

**Recombinant mucin-type proteins as tools
for studies on the interactions between
Helicobacter pylori and its carbohydrate
receptors**

Yolanda Hlamazi Mthembu

Department of Laboratory Medicine, Institute of Biomedicine

Sahlgrenska Academy at University of Gothenburg

Gothenburg, Sweden, 2020

Cover illustration: Production of PSGL-1/mIgG2b through recombinant DNA Technology by Yolanda Hlamazi Mthembu

Recombinant mucin-type proteins as tools for studies on the interactions between *Helicobacter pylori* and its carbohydrate receptors

© Yolanda Hlamazi Mthembu 2020
Yolanda.mthembu@gu.se

ISBN (Print) 978-91-7833-946-4
ISBN (e-pub) 978-91-7833-947-1
<http://hdl.handle.net/2077/64540>

Printed in Gothenburg, Sweden 2020

Printed by Ale Tryckteam AB, Bohus



Had it not been for the goodness of the lord in my life, where would I be?

To my parents, Mackenzie and Edith Gumede who always believed in me! You kept me knowledgeable and exposed me to the world of reading and learning, indulging me to opportunities and encouraged me to do my best in everything.

ABSTRACT

Glycan-protein interactions are important in pathogen adhesion and infections. *H. pylori* has adhesins which enables it to bind to glycans on the gastric mucosa and, in the long run, cause gastric cancer. The reported current antibiotic regimen used in the treatment to eradicate *H. pylori* fails in 20% of the patients. A multivalent glycan inhibitor could offer a suitable alternative to anti-biotics by acting as a competitive inhibitor for the cell receptors, leading to the binding and elimination of the microbe. This thesis is focused around the use of genetically engineered CHO-K1 cells producing a recombinant mucin-type fusion protein, P-selectin glycoprotein ligand-1/mouse IgG2b (PSGL-1/mIgG2b), which is used as a scaffold for multivalent presentation of engineered bioactive O-linked glycans. Through the engineering of carbohydrate determinants mediating attachment or affecting the growth of *H. pylori*, potential inhibitors of *H. pylori* infection were created (paper I, II and III).

In paper I, we show that B4GALNT3 added a β 1,4-linked GalNAc to GlcNAc (LDN) irrespective of whether the latter was carried by a core 2, core 3 or extended core 1 chain. There was no correlation between *H. pylori* binding and the expression of LDN determinants on gastric mucins or a mucin-type fusion protein carrying core 2, 3 and extended core 1 O-glycans.

In paper II, The *H. pylori* experiments demonstrated that only PSGL-1/mIgG2b proteins with Le^b on core 3 inhibited BabA-mediated binding. On the other hand, the series of sialylated PSGL-1/mIgG2b proteins all demonstrated various degrees of inhibition of SabA-mediated binding, suggesting that SabA accepts various substitution of sLe^x for binding.

In paper III, we show by Western blot and LC-MS/MS that core 1, core 2, core 3 and extended core 1 chains could all carry the GlcNAc α 4Gal determinant following transient transfection of CHO-K1 cells. Preliminary results showed that PSGL-1/mIgG2b carrying the GlcNAc α 4Gal-terminal on core 1 and core 2 O-glycans did not inhibit the growth of *H. pylori*.

In paper IV, we show that the interaction of galectin-3 with the lubricating protein, lubricin, derived from osteoarthritis as opposed to healthy joints is dependent on core 2 O-glycans.

In conclusion, we have shown that glyco-engineering of a mucin-type fusion protein in CHO-K1 cells generates a powerful tool for investigations on O-glycan biosynthesis and microbial, in this case *H. pylori*, adhesion. The use of a mucin-type fusion protein as a carrier of frequent O-glycan substitution not only may increase the avidity of the reporter protein for its binding partner under study, but in addition mimics the structural context in which bioactive carbohydrate determinants are presented and used as microbial attachment sites at our mucosal surfaces.

Keywords: O-glycans, mucins, glycosyltransferases, *Helicobacter pylori*, microbial adhesion

ISBN (Print) 978-91-7833-946-4

ISBN (e-pub) 978-91-7833-947-1

SAMMANFATTNING PÅ SVENSKA

En infektion orsakad av bakterien *Helicobacter pylori* ger upphov till magkatarr, magsår och anses också kunna orsaka magcancer. För att *H. pylori* skall orsaka infektion krävs att bakterien binder till celler i magsäcken. Bakteriens bindning till cellytan förmedlas ofta av att proteiner på bakterien, s.k. adhesiner, binder till sockermolekyler på cellytan. För att studera hur dessa adhesiner binder till cellytans sockermolekyler har vi använt en modell i vilken cellytans sockermolekyler har återskapats på mucinliknande glykoproteiner (sockerbärande proteiner) i så kallade cellfabriker genom rekombinant produktion. Då de rekombinant producerade mucinliknande glykoproteiner har speciella egenskaper, bland annat genom att de bär flera kopior av de önskade sockerstrukturerna, är de speciellt lämpliga för att studera bakteriens bindning till socker. Som cellfabrik har använts en cellinje som heter CHO, vars kapacitet att bilda olika sockerstrukturer är väl kartlagd. Genom att i cellfabrikerna uttrycka de enzym (glykosyltransferaser) som bygger upp önskade sockerstrukturer, kan CHO cellen fås att göra det mucinliknande protein, PSGL-1/mIgG2b, som vi använt för att i detalj studera bakteriens sockerspecificitet.

I arbete I, II och III modifierade vi CHO celler genetiskt så de producerade PSGL-1/mIgG2b med funktionella receptorer för olika adhesiner på *H. pylori*, nämligen LabA, BabA and SabA. I papper I visade vi att enzymet B4GALNT3 kunde återskapa den påstådda LabA-bindande sockerstrukturen, LDN (GalNAc4GlcNAc) på *O*-bundna socker baserade på olika prekursor kedjor. Däremot kunde vi inte reproducera den bindning till LDN som tidigare studier visat för LabA-bärande *H. pylori*. Varken PSGL-1/mIgG2b, nativt mucin från mage eller serumalbumin bärande LDN band olika stammar av *H. pylori*.

I arbete II återskapade vi de BabA- och SabA-bindande sockerstrukturerna Le^b respektive sLe^x på olika *O*-bundna prekursor kedjor. Genom att hämma bindningen mellan *H. pylori* och radioaktivt märkta albuminkonjugat av Le^b respektive sLe^x visade vi att PSGL-1/mIgG2b med Le^b på prekursor kedjan "core 3" hämmande BabA-medierad bindning bäst. Den SabA-medierade bindningen av sLe^x-konjugatet hämmades mer eller mindre av alla sockerformer av PSGL-1/mIgG2b, vilket skulle kunna förklaras av att hämningen korrelerade till mängden sialinsyra på PSGL-1/mIgG2b.

I arbete III visar vi att alla *O*-bundna prekursorstrukturer ("core 1, core 2, core 3 och extended core 1") på PSGL-1/mIgG2b kunde bära GlcNAc4Gal strukturen, vilket förklaras av att a4 *N*-acetylglucosaminyltransferaset accepterade galaktos på alla prekursor kedjor. I preliminära försök kunde vi inte påvisa en negativ effekt på växten av *H. pylori* med PSGL-1/mIgG2b bärande GlcNAc4Gal, vilket visats i tidigare studier. Fler studier krävs för att reda ut denna skillnad i resultat.

I arbete IV undersökte vi repertoaren av socker på proteinet lubricin dels tillverkat i cellfabriker genom rekombinant teknologi, dels isolerat från synovialvävnad från friska leder och leder drabbade av artros. Vi visade att ett kroppseget lektin, galectin-3, band *O*-bundet socker med en specifik prekursor kedja ("core 2), och att både mängden

galectin-3 såväl som dess bindning till lubricin var sänkt i artrosdrabbad led, vilket skulle kunna bidra till att förstå den molekylära mekanismen bakom broskskadan vid artros.

Genom att använda CHO celler som cellfabriker och rekombinant teknologi har ett antal av *H. pylori* sockerreceptorer återskapats i flera kopior på det mucinliknande proteinet, PSGL-1/mIgG2b, och dess bindning till *H. pylori* studerats i detalj. En förhoppning är att rekombinant PSGL-1/mIgG2b med skräddarsydd sockerbeklädnad i framtiden skall kunna användas terapeutiskt genom att hämma mikrobiell bindning och därmed förhindra infektion.

LIST OF PAPERS

This thesis is based on the following studies, referred to in the text by their Roman numerals.

- I. **Mthembu YH**, Jin C, Padra M, Liu J, Olofsson -Edlund J, Ma H, Padra J, Oscarson S, Borén T, Karlsson NG, Lindén SK and Holgersson J. Recombinant mucin-type proteins carrying LacdiNAc on different O-glycan core chains fail to support *H. pylori* binding. *Molecular Omics* (doi: 10.1039/C9MO00175A).

- II. **Mthembu YH**, Olofsson-Edlund J, Jin C, Cherian R, Liu J, Karlsson NG, Borén T, and Holgersson J. Identification of the O-glycomes of mucin-type receptors for BabA- and SabA-mediated *Helicobacter pylori* adhesion. Submitted.

- III. **Mthembu YH**, Jin C, Padra M, Liu J, Karlsson NG, Linden SK and Holgersson J. O-glycan core chain specificity of A4GNT and the effect of GlcNAc α 4Gal determinants on *Helicobacter pylori* growth. Manuscript.

- IV. Flowers SA, Thomsson KA, Ali L, Huang S, **Mthembu YH**, Gallini R, Holgersson J, Schmidt TA, Rolfson O, Björkman LI, Sundqvist M, Karlsson-Bengtsson A, Jay G, Kamali-Moghaddam M, Eisler T, Krawetz R and Karlsson NG. Core-2 O-glycans are required for galectin-3 interaction with the osteoarthritis related protein lubricin. Submitted.

Reprints were made with the permission from publisher

CONTENT

1.	INTRODUCTION	1
1.1.	Glycans	1
1.2.	Glycan nomenclature	2
1.3.	Protein glycosylation	3
1.3.1.	<i>N-glycosylation</i>	4
1.3.2.	<i>O-glycosylation</i>	5
	<i>O-GalNAc glycosylation (Mucin type)</i>	6
	<i>O-GalNAc biosynthesis</i>	6
	<i>Functional consequences of O-GalNAc glycosylation</i>	9
1.3.3.	<i>Additional forms of O-glycosylation</i>	10
	<i>O-GlcNAc glycosylation</i>	10
	<i>O-Man glycosylation</i>	10
	<i>O-Glc and O-Fuc glycosylation</i>	11
1.3.4.	<i>Glycosylphosphatidylinositol (GPI) anchors</i>	11
1.4.	Protein-carbohydrate interactions and methods of analysis	11
1.5.	Glycan-based pharmaceuticals	12
1.6.	Glyco-engineering technology	13
1.7.	Glyco-engineering of recombinant PSGL-1/mIgG2b proteins in CHO-K1 cells	14
	<i>CHO-K1 mammalian cell line</i>	15
1.8.	Glycan-pathogen interactions	16
1.9.	<i>Helicobacter pylori</i>	17
	<i>H. pylori pathogenesis</i>	17
	<i>H. pylori adhesins</i>	18

1.10.	Gastric glycans and <i>H. pylori</i> interactions.....	19
1.10.1.	<i>ABH and Lewis antigens</i>	20
1.10.2.	<i>The LacdiNAc determinant</i>	23
1.10.3.	<i>The GlcNAcα4Gal-terminal</i>	23
1.11.	Glycans in gastric cancer.....	24
	<i>Glycosylation alterations in gastric cancer</i>	25
1.12.	Glycosylation in other disease.....	26
1.12.1.	<i>Rheumatic arthritis (RA) & Osteoarthritis (OA)</i>	26
1.12.2.	<i>Galectin-3 as the major LDN-binding lectin</i>	27
2.	AIMS OF THE THESIS	29
3.	METHODOLOGICAL CONSIDERATION	31
3.1.	Cell culture and transfection of CHO cells.....	31
3.2.	Purification of secreted recombinant PSGL-1/mIgG2b (Paper I).....	32
3.3.	Quantification of PSGL-1/mIgG2b by anti-mouse IgG Fc enzyme-linked immunosorbent assay (ELISA).....	33
3.4.	Characterization of PSGL-1/mIgG2b and its carbohydrate determinant using SDS-PAGE and Western blot analysis.....	33
3.5.	Chemical release of <i>O</i> -linked glycans from purified PSGL-1/mIgG2b prior to LC-MS analysis	35
3.6.	LC-MS/MS analysis.....	35
3.7.	Analysis of <i>H. pylori</i> adhesion.....	36
4.	RESULTS AND DISCUSSION	39
4.1.	Production of recombinant PSGL-1/mIgG2b with tailored glycosylation (Papers I-IV).....	39

4.2.	Characterization of recombinant PSGL-1/mIgG2b with tailored glycosylation (Papers I-IV).....	39
4.3.	Carbohydrate-dependent inhibition of protein-carbohydrate interactions.....	40
4.3.1.	<i>LDN-containing glycoconjugates do not bind H. pylori (Paper I)</i>	42
4.3.2.	<i>Competition inhibition of H. pylori by glycoconjugates carrying Le^b, sLe^x, LDN and GlcNAcα4Gal on various core chain O-glycans (Paper II)</i>	42
4.3.3.	<i>Assessment of H. pylori growth in the presence of GlcNAcα4Gal-carrying mucin-type fusion proteins or purified human gastric mucins with or without GlcNAcα4Gal-terminals</i>	43
4.3.4.	<i>Core 2 type oligosaccharides aid in the recognition of lubricin by galectin-3</i>	43
5.	CONCLUSIONS	46
6.	FUTURE PROJECTS	48
7.	ACKNOWLEDGEMENTS	50
8.	REFERENCES	55

ABBREVIATIONS

B3GALT1	β 1,3galactosyltransferase;
B3GNT3	β 1,3galactosyltransferase 3;
B3GNT6	β 1,3-N-acetylglucosaminyltransferase;
B4GALNT3	β 1,4-N-acetylgalactosaminyltransferase 3;
B6GNT1	β 1,6-N-acetylglucosaminyltransferase 1;
BabA	blood group antigen-binding adhesion
BSA	Bovine serum albumin
CsCl	Caesium chloride
CHAPS	3-((3-cholamidopropyl)-dimethyl-ammonio)-1-propane sulfonate
CHO-K1	Chinese hamster ovary cells K1;
CO ₂	Carbon Dioxide
CV	Column volumes
dH ₂ O	Distilled water
°C	Degrees Celsius
DEAE	Dextran diethylaminoethyl ether of dextran
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic acid
ECL	Enhanced chemiluminescence
ELISA	Enzyme-linked immune sorbent assay
FBS	Fetal bovine serum
FUT	Fucosyltransferase
g	Gram
Gal	Galactose
GalNAc	N-Acetylgalactosamine
GalT	Galactosyltransferase
GuHCl	Guanidinium chloride
GM	Growth media
GlcNAc	N-acetylglucosamine
HCl	Hydrochloric acid
HexNAc	N-acetylhexosamine
HPLC	High pressure liquid chromatography

<i>H. pylori</i>	<i>Helicobacter pylori</i>
HRPO	Horse radish peroxidase
HSA	Human serum albumin
kDa	Kilo-Dalton
LabA	LacdiNAc binding adhesion
LDN	LacdiNAc (N,N'-diacetyllactosdiamine/GalNAc β 1,4GlcNAc);
LC	Liquid chromatography
Le ^b	Lewis b
μ l	Microliter
Min	Minute
ml	Milliliter
MUC	Mucin
MS	mass spectrometry
M _w	Molecular weight
Neg	Negative
Neu5Ac	N-acetylneuraminic acid
Neu5Gc	N-glycolylneuraminic acid
PAS	Periodic acid schiff
PBS	Phosphate Buffered Saline
PMSF	Phenylmethylsulfonylfluoride
Pos	Positive
ProCHO	Protein-free CHO Media;
PSGL-1/mIgG2b	P-selectin glycoprotein ligand-1/mouse IgG2b ;
RT	Room temperature
SabA	sialic acid-binding adhesion;
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis;
SDS	Sodium Dodecyl Sulphate
SDS-PAGE	Sodium Dodecyl Sulphate-Poly acrylamide-gel electrophoresis
SLe ^x	sialylated Lewis x;
TBST	Tris buffered saline-tween
Tween 20	Polyoxyethylene sorbitan monolaurate
Type 1	Gal β 1,3GlcNAc
Type 2	Gal β 1,4GlcNAc

V

Voltage

1. INTRODUCTION

Glycobiology is the study of saccharide (sugar chain or glycans) structure, biosynthesis, biology and evolution. With its roots in classical chemistry and biochemistry, glycobiology has been described as an extension of molecular biology, which emerged as a result of the development of many new technologies for exploring the structures and functions of these glycans and their role in biological systems (Varki A 2009).

1.1. Glycans

Glycans are oligo- or polysaccharide molecules (commonly known as carbohydrates), comprised of carbon, hydrogen and oxygen, and are important in many cell functions both physiologically and pathologically. Glycans append a wide variety of biological molecules and often contribute to physical and structural integrity, extracellular matrix formation, information exchange between cells and pathogen uptake (Stroh and Stehle 2014). For instance, cell surface glycans have been reported to facilitate attachment and entry of microbes, including viruses (Wasik et al. 2016) and bacteria (Karlsson 2001), into their target host cells. Sialic acid was one of the first glycans that has been known as a virus receptor.

Cell surface glycans have also been described to be important as adhesion receptors in cell-cell and cell-matrix interactions (Holgersson et al. 2005) such as in leucocyte extravasation during inflammation and lymphocyte recirculation. One such example of glycans in leucocyte extravasation is sialyl 6-sulfo Lewis x (terminal structure NeuNAc α 2,3Gal β 1,4[Fuca α 3]GlcNAc β [6-SO $_3^-$]1-R), a putative L-selectin ligand expressed on high endothelial venules (HEV) in human lymph nodes (Mitsuoka et al. 1998), which has been shown to play important roles in various aspects of lymphocyte homing (Kannagi 2002). These glycans are determinants of self/non-self and thus anti-carbohydrate antibodies can initiate a graft rejection following transplantation between individuals of different ABO blood groups, or between species (Holgersson et al. 2005).

Glycans can be secreted as free saccharides or be attached to a variety of biological molecules such as nucleic acids, proteins and lipids, to form glycoconjugates. Glycosylation is the enzymatic process by which the carbohydrate chain is established










and its function defined (Springer and Gagneux 2013). Such examples of macromolecule glycoconjugates include glycolipids, proteoglycans and glycoproteins. Briefly, glycolipids, also known as sphingolipids, have glucose or galactose attached to the terminal primary hydroxyl group of the lipid moiety ceramide composed of a long chain base (sphingosine) and a fatty acid. Proteoglycans have one or more glycosaminoglycan (GAG) chains which are linked to the hydroxyl group of a serine residue through a tetrasaccharide linker.

In this thesis, the glycans of glycoproteins will be discussed in detail. A glycoprotein is comprised of a protein carrying one or more glycans covalently attached to its polypeptide backbone, usually via *N*- or *O*-linkages. Typical glycoproteins include glycoproteins that have varying contents of glycans which are in the form of *N*- and/or *O*-linked oligosaccharide chains that can be linear or branched. These oligosaccharides attach covalently to amino acids via glycosidic linkages formed by glycosyltransferases in a process known as glycosylation.

1.2. Glycan nomenclature

Monosaccharides are the basic building blocks and the simplest structures of all oligo- and polysaccharides. Mammalian oligosaccharides are made up of a combination of hexoses [glucose (Glc), galactose (Gal), mannose (Man)], *N*-acetyl hexosamines [*N*-acetylgalactosamine (GlcNAc), *N*-acetylglucosamine (GalNAc)], fucose (Fuc), xylose (Xyl), and sialic acids [*N*-acetylneuraminic acids (Neu5Ac), *N*-glycolylneuraminic acid (Neu5Gc)]. Glycosyltransferases (GTs) are responsible for the formation of glycosidic linkages that occur between the anomeric carbon of one monosaccharide and the hydroxyl group of another. In glycobiology, multiple notation schemes have been defined to facilitate the identification and comparison of glycans. Two main annotation schemes are used, the Consortium for Functional Glycomics (CFG) nomenclature (Varki et al. 2009) and the Oxford system (Harvey et al. 2009). In this thesis, the CFG nomenclature (Table 1) is used to represent the glycan structures.

Table 1. Some common monosaccharides in mammals and their CFG symbols.

Monosaccharide	Abbreviation	CFG Symbol
Glucose	Glc	
Galactose	Gal	
Mannose	Man	
<i>N</i> -acetylglucosamine	GlcNAc	
<i>N</i> -acetylgalactosamine	GalNAc	
Fucose	Fuc	
Xylose	Xyl	
<i>N</i> -acetylneuraminic acid	NeuNAc (NANA)	
<i>N</i> -glycolylneuraminic acid	NeuNGc (NGNA)	

1.3. Protein glycosylation

Glycosylation is a post-translational modification involving the addition of sugar chains to membrane-anchored and secreted proteins as well as, in some cases, intracellular proteins. The glycosylation process plays an important role in regulating protein folding, targeting proteins to specific subcellular compartments, their interaction with ligands and other proteins as well as their overall function. Glycosylation is one of the most

common ways in which glycan side chains can attach to a polypeptide. Proteins can be both *N*- and *O*-glycosylated depending on the linkage of the oligosaccharide to the amino acid side chain of the protein. Other than the *N*- and *O*-linked glycosylation, the glycosylphosphatidylinositol (GPI) anchors is the third type of posttranslational modifications of proteins that also involve carbohydrates (Steen et al. 1998).

1.3.1. *N*-glycosylation

One distinction of *N*-glycosylation is recognized by the transfer of a common oligosaccharide (Glc3Man9GlcNAc2) sequence, which is pre-assembled on a lipid carrier, dolichol pyrophosphate, prior to its transfer to the nitrogen of asparagine (Asn) residues within polypeptides in the endoplasmic reticulum (Kornfeld and Kornfeld 1985). The Asn residues acting as acceptors for *N*-linked glycans are found in the sequence Asn-X-Ser/Thr (*N*-X-S/T), where X may be any amino acid except for proline (Kornfeld and Kornfeld 1985). *N*-glycan biosynthesis is initiated in the endoplasmic reticulum (ER) with additional monosaccharides added individually from nucleotide sugar donors in the Golgi complex. These *N*-glycans share a common core consisting of two *N*-acetylglucosamine (GlcNAc) and three mannose (Man) residues, and are classified into three types: high mannose (oligomannose), hybrid and complex type *N*-glycans (Varki 1993) as illustrated in Fig. 1.

O-GalNAc glycosylation (Mucin type)

O-linked GalNAc glycosylation is the most common *O*-linked modification initiated in the cis to trans Golgi apparatus. *O*-linked GalNAc is initiated by the post translational addition of GalNAc to the hydroxyl group (oxygen) of serine and threonine (Röttger et al. 1998) and is catalyzed by UDP-GalNAc-polypeptide GalNAc transferase (ppGalNAcTs).

O-GalNAc biosynthesis

The binding of GalNAc to Ser or Thr gives rise to the simplest known *O*-glycan structure and takes place in the Golgi complex. The GalNAc residue serves as an attachment site for further elongation and the generation of different *O*-glycan core structures (Beum et al. 2003). The innermost two or three sugars of the *O*-glycan chain define the core structures, which are used to classify the *O*-glycans (Fukuda 2002). At least eight different core chain types, of which cores 1-4 are more common than the rare cores 5-8, have been identified in mammalian glycoproteins (Table 2). All these core structures are based on the innermost β -GalNAc residue, which is further substituted at the C3, C6 or both positions with the monosaccharides β -Gal at C3, β -GlcNAc at C6 and/or C6, and α GalNAc at C3 or C4.

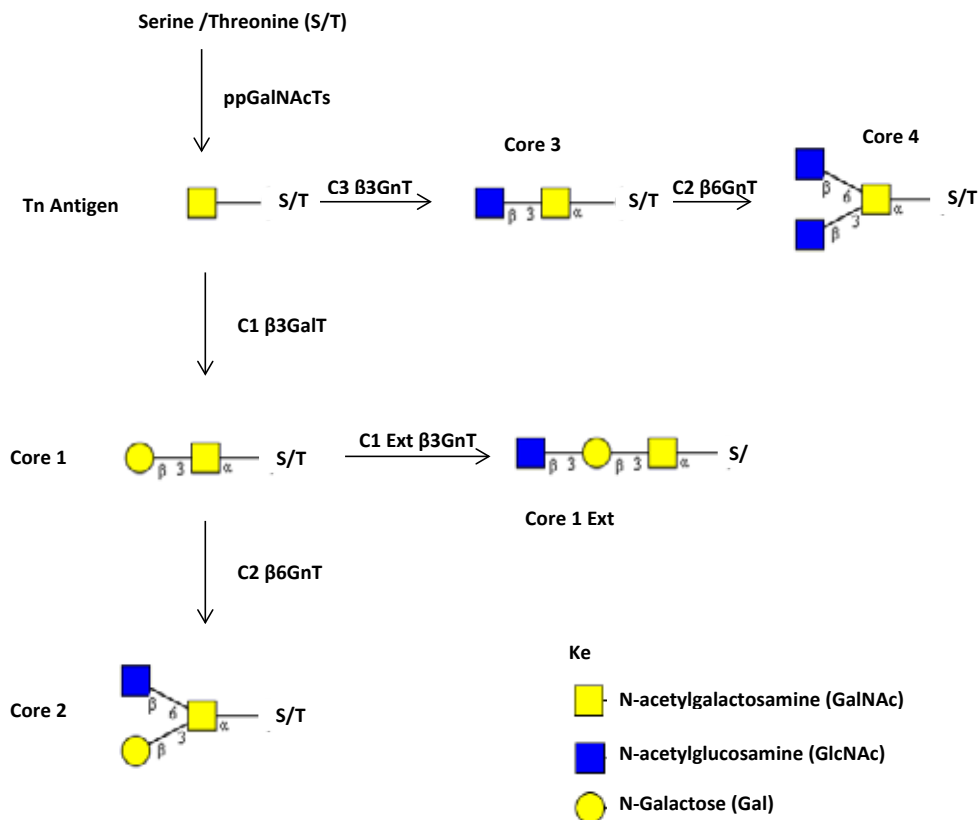


Figure 2: Biosynthesis of core 1-4 O-GalNAc glycans

As illustrated in Fig. 2, the O-glycan core 1 structure the most common and is catalyzed by the core 1 β 1,3galactosyltransferase (C1 β 3GalT1 or *B3GALT1*), which adds galactose in a β 1,3-linkage to the GalNAc residue (Ju et al. 2002). β 1,6-N-acetylglucosaminyltransferase (C2 β 6GnTs or *GCNT1*) produces the core 2 structure by the addition of an N-acetylglucosamine (GlcNAc) in a β 1,6-linkage to the GalNAc of the core 1 structure (Schwientek et al. 2000; Yeh et al. 1999). Other β 1,6-N-acetylglucosaminyltransferase (C2 β 6GnT1-3 or *GCNT13*) have been reported with C2

β 6GnT1 and C2 β 6GnT3 responsible for the synthesis of core 2 both *in vitro* and *in vivo*, while the C2 β 6GnT2 can also make core 4 (Schwientek 2000).

The core 3 structure produced from the addition of a GlcNAc in a β 1,3-linkage to the innermost GalNAc as directed by β 1,3-*N*-acetylglucosaminyltransferase 6 (C3 β 3GnT6 or *B3GNT6*) and competes with the C1 β 3GalT1 glycosyltransferase (Iwai et al. 2002).

The addition of a GlcNAc residue in a β 1,6-linkage to GalNAc of a core 3 structure by the GCNT2 enzyme results in the generation of a core 4 structure (Yeh et al. 1999). The core 3 and core 4 *O*-glycans are expressed in a more tissue-specific manner than the more abundant core 1 and core 2 structures (Hanisch 2001).

Oligosaccharide side chain can be elongated by repetitive backbone structures of different lengths. These backbone structures consist of β -linked GlcNAc and Gal forming three types, namely; Type 1 Gal β 3GlcNAc, type 2 Gal β 4GlcNAc and 3,6Gal-branched Gal β 4GlcNAc β 6(Gal β 4GlcNAc β 3) Gal structures.

The terminal determinants found at the end of core 1 to 4 glycans (Table 2) typically found in mammalian glycoproteins include sialic acid (linked α 2,3 and α 2,6), fucose (α 1,2 α 1,3 and α 1,4), *N*-acetylgalactosamine (GalNAc) (linked α 1,3 α 1,6 and β 1,4), *N*-acetylglucosamine (linked α 1,4), galactose (Gal) (linked α 1,3) and sulphate residues. Terminal sialic acid and sulphates that are attached to GalNAc impart negative charges to the mucins, whereas fucose imparts hydrophobicity (Forstner and Forstner, 1994). Other modifications include acetylation and methylation.

Table 2. Carbohydrate structures of mucin-type O-glycans

Peripheral Determinants	Backbones	Cores
Blood group H Fuca2Gal-	Type 1 Galβ3GlcNAc-	Tn antigen GalNAcαSer/Thr
Blood group A GalNAcα3(Fuca2)Gal-	Type 2 Galβ4GlcNAc-	Core 1 (T) Galβ3GalNAcαSer/Thr
Blood group B Galα3(Fuca2)Gal-	(3-6 Gal-branched) Galβ4GlcNAcβ6(Galβ4GlcNAcβ3)Gal-	Core 2 GlcNAcβ6(Galβ3)GalNAcαSer/Thr
Linear blood group B Galα3Gal-	Type 3 Galβ3GalNAc-	Core 3 GlcNAcβ3GalNAcαSer/Thr
Blood group i Galβ4GlcNAcβ3Gal-	Type 4 Galβ4GalNAc-	Core 4 GlcNAcβ6(GlcNAcβ3)GalNAcαSer/Thr
Blood group I Galβ4GlcNAcβ6(Galβ4GlcNAcβ3)Gal-		Core 5 GalNAcα3GalNAcαSer/Thr
Blood group Sd(a), Cad GalNAcβ4(Siaα2,3)Gal-		Core 6 GlcNAcβ6GalNAcαSer/Thr
Blood group Le^a Galβ3(Fuca4)GlcNAc-		Core 7 GalNAcα6GalNAcαSer/Thr
Blood group Le^b Fuca2Galβ1-3(Fuca4)GlcNAc-		Core 8 Galα3GalNAcαSer/Thr
Blood group Le^x Galβ4(Fuca3)GlcNAc-		Core 1 Extended GlcNAcβ3Galβ3GalNAcαSer/Thr
Blood group Sle^x Siaα2,3Galβ4(Fuca3)GlcNAc-		
Blood group Le^y Fuca2Galβ4(Fuca3)GlcNAc-		

Functional consequences of O-GalNAc glycosylation

The functional consequences of protein glycosylation are wide. The addition of glycans on a peptide/protein can modulate protein structure and stability. The mucin-type O-GalNAc glycosylation and subsequent elongation of the oligosaccharide chain influence protein conformation and leads to the formation of a 'bottle brush' structure caused by

frequent *O*-GalNAc substitution (sometimes every third to fourth amino acid in a mucin domain can be substituted). *O*-glycosylation also provides protection against proteolytic degradation and thermal disruptions (Wang et al. 1996). Along with *N*-glycans, the *O*-glycans play a role on a variety of recognition processes, such as cell growth/proliferation, signaling pathways, immunological recognition and glycoprotein trafficking/clearance (Varki 2017). Glycans also enable the glycoprotein to target specific tissue or cell types via glycan-binding receptors. Glycoproteins are abundant on the surface of both eukaryotic and prokaryotic cells and are important mediators of host-microbe interactions, especially in the gut (Keys and Aebi 2017).

1.3.3. Additional forms of *O*-glycosylation

***O*-GlcNAc glycosylation**

O-GlcNAc glycosylation involves linking of GlcNAc residues to serine and threonine residues in non-secretory cytoplasmic and nucleic proteins (Hart 1997). *O*-GlcNAc glycosylation has been shown to compete with phosphorylation (Hart et al. 2011) and therefore it regulates processes in the cell such as transcription, epigenetics, and cell signaling dynamics (Yang and Qian 2017).

***O*-Man glycosylation**

O-mannosylation is accomplished when a mannose from a dolichol-P-mannose donor molecule is transferred onto a serine or threonine residue of a protein. The *O*-mannosylation process is initiated in the endoplasmic reticulum, while sugar chain elongation occurs in the Golgi apparatus (Lommel and Strahl 2009). A well-characterized highly mannosylated protein is Dystroglycan (α -DG), which is a basement membrane receptor involved in a variety of physiological processes that maintain skeletal muscle membrane integrity and that plays an important role in central nervous system development (Inamori et al. 2012).

O-Glc and O-Fuc glycosylation

O-Glc and *O-Fuc* protein glycosylation involving the linkage of fucose or glucose to serine or threonine residues has been reported on epithelial growth factor (EGF) domains of proteins involved in the regulation of blood clotting (Spiro 2002). *O*-glucosylation and *O*-fucosylation are necessary for the proper folding of EGF domains in the Notch protein, a large single-pass transmembrane protein which control a variety of developmental processes (Takeuchi et al. 2012).

1.3.4. Glycosylphosphatidylinositol (GPI) anchors

This is a hybrid glycosylation in which a protein is attached to a lipid anchor via a glycan chain (Paulick and Bertozzi 2008). The GPI anchor contains a phosphoethanolamine linker, a glycan core, $\text{Man}\alpha 2\text{Man}\alpha 6\text{Man}\alpha 4\text{GlcNAc}\alpha 6\text{myo-inositol}$, and a phospholipid tail. GPI-anchored proteins are structurally and functionally diverse and play vital roles in numerous biological processes (Rajendran and Simons 2005). Some examples of functions include the involvement in lipid raft partitioning, signal transduction, targeting to the apical membrane, toxin binding and prion disease pathogenesis (Kinoshita 2016).

1.4. Protein-carbohydrate interactions and methods of analysis

Glycan structures, in the form of glycoconjugates, are presented on cell surfaces of cells to provide a dense structural code which is deciphered by glycan binding proteins other cells as well as from several viral, bacterial and fungal pathogens as illustrated in Fig. 3 (Imberty and Varrot 2008). For a protein-carbohydrate interaction to take place, it is essential that the specific glycan determinant is accessible. Proteins that can bind carbohydrates include antibodies, carbohydrate-specific enzymes, transport/sensor proteins for free sugars, and lectins. Different techniques have been used to monitor and measure protein-carbohydrate interactions. Some examples of these techniques include the use of X-Ray crystallography and nuclear magnetic resonance (NMR) spectroscopy

while the transferred nuclear Overhauser effect (NOE) can be used to obtain information on the bound-state conformation of the carbohydrate ligand (Pohl 2010). Calorimetric evaluations, mass spectrometry and surface plasmon resonance studies are also part of the arsenal used to investigate protein-carbohydrate interactions (del Carmen Fernández-Alonso et al. 2012). Glycan array is a technology which provides a powerful, high-throughput approach to examining the binding specificity of carbohydrate-binding proteins. The array is composed of a repertoire of structurally distinct oligosaccharides immobilized on a solid support (beads, silica or ELISA plates) in a spatially-defined arrangement (Oyelaran and Gildersleeve 2009).

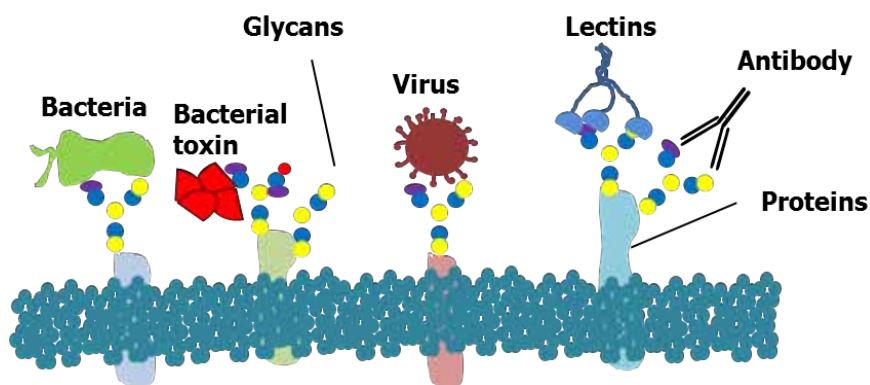


Figure 3: Schematic drawing illustrating protein-carbohydrate interaction at the cell surface mediating cell-microbe (bacterial, bacterial toxin and viral) attachment as well as lectin and antibody binding.

1.5. Glycan-based pharmaceuticals

The market for glycoprotein therapeutics is one of the fastest growing markets in the pharmaceutical industry. Some of these therapeutic glycoproteins such as clotting factors, hormones, cytokines, enzymes, enzyme inhibitors, IgG Fc fusion proteins and

monoclonal antibodies play a major role in the treatment of many diseases of great importance to humanity. They also represent an integral part of treatment protocols for various cancer types, autoimmune diseases, and as replacement therapies for diseases characterized by various types of deficiencies (Ghaderi et al. 2012). Some of the therapeutic proteins that are glycoproteins include erythropoietin (EPO; for the treatment of severe anemia) and Infliximab (for the treatment of autoimmune diseases). Tamiflu (Oseltamivir) is a sialidase inhibitor used for the treatment of influenza virus infection (Jefferson et al. 2014).

1.6. Glyco-engineering technology

Most therapeutic glycoproteins are produced in various host cells by recombinant DNA technology. There is a plethora of molecular techniques existing that allow the manipulation of the genetic codes at various levels. Such manipulations allow the programmable control of gene expression.

There are several nuclease-based genome editing tools that have been developed. To date, four main platforms for precision genome editing exist, including Zn-finger nucleases (ZFN), transcription activator-like effector nucleases (TALENs), meganucleases (homing endonucleases) and the RNA-guided clustered regularly interspaced short palindromic repeat/CRISPR-associated (CRISPR/Cas) nuclease system (Steentoft et al. 2014). Cys₂-His₂ Zn-finger proteins are DNA-binding proteins that contain sequences which binds a Zn(II) ion to form the structural domain termed a zinc finger, nucleases are designed to act as transcription activators, transcription inhibitors and restriction enzymes, thereby introducing genomic alterations such as point mutations, deletions, insertions, inversions, duplications and translocations (Kim et al. 1996). TALEs are engineered nucleases which comprise a non-specific nuclease domain fused to a customizable DNA-binding domain composed of highly conserved repeats derived from transcription activator-like effectors (TALEs).

These targetable nucleases are used to induce double-strand breaks (DSBs) at specific DNA sites which can then be exploited to create sequence alterations at the cleavage site

(Joung and Sander 2013; Mussolino and Cathomen 2012). Meganucleases are engineered versions of naturally occurring restriction enzymes that typically have extended DNA recognition sequences (Silva et al. 2011) CRISPR/Cas are RNA-guided nucleases (RGN) which use simple base pairing rules between an engineered RNA and the target DNA site, and dependent on RNA as the moiety that targets the nuclease to a desired DNA sequence (Sander and Joung 2014). These four tools are based on the capacity of modified and optimized nucleases to elicit sequence-specific in vivo double-strand DNA breaks, and resulting in various gene editing possibilities including gene disruption, deletion, insertion, correction, mutation and tagging (Steentoft et al. 2014).

1.7. Glyco-engineering of recombinant PSGL-1/mIgG2b proteins in CHO-K1 cells

This thesis is focused on the use of a recombinant mucin-type fusion protein, P-selectin glycoprotein ligand-1/mouse IgG2b (PSGL-1/mIgG2b) that is engineered to carry bioactive carbohydrate determinants in a multivalent fashion (Fig. 4). PSGL-1/mIgG2b is comprised of the extracellular part of P-selectin glycoprotein ligand-1, a ligand for the adhesion molecule P-selectin, and the Fc portion of mouse IgG2b and is expressed in CHO-K1 cells engineered to express also glycosyltransferases involved in the biosynthesis of the carbohydrate determinants whose interaction with various carbohydrate-binding protein is to be studied. In this thesis PSGL-1/mIgG2b has been used as a tool to study the interactions between *Helicobacter pylori* and its carbohydrate receptors. Previous research from our laboratory has used this technique to study a whole range of proteins with their carbohydrate counter-receptors, including antibodies and lectins (Lindberg et al. 2013; Liu et al. 2005), viral and bacterial proteins (Lofling et al. 2008) as well as bacterial toxins (Cherian et al. 2016; Maria Cherian et al. 2014).

PSGL-1 is a dimer (linked together by disulfide bonds) of type 1 transmembrane topology which is comprised of two identical subunits, each containing 402 amino acids as well as 53 sites for *O*-linked glycosylation and 3 sites for *N*-linked glycosylation, allowing for multivalent presentation of its bound carbohydrate ligands (Moore et al. 1995). The Fc portion of mouse IgG2b allow for its purification using affinity

chromatography. The glycosylation phenotype of PSGL-1/mIgG2b depends on the host cell line which has been used for its production in combination with additionally expressed glycosyltransferases.

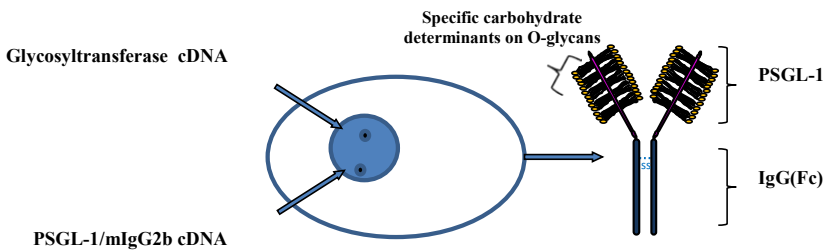


Figure 4: The production of recombinant mucin-type fusion protein PSGL-1/mIgG2b with specific carbohydrate structures as directed by the cell machinery and glycosyltransferase cDNA used.

CHO-K1 mammalian cell line

Mammalian cells are routinely used for the production of biopharmaceutical proteins an industrial scale. They support human-like glycosylation and most FDA-approved drugs of this category are produced in Chinese Hamster Ovary (CHO) cells (Ghaderi et al. 2012).

CHO-K1 cells carry a well-defined repertoire of glycans which is similar to that of human cells (Liu et al. 2005) and they are well-characterized from a regulatory point of view as they have been widely used for large-scale production of bio therapeutics (Barnes et al. 2003; Bäckström et al. 2003; Geisse et al. 1996). CHO cells also present low risk for the transmission of the major human viruses, high protein yield, and robustness towards pH, temperature, oxygen level, and pressure variations (Wurm 2004). They can carry out

post-translational modifications and decorate therapeutic glycoproteins with human-like *N*-glycans, leading to more compatible, stable, and bioactive therapeutic glycoproteins (Jacobs and Callewaert 2009). Even though CHO cell-based expression systems are currently the preferred host for the production of therapeutic glycoproteins, there are several disadvantages listed: a relatively high cost of goods; the potential for propagating infectious agents, such as viruses and prions; a long development time from gene to production cell line; and the inability to adequately control *N*-glycosylation (Sethuraman and Stadheim 2006). Regarding the *O*-glycans, sialylated core 1 *O*-glycans dominate (Liu et al. 2015) these CHO cells.

1.8. Glycan-pathogen interactions

Protein-carbohydrate interactions are important in pathogen adhesion and infections. There are a number of different pathogens which through their lectins can recognize specific glycan determinants presented on the epithelial cells of the airways, gastrointestinal and urinary tracts, supporting pathogen colonization and infection of the host (Karlsson 1998). Different forms of these lectins exist to facilitate entry and promote infection; they can be expressed directly on the bacteria or secreted as toxins. For example, *Escherichia coli* (*E. coli*) binds to host mannose (Mukhopadhyay et al. 2009), with other strains of *E. coli* showing specificities towards other host cell surface carbohydrate moieties such as galactose (Gal α 4Gal) (Khan et al. 2000) and NeuAc α 2,3Gal β 3GalNAc (Buts et al. 2003). Among other pathogens which bind carbohydrates, the genital pathogen *Neisseria gonorrhoeae* specifically binds to *N*-acetylglucosamine (Gal β 4GlcNAc, LacNAc) (Barthelson et al. 1998).

Examples of toxins from pathogens which bind carbohydrates include toxin A (*Clostridium difficile*) which binds to Gal α 3Gal β 4GlcNAc (Cherian et al. 2016) is responsible for antibiotic-associated diarrhoea. Shiga-like toxin (*E. coli*) that binds to globotriaosylceramide (Gal α 4Gal β 4Glc) causes diarrhoea and haemorrhagic colitis (Maria Cherian et al. 2014). A whole host of viruses which depend on carbohydrates (in addition to proteins and lipids) as a means for attachment to the host and infection have been studied. Some of these include influenza (Gaunitz et al. 2014), rotavirus (Yu and

Blanchard 2014), human immune-deficiency virus (HIV-1) (Shen et al. 2014), norovirus (Marionneau et al. 2002) and Herpes viruses (Mårdberg et al. 2004). This thesis focused on the glycans on gastric mucins, which are diverse and constitutes a vast number of different structures that are predominantly non-charged in healthy tissue (Kenny et al., 2012). These glycans can act as a functional binding site for *H. pylori*.

1.9. *Helicobacter pylori*

H. pylori is a major etiological agent of chronic gastritis and peptic ulcer disease and has been classified as a carcinogen by the WHO. Infection with *H. pylori* is the most important acquired risk factor for gastric cancer development (Zheng et al. 2009) with 1-3% of infected individuals developing the disease (Dunne et al. 2014). *H. pylori* is a gram negative, spiral shaped microaerophilic bacterium, measuring approximately 3-5µm in length. *H. pylori* was first isolated and cultured from gastric mucosa by Marshall and Warren in 1983 (Kenny et al. 2012; Marshall and Warren 1984). It is transmitted directly from person to person through feces (fecal-oral spread) or gastric content (gastric-oral spread). Almost 50% of the world's population is infected with *H. pylori*. However, only a subset of individuals will progress to have gastric cancer due to several factors at play. This includes the characteristics of the host as well as the genetic variation of the bacterium (Sokic-Milutinovic et al. 2015).

***H. pylori* pathogenesis**

H. pylori reside in the gastric mucus layer (Hazell SL 1986). *H. pylori* is able to adapt to various environmental dynamics in the stomach by producing adhesive molecules. The environment for *H. pylori* in the stomach can vary between hosts, gastric niches and as a result to changes due to infection and diseases (Linden et al. 2002). It is able to survive in the acidic environment of the stomach due to its ability to produce ammonia and carbon dioxide by releasing urease which hydrolyzes endogenous urea (Sachs et al. 2003). The bacteria moves toward host gastric epithelium cells using flagella-mediated motility, before colonizing the mucosal layer of the gastric epithelium through its

adhesins' binding to host cellular receptors (Kao et al. 2016b). The bacteria release effector proteins/toxins including cytotoxin-associated gene A (CagA) and vacuolating cytotoxin A (VacA), causing host tissue damage. When CagA gets injected into the gastric cells of the host upon binding, it becomes phosphorylated and cause the activation of a cascade of events which lead to a growth factor-like cellular response, production of cytokines and induction of junctional and polarity defects in the epithelial cells (Hatakeyama 2014). The bacteria also secrete VacA, which correlates with the expression of CagA, to elicit cellular vacuolation (Atherton et al. 1995; Stefano et al. 1996).

***H. pylori* adhesins**

H. pylori carry carbohydrate-binding adhesins which enable them to bind to the epithelial cells of the gastric mucosa (Kao et al. 2016a). The bacterium displays responsive adjustment with the ability to attach to non-acid Lewis antigens during health and to sialylated Lewis antigens post-infection. The most studied adhesins on *H. pylori* are the ABO blood group antigen-binding adhesin (BabA) and the sialic acid-binding adhesin (SabA). The BabA receptor on *H. pylori* binds to terminally fucosylated lacto-series type 1 (Gal β 3GlcNAc) chains. *H. pylori* inhibitory studies by Boren et al., identified Fuca α 1,2 epitope on Le^b and H type-1 determinants as the BabA receptors, whereas Le^a, H type-2, Le^x and Le^y did not display any receptor activity (Boren et al. 1993). *H. pylori* infection and the associated chronic inflammation induces expression of sLe^x in the gastric mucosa which is associated with gastric dysplasia and cancer. The *H. pylori* SabA mediates binding to sLe^x and sLe^a (Mahdavi et al. 2002).

LabA, is a third carbohydrate-binding adhesin which was described and identified to bind the LacdiNAc (N,N'-diacetyllactosdiamine/GalNAc β 4GlcNAc, LDN) determinant (Rossez et al. 2014). *H. pylori* has other known adhesins which allow it to adapt in its host such as neutrophil-activating proteins (NAP) (Teneberg et al. 1997), heat shock protein 60 (Hsp60) (Yamaguchi et al. 1997), adherence associated proteins (AlpA and AlpB) (Odenbreit et al. 1999), and *H. pylori* outer membrane protein (HopZ) (Peck et al. 1999).

NAP belongs to the DNA-protecting proteins (Dps) family, and it is able to stimulate high production of oxygen radicals from neutrophils as well as the induction of expression and release of macrophage inflammatory protein, leading to damage of local tissues during *H. pylori* infection. NAP also protect *H. pylori* DNA from being damaged from attack by free radicals (Kao et al. 2016b). HopZ is a member of the OMP family of *H. pylori*.

1.10. Gastric glycans and *H. pylori* interactions

The gastric mucosa is covered by two separable yet distinct mucus gel layers, with the shear-resistant mucus gel being the stronger and considered to protect the mucosa from acid, digestive enzymes and other damaging material, whereas the other loosely adherent mucus layer is continuously renewed after removal (Atuma et al. 2001). Mucin glycoproteins, which can be secreted into the mucus or bound to the cell, are responsible for the visco-elastic properties of mucus and are composed of protein domains that are highly glycosylated (Sheehan et al. 2004). These mucin glyco-proteins can be *N*- and/or *O*-glycosylated depending on how the protein domain is modified by glycans.

Several mucins have been reported to be present in the stomach. Two polymeric mucins are more prominent in gastric mucus with MUC5AC being found on the surface of gastric pits (foveolae) and MUC6 in the epithelial glands (neck) (Lindén et al. 2004; Skoog et al. 2012). MUC6 is produced from gastric fundic and antral glands (mucous neck and antral gland cells, respectively) as well as duodenal Brunner's glands and plays a role in the host defense against infection with *H. pylori* such that reduced expression of MUC6 is associated with increased gastric mucosal susceptibility to infection with *H. pylori* (Hoffmann 2015). An experimental model of guinea pig gastric epithelial cells *in vitro* and *in vivo* studies identified that an increased production of MUC5AC is in response to the presence of live *H. pylori* and also correlated to *H. pylori* adhesion to the gastric mucosa (Gonciarz et al. 2019).

Most of the *H. pylori* receptors on gastric cells are glycans. Some examples of these glycans include terminal β 3 linked galactose, ABH and Lewis antigens, glycosaminoglycans (GAGs) and sialic acid (Sia). The Le^b blood-group antigen mediated

the attachment of *H. pylori* to the gastric mucosa in a healthy stomach (Borén et al. 1993). In contrast, infection with *H. pylori* elicited gastric mucosal sialylation as part of the chronic inflammatory response suggesting that *H. pylori* bound to sialylated Lewis antigens in inflamed tissue (Mahdavi et al. 2002). The LDN (GalNAc β 4GlcNAc) determinant is found on secreted mucins present in the gastric mucosal lining, and is a unique terminal structure on the outer chains of *N*- and *O*-linked oligosaccharides (Ikehara et al. 2006; Kenny et al. 2012). The presence of LDN and Le^b blood group antigens have been shown to be expressed on MUC5AC and both have been suggested to aid in the colonization of the gastric mucosa by *H. pylori*. MUC6 has been shown to carry the specific carbohydrate moiety GlcNAc α 4Gal β 1-R at non-reducing terminals (Zhang et al. 2001). This determinant is biosynthesized by the α 1,4-N-acetylglucosaminyltransferase (α 4GnT) and is capable of exerting an inhibitory effect on *H. pylori* growth (Ishihara et al. 1996).

1.10.1. ABH and Lewis antigens

Karl Landsteiner discovered the ABO blood group antigenic system in 1901. He demonstrated that red blood cells from one individual could be agglutinated by serum factors from another individual and thus divided humans into different blood groups based on the finding (Landsteiner 1961). The distribution of ABO blood group determinants is cell- and tissue-specific; their expression on different carbohydrate precursor chain types (Fig. 5) has been reviewed (Clausen and Hakomori 1989; Holgersson et al. 1992). The biosynthesis of the ABH determinant is initiated by the α 1,2-fucosylation of the precursor chain. This process is catalyzed by the H (*FUT1*) or Se (*FUT2*) gene encoded α 1,2 fucosyltransferase (Pendur et al. 2001). *FUT1* (FucT1) is encoded for by the *H* gene and while the *FUT2* (FucT2) is encoded for by the *Se* gene. Both FucT1 and FucT2 produce the α 1,2-fucosylation on type 1 chains on both *N*- and *O*-linked structures (Holgersson and Lofling 2006) and on type 2 chains (Oriol et al. 1986). FucT1 is expressed in red blood cell precursors and prefers type 2 to type 1, while FucT2 prefers type 1 over type 2 and is mainly expressed on the epithelial lining of the gastrointestinal, respiratory and urinary tracts (Lowe 1993; Mollicone et al. 1985; Ravn and Dabelsteen 2000). The *FUT3* (Lewis) enzyme (FucT3), encoded by the *Le* gene

makes Lewis antigens by adding a fucose residue to the subterminal *N*-acetylglucosamine (Mollicone et al. 1985). The Lewis enzyme can act on both type 1 (Gal β 3GlcNAc) chains to make Le^a and Le^b (H type 1 as precursor) and type 2 (Gal β 4GlcNAc) chains to make Le^x and Le^y (H type 2 as precursor) (Liu et al. 2005). *FUT7* (FucT7) is the α 1,3-fucosyltransferase which has the highest specificity for sialylated lactosamine structures and produces SLe^x.

In healthy stomach epithelium, type 1 Lewis antigens Le^a and Le^b are mostly found in the superficial epithelium, while the type 2 structures Le^x and Le^y are located in the deep glands (Borén et al. 1993; Linden et al. 2004). The expression of Le^b correlates with MUC5AC whereas Le^y correlates with MUC6 (De Bolós et al. 1995).

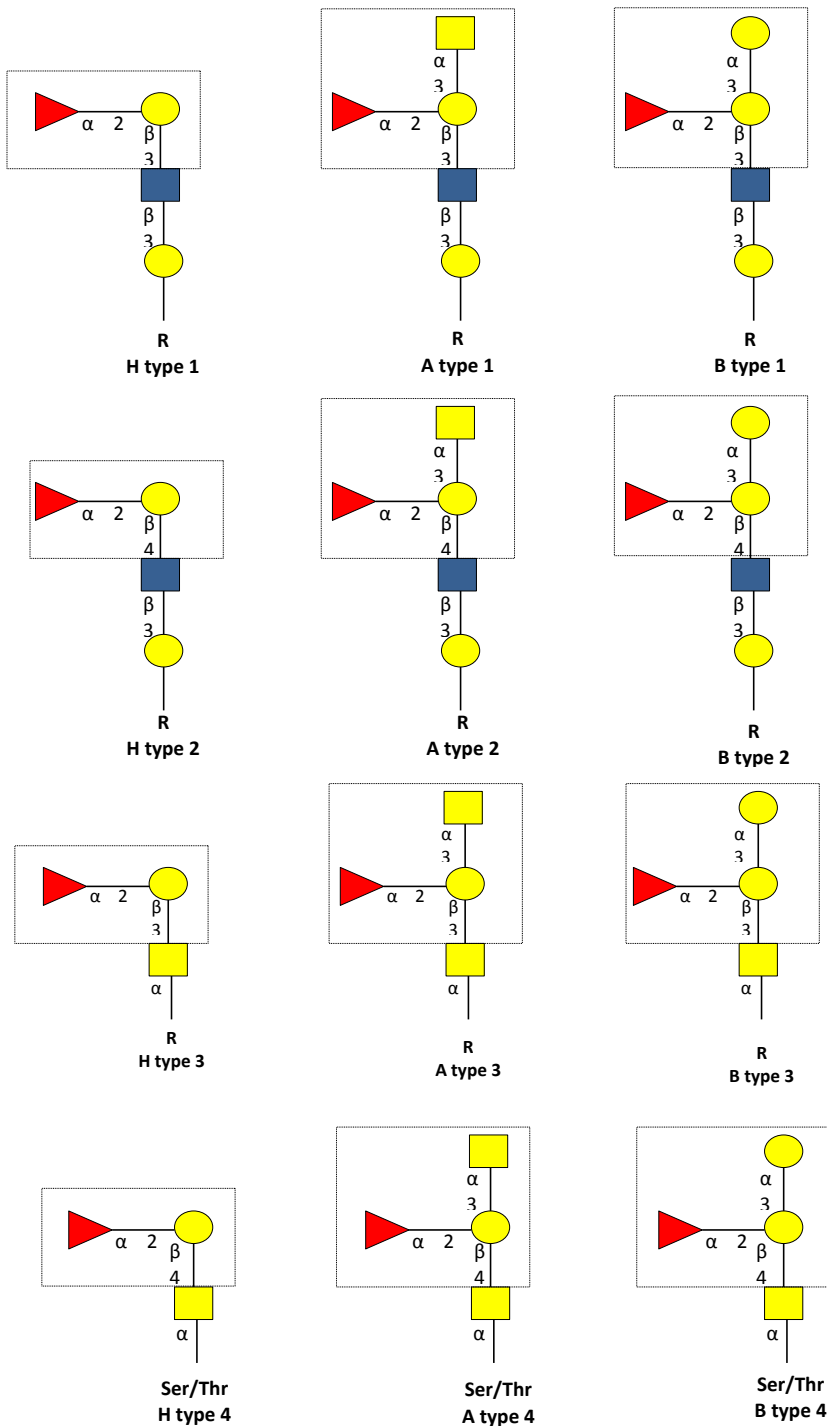


Figure 5: ABO (H) blood group determinants on different carbohydrate outer core chains

1.10.2. The LacdiNAc determinant

LDN is a terminal structure found in *N*- and *O*-glycans on vertebrate and invertebrate glycoproteins (Dell et al. 2003). LDN has been shown to be expressed on MUC5AC (Kenny et al. 2012) and on the epithelial surface of the stomach, heart, colon, small intestines and salivary glands (Gotoh et al. 2004). It is also highly expressed in parasitic helminthes such as *Schistosoma mansoni* (van Die and Cummings 2009). The LDN determinant contributes to the self-renewal of mouse embryonic stem cells (Che et al. 2014) and regulates half-life of lutropin (Hirano et al. 2014). Its expression is also high on colon cancer cells *in vitro*, and promotes tumor growth and metastasis *in vivo* (Huang et al. 2007). LDN is synthesized when GalNAc is transferred in a β 1,4-linkage to GlcNAc by human β 1,4-*N*-acetylgalactosaminyltransferase 3 (β 4GalNT3) and 4 (β 4GalNT4) (Gotoh et al. 2004; Sato et al. 2003).

These two β 4GalNAcTs show about 43% homology at the amino acid level. β 4GalNT3 contains 999 amino acids, four potential *N*-glycan sites and multiple *O*-glycan sites, while β 4GalNT4 contains 1039 amino acids, three potential *N*-glycan sites and multiple *O*-glycan sites (Gotoh et al. 2004; Sato et al. 2003). Their expression patterns are different; B4GALNT3 is abundantly expressed and limited to the surface mucous cells in human gastric mucosa (Ikehara et al. 2006), colon and testis. Its expressions in human colon cancer plays a role in promoting malignant behaviors. B4GALNT4 is predominantly expressed in ovary, fetal and adult brain, and fetal kidney and lung (Gotoh et al. 2004; Sato et al. 2003), its upregulation has also been reported in prostate cancer (Fukushima et al. 2009). Both β 4GalNT3 and 4 prefer the *O*-glycan core 6 over core 2 and 3 as the substrate (Gotoh et al. 2004; Sato et al. 2003).

1.10.3. The GlcNAc α 4Gal-terminal

GlcNAc α 4Gal-terminals have been suggested to have antimicrobial activity against *H. pylori* (Kawakubo et al. 2004; Nakayama et al. 1999). In an *in vivo* rat (Nakayama et al. 1999) and in an *in vitro* study (Kawakubo et al. 2004), it was shown that gastric mucins carry these GlcNAc α 4Gal-terminals on *O*-glycans. An α 4-*N*-acetylglucosaminyltransferase 1 (*A4GNT1*) is responsible for its biosynthesis by transferring GlcNAc from a UDP-GlcNAc

to a β Gal residue in an α -linkage on *O*-glycans. *O*-glycans with α 1,4GlcNAc was found in the gastric glands and were shown to be carried by MUC6 (Hidaka et al. 2001; Karasawa et al. 2012; Zhang et al. 2001).

The GlcNAc α 4Gal glycans are frequently expressed in gastric cancer cells and are therefore regarded as gastric cancer-associated carbohydrate antigens (Nakamura et al. 1998). GlcNAc α 4-carrying *O*-glycans have been shown to have antimicrobial activity against *H. pylori* by inhibiting the biosynthesis of cholesteryl- α -D-glucopyranoside, which is a major cell wall component (Kawakubo et al., 2004; Nakayama et al., 1999). The anti-microbial activity of GlcNAc α 4 *O*-glycans was demonstrated by culturing *H. pylori* in the presence of a recombinant soluble CD43 that was capped by GlcNAc α 4 glycans and whereby the growth of the bacteria was inhibited in a dose-dependent manner (Kawakubo et al. 2004).

Zheng *et al.* in an investigation of the role of sequence variation in the *A4GNT* gene with regards to risk of infection with *H. pylori* or gastric cancer found that some variants of *A4GNT* were related to *H. pylori* infection in the population (Zheng et al. 2009). Karasawa *et al.*, showed that this enzyme plays a significant role in gastric protection by demonstrating that mice lacking this enzyme (α 4gntKO) typically showed gastric mucosal inflammation and spontaneously developed adenocarcinoma exclusively in the gastric antrum through a hyperplasia-dysplasia-carcinoma sequence also in the absence of *H. pylori* infection (Karasawa et al. 2012).

1.11. Glycans in gastric cancer

More than 1 million individuals are gastric cancer afflicted per year, with the majority of patients dying within a year after diagnosis (Ferlay et al. 2010; Parsonnet et al. 1991). Gastric cancer takes second place as the leading cause of cancer-related death and the fourth most common malignancy in the world following cancers of the lung, breast and colon-rectum (Ferlay et al. 2010). High gastric cancer incidence rates are reported in Europe, East Asia and South America (Parkin et al. 1984), with the age-standardized incidence rates being about twice as high in men as in women (Carter 2014).

Two histopathophysiological subtypes of gastric cancer, with different clinicopathological characteristics, have been described by the Lauren classification (Lauren 1965). The intestinal, which represents 70% of the cases, and the diffuse representing nearly 30% of cases (Pinho et al. 2013). In comparison to the diffuse-type gastric cancer, the intestinal-type cancers are much more passive in terms of their growth, whereas the diffuse-type cancers are composed mainly of tumor cells displaying gastric differentiation and show poorer survival. The early detection of gastric cancer is critical to expect favorable outcome of the disease since the tumor can be completely removed by surgical procedures if detected in its early stages, whereas advanced gastric cancer with serosal invasion recurred after resection of the primary tumour (Shimizu et al. 2003). Carcinoembryonic antigen (CEA) and Carbohydrate antigen 19-9 (CA19-9) tests are useful adjunct for staging cases and monitoring therapy in patients with gastric cancer (Szymendera 1986).

However, the limitation of these methods in the early diagnosis of gastric cancer has been documented (Pectasides et al. 1997). Both markers are elevated in advanced cancers, but not in early stages of the cancer (Shimizu et al. 2003). Different factors are associated with the progression of gastric carcinogenesis, with diet being one of the etiological factors as well as other factors as smoking.

A combination of environmental factors and genetic alterations are thought to be the cause of gastric cancer (Milne et al. 2009). Persistent infection with *H. pylori* is also a causative agent of this disease. Treatment to eradicate this bacterium using antibiotics fail in 20% of the patients due to antibiotic bacterial resistance and poor patient compliance, potentially leaving 140 million people in the world without alternative therapy (Gonçalves et al. 2016), resulting in the justification if an alternative therapy.

Glycosylation alterations in gastric cancer

A shift in normal glycosylation to altered glycosylation as may be seen in cancer cells can be caused by over- or under-expression of glycosyltransferases (Taylor-Papadimitriou et al. 1999). Secondly, there could be changes in the conformation of the protein. Thirdly, differences in glycan expression may occur as a consequence of variability and

availability of precursor chains and sugar nucleotide donors (Ohtsubo and Marth 2006). Further, glycosyltransferase location in the Golgi apparatus may be deranged (Tuccillo et al. 2014). Some of the cancer-associated changes in glycosylation include sialylation, fucosylation, truncation and branching (Pinho and Reis 2015).

1.12. Glycosylation in other disease

Several patho-physiological processes can be controlled by glycosylation. Defect in such a key regulatory mechanism as glycosylation in humans and its links to disease show that the mammalian glycome contains biologically valuable information. The structural profile of a cell membrane or glycan determinants on glycoproteins can change in disease (Reily et al. 2019). Alterations in sialylation, galactosylation and/or fucosylation, as well as changes in glycan composition, can contribute to immune dysregulation and a range of autoimmune and chronic inflammatory diseases. The characterization of biological functions of glycans and glycan binding proteins such as galectins has provided important contributions in biology (Pinho and Reis 2015). The influence of glycans on the development of various diseases other than cancer, such as, congenital disorders of glycosylation (Al Teneiji et al. 2017), diabetes (Wright et al. 2017), and rheumatoid arthritis (Rombouts et al. 2015), among others, has been reported.

1.12.1. Rheumatic arthritis (RA) and osteoarthritis (OA)

RA is an autoimmune disease that results in chronic inflammation of joints and other associated tissues (Kourilovitch et al. 2014). Changes in the glycosylation of antigen-specific IgGs precede disease onset, and the degree of these changes correlates with disease severity. Altogether, these findings suggest that changes in the glycosylation of IgG are a critical component of RA pathogenesis.

OA is characterized by progressive loss of articular cartilage and formation of osteophytes, which lead to chronic pain and functional restrictions in the affected joints. There are many factors which could lead to the development of OA. Traumatic events

are causative but there are other factors like genetic predisposition, defective position of joints, ageing and malnutrition, which all lead to similar alterations in the joint cartilage (Lorenz and Richter 2006).

1.12.2. Galectin-3 as the major LDN-binding lectin

Galectins are conserved proteins from a family of lectins present in vertebrates, invertebrates and fungi (Diaz-Alvarez and Ortega 2017). Galectins have characteristic carbohydrate-recognition domains (CRD) of about 130 amino acids with specificity toward β -Gal and LacNAc residues (Varki A 2009). There are 15 galectins that have been identified in mammals with galectin-3 being widely distributed on gastric surface mucous cells (Nio-Kobayashi 2017; Nio-Kobayashi et al. 2009). Most galectins appear to be divalent proteins with two CRDs, while galectin-3 appears to be pentameric in its binding to multivalent glycans (Ahmad et al. 2004).

Myeloid cells, including macrophages, are responsible for the synthesis of Galectin-3 which then binds to a variety of endogenous glycoprotein ligands, including cell surface receptors and extracellular matrix proteins (Barondes et al. 1994). Galectin-3 has also been shown to be a major LDN-binding lectin in rat macrophages, and as such it can mediate interactions between macrophages and schistosome LDN glycans (van den Berg et al. 2004). The carbohydrate binding specificity of all the mammalian galectins examined have the ability to recognise the same determinants on lactose and related β -galactosides.

2. AIMS OF THE THESIS

The overall objective of the thesis was to use a glyco-engineered mucin-type fusion protein carrying frequent *O*-glycan substitution with defined glycan determinants as a tool to investigate carbohydrate-dependent *H. pylori* interactions.

Specific aims were to:

- I.** Determine the specificity of β 4GalNT3 on different *O*-glycan core chains using PGSL-1/mIgG2b as a mucin-type reporter protein and to assess the ability of the resulting LDN determinants to bind the *H. pylori* adhesin LabA.
- II.** Characterize the *O*-glycomes of PSGL-1/mIgG2b carrying Le^b and sLe^x, respectively, on different *O*-glycan core chains and to investigate their binding to the *H. pylori* adhesins BabA and SabA.
- III.** Investigate the *O*-glycan core chain specificity of α 4-N-acetylglucosaminyltransferase 1 (*A4GNT1*) and to evaluate a potential anti-proliferative effect of GlcNAc α 4Gal-terminals on *H. pylori*.
- IV.** Characterize glycoforms of synovial lubricin from OA patients and healthy individuals, and to compare it with the glycosylation of rhPRG4 expressed in CHO-cells, then analyse differences in their binding of galectin-3.

3. MEHODOLOGICAL CONSIDERATION

Detailed materials and methods utilized in the thesis have been described in each paper. Here, a brief discussion on the methodology used is provided.

3.1. Cell Culture and transfection of CHO cells

CHO cell lines have been used as the cells of choice since they are well characterized and often produce glycans like those in human cells. CHO-K1 cells were seeded in 75-cm² T-flasks in Dulbecco's Modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 2mM L-glutamine as well as 100 units/mL penicillin and 100 µg/mL streptomycin. FBS was the supplement used in the medium and provided carriers or chelators for labile and water-insoluble nutrients, hormones and growth factors, protease inhibitors and toxic substances. It is also a source of factors which may be necessary for the proper attachment and spreading of cells on the plastic culture substrate. The cells were maintained at 37°C and 5.0% CO₂ in a humidified incubator as a requirement for the cells to grow optimally. In the case of large scale culturing of stable transfectants, it was difficult to reach the high cell densities obtained in systems like the wave bioreactors where the levels of glucose, glutamine and pH could be monitored (Clincke et al. 2013).

In paper I, III and IV, transfected cells were adapted to serum-free medium (ProCHO-4 medium) supplemented with and 100 units/mL penicillin and 100 µg/mL streptomycin starting 24 hours post-transfection. Serum-free medium was used for culturing transiently transfected cells in order to decrease the level of serum proteins in the cell culture supernatant; large amounts of serum proteins makes purification of the fusion protein more difficult. For that same reason, stably transfected cells were also adapted to serum-free conditions. This is done in a stepwise manner in order to make sure they do not lose gene expression. Serum supplementation in media posed a challenge regarding purification and further downstream processing.

Table 3. Expression vectors used for transfections in CHO-K1 cells

Construct	Protein	Gene	Selection drug
PSGL-1/mIgG2b	Human PSGL-1 fused with mouse IgG2b	<i>PSGL-1/mIgG2b</i>	Puromycin
C2 β 6GnT1	β -1,3-galactosyl-O-glycosyl-glycoprotein β -1,6-N-acetylglucosaminyltransferase 1	<i>GCNT1</i>	G418
C3 β 3GnT6	Acetylgalactosaminyl-O-glycosyl-glycoprotein β -1,3-N-acetylglucosaminyltransferase	<i>B3GNT6</i>	Hygromycin
Extended C1 β 3GnT3	β 1,3-galactosyl-O-glycosyl-glycoprotein β -1,3-N-acetylglucosaminyltransferase	<i>B3GNT3</i>	Zeocin
α 1,2Fuc-T2	α 1,2 fucosyltransferase	<i>FUT2 (Secretor gene)</i>	G418
α 1,3 and α 1,3/4Fuc-Ts	α 1,3/4 fucosyltransferase	<i>FUT3 (Lewis gene)</i>	Hygromycin
α 1,3 FucTs	α 1,3-fucosyltransferase	<i>FUT7</i>	Zeocin
<i>GalT5</i>	β 1,3-galactosyltransferase-V	<i>GALT5</i>	Mycophenolic
β 4GnT3	β 1,4-N-acetylgalactosaminyltransferase 3	<i>GALNT3</i>	G418
α 4GnT	β 1,4-N-acetylgalactosaminyltransferase 3	<i>A4GNT1</i>	Hygromycin

For transfection purposes, we used lipofectamine; a cationic liposome which enabled the delivery of our CDM8 vectors carrying the cDNA of PSGL-1/mIgG2b or of different glycosyltransferases (Table 3). The expression vectors had EF1 α or CMV-IE promoters to drive expression. CHO-K1 were transiently or stably transfected in different experiments, with the transients yielding proteins in less time but in lower amounts.

3.2. Purification of secreted recombinant PSGL-1/mIgG2b

Large scale purification procedures were carried out using affinity chromatography columns containing immobilized Protein A Sepharose CL-4B. Protein A is a surface protein of *S. aureus*, which is very stable, has a molecular weight of 42kDa and binds to the Fc region of immunoglobulins through interactions with the heavy chain (Yang et al.

2003). Sepharose CL-4B is a cross-linked agarose gel filtration matrix on which the protein A was covalently coupled. PSGL-1/mIgG2b bound to the immobilized Protein A was eluted with a sodium citrate buffer with a pH of 3. Gel filtration was a second purification step used for the removal of proteins co-purified with PSGL-1/mIgG2b on the affinity chromatography column.

3.3. Quantification of PSGL-1/mIgG2b by anti-mouse IgG Fc enzyme-linked immunosorbent assay (ELISA)

An enzyme-linked immunosorbent assay was used to measure the concentration of PSGL-1/mIgG2b using a dilution series of an IgG2b antibody with known concentration as a standard. A polyclonal goat anti-mouse IgG Fc antibody was used to pre-coat 96-well plates and to capture the fusion protein of interest. Detection of captured fusion protein was accomplished by using an HRP-conjugated form of the same antibody. TMB (3,3',5,5'-tetramethylbenzidine) was used as the substrate for HRP and turned blue in color before stopping the reaction with sulphuric acid.

3.4. Characterization of PSGL-1/mIgG2b and its carbohydrate phenotype using SDS-PAGE and Western blot analysis

PSGL-1/mIgG2b was affinity purified from supernatants of transfected CHO cells using goat anti-mIgG (Fc-specific)-coupled agarose beads. In order to show that the fusion protein carried the glycans expected and as directed by the glycosyltransferase cDNAs expressed, Western blot (WB) analysis was performed using antibodies specific for PSGL-1 and IgG Fc, as well as antibodies or lectins specific for carbohydrate determinants as listed in Table 4. HRP-conjugated polyclonal goat anti-mouse IgG Fc for detection of the Fc portion of PSGL-1/mIgG2b, HRP-conjugated polyclonal goat anti-mouse IgG (Fab-specific) and HRP-conjugated goat anti-mouse IgM were used (Papers I-IV).

Mouse anti-human CD162 was used to detect the *N*-terminal part of PSGL-1 (Papers I-IV). An anti-LDN (anti-LDN; mouse IgG,) antibody specific for GalNAc β 4GlcNAc (*N,N'*-diacetyllectosdiamine/LDN) determinants was a gift from Prof. Richard Cummings, Harvard University (Nyame et al. 1999) (Papers I & II). The mouse anti-Le^b (IgM; T218) antibody recognizing Fuc α 2Gal β 3(Fuc α 4)GlcNAc and the mouse anti-CD15s antibody recognizing NeuAc α 3Gal β 4(Fuc3)GlcNAc were used (Paper II). For detection of GlcNAc α 4-R glycans, the cells were incubated with biotinylated *Griffonia (Bandeiraea) simplicifolia* lectin II (Nagano et al. 2005; Nakamura-Tsuruta et al. 2006; Wood et al. 1978) (Paper II & III).

Table 4. Primary antibodies and lectins used in Western blotting

Structure recognized	Antibody or Lectin	Isotype	Clone	Manufacturer	Dilution
IgG Fc	anti-mouse IgG Fc	polyclonal (affinity isolated)		Sigma	1:10 000
PSGL-1	anti-PSGL-1 (CD162)	mouse monoclonal IgG	KPL-1 (RUO)	BD Biosciences	1:1 000
Fuc α 2Gal β 3[Fuc α 4]GlcNAc-R	anti-Le ^b	mouse monoclonal IgM	(T218)	Santa Cruz Biotechnology	1:2 000
NeuAc α 3Gal β 4[Fuc α 3]GlcNAc-R	anti-SLe ^x (CD15s)	mouse monoclonal IgM	CSLEX 1 (RUO)	BD Biosciences	1:2 000
GalNAc β 4GlcNAc (LacdiNAc)	anti-LDN	mouse monoclonal IgG		Gift from Richard Cummings	1:2 00
GlcNAc α 1	biotinylated <i>Griffonia (Bandeiraea) simplicifolia</i> lectin II			Vector Laboratories	1:2 00

The recognition of a particular carbohydrate by the lectin is dependent on its presentation, linkage type to the penultimate sugar residue, core chain upon which the determinant is carried and sometimes the protein backbone on which the glycan is situated. Cross-reactivity was avoided by testing out the correct dilutions of both the

primary and secondary antibodies. Both positive and negative controls were included on the WB.

3.5. Chemical release of *O*-linked glycans from purified PSGL-1/mIgG2b prior to LC-MS analysis

O-linked glycans of purified PSGL-1/mIgG2b were released chemically by β -elimination and were reduced into alditols. This can be done on glycoproteins in solution or blotted onto PDVF membranes following separation by SDS-PAGE. The glycoprotein is incubated in 0.5 M NaBH₄ in 50 mM NaOH at 50°C for 16 hours. One disadvantage of the method is that there is no possibility to label the released oligosaccharides with for example aminobenzamide (2-AB). Its advantage is that it can reduce the destructive peeling of oligosaccharides which could occur with other release methods such as mild hydrazinolysis (Chiba et al. 1997). Sodium ions in the oligosaccharide samples were removed by cation exchange chromatography and repeated evaporation with 1% acetic acid in methanol was used to remove borate salts as borate esters.

3.6. LC-MS/MS analysis

In order to determine the composition and the sequence of the *O*-glycans in our recombinant protein samples, we used the mass spectrometer instrument and fragmentation techniques. The mass spectrometer is comprised of an ionization source which allows for the ionization of analytes and thereby forming ions in a gaseous phase before entering the MS. The MS also has a mass analyzer which can separate ions based on their (*m/z* value) size and isomeric structure; ions that get sent to the detector to output their *m/z* value. In this thesis, we used Electrospray Ionization (ESI) as an ionization mode. ESI allows for the ionization of the analytes in a solution, following their passage through a capillary held under high voltage leading to the nebulization of the solution, thus forming a gaseous mist. Through evaporation and a process of droplet subdivision resulting from high charge density (Coulomb explosion), the charged

droplets diminish in size (Goodarzi and Turner 1998). These ions then enter the MS in their gaseous form to be analyzed.

Negative ion mode was used and was preferred over positive ion mode, to analyze the reduced oligosaccharides, even though a less informative spectrum of the MSn fragmentation of sialylated oligosaccharides would be obtained as a result of the labile sialic acid. Oligosaccharides carrying negative charges because of the presence of sialic acid, hexuronic acid and/or sulfates/phosphates allow for efficient ionization of these oligosaccharides. The labile sialic acid is often lost during fragmentation in negative ion mode, but can be stabilized by derivatization in order to investigate the sialylated structures (Kang et al. 2005).

MS/MS spectra were manually interpreted for structural assignments. UniCarb-Dr was helpful for providing a reasonably quick MS2 spectral intensity comparison for annotating the structures. The relative amount of each oligosaccharide was retrieved from mass spectra averaged under the chromatographic peaks observed in the base peak trace.

3.7. Analysis of *H. pylori* adhesion

To assess the binding of *H. pylori* to our glycoconjugates and mucin samples *in vitro*, a microtiter well-based assay was used. In paper I, the mucins and the mucin-type fusion proteins (LDN carried by different *O*-glycan core structures on PSGL-1/mIgG2b) diluted in GuHCl were coated on 96-well polysorb plates and incubated with biotinylated *H. pylori*. Bound bacteria were detected by the reaction between biotin carried on the bacteria and horseradish peroxidase conjugated streptavidin with TMB (3,3',5,5'-tetramethylbenzidine) as peroxidase substrate to visualize the binding. Biotinylation of *H. pylori* was done in order to prevent any cross reactivity of the antisera with mucin samples. The biotinylation was mild and did not affect the bacteria's binding ability. Optical density at 450 nm was used to measure bacterial binding using a microplate reader.

In paper II, the *H. pylori* strains used were CCUG17875/Le^b and CCUG17875DM. The 17875/Le^b strain is a spontaneous mutant that binds to Le^b but does not bind to sialylated antigens, while the CCUG17875DM binds to sLe^x. Binding detection to our recombinant proteins was performed by using *H. pylori* radiolabelled with ¹²⁵I. The bound (pellet) and the free (supernatant) radiolabelled HSA-conjugate activity were measured in a Wizard gamma counter.

In paper III, we studied the effect on *H. pylori* growth by purified PSG-L/mIgG2b with GlcNAc α 4Gal-terminals on core 1 and core 2. The bacteria were cultured in the presence of the purified samples before measuring bacterial growth using optical density (OD), which is directly related to bacterial cell count. A disadvantage with the OD measurement method is that the viability of the bacteria (ratio of live and dead cells in the bacterial culture) is not provided.

4. RESULTS AND DISCUSSION

4.1. Production of recombinant PSGL-1/mIgG2b with tailored glycosylation (Papers I-IV)

This thesis focuses on using recombinant mucin-type fusion proteins PSGL-1/mIgG2b as a tool for elucidating different *O*-glycan phenotypes and how these different glycoforms of PSGL-1/mIgG2b interact with *H. pylori* and galectin-3(Gal-3). The mucin-type characteristic of PSGL-1 provides a multivalent presentation of its *O*-glycans. We have in all our papers used CHO-K1 as host cells for expression of PSGL-1/mIgG2b. The glycosylation machinery of wild type CHO-K1 cells supports the biosynthesis of mono- and disialylated core 1 *O*-glycans (Papers I-III). In addition to the plasmid encoding PSGL-1/mIgG2b, other plasmids encoding specific glycosyltransferases are expressed in CHO-K1 cells to allow for expression of specific carbohydrate determinants. We show in papers I-IV that CHO-K1 cells can be glyco-engineered to express *O*-glycans on PSGL-1/mIgG2b carrying LDN, Le^b, sLe^x, and GlcNAc α 4Gal determinants on different core structures.

The *O*-glycan biosynthesis pathways in CHO-K1 were studied by transiently and stably expressing *O*-glycan core chain glycosyltransferases C2 β 6GnT1, C3 β 3GnT6 or extended C1 β 3GnT3. In paper I-III, functional carbohydrate receptors of *H. pylori* were produced, whereas in paper IV, core 2 was produced to test if core 2 glycans compared to the simpler core 1 glycans were responsible for Gal-3 binding.

Paper I: PSGL-1/mIgG2b carrying *O*-glycans harbouring the LDN (GalNAc β 4GlcNAc β 1-R) determinants on different core chains were transiently produced to study the core chain dependence of the LDN producing β 4-*N*-acetylgalactosaminyltransferase 3 (β 4GalNT3).

Paper II: We transiently glyco-engineered CHO-K1 cells to express PSGL-1/mIgG2b carrying Le^b and sLe^x determinants on core 2, core 3 and extended core 1 *O*-glycan chains, thereby we could study the core chain dependence of glycosyltransferases responsible for producing Le^b and sLe^x determinants.

Paper III: We transiently produced GlcNAc α 4Gal-terminals on *O*-glycans carried on core 1, core 2, core 3 and extended core 1 precursors in order to study the core chain specificity of α GnT.

Paper IV: CHO-K1 cells were stably transfected with plasmids encoding PSGL-1/mIgG2b and the C2 GnT1 glycosyltransferase supporting the generation of core 2 *O*-glycans.

Further, in paper I and II, the inhibition of *H. pylori* binding was also studied using transiently produced LDN, Le^b, SLe^x determinants, respectively on core 2, core 3 and extended core 1 *O*-glycan chains. In paper III, stably produced GlcNAc α 4Gal-terminals on core 1 and core 2 *O*-glycans were used to study their effect on *H. pylori* growth.

4.2. Characterization of recombinant PSGL-1/mIgG2b with tailored glycosylation (Papers I-IV)

Western blot (WB) and liquid chromatography – mass spectrometry (LC-MS/MS) were employed in the characterization of recombinant PSGL-1/mIgG2b. WB analysis (Papers I-IV) using antibodies specific for mIgG Fc and the *N*-terminal of PSGL-1 showed that the fusion protein PSGL-1/mIgG2b was produced as a homodimer with an apparent MW between 268 and 460 kDa under non-reducing conditions, with additional stained bands migrating between the 117 and 268 kDa marker indicating the presence of PSGL-1/mIgG2b monomers and degradation products. The non-glycosylated PSGL-1/mIgG2b dimer has an expected molecular weight of about 117kDa (Gustafsson et al. 2011; Liu et al. 1997; Liu et al. 2003), indicating increased glycosylation.

Paper I: The antibody against the LDN revealed strong staining of the fusion protein expected to carry LDN on core 2 and core 3, while no or very weak staining of the fusion protein was seen when LDN was expected on extended core 1 or on core 1 *O*-glycans in wild-type CHO-K1 cells. However, the LC-MS/MS data did not only verify the presence of terminal LDN determinants on core 2 and core 3, but also on extended core 1 *O*-glycans, suggesting that the anti-LDN antibody recognizes the LDN determinant in a core chain-dependent manner. The LDN determinant was not detected on *O*-glycans on PSGL-

1/mIgG2b expressed in wild-type CHO-K1 cells upon co-transfection with the B4GALNT3 enzyme alone, which is likely explained by the fact that the core 1 *O*-glycans lack a terminal β -linked GlcNAc that could serve as a substrate for B4GALNT3.

Paper II: The Le^b determinant expression on PSGL-1/mIgG2b carrying core 2, core 3 and extended core 1 *O*-glycans was verified by anti-Le^b antibody staining. A more intense staining of the Le^b was observed on core 3 chains and less so on core 2 *O*-glycans. The absence of any definite Le^b *O*-glycans in the *O*-glycome of PSGL-1/mIgG2b co-expressed together with enzymes supporting Le^b and core 2 chain biosynthesis was assessed and verified by LC-MS/MS, suggesting that core 2 is not a very efficient precursor chain for type 1 extension and subsequent Le^b biosynthesis. Alternatively, the core 2 chain was consumed by efficient β 4galactosylation thereby competing with the type 1 chain-generating GALT5.

Still in paper II, a sLe^x antibody was also used in WB experiments to verify the presence of the sLe^x determinant on core 2, core 3 and extended core 1 chains. Differences in sLe^x antibody staining intensity between the different core chains were observed. sLe^x on core 2 and extended core 1 were more strongly stained than sLe^x on core 3. The LC-MS/MS data confirmed that FUT7 was active and could generate sLe^x determinants on core 2, core 3 and extended core 1 *O*-glycans.

As controls in the study, PSGL-1/mIgG2b carrying a LDN determinant on core 2 and a GlcNAc α 4Gal-terminal on core 1 were also generated. Their presence were confirmed by WB using an antibody against LDN and the biotinylated *Griffonia simplicifolia* (GSL-II) lectin, respectively, as well as by LC-MS/MS.

Paper III: The GSL-II lectin was used for the staining of terminal GlcNAc α -R moieties carried on core 1, core 2, core 3 and extended core 1 chains with the staining intensity being similar between the core chains, suggesting that α 4GnT is not core chain-specific. LC-MS/MS and immunocytochemistry also confirmed the successful transfection of α 4GnT and other glycosyltransferases supporting the biosynthesis of core 2, core 3 and extended core 1.

Paper IV: We used biotinylated Gal-3 lectin staining to show that Gal-3 binds to core 2 *O*-glycans. A strong Gal-3 staining of PSGL-1/mIgG2b carrying core 2 but not core 1 *O*-glycans was revealed.

4.3. Carbohydrate-dependent inhibition of protein-carbohydrate interactions

4.3.1. LDN-containing glycoconjugates do not bind *H. pylori* (Paper I)

Previous experiments had showed that free LDN saccharides could inhibit the binding of several *H. pylori* strains (B128, 26695, J99, G27, P12 and B38) to human gastric mucosa tissue sections, suggesting a novel LDN-binding adhesin in those strains (Rossez et al. 2014). We wanted to further characterize this LDN-binding specificity of the LabA adhesin, as well as determine any potential *O*-glycan core chain-dependent binding. Upon confirming the identity of these LDN containing glycoconjugates on different core chains, binding to *H. pylori* strains 26695 and J99 was assessed. We did not observe a correlation between *H. pylori* binding and the expression of LDN determinants on PSGL-1/mIgG2b carrying core 2, 3 and extended core 1 *O*-glycans, or on gastric mucins. We also did not identify binding of a ¹²⁵I-labelled LDN-BSA neoglycoconjugate to any clinical *H. pylori* isolates

4.3.2. Competition inhibition of *H. pylori* by glycoconjugates carrying Le^b, sLe^x, LDN and GlcNAc α 4Gal on various core chain *O*-glycans (Paper II)

H. pylori binds mucin-carried glycans through its blood group antigen-binding adhesin, BabA, which recognizes the Lewis b (Le^b) and H type 1, while its sialic acid-binding adhesin, SabA binds to sLe^x and sLe^a. We investigated whether the type of core chain carrying Le^b or sLe^x influenced their binding to *H. pylori* strains holding the BabA or

SabA adhesins, respectively. Different concentrations of PSGL1/mlgG2b fusion proteins were competed against ^{125}I -HSA Le^b for the binding to BabA adhesin in *H. pylori* strain CCUG 17875/Leb. The fusion protein carrying Le^b on core 3 showed stronger inhibition properties towards the Le^b – BabA binding compared with the lower expression of Le^b on either core 2 or extended core 1. In addition, none of the mucin-type fusion proteins presenting sLe^x were recognized by BabA. The mucin-type fusion proteins were also competed against ^{125}I -HSA sLe^x for binding to the SabA adhesin in *H. pylori* strain CCUG 17875DM. Our data showed some degree of inhibition properties against SabA-mediated binding for all the PSGL1/mlgG2b glycoforms, suggesting inhibition properties correlating to the degree of sialylation of the fusion protein.

4.3.3. Assessment of *H. pylori* growth in the presence of GlcNAc α 4Gal-carrying mucin-type fusion proteins or purified human gastric mucins with or without GlcNAc α 4Gal-terminals (Paper III)

There is contradictory data regarding the antimicrobial activity of the GlcNAc α 4Gal-terminals synthesized by the α 1, 4-*N*-acetylglucosaminyltransferase (α 4GnT). While Kawakubo *et al*, has suggested that GlcNAc α 4Gal-terminals have antimicrobial activity against *H. pylori* (Kawakubo *et al.* 2004; Nakayama *et al.* 1999) a study by Skoog *et al.* showed that gland-derived mucins which contained this GlcNAc α 4-glycostructure did not inhibit the *H. pylori* strains which were tested (Skoog *et al.* 2012). We produced and characterized PSGL1/mlgG2b stably produced in CHO-K1 cells and carrying GlcNAc α 4Gal-terminals on core 1 and core 2, and then assessed their effect on *H. pylori* growth. Our PSGL-1/mlgG2b carrying GlcNAc α 4Gal-terminals on core 1 and on core 2 did not show any growth inhibitory activity against *H. pylori*, however, the gastric mucins with and without GlcNAc α 4Gal-terminals, used as controls, inhibited *H. pylori* growth. Further investigations of the potential antimicrobial activity of the GlcNAc α 4Gal-glycostructure are justified.

4.3.4. Core 2 type oligosaccharides aid in the recognition of lubricin by galectin-3(Paper IV)

We generated a fusion protein carrying core 2. Galectin-3 could bind to the core 2 proteins *O*-linked glycans. Without the transfection of the glycosyltransferase, no binding was observed. This demonstrated that core-2 type oligosaccharides are key for recognition by Gal-3. Using surface plasmon resonance (SPR), we could show that recombinant lubricin (rhPRG4) was not able to bind Gal-3. The low GlcNAc β 1-6 glycosyltransferase activity in native CHO-cells suggests that recombinant expression of lubricin in CHO-cells is not sufficient for achieving Gal-3 binding to *O*-glycans.

Galectin-3 has been proposed to be involved in stabilizing the boundary lubricating glyco surface of the joint by binding to lubricin. It has also been shown that patients with RA have increased free Gal-3 in the synovial fluid (SF), while the literature is less conclusive about the situation in OA. Our data shows clearly that compared to the normal SF, the Gal-3 level is down in late stage OA. An altered ability of lubricin to bind Gal-3 together with a decreased Gal-3 in OA will potentially contribute to destabilization of the boundary lubrication that could spiral the joint degradation. The data together indicates that the avidity of the multimeric Gal-3 relies on numerous epitopes within lubricin and its mucin domain.

5. CONCLUSIONS

- CHO-K1 cells could be modified to produce various carbohydrate determinants and are thus suitable for glyco-engineering.
- B4GALNT3 can efficiently transfer GalNAc to GlcNAc-terminated *O*-glycans to carry LDN on the core 2, core 3 and extended core 1 structures, but not the core 1 *O*-glycans.
- Further, none of our fusion proteins carrying LDN or mucins Carrying LDN could mediate adhesion of *H. pylori* strains 26695 and J99 in a microtiter well-based adhesion assay.
- Only PSGL-1/mIgG2b proteins with Le^b on core 3 inhibited BabA-mediated binding while the series of sialylated PSGL-1/mIgG2b proteins all demonstrated various degrees of inhibition of SabA-mediated binding.
- GlcNAc α 4Gal determinant can be carried on core 1, core 2, core 3 and extended core 1.
- Galectin-3 binding to lubricin is dependent on core-2 *O*-linked glycans.

6. FUTURE PROJECTS

- Verify that LDN constitutes a significant binding receptor for *H. pylori* adhesion to the gastric mucosa.
- These proteins have allowed us to evaluate the role of the *O*-glycan core for correct presentation of the binding determinant. In the long run these findings may also serve as a template for development of inhibitors of *H. pylori* adhesion.
- Investigate the presence of other factors, other than GlcNAc α 4Gal which might be responsible, for the antimicrobial effect against *H. pylori*.
- Knock-off core 2 glycan production and study the protective ability of Lubricin *in vivo*.

7. ACKNOWLEDGEMENTS

To my supervisor, **Prof. Jan Holgersson**, my heartfelt gratitude for your help in producing this PhD thesis. I have learnt a lot from you, you are one of the smartest people I know, yet so humble! You have a magical brain which seems to remember everything, something I really admire from you. Your knowledge in Science is deep, thank you for teaching me glycobiology! You were so patient as a teacher, throwing in some jokes here and there while teaching me glycan structures. You are an excellent writer; I have grown to be aware of my own writing, thank you! You listened to my ideas regarding my project, which gave me confidence to be an independent researcher. You have incredible communication capabilities, though you were busy, you made the time to meet with us regularly and you were always attentive. One of the other things that I am grateful for was your availability to coach me in my career path and connected me to the right people. I am grateful that I had you as my supervisor; you were a very good one!

To my co-supervisor, **Dr. Niclas Karlsson**, I am grateful that I met you in 2011 as part of a collaboration I had with you while still studying towards my masters at the University of Cape Town. I don't think that I would have been in Sweden, had I not met you. I remember when I first did my experiments in your lab and had no idea how to analyze my mass spec data, later that day, I came to you and I told you I wanted to do my PhD in glycobiology, you smiled and gave me a nod and told me that you would let me know if anything came up. Honestly, I thought you would never contact me again, to my surprise, you did! Thank you for contributing knowledge and your expertise towards my thesis, and for always availing yourself whenever I needed help!

Dr. Jining Liu, thank you so much for teaching me molecular biology skills, you were such a patient teacher, and I have learnt a lot from you! I have enjoyed working with you and I will miss you☺ I totally love your sense of humor. Thank you that I could rely on you whenever I needed your expertise.

Dr. Chungsheng Jin, Thank you so much for the many hours that you spend teaching me how to analyze mass spec data. I am grateful to have worked with you Jin; you have

taught me a lot. I have often asked myself how you were able to remember the things you remember, but then I have seen your diligence, focus and excellence in what you do. You are brilliant and thank you for your help, thank you for your patience.

Prof. Sara Linden, Thank you for all the invaluable discussions during our collaborations. You are very thoughtful on the execution of experiments; I appreciate your informative comments and feedbacks. Thank you for letting me use your lab space for my *H. Pylori* studies, I am eternally grateful for all that I have learnt. **Dr. Medea Padra**, I can never thank you enough, I am so appreciative for your help whenever I needed it. Thank you for your patience and for the encouragements when experiments did not work. I enjoyed collaborating with you, thank you from my heart☺

Prof. Thomas Boren, Thank you for your collaboration and for being available when I needed help with the *H Pylori* studies. A sincere thank you to **Johan Edlund-Olofsson**, for an awesome collaboration, it was a pleasure doing science with you. I appreciate all that you have done, but most importantly for being kind.

Holgersson research group: **Dr. Reeja Maria Cherian**, I have LOVED working with you, and quite honestly, I miss you, I missed you even more when I had just returned from parental leave and you were gone. I liked how we connected in the lab. You have an amazing work ethic. I have learnt the importance of time management from you! Surprised? Thank you for all that you taught me in the lab Reeja, I still think about the possibility of us working together in the future☺. **Dr. Anki Nilsson**, Thank you for every second you spent mentoring me, for your travels from Stockholm to make sure that I was always on the right path, not only in my PhD journey but also providing me with advice regarding my career. You are so knowledgeable and focused, I am grateful to have you alongside me to mentor me. You walk the track in excellence, you inspire me. I know that this is the beginning of an awesome mentorship, I appreciate you Anki. **Dr. Lynda Mottram**, it was great to see you in our lab every now and then. Thank you for the critical reading of my thesis! All the best to you!

Prof. Suchitra Sumitran and group: It was a great pleasure to work alongside your group Suchi, I loved your excitement whenever you got amazing results in the lab. I

totally admire your resilience, despite whatever you are surrounded with. Your laughter was too contagious, sounded like a 'thousand birds'. **Dr. Vijay Kumar Kuna**, I am glad that you got married and did not sleep in the lab anymore; you had a way of making everybody feel that they were not working. Thank you and Supreja for being great friends, keep up the good work in raising your beautiful children! **XuBo**, I don't ever recall a day when you did not smile, it was a pleasure to work with you.

Niclas Karlsson Group: I think a part of me belonged to this group. **Drs. Jessica Örnros, Shika Acharya, Barbara Adamczyk and Varvara Vitiازهva**, thank you for contributing towards a wonderful journey throughout my PhD. **Dr. Kristina Thomsson (Tina)**, thank you for your encouraging words every now and then, and for 'sharing' the HPLC even when you had booked it ☺.

Olausson research group: a special thank you to **Dr. Ketaki Methe**, you have an amazing heart, thank you for those special little moments during lunch and our heartfelt motherhood conversations. **Dr. Nikhil Nayakawde, Galya Travnikova, Dr. Goditha Premaratne, Dr. Debashish Banerjee, Deepti Antony**, thank you for your friendship and conversations during lunch breaks.

Hellström Research Group: it has been such a pleasure to share the same facilities with you! **Dr. Mats Brännstrom**, I am grateful for the conversations and guidance, I like how you always think a step ahead. You were a wonderful positive influence to have, thank you! **Arvind manikantan padma**, I will miss you! You and Verona have become close to Mfundo and I, and we are grateful for you ☺

I would like to acknowledge **Prof. Anwar Mall, Marilyn Tyler (Mrs. T) and Prof. Hedrik Denver** from the University of Cape Town (South Africa), for your tremendous support throughout my research journey as an Honors and Masters student and continuing in Sweden. Thank you Prof Mall, for introducing to the world of mucins, your passion for it was admirable. Mrs. T, I will never forget the kindness and support you showed me during all 3 of my pregnancies while I was in your lab, I still make amazing cartoon character birthday cakes for my children. When I was hesitant to relocate and pursue this PhD in Sweden, you encouraged me, thank you! Prof. Denver Hendricks, I

never would have made it this far in my science journey had it not be for you, thank you for believing in me! I would like to thank **Prof. Seeham Khan** who introduced me to Molecular biology and did so with passion and excitement.

To my friends, I want to name you, but I am afraid I might get myself in trouble by forgetting some.....you most certainly held my hand in this journey, I am thankful!

Mama na papa, the simple yet profound truths you taught me in life have formed the essence of who I am today, you are simply the best!! To my late big brother, **Dungamanzi Gumede**, and my younger sister, **Nobantu Gumede**, you will always be special to me, I love you both always!! To **kokwana mama (Babsy Shilubane)** na **Mhake (Sara n'wa Makamu)**, ni mi rhandza randzi, ni khensa Xikwembu kuva xi ni nyike n'wina!!!

To the most positive person in my world, my husband, **Mfundo Mthembu**! Thank you for your sacrifice, leaving everything in South Africa so that I could pursue this PhD. You are such a pillar of strength in my life. You always bring out the best in me. You believe so much in me and our family. Thank you for all your support. I love you dearly!

To my beautiful children **Buhlebenkosi, Mpilwenhle, Okuhlekodwa and Kwazinkosi**, you are my joy and strength, my absolute 'happy place'! You inspire me to be a better human being every day! My greatest achievement would be raising you to be all that God has made you to be!

Finally, I would like to extend my sincere appreciation to the **South African National Research Foundation (NRF)** for the doctoral scholarship and to the **European Union FP7 GastricGlycoExplorer ITN** for the Marie Curie Fellowship

Where would I be, had it not been for the goodness of the Lord... **Jehovah El Shaddai**!?

8. REFERENCES

- Ahmad N, Gabius H-J, André S, Kaltner H, Sabesan S, Roy R, Liu B, Macaluso F, Brewer CF. 2004. Galectin-3 precipitates as a pentamer with synthetic multivalent carbohydrates and forms heterogeneous cross-linked complexes. *The Journal of biological chemistry*. 279:10841.
- Al Teneiji A, Bruun TU, Sidky S, Cordeiro D, Cohn RD, Mendoza-Londono R, Moharir M, Raiman J, Siriwardena K, Kyriakopoulou L et al. 2017. Phenotypic and genotypic spectrum of congenital disorders of glycosylation type i and type ii. *Mol Genet Metab*. 120:235-242.
- Andersch-Björkman Y, Thomsson KA, Holmén Larsson JM, Ekerhovd E, Hansson GC. 2007. Large scale identification of proteins, mucins, and their o-glycosylation in the endocervical mucus during the menstrual cycle. *Mol Cell Proteomics*. 6:708-716.
- Atherton JC, Cao P, Peek RM, Tummuru MK, Blaser MJ, Cover TL. 1995. Mosaicism in vacuolating cytotoxin alleles of helicobacter pylori. Association of specific vaca types with cytotoxin production and peptic ulceration. *The Journal of biological chemistry*. 270:17771.
- Atuma C, Strugala V, Allen A, Holm L. 2001. The adherent gastrointestinal mucus gel layer: Thickness and physical state in vivo. *Am J Physiol Gastrointest Liver Physiol*. 280:G922-929.
- Barnes LM, Bentley CM, Dickson AJ. 2003. Stability of protein production from recombinant mammalian cells. *Biotechnol Bioeng*. 81:631-639.
- Barondes SH, Cooper DN, Gitt MA, Leffler H. 1994. Galectins. Structure and function of a large family of animal lectins. *The Journal of biological chemistry*. 269:20807.
- Barthelson R, Mobasser A, Zopf D, Simon P. 1998. Adherence of streptococcus pneumoniae to respiratory epithelial cells is inhibited by sialylated oligosaccharides. *Infect Immun*. 66:1439-1444.
- Beum PV, Bastola DR, Cheng PW. 2003. Mucin biosynthesis: Epidermal growth factor downregulates core 2 enzymes in a human airway adenocarcinoma cell line. *Am J Respir Cell Mol Biol*. 29:48-56.
- Borén T, Falk P, Roth KA, Larson G, Normark S. 1993. Attachment of helicobacter pylori to human gastric epithelium mediated by blood group antigens. *Science*. 262:1892-1895.
- Buts L, Bouckaert J, De Genst E, Loris R, Oscarson S, Lahmann M, Messens J, Brosens E, Wyns L, De Greve H. 2003. The fimbrial adhesin fl17-g of enterotoxigenic escherichia coli has an immunoglobulin-like lectin domain that binds n-acetylglucosamine. *Mol Microbiol*. 49:705-715.
- Bäckström M, Link T, Olson FJ, Karlsson H, Graham R, Picco G, Burchell J, Taylor-Papadimitriou J, Noll T, Hansson GC. 2003. Recombinant muc1 mucin with a breast cancer-like o-glycosylation produced in large amounts in chinese-hamster ovary cells. *Biochem J*. 376:677-686.
- Carter D. 2014. New global survey shows an increasing cancer burden. *AJN The American Journal of Nursing*. 114:17.
- Che MI, Huang J, Hung JS, Lin YC, Huang MJ, Lai HS, Hsu WM, Liang JT, Huang MC. 2014. Beta1, 4-n-acetylgalactosaminyltransferase iii modulates cancer stemness through egfr signaling pathway in colon cancer cells. *Oncotarget*. 5:3673-3684.

- Cherian RM, Jin C, Liu J, Karlsson NG, Holgersson J. 2016. Recombinant mucin-type fusion proteins with a gal α 1,3gal substitution as clostridium difficile toxin a inhibitors. *Infection And Immunity*, 2016, Vol 84, Iss 10, pp 2842-52. 84:2842-2852.
- Chiba A, Matsumura K, Yamada H, Inazu T, Shimizu T, Kusunoki S, Kanazawa I, Kobata A, Endo T. 1997. Structures of sialylated o-linked oligosaccharides of bovine peripheral nerve α -dystroglycan. The role of a novel o-mannosyl-type oligosaccharide in the binding of α -dystroglycan with laminin. *J Biol Chem*. 272:2156-2162.
- Clausen H, Hakomori S. 1989. Abh and related histo-blood group antigens; immunochemical differences in carrier isotypes and their distribution. *Vox Sang*. 56:1-20.
- Clincke MF, Mölleryd C, Zhang Y, Lindskog E, Walsh K, Chotteau V. 2013. Very high density of cho cells in perfusion by atf or tff in wave bioreactor™. Part i. Effect of the cell density on the process. *Biotechnol Prog*. 29:754-767.
- De Bolós C, Garrido M, Real FX. 1995. Muc6 apomucin shows a distinct normal tissue distribution that correlates with lewis antigen expression in the human stomach. *Gastroenterology*. 109:723-734.
- del Carmen Fernández-Alonso M, Díaz D, Berbis MÁ, Marcelo F, Cañada J, Jiménez-Barbero J. 2012. Protein-carbohydrate interactions studied by nmr: From molecular recognition to drug design. *Curr Protein Pept Sci*. 13:816-830.
- Dell A, Chalabi S, Easton RL, Haslam SM, Sutton-Smith M, Patankar MS, Lattanzio F, Panico M, Morris HR, Clark GF. 2003. Murine and human zona pellucida 3 derived from mouse eggs express identical o-glycans. *Proc Natl Acad Sci U S A*. 100:15631-15636.
- Diaz-Alvarez L, Ortega E. 2017. The many roles of galectin-3, a multifaceted molecule, in innate immune responses against pathogens. *Mediators Inflamm*. 2017:10.
- Dunne C, Dolan B, Clyne M. 2014. Factors that mediate colonization of the human stomach by helicobacter pylori. *World J Gastroenterol*. 20:5610-5624.
- Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. 2010. Estimates of worldwide burden of cancer in 2008: Globocan 2008. *Int J Cancer*. 127:2893-2917.
- Fukuda M. 2002. Roles of mucin-type o-glycans in cell adhesion. *Biochim Biophys Acta*. 1573:394-405.
- Fukushima K, Satoh T, Baba S, Yamashita K. 2009. A1,2-fucosylated and β -n-acetylgalactosaminylated prostate-specific antigen as an efficient marker of prostatic cancer. *Glycobiology*. 20:452-460.
- Gaunitz S, Liu J, Nilsson A, Karlsson N, Holgersson J. 2014. Avian influenza h5 hemagglutinin binds with high avidity to sialic acid on different o-linked core structures on mucin-type fusion proteins. *Glycoconj J*. 31:145-159.
- Geisse S, Gram H, Kleuser B, Kocher HP. 1996. Eukaryotic expression systems: A comparison. *Protein Expr Purif*. 8:271-282.
- Ghaderi D, Zhang M, Hurtado-Ziola N, Varki A. 2012. Production platforms for biotherapeutic glycoproteins. Occurrence, impact, and challenges of non-human sialylation. *Biotechnol Genet Eng Rev*. 28:147-176.
- Gonçalves IC, Magalhães A, Costa AMS, Oliveira JR, Henriques PC, Gomes P, Reis CA, Martins MCL. 2016. Bacteria-targeted biomaterials: Glycan-coated microspheres to bind helicobacter pylori. *Acta Biomaterialia*. 33:40-50.
- Gonciarz W, Walencka M, Moran AP, Hinc K, Obuchowski M, Chmiela M. 2019. Upregulation of muc5ac production and deposition of lewis determinants by helicobacter pylori facilitate gastric tissue colonization and the maintenance of infection. *J Biomed Sci*. 26:23-23.

- Goodarzi M, Turner G. 1998. Reproducible and sensitive determination of charged oligosaccharides from haptoglobin by pngase f digestion and hpaec/pad analysis: Glycan composition varies with disease. *Glycoconj J*. 15:469-475.
- Gotoh M, Sato T, Kiyohara K, Kameyama A, Kikuchi N, Kwon YD, Ishizuka Y, Iwai T, Nakanishi H, Narimatsu H. 2004. Molecular cloning and characterization of beta1,4-n-acetylgalactosaminyltransferases iv synthesizing n,n'-diacetyllactosediamine. *FEBS Lett*. 562:134-140.
- Gustafsson A, Sjoblom M, Strindelius L, Johansson T, Fleckenstein T, Chatzissavidou N, Lindberg L, Angstrom J, Rova U, Holgersson J. 2011. Pichia pastoris-produced mucin-type fusion proteins with multivalent o-glycan substitution as targeting molecules for mannose-specific receptors of the immune system. *Glycobiology*. 21:1071-1086.
- Hart GW. 1997. Dynamic o-linked glycosylation of nuclear and cytoskeletal proteins. *Annu Rev Biochem*. 66:315-335.
- Hart GW, Slawson C, Ramirez-Correa G, Lagerlof O. 2011. Cross talk between o-glcnaacylation and phosphorylation: Roles in signaling, transcription, and chronic disease. *Annu Rev Biochem*. 80:825-858.
- Harvey DJ, Merry AH, Royle L, P. Campbell M, Dwek RA, Rudd PM. 2009. Proposal for a standard system for drawing structural diagrams of n - and o -linked carbohydrates and related compounds. *PROTEOMICS*. 9:3796-3801.
- Hatakeyama M. 2014. Helicobacter pylori caga and gastric cancer: A paradigm for hit-and-run carcinogenesis. *Cell Host & Microbe*. 15:306-316.
- Hazell SL LA, Brady L, Hennessy W 1986. Campylobacter pyloridis and gastritis: Association with intercellular spaces and adaptation to an environment of mucus as important factors in colonization of the gastric epithelium. *J Infect Dis*.658–663.
- Hidaka E, Ota H, Hidaka H, Hayama M, Matsuzawa K, Akamatsu T, Nakayama J, Katsuyama T. 2001. Helicobacter pylori and two ultrastructurally distinct layers of gastric mucous cell mucins in the surface mucous gel layer. *Gut*. 49:474-480.
- Hirano K, Matsuda A, Shirai T, Furukawa K. 2014. Expression of lacdinac groups on n-glycans among human tumors is complex. *BioMed Research International*. 2014:981627.
- Hoffmann W. 2015. Tff2, a muc6-binding lectin stabilizing the gastric mucus barrier and more (review). *Int J Oncol*. 47:806-816.
- Holgersson J, Breimer M, Samuelsson BE. 1992. Basic biochemistry of cell surface carbohydrates and aspects of the tissue distribution of histo-blood group abh and related glycosphingolipids. *APMIS Supplementum*, 1992, Vol 27, pp 18-27. 27:18-27.
- Holgersson J, Gustafsson A, Breimer ME. 2005. Characteristics of protein-carbohydrate interactions as a basis for developing novel carbohydrate-based antirejection therapies. *Immunol Cell Biol*. 83:694-708.
- Holgersson J, Lofling J. 2006. Glycosyltransferases involved in type 1 chain and lewis antigen biosynthesis exhibit glycan and core chain specificity. *Glycobiology*. 16:584-593.
- Huang J, Liang JT, Huang HC, Shen TL, Chen HY, Lin NY, Che MI, Lin WC, Huang MC. 2007. Beta1,4-n-acetylgalactosaminyltransferase iii enhances malignant phenotypes of colon cancer cells. *Mol Cancer Res*. 5:543-552.
- Ikehara Y, Sato T, Niwa T, Nakamura S, Gotoh M, Ikehara SK, Kiyohara K, Aoki C, Iwai T, Nakanishi H et al. 2006. Apical golgi localization of n,n'-diacetyllactosediamine synthase, beta4galnac-t3, is responsible for lacdinac expression on gastric mucosa. *Glycobiology*. 16:777-785.

- Imberty A, Varrot A. 2008. Microbial recognition of human cell surface glycoconjugates. *Curr Opin Struct Biol.* 18:567-576.
- Inamori K-i, Yoshida-Moriguchi T, Hara Y, Anderson ME, Yu L, Campbell KP. 2012. Dystroglycan function requires xylosyl- and glucuronyltransferase activities of large. *Science (New York, NY).* 335:93-96.
- Ishihara K, Kurihara M, Goso Y, Urata T, Ota H, Katsuyama T, Hotta K. 1996. Peripheral alpha-linked n-acetylglucosamine on the carbohydrate moiety of mucin derived from mammalian gastric gland mucous cells: Epitope recognized by a newly characterized monoclonal antibody. *Biochem J.* 318 (Pt 2):409-416.
- Jacobs PP, Callewaert N. 2009. N-glycosylation engineering of biopharmaceutical expression systems. *Curr Mol Med.* 9:774-800.
- Jefferson T, Jones MA, Doshi P, Del Mar CB, Hama R, Thompson MJ, Spencer EA, Onakpoya I, Mahtani KR, Nunan D et al. 2014. Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children. *Cochrane Database Syst Rev.* 2014:CD008965-CD008965.
- Joung JK, Sander JD. 2013. Talens: A widely applicable technology for targeted genome editing. *Nat Rev Mol Cell Biol.* 14:49-55.
- Kang P, Mechref Y, Klouckova I, Novotny MV. 2005. Solid-phase permethylation of glycans for mass spectrometric analysis. *Rapid Commun Mass Spectrom.* 19:3421-3428.
- Kannagi R. 2002. Regulatory roles of carbohydrate ligands for selectins in the homing of lymphocytes. *Curr Opin Struct Biol.* 12:599-608.
- Kao C-Y, Sheu B-S, Wu J-J. 2016a. Helicobacter pylori infection: An overview of bacterial virulence factors and pathogenesis. *Biomedical journal.* 39:14-23.
- Kao C-Y, Sheu B-S, Wu J-J. 2016b. Helicobacter pylori infection: An overview of bacterial virulence factors and pathogenesis. *Biomedical journal.* 39:14.
- Karasawa F, Shiota A, Goso Y, Kobayashi M, Sato Y, Masumoto J, Fujiwara M, Yokosawa S, Muraki T, Miyagawa S et al. 2012. Essential role of gastric gland mucin in preventing gastric cancer in mice. *J Clin Invest.* 122:923-934.
- Karlsson KA. 1998. Meaning and therapeutic potential of microbial recognition of host glycoconjugates. Oxford BSL. p. 1-11.
- Karlsson KA. 2001. Pathogen-host protein-carbohydrate interactions as the basis of important infections. *Adv Exp Med Biol.* 491:431-443.
- Kawakubo M, Ito Y, Okimura Y, Kobayashi M, Sakura K, Kasama S, Fukuda MN, Fukuda M, Katsuyama T, Nakayama J. 2004. Natural antibiotic function of a human gastric mucin against helicobacter pylori infection. *Science.* 305:1003-1006.
- Kenny DT, Skoog EC, Linden SK, Struwe WB, Rudd PM, Karlsson NG. 2012. Presence of terminal n-acetylgalactosaminebeta1-4n-acetylglucosamine residues on o-linked oligosaccharides from gastric muc5ac: Involvement in helicobacter pylori colonization? *Glycobiology.* 22:1077-1085.
- Keys TG, Aebi M. 2017. Engineering protein glycosylation in prokaryotes. *Current Opinion in Systems Biology.* 5:23-31.
- Khan AS, Kniep B, Oelschlaeger TA, Van Die I, Korhonen T, Hacker J. 2000. Receptor structure for f1c fimbriae of uropathogenic escherichia coli. *Infect Immun.* 68:3541-3547.
- Kim YG, Cha J, Chandrasegaran S. 1996. Hybrid restriction enzymes: Zinc finger fusions to fok i cleavage domain. *Proceedings of the National Academy of Sciences.* 93:1156-1160.
- Kinoshita T. 2016. Glycosylphosphatidylinositol (gpi) anchors: Biochemistry and cell biology: Introduction to a thematic review series. *J Lipid Res.* 57:4-5.

- Kornfeld R, Kornfeld S. 1985. Assembly of asparagine-linked oligosaccharides. *Annu Rev Biochem.* 54:631-664.
- Kourilovitch M, Galarza-Maldonado C, Ortiz-Prado E. 2014. Diagnosis and classification of rheumatoid arthritis. *J Autoimmun.* 48-49:26-30.
- Landsteiner K. 1961. On agglutination of normal human blood. *Transfusion (Paris).* 1:5-8.
- Lauren P. 1965. The two histological main types of gastric carcinoma: Diffuse and so-called intestinal-type carcinoma. An attempt at a histo-clinical classification. *Acta Pathol Microbiol Scand.* 64:31-49.
- Lindberg L, Liu J, Gaunitz S, Nilsson A, Johansson T, Karlsson NG, Holgersson J. 2013. Mucin-type fusion proteins with blood group a or b determinants on defined o-glycan core chains produced in glycoengineered chinese hamster ovary cells and their use as immunoaffinity matrices. *Glycobiology.* 23:720-735.
- Linden S, Boren T, Dubois A, Carlstedt I. 2004. Rhesus monkey gastric mucins: Oligomeric structure, glycoforms and helicobacter pylori binding. *Biochem J.* 379:765-775.
- Lindén S, Borén T, Dubois A, Carlstedt I. 2004. Rhesus monkey gastric mucins: Oligomeric structure, glycoforms and helicobacter pylori binding. *The Biochemical journal.* 379:765.
- Linden S, Nordman H, Hedenbro J, Hurtig M, Boren T, Carlstedt I. 2002. Strain- and blood group-dependent binding of helicobacter pylori to human gastric muc5ac glycoforms. *Gastroenterology.* 123:1923-1930.
- Liu J, Gustafsson A, Breimer ME, Kussak A, Holgersson J. 2005. Anti-pig antibody adsorption efficacy of {alpha}-gal carrying recombinant p-selectin glycoprotein ligand-1/immunoglobulin chimeras increases with core 2 {beta}1, 6-n-acetylglucosaminyltransferase expression. *Glycobiology.* 15:571-583.
- Liu J, Jin C, Cherian RM, Karlsson NG, Holgersson J. 2015. O-glycan repertoires on a mucin-type reporter protein expressed in cho cell pools transiently transfected with o-glycan core enzyme cdnas. *J Biotechnol.* 199:77-89.
- Liu J, Qian Y, Holgersson J. 1997. Removal of xenoreactive human anti-pig antibodies by absorption on recombinant mucin-containing glycoproteins carrying the gal alpha1,3gal epitope. *Transplantation.* 63:1673-1682.
- Liu J, Weintraub A, Holgersson J. 2003. Multivalent galalpha1,3gal-substitution makes recombinant mucin-immunoglobulins efficient absorbers of anti-pig antibodies. *Xenotransplantation.* 10:149-163.
- Lofling J, Diswall M, Eriksson S, Boren T, Breimer ME, Holgersson J. 2008. Studies of lewis antigens and h. Pylori adhesion in cho cell lines engineered to express lewis b determinants. *Glycobiology.* 18:494-501.
- Lommel M, Strahl S. 2009. Protein o-mannosylation: Conserved from bacteria to humans*. *Glycobiology.* 19:816-828.
- Lorenz H, Richter W. 2006. Osteoarthritis: Cellular and molecular changes in degenerating cartilage. *Prog Histochem Cytochem.* 40:135-163.
- Lowe JB. 1993. The blood group-specific human glycosyltransferases. *Baillieres Clin Haematol.* 6:465.
- Mahdavi J, Sonden B, Hurtig M, Olfat FO, Forsberg L, Roche N, Angstrom J, Larsson T, Teneberg S, Karlsson KA et al. 2002. Helicobacter pylori saba adhesion in persistent infection and chronic inflammation. *Science.* 297:573-578.
- Maria Cherian R, Gaunitz S, Nilsson A, Liu J, Karlsson NG, Holgersson J. 2014. Shiga-like toxin binds with high avidity to multivalent o-linked blood group p1 determinants on mucin-type fusion proteins. *Glycobiology.* 24:26-38.
- Marionneau S, Ruvoën N, Le Moullac-Vaidye B, Clement M, Cailleau-Thomas A, Ruiz-Palacois G, Huang P, Jiang X, Le Pendu J. 2002. Norwalk virus binds to histo-blood

- group antigens present on gastroduodenal epithelial cells of secretor individuals. *Gastroenterology*. 122:1967-1977.
- Marshall BJ, Warren JR. 1984. Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. *Lancet*. 1:1311-1315.
- Milne AN, Carneiro F, O'Morain C, Offerhaus GJ. 2009. Nature meets nurture: Molecular genetics of gastric cancer. *Hum Genet*. 126:615-628.
- Mitsuoka C, Sawada-Kasugai M, Ando-Furui K, Izawa M, Nakanishi H, Nakamura S, Ishida H, Kiso M, Kannagi R. 1998. Identification of a major carbohydrate capping group of the I-selectin ligand on high endothelial venules in human lymph nodes as 6-sulfo sialyl lewis x. *J Biol Chem*. 273:11225-11233.
- Mollicone R, Bara J, Le Pendu J, Oriol R. 1985. Immunohistologic pattern of type 1 (lea, leb) and type 2 (x, y, h) blood group-related antigens in the human pyloric and duodenal mucosae. *Lab Invest*. 53:219-227.
- Moore KL, Patel KD, Bruehl RE, Li F, Johnson DA, Lichenstein HS, Cummings RD, Bainton DF, McEver RP. 1995. P-selectin glycoprotein ligand-1 mediates rolling of human neutrophils on p-selectin. *J Cell Biol*. 128:661-671.
- Mukhopadhyay B, Martins MB, Karamanska R, Russell DA, Field RA. 2009. Bacterial detection using carbohydrate-functionalised cds quantum dots: A model study exploiting e. Coli recognition of mannosides. *Tetrahedron Lett*. 50:886-889.
- Mussolino C, Cathomen T. 2012. Tale nucleases: Tailored genome engineering made easy. *Curr Opin Biotechnol*. 23:644-650.
- Mårdberg K, Nyström K, Tarp MA, Trybala E, Clausen H, Bergström T, Olofsson S. 2004. Basic amino acids as modulators of an o-linked glycosylation signal of the herpes simplex virus type 1 glycoprotein gc: Functional roles in viral infectivity. *Glycobiology*. 14:571-581.
- Nagano M, Stübiger G, Marchetti M, Gmeiner G, Allmaier G, Reichel C. 2005. Detection of isoforms of recombinant human erythropoietin by various plant lectins after isoelectric focusing. *ELECTROPHORESIS*. 26:1633-1645.
- Nakamura-Tsuruta S, Kominami J, Kamei M, Koyama Y, Suzuki T, Isemura M, Hirabayashi J. 2006. Comparative analysis by frontal affinity chromatography of oligosaccharide specificity of glcnac-binding lectins, griffonia simplicifolia lectin-ii (gsl-ii) and boletopsis leucomelas lectin (bll). *The Journal of Biochemistry*. 140:285-291.
- Nakamura N, Ota H, Katsuyama T, Akamatsu T, Ishihara K, Kurihara M, Hotta K. 1998. Histochemical reactivity of normal, metaplastic, and neoplastic tissues to alpha-linked n-acetylglucosamine residue-specific monoclonal antibody hik1083. *J Histochem Cytochem*. 46:793-801.
- Nakayama J, Yeh JC, Misra AK, Ito S, Katsuyama T, Fukuda M. 1999. Expression cloning of a human alpha1, 4-n-acetylglucosaminyltransferase that forms glcnacalpha1-->4galbeta-->r, a glycan specifically expressed in the gastric gland mucous cell-type mucin. *Proc Natl Acad Sci U S A*. 96:8991-8996.
- Nio-Kobayashi J. 2017. Tissue- and cell-specific localization of galectins, beta-galactose-binding animal lectins, and their potential functions in health and disease. *Anat Sci Int*. p. 25-36.
- Nio-Kobayashi J, Takahashi-Iwanaga H, Iwanaga T. 2009. Immunohistochemical localization of six galectin subtypes in the mouse digestive tract. *J Histochem Cytochem*. 57:41-50.
- Nyame AK, Leppanen AM, DeBose-Boyd R, Cummings RD. 1999. Mice infected with schistosoma mansoni generate antibodies to lacdinac (galnac beta 1-->4glcnac) determinants. *Glycobiology*. 9:1029-1035.

- Odenbreit S, Till M, Hofreuter D, Faller G, Haas R. 1999. Genetic and functional characterization of the *alpab* gene locus essential for the adhesion of helicobacter pylori to human gastric tissue. *Mol Microbiol.* 31:1537-1548.
- Ohtsubo K, Marth JD. 2006. Glycosylation in cellular mechanisms of health and disease. *Cell.* 126:855-867.
- Oriol R, Pendu J, Mollicone R. 1986. Genetics of abo, h, lewis, x and related antigens. Oxford, UK. p. 161-171.
- Oyelaran O, Gildersleeve JC. 2009. Glycan arrays: Recent advances and future challenges. *Curr Opin Chem Biol.* 13:406-413.
- Parkin DM, Stjernsward J, Muir CS. 1984. Estimates of the worldwide frequency of twelve major cancers. *Bull World Health Organ.* 62:163-182.
- Parsonnet J, Friedman GD, Vandersteen DP, Chang Y, Vogelman JH, Orentreich N, Sibley RK. 1991. Helicobacter pylori infection and the risk of gastric carcinoma. *The New England Journal of Medicine.* 325:1127-1131.
- Paulick MG, Bertozzi CR. 2008. The glycosylphosphatidylinositol anchor: A complex membrane-anchoring structure for proteins. *Biochemistry.* 47:6991-7000.
- Peck B, Ortkamp M, Diehl KD, Hundt E, Knapp B. 1999. Conservation, localization and expression of hopz, a protein involved in adhesion of helicobacter pylori. *Nucleic Acids Res.* 27:3325.
- Pectasides D, Mylonakis A, Kostopoulou M, Papadopoulou M, Triantafillis D, Varthalitis J, Dimitriades M, Athanassiou A. 1997. Cea, ca 19-9, and ca-50 in monitoring gastric carcinoma. *Am J Clin Oncol.* 20:348-353.
- Pendu J, Marionneau S, Cailleau-Thomas A, Rocher J, Moullac-Vaidye B, Clément M. 2001. Abh and lewis histo-blood group antigens in cancer. *APMIS.* 109:9-26.
- Pinho SS, Carvalho S, Marcos-Pinto R, Magalhaes A, Oliveira C, Gu J, Dinis-Ribeiro M, Carneiro F, Seruca R, Reis CA. 2013. Gastric cancer: Adding glycosylation to the equation. *Trends Mol Med.* 19:664-676.
- Pinho SS, Reis CA. 2015. Glycosylation in cancer: Mechanisms and clinical implications. *Nat Rev Cancer.* 15:540-555.
- Pohl N. 2010. The sugar code: Fundamentals of glycosciences. Edited by hans-joachim gabius. *ChemBiochem.* 11:1147-1147.
- Rajendran L, Simons K. 2005. Lipid rafts and membrane dynamics. *J Cell Sci.* 118:1099-1102.
- Ravn V, Dabelsteen E. 2000. Tissue distribution of histo-blood group antigens. Oxford, UK. p. 1-28.
- Reily C, Stewart TJ, Renfrow MB, Novak J. 2019. Glycosylation in health and disease. *Nat Rev Nephrol.* 15:346-366.
- Rombouts Y, Ewing E, van de Stadt LA, Selman MH, Trouw LA, Deelder AM, Huizinga TW, Wuhler M, van Schaardenburg D, Toes RE et al. 2015. Anti-citrullinated protein antibodies acquire a pro-inflammatory fc glycosylation phenotype prior to the onset of rheumatoid arthritis. *Ann Rheum Dis.* 74:234-241.
- Rossez Y, Gosset P, Boneca IG, Magalhaes A, Ecobichon C, Reis CA, Cieniewski-Bernard C, Joncquel Chevalier Curt M, Leonard R, Maes E et al. 2014. The ladinac-specific adhesin laba mediates adhesion of helicobacter pylori to human gastric mucosa. *J Infect Dis.* 210:1286-1295.
- Röttger S, White J, Wandall HH, Olivo JC, Stark A, Bennett EP, Whitehouse C, Berger EG, Clausen H, Nilsson T. 1998. Localization of three human polypeptide galnac-transferases in hela cells suggests initiation of o-linked glycosylation throughout the golgi apparatus. *J Cell Sci.* 111 (Pt 1):45-60.

- Sachs G, Weeks DL, Melchers K, Scott DR. 2003. The gastric biology of helicobacter pylori. *Annu Rev Physiol.* 65:349-369.
- Sander JD, Joung JK. 2014. Crispr-cas systems for editing, regulating and targeting genomes. *Nat Biotechnol.* 32:347-355.
- Sato T, Gotoh M, Kiyohara K, Kameyama A, Kubota T, Kikuchi N, Ishizuka Y, Iwasaki H, Togayachi A, Kudo T et al. 2003. Molecular cloning and characterization of a novel human beta 1,4-n-acetylgalactosaminyltransferase, beta 4galnac-t3, responsible for the synthesis of n,n'-diacetyllactosediamine, galnac beta 1-4glcnac. *J Biol Chem.* 278:47534-47544.
- Sethuraman N, Stadheim TA. 2006. Challenges in therapeutic glycoprotein production. *Curr Opin Biotechnol.* 17:341-346.
- Sheehan JK, Kirkham S, Howard M, Woodman P, Kutay S, Brazeau C, Buckley J, Thornton DJ. 2004. Identification of molecular intermediates in the assembly pathway of the muc5ac mucin. *J Biol Chem.* 279:15698-15705.
- Shen R, Raska M, Bimczok D, Novak J, Smith PD. 2014. Hiv-1 envelope glycan moieties modulate hiv-1 transmission. *J Virol.* 88:14258-14267.
- Shimizu F, Nakayama J, Ishizone S, Zhang MX, Kawakubo M, Ota H, Sugiyama A, Kawasaki S, Fukuda M, Katsuyama T. 2003. Usefulness of the real-time reverse transcription-polymerase chain reaction assay targeted to alpha1,4-n-acetylglucosaminyltransferase for the detection of gastric cancer. *Lab Invest.* 83:187-197.
- Silva G, Poirot L, Galetto R, Smith J, Montoya G, Duchateau P, Paques F. 2011. Meganucleases and other tools for targeted genome engineering: Perspectives and challenges for gene therapy. *Curr Gene Ther.* 11:11-27.
- Skoog EC, Sjolting A, Navabi N, Holgersson J, Lundin SB, Linden SK. 2012. Human gastric mucins differently regulate helicobacter pylori proliferation, gene expression and interactions with host cells. *PLoS One.* 7:e36378.
- Sokic-Milutinovic A, Alempijevic T, Milosavljevic T. 2015. Role of helicobacter pylori infection in gastric carcinogenesis: Current knowledge and future directions. *World J Gastroenterol.* 21:11654-11672.
- Spiro RG. 2002. Protein glycosylation: Nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology.* 12:43R-56R.
- Springer SA, Gagneux P. 2013. Glycan evolution in response to collaboration, conflict, and constraint. *J Biol Chem.* 288:6904-6911.
- Steen Pvd, Rudd PM, Dwek RA, Opendakker G. 1998. Concepts and principles of o-linked glycosylation. *Crit Rev Biochem Mol Biol.* 33:151-208.
- Steenftoft C, Bennett EP, Schjoldager KT, Vakhrushev SY, Wandall HH, Clausen H. 2014. Precision genome editing: A small revolution for glycobiology. *Glycobiology.* 24:663-680.
- Stefano C, Christina L, Zhaoying X, Jean EC, Paolo G, Mark B, Rino R, Antonello C. 1996. Cag, a pathogenicity island of helicobacter pylori, encodes type i-specific and disease-associated virulence factors. *Proc Natl Acad Sci U S A.* 93:14648.
- Stroh LJ, Stehle T. 2014. Glycan engagement by viruses: Receptor switches and specificity. *PALO ALTO: ANNUAL REVIEWS.* p. 285-306.
- Szymendera JJ. 1986. Clinical usefulness of three monoclonal antibody-defined tumor markers: Ca 19-9, ca 50, and ca 125. *Tumour Biol.* 7:333-342.
- Takeuchi H, Kantharia J, Sethi MK, Bakker H, Haltiwanger RS. 2012. Site-specific o-glycosylation of the epidermal growth factor-like (egf) repeats of notch: Efficiency of glycosylation is affected by proper folding and amino acid sequence of individual egf repeats. *The Journal of biological chemistry.* 287:33934.

- Taylor-Papadimitriou J, Burchell J, Miles DW, Dalziel M. 1999. Muc1 and cancer. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*. 1455:301-313.
- Teneberg S, Miller-Podraza H, Lampert HC, Evans DJ, Evans DG, Danielsson D, Karlsson KA. 1997. Carbohydrate binding specificity of the neutrophil-activating protein of helicobacter pylori. *The Journal of biological chemistry*. 272:19067.
- Tuccillo FM, de Laurentis A, Palmieri C, Fiume G, Bonelli P, Borrelli A, Tassone P, Scala I, Buonaguro FM, Quinto I et al. 2014. Aberrant glycosylation as biomarker for cancer: Focus on cd43. *Biomed Res Int*. 2014:742831.
- van den Berg TK, Honing H, Franke N, van Remoortere A, Schiphorst WECM, Liu F-T, Deelder AM, Cummings RD, Hokke CH, van Die I. 2004. Lactinac-glycans constitute a parasite pattern for galectin-3-mediated immune recognition. *The Journal of Immunology*. 173:1902-1907.
- van Die I, Cummings RD. 2009. Glycan gimmickry by parasitic helminths: A strategy for modulating the host immune response? *Glycobiology*. 20:2-12.
- Wang C, Eufemi M, Turano C, Giartosio A. 1996. Influence of the carbohydrate moiety on the stability of glycoproteins. *Biochemistry*. 35:7299.
- Varki A. 1993. Biological roles of oligosaccharides: All of the theories are correct. *Glycobiology*. 3:97-130.
- Varki A. 2017. Biological roles of glycans. *Glycobiology*. 27:3-49.
- Varki A CR, Esko JD, et al., . 2009. Essentials of glycobiology Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.
- Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Marth JD, Bertozzi CR, Hart GW, Etzler ME. 2009. Symbol nomenclature for glycan representation. *PROTEOMICS*. 9:5398-5399.
- Wasik BR, Barnard KN, Parrish CR. 2016. Effects of sialic acid modifications on virus binding and infection. *Trends Microbiol*. 24:991-1001.
- Wood C, Kabat EA, Ebisu S, Goldstein IJ. 1978. An immunochemical study of the combining sites of the second lectin isolated from *bandeiraea simplicifolia* (bs ii). *Ann Immunol (Paris)*. 129:143.
- Wright JN, Collins HE, Wende AR, Chatham JC. 2017. O-glcNacylation and cardiovascular disease. *Biochem Soc Trans*. 45:545-553.
- Wurm FM. 2004. Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat Biotechnol*. 22:1393-1398.
- Yamaguchi H, Belk JG, Zhang XM, Sudijono JL, Fahy MR, Jones TS, Joyce BA. 1997. Rheed and stm study of the two-dimensional growth of inas on gaas (111)a. *Microelectronics Journal*. 28:825-831.
- Yang L, Biswas ME, Chen P. 2003. Study of binding between protein a and immunoglobulin g using a surface tension probe. *Biophys J*. 84:509-522.
- Yang X, Qian K. 2017. Protein o-glcNacylation: Emerging mechanisms and functions. *Nature reviews Molecular cell biology*. 18:452-465.
- Yu X, Blanchard H. 2014. Carbohydrate recognition by rotaviruses. *Journal of structural and functional genomics*. 15:101-106.
- Zhang MX, Nakayama J, Hidaka E, Kubota S, Yan J, Ota H, Fukuda M. 2001. Immunohistochemical demonstration of alpha1,4-n-acetylglucosaminyltransferase that forms glcnacalpha1,4galbeta residues in human gastrointestinal mucosa. *J Histochem Cytochem*. 49:587-596.
- Zheng Z, Jia Y, Hou L, Persson C, Yeager M, Lissowska J, Chanock SJ, Blaser M, Chow WH, Ye W. 2009. Genetic variation in a4gnt in relation to helicobacter pylori serology and gastric cancer risk. *Helicobacter*. 14:120-125.

