CH₃OH:EtOAc). ¹H NMR (400 MHz, CDCl₃) δ 4.28-4.20 (1H, m, β-H), 3.92 (1H, s, α-H), 3.37 (3H, s, CH₃), 1.34 (3H, d, *J* = 6.7, CH₃). ¹³C NMR (100 MHz, CDCl₃) δ 169.5 (C=O), 66.3 (CHOH), 59.8 (CHNH₂), 53.9 (OCH₃), 20.7 (CH₃)



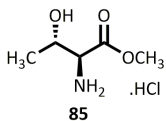
(2S)-Methyl 2-benzamido-3-hydroxybutanoate (83, BzHN-L-Thr-OCH₃). Methyl ester threonine (1.4 g, 8.4 mmol) was dissolved in CH₂Cl₂ (20 ml) and cooled to 0 °C. With a dropping funnel, benzoyl chloride (0.87 ml, 7.57 mmol) dissolved in CH₂Cl₂ was added over 10 min. The reaction mixture was stirred for 1 hour at 0 °C before it was stirred overnight at room temperature. The reaction mixture was diluted with water (50 ml) and extracted with CH₂Cl₂ (3 × 20 ml). The organic layer was washed with NH₄Cl (3 × 20 ml) and Na₂HCO₃ (3 × 20 ml), dried over Na₂SO₄, and concentrated *in vacuo*. Flash column chromatography (25% EtOAc in petroleum ether (v/v)) gave protected threonine (**83**, 0.50 g, 25%) as white crystals. *R*_f = 0.31 (95:5; CH₂Cl₂:CH₃OH). ¹H NMR (400 MHz, CDCl₃) δ 7.88-7.85 (2H, m, 2 × CH_{ar}), 7.57-7.53 (1H, m, CH_{ar}), 7.49-7.45 (2H, m, 2 × CH_{ar}), 6.98-6.96 (1H, d, *J* = 8.8, NH), 4.87-4.84 (1H, dd, *J* = 8.8, 2.4, βH), 4.50-4.48 (1H, m, α-H), 3.80 (3H, s, CH₃), 2.41-2.40 (1H, d, *J* = 4.8, OH), 1.33-1.31 (3H, d, *J* = 6.4, CH₃). ¹³C NMR (100 MHz, CDCl₃) δ = 171.6 (C=O), 167.8 (C=O), 133.8 (C_{ar}), 131.9 (C_{ar}), 128.6 (C_{ar}), 127.2 (C_{ar}), 68.3 (βC), 57.5 (αC), 52.0 (OCH₃), 20.1 (CH₃). **MS**: found 260.08 [M+Na]⁺, calculated for [C₁₂H₁₅NO₄ + Na]⁺ 260.09.

**(4R,5S)-Methyl****5-methyl-2-phenyl-4,5-dihydrooxazole-4-carboxylate**

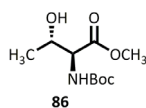
hydrochloride 84. Amide (**83**, 0.50 mg, 2.1 mmol) was dissolved in anhydrous CH_2Cl_2 and cooled to 0 °C. Thionyl chloride (1.4 ml, 19.33 mmol) was slowly added to the solution over 10 min. The reaction mixture was stirred overnight at 0 °C before it was concentrated *in vacuo* and dissolved in CH_2Cl_2 . The reaction mixture was washed with Na_2HCO_3 (2 × 10 ml), and brine (2 × 10 ml), dried over Na_2SO_4 , and concentrated *in vacuo* to give oxazole **84** as an oil (350 mg, 65%). R_f = 0.75 (90:10; CH_2Cl_2 : CH_3OH). $^1\text{H NMR}$ (400 MHz, CDCl_3) δ = 8.09-8.08 (2H, m, 2 × CH_{ar}), 7.52-7.55 (1H, m, CH_{ar}), 7.44-7.43 (2H, m, 2 × CH_{ar}), 5.14-5.09 (1H, m, β -H), 5.03-5.01 (1H, m, α -H), 3.85 (3H, s, OCH_3), 1.43-1.41 (3H, d, J = 6.3, CH_3). $^{13}\text{C NMR}$ (100 MHz, CDCl_3) δ = 168.0 (C=O), 167.3 (C=N), 134.8 (C_{ar}), 131.7 (C_{ar}), 130.8 (C_{ar}), 129.8 (C_{ar}), 70.6 (CHCH_3), 58.5 (CH), 54.9 (OCH_3), 17.9 (CH_3).

**L-allo-Threonine hydrochloride (81, $\text{H}_2\text{N-L-allo-Thr-OH} \cdot \text{HCl}$).**

Oxazole (**84**, 350 mg, 1.37 mmol) was added to a stirred solution of aqueous hydrochloric acid at room temperature. The reaction mixture was refluxed for 3 h, stirred at room temperature overnight and diluted with H_2O (10 ml). The aqueous layer was washed with diethylether (3 × 10 ml) and concentrated *in vacuo* to give a brown precipitate (**81**, 212 mg, 99%) which was used without any further purification. $^1\text{H NMR}$ (400 MHz, CDCl_3) δ = 4.35-4.25 (1H, m, β -H), 4.10-4.09 (1H, d, J = 3.5, α -H), 1.28-1.26 (3H, d, J = 6.7, CH_3). $^{13}\text{C NMR}$ (100 MHz, CDCl_3) δ = 171.8 (C=O), 65.4 (CHOH), 59.7 (CHNH_2), 16.2 (CH_3). $[\alpha]_{\text{D}}^{23}$ +7.98 (c 0.01, H_2O).

**Methyl-L-allo-threonine.HCl (85, $\text{H}_2\text{N-L-allo-Thr-OCH}_3$).**

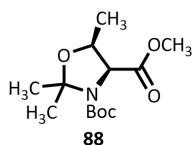
To a roundbottom flask with anhydrous methanol (20 ml), thionyl chloride (128 μl , 1.77 mmol) was added dropwise over 5 min. The reaction mixture was cooled to 0 °C after which L-allo-threonine (**81**, 212 mg, 1.37 mmol) was added over 5 min in several smaller portions. The reaction mixture was stirred for 1 hour at 0 °C, followed by refluxing for 4 h and stirring overnight at room temperature. A water aspirator was used to concentrate the reaction mixture under reduced pressure. The resulting viscous oil (**85**, 220 mg, 95%) was used without any further purifications. R_f = 0.23 (3:1:2; CH_2Cl_2 : CH_3OH : EtOAc). $^1\text{H NMR}$ (400 MHz, CDCl_3) δ = 4.33-4.31 (1H, m, β -H), 4.03-4.02 (1H, m, α -H), 3.37 (3H, s, OCH_3), 1.30-1.28 (3H, d, J = 6.6, CH_3). $^{13}\text{C NMR}$ (100 MHz, CDCl_3) δ = 168.8 (C=O), 66.6 (CH), 59.5 (CH), 53.7 (OCH_3), 18.8 (CH_3).

**Methyl (tert-butoxycarbonyl)-L-allo-threonine (86, $\text{BocHN-L-allo-Thr-OCH}_3$).**

Methyl-L-allo-threoninate (**85**, 2.99 g, 17.27 mmol) was dissolved in water: methanol (50 ml, 1:1) and NaHCO_3 (4.39 g, 52.26 mmol). Boc_2O (5.85 g, 26.80 mmol) was added slowly and a white cloudy suspension was observed. The reaction was left stirring overnight. After full completion of the reaction, the reaction mixture was concentrated and subsequently diluted with ethyl acetate (100 ml). Sulfate buffer (1M, 100 ml) was added to neutralise the sodium bicarbonate. The aqueous layer was extracted with ethyl acetate (3 × 30 ml) and the organic layers were combined, dried over Na_2SO_4 and concentrated under reduced

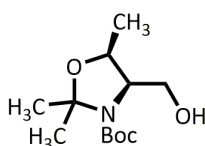
Total Synthesis towards Pseudaminic Acid Probes

pressure to obtain yellow oil **86** in 72% yield (2.98 g, 12.78 mmol). $R_f = 0.46$ (3:1; Hept.:EtOAc). $^1\text{H NMR}$ (400 MHz, CDCl_3) δ 5.46 (1H, d, $J = 7.6$, NH), 4.40 – 4.34 (1H, m, αH), 4.12 (1H, m, $J = 7.1$, βH), 3.77 (3H, d, $J = 1.4$, OCH_3), 1.48 (9H, s, $3 \times \text{CH}_3$), 1.19 (3H, d, $J = 6.4$, CH_3). $^{13}\text{C NMR}$ (100 MHz, CDCl_3) δ 171.0 (C=O), 156.1 (C=O), 80.5 (C_q), 69.0 (CH), 59.1 (CH), 52.4 (OCH_3), 28.2 ($3 \times \text{CH}_3$), 18.8 (CH_3). **MS**: found 256.11 $[\text{M}+\text{Na}]^+$, calculated for $[\text{C}_{10}\text{H}_{19}\text{NO}_5 + \text{Na}]^+$ 256.12.



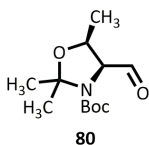
3-(*tert*-Butyl) 4-methyl (4*S*, 5*S*)-2,2,5-trimethyloxazolidine-3,4-dicarboxylate

88. BocHN-*L*-*allo*-threonine (2.98 g, 12.79 mmol) was dissolved in anhydrous toluene (50 ml) and camphor sulfonic acid was added (315.6 mg, 1.35 mmol). The pH was set to 4-5. Dimethoxypropane **87** was added slowly (4.4 ml, 35.91 mmol) and the reaction mixture was stirred overnight. After stirring overnight, the reaction mixture was quenched with triethylamine (1 ml), concentrated under reduced pressure and the procedure was repeated: the oil was dissolved in anhydrous toluene (50 ml), camphor sulfonic acid (30.5 mg, 0.13 mmol) was added, the pH was set to 5 and dimethoxypropane (4.4 ml, 35.91 mmol) was added. The reaction was left stirring overnight, isolated and purified by flash column chromatography (eluent: EtOAc: Heptane; 1: 8 (v/v) for fraction 1 to 19 and EtOAc: Heptane; 1: 4 (v/v) for fraction 20-24) was performed. A colourless oil in 40% yield (**88**, 1.47 g, 5.38 mmol) was obtained. $R_f = 0.45$ (3:1; Hept.:EtOAc). $^1\text{H NMR}$ (400 MHz, CDCl_3) δ 4.40 – 4.23 (2H, m, αH and βH), 3.70 (3H, s, OCH_3), 1.69 (6H, s, $2 \times \text{CH}_3$), 1.46 (9H, s, $3 \times \text{CH}_3$), 1.31 – 1.16 (3H, m, CH_3). $^{13}\text{C NMR}$ (100 MHz, CDCl_3) δ 170.5 (C=O), 151.0 (C=O), 94.4 (C_q), 80.7 (C_q), 71.4 (CHNBoc), 63.6 (CH), 51.7 (OCH_3), 28.2 ($3 \times \text{CH}_3$), 25.5 ($2 \times \text{CH}_3$), 15.4 (CH_3). **MS**: found 296.17 $[\text{M}+\text{Na}]^+$, calculated for $[\text{C}_{13}\text{H}_{23}\text{NO}_5 + \text{Na}]^+$ 296.15.



((4*R*, 5*S*)-2,2,5-Trimethyloxazolidin-4-yl)methanol

The Garner ester (**88**, 1.47 g, 5.38 mmol) was dissolved in anhydrous THF (50 ml) and cooled to 0 °C in an ice bath. LiAlH_4 solution (8.3 ml, 1 M) was added slowly via a syringe and the reaction was left stirring for 2 h. After completion of the reaction, ethyl acetate (20 ml) was slowly added to the mixture and then a saturated solution of Rochelles salt (20 ml) was added. The reaction was left stirring vigorously for circa 15 min and was filtered through celite. The aqueous layer was extracted (3×30 ml) with ethyl acetate and after the second extraction some brine (10 ml) was added. The organic layers are combined, dried over Na_2SO_4 , and concentrated under reduced pressure to obtain a yellow oil and a white solid in quantitative yield. $R_f = 0.50$ (1:1; Hept.:EtOAc). $^1\text{H NMR}$ (400 MHz, CDCl_3) δ 4.25 (1H, m, βH), 3.99 (1H, m, αH), 3.78-3.54 (2H, m, $J = 10.8$, 4.2, CH_2), 1.53 (6H, s, $2 \times \text{CH}_3$), 1.47 (9H, s, $3 \times \text{CH}_3$), 1.26 (3H, d, $J = 6.5$, CH_3). $^{13}\text{C NMR}$ (100 MHz, CDCl_3) δ 154.5 (C=O), 92.7 (C_q), 81.1 (C_q), 71.3 (CHCH_3), 63.1 (CH_2), 61.6 (CHCHCH_2), 28.4 ($3 \times \text{CH}_3$), 27.8 (CH_3C), 24.5 (CH_3C), 14.4 (CH_3). **MS**: found 268.17 $[\text{M}+\text{Na}]^+$, calculated for $[\text{C}_{12}\text{H}_{23}\text{NO}_4 + \text{Na}]^+$ 268.15.



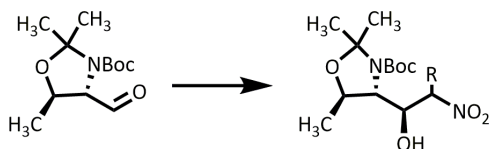
(4S, 5S)-2,2,5-Trimethyloxazolidin-4-carbaldehyde 80. The Garner alcohol (0.24 g, 1.04 mmol) was dissolved in anhydrous dichloromethane and Dess-Martin reagent was added (0.48 g, 1.14 mmol). After 2 h 0.1 eqv. Dess Martin reagent (0.07 g, 0.17 mmol) was added in addition. The reaction mixture was quenched after 5 h with 25 ml of $\text{Na}_2\text{S}_2\text{O}_3(\text{aq})$: $\text{NaHCO}_3(\text{aq})$; 1:1 (v/v) for 40 min.

The product was extracted with CH_2Cl_2 (3×75 ml), dried over Na_2SO_4 and concentrated *in vacuo* to obtain an orange oil in quantitative yield (0.24 g, 1.04 mmol).

Or:

A 250 ml three necked round bottom flask was equipped two dropping funnels. One dropping funnel was charged with DMSO (2.0 ml, 28.6 mmol, 3 eqv., in 2.5 ml CH_2Cl_2) the other one with *tert*-butyl (4*R*,5*S*)-4-(hydroxymethyl)-2,2,5-trimethyloxazolidine-3-carboxylate (2.32 g, 9.54 mmol in 15 ml CH_2Cl_2). The three necked round bottom flask was charged with oxalyl chloride (1.20 ml, 14.30 mmol, 1.5 eqv., in 20 ml CH_2Cl_2). The flask was cooled to -78°C and the DMSO solution was dropwise added over a period of 15 min. Next the solution of alcohol was dropwise added followed by rinsing of the dropping funnel with CH_2Cl_2 (2×5 ml) and the dropping funnel was subsequently charged with DIEA (9.0 ml, 51.6 mmol, 5.5 eqv.). The reaction mixture was allowed to warm up to -40°C and after 30 min DIEA was dropwise added. The reaction mixture was allowed to warm up to -0°C and 1M HCl (35 ml) was added. The aqueous layer was extracted with CH_2Cl_2 (3×10 ml) and the combined organic phases were washed with phosphate buffer (pH 7, 4×20 ml). The organic phase was dried (Na_2SO_4) and concentrated *in vacuo* to afford *tert*-butyl (4*S*,5*S*)-4-formyl-2,2,5-trimethyloxazolidine-3-carboxylate (1.83 g, 7.52 mmol) as a yellow oil in 80% yield. $R_f = 0.38$ (4:1; Hept.:EtOAc). $^1\text{H NMR}$ (400 MHz, CDCl_3) δ 9.49 (1H, dd, $J = 20.9, 3.4$, COH), 4.25 (1H, p, $J = 6.6$, αH), 4.04-3.88 (1H, m, βH), 1.53 (6H, s, $2 \times \text{CH}_3$), 1.47 (9H, s, $3 \times \text{CH}_3$), 1.26 (3H, d, $J = 6.5$, CH_3). $^{13}\text{C NMR}$ (100 MHz, CDCl_3) δ 200.6 (COH), 152.5 (C=O), 94.5 (C_q), 80.8 (C_q), 72.3 (CH), 68.0 (CHCH_3), 28.3 ($3 \times \text{CH}_3$), 26.7 ($2 \times \text{CH}_3$), 15.2 (CH_3). **MS:** found 298.25 $[\text{M} + \text{CH}_3\text{OH} + \text{Na}]^+$, calculated for $[\text{C}_{12}\text{H}_{21}\text{NO}_4 + \text{CH}_3\text{OH} + \text{Na}]^+$ 298.16; found 266.17 $[\text{M} + \text{Na}]^+$, calculated for $[\text{C}_{12}\text{H}_{21}\text{NO}_4 + \text{Na}]^+$ 266.14.

General procedure Henry reaction

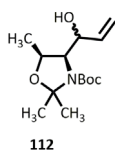


For analysis Henry reaction: A visual reaction was clearly seen when base was added, the colourless solution turned into brown orange like solutions.

inorganic bases. The aldehyde (1.0 eqv) was dissolved in THF:water mixtures (3:2 / 1 gram aldehyde in 150 ml) and cooled to the desired temperature. Inorganic base (0.1 eqv) was added followed by nitro-acetal (7 eqv). The reaction mixture was stirred overnight at the set temperature. The solvent was removed under reduced pressure and the residue was extracted with CH_2Cl_2 and washed with water (3×50 ml). The organic phase was dried over Na_2SO_4 and the solvent was removed *in vacuo*. Column

chromatography was performed with EtOAc in heptane (1:4) to give the desired product typically as a mixture of diastereomers.

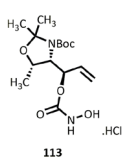
Organic bases or catalyst. The aldehyde (1.0 eqv) was dissolved in anhydrous THF and cooled to -40 °C. Nitro-acetal (7 eqv) was added followed by organic base or catalyst. The reaction mixture was stirred at -40 °C overnight or longer if needed. The reaction mixture was diluted with EtOAc. The organic layer was washed with NH₄Cl (3 × 10 ml) and water (3 × 10 ml). The organic layer was dried over Na₂SO₄ and concentrated *in vacuo*. Column chromatography was performed with EtOAc in heptane (1:4) to give the desired product typically as a mixture of diastereomers.



112

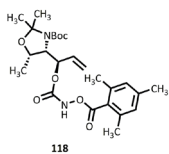
tert-Butyl (4R,5S)-4-((R)-1-hydroxyallyl)-2,2,5-trimethyloxazolidine-3-carboxylate 112. A round bottom flask was charged with *tert*-butyl (4S,5S)-4-formyl-2,2,5-trimethyloxazolidine-3-carboxylate (**80**, 1.83 g, 7.52 mmol) in anhydrous THF (15 ml) and cooled to -78 °C. After 15 min, Vinylmagnesium bromide solution (1.0 M, 8.27 ml, 1.1 eqv.) was dropwise added. After 5 h, sat.

aq. NH₄Cl (20 ml) was added. The aqueous layer was extracted with EtOAc (4 × 50 ml) and the combined organic phases were washed with brine (1 × 50 ml). The organic phase was dried over Na₂SO₄ and concentrated *in vacuo*. Flash column chromatography (SiO₂, 1:1 Hept.:EtOAc » 1:3 Hept.:EtOAc) afforded *tert*-butyl (4R,5S)-4-((R)-1-hydroxyallyl)-2,2,5-trimethyloxazolidine-3-carboxylate (**112**, 1.39 g, 5.12 mmol) as a yellow oil in 68% yield. *R_f* = 0.43 (2:1; Hept.:EtOAc). ¹H NMR (400 MHz, CDCl₃) (mixture of rotamers, main / minor) δ 5.99-5.83 (1H, m, CHCH₂), 5.31/5.38 (1H, dt, *J* = 17, 1.5, CH₂), 5.22-5.14 (1H, m, CH₂), 4.33 (1H, t, *J* = 6.5, CHCH₃), 4.26 (2H, m, 2 × CH), 1.53 (3H, s, CH₃C), 1.45 (12H, m, CH₃C and 3 × CH₃), 1.34 (3H, d, *J* = 6.5, CH₃). ¹³C NMR (100 MHz, CDCl₃) (mixture of rotamers, main / minor) δ 155.4 (C=O), 138.5/139.3 (CCH₂), 115.0/116.0 (CH₂CH), 93.1 (C_q), 81.4 (C_q), 72.5 (CH₃CH), 71.4 (HOCH), 65.8 (CHCHCH), 26.5 (CH₃), 24.4 (CH₃), 28.4 (3 × CH₃), 15.3 (CH₃). **MS:** found 294.17 [M+Na]⁺, calculated for [C₁₄H₂₅NO₄ + Na]⁺ 294.17.



113

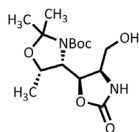
tert-Butyl (4R,5S)-4-((R)-1-((hydroxycarbamoyl)oxy)but-3-en-2-yl)-2,2,5-trimethyloxazolidine-3-carboxylate hydrochloride 113. The vinyl alcohol (**112a**, 800 mg, 2.95 mmol) was dissolved in anhydrous pyridine (20 ml) and stirred at 40 °C. To the reaction mixture was added 1,1-carbodiimidazole (2.87 g, 17.69 mmol, 6 eqv). After 4 h, the reaction mixture was cooled to 0 °C before hydroxylamine.HCl (4 gram, 20 eqv) was added. The reaction mixture was stirred overnight at room temperature and quenched with water. The product was extracted with EtOAc. The organic layer was washed with water and dried over Na₂SO₄ before it was concentrated *in vacuo*. Column chromatography with heptane and EtOAc (1:1) gave the product in moderate yield (51%, 550 mg, 1.50 mmol). ¹H NMR (400 MHz, CDCl₃) δ 5.99 – 5.74 (1H, m, CHCH₂), 5.37 – 5.17 (2H, m, CH₂), 5.16 – 5.06 (1H, m, NH), 4.31 – 4.23 (1H, m, CHO), 4.07 (1H, dd, *J* = 6.6, 1.7, CH), 3.40 (1H, s, OH), 1.47-1.40 (15H, m, 2 × CH₃ + 3 × CH₃), 1.27 (3H, t, *J* = 4.9, CH₃). ¹³C NMR (100 MHz, CDCl₃) δ 154.9 (C=O), 138.3 (CHCH₂), 114.9 (CH₂), 93.0 (C_q), 81.3 (C_q), 72.3 (CH), 71.3 (CH), 65.7 (CH), 28.3 (3 × CH₃), 24.8 (2 × CH₃) 15.2 (CH₃). **HRMS:** found 413.2658 calculated for [C₁₆H₂₉N₂O₆Cl + CH₃OH + H]⁺ 413.2035.



118

tert-Butyl (4R,5S)-2,2,5-trimethyl-4-((R)-1-(((2,4,6-trimethylbenzoyl)oxy)carbamoyl)oxy)but-3-en-2-yl)oxazolidine-3-carboxylate 118.

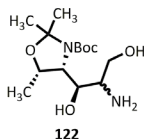
The hydroxylamine (**113**, 550 mg, 1.50 mmol) was dissolved in anhydrous CH_2Cl_2 and cooled to 0 °C. Triethylamine (275 μl , 1.65 mmol, 1.1 eqv) and 2,4,6-trimethylbenzoyl chloride (460 μl , 3.30 mmol, 2.2 eqv) were slowly added. The reaction mixture was stirred overnight at room temperature. The reaction mixture was quenched with NaHCO_3 (20 ml) and extracted with CH_2Cl_2 (3×10 ml). The organic layer was washed with brine, water (both 2×5 ml), and dried over Na_2SO_4 before it was concentrated *in vacuo*. Column chromatography was performed with heptane and EtOAc (H4:E1) to give the product **118**, as a white foam (600 mg, 84% yield). ^1H NMR (400 MHz, CDCl_3) δ 6.87 (2H, s, $2 \times \text{CH}$), 5.97–5.84 (1H, m, CH_2CH), 5.62–5.33 (2H, m, CH_2), 4.31–4.26 (1H, m, CH_3CH), 4.19–4.17 (1H, m, CHO), 4.04–3.95 (CHCHCH_3), 2.36 (3H, s, $1 \times \text{CH}_3$), 2.29 (6H, s, $2 \times \text{CH}_3$), 1.56 (6H, m, $2 \times \text{CH}_3$), 1.49 (9H, m, $3 \times \text{CH}_3$), 1.35 (3H, d, $J = 4$, CH_3). ^{13}C NMR (100 MHz, CDCl_3) δ 168.9 (C=O), 155.4 (C=O), 152.8 (C=O), 136.7 (C_q), 133.3 (CHCH_2), 128.6 (CH_{ar}), 126.5 (C_q), 119.2 (CH_2), 92.6 (C_q), 80.5 (C_q), 72.3 (CH_3CH), 61.0 (CHCHCH_2), 60.8 (CHN), 28.4 ($3 \times \text{CH}_3$), 24.9 ($2 \times \text{CH}_3$), 19.9 ($3 \times \text{CH}_3$), 14.7 (CH_3). For the synthesis of benzoyl **117** and pentafluorobenzoyl **119** the same procedure was used. The products were obtained in 77% and 40% respectively.



120

tert-Butyl (4S,4'R,5S,5'R)-4'-(hydroxymethyl)-2,2,5-trimethyl-2'-oxo-[4,5'-bioxazolidine]-3-carboxylate 120. The trimethylene benzoyl carbamate (**118**, 200 mg, 0.42 mmol) was dissolved in $t\text{-BuOH}:\text{CH}_3\text{CN}:\text{H}_2\text{O}$, (4:4:1, 2 ml), a catalytic amount of potassium osmate (VI) dihydrate (5 mol%) was added and the reaction mixture was stirred overnight. The reaction was quenched with

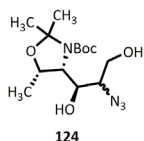
sat. aq. Na_2SO_3 . The aqueous layer was extracted with EtOAc (3×10 ml) and the organic phases were dried over MgSO_4 and concentrated *in vacuo*. Column chromatography was performed (2.5% CH_3OH in CH_2Cl_2) to obtain the product as a mixture of diastereoisomers (**120**, 70 mg, 50% yield). ^1H NMR (400 MHz, CDCl_3) δ 4.38 (1H, s, CHO), 4.27 – 4.06 (4H, m, CH_3CH , CH_2 , and CHN), 3.62 (1H, dd, $J = 11.4$, 3.7, CHNH), 3.47 (1H s, OH), 1.41 (6H, s, $2 \times \text{CH}_3$), 1.32 (3H, d, $J = 6.5$, CH_3), 1.29 – 1.12 (9H, m, $3 \times \text{CH}_3$). ^{13}C NMR (100 MHz, CDCl_3) δ 159.7 (C=O), 153.8 (C=O), 93.4 (C_q), 81.3 (C_q), 75.9 (CHO), 72.2 (CH_3CH), 63.6 (CHNH), 60.5 (CHN), 55.9 (CH_2), 28.3 ($3 \times \text{CH}_3$), 24.9 ($2 \times \text{CH}_3$), 14.4 (CH_3). HRMS: found 353.1680 $[\text{M}+\text{Na}]^+$ calculated for $[\text{C}_{15}\text{H}_{26}\text{N}_2\text{O}_6]^-$ 330.1689.



122

tert-Butyl (4R,5S)-4-((1R)-2-amino-1,3-dihydroxypropyl)-2,2,5-trimethyloxazolidine-3-carboxylate 122.

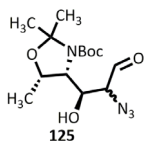
tert-butyl (4S,5S,5'R)-4'-(hydroxymethyl)-2,2,5-trimethyl-2'-oxo-[4,5'-bioxazolidine]-3-carboxylate (**120**, 28 mg, 0.09 mmol, 1 eqv) was dissolved in ethanol (4 ml) and refluxed at 66 °C. NaOH (2M, 2 ml) was slowly added and stirring was continued for 6 h. The reaction mixture was concentrated *in vacuo* and used in the next step without any further purification. MS: found 305.33 $[\text{M}+\text{H}]^+$, calculated for $[\text{C}_{14}\text{H}_{29}\text{N}_2\text{O}_5]^+$ 305.39.



124

tert-Butyl
(4R,5S)-4-((1S)-2-azido-1,3-dihydroxypropyl)-2,2,5-
trimethyloxazolidine-3-carboxylate 124.

tert-butyl (4*R*,5*S*)-4-((1*R*)-2-amino-1,3-dihydroxypropyl)-2,2,5-trimethyloxazolidine-3-carboxylate (**122**, 30 mg, 0.09 mmol, 1 eqv) was dissolved in CH₃OH – H₂O (8:2, 10 ml). NaHCO₃ (16.60 mg, 0.20 mmol, 2.2 eqv) and CuSO₄·5H₂O (5 mg, 0.018 mmol, 0.2 eqv) were added and the pH was adjusted to 8-9. Diazotransfer reagent (**123**, 30 mg, 0.11 mmol, 1.2 eqv) was added and the reaction mixture was stirred at room temperature for 4h. The reaction mixture was diluted with H₂O and the product was extracted with CH₂Cl₂ (3 × 10 ml). The organic layer was dried over Na₂SO₄ and concentrated *in vacuo*. Column chromatography was performed with 5% CH₃OH in CH₂Cl₂ to give the desired product (20 mg, 67% yield). ¹H NMR (400 MHz, CDCl₃) 5.83 (1H, bs, OH), 4.55 (1H, m, CHOH), 4.36 – 4.14 (3H, m, CH₃CH, CH₃CHCH, and CHN₃), 3.53 (2H, m, CH₂), 1.58-1.40 (18H, m, 2 × CH₃, 3 × BocCH₃, and CH₃). ¹³C NMR (100 MHz, CDCl₃) δ 159.3 (C=O), 93.6 (CH₃Cq), 81.2 ((CH₃)₃Cq), 71.4 (CH), 63.8, (CH) 60.2 (CH₂), 55.9 (CH), 53.4 (CH), 28.3 (3 × CH₃), 26.8, 24.9 (2 × CH₃Cq), 14.6 (CH₃). MS: found 353.25 [M+Na]⁺, calculated for [C₁₄H₂₆N₄O₅]⁺ 330.19.



125

tert-Butyl
(4R,5S)-4-((1S)-2-azido-1-hydroxy-3-oxopropyl)-2,2,5-
trimethyloxazolidine-3-carboxylate 125.

The azido alcohol (**124**, 15 mg, 0.045 mmol, 1 eqv) was dissolved in anhydrous CH₂Cl₂ (3 ml). Dess-Martin reagent (0.3 M, 10 μl) was added and the reaction mixture was stirred at room temperature for 4.5 h. The reaction mixture was diluted with CH₂Cl₂ (5 ml) and quenched with Na₂S₂O₃ and NaHCO₃ (both 2 ml). The product was extracted with CH₂Cl₂ (3 × 5 ml), dried over Na₂SO₄, and concentrated *in vacuo* to give a crude mixture (**125**, 10 mg, 67% yield).

References

1. Rangarajan, E. S.; Proteau, A.; Cui, Q.; Logan, S. M.; Potetinova, Z.; Whitfield, D.; Purisima, E. O.; Cygler, M.; Matte, A.; Sulea, T.; Schoenhofen, I. C., *J. Biol. Chem.* **2009**, *284* (31), 20989-1000.
2. Schoenhofen, I. C.; Vinogradov, E.; Whitfield, D. M.; Brisson, J. R.; Logan, S. M., *Glycobiology* **2009**, *19* (7), 715-25.
3. Hassan, M. I.; Lundgren, B. R.; Chaumun, M.; Whitfield, D. M.; Clark, B.; Schoenhofen, I. C.; Boddy, C. N., *Angew. Chem.* **2016**, *55* (39), 12018-21.
4. Lee, Y. J.; Kubota, A.; Ishiwata, A.; Ito, Y., *Tetrahedron Lett.* **2011**, *52* (3), 418-421.
5. Gross, P. H.; Jeanloz, R. W., *J. Org. Chem.* **1967**, *32* (9), 2759-63.
6. Sharma, M.; Bernacki, R. J.; Paul, B.; Korytnyk, W., *Carbohydr. Res.* **1990**, *198* (2), 205-21.
7. Rosenberg, A., *Biology of the Sialic Acids*. Plenum Press: New York, **1995**.
8. Barbier, P., *Compt. Rend.* **1899**, *128* (110).
9. Csuk, R.; Hugener, M.; Vasella, A., *Helv. Chim. Acta* **1988**, *71* (3), 609-618.
10. Gordon, D. M.; Whitesides, G. M., *J. Org. Chem.* **1993**, *58* (27), 7937-7938.
11. Zunk, M.; Williams, J.; Carter, J.; Kiefel, M. J., *Org. Biomol. Chem.* **2014**, *12* (18), 2918-25.
12. Glatzhofer, D. T.; Roy, R. R.; Cossey, K. N., *Org. Lett.* **2002**, *4* (14), 2349-52.
13. Schreiner, E.; Zbiral, E., *Justus Liebigs Ann. Chem.* **1990**, (6), 581-586.
14. Williams, J. T.; Corcilius, L.; Kiefel, M. J.; Payne, R. J., *J. Org. Chem.* **2016**, *81* (6), 2607-11.
15. Knirel, Y. A.; Shashkov, A. S.; Tsvetkov, Y. E.; Jansson, P. E.; Zahringer, U., *Adv. Carbohydr. Chem. Biochem.* **2003**, *58*, 371-417.
16. Tsvetkov, Y. E.; Shashkov, A. S.; Knirel, Y. A.; Backinowsky, L. V.; Zahringer, U., *Mendeleev Commun.* **2000**, (3), 90-91.
17. Tsvetkov, Y. E.; Shashkov, A. S.; Knirel, Y. A.; Zahringer, U., *Carbohydr. Res.* **2001**, *335* (4), 221-43.
18. Tsvetkov, Y. E.; Shashkov, A. S.; Knirel, Y. A.; Zahringer, U., *Carbohydr. Res.* **2001**, *331* (3), 233-7.
19. Kochetkov, N. K.; Byramova, N. E.; Tsvetkov, Y. E.; Backinowsky, L. V., *Tetrahedron* **1985**, *41* (16), 3363-3375.
20. Eis, M. J.; Ganem, B., *Carbohydr. Res.* **1988**, *176* (2), 316-323.
21. Mas Pons, J.; Dumont, A.; Sautejeau, G.; Fugier, E.; Baron, A.; Dukan, S.; Vauzeilles, B., *Angew. Chem.* **2014**, *53* (5), 1275-8.
22. Matthies, S.; Stallforth, P.; Seeberger, P. H., *J. Am. Chem. Soc.* **2015**, *137* (8), 2848-51.
23. Liang, X.; Lee, C. J.; Chen, X.; Chung, H. S.; Zeng, D.; Raetz, C. R.; Li, Y.; Zhou, P.; Toone, E. J., *Bioorg. Med. Chem.* **2011**, *19* (2), 852-60.
24. Candeias, N. R.; Montalbano, F.; Cal, P. M.; Gois, P. M., *Chem. Rev.* **2010**, *110* (10), 6169-93.
25. Szechner, B.; Achmatowicz, O.; Galdecki, Z.; Fruzinski, A., *Tetrahedron* **1994**, *50* (25), 7611-7624.
26. Anelli, P. L.; Biffi, C.; Montanari, F.; Quici, S., *J. Org. Chem.* **1987**, *52* (12), 2559-2562.
27. Soule, J. F.; Mathieu, A.; Norsikian, S.; Beau, J. M., *Org. Lett.* **2010**, *12* (22), 5322-5.
28. Choi, S. K.; Lee, S.; Whitesides, G. M., *J. Org. Chem.* **1996**, *61* (25), 8739-8745.
29. Lall, M. S.; Ramtohul, Y. K.; James, M. N.; Vederas, J. C., *J. Org. Chem.* **2002**, *67* (5), 1536-47.
30. Ait-Haddou, H.; Hoarau, O.; Cramailere, D.; Pezet, F.; Daran, J. C.; Balavoine, G. G., *Chemistry* **2004**, *10* (3), 699-707.
31. Elliott, D. F., *J. Chem Soc.* **1950**, (Jan), 62-68.
32. Kelly, G. T.; Sharma, V.; Watanabe, C. M., *Bioorg. Chem.* **2008**, *36* (1), 4-15.
33. Henry, L., *C. R. Acad. Sci. Ser. C* **1895**, *120*, 1265.

34. Henry, L., *Bull. Soc. Chim. France* **1895**, 13, 999.
35. Luzzio, F. A., The Henry reaction: recent examples. *Tetrahedron* **2001**, 57 (6), 915-945.
36. Sasai, H.; Itoh, N.; Suzuki, T.; Shibasaki, M., *Tetrahedron Lett.* **1993**, 34 (5), 855-858.
37. Sasai, H.; Suzuki, T.; Itoh, N.; Shibasaki, M., *Tetrahedron Lett.* **1993**, 34 (5), 851-854.
38. Sasai, H.; Suzuki, T.; Itoh, N.; Tanaka, K.; Date, T.; Okamura, K.; Shibasaki, M., *J. Am. Chem. Soc.* **1993**, 115 (22), 10372-10373.
39. Shibasaki, M.; Kanai, M.; Matsunaga, S.; Kumagai, N., *Acc. Chem. Res.* **2009**, 42 (8), 1117-27.
40. Sasai, H.; Watanabe, S.; Suzuki, T.; Shibasaki, M., *Org. Synth.* **2002**, 78 (14).
41. Sasai, H.; Kim, W. S.; Suzuki, T.; Shibasaki, M.; Mitsuda, M.; Hasegawa, J.; Ohashi, T., *Tetrahedron Lett.* **1994**, 35 (33), 6123-6126.
42. Rene, L.; Royer, R., *Synthesis-Stuttgart* **1981**, (11), 878-878.
43. Weberg, R. T.; Haltiwanger, R. C.; Laurie, J. C. V.; Dubois, M. R., *J. Am. Chem. Soc.* **1986**, 108 (20), 6242-6250.
44. Schmölzer, C.; Nowikow, C.; Kahlig, H.; Schmid, W., *Carbohydr. Res.* **2013**, 367, 1-4.
45. Roskamp, E. J.; Pedersen, S. F., *J. Am. Chem. Soc.* **1987**, 109 (21), 6551-6553.
46. Hanessian, S.; Devasthale, P. V., *Bioorg. Med. Chem. Lett.* **1996**, 6 (18), 2201-2206.
47. Hanessian, S.; Devasthale, P. V., *Tetrahedron Lett.* **1996**, 37 (7), 987-990.
48. Menzel, A.; Ohrlein, R.; Griesser, H.; Wehner, V.; Jager, V., *Synthesis-Stuttgart* **1999**, (11), 2000-2000.
49. Kitayama, T., *Tetrahedron* **1996**, 52 (17), 6139-6148.
50. Gruber-Khadjawi, M.; Purkharthofer, T.; Skranc, W.; Griengl, H., *Adv. Synth. Catal.* **2007**, 349 (8-9), 1445-1450.
51. Dondoni, A.; Perrone, D., *Org. Synth.* **2000**, 77, 64.
52. Donohoe, T. J.; Johnson, P. D.; Pye, R. J., *Org. Biomol. Chem.* **2003**, 1 (12), 2025-8.
53. Li, G. G.; Chang, H. T.; Sharpless, K. B., *Angew. Chem.* **1996**, 35 (4), 451-454.
54. Donohoe, T. J.; Chughtai, M. J.; Klauber, D. J.; Griffin, D.; Campbell, A. D., *J. Am. Chem. Soc.* **2006**, 128 (8), 2514-2515.
55. Donohoe, T. J.; Callens, C. K. A.; Lacy, A. R.; Winter, C., *Eur. J. Org. Chem.* **2012**, (4), 655-663.
56. Wai, J. S. M.; Marko, I.; Svendsen, J. S.; Finn, M. G.; Jacobsen, E. N.; Sharpless, K. B., *J. Am. Chem. Soc.* **1989**, 111 (3), 1123-1125.
57. Donohoe, T. J.; Bataille, C. J.; Gattrell, W.; Kloesges, J.; Rossignol, E., *Org. Lett.* **2007**, 9 (9), 1725-8.
58. Pullin, R. D.; Rath, A. H.; Melikhova, E. Y.; Winter, C.; Thompson, A. L.; Donohoe, T. J., *Org. Lett.* **2013**, 15 (21), 5492-5.
59. Gao, J. M.; Harter, R.; Gordon, D. M.; Whitesides, G. M., *J. Org. Chem.* **1994**, 59 (13), 3714-3715.
60. Warwel, M.; Fessner, W. D., *Synlett* **2000**, (6), 865-867.
61. Ireland, R. E.; Liu, L. B., *J. Org. Chem.* **1993**, 58 (10), 2899-2899.
62. Frigerio, M.; Santagostino, M.; Sputore, S., (IBX). *J. Org. Chem.* **1999**, 64 (12), 4537-4538.
63. Spielvogel, D.; Kammerer, J.; Keller, M.; Prinzbach, H., *Tetrahedron Lett.* **2000**, 41 (41), 7863-7867.
64. Goddard-Borger, E. D.; Stick, R. V., *Org. Lett.* **2007**, 9 (19), 3797-800.
65. Fischer, N.; Goddard-Borger, E. D.; Greiner, R.; Klapotke, T. M.; Skelton, B. W.; Stierstorfer, J., *J. Org. Chem.* **2012**, 77 (4), 1760-1764.
66. Potter, G. T.; Jayson, G. C.; Miller, G. J.; Gardiner, J. M., *J. Org. Chem.* **2016**, 81 (8), 3443-6.
67. Still, W. C.; Kahn, M.; Mitra, A., *J. Org. Chem.* **1978**, 43 (14), 2923-2925.



Chapter 6

General Discussion

"Keep talking"

The Division Bell, Pink Floyd

General discussion

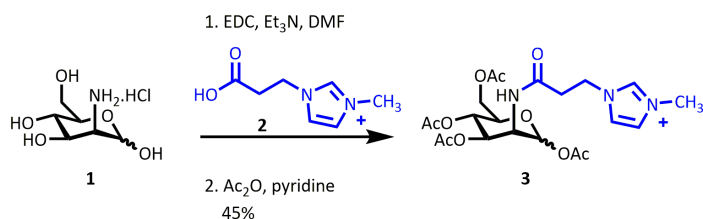
The previous three research chapters in this thesis each focussed on a specific study within the field of chemical microbiology. Furthermore, **chapter 2** reviewed the progress made in the field of metabolic oligosaccharide engineering over the last five years. A chemistry-based approach was chosen to contribute to the fascinating world of bacterial carbohydrates in the human gut. This combination between chemical biology and microbiology - chemical microbiology - is rather challenging as this thesis clearly illustrates. Many outstanding research questions are yet to be answered. In the present chapter we further discuss the results presented in the previous chapters, elaborate on these, suggest alternatives, and provide an outlook towards further research and/or applications.

Chapter 2: Getting a Grip on Glycans: a Current Overview of the Metabolic Oligosaccharide Engineering Toolbox

In **chapter 2** a review of advances in the field of metabolic oligosaccharide engineering (MOE) over the last five years is reported, with a focus on the synthesis and structure-activity relationship of the chemical probes used in this technique. The most recent key developments (2016-2017) in this field, since publication of the review, are discussed below.

New bioconjugation techniques

Non-covalent interactions between a carbohydrate and a fluorescent marker would offer an interesting alternative to covalent interactions. Galan and co-workers recently reported an imidazolium-based ionic liquids tags (ITags) for use on *N*-acetylmannosamine derivatives (Scheme 1).¹ The ITag was tolerated by aldolases since the imidazolium *N*-acetylmannosamine probe **3** was *in situ* successfully transformed into the corresponding neuraminic acid probe. Different cell lines, such as Jurkat and HeLa, were incubated with probe **3** and fluorescence was detected after incubation and fluorescent labelling. When the cells were exposed to sialidases after incubation with **3**, the fluorescent signal was diminished. It remains to be seen if this method is applicable in bacteria.

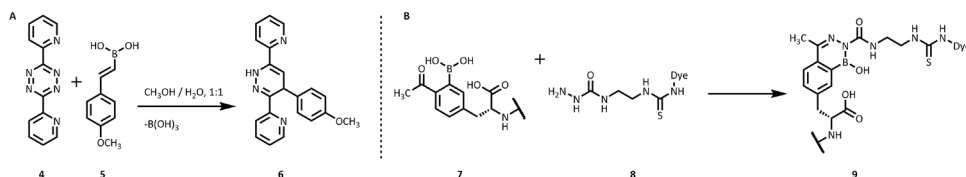


Scheme 1. Synthesis of imidazolium-based ionic liquid tagged probe **3** from *N*-acetylmannosamine **1**.

As reviewed in **chapter 2**, the last decade saw a rapid development of new bioconjugation reactions and two recent discoveries in this field rely on the element boron. Boronic acid-containing molecules remain highly interesting for chemical biological applications due to their hydrophilicity, stability, low

toxicity and as bioisosteres of carboxylic acids. In 2016, Bonger and co-workers reported the Carboni-Lindsey reaction, an inverse electron-demand Diels–Alder cycloaddition between a tetrazine **4** and vinylboronic acid **5** (Scheme 2A).² The method was successfully applied in protein bioconjugation and might be an interesting candidate for application in metabolic oligosaccharide engineering (MOE).

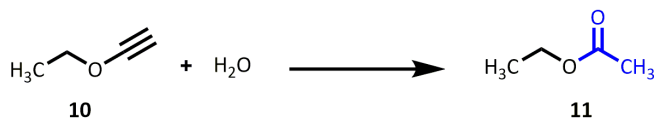
One of the earliest reported bioconjugation reactions is the formation of an oxime and hydrazone between aldehydes/ketones and (hydroxy)amines. The slow reaction kinetics, however, make this system not suitable for live cell imaging to label and image fast biological processes. Inspired by a report from the 1960s, on the conjugation of phenylhydrazine and boronic acid, Gao and co-workers further explored the possibility of diazaborine heterocycles (Scheme 2B).^{3,4} 2-Acetylphenylboronic acid **7** readily conjugated with semicarbazide **8** to form diazaborine **9**. This method was successfully applied in bacterial cell wall labelling of **7** in *Escherichia coli*.



Scheme 2. A) The Carboni-Lindsey reaction with vinylboronic acids and tetrazine. B) iminoboronate ligation.

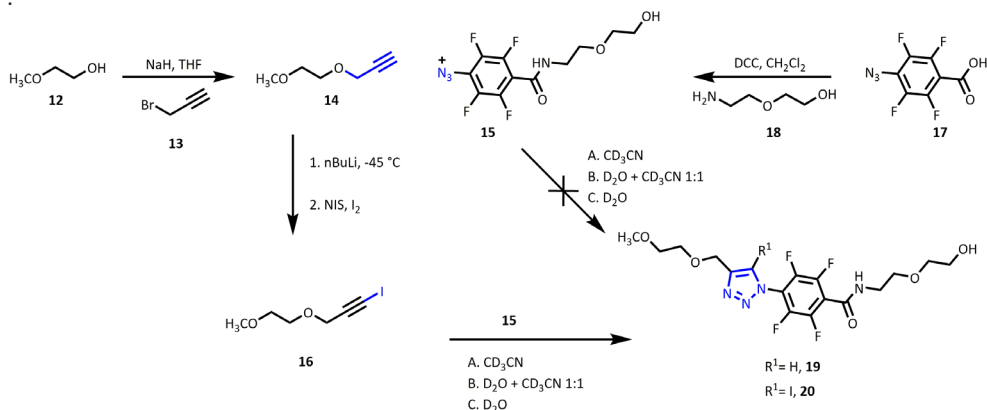
New bioconjugation reactions

There is always room for better bioconjugation reactions, meaning faster and still bioorthogonal, electron-deficient azides might offer a new class and were therefore explored in more detail.⁷ These azides are electron-poor and are easy to make from commercially available perfluorophenylazides.⁸ The mechanism of action is via an inverse electron-demand cycloaddition of electron-deficient azides with aliphatic cyclooctynes as published by Van Delft and co-workers.⁹ They recently showed that electron-deficient aryl azides reacts 29 times faster with bicyclo[6.1.0]nonyne (BCN) than their aliphatic azide counterparts. Based on these findings, we hypothesised that a terminal alkoxy-alkyne functional group might also lead to a cycloaddition with an electron-deficient azide. Furthermore, we hypothesised that this reaction will proceed without the cytotoxic copper catalyst normally needed for this reaction with terminal alkynes. To test this hypothesis we synthesised an alkyne attached to a methoxyethylene glycol. The glycol moiety was added to increase the water solubility of the compound. 2-Methoxyethanol-1-ol **12** was reacted with propargylbromide in the presence of a strong base such as sodium hydride in THF to produce the desired alkyne (scheme 4). For comparison, we also tested the commercially available ethoxyethyne **10**, but it was quickly abandoned as during the course of the reaction exposure to water efficiently transformed it into ethyl acetate **11** (Scheme 3).



Scheme 3. Reaction of ethoxyethyne and water.

The electron-deficient azide was commercially acquired as the tetrafluoro azido benzoic acid **17**. A peptide coupling with 2-(2-aminoethoxy)ethan-1-ol **18** gave the desired product **15**. We also installed an ethyleneglycol group on this reagent to improve its water solubility (scheme 4). Alkoxy-alkyne **14** and electron deficient azide **15** were combined in a series of test tubes containing various deuterated solvents. One with only acetonitrile, one with a 1:1 mixture of acetonitrile and water, and one with only deuterated water. No reaction towards product **19** at room temperature was observed in any of the test tubes.

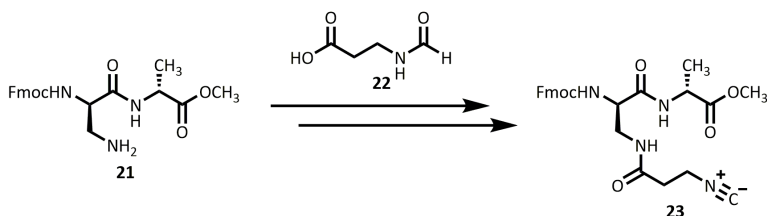


Scheme 4. Click reaction between an electron deficient azide species and aliphatic alkyne.

Inspired by the work of Fokin and Sharpless we also tested iodo-alkynes as suitable reagents for this reaction. Iodo-alkynes are easily accessible alkynes, which are more electron-rich than standard alkynes and thereby increase the reaction kinetics.¹⁰ 1-Iodobut-1-yne was synthesised in two steps involving a lithium halogen exchange reaction (Scheme 4).¹¹ Reacting the iodo-alkyne **16** with **15** also did not result in any product formation **20**. Unfortunately, these two attempts towards new variations of the copper-free click reaction between electron-poor azides and aliphatic alkynes failed, but if a sterically small alkyne can be identified that is capable of this copper-free click reaction it would be a great addition as a chemical reporter on probes in the MOE/bioorthogonal reaction toolbox.

Chapter 3: Probing Peptidoglycan Synthesis and L-fucose Salvage in the Gut Commensal *Akkermansia muciniphila* with Bioorthogonal Chemical Reporters

In **chapter 3**, peptidoglycan synthesis in the bacterial cell wall of *Akkermansia muciniphila* was probed with D-Ala-D-Ala dipeptide derivatives with isonitrile and cyclopropene chemical reporters. Furthermore, metabolic labelling of cell envelope glycans with azido-fucose was also performed in *Akkermansia muciniphila*. MOE relies heavily on active or passive transport of the chemical probes into the cytoplasm. To facilitate passive uptake in MOE for instance acetylated carbohydrates are often used. Non-specific esterases in the cytoplasm hydrolyse the acetylated carbohydrates, followed by metabolism of the unmasked probe in the salvage pathway.¹² Although these non-specific esterases are considered as an advantage for this application, potential challenges in the design of chemical probes need to be taken into account. For example, the cyclopropene peptidoglycan (PG) dipeptide contains a carbamate group which might be sensitive to hydrolysis by non-specific esterases.¹³ However, this phenomenon has not yet been reported for carbamate-linked cyclopropene-glycans.¹⁴⁻¹⁶ The diminished labelling of the cyclopropene PG dipeptide probe compared to the isonitrile PG dipeptide probe might thus be the result of non-specific esterases which cleave the carbamate linked cyclopropene from the core-dipeptide. Another factor might be the size of the chemical reporter group. Our new isonitrile PG dipeptide probe, is, compared to the cyclopropene probe, relatively small and only slightly bigger than the natural D-Ala dipeptide probe. Others, however, showed that larger substituents, such as fluorescent dyes, are also tolerated by the PG biosynthesis pathway in bacteria. As Leeper and co-workers showed on glycan analogues, the isonitrile dipeptide probe can also be elongated.¹⁷ Our synthesis of the PG probes can easily give access to these or other PG probes from amino dipeptide intermediate **21**.



Scheme 5. Proposed alternative isonitrile PG dipeptide probe.

An elongated isonitrile PG dipeptide probe, such as **23**, has extra options in terms of stability. A secondary or tertiary isonitrile version of the probe can for instance be made. Leeper et al.¹⁷ showed that the tertiary isonitriles on *N*-acetylglucosamine sugars are more stable after the click reaction compared to the primary or secondary isonitrile on *N*-acetylglucosamine sugars. Like the here proposed elongated isonitrile PG probe **23**, the isonitrile PG probe reported in **Chapter 3** does already contain a carbonyl group on the γ -position, which should benefit the stability of the tetrazine-clicked product by forming a stable α,β -unsaturated system.¹⁸

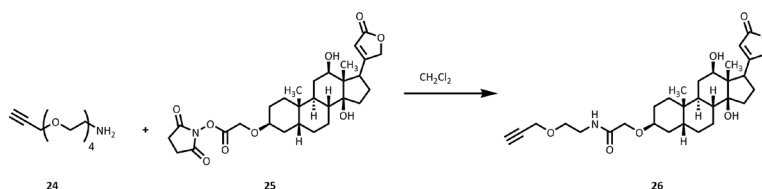
General Discussion

From the microbiology point of view further investigation is also possible. *Akkermansia muciniphila* can easily grow on a synthetic (CP) medium with the addition of vitamins and a carbon source, such as mucin or *N*-acetylglucosamine. Cultures that can grow on GlcNAc might also grow on the azido-variant of GlcNAc, GlcNAz. In contrast, when we treated *A. muciniphila* with Ac₄GlcNAz in CP medium we did not observe any growth. The acetylated sugar can probably not enter the cell via the standard active transport. When 10% of GlcNAc was present, however, we did observe slow culture growth that never reached exponential phase. The addition of more GlcNAc was necessary to reach standard culture density after overnight growth. The optimisation of the medium for *A. muciniphila* is an interesting step to improve the labelling efficiency. For example, addition of probe over a certain time window or using the non-acetylated version of the probe might be explored in future studies.

Chapter 4. Metabolic Glycan Labelling with Azido-Monosaccharides in Human-gut Related Microbes

The study of metabolic labelling of peptidoglycan (PG) in **chapter 3** was further continued towards other human-gut related microbes in **chapter 4**. In all tested species, azido-glycans were added immediately to freshly inoculated samples. Since glycans are important carbon sources for the energy generation by bacteria the glycan-based probes are likely to be directly metabolised, catabolised, or incorporated into other cellular glycans. Moreover, they can be incorporated into internal or external (cell-surface) glycans. For PG labelling this may be slightly different since PG-glycans are an essential component of the cell envelope and may be primarily needed when cells start to divide. Based on these assumptions, we always added the PG probes after a few hours of initial growth and the glycan probes immediately from the start. Further studies should aim to unravel the kinetics of the probe incorporation and provide further insight in the processes underlying the labelling. Furthermore, more detailed analyses on the outer surface, e.g. involving mass spectrometry, is needed to elucidate the chemical structures of labelled cell-surface glycans.

For glycan-based metabolic labelling and visualisation on a blot, a common combination is biotin-alkyne for labelling and streptavidin to visualise on blot.¹⁹ A more preferred combination is, however, digoxigenin (DIG) and the commercially available *anti*-digoxigenin antibody.²⁰ DIG is a steroid found exclusively in the *Digitalis* plant, and the *anti*-DIG antibody does not show any cross-reactivity with other steroids. For that reason DIG – *anti*-DIG was chosen above the more common biotin streptavidin labelling. Tempted by this more selective labelling method, we decided to investigate it and first prepared DIG-alkyne **26** via a one-step synthesis from commercially available DIG-NHS **25** and amino-PEG4-alkyne **24** (Scheme 6). Alkyne **24** can be added to azido-glycans present on the surface of bacterial cells. The attached DIG ligand can after separation then react with *anti*-DIG to allow visualisation.



Scheme 6. Synthesis of DIG-Alkyne.

After labelling of cell envelope glycans using MOE, direct analysis of the exact glycoconjugate composition is not possible. One technique which enables separation of cell surface glycans is sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). SDS-PAGE can serve as an additional separation or isolation technique for the analysis of glycans on glycoproteins. The thus separated glycans can subsequently be transferred to a blot and visualised by staining using dye-labelled lectins or specific glycoprotein stains. The mechanism of action of the latter glycoprotein staining is based on periodate-based oxidative cleavage of vicinal diols in glycoconjugates to aldehydes, followed by Schiff base formation with a reporter group (e.g. an amine with fluorescent dye). When clear separation is observed mass spectrometry can reveal the glycan composition in more detail. Our work with glycoprotein staining was focussed on *A. muciniphila* and *B. fragilis* with Ac₄FucAz. When *A. muciniphila* or *B. fragilis* were both incubated with Ac₄FucAz following the standard protocol as described in **chapter 3** we observed labelling as controlled by confocal microscopy. Next, after washing, DIG-alkyne was added and the culture was incubated for one hour at room temperature. After three more washing steps the cultures were heated at 95 °C for five minutes and loaded onto SDS-PAGE gels. After separation we transferred the gel to a blot and after blocking labelling with anti-DIG was performed. Even after thorough optimisation we observed mainly α-specific binding to our samples. α-specific binding was even seen in samples without any azido-probe present and also with pure cultures. When biotin-streptavidin combinations were used we again observed α-specific binding. Despite known methods to perform glycoproteins-labelling and separation via SDS-PAGE we never observed positive results.²¹ SDS-PAGE is optimised for protein separation based on molecular mass but since glycoproteins may be heterogeneous and show differential SDS binding, some tailing is often seen when glycoproteins are analysed in this way. This was observed when we performed initial tests with SDS-PAGE. To optimise this process other gel types or denaturation techniques need to be explored. Future experiments should investigate the optimisation of steps prior to loading on the gel. Furthermore, as showed by confocal microscopy and flow cytometry analysis the efficiency of incorporation of azido-glycans onto surface glycans is sometimes below 10%. This will also negatively influence a successful detection on SDS-PAGE and subsequent analysis via mass spectrometry.

For our analysis of the MOE results we heavily relied on fluorescence imaging. Fluorescent imaging of gut microbiota is however rather challenging. Molecular biology-based tagging methods, such as fluorescent proteins, cannot be installed in peptidoglycan, glycolipids, and other glycoconjugates since these are not directly encoded in the genes. Furthermore, most gut-bacteria are anaerobic

General Discussion

which causes problems for live cell imaging fluorescence measurements since this often exposes the sample to oxygen. Live cell imaging provides another challenge. The constant movement of the bacteria in the media may result in unfocussed images. A variety of techniques to immobilise cells without killing them are available such as poly-L-lysine coating²² and cell tak²³. Poly-L-lysine coating is based on electrostatic interactions with the cell membrane of bacteria and the positively-charged L-lysine ions on the surface. Cell tak is a commercially available mixture of compounds that is based on a natural extract from mussels where it serves as an adhesive protein. To improve unfocussed microscopy images such commercially available immobilisation techniques need to be implemented and standardised.

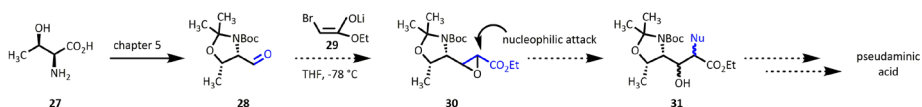
Personal Note on Metabolic Oligosaccharide Engineering in Bacteria

Metabolic oligosaccharide engineering (MOE) is widely adapted as a tool to explore glycans on the outer surface of cells. It is nowadays a well-established method in glycan-labelling of mammalian cell lines. For bacteria, however, labelling is not always straightforward. For mammalian cells versus bacterial cells, different MOE protocols need to be employed. Especially in bacteria an extra set of important parameters need to be considered: growth rate, metabolism, culture density, point of inoculum, concentration of the probe, and the nutrients present in the medium. Compared to MOE in eukaryotes a higher concentration of chemical probe is added to perform MOE in bacterial cells (minimum 10 times more concentrated compared to eukaryotes). Inoculation for MOE experiments is often performed from a pre-culture which was allowed to reach full growth. Inoculation from a late-stage exponential phase, thus from an active culture, can result in direct active metabolism after inoculation and thereby benefit metabolic uptake of the probe. Another important factor for MOE in bacteria is the medium. A variety of nutrients are available for the bacteria. If there is too much nutrient present, especially carbohydrate-based, the chance the chemical probe will be consumed is less likely compared to a minimal medium. Extensive optimisation is needed to further explore all these possibilities. Only after that, MOE can be a powerful tool in studying bacteria. Besides this experimental design, the actual experimental procedures are also crucial. Important parameters such as final concentration of the probe after inoculation, the exact washing procedure (volumes, time and repetition), the incubation time and the exact settings of fluorescent imaging equipment are often not unambiguously reported in current scientific literature.²⁴⁻²⁶ A more detailed experimental procedure would speed up the rate at which MOE in bacteria is further developed and thereby improve our insight into the role of glycans in the human microbiome. Despite the need for a more standard method the wide variety of bacterial species will remain challenging. Moreover, bacteria live in a mixed environment and analysis in these mixed communities, as illustrated in **chapter 4**, is even more challenging than MOE in single cultures.

Chapter 5: Total Synthesis towards Pseudaminic Acid Probes

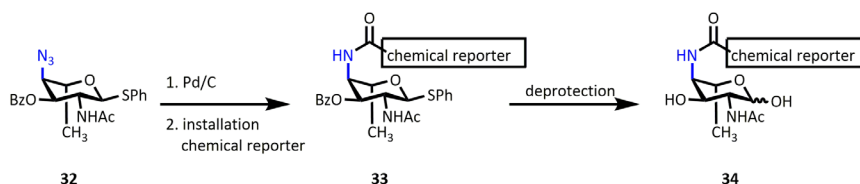
A major goal of the studies presented in this thesis was the total synthesis of pseudaminic acid (Pse) and the subsequent development of chemical probes based on its structure. Results towards this goal can be found in **chapter 5**. In a total synthesis route important choices need to be made all the time as planned transformations fail and alternative routes can always be devised on paper to obtain the target structure. At the start of such a project it is hard to predict beforehand which route will turn out to be most efficient or better yielding. In **chapter 5** we explored both the Henry reaction and tethered aminohydroxylation in great detail as the key transformation to obtain the hexose precursor of pseudaminic acid. Several suggestions for alternative routes are discussed below. Furthermore, we elaborate on how to synthesise chemical probe versions of the Pse hexose precursor and pseudaminic acid itself.

The Garner aldehyde derivative **28** served as a key intermediate in our pursued total synthesis route towards pseudaminic acid (**chapter 5**). An alternative strategy towards pseudaminic acid from this aldehyde **28** is an aldol-Darzens reaction with lithium olefin **29** (Scheme 7).²⁷ The resulting epoxide **30** can be opened using a nucleophile such as sodium azide in order to install the amine alcohol moiety. More steps are needed to further transform this intermediate to pseudaminic acid.



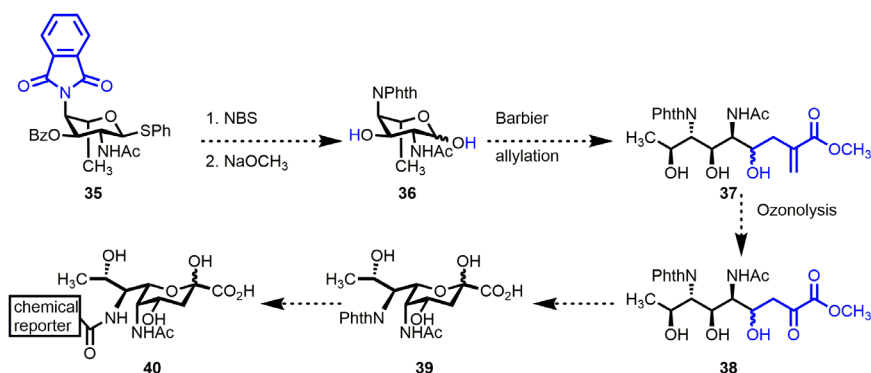
Scheme 7. Proposed alternative synthesis from towards amino-alcohol motif **31**.

In order to perform metabolic labelling with Pse-derivatives a chemical reporter needs to be installed. A few considerations need to be made when selecting a chemical reporter. First, a bioorthogonal strategy needs to be employed to perform efficient metabolic labelling experiments on live bacteria. Next, the selection of the preferred chemical reporter (e.g. azide or isonitrile) needs to be made.²⁸ The Pse hexose precursor **32** can be turned into a chemical probe on the C4-position. The azide can be reduced to an amine and coupled to a cyclopropene group or transformed into an azide. After deprotection of the benzoyl and thiophenol glycoside, followed by acetylation, an acetyl-protected Pse hexose precursor-based probe **33** can be obtained (Scheme 8).



Scheme 8. Proposed synthesis of Pse hexose chemical probe **34**.

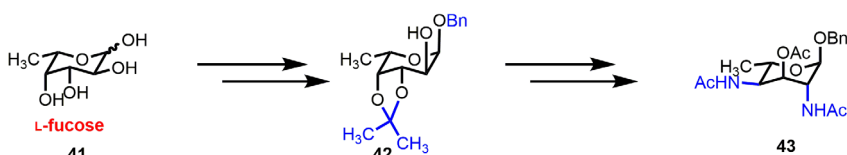
The orthogonal protected precursor **32** can, of course, at this stage be modified into a Pse-based chemical probe (Scheme 9). It is not fully clear whether the remaining synthetic manipulations will affect the activity of the chemical probe. Since the C4-azido group in **32** will not survive the Barbier allylation it can be reduced to an amine and protected with a phthalimide **35**.²⁹ After the Barbier allylation ozonolysis of the olefin produces product **38**.³⁰ In the final steps the two orthogonally protected amines can be turned into chemical reporters. After a final deprotection step the Pse chemical probe should be ready for metabolic labelling experiments (Scheme 9).



Scheme 9. Proposed synthesis of Pse chemical probe **40**.

The usefulness of Pse-based chemical probes for MOE is illustrated by a striking chemical biology experiment with azido-modified Pse, performed by Tanner and co-workers in 2009. A pseudo-chemoenzymatic synthesis allowed them to produce 6-deoxy-AltNac4Az.³¹ This azido-Pse precursor was fed to *Campylobacter jejuni* and metabolic uptake resulting in azido-Pse on the cell surface glycans. *C. jejuni* was genetically modified in such a way that the flagella was not present. After the addition of azido-Pse the bacteria started producing flagella and restored their motility. Hence, the importance of Pse for the motility of *C. jejuni* was proven. In *H. pylori*, another pathogenic gastric bacteria, Pse also plays a pivotal role. Dube and co-workers identified a list of glycosylated proteins containing Pse.³² They also showed that Pse is essential for functional flagella in *H. pylori*. The glycosylated proteins containing azido-Pse on the flagella of *H. pylori* and *C. jejuni* have the potential to elucidate the function of bacterial glycoproteins in bacterial-host interactions.³³⁻³⁶

Pseudaminic acid is a difficult target for total synthesis due to a axial methyl group in the Pse precursor. To install the 6-methyl in an axial position, Ito and co-workers performed a hydrogenation using Wilkinson's catalyst on an olefin. This reaction was performed with remarkable selectivity to yield the axial 6-methyl group. Synthetic challenges in the hydrogenation step of the olefin on similar substrates can however be expected. To circumvent the potential difficult hydrogenation step on our slightly different olefin another route starting from L-fucose **41** can be chosen following a route of Sharon and co-workers.³⁷⁻³⁸ Although L-fucose is an expensive starting point (25 gram ~ € 700,-) for a total synthesis, routes are known from this building block.³⁷ The benefit of L-fucose is the fact that it already contains the 6-methyl group with the right stereochemistry. Further functionalisation to install the two chiral amine moieties does, however, requires quite a few synthetic steps as showed by Sharon and co-workers.³⁷⁻³⁸

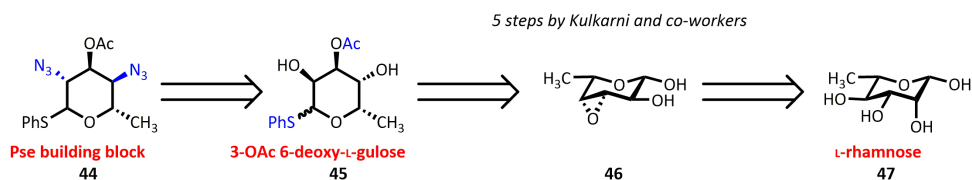


Scheme 10. Synthesis towards Pse from Sharon et al.³⁷

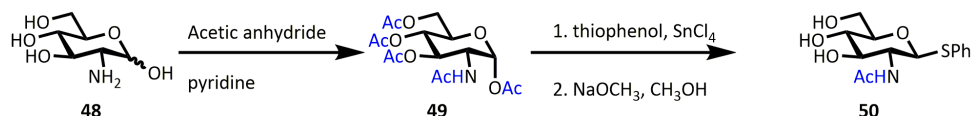
In order to obtain a probe suitable for MOE, the azides need to be transformed into azido-acetamide groups. Differentiation between the two azides for attachment of a chemical reporter can, of course, not be achieved via this method. However, for initial studies in Pse-containing species this synthetic route might be interesting.

Lengthier routes towards Pse starting from cheaper starting sources can also be designed. 6-Deoxy-L-gulose is an interesting starting point towards Pse, although many synthetic manipulations are needed (Scheme 11).³⁹ In 1980, Ireland and co-workers reported the synthesis of unprotected 6-deoxy-L-gulose from D-(+)-Glucuronic acid γ -lactone (100 gram, €60,-) in six steps. 6-Deoxy-L-gulose has the correct stereochemistry on the C-3 and C-5 position. Furthermore the 2,4-diol has the opposite configuration making it possible to perform after activation a S_N2 type nucleophilic addition with an azide source. An improved synthesis of protected 3-OAc 6-deoxy-L-gulose **45** was reported by Kulkarni and co-workers in 2016 (Scheme 11).⁴⁰ From L-rhamnose (100 gram, €400,-) **47**, 3-OAc 6-deoxy-L-gulose **45** was prepared in five steps. At this point, the synthesis can continue with activation of the hydroxyl groups and S_N2 displacement with sodium azide to install the two protected amine functionalities and obtain the Pse building block.

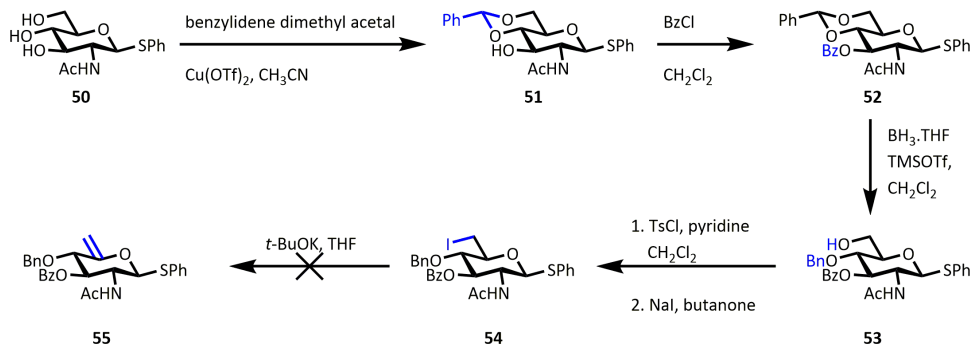
General Discussion



Inspired by the route published by Ito and co-workers we also worked on the total synthesis of pseudaminic acid starting from D-glucosamine **48**.⁴¹ β -thiophenol donor **50** was made in three steps from D-glucosamine **48** (Scheme 12).⁴²

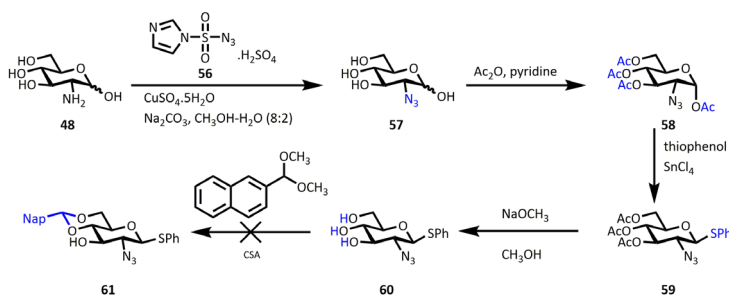


A 4,6-benzylidene moiety was installed to enable modification of the C3-position. The standard method, using camphor sulfonic acid and dimethylbenzaldehyde acetal, did not work in our hands. First of all the solubility of the starting material caused problems, and in fact in standard organic solvents modified *N*-acetylglucosamine sugars are known to cause solubility issues. Secondly, no full conversion was seen using this method. Catalytic amounts of copper triflate in the presence of dimethylbenzaldehyde acetal was used and the reaction mixture was sonicated for 1 hour to obtain the desired product **51** in 94% yield.⁴³ We envisioned that a benzoyl group on the C3-position would be ideal, compared to the benzyl group used by Ito and co-workers.⁴¹ Selective ring opening of the benzylidene proved challenging. A variety of procedures was tested ranging from aluminium trichloride with borane trimethylamine⁴⁴ to TMSOTf and borane in THF.⁴⁵ Unfortunately, no selective opening was observed and only mixtures were isolated. We could, however, retrieve the desired 6-OH product **53** from the reaction mixture. The deoxygenation of C-6 and inversion of C-5 also proved challenging. 6-OH activation was performed with either TsCl, triflic anhydride, or NBS followed by iodination.⁴⁶ The small amount of product **54** that could be isolated was treated with *t*-BuOK, but only trace amounts of the desired alkene **55** were observed.⁴¹ Also throughout the whole synthesis the low solubility of the intermediates was a practical issue that caused loss of products during work up and purification. At this stage, due to these challenges and a lack of precursors, this synthetic route towards Pse was therefore discontinued (Scheme 13).



Scheme 13. Synthesis of β -thiophenol N-acetylglucosamine alkene **55**.

At that stage we reasoned that the low solubility of the GlcNAc-based intermediates in this synthesis could perhaps be solved by the conversion of the amine into an azide. Diazotransfer reagent **56** was successfully applied on commercially available D-glucosamine **48**. After conversion to the azide and acetylation of the alcohols the corresponding β -thiophenol **59** was synthesised. Zémpfen de-acetylation was followed by installation of a naphthyl acetal on the 4,6-position. The naphthyl acetal was preferred over the benzyl acetal, since the naphthyl group can also be removed selectively with DDQ.⁴⁷ This naphthyl protection, however, never reached full conversion and due to lack of precursor this route was also not further pursued (Scheme 14).



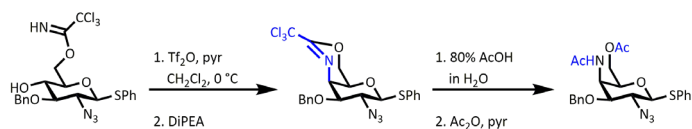
Scheme 14. Synthesis of β -thiophenol azidoglucosamine **60**.

Future directions

I propose to continue the synthesis towards **62** following the route of Ito and co-workers.⁴¹ Alkene **62** can be selectively reduced to the axial 6-methyl **63** using Wilkinson's catalyst. The synthetic challenges related to this step might be solved using another approach. To install the 4-amino moiety I propose

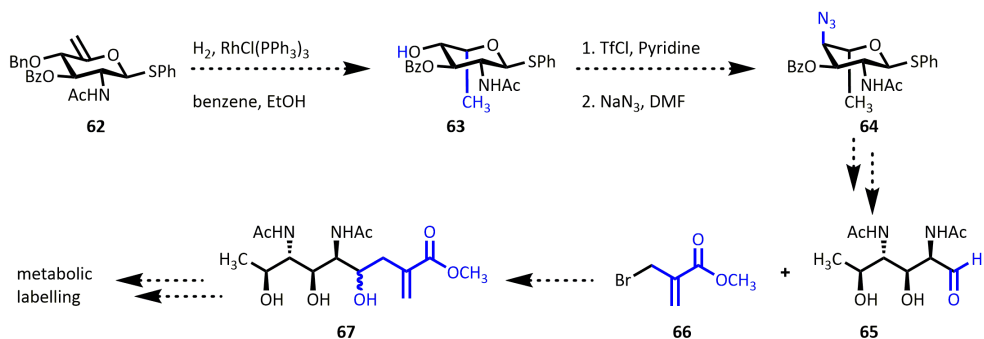
General Discussion

to deprotect the benzyl group under hydrogenation conditions followed by activation and nucleophilic displacement with sodium azide to obtain thio-glycoside **64**. On the related building block used by Ito and co-workers this was however not possible as discussed in **chapter 5**.⁴¹ An alternative route is the installation of an oxazine on the 4,6-position (Scheme 15). Opening of the oxazine with acetic acid gives a 4-amino and 6-hydroxy group as reported earlier.⁴⁸



Scheme 15. Proposed synthesis of 4-NHAc azido glucosamine.

This Pse hexose **65** can be seen as a precursor towards Pse if the 6-OAc group is converted into an axial deoxy 6-methyl using the method of Ito and co-workers using Wilkinson's catalyst. One can either transform the 4-azido into a 4-azidoacetamide group, deprotect the benzoyl and thiophenol group and do metabolic labelling or perform Barbier allylation on the hydrolysed glycoside to obtain alkene **67**. After some more additional steps it should be feasible to obtain Pse probes to do metabolic labelling (Scheme 16).



Scheme 16. Proposed synthesis of Pse precursor **65**.

What is next?

The era of chemical (micro)biology has just started. Numerous bioconjugation reactions have been developed and applied in a variety of (biological) systems and organisms, e.g. plants and microbes. This thesis for example showed the synthesis of two novel PG dipeptide probes which were successfully imaged in *A. muciniphila* and *E. coli*. The imaging of other cellular glycans in human gut related microbes is however rather challenging as **chapter 4** illustrated. In the next decade all these (biological) systems will be mapped in even more detail, especially focussing on the diversity

of glycans. Metabolic labelling techniques such as metabolic oligosaccharide engineering (MOE) and selective exo-enzymatic labelling (SEEL) will be further improved to obtain better selectivity in MOE and the easier excess to complex molecules needed for SEEL. These methods are needed in order to elucidate the biological role of important microbial glycans such as pseudaminic and legionaminic acid. Synthetic chemistry can advance this by developing chemical probes of bacterial glycans as we showed in **chapter 5**. More knowledge about bacterial glycans will help to understand in even more detail the important role of glycans in microbes related to human life.

Experimental section

General

For reactions conducted under anhydrous conditions glassware was dried overnight in an oven at 150 °C and was subsequently allowed to cool in a desiccators over anhydrous KOH. Anhydrous reactions were carried out under nitrogen. Reagents/solvents for anhydrous reactions were obtained from a Pure Process Technology Glass Contour solvent purification system (SPS) (CH_2Cl_2 , THF, diethyl ether, toluene), with 3Å molecular sieves (DMF, CH_3CN , CH_3OH , toluene, pyridine) or Na_2SO_4 (acetone). Commercially acquired chemicals were used without further purification unless stated otherwise.

Flash Column Chromatography

Flash column chromatography was carried out according to standard procedures using silica gel 60 (40-63 μm mesh) from Screening Devices, Amersfoort. Prior to loading an appropriate eluent was chosen by running a TLC of the crude mixture which was to be purified. Loading was performed with celite or small volumes of the eluent.⁴⁷

Thin Layer Chromatography (TLC)

Thin layer chromatography was carried out on and visualised using UV (254 nm), Hanessian stain, sulfuric acid, KMnO_4 , I_2 or ninhydrin stain. The quoted R_f values are rounded to the nearest 0.05.

Nuclear Magnetic Resonance (NMR)

^1H - and ^{13}C -NMR were run on a Bruker Avance 400 NMR spectrometer (400 MHz). Chemical shifts are reported in parts per million (ppm) with reference to tetramethylsilane (TMS) as internal standard. Coupling constants (J) are reported in Hertz (Hz), rounded to the nearest 0.5 Hz. Signal assignment was based on unambiguous chemical shifts in combination with COSY, HSQC, HMBC, and/or DEPTQ.

High Resolution Mass Spectroscopy (HRMS)

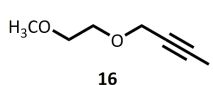
High-resolution ESI mass analyses were recorded on a Orbitrap high-resolution mass spectrometer. The orbitrap high-resolution mass spectrometer system consisted of a ESI ion source (Ion-Sense, Saugus, USA) coupled to an Exactive high resolution MS system (Thermo Fisher Scientific, San Jose, CA,

General Discussion

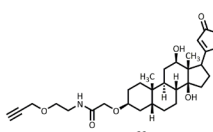
USA). The MS was calibrated daily using ProteoMass™ LTQ/FT-Hybrid ESI (positive and negative mode) Cal Mix (Sigma Aldrich), which is applicable for the m/z range 100–2000. XCalibur software (version 2.1) was used for instrument control, data acquisition and data processing.

Infrared Spectroscopy (IR)

Infrared spectra were recorded on a FT-IR spectrometer apparatus (Bruker Tensor) with a universal ATR accessory. Both solid and oily compounds were recorded. FT-IR signals are reported in cm^{-1} .

**1-iodo-3-(2-methoxyethoxy)prop-1-yne** Propargylmethoxyethanol (50 mg, 0.44 mmol) was dissolved in anhydrous THF (10 ml) and stirred at room temperature. Copperiodide (4 mg, 0.021 mmol, 0.05 eqv) and *N*-iodomorpholine. KI (164 mg, 0.48, 1.1 eqv) was added and the reaction mixture was stirred overnight at room temperature. The reaction mixture was filtered over neutral alumina and washed with CH_2Cl_2 (3×5 ml). The organic layer was dried over Na_2SO_4 and concentrated *in vacuo*. The crude mixture was purified using flash column chromatography (heptane 4:1 EtOAc) to give the product **16**. ^1H NMR (400 MHz, CDCl_3) δ 4.22 (2H, s, CH_2), 3.42 (3H, s, OCH_3), 1.3–1.24 (4H, m, $2 \times \text{CH}_2$).

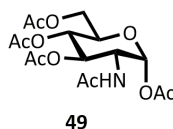
DIG-NHS ester was purchased from Enzo lifesciences. Amino-PEG4-alkyne was purchased from Sigma-Aldrich.

**26** DIG-NHS ester (5 mg, 0.075 mmol) was dissolved in anhydrous DMF (1 ml) and triethylamine (16 μl). Amino-PEG4-alkyne (16 mg, 0.068 mmol) was added and the reaction mixture was stirred overnight at room temperature. The reaction mixture was concentrated *in vacuo* and used without any further purification. Product **26** was obtained as a colourless oil (6 mg, quantitative yield).

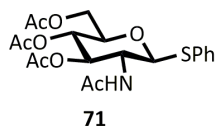
MS: found 797.5 $[\text{M}^+\text{Na}^+]$, calculated for $[\text{C}_{42}\text{H}_{66}\text{N}_2\text{O}_{11} + \text{H}^+]$ 774.99.

Reference: *Bioconjugate Chemistry* **2011**, 22 (7), 1422-1432.

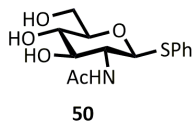
Chapter 5: additional experimental

**49**

(2R,3R,4R,5S,6R)-3-Acetamido-6-(acetoxymethyl)tetrahydro-2H-pyran-2,4,5-triyl triacetate D-Glucosamine.HCl (**48**, 22 g, 102 mmol, 1 eqv) was suspended in pyridine (100 ml) at room temperature. Acetic anhydride (94 ml, 998 mmol, 9.8 eqv) was added and the mixture was stirred overnight at room temperature. The reaction mixture was poured into H₂O (500 ml) and extracted with CH₂Cl₂ (3 × 100 ml). The organic layers were dried over Na₂SO₄ and concentrated *in vacuo*. The oily residue was crystallised by adding CH₂Cl₂:Et₂O (1:10, 200 ml). The α-anomer **49** was obtained in 52% yield. ¹H NMR (400 MHz, CD₃OD) δ 6.01 (1H, d, *J* = 3.6, C-1), 5.24 (1H, dd, *J* = 11.1, 9.3, C-3), 5.03 (1H, t, *J* = 4.0, 10.0, C-4), 4.40 – 4.26 (1H, m, C-2), 4.22 (1H, m, C-6b), 4.08 (1H, m, C-5), 3.99 (dd, *J* = 12.4, 2.4, C-6a), 2.13 (3H, s, CH₃), 2.05 – 1.89 (9H, m, 3 × CH₃), 1.85 (3H, s, CH₃). ¹³C NMR (100 MHz, CD₃OD) δ 174.0 (C=O), 173.9 (C=O), 172.7 (C=O), 172.3 (C=O), 171.6 (C=O), 171.1 (C=O), 91.6 (C-1), 71.8 (C-5), 70.9 (C-4), 69.8 (C-3), 62.9 (C-6), 51.94 (C-2), 20.7 (5 × CH₃).

**71**

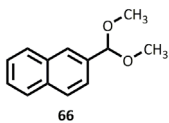
(2R,3S,4R,5R,6S)-5-Acetamido-2-(acetoxymethyl)-6-(phenylthio) tetrahydro-2H-pyran-3,4-diyl diacetate α-D-Glucosaminepentaacetate **49** (20 g, 51 mmol, 1 eqv) was dissolved in CH₂Cl₂ (100 ml) and charged with thiophenol (6 ml, 61 mmol, 1.2 eqv). Tintetrachloride (1M, 4.21 ml, 36 mmol, 0.7 eqv) was added and the reaction mixture was refluxed overnight. The reaction mixture was quenched with NaHCO₃ (250 ml) and extracted with CH₂Cl₂ (3 × 75 ml). The organic layer was washed with brine (3 × 20 ml), dried over Na₂SO₄ and concentrated *in vacuo*. The product was crystallised from Et₂O and hexanes (10:1) to obtain the product **71** (21 g, 98% yield). ¹H NMR (400 MHz, CD₃OD) δ 7.55 – 7.44 (2H, m, CH_{ar}), 7.36 – 7.27 (2H, m, CH_{ar}), 5.65 (1H, d, *J* = 9.2, C-1), 5.22 (1H, dd, *J* = 10.2, 9.3, C-3), 5.04 (1H, t, *J* = 9.7, C-5), 4.85 (1H, d, *J* = 10.4, C-4), 4.28 – 4.11 (2H, m, C-2 and C-6a), 3.72 (1H, ddd, *J* = 10.0, 5.3, 2.7, C-6b), 2.31 – 1.76 (12H, m, 4 × CH₃). ¹³C NMR (100 MHz, CD₃OD) δ 171.0 (C=O), 170.6 (C=O), 170.0 (C=O), 169.3 (C=O), 132.5 (Ar-CH), 132.4 (Ar-CH), 128.9 (Ar-CH), 128.0 (C-5), 86.6 (C-1), 75.8 (C-5), 73.7 (C-4), 68.4 (C-3), 62.4 (C-6), 53.4 (C-2), 23.3 (CH₃), 20.7 (CH₃), 20.7 (CH₃), 20.6 (CH₃).

**50**

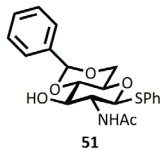
N-((2S,3R,4R,5S,6R)-4,5-Dihydroxy-6-(hydroxymethyl)-2-(phenylthio) tetrahydro-2H-pyran-3-yl)acetamide Thiophenol **68** (270 mg, 0.6 mmol, 1 eqv) was dissolved in anhydrous methanol (5 ml). NaOCH₃ (25% wt, 2 drops, 0.1 eqv) was added and the reaction mixture was stirred at room temperature for 2 hours.

The reaction mixture was neutralised by adding Dowex H⁺, filtered, and concentrated *in vacuo*. Product **50** was obtained in quantitative yield. ¹H NMR (400 MHz, CD₃OD) δ 7.49 – 7.35 (2H, m, CH_{ar}), 7.28 – 7.10 (2H, m, CH_{ar}), 4.73–4.70 (1H, d, *J* = 12, C-1), 3.82 – 3.78 (1H, m, C-3), 3.70–3.66 (1H, m, C-5), 3.63–3.59 (1H, m, C-4), 3.42–3.37 (1H, m, C-2), 3.30–3.25 (2H, m, C-6), 1.92 (3H, s, CH₃).

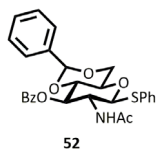
General Discussion



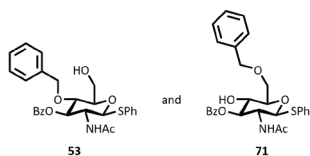
2-(Dimethoxymethyl)naphthalene 2-Naphtylaldehyde (1 g, 6 mmol) was dissolved in anhydrous methanol (10 ml) and stirred at room temperature. Trimethylorthoformate (770 μ l, 7 mmol, 1.1 eqv) and camphor sulfonic acid (0.01 eq, 1-10 mg) were added before the reaction was stirred for 15 minutes. The reaction mixture was quenched with Et₃N and concentrated *in vacuo* to give dimethylnaphtylacetal **66** in quantitative yield (1.21 gram). ¹H NMR (400 MHz, CDCl₃) δ : 8.00 – 7.78 (4H, m, CH_{ar}), 7.64 – 7.41 (3H, m, CH_{ar}), 5.56 (1H, s, CH), 3.38 (6H, s, 2 \times CH₃).



N-((2S,4aR,6S,7R,8R,8aS)-8-Hydroxy-2-phenyl-6-(phenylthio)hexahydropyrano[3,2-d][1,3]dioxin-7-yl)acetamide Diol **50** (3.72 mg, 11.87 mmol, 1 eqv) was dissolved in anhydrous acetonitrile (2 ml). Dimethylbenzaldehyde acetal (2.15 ml, 14.25 mmol, 1.2 eqv) and coppertriflate (214 mg, 0.59 mmol, 0.05 eqv) were added, and the reaction mixture was sonicated for 60 minutes. The reaction mixture was concentrated *in vacuo* and purified by column chromatography (5% CH₃OH in CH₂Cl₂) to obtain the 3-OH product **51** (4.40 g, 94%). ¹H NMR (400 MHz, CDCl₃) δ : 7.51-7.22 (10H, m, CH_{ar}), 5.10 (1H, d, *J* = 2, CH), 4.80-4.77 (1H, d, *J* = 12, C-1), 3.91-3.66 (3H, m, C-3, C-5, C-4), 3.58-3.35 (3H, m, C-2, C-6), 2.00 (3H, s, CH₃). The naphtyl acetal **70** was installed by the same procedure.

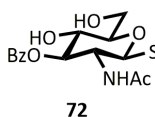


(2S,4aR,6S,7R,8R,8aS)-7-Acetamido-2-phenyl-6-(phenylthio)hexahydropyrano[3,2-d][1,3] dioxin-8-yl benzoate 3-OH Thiophenol glycoside **51** (70 mg, 0.17 mmol, 1 eqv) was suspended in anhydrous CH₂Cl₂. Benzoylchloride (22 μ l, 0.19 mmol, 1.1 eqv) and pyridine (28 μ l, 0.35 mmol, 2 eqv) were added and the reaction mixture was stirred overnight at room temperature. The reaction mixture was diluted with CH₂Cl₂ (5 ml) and washed with brine (3 \times 5 ml). The organic layer was dried over Na₂SO₄ and concentrated *in vacuo*. The product was purified by column chromatography (EtOAc in heptane) to give traces of product **52**. ¹H NMR (400 MHz, CDCl₃) δ 8.06 – 7.87 (2H, m, 2 \times CH_{ar}), 7.67 – 7.30 (3H, m, 3 \times CH_{ar}), 5.69 (1H, d, *J* = 9.5, C-1), 5.61 – 5.45 (3H, m, CCHO and C-5), 4.59 – 4.27 (2H, m, CH₂), 3.89 (2H, m, C-3 and C-4), 3.66 (1H, m, C-2), 1.91 (3H, s, NHAc). ¹³C NMR (100 MHz, CDCl₃) δ 169.7 (C=O), 165.3 (C=O), 137.3 (C_q), 129.0 (CH_{ar}), 128.8 (CH_{ar}), 128.4 (CH_{ar}), 128.1 (CH_{ar}), 126.0 (CH_{ar}), 100.4 (C-1), 78.0 (C-4), 73.6 (C-5), 72.1 (C_qCO), 69.8 (C-2), 67.6 (C-3), 52.4 (CH₂). **MS**: found 528.33 [M+Na⁺], calculated for [C₂₈H₂₇NO₆S + Na⁺] 528.15.



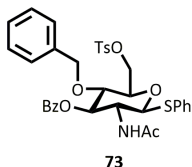
(2S,3R,4R,5S,6R)-3-Acetamido-5-(benzyloxy)-6-(hydroxymethyl)-2-(phenylthio)tetrahydro-2H-pyran-4-yl benzoate

Thiophenolglycoside **52** (32 mg, 0.06 mmol, 1 eqv) was suspended in borane.THF (0.1 ml, 1 M). The reaction mixture was cooled to 0 °C before TMSOTf (16 µl, 0.08 mmol, 1.4 eqv) was added slowly. The reaction mixture was quenched with CH₃OH (0.5 ml) and Et₃N (0.1 ml) and concentrated *in vacuo*. Flash column chromatography was performed to give the opened 4-OH **53** and 6-OH **71** products. ¹H NMR (400 MHz, CDCl₃) δ 7.38–7.18 (6H, m, 6 × CH_{ar}), 6.88–6.76 (3H, m, 3 × CH_{ar}), 6.76–6.63 (6H, m, 6 × CH_{ar}), 5.11 (1H, dd, *J* = 15.1, 9.1 Hz, CH), 4.58 (2H, m, CH₂), 4.77–4.61 (1H, m, αH), 4.00–3.69 (3H, m, C-3, C-5, C-4), 3.66–3.50 (3H, m, C-2, C-6). ¹³C NMR (100 MHz, CDCl₃) δ 170.0 (C=O), 167.0 (C=O), 137.1 (Ar-C_q), 133.6 (Ar-C_q), 133.2 (Ar-C_q), 131.7 (Ar-C_q), 129.8 (CH_{ar}), 129.1 (CH_{ar}), 129.0 (CH_{ar}), 128.6 (CH_{ar}), 128.5 (CH_{ar}), 128.4 (CH_{ar}), 128.1 (CH_{ar}), 128.0 (CH_{ar}), 127.8 (CH_{ar}), 87.2 (α-CH), 79.4 (CH₂), 77.3 (C-3), 77.2 (C-5), 76.8 (C-4), 74.9 (CH₂), 53.3 (C-2), 23.2 (CH₃).



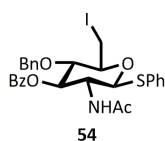
(2S,3R,4R,5S,6R)-3-Acetamido-5-hydroxy-6-(hydroxymethyl)-2-(phenylthio)tetrahydro-2H-pyran-4-yl benzoate

Acetal **52** (50 mg, 0.09 mmol, 1 eqv) was suspended in 60% AcOH (1 ml) in H₂O (5 ml) and stirred for three hours at room temperature. The reaction mixture was concentrated *in vacuo* and purified by column chromatography (5% CH₃OH in CH₂Cl₂) to give the desired product **72** (35 mg, 98% yield). ¹H NMR (400 MHz, CDCl₃) δ 8.08–7.78 (2H, m, CH_{ar}), 7.67–7.23 (8H, m, CH_{ar}), 5.64 (1H, d, *J* = 9.5 Hz, C-1), 5.55–5.36 (2H, m, CH₂), 4.85 (1H, d, *J* = 10.5, C-3), 4.43–4.07 (2H, m, C-4 and C-5), 3.85–3.75 (2H, m, *J* = 15.2, 9.9, C-6), 3.58 (1H, m, C-2), 1.82 (3H, s, CH₃).



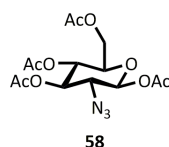
(2S,3R,4R,5S,6R)-3-Acetamido-5-(benzyloxy)-2-(phenylthio)-6-((tosyloxy)methyl)tetrahydro-2H-pyran-4-yl benzoate

6-OH Thiophenol glycoside **53** (11.30 mg, 0.02 mmol, 1 eqv) was dissolved in anhydrous CH₂Cl₂. Tosylchloride (4.61 mg, 0.02 mmol, 1.1 eqv) and pyridine (18 µl, 0.22 mmol, 10 eqv) were added and the reaction mixture was stirred at room temperature overnight. The reaction mixture was concentrated *in vacuo*. Analysis by mass spectrometry showed traces of the desired product **73**. ¹H NMR (400 MHz, CDCl₃) δ 7.99 (2H, d, *J* = 7.9 Hz, 2 × CH_{ar}), 7.81 (2H, d, *J* = 8.4 Hz, 2 × CH_{ar}), 7.58 (2H, m, *J* = 16.1, 11.8 Hz, 2 × CH_{ar}), 7.49–7.35 (4H, m, 4 × CH_{ar}), 7.35–7.28 (4H, m, 4 × CH_{ar}), 7.18 (3H, m, 3 × CH_{ar}), 7.11–6.97 (2H, m, 2 × CH_{ar}), 5.61 (1H, d, *J* = 9.5 Hz, C-1), 5.31 (2H, s, CH₂), 4.70 (1H, d, *J* = 10.5, C-3), 4.43–4.11 (2H, m, C-4 and C-5), 3.98–3.66 (2H, m, C-6), 3.61 (1H, d, *J* = 9.6, C-2), 2.41 (3H, s, CH₃), 1.87 (3H, m, CH₃). MS: found 684.25 [M+Na⁺], calculated for [C₃₅H₃₅NO₈S₂ + Na⁺] 684.77



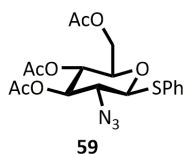
(2S,3R,4R,5S,6S)-3-Acetamido-5-hydroxy-6-(iodomethyl)-2-(phenylthio)tetrahydro-2H-pyran-4-yl benzoate Tosyl alcohol **73** (150 mg, 0.26 mmol, 1 eqv) was dissolved in anhydrous DMF. Sodiumiodine (**87** mg, 0.58 mmol, 2.2 mg) was added and the reaction mixture was stirred for 3 hours at 100 °C. The reaction mixture was cooled to r.t. before it was diluted with toluene and concentrated *in vacuo*. Column chromatography was performed with 5% CH₃OH in CH₂Cl₂ to give the product **54** with minor impurities.

¹H NMR (400 MHz, CD₃OD) δ 8.03 – 7.88 (1H, m, CH_{ar}), 7.69 – 7.60 (1H, m, CH_{ar}), 7.60 – 7.50 (1H, m, CH_{ar}), 7.50 – 7.43 (2H, m, 2 × CH_{ar}), 7.40 (2H, m, 2 × CH_{ar}), 7.33 – 7.10 (3H, m, 3 × CH_{ar}), 5.29 – 5.16 (1H, t, *J* = 8 and 20, C-3), 4.90 (1H, d, *J* = 10.4, C-1), 4.15–4.07 (1H, m, C-4), 3.85–3.82 (1H, m, C-5), 3.72 – 3.57 (1H, m, C-2), 3.45–3.40 (2H, m, CH₂) 2.64 (3H, s, NHAc).



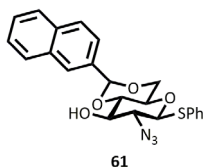
(3R,4R,5S,6R)-3-Azido-6-(hydroxymethyl)tetrahydro-2H-pyran-2,4,5-triol D-Glucosamine **48** (1.50 g, 6.96 mmol, 1 eqv) was suspended in CH₃OH (30 ml), K₂CO₃ (3.17 g, 23.96 mmol, 3.3 eqv), and CuSO₄·5H₂O (17.37 mg, 0.07 mmol, 0.01 eqv). Sulfonyl diazotransfer reagent **56** (2.26 g, 8.35 mmol, 1.2 eqv) was added followed by stirring at room temperature overnight. The suspension was filtered

over celite and concentrated *in vacuo*. To the crude reaction mixture was added acetic anhydride (0.28 ml, 31.32 mmol, 4.5 eqv), triethylamine (0.38 ml, 27.84 mmol, 4 eqv) and the reaction mixture was stirred for 1.5 hour at room temperature. The reaction mixture was purified by flash column chromatography (EtOAc / Heptane) to give β-anomer **58** (600 mg, 41% yield). ¹H NMR (400 MHz, CDCl₃) δ 5.45 (1H, d, *J* = 8.6, C-1), 5.14 – 4.60 (2H, m, C-3 and C-5), 4.18 (1H, dd, *J* = 12.5, 4.5, C-4), 3.95 (1H, dd, *J* = 12.5, 2.1, C-2), 3.70 (1H, m, C-6a), 3.54 (1H, dd, *J* = 9.9, 8.5 Hz, C-6b), 2.34 – 1.70 (12H, m, 4 × CH₃).



(2R,3S,4R,5R,6S)-2-(Acetoxymethyl)-5-azido-6-(phenylthio)tetrahydro-2H-pyran-3,4-diyl diacetate Tetracetate **58** (680 mg, 1.82 mmol) was dissolved in anhydrous

CH₂Cl₂ (10 ml) and stirred at room temperature. Thiophenol (224 μl, 2.18 mmol, 1.2 eqv) was added followed by the addition of SnCl₄ (1.27 ml, 1 M, 0.7 eqv) over 5 minutes. The reaction mixture was refluxed overnight. The reaction mixture was quenched with NaHCO₃ (100 ml) and extracted with CH₂Cl₂ (3 × 25 ml). The organic layer was washed with brine (3 × 20 ml), dried over Na₂SO₄ and concentrated *in vacuo*. The reaction mixture was purified by flash column chromatography (5% CH₃OD in CH₂Cl₂) to give thioglycoside **59** (400 mg, 52% yield). ¹H NMR (400 MHz, CDCl₃) δ 7.59 – 7.16 (5H, m, 5 × CH_{ar}), 5.56 (1H, d, *J* = 5.6, C-1), 5.26 (1H, t, *J* = 9.9, C-3), 5.11 – 4.91 (2H, m, C-4 and C-5), 4.51 (1H, ddd, *J* = 10.3, 5.2, 2.1, C-2), 4.29 – 3.84 (2H, m, CH₂), 2.07 – 1.85 (9H, m, 3 × CH₃).



61

(2*S*,4*aR*,6*S*,7*R*,8*R*,8*aS*)-7-Azido-2-(naphthalen-2-yl)-6-(phenylthio)hexahydropyrano[3,2-*d*][1,3]dioxin-8-ol β -Thiophenol 2-azido triacetate **59** (81 mg, 0.19 mmol, 1 eqv) was dissolved in anhydrous CH_3OH (2 ml) and stirred at room temperature. Sodium methoxide (2 drops) were added and the reaction mixture was stirred for 1 hour at room temperature. The reaction mixture was quenched with Dowex H^+ , filtered, and concentrated *in vacuo* to give the de-acetelated azido sugar **60** in quantitative yield. The crude diol **60** (65 mg, 0.19 mmol, 1 eqv) was dissolved in anhydrous DMF (0.5 ml) and stirred at room temperature. Molsieves (3 Å, 3 pcs), 2-naphtyldimethylacetal (42 mg, 0.21 mmol, 1.1 eqv), and camphor sulfonic acid (cat.) were added, followed by stirring at room temperature overnight. The reaction mixture was quenched with Et_3N (1 ml) and concentrated *in vacuo* to give the desired product ($R_f = 0.57$ in 10% CH_3OH in CH_2Cl_2 , 8 mg, 9% yield). No further analysis were performed.

References

- Benito-Alifonso, D.; Tremell, S.; Sadler, J. C.; Berry, M.; Galan, M. C., *Chem. Commun.* **2016**, 52 (27), 4906-9.
- Eising, S.; Lelivelt, F.; Bongers, K. M., *Angew. Chem.* **2016**, 55 (40), 12243-12247.
- Dewar, M. J. S.; Dougherty, R. C., *J. Am. Chem. Soc.* **1962**, 84 (13), 2648-8.
- Bandyopadhyay, A.; Cambray, S.; Gao, J., *J. Am. Chem. Soc.* **2017**, 139 (2), 871-878.
- Mbua, N. E.; Li, X.; Flanagan-Steet, H. R.; Meng, L.; Aoki, K.; Moremen, K. W.; Wolfert, M. A.; Steet, R.; Boons, G. J., *Angew. Chem.* **2013**, 52 (49), 13012-5.
- Yu, S. H.; Zhao, P.; Sun, T.; Gao, Z.; Moremen, K. W.; Boons, G. J.; Wells, L.; Steet, R., *J. Biol. Chem.* **2016**, 291 (8), 3982-9.
- Xie, S.; Lopez, S. A.; Ramstrom, O.; Yan, M.; Houk, K. N., *J. Am. Chem. Soc.* **2015**, 137 (8), 2958-66.
- Bailey, A. S.; White, J. E., *J. Chem. Soc. B.* **1966**, (9), 819-821.
- Dommerholt, J.; van Rooijen, O.; Borrmann, A.; Guerra, C. F.; Bickelhaupt, F. M.; van Delft, F. L., *Nat. Commun.* **2014**, 5, 5378.
- Hein, J. E.; Fokin, V. V., *Chem. Soc. Rev.* **2010**, 39 (4), 1302-15.
- Hein, J. E.; Tripp, J. C.; Krasnova, L. B.; Sharpless, K. B.; Fokin, V. V., *Angew. Chem.* **2009**, 48 (43), 8018-21.
- Prescher, J. A.; Dube, D. H.; Bertozzi, C. R., *Nature* **2004**, 430 (7002), 873-7.
- Siegrist, M. S.; Swarts, B. M.; Fox, D. M.; Lim, S. A.; Bertozzi, C. R., *FEMS Microbiol. Rev.* **2015**, 39 (2), 184-202.
- Yang, J.; Seckute, J.; Cole, C. M.; Devaraj, N. K., *Angew. Chem.* **2012**, 51 (30), 7476-9.
- Patterson, D. M.; Jones, K. A.; Prescher, J. A., *Mol. BioSystems.* **2014**, 10 (7), 1693-7.
- Spate, A. K.; Scharf, V. F.; Hafner, J.; Niederwieser, A.; Mayer, T. U.; Wittmann, V.; Beilstein *J. Org. Chem.* **2014**, 10, 2235-42.
- Stairs, S.; Neves, A. A.; Stockmann, H.; Wainman, Y. A.; Ireland-Zecchini, H.; Brindle, K. M.; Leeper, F. J., *ChemBioChem* **2013**, 14 (9), 1063-7.
- Stockmann, H.; Neves, A. A.; Stairs, S.; Brindle, K. M.; Leeper, F. J., *Org. Biomol. Chem.* **2011**, 9 (21), 7303-5.

General Discussion

19. Besanceney-Webler, C.; Jiang, H.; Zheng, T.; Feng, L.; Soriano del Amo, D.; Wang, W.; Klivansky, L. M.; Marlow, F. L.; Liu, Y.; Wu, P., *Angew. Chem.* **2011**, *50* (35), 8051-6.
20. Tytgat, H. L.; Schoofs, G.; Driesen, M.; Proost, P.; Van Damme, E. J.; Vanderleyden, J.; Lebeer, S., *Microb. Biotechnol.* **2015**, *8* (1), 164-8.
21. Royle, L.; Radcliffe, C. M.; Dwek, R. A.; Rudd, P. M., *Methods Mol. Biol.* **2006**, *347*, 125-43.
22. Mazia, D.; Schatten, G.; Sale, W., *J. Cell. Biol.* **1975**, *66* (1), 198-200.
23. Notter, M. F., *Cell. Res.* **1988**, *177* (2), 237-46.
24. Besanceney-Webler, C.; Jiang, H.; Wang, W.; Baughn, A. D.; Wu, P., *Bioorg. Med. Chem. Lett.* **2011**, *21* (17), 4989-92.
25. Liechti, G. W.; Kuru, E.; Hall, E.; Kalinda, A.; Brun, Y. V.; VanNieuwenhze, M.; Maurelli, A. T., *Nature* **2014**, *506* (7489), 507-10.
26. MacKenzie, D. A.; Sherratt, A. R.; Chigrinova, M.; Kell, A. J.; Pezacki, J. P., *Chem. Commun.* **2015**, *51* (62), 12501-4.
27. Mordini, A.; Valacchi, M.; Nardi, C.; Bindi, S.; Poli, G.; Reginato, G., *J. Org. Chem.* **1997**, *62* (24), 8557-8559.
28. Sminia, T. J.; Zuilhof, H.; Wennekes, T., *Carbohydr. Res.* **2016**, *435*, 121-141.
29. Green, T. W., *Protective Groups in Organic Synthesis*. Wiley: **1999**.
30. Hemeon, I.; Bennet, A. J., *Synthesis-Stuttgart* **2007**, (13), 1899-1926.
31. Liu, F.; Aubry, A. J.; Schoenhofen, I. C.; Logan, S. M.; Tanner, M. E., *ChemBioChem* **2009**, *10* (8), 1317-20.
32. Champasa, K.; Longwell, S. A.; Eldridge, A. M.; Stemmler, E. A.; Dube, D. H., *Mol. Cell. Proteomics* **2013**, *12* (9), 2568-2586.
33. Goon, S.; Kelly, J. F.; Logan, S. M.; Ewing, C. P.; Guerry, P., *Mol. Microbiol.* **2003**, *50* (2), 659-71.
34. Schirm, M.; Soo, E. C.; Aubry, A. J.; Austin, J.; Thibault, P.; Logan, S. M., *Mol. Microbiol.* **2003**, *48* (6), 1579-92.
35. Thibault, P.; Logan, S. M.; Kelly, J. F.; Brisson, J. R.; Ewing, C. P.; Trust, T. J.; Guerry, P., *J. Biol. Chem.* **2001**, *276* (37), 34862-70.
36. Zunk, M., *An investigation into the synthesis of 5,7-diamino-3,5,7,9-tetradeoxy-nonulosonic acid analogues*. thesis **2014**.
37. Liav, A. L.; Sharon, N., *Carbohydr. Res.* **1973**, *30* (1), 109-126.
38. Zehavi, U.; Sharon, N., *J. Org. Chem.* **1972**, *37* (13), 2141-5.
39. Ireland, R. E.; Wilcox, C. S., *J. Org. Chem.* **1980**, *45* (2), 197-202.
40. Sanapala, S. R.; Kulkarni, S. S., *Org. Lett.* **2016**, *18* (15), 3790-3.
41. Lee, Y. J.; Kubota, A.; Ishiwata, A.; Ito, Y., *Tetrahedron Lett.* **2011**, *52* (3), 418-421.
42. Gammon, D. W.; Steenkamp, D. J.; Mavumengwana, V.; Marakalala, M. J.; Mudzungu, T. T.; Hunter, R.; Munyololo, M., *Bioorg. Med. Chem.* **2010**, *18* (7), 2501-2514.
43. Tran, A. T.; Jones, R. A.; Pastor, J.; Boisson, J.; Smith, N.; Galan, M. C., *Adv. Synth. Catal.* **2011**, *353* (14-15), 2593-2598.
44. Ghosh, M.; Dulina, R. G.; Kakarla, R.; Sofia, M. J., *J. Org. Chem.* **2000**, *65* (24), 8387-8390.
45. Lin, Y. H.; Ghosh, B.; Mong, K. K., *Chem. Commun.* **2012**, *48* (88), 10910-2.
46. Zunk, M.; Kiefel, M. J., *Tetrahedron Lett.* **2011**, *52* (12), 1296-1299.
47. Akaji, K.; Aimoto, S., *Tetrahedron* **2001**, *57* (9), 1749-1755.
48. van den Bos, L. J.; Codee, J. D.; van Boom, J. H.; Overkleeft, H. S.; van der Marel, G. A., *Org. Biomol. Chem.* **2003**, *1* (23), 4160-5.
49. Still, W. C.; Kahn, M.; Mitra, A., *J. Org. Chem.* **1978**, *43* (14), 2923-2925.



Appendix A

Supporting Information for Chapter 3

"Another brick in the wall, Part 2"

The Wall, Pink Floyd"

Chemical procedures

General considerations and materials. All moisture sensitive reactions were carried out under an argon atmosphere, using oven-dried glassware and allowed to cool in a dessicator over anhydrous KOH, unless otherwise stated. Solvents were either purified over aluminum oxide under argon using a Pure Solv 400 solvent purification system (Innovative Technology, Amesbury, USA) (CH_2Cl_2 , Et_2O , toluene, and THF) or dried over 3Å molecular sieves (CH_3CN , CH_3OH , Pyridine). Commercially acquired chemicals were used without further purification unless stated otherwise. Burgess reagent was purchased from VWR. 6-methyl-tetrazine-ATTO-488, 6-methyl-tetrazine-Sulfo-cy3, 6-methyl-tetrazine-Sulfo-cy5, and DBCO-cy7 were purchased from Jena Bioscience. Dipyrindyl tetrazine, DBCO-PEG4-Fluor 545, DBCO-cy3, and DBCO-cy5 were purchased from Sigma-Aldrich. All chemicals used for biological experiments were dissolved in DMSO (1mM stock solution).

TLC. Analytical TLC was performed using plates of silica gel (Merck 60 F-254 on aluminium) and visualised using UV light (254 nm), ninhydrin, or KMnO_4 . Silica gel 60 (70–230 mesh) from Screening Devices, Amersfoort was used for flash chromatography.

NMR. ^1H NMR and ^{13}C NMR spectra were recorded on a Bruker Avance III 400 spectrometer (observation of ^1H nucleus 400 MHz, and of ^{13}C nucleus 100 MHz). Chemical shifts are reported in parts per million (ppm), calibrated on the residual peak of the solvent, whose values are referred to tetramethylsilane (TMS, $\delta\text{TMS} = 0$), as the internal standard. Coupling constants (J) are reported in Hertz (Hz), rounded to the nearest 0.5 Hz. Signal assignment was made from unambiguous chemical shifts and COSY, HSQC, HMBC, and DEPT.

ESI. Electrospray ionization (ESI) mass analyses were performed on a Finnigan LXQ, while high-resolution ESI mass analyses were recorded on a Thermo Scientific Q Exactive High-Resolution mass spectrometer.

IR. Infra-red analyses were performed on a Bruker FT-IR spectrometer.

2. Biological procedures

Bacterial strains and growth conditions. *Escherichia coli* K12 MG1655 was grown in Lysogeny Broth (LB) at 37 °C with aeration. *Akkermansia muciniphila* MucT (ATTC BAA-835)² was cultivated anaerobically at 37 °C in CP medium supplemented with mucin² or in CP-PT synthetic medium as described earlier.³ Specific Growth Rate (SGR, μ) was calculated by measuring OD_{600} over a two hour growth interval during exponential growth of the respective bacteria. Colony forming units (CFUs) were determined by plating the *E. coli* on LB agar plates and *A. muciniphila* on CP-mucin agar plates.

Application of bioorthogonal probes. PG probes were filter-sterilised before application in microbiological experiments. The probes were added in a 50 μM final concentration to early exponential bacterial cultures ($\text{OD}_{600} = 0.3$). The cultures were incubated for 1-2 hours in the presence of the

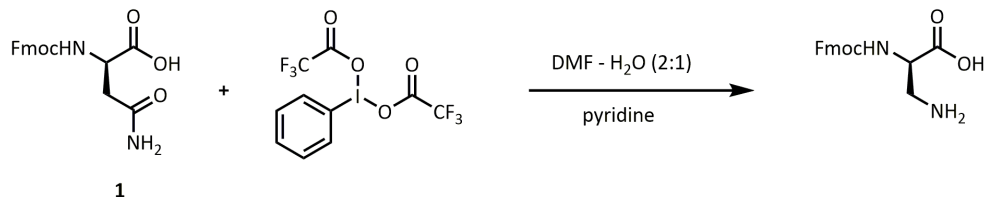
probes. Glycan probes were added at inoculation and left to grow for a maximum of 3 hours in the case of *E. coli* and a maximum of 10 hours for *A. muciniphila*. After growth in the presence of the probes, cells were harvested and washed twice with PBS. Cells were resuspended in click buffer (CuAAC: 100 μ M CuSO₄, 100 mM ascorbic acid, 5 μ l of 1 mM dye dissolved in PBS; SPAAC: 5 μ l of 1 mM dye in PBS; tetrazine: 5 μ l of 1 mM dye in PBS). Cells were incubated for 1 hour at room temperature in the dark and then washed three times with PBS. Supernatant was resuspended in PBS prior to confocal or FCM analysis.

Confocal microscopy. Ibidi μ -slide 8 well glass bottom slides were used for confocal microscopy. Solutions for CuAAC or SPAAC labelling were freshly prepared with the following concentrations: CuSO₄ (100 mM) and Ascorbic acid (5 mM). Live imaging was performed on a confocal laser scanning microscope (Leica TCS SP8X system). Alexa Fluor 488 was excited using an argon laser (488nm), and fluorescence emission was detected from 500 to 540 nm. Cy-3, Cy-5, and Cy-7 were excited using a white light continuum laser. Images were captured using a 633 water-immersion objective with a numeric aperture of 1.2 and a pinhole set to 1 Airy unit. Images were processed by ImageJ or LAS AF Lite.

Flow cytometry. Analysis of isonitrile PG probe **6** and Ac₄FucAz labelled cells was performed using a BD FACSaria III Cell Sorter fitted with BD FACSDIVA 6.1 software. Single cell populations were localised using FCS and SSC (488-nm 20 mW laser), gating pulse Area over pulse Width. Furthermore, Syto-9 (488-nm, 502 LP, 530/30 nm) was used to counterstain and distinguish cells from background. The Cy-3 tetrazine labelled PG isonitrile probe was detected using a 561-nm (50 mW) laser including a 582 nm band pass filter with a bandwidth of 15 nm. Cy-7 DBCO labelled Ac₄FucAz was identified using a 633-nm (18 mW) laser using a 755 nm Long Pass filter followed by a 780 nm band pass filter with a bandwidth of 60 nm. Data analysis was performed using FlowJo X (Treestar, USA).

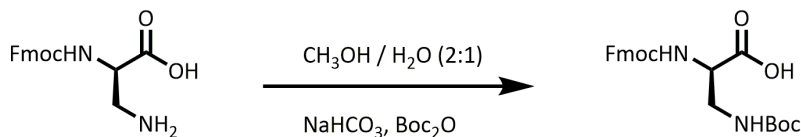
Supplementary details chemical synthesis

FmocHN-D-Dap-OH



FmocHN-D-Asn-OH (1: 1.00 g, 2.82 mmol) was suspended in a solution of [bis(trifluoroacetoxy)iodo] benzene (1.82 g, 4.23 mmol) in DMF – H₂O (20 ml, 2:1 (v/v)). After 15 minutes, Pyridine (0.45 ml, 5.64 mmol) was added and the reaction mixture was stirred at room temperature for 16h. The solvent was removed under reduced pressure, and the oily residue was dissolved in H₂O (20 ml). Concentrated HCl (1 ml) was added, and the acidified solution was washed with Et₂O (3 × 10 ml). The aqueous phase was adjusted to pH 6 with 2 M NaOH solution, and the resulting precipitate was filtered, washed with H₂O (3 × 10 ml), ice-cold EtOH (3 × 10 ml), and Et₂O (3 × 10 ml). Product (0.92 g, 99%) was obtained after drying *in vacuo* as a beige powder. ¹H NMR (400 MHz, DMSO-d₆): δ 7.83 (d, *J* = 7.5, 2H, CH_{ar}), 7.64 (d, *J* = 7.4, 2H, CH_{ar}), 7.36 (t, *J* = 7.5, 2H, CH_{ar}), 7.28 (t, *J* = 7.5, 2H, CH_{ar}), 6.78 (d, *J* = 6.7, 1H, CONH), 4.26–4.16 (m, 3H, Fmoc CH and CH₂), 2.98–2.60 (m, 3H, α-H and β-CH₂). ¹³C NMR (100 MHz, DMSO-d₆): δ 171.1 (C=O), 143.8 (C_q), 143.7 (C_q), 140.6 (C_q), 127.6 (CH aryl), 127.1 (CH aryl), 125.1 (CH aryl), 120.0 (CH aryl), 99.5 (α-H), 46.6 (CH₂), 40.4 (CH), 38.8 (CH₂). ESI (+) *m/z* [M + H] calculated for C₁₈H₁₉N₂O₄: 327.135; found: 327.3 IR (FT-ATR) 2948, 1722, 1686, 1636, 1577, 1532, 1491. **Reference:** Y.H. Lau & D. R. Spring, Synlett 2011, 13, 1917–1919.4

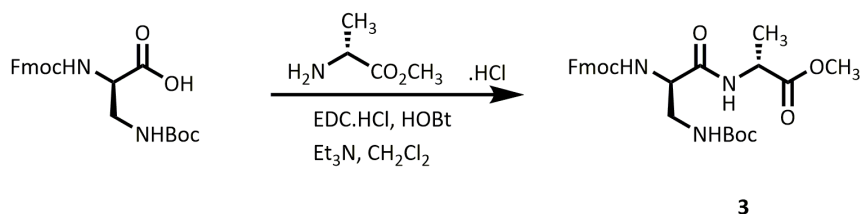
FmocHN-D-DapNHBOC-OH



FmocHN-D-Dap-OH (1.0 eqv, 0.86 g, 2.64 mmol) was dissolved in CH₃OH / H₂O (2:1, 10 ml) and NaHCO₃ (1.1 eqv, 0.24 g, 2.89 mmol). A solution of Boc₂O (1.1 eqv, 0.63 g, 2.89 mmol) in CH₃OH (45 ml) was added over 10 minutes. The reaction mixture was stirred and after TLC indicated completion it was diluted with aq. Na₂CO₃ (sat., 30 ml). The aqueous layer is washed with EtOAc (3 × 10 ml) and then acidified using 10% NaHSO₄ solution until pH = 1. The acidified aqueous layer was extracted with EtOAc (3 × 10 ml) and the combined organic layers were dried over Na₂SO₄ and concentrated *in vacuo*. The product was isolated as a white foam in 79% yield that was used without any further purification. R_f 0.28 (10% CH₃OH in CH₂Cl₂ (v/v)). ¹H NMR (400 MHz, CDCl₃): δ 7.76 (d, *J* = 7.5, 2H, CH_{ar}), 7.59 (d, *J*

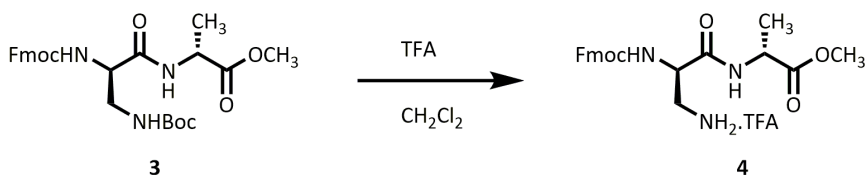
= 7.5, 2H, CH_{ar}), 7.40 (t, *J* = 7.4, 2H, CH_{ar}), 7.30 (t, *J* = 7.4, 2H, CH_{ar}), 6.38 (s, 1H, NH), 5.15 (s, 1H, NH), 4.37 (m, 3H, CH and CH₂ Fmoc), 4.22 (t, *J* = 6.9, 1H, α-H), 3.59 (d, *J* = 45.6, 2H, β-CH₂), 1.45 (s, 9H, 3 × CH₃). ¹³C NMR (100 MHz, CDCl₃): δ 173.2 (C=O Fmoc), 157.2 (C=O Boc), 143.8 (C_q Fmoc), 143.6 (C_q Fmoc), 141.3 (C_q Fmoc), 141.3 (C_q Fmoc), 127.7 (Ar-CH), 127.1 (Ar-CH), 125.2 (Ar-CH), 119.9 (Ar-CH), 80.6 (C_q-Boc), 67.4 (CH₂-Fmoc), 55.4 (CH-Fmoc), 47.1 (α-CH), 42.1 (β-CH₂), 28.2 (3 × CH₃). For detailed experimental analysis see Ichikawa et al.⁵

FmocHN-D-DapNH₂Boc-D-Ala-OCH₃



FmocHN-D-DapNH₂Boc-OH (1 eqv, 0.86 g, 2.02 mmol), EDC.HCl (1.1 eqv, 0.43 g, 2.22 mmol), D-Ala-OCH₃.HCl (1 eqv, 0.37 g, 2.02 mmol), HOBT (1.2 eqv, 0.3 g, 2.24 mmol), and DIPEA (3.4 eqv, 1.10 ml, 6.9 mmol) were charged to a roundbottomflask containing CH₂Cl₂ (20 ml). The reaction mixture was stirred for 16h before it was washed with NH₄Cl (3 × 10 ml). The organic layer was washed with NaHCO₃ (3 × 10 ml) and brine (3 × 10 ml). The organic layer was dried over Na₂SO₄ and concentrated *in vacuo*. The product was purified by flash column chromatography with 2.5% CH₃OH in CH₂Cl₂ to afford dipeptide **3** (0.91 g, 82%). *R*_f 0.20 (10% CH₃OH in CH₂Cl₂ (v/v)). ¹H NMR (400 MHz, CDCl₃): δ 7.77 (2H, d, *J* = 7.4, CH_{ar}), 7.61 (2H, d, *J* = 7.4, CH_{ar}), 7.40 (2H, t, *J* = 7.4, CH_{ar}), 7.31 (2H, t, *J* = 7.5, CH_{ar}), 6.22 (1H, s, NH), 5.14 (1H, s, NH), 4.54 – 4.32 (3H, m, FmocCH₂ and FmocCH), 4.24 (2H, t, *J* = 7.2, 2 × α-H), 3.67 (3H, s, OCH₃) 3.54 (2H, d, *J* = 28.1, CH₂), 1.46 (9H, s, BocCH₃), 1.38 (3H, d, *J* = 7.1, CH₃). ¹³C NMR (100 MHz, CDCl₃): δ 171.1 (4 × C=O), 143.7 (C_q), 141.3 (C_q), 127.7 (CH_{ar}), 126.6 (CH_{ar}), 125.1 (CH_{ar}), 120.0 (CH_{ar}), 67.6 (FmocCH₂), 52.5 (OCH₃), 48.3 (CH), 47.6 (CHCH₃), 47.3 (FmocCH), 42.8 (CH₂), 28.3 (3 × CH₃), 14.1 (CH₃). HRMS: found 534.2213 [M + Na⁺], calculated for [C₂₇H₃₃N₃O₇ + Na⁺] 534.2216

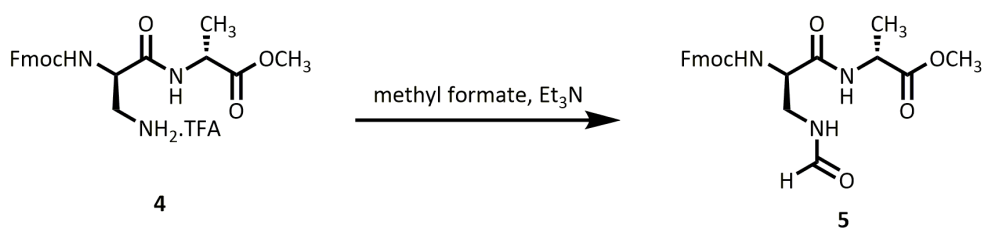
FmocHN-D-Dap-D-Ala-OH



The protected dipeptide (**3**: 0.91 g, 0.5 mmol) was dissolved in anhydrous CH₂Cl₂ (20 ml) and cooled in an ice-water bath. TFA (5.0 ml, excess) was added dropwise to the reaction mixture over 5 minutes.

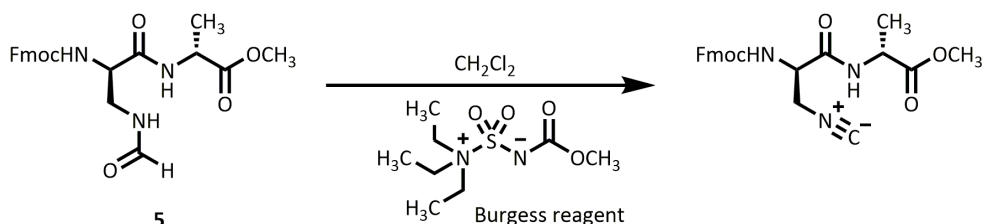
The reaction mixture was stirred for 2 h and afterwards concentrated *in vacuo* to obtain the product **4** as an oily residue that was used without any further purification. R_f 0.6 (10% CH₃OH in CH₂Cl₂ + Et₃N (v/v)). ¹H NMR (400 MHz, CDCl₃): δ 7.94 – 7.80 (2H, d, J = 8.0, CH_{ar}), 7.71 (2H, t, J = 7.0, CH_{ar}), 7.60 – 7.12 (4H, m, CH_{ar}), 4.45 – 4.07 (3H, m, FmocCH₂ and FmocCH), 3.91 (2H, m, 2 × α-H), 3.74 (3H, s, OCH₃), 3.13 – 2.66 (2H, m, CH₂), 1.21 (3H, d, J = 7.4 Hz, CH₃). ¹³C NMR (100 MHz, CDCl₃): δ 186.6 (C=O), 170.2 (C=O), 143.8 (C_q), 141.3 (C_q), 127.7 (CH_{ar}), 127.1 (CH_{ar}), 125.1 (CH_{ar}), 120.0 (CH_{ar}), 54.3 (CH₂), 52.6 (OCH₃), 48.4 (CHCH₃), 47.1 (CH), 43.6 (CH₂), 28.3 (3 × CH₃), 17.6 (CH₃). HRMS: found 412.1872 [M + H⁺], calculated for [C₂₂H₂₆N₃O₅ + H⁺] 412.1872

FmocHN-D-DapNHCOH-D-Ala-OCH₃



Amine **4** (1.0 eqv, 1.38 g, 2.62 mmol) was dissolved in a solution of Et₃N (1.5 eqv, 0.55 ml, 3.93 mmol) and methyl formate (excess, 20 ml). The reaction mixture was refluxed for 16h after which it was concentrated *in vacuo*. The product was purified by column chromatography (5% CH₃OH in CH₂Cl₂) to give 124 mg of **5** (11%). R_f 0.46 (10% CH₃OH in CH₂Cl₂ (v/v)). ¹H NMR (400 MHz, DMSO-d₆): δ 7.87 (d, J = 7.5, 2H, CH_{ar}), 7.70 (d, J = 7.4, 2H, CH_{ar}), 7.46 (d, J = 8.4, 2H, CH_{ar}), 7.31 (t, J = 7.5, 2, 2H, CH_{ar}), 4.34 – 4.16 (m, 3H, FmocCH, FmocCH₂), 3.64 (m, 2H, 2 × α-CH), 3.15 (s, 3H, OCH₃), 2.84 (q, J = 7.2, 2H, β-CH₂), 1.28 (d, J = 7.3, 2.8, 3H, CH₃). ¹³C NMR (100 MHz, DMSO-d₆): δ 172.8 (C=O Fmoc), 162.3 (C=O Ester), 161.5 (C=O formamide), 143.7 (C_{ar}), 139.4 (C_{ar}), 128.8 (C_{ar}), 127.6 (C_{ar}), 125.2 (C_{ar}), 120.0 (C_{ar}), 65.6 (FmocCH₂), 48.5 (OCH₃), 47.6 (CHCH₃), 47.3 (FmocCH), 46.6 (CH₂), 45.7 (CH₂CH), 45.4 (β-CH₂CH), 16.7 (CH₃). HRMS: found 462.1635 [M + Na⁺], calculated for [C₂₃H₂₅N₃O₆ + Na⁺] 462.1641

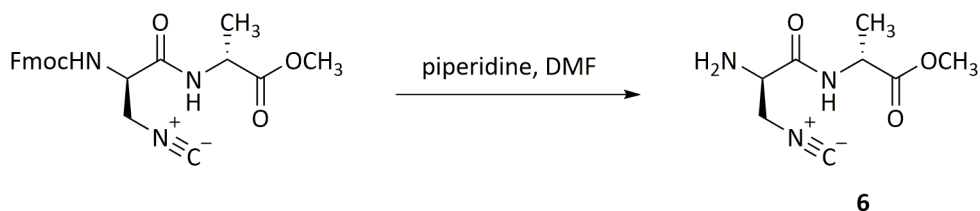
FmocHN-D-DapNC-D-Ala-OCH₃



Formamide dipeptide **5** (1 eqv, 124 mg, 0.28 mmol) was co-evaporated three times with toluene before it was dissolved in CH₂Cl₂ (5 ml). Burgess reagent (1.5 eqv, 100 mg, 0.42 mmol) was added and stirring was continued for 16h at room temperature. Another 1.5 equivalents of Burgess reagent was added

before it was stirred for an additional 16h at room temperature. The reaction mixture was diluted with CH_2Cl_2 (5 ml) and quenched with H_2O (5ml). The organic layer was isolated and dried over Na_2SO_4 and subsequently concentrated *in vacuo*. The crude product was purified by column chromatography (2.5% CH_3OH in CH_2Cl_2) to obtain 26 mg (22% yield) of the isonitrile as a white powder. R_f 0.30 (5% CH_3OH in CH_2Cl_2 (v/v)). $^1\text{H NMR}$ (400 MHz, DMSO-d_6): δ 7.77 (d, $J = 7.4$, 2H, CH_{ar}), 7.58 (d, $J = 7.4$, 2H, CH_{ar}), 7.41 (t, $J = 7.4$, 2H, CH_{ar}), 7.32 (t, $J = 7.4$, 1.6, 2H, CH_{ar}), 6.91 (d, $J = 6.8$, 1H, NH), 5.90 – 5.60 (m, 1H, NH), 4.72 – 4.30 (m, 2H, FmocCH_2), 4.23 (t, $J = 6.7$, 1H, FmocCH), 3.75 (s, 3H, OCH_3), 3.82 – 3.62 (m, 5H, CH_2CH), 3.44 (q, $J = 21.6$, 7.3, 1H, CHCH_2), 3.14 (q, $J = 7.3$ Hz, 1H, CHCH_3), 1.59 – 1.36 (m, 3H, CH_3). $^{13}\text{C NMR}$ (100 MHz, DMSO-d_6): δ 170.7 (C=O Fmoc), 165.1 (C=O OCH_3), 157.8 (NC), 141.5 (C_q), 139.4 (C_q), 125.9 (CH_{ar}), 125.2 (CH_{ar}), 123.0 (CH_{ar}), 118.1 (CH_{ar}), 65.6 (FmocCH_2), 51.4 (OCH_3), 50.7 (CHCH_2), 46.6 (FmocCH), 46.5 (CHCH_3), 44.2 (CH_2CH), 16.1 (CH_3). **HRMS**: found 444.1528 [$\text{M} + \text{Na}^+$], calculated for [$\text{C}_{23}\text{H}_{23}\text{N}_3\text{O}_5 + \text{Na}^+$] 422.1530

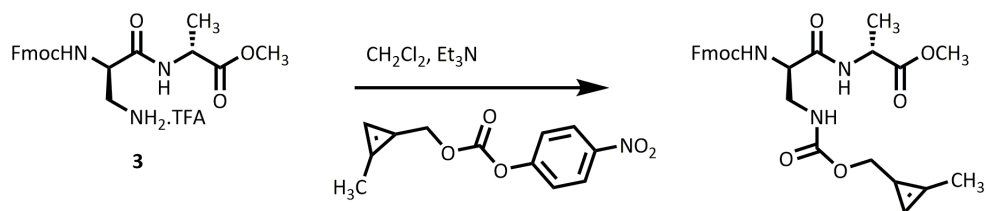
$\text{H}_2\text{N-D-DapNC-D-Ala-OCH}_3$



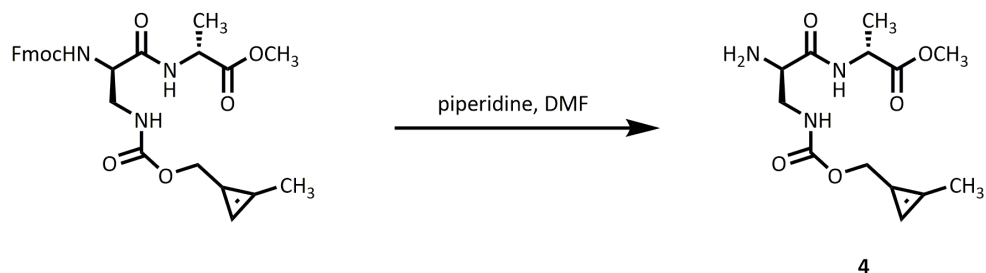
The Fmoc protected starting material (1 eqv, 20 mg, 0.047 mmol) was dissolved in anhydrous DMF (0.5 ml). Piperidine (1.1 eqv, 5 μl , 0.052 mmol) was added and the reaction mixture was stirred at room temperature for 1 hour. The reaction mixture was transferred to an aqueous layer (10 mL) which was washed with diethylether (3 \times 5 ml). The aqueous layer was lyophilised overnight to obtain the product in quantitative yield as a white solid. **IR** ν_{max} (neat) 2152, 1650 cm^{-1} ; $^1\text{H NMR}$ (400 MHz, DMSO-d_6): δ 6.91 (d, $J = 6.8$, 1H, NH_2), 5.90 – 5.60 (m, 1H, NH), 3.75 (s, 3H, OCH_3), 3.82 – 3.62 (m, 5H, CH_2CH), 3.44 (q, $J = 21.6$, 7.3, 1H, CHCH_2), 3.14 (q, $J = 7.3$, 1H, CHCH_3), 1.59 – 1.36 (m, 3H, CH_3). **MS**: found 200.19 [$\text{M} + \text{H}^+$], calculated for [$\text{C}_8\text{H}_{13}\text{N}_3\text{O}_3 + \text{H}^+$] 200.22 **HRMS**: found 200.1028 [$\text{M} + \text{H}^+$], calculated for [$\text{C}_8\text{H}_{13}\text{N}_3\text{O}_3 + \text{H}^+$] 200.1030

Cyclopropene dipeptide

The *p*-nitrophenol-activated cyclopropene carbonate was prepared as previously described.⁶⁻⁸

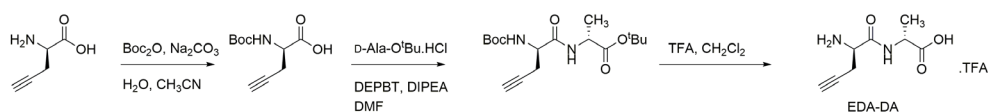
FmocHN-D-cyclopropene-D-Ala-OCH₃

FmocHN-amino-D-Ala-OCH₃ **3** (78 mg, 0.15 mmol, 1 eqv) was dissolved in anhydrous CH₂Cl₂ (2 ml) and triethylamine (52 μ l). Cyclopropene carbonate (42 mg, 0.17 mmol, 1.1 eqv) was slowly added and the reaction mixture was stirred overnight at room temperature. The reaction mixture was quenched with water and extracted with CH₂Cl₂ (3 \times 5 ml). The organic layer was washed with NH₄Cl (3 \times 5 ml), dried over Na₂SO₄ and concentrated *in vacuo*. Column chromatography performed with 2.5% CH₃OH in CH₂Cl₂ gave the product (40 mg, 50% yield) as a white solid. **¹H NMR** (400 MHz, CDCl₃): δ 8.30–8.06 (1H, m, CH_{cyc}), 7.76 (2H, d, *J* = 7.5, CH_{ar}), 7.60 (2H, d, *J* = 7.6, CH_{ar}), 7.40 (2H, t, *J* = 7.5, CH_{ar}), 7.31 (2H, d, *J* = 7.5, CH_{ar}), 6.91 (1H, d, *J* = 8, CH), 6.55 (1H, s, NH), 4.55 (1H, *J* = 7.2, α HCH₃), 4.40 (2H, d, *J* = 7.2, FmocCH₂), 4.32 (1H, m, α CH), 4.23 (1H, t, *J* = 7.2, FmocCH), 4.00–3.93 (2H, m, CH₂), 3.74 (3H, s, OCH₃), 3.69–3.50 (2H, m, CH₂), 3.57 (1H, s, α CH), 2.12 (3H, s, CH₃), 1.65 (1H, m, CH), 1.42 (3H, d, *J* = 7.3, CH₃), 1.26 (1H, s, CH). **¹³C NMR** (100 MHz, CDCl₃): δ 205.9 (C=O), 172.9 (C=O), 169.8 (C=O), 162.3 (C=O), 143.7 (C_q), 141.3 (C_q), 127.75 (CH_{ar}), 127.1 (CH_{ar}), 126.2 (C_q), 125.1 (CH_{ar}), 119.9 (CH_{ar}), 115.6 (CH_{cyc}), 72.9 (CH₂), 67.5 (FmocCH₂), 53.4 (OCH₃), 52.7 (CHCH₂), 48.3 (α HCH₃), 47.1 (FmocCH), 43.3 (CH₂CH), 17.9 (CH₃), 17.1 (CH), 11.6 (CH₃). **MS**: found 544.61 [M+Na⁺], calculated for [C₂₈H₃₁N₃O₇ + Na⁺] 544.21 **HRMS**: found 544.2041 [M+Na⁺], calculated for [C₂₈H₃₁N₃O₇ + Na⁺] 544.2060

H₂N-D-cyclopropene-D-Ala-OCH₃

FmocHN-D-cyclopropene-D-Ala-OCH₃ (40 mg, 0.07 mmol) was dissolved in anhydrous DMSO (1 ml). Piperidine (100 μ l) were added and the reaction mixture was stirred at room temperature for 1h. After

full conversion the reaction mixture was diluted with water (10 ml) and washed with diethyl ether (3 × 5 ml). The aqueous phase was lyophilised to give product **4** as a solid (20 mg, 87% yield). The dipeptide was dissolved in DMSO (5 mM stock solution), divided over 13 Eppendorf vials and stored at -20 °C. **MS**: found 300.1 [M+H⁺], calculated for [C₁₃H₂₁N₃O₅ + H⁺] 300.16 **¹H NMR** (400 MHz, D₂O): δ 6.58 (1H, d, *J* = 14.3, CH), 4.16 – 4.08 (3H, m, αH and CH₂), 3.92-3.51 (3H, m, αH + CH₂), 2.67 (3H, s, OCH₃), 1.43 (3H, bd, CH₃), 1.41 (3H, s, CH₃), 1.31 (1H, d, *J* = 7.2, CH). **¹³C NMR** (100 MHz, D₂O): δ 174.2 (C=O), 167.3 (C=O), 162.3 (C=O), 126.7 (C_q), 116.0(CH_{cyc}), 101.6 (αCHCH₂), 73.4 (CH₂NH), 53.1 (OCH₃), 48.9 (αCHCH₃), 44.6 (CH₂CH), 22.4 (CH₃), 16.6 (CH), 11.1 (CH₃). **HRMS**: found 300.1550 [M+H⁺] calculated for [C₁₃H₂₁N₃O₅ + H⁺] 300.1554



Scheme S1. The alkyne-dipeptide EDA-DA was synthesised according to the procedure described by Liechti and co-workers.¹ The dipeptide was filter-sterilised before application in microbiological experiments.

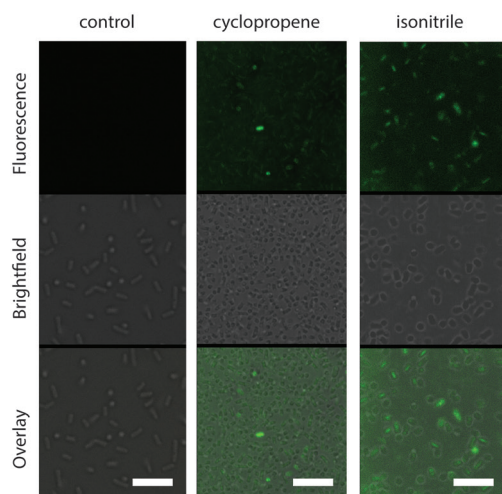


Figure S1. Confocal microscopy images of *E. coli* with control, cyclopropene, and isonitrile dipeptide PG probe. Cy-3 azide or tetrazine dyes were used for staining. Scale bar = 5 µm.

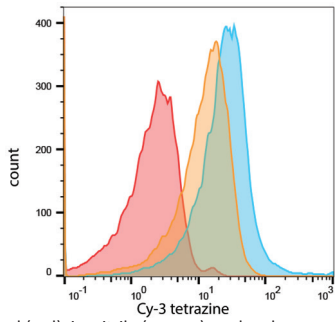


Figure S2. Flow cytometry analysis of control (red), isonitrile (orange), and cyclopropene (blue) PG dipeptide probes in *E. coli*.

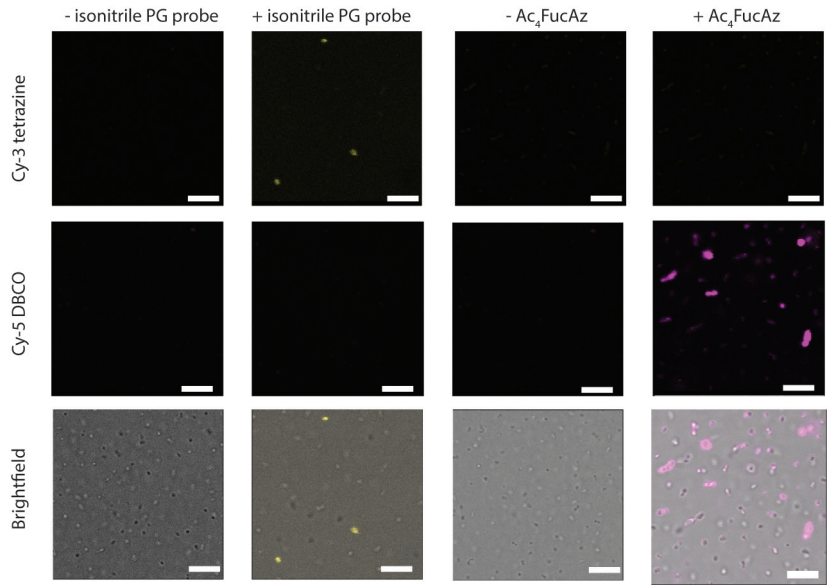


Figure S3. Overview labelling with control experiments in *A. muciniphila*. Scale bar = 5 μ m.

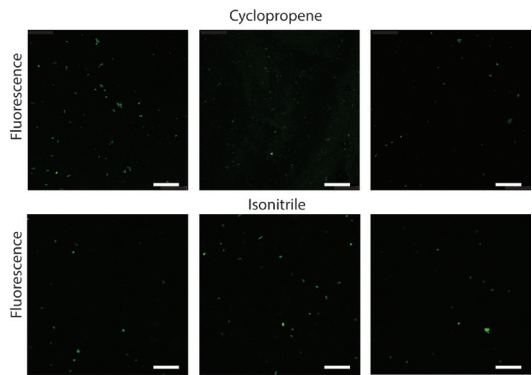


Figure S4. Additional confocal images of PG labelling with cyclopropene and isonitrile PG dipeptide probe. Scale bar = 25 μm .

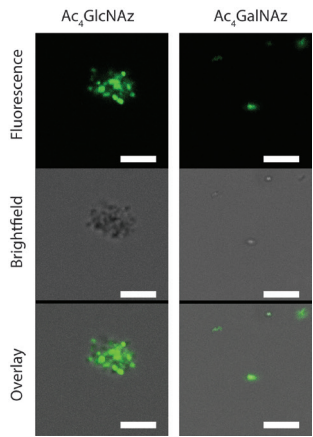


Figure S5. Overview of labelling with Ac₄GlcNAz and Ac₄GalNAz in *A. muciniphila*. Scale bar = 5 μm .

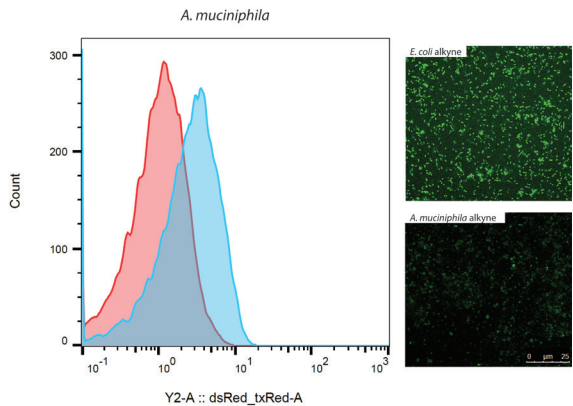


Figure S6. FCM analysis and confocal microscopy images of alkyne PG probe in *A. muciniphila*.

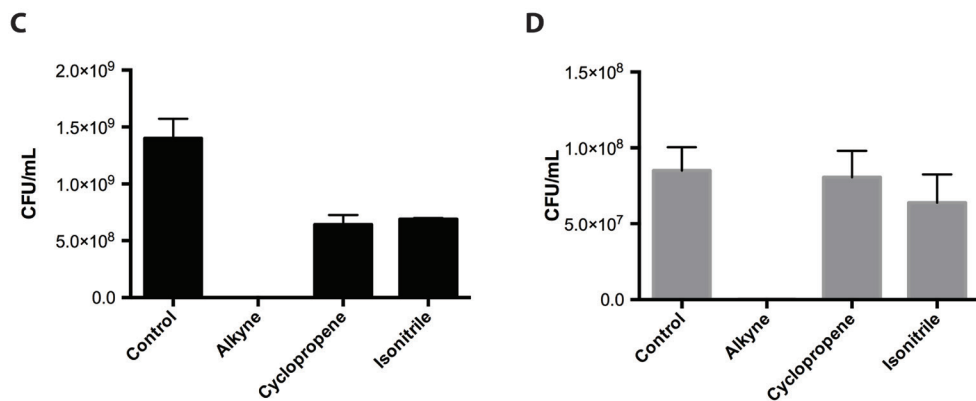


Figure S7. Colony-Forming Units (CFU) of *E. coli* (C) and *A. muciniphila* (D) after incubation with PG probes and their click reaction (CuAAC with azide for Alkyne PG probe and invDA or [4+1] with tetrazine for cyclopropene or isonitrile probes).



Summary

"Main theme"

More, Pink Floyd"

Summary

After DNA and proteins, carbohydrates are the third language of life. **Chapter 1** introduces the reader to this class of biomolecules, also called sugars or glycans, that can be found on the outer surface of almost all cells and plays a critical role as the social messengers of a cell. Although our knowledge about the role of glycans in eukaryotic cells has increased considerably in recent decades, our understanding of the glycan layer on bacterial cells is still very limited. Besides the carbohydrates that are present in both eukaryotes and prokaryotes an additional wide range of unique (e.g. microbial sialic acid), often very complex (e.g. pseudaminic acid), carbohydrates are present in prokaryotes. This chapter briefly introduces two research fields, carbohydrate chemistry and chemical biology, that when combined provide a powerful way to investigate the biological role of these unique bacterial carbohydrates at the molecular level. This chemistry-based approach, termed chemical microbiology, often starts with the development of a chemical synthesis for a target bacterial carbohydrate. Subsequently, the synthetic route towards this target allows for the introduction of unnatural functional groups, like chemical reporters, that result in the molecular tools needed to study their biological function. The studies described in this thesis, focus on developing such molecular tools to study the role of glycans and glycoconjugates in human gut bacteria and human-associated bacteria.

Chapter 2 provides an overview of metabolic oligosaccharide engineering (MOE) a popular chemical biology technique to label glycans in living cells. In MOE, carbohydrate derivatives are synthesised with unnatural chemical reporters and used to study their incorporation in glycans of eukaryote to prokaryote species. The progress in this field over the last 6 years is reviewed in detail with a special emphasis on the synthesis of the unnatural carbohydrates from commercially available sources. The principle behind MOE is that these unnatural carbohydrates with e.g. azide, alkyne, cyclopropene, or isonitrile chemical reporter groups, are still recognised by the endogenous enzymes in the cell that salvage this new carbohydrate. In this way they can enter the associated biochemical pathways and end up in newly biosynthesised cellular glycans. Subsequent labelling techniques, such as strain promoted azide alkyne cycloaddition or tetrazine ligation, enable the visualisation of these incorporated unnatural carbohydrates with for instance fluorescence microscopy.

Metabolic labelling is further explored in **chapter 3**. Key cell envelope glycoconjugates in the mucin-degrading gut microbiota member, *Akkermansia muciniphila*, were subjected to chemistry-based functional analysis, with *Escherichia coli* being used as a control species. Two novel non-toxic peptidoglycan (PG) probes were designed and synthesised to investigate the presence of PG in this species. Their design was based on the natural D-alanine dipeptide motif found in PG. Inspired by the fact that D-alanine dipeptide-derivatives were previously reported to be incorporated in newly synthesised PG, we synthesised a cyclopropene and isonitrile D-alanine dipeptide. Our probes proved to be non-toxic, as shown by growth and viable count analysis, and were therefore superior over existing PG probes. Another beneficial property was that the probes also did not influence the specific growth rate of *A. muciniphila* or *E. coli*. The PG probes were successfully incorporated into the peptidoglycan layer of *A. muciniphila* and visualised using a tetrazine click-ligation with a fluorophore. Our analysis proved for the first time that *A. muciniphila* has a PG layer. Besides PG labelling, we also investigated metabolic labelling of other glycoconjugates on the outer surface of *A. muciniphila*. This part of the

study showed that azido-monosaccharide derivatives of *N*-acetylglucosamine, *N*-acetylgalactosamine, and fucose are successfully processed by *A. muciniphila* salvage pathways and incorporated into its surface glycoconjugates. Especially Ac₄FucAz was readily processed by the recently discovered L-fucose salvage pathway of *A. muciniphila*. The two compatible labelling techniques were next combined in a dual labelling experiment. Our isonitrile dipeptide peptidoglycan probe and Ac₄FucAz were successfully incorporated into *A. muciniphila*. Subsequent fluorescent labelling with bio-orthogonal techniques resulted in dual labelling of peptidoglycan and fucose-containing glycans in live *A. muciniphila* cells.

With the positive results of MOE in *A. muciniphila* in hand, **chapter 4** describes the further investigation of MOE. After successful validation of our Ac₄FucAz probe for MOE in *Bacteroides fragilis* we continued their application in other human gut microbiota members, including the butyrate-producing *Anaerostipes rhamnosivorans*, *Intestimonas butyriciproducens*, and *Eubacterium hallii*. Labelling of these human gut microbes proved to be rather challenging with a-specific cellular labelling with the fluorophore being the major problem. Initial results, however, did show that a 6-azido-L-rhamnose probe resulted in fluorescent labelling of *A. rhamnosivorans*, which provides initial evidence for the existence of an as of yet undocumented salvage pathway. In this species the 6-azido-fucose probe was not salvaged. Via confocal microscopy and flow cytometry analysis we observed that the 6-azido-rhamnose probe was selective for *A. rhamnosivorans* in the presence of *A. muciniphila*. Such a co-culture experiment is a first step in mimicking the complex human gut microbiome. For *E. hallii* Ac₄GalNAz gave clear metabolic labelling and the majority of the cell population could be labelled with the fluorescent dye after a strain-promoted azide alkyne cycloaddition. Other glycan probes (Ac₄GlcNAz, Ac₄FucAz, and Neu5Az) also resulted in labelling, but not as prominent as Ac₄GalNAz. Surprisingly, MOE has never been reported for the common lab strain *Escherichia coli* MG1655. Curious to investigate this in more detail we started MOE in *E. coli*. However, no labelling was obtained when Ac₄GlcNAz probe was added to *E. coli*, most likely due to the fast growth, metabolism and turnover. Only, when fresh Ac₄GlcNAz probe was added every 30 minutes, metabolic labelling in *E. coli* was observed. To further investigate the influence of GlcNAc metabolism in *E. coli* on MOE, single-gene knock-outs of *E. coli* GlcNAc metabolism from the Keio collection were investigated. Labelling was observed for NagA (*N*-acetyl glucosamine 6 P deacetylase) and NagK (*N*-acetyl-D-glucosamine kinase) *E. coli* mutants. Both enzymes are involved in the last step of the biosynthesis towards UDP-*N*-acetylglucosamine. When the overall *E. coli* metabolism was inhibited, after addition of the respiration inhibitor sodium azide, no metabolic labelling was observed. These results indicate that MOE in *E. coli* is possible, but challenging and can only be performed under specific circumstances.

An investigation into the total synthesis of pseudaminic acid, a sialic acid produced by specific human-associated prokaryotes, is described in **chapter 5**. Sialic acids are typically found at the terminal positions of surface glycoconjugates in both eukaryotes and prokaryotes. Other related microbial sialic acids are legionaminic and acinetaminic acid. The total synthesis of these microbial sialic acids is notoriously difficult, as exemplified by the fact that only a few chemical synthesis routes towards them are currently known. Our total synthesis of pseudaminic acid started from the readily available amino acid L-threonine that was transformed into a key versatile Garner aldehyde derivative intermediate.

Summary

With this aldehyde in hand, the Henry nitro-aldol condensation reaction was investigated. After studying numerous conditions, such as asymmetric catalysis or elongated reaction times, and extensive optimisation efforts we were never able to obtain the Henry reaction product to continue with this route. As an alternative, a tethered aminohydroxylation was investigated for its ability to introduce the key functional group and stereochemistry onto an intermediate obtained from the Garner aldehyde derivative. This reaction indeed gave the desired amino-alcohol motif in the correct stereochemistry, but another diastereomer proved very difficult to separate from the desired product. After some additional transformations and protection steps we obtained a derivative in which the primary alcohol could be oxidised to provide a hexose intermediate that resembles the hexose intermediate present in pseudaminic acid biosynthesis. This key hexose intermediate will likely enable a subsequent Barbier reaction, a chain elongation step, in future studies. With most of the key transformations accomplished, the completion of a pseudaminic total synthesis based on L-threonine should soon be possible. Besides finishing the total synthesis, future work should also focus on adapting this synthesis route to allow installation of chemical reporter groups on pseudaminic acid for its application in MOE.

Chapter 6 is the general discussion about all the work mentioned in the other chapters. It also contains additional information and suggestions for further research in the field of chemical microbiology.



Samenvatting

"Backs to the Wall

The Wall, Pink Floyd"

DNA, eiwitten, lipiden en suikers zijn de bouwstenen van het leven. Deze thesis gaat specifiek in op de laatste klasse van biomoleculen: suikers, ook wel koolhydraten of glycanen genoemd. Suikers komen onder andere voor op de buitenkant van cellen waar ze een belangrijke rol vervullen als boodschapper van de cel. De laatste 10-tallen jaren is er veel bekend geworden over de rol van suikers in eukaryoten cellen, hoe anders is dat in prokaryoten cellen waar de rol van suikers nog niet is verklaard. Prokaryoten, zoals bacteriën, hebben naast suikers die ook voorkomen in eukaryoten cellen een eigen klasse van suikers. Deze suikers zijn vaak zeer complex (e.g. pseudamine zuur) en vormen dan ook de aanleiding tot het onderzoek beschreven in deze thesis. In **Hoofdstuk 1** worden er twee vakgebieden naarder verklaard: suiker chemie en chemische biologie. Technieken en kennis uit deze twee vakgebieden kunnen samen de biologische rol van deze unieke suikers in bacteriën nader bestuderen en verklaren. Deze chemische aanpak van het bestuderen van biologische processen, ook wel chemische microbiologie genoemd, start vaak met de organische synthese van zo'n complex bacteriële suiker. Door middel van organische synthese kunnen er onnatuurlijke chemische groepen aan het suiker gezet worden, moleculair gereedschap, die het mogelijk maken om de suikers te bestuderen in bacteriën. In deze thesis wordt de synthese van dit moleculair gereedschap en suikers nader beschreven waarna er wordt gekeken hoe suikers zich gedragen in menselijke darmbacteriën.

In **Hoofdstuk 2** wordt er een overzicht gegeven van een populaire chemische biologische techniek die het mogelijk maakt om suikers te labelen in levende cellen: metabool oligosacharide architectuur. In deze techniek worden suiker derivaten gesynthetiseerd met onnatuurlijke chemische handvaten die het mogelijk maken om glycanen te bestuderen in eukaryoten of prokaryoten cellen. Een overzicht van de afgelopen zes jaar zijn in dit hoofdstuk uiteengezet. De focus van dit overzicht is de beschrijving van de synthese van de onnatuurlijke suikers en het moleculaire gereedschap vanuit commercieel verkrijgbare bouwstenen. Voorbeelden van moleculair gereedschap zijn azide, alkyn, cyclopropeen, of isocyanide groepen. Dit moleculair gereedschap wordt volgens de standaard biochemische routes ingebouwd door de aanwezige enzymen ondanks de chemische aanpassingen aan de suikers. Na inbouw in de glycanen kan het moleculair gereedschap gelabeld worden met bijvoorbeeld fluorescerende groepen zodat dit bekeken kan worden onder fluorescentie microscopie. Voor deze laatste stap maakte we gebruik van een tweetal chemische reacties: de spanning gestuurde azide alkyn cycloadditie of de tetrazine ligatie.

Metabool labelen wordt ook toegepast in **Hoofdstuk 3**. Essentiële glycaan conjugaten in de belangrijke darmbacterië *Akkermansia muciniphila* werden nader onderzocht en geanalyseerd vanuit een chemische invalshoek. Als controle organisme werd *Escherichia coli* gebruikt. Twee nieuwe niet-toxische peptidoglycaan (PG) probes zijn ontwikkeld en gesynthetiseerd om PG aan te tonen in deze soorten. Het ontwerp was gebaseerd op het natuurlijke aminozuur D-alanine dipeptide wat aanwezig is in natuurlijk PG. In de literatuur waren al voorbeelden aanwezig van D-alanine dipeptide probes om PG aan te tonen. Als moleculair gereedschap hebben wij isocyanide en cyclopropeen groepen op het D-alanine dipeptide probe geplaatst. Na inbouw in PG is aangetoond dat onze probes niet toxisch waren en dus een verbetering zijn op de bestaande probes. Verder werd de groei van de bacteriën niet aangetast door de aanwezigheid van onze probes. De inbouw van de D-alanine dipeptide

probes werd aangetoond met de tetrazine ligatie met een fluorescerende groep aan het tetrazine conjugaat. Hiermee toonde we voor het eerst aan dat *A. muciniphila* een PG laag heeft. Naast de PG labelen hebben we ook andere suiker conjugaten gelabeld op de buitenkant van *A. muciniphila*. Verschillende monosaccharide met azide als molecuulair gereedschap, gebaseerd op de natuurlijke suikers *N*-acetylglucosamine, *N*-acetylgalactosamine, en fucose, werden succesvol ingebouwd in *A. muciniphila*. Met name azide-fucose werd goed ingebouwd wat de recent ontdekte fucose recycling synthese in *A. muciniphila* verder onderbouwd. Tot slot werden beide label technieken tegelijk toegepast in *A. muciniphila*. Zowel het isocyanide D-alanine dipeptide probe als de azide fucose werden ingebouwd in *A. muciniphila* en daarna met behulp van fluorescentie zichtbaar gemaakt.

De positieve resultaten van metabool labelen in *A. muciniphila* gaven aanleiding om dit verder door te zetten in andere darm-bacteriën in **Hoofdstuk 4**. Na validatie van de fucose azide probe in *Bacteroides fragilis* hebben we een selectie gemaakt van andere relevante darmbacteriën om metabool te labelen: *Anaerostipes rhamnosivorans*, *Intestimonas butyriciproducens*, en *Eubacterium hallii*. Het labelen van deze bacteriën was zeer problematisch gezien het hoge achtergrond signaal wat we vaak kregen. De eerste resultaten laten echter zien dat de 6-azide rhamnose probe opgenomen werd door *A. rhamnosivorans* en succesvol werd ingebouwd wat duidt op de aanwezigheid van een rhamnose recycling synthese. Ter vergelijking, 6-azide fucose werd niet ingebouwd in *A. rhamnosivorans*. Bovendien konden we aantonen dat de 6-azide rhamnose probe selectief was voor *A. rhamnosivorans* in een mengsel met *A. muciniphila*. Dit resultaat zet een eerste stap in het nabootsen van ons complexe darmsysteem. In *E. hallii* metabool labelen met azido-*N*-acetylgalactosamine gaf duidelijk fluorescente gelabeld bacteriën na een spanning gestuurde azide alkyn cycloadditie. Andere azide probes (Ac_4GlcNAz , Ac_4FucAz , and Neu5Az) gaven minder labelling vergeleken met Ac_4GalNAz . Metabool labelen in *Escherichia coli* MG1655 was, tot onze verbazing, nog nooit aangetoond. Daarvoor ondernamen we verder proeven om dit nader te bestuderen. De snelle groei, een hoog metabolisme, en de duplicatiesnelheid waren echter waarschijnlijke oorzaken waarom metabool labelen met Ac_4GlcNAz niet lukte. Alleen bij toevoeging van Ac_4GlcNAz probe elke 30 minuten was er labeling te zien in *E. coli*. Voor nader onderzoek hebben we ons gericht op mutanten van *E. coli* waarbij belangrijke metabole routes waren uitgeschakeld. Hierbij was labeling te zien voor NagA (*N*-acetyl glucosamine 6 P deacetylase) en NagK (*N*-acetyl-D-glucosamine kinase) *E. coli* mutanten. Wanneer het metabolisme volledig werd platgelegd, door toevoeging van natriumazide, was er geen metabool labeling te zien. Deze resultaten laten zien dat metabool labelen mogelijk is in *E. coli* maar dat het zeer uitdagend is en alleen onder specifieke niet natuurlijke condities werkt.

In **Hoofdstuk 5** komt de totaal synthese van pseudamine zuur aan bod. Dit is een suiker behorende bij de klasse van siaalzuren. Pseudamine zuur is een microbiële siaalzuur wat alleen gemaakt wordt door bacteriën die in contact staan met de mens. Siaalzuren zijn vaak te vinden op de uiteinden van glycoconjugaten in eukaryoten of prokaryoten. Naast pseudamine zuur is er ook legionamine en acinetamine zuur recentelijk ontdekt in bacteriën. De totaal synthese van deze microbiële siaalzuren is niet triviaal, desondanks zijn er enkele synthese routes bekend naar de siaalzuren. Onze route naar pseudamine zuur begon vanaf het commercieel verkrijgbare aminozuur L-threonine. Na enkele stappen

Samenvatting

was een belangrijk veelzijdig Garner aldehyde derivaat bereikt. Vanuit hier werd er een Henry reactie, een nitro-aldol condensatie, geprobeerd. Na verscheidene pogingen, zoals asymmetrische katalyse of lange reactie tijden, werd het product van deze reactie nooit volledig gevormd. Op zoek naar een alternatief om het amino-alcohol motief in het molecuul in te bouwen hebben we de gestuurde aminohydroxylatie geprobeerd. Tijdens deze reactie werd er een diastereomeer gevormd die lastig te scheiden bleek. Na enkele reactie waren we aangekomen bij de oxidatie van een primair alcohol. Na oxidatie zou dit product een precursor opleveren voor pseudamine zuur. Door tijdsgebrek zijn deze transformaties helaas niet voltooid en konden we de totaal synthese van pseudamine zuur tot op heden niet voltooien. Bij een succesvolle totaal synthese zou de volgende stap zijn het inbouwen van het moleculair gereedschap wat metabool labelen mogelijk moet maken.

Hoofdstuk 6 is een algemene discussie naar aanleiding van alle hoofdstukken. Verder geeft het aanvullende informatie en toekomst perspectieven in het vakgebied chemische microbiologie.



Acknowledgement

"Hey you"

The Wall, Pink Floyd

Acknowledgement

Childhood's End: Almost five years ago I entered Wageningen to start my PhD at the Laboratory of Organic Chemistry. I would like to thank many people for this unforgettable experience!

Han Zuilhof, my promotor: thank you very much for all the possibilities over the last 5 years to perform my research at the Laboratory of Organic Chemistry (ORC). I greatly enjoyed our talks about our common interest (beside science): chess!

Tom Wennekes, my co-promotor, neighbour at the Mörfelden-Walldorfplein in Wageningen, and most important my daily supervisor: your enthusiasm for science, and in particular sugars, is very inspiring. You always had time to discuss my project in more detail or to improve the general working and social atmosphere at ORC. Unfortunately, this changed quite a bit when you found an even nicer place at Utrecht University. The last years were hard for me without this daily face-to-face contact but we managed! Wish you all the best in your future career.

Lost for Words: I would like to thank all past and present members of the Laboratory of Organic Chemistry (ORC) for their help, scientific and non-scientific discussions, and nice working atmosphere: Sweccha Joshi, Ton Marcelis, Fred van Geenen, Jorin Hoogenboom, Stefanie Lange, Digvijay Gahtory, Maurice Franssen, Wilco Duivier, Willem van Heugten, Bas van den Berg, Teris van Beek, Satesh Gangarapu, Radostina Manova, Elly Geurtsen, Maarten Smulders, Rui Rijo da Costa Carvalho, Esther Roeven, Bram Bielen, Jaime Garcia Hartjes, Aleida Ruisch, Frank Versluis, Nagendra Bhairamadgi, Esther van Andel, Jorge Escorihuela Fuentes, Andrada But, Milou Santbergen, Jorick Bruins, Jose Maria Alonso Carnicero, Erik van Rozendaal, Judith Firet, Aline Debrassi, Quyen Nguyen, Anne-Marie Franssen, Bauke Albada, Florine Duval, Alexandre Villela, Carel Weijers, Cees van Rijn, Rickdeb Sen, Zhanhua Wang, Sjoerd Slagman, Fatima Garcia Melo, Floris van Delft, Sidharam Pujari, Medea Kosian, Annemieke van Dam, Cees van der Haar, Hendra Willemen, Umesh Chinnaswamy, Anke Trilling, Pepijn Beekman, Ian de Bus, Louis de Smet, Iris van Marwijk, Barend van Lagen, Pepijn Geutjes, Wouter Biesta, Yao Shen, Elbert van der Klift, Nathalja Berghuis, Jelmer van der Rijst, and Frank Claassen.

During the two years at ORC, I also spent a lot of time at the Laboratory of Microbiology headed by my co-promotor Willem de Vos. The field of microbes was new for me and was definitely not easy. Thanks to the great expertise, patience, and commitment of Hanne Tytgat I felt comfortable to work with all types of microbes and was able to learn new techniques every day. Thank you Willem and Hanne for all your help, guidance, and patience. Furthermore, I would like to thank Steven Aalvink, Nam Bui, Nico Claassens, Marcel Tempelaars, Jan Willem Borst, and Tjakko Abee for all their help to perform my experiments and my transformation into a microbiologist. In the medical microbiology group at the medical centre in Utrecht I furthermore received great help from Bart Bardoel during the flow cytometry measurements. Without all your help especially for chapter 3 and 4 I would never have been able to call myself a chemical microbiologist.

The PhD study-trips to Germany and Switzerland in 2013, and Canada in 2015 were great fun. Thank you all, especially Sweccha, Medea, and Digvijay for all the good moments and the organisation of the Canada trip.

During my PhD I received great help from my students Marijn Schrage, Hanna de Jong, Connie de Kock, Jona Merx, Igor Vozny, and Michiel Uiterweerd. I enjoyed guiding you all and the many great findings explored by all of you helped me in finding answers in the mysterious world of chemical microbiology. All the best in your future careers.

Outside The Wall: op de donderdagavonden was het tijd voor ontspanning en plezier bij de Wageningse Schaakvereniging. De absolute hoogtepunten waren de maandelijkse teamcompetitie wedstrijden met het 2^e team. Schakers, hartelijk dank voor alle leuke partijen, discussies, trainingen, etentjes, en alle activiteiten in en om Wageningen.

Round and Around: dat hardlopen net zo verslavend is als schaken had ik nooit gedacht. Anne Vissers en Jorin Hoogenboom hartelijk dank voor de vele kilometers die we samen hebben gelopen. Op naar de volgende kilometers.

Bas and Jorin, my two paranimphs and proud members of *The Usual Suspects*. Thank you for all the fun we had over the last years in and outside ORC: hiking (from the water in the Bay of Fundy to unicorns in Rhenen), beers from all over the world ("*Those who like it, like it a lot*"), movie- and game nights (*baaaa...*) and our FSB-meetings with Sjoerd and Jorick.

Tot slot mijn lieve familie: Jelmer en Marije, wat kom ik graag langs in Utrecht. Jelmer, grote broer, jouw aanpakkers-mentaliteit bewonder ik. Pap en Mam, jullie steun en liefde zijn niet te beschrijven in woorden, elke dag weer vormen jullie mijn inspiratiebron!

To the very best of times,

Tjerk

About the Author



Tjerk Jacco Sminia was born on April 15, 1989 in Amsterdam, The Netherlands. After graduation from secondary school (Trinitas College, Heerhugowaard) he started the bachelor pharmaceutical sciences, at VU University in Amsterdam, in September 2007. His bachelor degree was focussed on synthesis and analysis of derivatisation reagents for citric acid under the supervision of dr. Maikel Wijtmans and dr. Jeroen Kool. In 2010, he continued at VU University in a MSc Double Degree Program focussing on Drug Design and Synthesis. In the first year of this program a 5-month internship at the Drug Design and Synthesis group was conducted working on the synthesis of phosphodiesterase inhibitors under the supervision of dr. Kristina Orrling. In 2011, the Double Degree Program was continued at Copenhagen University, Denmark. His MSc thesis was performed in the group of dr. Daniel Sejer Pedersen working on the synthesis of unnatural α - and β -amino acids for peptidomimetics. To finalise the MSc Double Degree Program a literature thesis was performed on site-selective palladium cross-couplings reactions of poly-halogenated heterocycles. In October 2012 he obtained his MSc Double Degree from VU University and Copenhagen University. In January 2013 he started his PhD at the Laboratory of Organic Chemistry at Wageningen University under the supervision of dr. Tom Wennekes and Prof. dr. Han Zuilhof. The results of this work are described in this thesis. Tjerk currently works as a scientist at ChemConnection, focusing on process research of active pharmaceutical ingredients.

List of Publications

Probing peptidoglycan synthesis and L-fucose salvage in the gut commensal *Akkermansia muciniphila* with bioorthogonal chemical reporters

T. J. Sminia, H. L.P. Tytgat, S. Aalvink, H. de Jong, M. Tempelaars, H. Zuilhof, T. Abee, W. M. de Vos, and T. Wennekes

Manuscript in preparation

G_s Protein Peptidomimetics as Allosteric Modulators of the β_2 -Adrenergic Receptor

L-M. Boyhus, M. Danielsen, N. S. Bengtson, M. B. A. Kunze, X. Kubiak, T. J. Sminia, J. H. Løper, P. T. Tran, K. Lindorff-Larsen, S. G. F. Rasmussen, J. M. Mathiesen, D. S. Pedersen

Submitted to RSC Advances

Getting a grip on glycans: A current overview of the metabolic oligosaccharide engineering toolbox

T. J. Sminia, H. Zuilhof, and T. Wennekes,

Carbohydrate Research, **2016**, 435, 121-141.

Azide- and Alkyne-Functionalised α - and β^3 -Amino Acids

Tjerk J. Sminia and D. S. Pedersen

Synlett **2012**, 23, 2643-2646

Overview of completed training activities & education

| Discipline specific activities | organizing institute | year | oral/poster |
|---|----------------------|---------|-------------|
| Synthesis Summer School | HRSMC | 2013 | poster |
| Organic Chemistry symposium | NWO | 2013 | poster |
| SOC spring symposium | KNCV / SOC | 2013 | |
| SOC fall symposium | KNCV / SOC | 2014 | |
| SOC international symposium | KNCV / SOC | 2014 | |
| CHAINS symposium | NWO | 2014 | |
| CHAINS symposium | NWO | 2015 | poster |
| SOC spring symposium | KNCV / SOC | 2015 | |
| Advanced Organic Chemistry | ORC | 2013-16 | |
| Advanced Chemistry | ORC | 2013-16 | |
| SOC international symposium | KNCV / SOC | 2016 | oral |
| International carbohydrate symposium | ACS | 2016 | poster |
| CHAINS symposium | NWO | 2016 | poster |
| SOC spring symposium | KNCV / SOC | 2017 | oral |
| General courses | | | |
| PhD week | VLAG | 2013 | |
| Information literacy endnote course | WGS | 2013 | |
| Advanced Artwork | WGS | 2014 | |
| ACS on campus | ACS | 2014 | |
| PhD carousel | WGS | 2014-15 | |
| Scientific writing and presenting | WGS | 2014 | |
| Career Perspectives | WGS | 2016 | |
| NMR course - Mnova | Mestrelab | 2015 | |
| Optionals | | | |
| Preparation of research proposal | | 2013 | |
| PhD study tour to Germany and Switzerland | | 2013 | |
| Weekly group meetings | | 2013-16 | |
| Assistant open dag WUR | | 2014-16 | |
| Member VLAG PhD council | | 2014-15 | |
| Literature club Wennekes laboratory | | 2014-15 | |
| PhD study tour to Canada | | 2015 | Chairman |
| Assistant kinduniversiteit | | 2017 | |
| Education | | | |
| Bio-organic chemistry for life sciences | | 2013-14 | |
| Organic chemistry II | | 2013-14 | |
| Bio-organic chemistry | | 2013-15 | |
| Research methods | | 2014-17 | |
| Supervision of BSc and MSc students | | 2013-16 | |

The research presented in this thesis was financially supported by The Netherlands Organisation for Scientific Research (NWO) via a Veni grant (722.011.006).

Financial support from Wageningen University, Screening Devices B.V. , and Euro-Logo B.V. for printing this thesis are gratefully acknowledged.

Cover design: Tjerk J. Sminia

Printed by GVO drukkers & vormgevers B.V., Ede, The Netherlands