University of Alberta

Characterization of the bacterial oligosaccharyltransferase for *N*-linked protein glycosylation

by

Yasmin Barre

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science
in
Microbiology and Biotechnology

Department of Biological Sciences

Spring 2012
Edmonton, Alberta

Your file Votre référence
ISBN: 978-0-494-87862-0

Our file Notre référence
ISBN: 978-0-494-87862-0

For my parents, Osman and Faduma

Abstract

In *Campylobacter jejuni*, the oligosaccharyltransferase, PglB, transfers a heptasaccharide from a lipid donor to asparagine within the D/E-X-N-X-S/T consensus sequon (X≠P) or releases this glycan directly into the periplasm. Using a bioinformatic approach, we identified a conserved DXXK motif within *C. jejuni* PglB and generated point mutants at the D and K sites. We demonstrated that the DXXK motif plays an important structural role for *N*-glycosylation and free-oligosaccharide (fOS) production. Further downstream of the DXXK sequence, we analyzed the PglB enzymes from *C. jejuni* and *C. lari* and showed that a single amino acid substitution improved *C. lari* PglB activity. We also explored the possible autoregulatory mechanism of PglB by generating a point mutation within the *N*-glycosylation consensus sequence within the soluble domain of PglB. We demonstrated that the asparagine mutation does not influence protein glycosylation, but shows reduced fOS production, suggesting that PglB auto-glycosylation may regulate its hydrolase activity.

Acknowledgments

First and foremost, I wish to express my deepest gratitude to my supervisor, Dr. Christine Szymanski, who has supported me throughout my thesis and has offered me invaluable advice and guidance from the very early stage of my research. I want to thank her for providing me encouragement and immense support over the years, which has allowed me to grow as a student, a researcher, and a scientist. Her passion for science and her perpetual energy and enthusiasm in research inspires me. Thank you for making my experience in the lab very rewarding.

I especially want to thank Dr. Harald Nothaft in the lab for all of his guidance and expertise. He has played a crucial role in my research and skills development. I want to sincerely thank him for his patience, advice, and support from the preliminary to the concluding level of my research. I would also like to thank all of the members of the Szymanski group. Special thanks to Bernadette Beadle for being very accommodating and making the operations in the lab run very smoothly. We had good times in the lab! I also would not have had all the fun I did in the lab without my partner-in-crime, Rajinder Dubb! Thank you for all the laughter.

I would also like to thank Dr. Mario Feldman without whose knowledge and assistance in my research would not have been successful. He has made his support available to me in a number of ways and has especially taught me the importance of controls, controls, controls! Thanks for making science fun. Muchas gracias por todo!

It is also my pleasure to thank Dr. John Klassen for his insight and helpful suggestions during my committee meetings. In addition, I would like to thank Veronica Ielmini who has helped me tremendously in the final stages of my research with optimizing experiments. And last but certainly not least, thank you to my beautiful family and friends who have stood by my side through this journey. I could not have done this without you!

TABLE OF CONTENTS

Heading **Page Number**

Chapter 1. General Introduction.....1

1.1. Overview of protein glycosylation.....2

1.2. *N*-linked protein glycosylation in Eukaryotes.....4

1.2.1. The dolichol pathway.....4

1.2.2. Eukaryotic OTase.....7

1.2.3. OTase substrate and acceptor specificity.....9

1.3. *N*-linked protein glycosylation in Archaea.....10

1.3.1. The archaeal pathway.....11

1.3.2. Archaeal OTase, AglB.....13

1.3.3. AglB substrate and acceptor specificity.....15

1.4. Protein glycosylation in Bacteria.....16

1.4.1. Discovery of the *N*-glycosylation pathway.....17

1.4.2. The *C. jejuni* PglB pathway.....18

1.4.3. Importance of the *C. jejuni* Pgl pathway.....22

1.4.4. Functional characteristics of *C. jejuni* PglB.....23

1.4.5. Structural biology of *C. jejuni* PglB.....26

1.4.6. *C. jejuni* PglB substrate specificity.....30

1.4.7. *C. jejuni* PglB acceptor specificity.....33

1.4.8. Exploitation of PglB for glycoprotein engineering.....36

1.4.9. Putative Pgl pathways in proteobacteria.....39

1.4.10. Comparison of PglB to other OTases.....41

1.5. Conclusions and thesis objectives.....47

1.5.1. Concluding remarks.....47

1.5.2. Thesis objectives.....48

1.6. Literature cited.....50

Chapter 2. Identification of important amino acid residues required for the function of the *Campylobacter* oligosaccharyltransferase for *N*-linked protein glycosylation.....62

2.1 Introduction.....63

2.2. Materials and Methods.....67

2.2.1. Bacterial strains, plasmids, and growth conditions.....67

2.2.2. Construction of plasmids.....67

2.2.3. Expression of PglB point mutants in *E. coli* and purification of AcrA.....72

2.2.4. Preparation of *C. jejuni* whole cell lysates.....73

2.2.5. Analysis of fOS production.....73

2.2.6. Immunodetection by Western blotting.....74

2.3. Results.....74

2.3.1. Site-directed mutagenesis and expression of PglB enzymes.....74

2.3.2. Investigation of the DXXK motif in PglB expressed in *E. coli*.....75

2.3.3. Activity of PglB DXXK point mutants in *C. jejuni* and analysis of fOS production.....77

2.3.4. Analysis of the *C. jejuni* and *C. lari* PglB proline / serine point mutants.....79

2.4. Discussion.....80

2.5. Acknowledgements.....88

2.7. Literature cited.....89

Chapter 3. Autoregulation of the bacterial oligosaccharyltransferase.....93

3.1. Introduction.....94

3.2. Materials and Methods.....97

3.2.1. Bacterial strains, plasmids, and growth conditions.....97

3.2.2. Construction of plasmds.....98

3.2.3. Expression of the PglB point mutant in *E. coli* and purification of AcrA.....101

3.2.4. Complementation of $PglB_{N534Q}$ in *C. jejuni pglB* mutant.....102

3.2.5. Semi-quantitative mass spectrometry (sq-MS-based fOS analysis).....103

3.2.6. PglB overexpression and purification.....103

3.2.7. Purification of $AcrA_{His6}$.....104

3.2.8. Preparation of LLOs.....105

3.2.9. *In vitro* PglB activity assay.....106

3.2.10. Western blot analysis.....106

3.3. Results.....106

3.3.1. Site-directed mutagenesis and expression of $PglB_{N534Q}$.....106

3.3.2. $PglB_{N534Q}$ activity in *E. coli* and *C. jejuni*.....107

3.3.3. Analysis of fOS production in *C. jejuni*.....109

3.3.4. *In vivo* PglB self-glycosylation.....110

3.3.5. *In vitro* protein glycosylation.....113

3.4. Discussion.....114

3.5. Acknowledgements.....120

3.6. Literature cited.....121

Chapter 4. General Discussion and Conclusions.....126

4.1. Research purpose.....127

4.2. Summary and future directions.....127

4.3. Conclusions and perspectives.....134

4.4. Literature cited.....136

Appendix.....137

A.1. Development of an ELISA-based quantitation assay for protein glycosylation.....138

A.1.1. Introduction.....138

A.1.2. Methods.....138

ELISA.....139

SDS-PAGE and Western blotting.....141

A.1.3. Results and Discussion.....141

A.1.4. Literature cited.....146

A.2. Antibody generation against the soluble domain of *C. jejuni* PglB.....147

A.2.1. Introduction.....147

A.2.2. Methods.....148

Construction of plasmid.....148

Expression and purification of the PglB soluble domain.....148

Rabbit immunization.....150

Immunodetection by Western blotting.....150

Expression of full-length $PglB_{His6}$.....151

A.2.3. Results and Discussion.....151

A.2.4. Literature cited.....158

LIST OF TABLES

Table 2.1. Bacterial strains and plasmids used in this study.....68

Table 2.2. Oligonucleotides used in this study.....69

Table 3.1. Bacterial strains and plasmids used in this study.....99

Table A-1.1. Bacterial strains and plasmids used in this study.....139

Table A-2.1. Bacterial strains and plasmids used in this study.....149

LIST OF FIGURES

Figure 1.1. *N*-linked protein glycosylation pathway in Eukaryotes.....6

Figure 1.2. Two characterized *N*-linked protein glycosylation pathways in Archaea.....13

Figure 1.3. *N*-linked protein glycsosylation pathway in Bacteria.....21

Figure 1.4. Phylogenetic tree of PglB protein sequences.....40

Figure 2.1. Multiple sequences alignment of *Campylobacter* PglB protein sequences.....75

Figure 2.2. Western blot analysis of DXXK mutants in *E. coli*.....76

Figure 2.3. Complementation of *C. jejuni pglB* knock-out with the DXXK *pglB* point mutants.....77

Figure 2.4. The amount of fOS release in *C. jejuni* was determined using semi-quantitative mass spectrometry.....78

Figure 2.5. Close-up view of the conserved WWDYG motif in *C. jejuni* PglB...79

Figure 2.6. Amino acid exchanges of proline and serine residues were made to examine the importance of these sites in PglB.....83

Figure 2.7. Close-up view of the *C. lari* PglB catalytic site.....86

Figure 3.1. The crystal structure of the *C. jejuni* PglB soluble domain.....97

Figure 3.2. Western blot analysis of the *pglB* N534Q mutant in *E. coli* and *C. jejuni*.....108

Figure 3.3. The role of the ^{532}DYNQS^{536} consensus sequence in *C. jejuni* PglB fOS production.....109

Figure 3.4. Purification of His6-tagged *C. jejuni* PglB by Ni-NTA affinitiy chromatography.....111

Figure 3.5. Glycosylation of *C. jejuni* PglB.....112

Figure 3.6. PglB activity assay.....113

Figure A-1.1. Whole-cell lysates prepared from *E. coli* cells were applied to the nickel-coated ELISA-plate to determine the relative levels of AcrA glycosylation.....143

Figure A-1.2. Electrophoretic mobility shift assay and immunodetection of *E. coli* periplasmic extracts.....145

Figure A-2.1. Purification of His6-tagged *C. jejuni* PglB soluble domain (SD) by Ni-NTA affinity chromatography.....152

Figure A-2.2. Western blot analysis of the polyclonal rabbit-anti-PglB(SD) antibodies for use against *C. jejuni* whole-cell extracts.....153

Figure A-2.3. Immunodetection of full-length $PglB_{His6}$ with His-specific and PglB(SD)-specific antisera.....155

LIST OF SYMBOLS, NOMENCLATURE, AND ABBREVIATIONS

Bac: bacillosamine (2,4-diacetamido-2,4,6-trideoxyglucopyranose)

Dol-P: dolichol-phosphate

Dol-PP: dolichol-pyrophosphate

fOS: free oligosaccharides

GalNAc: *N*-acetylgalactosamine

Glc: glucose

GlcNAc: *N*-acetylglucosamine

GTase: glycosyltransferase

Hex: hexose

HexNAc: *N*-acetylhexosamine

IPTG: isopropyl-β-D-thiogalactoside

LLO: lipid-linked oligosaccharide

MS: mass spectrometry

NMR: nuclear magnetic resonance

OTase: oligosaccharyltransferase

Pgl: protein glycosylation

SBA: soybean-agglutinin

SDS-PAGE: sodium dodecyl sulfate polyacrylamide gel electrophoresis

Stt3: staurosporine and temperature sensitivity 3 protein

STT3: staurosporine and temperature sensitivity 3 gene

Und-PP: undecaprenyl-pyrophosphate

CHAPTER 1

General Introduction

1.1. Overview of protein glycosylation

Glycosylation is a common modification of many proteins. Glycans play important structural and functional roles for cellular proteins including: protection from proteolysis, evasion from antigenic recognition, recruitment of molecular chaperones and sorting receptors, proper folding, as well as protein stability and localization (Hebert *et al.*, 2005). Oligosaccharides are also crucial for the development, growth or survival of an organism (Varki, 1993). Various modifications occur in proteins that may be secreted by cells or remain in the cellular membrane or cytoplasm. Protein glycosylation involves the covalent attachment of a carbohydrate unit to a specific amino acid residue. The central step in the process is the enzymatic formation of the sugar-amino acid bond. A range of glycosylation types have been identified in several organisms that consist of *N*-glycosylation, *O*-glycosylation, *C*-mannosylation, phospho-glycosylation, and glypiation (Spiro, 2002). It has been predicted that more than half of all proteins in nature are glycosylated, either *O*- or *N*-linked (Apweiler *et al.*, 1999).

In eukaryotic cells, *O*-glycosylation normally starts in the Golgi apparatus and involves the transfer of a variety of sugars at the reducing end, such as *N*-acetylglucosamine (GlcNAc), *N*-acetylgalactosamine (GalNAc), glucose (Glc), mannose (Man), fucose, xylose, galactose, and arabinose, to the hydroxyl group of a serine (S) or threonine (T) amino acid to which further elongation and modification occurs (Steen *et al.*, 1998). In prokaryotes, this process occurs in the cytoplasm or inner membrane, where nucleotide-activated monosaccharides are attached to S or T residues as well (Szymanski & Wren, 2005). *O*-glycans are

added to proteins either through sequential addition or *en bloc* transfer of oligosaccharides (Feldman, 2009). *N*-linked protein glycosylation involves the transfer of a glycan chain with a terminal GlcNAc residue at the reducing end to the amino group of an asparagine (N) residue (Spiro, 2002). This process has been most extensively studied in eukaryotes, particularly in the yeast *Saccharomyces cerevisiae* (Kornfeld & Kornfeld, 1985; Burda & Aebi, 1999; Kelleher & Gilmore, 2006). *N*-glycosylation was first observed to exist outside of the eukaryotic domain in the archaeon *Halobacterium salinarium* (Mescher & Strominger, 1976). Approximately one decade ago, a system of general *N*-glycosylation was discovered in the Gram-negative bacterium, *Campylobacter jejuni* (Szymanski *et al.*, 1999). *N*-glycoproteins that have been characterized to date are synthesized by *en bloc* transfer of an oligosaccharide assembled on a polyisoprenyl-pyrophosphate carrier to an acceptor protein (Weerapana & Imperiali, 2006), with the exception of *Haemophilus influenzae* where the proteins are modified by the sequencial addition of sugars (Gross *et al.*, 2008).

Altered glycosylation patterns in humans have been observed to be associated with several diseases. Congenital disorders of glycosylation (CDGs) are inherited disorders involving the central and peripheral nervous system caused by defects in the attachment of *N*-linked carbohydrates to protein (Bhatia & Mukhopadhyay, 1999). CDGs result from enzyme or transport defects in the endoplasmic reticulum, Golgi compartment, or cytoplasm (Grunewald *et al.*, 2002). These diseases do not result only from the absence of glycosylation, but altered glycan structures can also cause CDGs with a similar effect (Freeze &

Aebi, 2005). Mutations in glycosyltransferase genes that are responsible for further processing of glycan structures have also been shown to result in disorders of glycosylation (Dennis *et al.*, 1999). *O*-linked glycans, such as *O*-mannosyl, are abundant in the mammalian brain. Mutations in genes involved in the modification of brain glycoproteins have been associated with neuromuscular diseases, such as congenital muscular dystrophy and brain malformations (Willer *et al.*, 2003). Therefore, glycosylation is an important modification essential to a variety of biological processes in eukaryotes.

Based on the high degree of conservation of the *N*-linked protein glycosylation systems of Eukarya, Archaea, and Bacteria, it is assumed that the mechanism of *N*-glycosylation may be similar. The defining event in the process is the catalytic transfer of an oligosaccharide onto a protein acceptor catalyzed by the key oligosaccharyltransferase (OTase). Herein, the OTase in all three domains of life will be discussed and its functional and structural characteristics evaluated. I will also discuss the similarities and differences of the *N*-glycosylation machinery with regards to substrate and acceptor specificity. Furthermore, information will be presented regarding the use of OTases for biotechnological applications in the development of glycan-based therapeutics.

1.2. *N*-linked protein glycosylation in Eukaryotes

1.2.1. The dolichol pathway

The *N*-linked protein glycosylation pathway is a co-translational process that functions on the lumenal membrane of the endoplasmic reticulum (ER) (Kornfeld & Kornfeld, 1985). A tetradecasaccharide, $Glc_3Man_9GlcNAc_2$, is

ultimately transferred to the asparagine side chains of nascent polypeptides. First, the oligosaccharide is assembled on a polyisoprenyl lipid carrier, dolichol-phosphate (Dol-P), through a stepwise addition of monosaccharides catalyzed by glycosyltransferases (Kukuruzinska & Lennon, 1998). *N*-linked protein glycosylation relies significantly on the availability of Dol-P for lipid intermediate synthesis (Rush *et al.*, 2008).

The pathway begins with synthesis of the dolichol-linked oligosaccharide at the cytoplasmic side of the ER membrane, which requires three different nucleotide-activated sugar donors: guanidine diphosphate (GDP)-Man, uridine diphosphate (UDP)-GlcNAc, and UDP-Glc (Burda & Aebi, 1999). The first step involves an *N*-acetylglucosaminyl phosphate transferase that transfers a GlcNAc-P from UDP-GlcNAc onto Dol-P (Barnes *et al.*, 1984), which is then followed by the transfer of a second GlcNAc residue to form $GlcNAc_2$-PP-Dol. Subsequently, a β-1,4-mannosyltransferase adds the first mannosyl residue to the lipid-linked disaccharide, followed by four additional mannosyltransferases that are required for the synthesis of $Man_5GlcNAc_2$-PP-Dol (Burda & Aebi, 1999). The partially assembled heptasaccharide is flipped to the lumenal side of the ER membrane by an ATP-independent flippase (Helenius *et al.*, 2002), where additional modification of the glycan occurs (Figure 1.1).

The late steps of the assembly of the lipid-linked oligosaccharide (LLO) require Dol-P-activated saccharides as sugar donors, Dol-P-Man and Dol-P-Glc (Weerapana & Imperiali, 2006). The addition of four mannose and three glucose residues complete the tetradecasaccharide ($Glc_3Man_9GlcNAc_2$-PP-Dol), which is

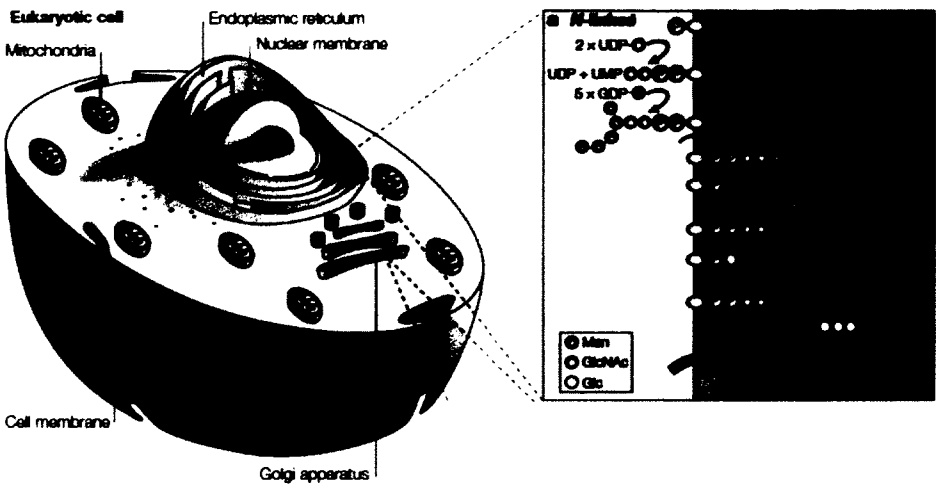

Figure 1.1. *N*-linked protein glycosylation pathway in Eukaryotes. The process begins with the synthesis of dolichol-linked oligosaccharides in the cytoplasmic face of the lumenal membrane of the endoplasmic reticulum. After translocation into the lumen, the oligosaccharide is further processed to a complete tetradecasaccharide where it then gets transferred to asparagine residues within the N-X-S/T consensus sequence of protein acceptors by the oligosaccharyltransferase, OTase complex. Modified from Szymanski & Wren (2005).

transferred by an OTase from the lipid carrier to asparagine residues of protein acceptors found within the N-X-S/T consensus sequon, where X can be any amino acid except proline (Aebi *et al.*, 2009). The transfer of the glycan generates dolichol-pyrophosphate (Dol-PP) in the lumenal monolayer of the ER followed by cleavage to Dol-P by Dol-PP phosphatase, recycling the lipid carrier for the assembly of a new oligosaccharide (Rush *et al.*, 2008). Initial modification of

nascent polypeptides with the core oligosaccharide results in a homogeneous display of glycans. However, further processing of the *N*-linked glycans occurs in the ER, where the glycans undergo trimming, followed by a series of non-uniform modifications that occur in the medial stacks of the Golgi apparatus (Helenius & Aebi, 2001). Thus, a tremendous diversity of *N*-glycan structures is observed, particularly in higher level eukaryotes. It has been proposed that the heterogeneous mix of *N*-glycans that result from further modifications serve as signals in the quality control of protein folding (Helenius, 1994). However, the mechanism by which the signals are generated and interpreted is still unknown (Aebi *et al.*, 2009).

1.2.2. Eukaryotic OTase

In the yeast *Saccharomyces cerevisiae*, the OTase is a multimeric enzyme complex that consists of at least eight different transmembrane protein subunits, each of which is required for enzymatic activity (Yan & Lennarz, 1999). The protein subunits of the OTase include: Wbp1 (Te Heesen *et al.*, 1991), Swp1 (Te Heesen *et al.*, 1993), Ost1 (Pathak *et al.*, 1995), Ost2 (Silberstein *et al.*, 1995), Ost3 (Karaoglu *et al.*, 1995), Ost4 (Chi *et al.*, 1996), Ost5 (Reiss *et al.*, 1997), Stt3 (Zufferey *et al.*, 1995), and Ost6 (Roos *et al.*, 1997). The Ost3 and Ost6 proteins have redundant functions and, therefore, can complement each other for OTase activity (Knauer & Lehle, 1999). Among all the protein subunits in the OTase required for oligosaccharide transfer, Stt3 (staurosporine and temperature sensitivity 3) is the most conserved, displaying more than 50% identity in amino acid sequence among eukaryotes (Spirig *et al.*, 1997).

Stt3 is a 78-kDa protein with a hydrophobic N-terminal domain spanning 11-13 transmembrane domains and a soluble, lumenal domain at the C-terminus (Spirig *et al.*, 1997). Through site-directed mutagenesis combined with photoactivated-crosslinking experiments, Yan and Lennarz (2002) demonstrated that the Stt3 subunit of the yeast OTase is the catalytic center. Block mutations that resulted in a lethal phenotype revealed a ^{516}WWDXG^{520} sequence that shares the highest sequence similarity among eukaryotic Stt3 proteins (Yan & Lennarz, 2002). Among the five residues in the identified motif, only W517 and D518 caused lethality in the yeast cells. The authors concluded that Stt3 is involved in peptide recognition and/or catalysis and that the ^{516}WWDXG^{520} motif plays a central role in the OTase activity. In addition, a new motif, ^{583}DXXK^{586}, was recently identified within the *STT3* sequence in yeast (Igura *et al.*, 2008). Site-directed mutagenesis of this short sequence revealed the essential role of the motif in catalysis. Igura and colleagues (2008) replaced the D and K residues with A and found that these mutants displayed a lethal phenotype, demonstrating that the newly identified motif is also necessary for OTase transfer activity. The authors propose that the DXXK motif is the binding site for the pyrophosphate group of the lipid-linked glycan, through a transiently bound cation.

In lower eukaryotes, such as *Giardia* and kinetoplastids, the OTase is composed of a single membrane protein that consists of Stt3 alone (Kelleher & Gilmore, 2006). In contrast, four subunit OTases that contain genes homologous to those in yeast encoding Stt3, Ost1, Ost2, and Wbp1 are present in diplomonads, entamoebas, and apicomplexan species (Nasab *et al.*, 2008). The genomes of

Trypanosoma brucei and *Leishmania major* reveal the presence of three and four *STT3* paralogous genes, respectively (Berriman *et al.*, 2005; Ivens *et al.*, 2005). The *L. major* OTase was determined to be composed of a single subunit, and three of its four Stt3 proteins were found to complement a yeast OTase mutant generated by deleting the *STT3* gene (Nasab *et al.*, 2008). The yeast Stt3 protein was also replaced by the *Trypanosoma cruzi* homolog and showed complementation for OTase activity (Castro *et al.*, 2006). However, the substitution of the whole OTase complex in yeast by homologous Stt3 proteins indicates that yeast Stt3 cannot function on its own, but requires interaction with the other subunits in the OTase complex for catalytic activity (Nasab *et al.*, 2008).

1.2.3. OTase substrate and acceptor specificity

The complete tetradecasaccharide linked to Dol-PP ($Glc_3Man_9GlcNAc_2$) is the preferred substrate donor in all eukaryotes, except in the trypanosomatid protozoa that transfer $Man_9GlcNAc_2$ and $Man_5GlcNAc_2$ (Silberstein & Gilmore, 1996; Izquierdo *et al.*, 2009). Studies with yeast OTase mutants demonstrated that incomplete Dol-PP-oligosaccharides could be transferred but with lower efficiency (Sharma *et al.*, 1981; Huffaker & Robbins, 1983). Typically, the specificity of the OTase for the glycan lies within the reducing end sugar (Kelleher & Gilmore, 2006).

N-glycoproteins are synthesized by the covalent attachment of oligosaccharides derived from lipid-linked glycans to asparagine residues of protein acceptors. Sequencing of eukaryotic *N*-glycoproteins showed that modification occurs at asparagine residues that are found within an N-X-S/T

consensus sequon, where X can be any amino acid except proline (Marshall, 1972). This glycoyslation site has been established as the recognition signal by the OTase for *N*-glycosylation activity (Bause, 1983). Although proline completely inhibits protein modification, four amino acid residues (Trp, Asp, Glu and Leu) were identified to be unfavourable in the X position of the consensus sequon (Kasturi *et al.*, 1997). Using a cell-free translation/glycosylation system, the authors examined the influence of local amino acid residues on glycosylation and found that *N*-glycan transfer is possible at glycosylation sites containing these residues, but OTase efficiency is considerably reduced.

Moreover, the secondary structure of the acceptor protein greatly influences the activity of the OTase. The process of *N*-glycosylation in eukaryotes occurs co-translationally. Protein translocation and folding is highly coupled to this process, as the acceptor site is presented to the OTase in a flexible form for glycan transfer before folding (Kowarik *et al.*, 2006a). Not all glycosylation sites present in nascent polypeptides are modified. The consensus site cannot be closer than 12-14 residues to the N-terminal transmembrane domain of a protein (Knauer & Lehle, 1999). In addition, acceptor sites close to the C-terminal domain of a protein are less effectively modified (Kukuruzinska & Lennon, 1998). Studies on the structural requirements of the OTase suggest that the acceptor protein may adopt an Asx-turn (Asx standing for Asn or Asp) at the recognition motif in order for efficient glycan transfer to occur (Imperiali *et al.*, 1992). The Asx-turn is a local conformation that is topologically similar to the β-turns typically found in peptide chains (Abbadi *et al.*, 1991).

N-linked glycosylation in eukaryotic systems is a multifaceted process involving many enzymes required to synthesize the glycans on a Dol-P lipid carrier, transfer the sugars onto protein acceptors, and modify the glycans further for proper folding and signaling. More studies are required to understand the mechanism of the transfer reaction catalyzed by the OTase. However, finding homologous pathways outside of the Eukaryotic domain has shed new light on this complicated process.

1.3. *N*-linked protein glycosylation in Archaea

1.3.1. The archaeal pathway

The first system of glycosylation that was found outside of the Eukaryotic domain was in the halophilic archaeon *Halobacterium salinarum*, which glycosylates its surface layer (S-layer) proteins (Mescher & Strominger, 1976). Much of the current knowledge of archaeal *N*- and *O*-glycosylated S-layer proteins builds from this initial work. The genes involved in S-layer protein glycosylation were putatively identified by Abu-Qarn and colleagues (2006) in the haloarchaeon *Haloferax volcanii*. Gene deletions revealed certain genes to be important for archaeal glycosylation, but loss of glycoprotein synthesis did not affect cell viability (Abu-Qarn & Eichler, 2006). Studies on the methanoarchaeon *Methanococcus voltae* revealed a unique *N*-glycan decorating several flagellin proteins in addition to the S-layer protein (Voisin *et al.*, 2005). The *M. voltae* glycan was structurally characterized using mass spectrometry (MS) in combination with nuclear magnetic resonance (NMR) analysis. The *N*-glycan structure was determined to be a trisaccharide composed of β-Man*p*NAcA6Thr-

(1–4)-β-Glc-*p*NAc3NAcA-(1–3)-β-Glc*p*NAc linked to asparagine (Voisin *et al.*, 2005).

The archaeal glycosylation (*agl*) genes encode enzymes involved in the assembly of the archaeal *N*-linked glycan and its eventual transfer onto protein acceptors in *H. volcanii* and *M. voltae* (Chaban *et al.*, 2006). Archaeal *N*-glycosylation is a post-translational process that occurs in the exterior face of the plasma membrane (Yurist-Doutsch *et al.*, 2008). Similar to Eukaryotes, nucleotide-activated sugars are assembled on the lipid carrier, Dol-P (Lechner *et al.*, 1985). Several archaeal enzymes involved in the biosynthesis of nucleotide-activated sugars functionally characterized in the archaeon *Methanococcus maripaludis* were found to convert a range of sugars (Namboori & Graham, 2008). Archaeal glycosyltransferases for oligosaccharide biosynthesis accept a broader range of monosaccharides with efficient turnover rates than their eukaryotic counterpart, which only use GlcNAc as the reducing end sugar (Lechner & Wieland, 1989; Mizanur *et al.*, 2004).

The pathway begins with the sequential addition of nucleotide-activated sugars to the Dol-P lipid carrier by glycosyltransferases on the cytoplasmic side of the membrane (Yurist-Doutsch *et al.*, 2008). The lipid-linked glycan is then flipped across the membrane to the exterior side of the cell through an unknown mechanism. Search of completed archaeal genomes and targeted deletions of possible candidates have failed to reveal the *agl* enzyme responsible for the translocation of the glycan (Abu-Qarn & Eichler, 2006; Chaban *et al.*, 2006). However, it is presumed to follow a similar method found in eukaryotes where the

translocation of the LLO results from the action of a 'flippase' enzyme. The final step in the pathway involves the transfer of the complete glycan to acceptor proteins by the archaeal OTase, AglB (Figure 1.2).

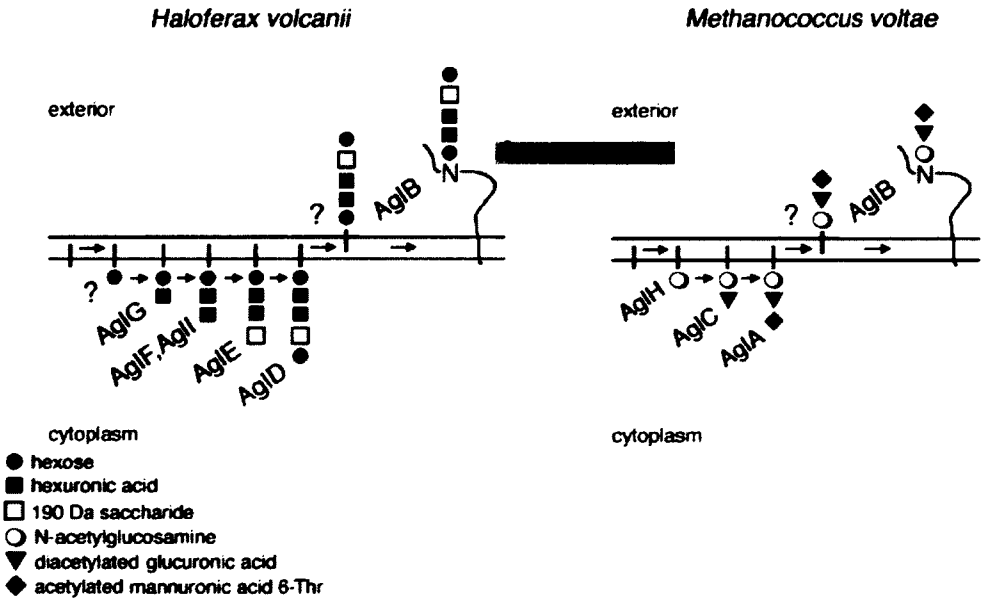

Figure 1.2. Two characterized *N*-linked protein glycosylation pathways in Archaea. Nucleotide-activated monosaccharides are assembled on dolichol phosphate in the cytoplasmic face of the membrane. The oligosaccharide is then translocated to the exterior face of the plasma membrane by an unknown mechanism. Transfer of the oligosaccharide is catalyzed by the oligosaccharyl-transferase, AglB, onto asparagine residues within the N-X-S/T consensus sequence of target proteins. Modified from Yurist-Doutsch *et al.* (2008).

1.3.2. Archaeal OTase, AglB

AglB is an Stt3 homologue comprised of a single subunit, which acts alone in mediating OTase activity. Similar to the OTase found in lower level

Eukarya, AglB is composed of an N-terminal transmembrane domain and a soluble domain at the C-terminal end (Maita *et al.*, 2010). Deletion of *H. volcanii* and *M. voltae aglB* genes resulted in loss of *N*-linked glycans from reporter glycoproteins (Chaban *et al.*, 2006; Abu-Qarn & Eichler, 2007). The OTase from the thermophilic archaeon *Pyrococcus furiosus* is also composed of the Stt3 protein alone and contains *STT3* paralogues in its genome. *In vitro* assays of purified *P. furiosus* AglB confirmed that the single subunit protein is capable of transferring glycans onto peptide acceptors (Kohda *et al.*, 2007).

The 2.7 Å resolution crystal structure of the C-terminal soluble domain of *P. furiosus* AglB was determined and provided the first structural insight into OTase activity (Igura *et al.*, 2008). A structure-guided sequence alignment that included the yeast STT3 and archaeal AglB sequences revealed a newly identified ^{571}DXXK^{574} motif in addition to the well-conserved ^{511}WWDXG^{515} sequence that was previously implicated in yeast OTase catalysis by Yan and Lennarz (2002). Since the soluble domain of the *P. furiosus* AglB is inactive in OTase activity, a mutagenesis study of the DXXK motif was carried out using the yeast Stt3 protein in yeast cells and demonstrated the essential role of the domain in catalysis (Igura *et al.*, 2008). Furthermore, a multiple sequence alignment using recently published OTase sequences revealed a third motif, DXD, which exists in the first extracellular loop of the transmembrane domain (Maita *et al.*, 2010). Maita and colleagues (2010) proposed that the WWDXG, DXXK and DXD motifs compose the three distinct catalytic domains of AglB, which coordinate together for function.

1.3.3. AglB substrate and acceptor specificity

A wide diversity of glycan structures have been found to be linked to asparagine residues of archaeal protein acceptors (Lechner & Wieland, 1989). Unlike Eukaryotes, the variability in the chemical structure of the glycan donor does not result from additional modifications after transfer onto protein, but rather from the original assembly of these sugars onto lipid carriers. For instance, the oligosaccharide in *H. volcanii* consists of a pentasaccharide comprising two hexoses, two hexuronic acids, and a 190 Da species yet to be characterized by MS (Abu-Qarn *et al.*, 2008); whereas, the glycoproteins in *P. furiosus* are modified with a heptasaccharide composed of two *N*-acetylhexosamines, two hexoses, one hexuronic acid, and two pentoses (Igura *et al.*, 2008). Nevertheless, the oligosaccharide must be lipid-linked for transfer to occur. And modification at asparagine residues has been found to rely on a Dol-P-linked oligosaccharide more frequently than a Dol-PP-linked oligosaccharide, which is involved in the eukaryotic system (Lechner *et al.*, 1985).

Analysis of archaeal glycoproteins has revealed that modification occurs at asparagine residues that are within the N-X-S/T (X $\neq$ proline) consensus sequon, also found in Eukarya (Abu-Qarn & Eichler, 2007). Interestingly, the consensus sequence of *Halobacterium halobium* is less stringent because it does not follow the typical glycosylation site known so far. Zeitler and colleagues (1998) exchanged the serine residue, a part of an N-A-S sequon, with various amino acids and found that glycosylation can still occur when serine is replaced with asparagine, leucine or valine. These results suggest that two different OTases exist

in *H. halobium*, one of which utilizes a different consensus sequon than the typical N-X-S/T used for all other eukaryotic and archaeal OTases described so far (Zeitler *et al.*, 1998). The factors mentioned above contribute to *N*-glycosylation in Archaea, since all of the sequons in the acceptor protein are not utilized for modification. Whether a particular protein conformation is favoured for *N*-glycosylation is yet to be determined.

The observation that *aglB* mutants in *H. volcanii* and *M. voltae* are still viable suggests that *N*-glycosylation is not required for archaeal cell survival, but possibly represents an adaptation to the extreme environments that archaeal microorganisms can inhabit (Chaban *et al.*, 2006; Yurist-Doutsch *et al.*, 2008). Similar to the misleading assumption that protein *N*-glycosylation was limited to eukaryotes, the notion that *N*-glycoproteins in prokaryotes were only present in Archaea also took hold. The lack of awareness about bacterial glycoprotein systems was due to the fact that the most frequently studied organisms, such as *Escherichia coli, Salmonella* sp., and *Bacillus subtilis*, did not produce *N*-glycosylated proteins (Messner, 2004). However, since the discovery of a protein *N*-glycosylation system in Bacteria, the field has moved at an extraordinary pace.

1.4. Protein glycosylation in Bacteria

It is currently well recognized that bacteria also glycosylate their proteins in a similar manner to Eukaryotes and Archaea. Several bacterial protein glycosylation pathways have been characterized thus far. Glycoproteins in *Mycobacterium tuberculosis* were identified in early studies as a result of their binding interaction with highly specific sugar-binding proteins called lectins

(Espitia & Mancilla, 1989). Other important pathogens, including *Streptococcus* and *Staphylococcus* species, synthesize glycoproteins that are involved in essential physiological processes (Chia *et al.*, 2001; Lee *et al.*, 2002). Glycosylation of flagellin and pilin, the major structural proteins in many bacterial species, has been extensively studied. Several genera, such as *Campylobacter*, *Helicobacter,* and *Listeria*, attach a single monosaccharide to the hydroxyl group of S or T residues in their flagellins (Doig *et al.*, 1996; Josenhans *et al.*, 2002; Logan, 2006); whereas in certain *Pseudomonas aeruginosa* strains and *Neisseria* spp., oligosaccharides are linked to the hydroxyl groups of Ser or Thr residues in the pilin (Castric, 1995; Stimson *et al.*, 1995). However, the sugars in these glycoproteins were all shown to be *O*-linked.

1.4.1. Discovery of the *N*-glycosylation pathway

A cluster of genes located on a 16 kb chromosomal DNA fragment in *Campylobacter jejuni* was originally reported to be involved in lipopolysaccharide (LPS) biosynthesis (Fry *et al.*, 1998). Subsequent work by Szymanski and colleagues (1999) demonstrated through site-specific mutations that the genes of this locus had no role in LPS biosynthesis, but rather were part of a system of general protein glycosylation in *Campylobacter*. This cluster was termed the protein glycosylation (*pgl*) gene locus and was found to be involved in the biosynthesis of a number of highly immunogenic glycoproteins (Szymanski *et al.*, 1999). *C. jejuni* glycoproteins were characterized using the soybean agglutinin (SBA) lectin known to bind terminal GalNAc residues (Linton *et al.*, 2002). MS analysis of the SBA-purified *C. jejuni* proteins led to the identification of the first

non-flagellin *Campylobacter* glycoproteins. In addition, MS/MS collision-induced dissociation of *C. jejuni* PEB3 and AcrA glycoproteins showed them to be modified with an *N*-linked glycan (Young *et al.*, 2002; Wacker *et al.*, 2002).

Using nano-NMR techniques, the structure of the glycan was determined to be GalNAc-α1,4-GalNAc-α1,4-[Glcβ1,3]GalNAc-α1,4-GalNAc-α1,4-GalNAc-α1,3-Bac-β1,*N*-Asn-Xaa, where Bac is bacillosamine (2,4-diacetamido-2,4,6-trideoxyglucopyranose) (Young *et al.*, 2002). Bac is a deoxy-sugar unique to bacteria (Maki & Renkonen, 2004). In *C. jejuni*, it serves as the linking saccharide at the reducing end (Wacker *et al.*, 2002; Young *et al.*, 2002), while the first sugar added to the lipid carrier is almost always GlcNAc in eukaryotes (Burda & Aebi, 1999) and a hexose or a GlcNAc in Archaea (Abu-Qarn *et al.*, 2008). Methods for studying *N*-glycans in bacteria are still limited because well-established techniques are predominantly based on the release of GlcNAc from protein. Peptide-*N*-glycosidase F is a widely used sugar-specific enzyme that cleaves between the innermost GlcNAc and asparagine residues from *N*-linked glycoproteins (Mussar *et al.*, 1989), a method not applicable to bacterial proteins.

1.4.2. The *C. jejuni* Pgl pathway

The bacterial process of *N*-linked protein glycosylation shares many features with its eukaryotic and archaeal counterparts, but significant differences also exist. *N*-linked protein glycosylation in *C. jejuni* is a post-translational modification that occurs in the periplasm. Although mutational studies in *C. jejuni* have demonstrated that the *pgl* genes are involved in *N*-glycosylation (Szymanski *et al.*, 1999; Linton *et al.*, 2002; Young *et al.*, 2002), bioinformatic analysis

suggested that the *pgl* gene locus encodes three sugar-modifying enzymes and five glycosyltransferases. This analysis was greatly facilitated by the fact that there are several genes present in the *N. meningitidis* cluster involved in pilin glycosylation that are homologous to those in the *C. jejuni* gene cluster (Power *et al.*, 2000; Weerapana & Imperiali, 2006). The functional transfer of the protein glycosylation pathway into *E. coli* was demonstrated for the first time by Wacker and colleagues (2002) through co-expression of the complete *C. jejuni pgl* locus with a protein acceptor in the heterologous host. The resulting production of recombinant glycoproteins suggested that the *pgl* locus contains all of the genes required for the biosynthesis of the LLO and its eventual transfer onto an acceptor protein.

Polyisoprenyl phosphates are the universal glycan lipid carriers and the assembly of glycans onto these lipid carriers is a highly conserved process in all cells for the biosynthesis of glycoproteins and cell surface polysaccharides (Burda & Aebi, 1999; Helenius *et al.*, 2002; Valvano, 2008). It has been postulated that the bacterial lipid carrier is undecaprenyl-phosphate (Und-P), a polyisoprene that contains 11 isoprene units that are fully unsaturated compared to the 15-20 saturated units of the corresponding eukaryotic and archaeal dolichols (Wacker *et al.*, 2002; Weerapana & Imperiali, 2006). Linton *et al.* (2005) provided experimental evidence for the involvement of the lipid carrier as a precursor for *N*-glycan assembly prior to transfer onto protein. The biosynthesis of the *C. jejuni* heptasaccharide on the lipid carrier was also confirmed through an *in vitro* biochemical analysis of the individual enzymes in the *pgl* locus, where chemically

synthesized undecaprenyl-pyrophosphate (Und-PP)-linked Bac was used to deduce the functions of putative glycosyltransferases (Glover *et al.*, 2005a).

The pathway begins in the cytoplasm with the conversion of UDP-GlcNAc to UDP-Bac by the activities of a dehydratase (PglF) and an aminotransferase (PglE) (Schoenhofen *et al.*, 2006), followed by an acetyltransferase (PglD) (Linton *et al.*, 2005) (Figure 1.3). PglC transfers Bac-P onto the lipid carrier on the cytoplasmic face of the inner membrane to produce Und-PP-Bac. Then PglA, PglH, and PglJ sequentially add a total of five GalNAc residues to produce the hexasaccharide followed by PglI, which transfers a single glucose branch to complete the heptasaccharide (Glover *et al.*, 2005a). The complete LLO is then translocated from the cytoplasmic side to the periplasmic side of the inner membrane by the ATP-binding cassette (ABC) transporter protein, PglK (Alaimo *et al.*, 2006). Next, the heptasaccharide is transferred from the lipid carrier to asparagine residues of protein acceptors solely by the OTase, PglB.

Young and colleagues (2002) identified 38 putative *C. jejuni* glycoproteins by means of SBA lectin affinity chromatography and two-dimensional gel proteomics analysis. Recently, over 150 *C. jejuni* proteins have been proposed to be modified through the Pgl pathway as predicted by the presence of glycosylation sites and signal peptides that target the proteins to the periplasm, where the glycosylation machinery exists (Nothaft *et al.*, 2010). Furthermore, 130 glycopeptides in *C. jejuni* were experimentally identified using several proteomic approaches, such as two-dimensional gel electrophoresis, SBA affinity and

zwitterionic hydrophilic interaction chromatography coupled to MS fragmentation techniques (Scott *et al.*, 2011). The authors also showed that, in *C. jejuni*, the hydrophobic integral membrane proteins are the predominant targets for *N*-glycosylation.

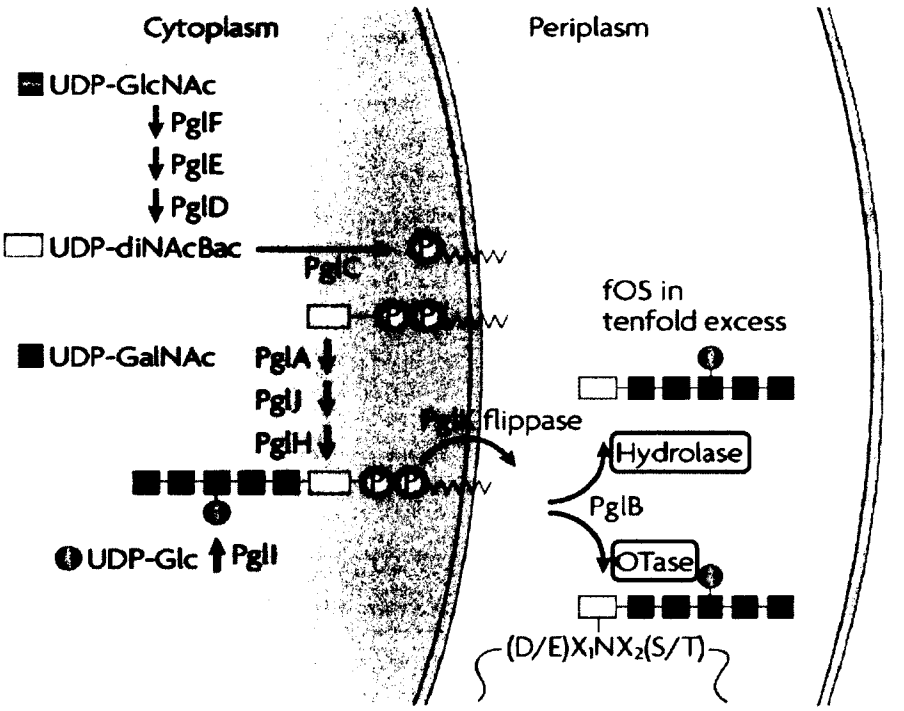

Figure 1.3. *N*-linked protein glycosylation pathway in Bacteria. In *C. jejuni*, glycosylation occurs post-translationally in the periplasm where nucleotide-activated sugars are assembled on undecaprenyl-pyrophosphate in the cytoplasmic side of the inner membrane. The heptasaccharide is then flipped into the periplasm and transferred by the oligosaccharyl-transferase (OTase), PglB, onto asparagine residues within the extended D/E-X-N-X-S/T consensus sequence. In addition to its OTase activity, PglB also exhibits hydrolase activity where it releases free oligosaccharides (fOS) into the periplasm in a fOS:*N*-glycan ratio of approximately 10:1. Modified from Nothaft & Szymanski, 2010.

1.4.3. Importance of the *C. jejuni* Pgl pathway

C. jejuni is a commensal organism naturally colonizing the gastrointestinal tracts of many birds and animals, but is pathogenic to humans, causing gastroenteritis. The glycan of *Campylobacter* has been demonstrated to play an important role in host immunity during infection (Szymanski *et al.*, 1999; Szymanski *et al.*, 2005). The glycosyl moieties on *C. jejuni* glycoproteins were found to be immunodominant, suggested by the dramatic reduction in reactivity with antisera upon loss of the carbohydrate components. Disruption of the *N*-glycosylation pathway does not affect the viability of the organism *in vitro*, but rather affects many cellular processes. *N*-linked glycosylation in *C. jejuni* is also important in the adherence and invasion of host epithelial cells and the colonization of intestinal tracts of mice (Szymanski *et al.*, 2002). Using *in vitro* cultured human intestinal epithelial cells, *pglB* and *pglE* mutants showed considerable decreases in adherence and invasion. Likewise, both *pglB* and *pglE* mutant strains that were tested in mice exhibited a significant reduction in colonization, whereas the strains adhered, invaded, and colonized at levels comparable to those of the wild type when complemented in *trans*. A *C. jejuni pglH* mutant was also shown to be severely affected in its ability to colonize chicks (Karlyshev *et al.*, 2004). Additional *C. jejuni pgl* genes were identified by signature-tagged transposon mutagenesis and shown to be involved in colonization of the chick gastrointestinal tract (Hendrixson & DiRita, 2004).

The *N*-linked glycan was also demonstrated to have a role in protein complex formation. In *C. jejuni* strain 81-176, the type IV secretion system was

shown to include a glycoprotein, VirB10, which contains two glycosylation sites (Larsen *et al.*, 2004). Modification at only one of the asparagine residues resulted in the loss of type IV secretion system assembly and competence. Since most identified glycoproteins are parts of protein complexes, it is proposed that glycosylation is involved in stabilization of these complexes. In addition, loss of *N*-glycosylation in *C. jejuni* was found to affect the expression of iron acquisition genes (Nothaft, H., personal communication). Conversely, iron was suggested to play a role in the regulation of protein glycosylation in *C. jejuni*, as demonstrated by the hyperglycosylation of proteins produced under iron-limited conditions (Palyada *et al.*, 2004). Crosstalk between the Pgl pathway and other carbohydrate pathways, such as *O*-linked glycosylation and capsular polysaccharide, exist in *C. jejuni.* Loss of PglF can affect *O*-glycosylation of flagellin proteins (Schoenhofen *et al.*, 2006), while disruption in the enzyme function of PglE, PglA and PglJ has been demonstrated to downregulate the expression of heptose biosynthesis genes for capsular polysaccharide (Nothaft, H., personal communication). In addition, several genes encoding epimerases involved in CPS heptose biosynthesis were downregulated in a *C. jejuni pglB* mutant background (Dwivedi, R., personal communication).

1.4.4. Functional characteristics of *C. jejuni* PglB

C. jejuni PglB is an 82-kDa single-subunit OTase that acts solely to transfer a conserved heptasaccharide to asparagine residues of acceptor proteins. Topology mapping of PglB predicts that the OTase comprises 11-13 hydrophobic N-terminal domains containing two large periplasmic loops within the

transmembrane region in addition to a soluble C-terminal domain oriented towards the periplasmic space (Li *et al.*, 2010; Jaffee & Imperiali, 2011). The *pglB* gene shares sequence similarity with the *STT3* gene that encodes the largest subunit of the eukaryotic OTase complex (Wacker *et al.*, 2002; Yan & Lennarz, 2002). The OTases of all three domains of life contain a conserved amino acid motif, WWDXG. PglB was proposed to fulfill the OTase function in *C. jejuni* since it contains this sequence within the soluble C-terminal domain of PglB (Wacker *et al.*, 2002). A *pglB* mutant constructed by antibiotic cassette mutagenesis demonstrated dramatic changes in immunoreactivity when analyzed by Western blotting using *C. jejuni* whole-cell extract antisera (Szymanski *et al.*, 1999; Young *et al.*, 2002; Szymanski *et al.*, 2005; Linton *et al.*, 2005). The *C. jejuni* PEB3 glycoprotein that was previously shown to be glycosylated was also shown by MS analysis to completely lack the *N*-glycan in this mutant (Young *et al.*, 2002). Loss of glycosylation of *C. jejuni* PEB3 and AcrA glycoproteins was also observed in a *pglB* mutant strain that was generated by point mutations within the ^{457}WWDYG^{462} motif of PglB (W458A, D459A), demonstrating the importance of this site for activity and confirming the essentiality of PglB as an OTase (Wacker *et al.*, 2002; Nita-Lazar *et al.*, 2005).

The archaeal and bacterial Stt3 orthologues show very limited sequence identity to the eukaryotic Stt3s. However, the highest similarity lies within the region extending 20 residues toward the N-terminus and 70 residues toward the C-terminus from the well-conserved WWDXG motif (Igura *et al.*, 2008). From this region, Igura and colleagues (2008) identified a short motif conserved in all

the OTases from the three kingdoms. The DXXK motif, first discovered within the C-terminal domain of the functionally equivalent archaeal AglB of *P. furiosus*, was shown to be essential for OTase activity in yeast as identified through the multiple sequence alignment of the eukaryotic and archaeal *STT3* sequences (Igura *et al.*, 2008). More recently, three appearances of the DXXK motif were found to be present within the soluble domain of *C. jejuni* PglB (Jaffee & Imperiali, 2011). Alanine mutations of the D and K residues within PglB showed comparable levels of activity to wildtype PglB when tested *in vitro*, which led the authors to conclude that the proposed DXXK motif is not important for OTase activity in *C. jejuni* (Jaffee & Imperiali, 2011).

In addition to *N*-glycosylation, PglB was shown to be necessary for the release of free oligosaccharides (fOS) in *C. jejuni* (Nothaft *et al.*, 2009). The authors demonstrated that fOS are produced in response to changes in the osmolarity of the environment and bacterial growth phase. In *C. jejuni*, the periplasmic fOS were detected by MS analysis (Liu *et al.*, 2006) and shown to be structurally identical to the *N*-linked heptasaccharide that is added to periplasmic and membrane proteins (Young *et al.*, 2002; Nothaft *et al.*, manuscript in preparation). Evidence showing that the conserved ^{457}WWDYG^{462} motif of PglB is necessary for the release of fOS was provided through an *in vitro* PglB activity assay, where PglB and LLOs were incubated in the presence or absence of osmolytes, such as NaCl, followed by semi-quantitative MS analysis (Nothaft *et al.*, 2009). Considerable decrease in fOS formation was observed in the presence of NaCl. Decrease in the release of fOS was also shown to be dependent on the

ionic strength of the medium, but independent of the presence of the *N*-glycan acceptor protein, AcrA. In addition, only full length heptasaccharides were detected *in vitro* indicating that PglB is highly selective for the release of the full length wild-type glycan (Nothaft *et al.*, 2009). The authors suggest that PglB demonstrates hydrolytic activity, whereby it cleaves the oligosaccharides from their lipid carriers and transfers them to water. In other Gram-negative Proteobacterial species, a common feature in response to low osmotic environments is the production of periplasmic glucans, which has been shown to maintain equilibrium in the periplasm (Bohin, 2000; Bohin & Lacroix, 2007). In eukaryotes, fOS production has been reported as a result of cleavage from Dol-PP-oligosaccharides (Kmiecik *et al.*, 1995) or from *N*-linked glycoproteins by peptide-*N*-glycosidase within the cytosol (Belard *et al.*, 1988). The mechanism of PglB cleaving the *N*-glycan directly from the glycoprotein is unfavoured as demonstrated through an *in vitro* assay where co-incubation of *C. jejuni* PglB and fully glycosylated AcrA did not result in fOS formation (Nothaft, H., personal communication). This result suggests that the lipid carrier not only serves as a membrane achor upon which nucleotide-activated sugars are assembled, but plays a crucial role in the function of PglB.

1.4.5. Structural biology of *C. jejuni* PglB

The 2.8 Å resolution crystal structure of the C-terminal soluble domain (SD) of *C. jejuni* PglB has been reported (Maita *et al.*, 2010). Based on a multiple sequence alignment between the OTases of yeast, *Pyrococcus,* and *Campylobacter* published by the same group (Igura *et al.*, 2008), Maita and

colleagues (2010) compared the crystal structure of the *C. jejuni* PglB(SD) with the previously determined AglB structure from *P. furious*. Initial sequence alignments of the archaeal and bacterial OTases showed that the DXXK motifs in both aligned (Igura *et al.*, 2008), but after solving the crystal structure, the authors realized that the alignments were dramatically incorrect. Although the structure comparison between AglB and PglB did not show the DXXK counterpart in PglB, it did however reveal a new catalytic motif in PglB, where ^{568}MXXI^{571} residues occupy the same positions (Maita *et al.*, 2010). A much larger sequence alignment of eukaryotic, archaeal and bacterial OTases revealed three types of conserved motifs. With reference to the crystal structures of AglB and PglB, the authors proposed that PglB consists of the previously studied ^{457}WWDYG^{461} motif, the ^{568}MXXI^{571} motif that resides on the kinked helix immediately below the signature sequence, and a highly conserved aspartic acid that exists in the first loop in the transmembrane region of the protein (D54 in *C.jejuni* PglB). Site-directed mutagenesis of the ^{568}MXXI^{571} motif and the D54 residue were carried out to address the role of these newly identified amino acids. PglB activity was tested using the full-length PglB protein expressed recombinantly in *E. coli* membrane fractions, since the C-terminal domain alone has no catalytic activity. Wild type PglB activity was observed in the alanine mutation of the M residue, but led to substantially reduced activity in the alanine mutation of the I residue. Alanine mutations of the D54 residue also resulted in nearly complete loss of PglB activity. Therefore, Maita and colleagues (2010) propose that PglB

comprises three distinct types of catalytic sequences that intimately coordinate with each other for function.

Recently, the 3.4 Å resolution crystal structure of the full-length active *Campylobacter lari* PglB has been solved (Lizak *et al.*, 2011). The PglB protein sequence of *C. lari* is 56% identical to that of *C. jejuni*. In accordance with the topological predictions made for *C. jejuni* PglB, the crystal structure of the *C. lari* PglB revealed 13 N-terminal transmembrane domains with two long external loops facing the periplasmic region, and a soluble C-terminal periplasmic domain (Lizak *et al.*, 2011). The C-terminal soluble domain of the *C. jejuni* PglB that has been crystallized was shown to have no activity (Maita *et al.*, 2010), but this can be explained by the crystal structure of the *C. lari* PglB where many of the transmembrane domains interact with the periplasmic region to form the binding and catalytic sites (Lizak *et al.*, 2011). Co-crystallization of the *C. lari* PglB with a short acceptor peptide revealed two cavities at opposite sides of PglB, the peptide-binding site and the LLO-binding site, which are connected by a channel that houses the acceptor asparagine (Lizak *et al.*, 2011).

For nearly a decade, the highly conserved WWDYG motif in PglB was regarded as the putative catalytic site (Wacker *et al.*, 2002). However, the *C. lari* PglB structure reveals that the WWDYG motif is not directly involved in catalysis, but defines the acceptor peptide substrate specificity (Lizak *et al.*, 2011). Specifically, the β-hydroxyl group of the +2 Thr of the bound acceptor peptide forms three hydrogen bonds with the side chains of the WWD amino acids. Moreover, Ile in the MXXI motif (Ile 572 in *C. lari* PglB), which is

conserved in bacteria and shown to be important for OTase activity (Maita *et al.*, 2010), was suggested to provide contacts to the +2 Thr (Lizak *et al.*, 2011). Despite the molecular explanation provided by the PglB structure for OTase activity, the authors do not discuss the mechanism for fOS production, which is likely due to the absence of the LLO present in the structure. The WWDYG motif was shown to be necessary for fOS release both *in vivo* and *in vitro* (Nothaft *et al.*, 2009). Particularly, PglB can release fOS *in vitro* in the absence of an acceptor protein, but no significant fOS is observed with an inactive PglB mutated at the WWDYG motif, suggesting that this signature motif plays a larger role than simply peptide binding or recognition (Nothaft *et al.*, 2009).

Furthermore, PglB is a metal-dependent glycosyltransferase and is only functional when it is bound to a divalent cation (M^{2+}), either Mn^{2+} or Mg^{2+} (Imperiali & Rickert, 1995). Since a bound metal cation can be found in the LLO-binding site of the PglB structure, it was proposed that this binding cavity serves as the catalytic pocket where the asparagine residue is modified (Lizak *et al.*, 2011). From this region of PglB, the authors identified three acidic side chains (D56, D154, E319) located in the transmembrane domain and, through alanine substitutions of these residues, they demonstrated a significant reduction in PglB activity. The crystal structure also revealed that the carboxyl groups of D56 and E319 interact with the M^{2+} cation and the amido group of the asparagine residue, suggesting that M^{2+} stabilizes the lipid-pyrophosphate leaving group (Lizak *et al.*, 2011). Recent biochemical and mutagenesis studies of PglB confirm the importance of the aforementioned and corresponding residues (D54, D152, E316

in *C. jejuni* PglB) that are involved in catalysis and enzyme activity (Jaffee & Imperiali, 2011).

1.4.6. *C. jejuni* PglB substrate specificity

The oligosaccharide substrate for *N*-glycosylation must be lipid-linked. Although the undecaprenyl carrier is the native substrate for *N*-linked protein glycosylation that is accepted by PglB, polyisoprenyl carriers of shorter length also display similar turnover to the native substrate (Chen *et al.*, 2007b). PglB does not have stringent specificity for the number of isoprene units, but its activity significantly decreases in the presence of dolichols, which include a saturated terminal α-isoprene unit (Chen *et al.*, 2007b). This suggests that PglB is highly sensitive to the degree of lipid saturation near the site of enzymatic action.

The *N*-glycan in *C. jejuni* is a heptasaccharide consisting of five GalNAc residues, a glucose branch and bacillosamine at the reducing end, as determined by nano-NMR techniques (Young *et al.*, 2002). However, the *C. jejuni* glycoprotein AcrA produced in an *E. coli* background containing the *pgl* operon displayed two separate fragmentation patterns when analyzed by collision induced dissociation MS-MS (Wacker *et al.*, 2002). The glycan structures of the two recombinant AcrA proteins were composed of a heptasaccharide that showed very similar patterns with the exception that all major fragment ions from one sample were shifted to a lower mass by 25 daltons. Wacker and colleagues (2002) determined that the 25-dalton increment was also the difference between bacillosamine and an *N*-acetylhexosamine (HexNAc) residue, which suggests that

the glycan on one of the recombinant AcrA glycoproteins is attached through a HexNAc rather than bacillosamine.

Similarly, the *C. jejuni* glycoprotein PEB3 produced in *E. coli* in the presence of the *pgl* locus also showed two structural variants of the glycan attached to the protein (Linton *et al.*, 2005). Tandem MS-MS analysis of the glycans attached to PEB3 also showed a second fragmentation pattern that was shifted down by 25 daltons, as previously observed for the AcrA glycoprotein analysis. Linton and colleagues (2005) demonstrated that the variant glycan containing a HexNAc at the reducing end results from the activity of *E. coli* WecA, a UDP-GlcNAc transferase, in lieu of *C. jejuni* PglC, the glycosyltransferase responsible for the transfer of UDP-Bac onto the lipid carrier Und-P. When the *C. jejuni* glycoprotein AcrA and the *pgl* locus with an inactivated *pglC* gene were expressed in *E. coli* lacking the *wecA* gene, production of non-glycosylated AcrA resulted. However, complementation of the *pglC* mutant by a wild-type *wecA* gene *in trans* restored glycosylation of AcrA, but only with the glycoform with reduced mass. Therefore, the *E. coli* WecA can complement the function of PglC (Linton *et al.*, 2005). More importantly, these results show that PglB can transfer a heptasaccharide that contains a reducing HexNAc residue, which varies from the native glycan found in *C. jejuni*.

The ability of PglB to transfer glycans of various lengths was examined by preparing chemically synthesized sugar substrates using enzymes from the *C. jejuni pgl* locus (Glover *et al.*, 2005b). *In vitro* analysis showed that PglB is capable of transferring oligosaccharides of varying lengths (two to seven

saccharides) to protein acceptors (Glover *et al.*, 2005b). The transfer of the truncated *C. jejuni* glycans by PglB was also examined *in vivo* through site-directed mutagenesis of individual genes in the *pgl* operon expressed in *E. coli* (Linton *et al.*, 2005). Although the primary purpose of the *pgl* gene knockout mutations was to elucidate the function of the various glycosyltransferase and sugar-modifying enzymes involved in the pathway, these studies have also revealed that PglB can transfer truncated glycans down to a single residue of bacillosamine attached to the peptide (Linton *et al.*, 2005).

To further examine the substrate specificity of PglB, Feldman and colleagues (2005) tested whether PglB can modify acceptor proteins with the O-antigens from *E. coli* and *Pseudomonas aeruginosa,* an outer component of the lipopolysaccharide synthesized on the Und-P lipid carrier. PglB was found to transfer diverse oligosaccharides from *E. coli*, in addition to the *C. jejuni* glycan. All of the O-antigens transferred by PglB were confirmed to contain Bac, GlcNAc, GalNAc, or N-acetyl fucosamine (FucNAc), all of which contain acetyl groups at the C-2 position of the sugar residues at the reducing ends (Feldman *et al.*, 2005). Furthermore, Wacker and colleagues (2006) investigated the influence of a hexose branch at the C-6 position of the reducing sugar from the *E. coli* O16 antigen and found that it did not prevent PglB from covalently linking it to the protein acceptor. However, when PglB was expressed in a *Salmonella enterica* LT2 ligase mutant, glycosylation of the AcrA acceptor protein was lacking. The coexpression of the *C. jejuni N*-glycosylation machinery and AcrA in *S. enterica* LT2 cells, which resulted in glycosylated protein, confirmed the activity of PglB.

The *Salmonella* LT2 O-antigen contains a galactose residue at the reducing end, which strongly suggests that PglB has specificity for HexNAc residues at that position. It was determined that an acetyl group at the C-2 position of the reducing-end sugar is required for transfer by PglB (Wacker *et al.*, 2006). Although the glycan substrate of PglB is not required to comprise the exact *C. jejuni* heptasaccharide structure, its specificity lies within the reducing monosaccharide. All of the PglB substrates identified thus far contain a sugar at the reducing end with an acetyl group at the C-2 position. Structural evaluation of all the oligosaccharides accepted by PglB also suggests that the first two sugars at the reducing end cannot have a β-1,4-linkage since the native linkage in *C. jejuni* is a-1,3 (Chen *et al.*, 2007a).

1.4.7. *C. jejuni* PglB acceptor specificity

Similar to eukaryotic glycoproteins, the consensus sequon (N-X-S/T) was also found at glycan attachment sites within the *C. jejuni* identified glycoproteins (Linton *et al.*, 2002. Young *et al.*, 2002; Wacker *et al.*, 2002). The sequence requirements of the acceptor proteins for bacterial *N*-glycosylation were analyzed *in vivo* in the heterologous *E. coli* host expressing the *C. jejuni pgl* locus and the model glycoprotein AcrA (Nita-Lazar *et al.*, 2005). Site-directed mutagenesis of the putative bacterial N-X-S/T consensus sequence showed loss of AcrA glycosylation when the asparagine residue and the serine or threonine residues at the +2 position were replaced with any other amino acids. The amino acid at the +1 position was determined to be any residue except proline, as it impaired glycosylation (Nita-Lazar *et al.*, 2005). Subsequent work by Kowarik and

colleagues (2006) demonstrated that an extended consensus sequence (D/E-X_1-N-X_2-S/T) is required for bacterial *N*-linked protein glycosylation, where X_1 and X_2 cannot be proline. The consensus sequence requirements were determined by truncation and mutation experiments with the *C. jejuni* protein AcrA expressed in *E. coli* bearing the *pgl* system (Kowarik *et al.*, 2006b). Stepwise reduction of the AcrA protein was confirmed by the decreasing size of the protein as detected by SDS-polyacrylamide gel electrophoresis. The truncated forms of AcrA were analyzed by Western Blot using sera raised against *C. jejuni* whole-cell extracts shown to preferentially detect *C. jejuni N*-glycoproteins (Wacker *et al.*, 2002). This allowed the authors to determine the minimal requirements of the bacterial consensus sequence to contain an amino acid with a negatively charged side chain at position -2, either an aspartic acid or a glutamic acid residue. Furthermore, comparison of the amino acid residues in the consensus sequence revealed that PglB has a significant preference for threonine as the hydroxyamino acid at position +2 and aspartic acid as the acidic amino acid at position -2 (Chen *et al.*, 2007a).

The acceptor proteins must also contain a signal sequence that will target the proteins to the periplasm for modification by PglB (Nita-Lazar *et al.*, 2005). This was confirmed by deletion of the *C. jejuni* AcrA signal sequence, which resulted in cytoplasmic expression of the protein without modification. However, there are proteins in the *C. jejuni* genome that contain the extended consensus sequence in addition to a signal sequence, but do not appear to be glycosylated (Weerapana & Imperiali, 2006). Thus, these sequences are necessary, but not

sufficient for *N*-glycosylation. The *C. jejuni* AcrA glycoprotein contains five potential *N*-glycosylation sites; however, when analyzed by Western Blot using anti-AcrA antisera, only three bands are observed indicating the unglycosylated, monoglycosylated, and diglycosylated forms of AcrA (Wacker *et al.*, 2002; Feldman *et al.*, 2005; Kowarik *et al.*, 2006b). Glycosylated peptides derived from tryptic cleavage of *C. jejuni* AcrA and site-directed mutagenesis of other potential glycosylation sites in AcrA demonstrated that the modified sites are ^{123}NRS^{125} and ^{273}NNS^{275} (Nita-Lazar *et al.*, 2005; Wacker *et al.*, 2006).

The conformation of the acceptor protein was proposed to be important for *N*-linked glycosylation. Structural studies indicate that a specific secondary structure, named the Asn-turn, is adopted by the acceptor protein at the consensus sequence (Imperiali *et al.*, 1992; Imperiali & Hendrickson, 1995). The Asn-turn is proposed to form by the interaction of the amide group of the acceptor asparagine with the hydroxyl group of the +2 Thr or Ser, which has been suggested to be necessary for interaction with the OTase during the glycosylation process. However, the *C. lari* PglB crystal structure reveals that the +2 Thr of the acceptor peptide binds firmly to the WWD residues of the WWDYG motif in PglB, disproving the involvement of the +2 Thr (Lizak *et al.*, 2011). Instead, the molecular explanation provided by the PglB structure is that the catalytically essential residues D56 and E319 of the *C. lari* PglB form hydrogen bonds with the two amide protons of the acceptor asparagine. This is proposed to result in the rotation of the N-C bond of the amido group, which would increase the

electronegative nature of the amide nitrogen making it optimal for a nucleophilic attack (Lizak *et al.*, 2011).

Furthermore, the folding state of acceptor proteins during glycosylation was demonstrated to play a role in the efficiency of PglB (Kowarik *et al.*, 2006a). In Eukaryotes, the glycosylation process occurs co-translationally, where the OTase modifies the acceptor proteins in the lumen of the ER prior to folding of the proteins (Kelleher & Gilmore, 2006). An *in vitro* glycosylation assay using PglB expressed in and purified from *E. coli*, co-incubated with *C. jejuni* AcrA and LLO, resulted in fully glycosylated AcrA (Kowarik *et al.*, 2006a). However, analysis of the PglB-dependent *in vitro* glycosylation of the eukaryotic glycoprotein, bovine ribonuclease A, showed weak glycosylation when the protein was fully folded. Different folding variants of bovine ribonuclease A demonstrated the efficiency of PglB, where partial or complete unfolding improved glycosylation activity. Therefore, in the bacterial system, it is strongly suggested that glycosylation sites are located in the flexible loops of folded proteins since the Pgl process occurs post-translationally (Kowarik *et al.*, 2006a).

1.4.8. Exploitation of PglB for glycoprotein engineering

The *C. jejuni pgl* gene locus contains all of the genes necessary for the biosynthesis of LLO and the transfer of the heptasaccharide onto asparagine residues of acceptor proteins. And as mentioned above, the *C. jejuni N*-linked protein glycosylation pathway, encoded by the *pgl* operon, has been functionally reconstituted in *E. coli* to produce recombinant glycoproteins (Wacker *et al.*, 2002). This has opened up the possibility to exploit the bacterial glycosylation

system for biotechnological applications. The study of the glycosylation machinery has been accelerated by the use of *E. coli* laboratory strains, given that this well-characterized bacterium is easy to genetically manipulate. In addition to the full mutant library available for this organism, *E. coli* constitutes a perfect toolbox for glycoengineering due to its lack of native glycoproteins and its ability to tolerate incorportation of foreign glycosylation pathways (Feldman, 2009). Elucidation of the individual roles for the enzymes encoded by the *pgl* locus has provided understanding of the bacterial *N*-glycosylation pathway and has allowed the development of strategies to apply this system for glycoengineering.

The demonstration that PglB can transfer glycans other than the *C. jejuni* heptasaccharide to acceptor proteins has also advanced the field such that glycoproteins can be engineered to contain a diverse range of bacterial glycans (Feldman *et al.*, 2005; Wacker *et al.*, 2006). Moreover, engineering additional glycosylation sites into an acceptor protein has been demonstrated (Kowarik *et al.*, 2006b). The extended consensus sequence (D/E-X_1-N-X_2-S/T) was artificially introduced into *C. jejuni* AcrA at a flexible domain, resulting in a triglycosylated form of AcrA compared to the wild-type diglycosylated form. Therefore, modification by PglB does not require the acceptor protein to be a natural glycoprotein, but rather an acceptor site and a signal sequence can be engineered into any protein of interest for the production of novel recombinant glycoproteins. The establishment of glycoprotein engineering in *E. coli* shows particular promise in the area of vaccine development against bacterial pathogens.

The first line of defense against bacterial infections in humans is through circumvention by the innate immune system that consists of nonspecific host defenses existing prior to bacterial exposure, whereas the adaptive immune response mediated by B and T cells following exposure to bacterial antigens provides long-lasting protection against pathogens (Kindt *et al.*, 2007). Therefore, vaccination in humans prepares the adaptive immune system to deal efficiently with pathogens that would not be easily cleared by innate immunity. Bacterial polysaccharides are the most abundant and accessible molecules on the cell surface and have been used as vaccines against diseases caused by *Streptococcus pneumonia, Neisseria meningitidis,* and *Haemophilus influenzae* type b (Zou & Jennings, 2009). However, a limitation of polysaccharide vaccines is their inability to activate T cells that differentiate into memory cells, which are antigen-generated, long-lived cells that respond with heightened reactivity to subsequent challenge with the same antigen (Kindt *et al.*, 2007). Polysaccharides alone have been shown to be poor immunogens because they fail to elicit a long-term antibody response (Jennings & Lugowski, 1981). A solution to involve T cells directly for long-lived, antibody-mediated protection is to develop glycoconjugate vaccines, which are composed of bacterial glycans that are covalently attached to an appropriate carrier protein (Feldman, 2009). The first and most successful example is the infant meningitis vaccine against group B meningococcal infection caused by *H. influenzae* type b (Jennings *et al.*, 1986). The vaccine was developed by conjugating the group B meningococcal polysaccharide to the protein carrier, tetanus toxoid, which yielded high levels of group B meningococcal

polysaccharide-specific IgG antibodies, known as memory cells. The rate of *H. influenzae* type b infections has dramatically fallen since the introduction of this vaccine in the United States (Kindt *et al.*, 2007).

The traditional approach to generate glycoconjugate vaccines include several steps that involve the purification of the acceptor protein, chemical treatment of the bacterial polysaccharide to remove any toxic components, and chemical-crosslinking of the purified polysaccharide with the acceptor protein (Feldman, 2009). However, using the glycoprotein engineering method in bacteria to produce glycoconjugates has several advantages including minimal manipulation and simpler purification. In this case, the *C. jejuni* PglB can be exploited for the production of glycoconjugated vaccines. The approach would simply involve the co-expression of PglB, an appropriate acceptor protein and a plasmid carrying genes encoding specific oligosaccharides into *E. coli* to produce recombinant glycoproteins. Glycoconjugates can also be produced directly in bacteria other than *E. coli*.

1.4.9. Putative Pgl pathways in proteobacteria

The *N*-glycosylation pathway first discovered in *C. jejuni,* a microaerophilic bacterium belonging to the epsilon class of Proteobacteria, has now been identified in an increasing number of organisms (Nothaft and Szymanski, 2010). For undetermined reasons, the pathway is confined to the epsilon and delta class of Proteobacteria (Figure 1.4). The *pgl* gene clusters are conserved in all *Campylobacter* species sequenced to date, the delta-Proteobacterium *Desulfovibrio desulfuricans* (Ielmini & Feldman, 2011), and in

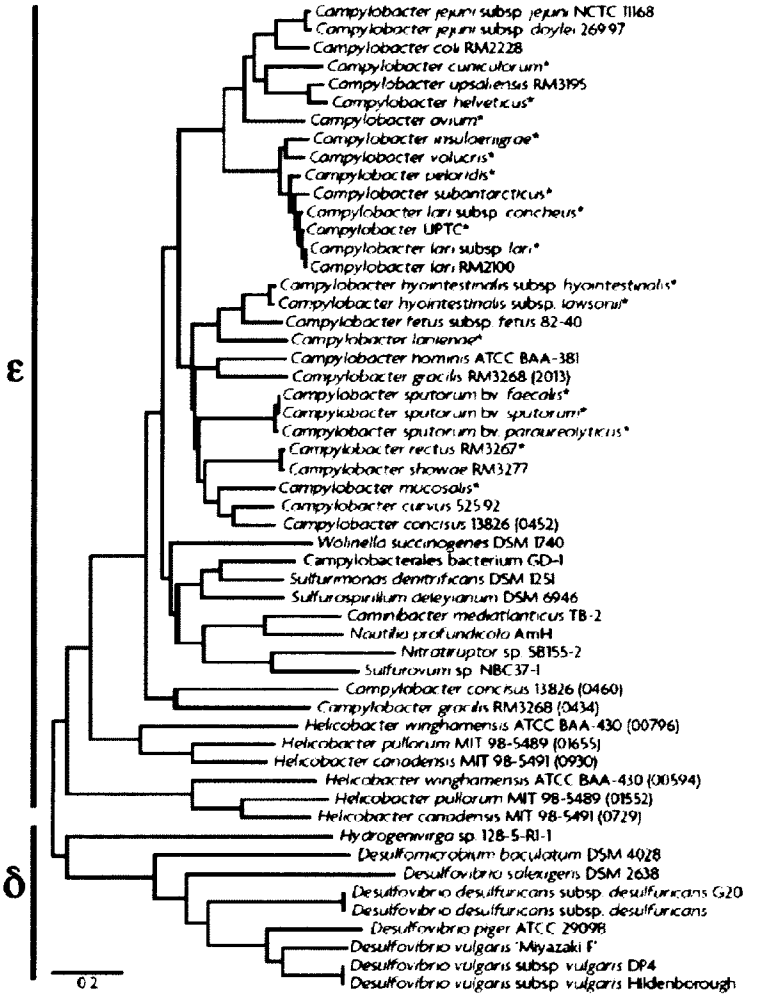

Figure 1.4. Phylogenetic tree of PglB protein sequences. A phylogenetic tree containing all the available oligosaccharyltransferase sequences for *N*-linked protein glycosylation. The pathway appears to be confined to the epsilon (ε) and delta (δ) class of Proteobacteria. Modified from Nothaft & Szymanski, 2010.

the epsilon-Proteobacteria *Wolinella succinogenes* (Baar *et al.*, 2003), *Sulfurovum* sp. and *Nitratiruptor* sp. (Nothaft *et al.*, 2008). There are a few *Helicobacter* species that also contain *pgl* gene orthologues found in the genomes of *H.*

canadensis, *H. winghamensis*, and *H. pullorum*, each carrying two *pglB* orthologues (Jervis *et al.*, 2010). Some *Campylobacter* species also carry two *pglB* orthologues, including *C. concisus*, *C. curvus*, and *C. gracilis* (Nothaft & Szymanski, 2010). It has not been shown that the *N*-glycosylation pathway is functional in all of these organisms yet, but all of them contain the key enzyme, PglB. Therefore, it is predicted that these organisms also *N*-glycosylate their proteins (Nothaft *et al.*, 2008). Interestingly, *Sulfurovum* sp., and *Nitratiruptor* sp. are two non-pathogenic, deep-sea vent strains that share virulence genes with their pathogen relatives, *Helicobacter* and *Campylobacter* species (Nakagawa *et al.*, 2007). Since the deep-sea vent epsilon-Proteobacteria represent the deepest branching Bacteria in the PglB/Stt3 phylogenetic tree, it has been proposed that the Pgl pathway most likely arose in these organisms to maintain a symbiotic relationship with hydrothermal vent invertebrates (Nakagawa *et al.*, 2007), although another explanation may be that the pathway was aquired to allow the organisms to survive in saline environments.

1.4.10. Comparison of PglB to other OTases

Bacterial *N*-linked protein glycosylation is best characterized in *C. jejuni*; however, this system has recently been characterized in *C. lari*, *H. pullorum*, and *D. desulfuricans* (Schwarz *et al.*, 2010; Jervis *et al.*, 2010; Ielmini & Feldman, 2010). The genome of *C. lari* contains a *pgl* gene locus that shares a similar genetic organization to that of *C. jejuni*, with the exception of a *pglI* gene encoding for the glucosyltransferase that is missing in *C. lari* (Nothaft & Szymanski, 2010). The *C. lari* Pgl pathway has been reconstituted in *E. coli* to

produce recombinant glycoproteins (Schwarz *et al.*, 2010). The authors demonstrated that *C. lari* PglB (*Cl*PglB) modifies the *C. jejuni* acceptor protein AcrA with a linear hexasaccharide that consisted of five HexNAc residues and a Bacillosamine at the reducing end. *Cl*PglB was also shown to exhibit a broader specificity toward the acceptor protein (Schwarz *et al.*, 2010). The authors found that *Cl*PglB does not require an acidic group at the -2 position of the consensus sequence, but that it can modify asparagine residues found within the shorter eukaryotic sequence N-X-S/T, albeit with low efficiency. When they tested the activity of *Cl*PglB towards the eukaryotic protein immunoglobulin G, which carries the N-X-S/T sequence, the authors found that glycosylation did not result. However, when an acidic group was engineered into the protein at the -2 position, glycosylation was observed by Western blot analysis using anti-*N*-glycan specific antibodies. *Cl*PglB was also shown to glycosylate AcrA at NNNST and at an unusual DANSG sequon (Schwarz *et al.*, 2010). This study extended the understanding of *Campylobacter* PglBs in regards to substrate and sugar specificity.

In *H. pullorum*, the *pgl* gene orthologues are not encoded in a single genetic locus as seen for *C. jejuni*, but distributed throughout the genome (Nothaft & Szymanski, 2010). *H. pullorum* contains two unrelated *pglB* genes (*pglB1* and *pglB2*), where they could only confirm that PglB1 displayed OTase activity (Jervis *et al.*, 2010). The researchers attempted to clone the *pglB2* gene from *H. pullorum*, but were unsuccessful suggesting that this gene is essential for the viability of the organism. *In vivo* glycosylation of the *C. jejuni* glycoprotein,

*Cj*0114, was observed in *E. coli* upon coexpression of PglB1 from *H. pullorum* with the *C. jejuni pgl* operon carrying a mutated *pglB* by insertion of a kanamycin-cassette (Jervis *et al.*, 2010). *In vitro* peptide glycosylation by PglB1 was assayed using detergent-solubilized membrane preparations of *H. pullorum* wildtype and a *pglB1* insertional knock-out mutant. The authors demonstrated that the *H. pullorum* wildtype, but not the *pglB1* knock-out mutant, modified the fluorescent peptide carrying the DQNAT sequon. Unlike *Cj*PglB, PglB1 of *H. pullorum* shares similar acceptor specificity with *Cj*PglB and requires the negatively charged Asp at the -2 position of the consensus sequence (Jervis *et al.*, 2010). However, the *N*-linked glycan structure of the PglB1-generated glycopeptide was analyzed by MALDI-MS/MS and determined to comprise a pentasaccharide with a HexNAc at the reducing end, a structure that differs from the *C. jejuni N*-linked heptasaccharide (Jervis *et al.*, 2010).

N-glycosylation has also been characterized in the delta class of Proteobacteria. Structural and MS analyses led to the identification of the *N*-glycoprotein high-molecular-mass cytochrome *c* (HmcA), a 16 heme cytochrome of *Desulfovibrio gigas*, modified at the asparagine residue within a TANGT sequon lacking the negatively charged amino acid at the -2 position (Santos-Silva *et al.*, 2007). The authors proposed that a disaccharide is attached to Asn261 of HmcA, which comprises GlcNAc followed by its epimer *N*-acetylallosamine. Recently, PglB of *D. desulfuricans* ($PglB_{Dd}$) was demonstrated to exhibit OTase activity when it was functionally expressed in *E. coli* (Ielmini & Feldman, 2010). Similar to *H. pullorum*, the *pgl* gene orthologues in *D. desulfuricans* are scattered

across the genome (Nothaft & Szymanski, 2010), but non-homologous genes potentially involved in glycan biosynthesis were found to be clustered around *pglB* (Ielmini & Feldman, 2010). Unlike *C. jejuni* and *H. pullorum*, $PglB_{Dd}$ was shown to preferentially modify asparagine residues found within the eukaryotic-like sequon N-X-S/T (Ielmini & Feldman, 2010). The authors further demonstrated that $PglB_{Dd}$ transfers the *C. jejuni* heptasaccharide, a single *N,N'*-diacetyl-bacillosamine, and the *E. coli* O7 O-antigen. The structure of the endogenous glycan in *D. desulfuricans* is yet to be fully elucidated.

A common theme for the bacterial *N*-glycosylation system is the modification of acceptor asparagine residues of glycopeptides which takes place in the periplasm through a post-translational process. Interestingly, an unusual system of bacterial *N*-glycosylation has been recently described in the gammaproteobacterium *Haemophilus influenzae*, where modification occurs in the cytoplasm through a sequential transfer of sugars (Gross *et al.*, 2008). The *H. influenzae* high-molecular-weight adhesin (HMW1) was shown to be glycosylated at 31 sites, where all the acceptor asparagine residues were found within the eukaryotic-like N-X-S/T consensus sequence, with the exception of one site (Gross *et al.*, 2008). MS analyses determined the structure of the modifying glycans decorated on HMW1 to comprise monohexoses or dihexoses. Further analysis revealed that the *H. influenzae* HMW1C protein is the glycosyltransferase (GTase) responsible for modifying asparagine residues in the HMW1 adhesin, particularly with glucose and galactose residues as demonstrated in an *in vitro* transferase assay with purified components (Grass *et al.*, 2010).

Surprisingly, HMW1C exhibits both OTase and GTase activities, where asparagine residues are modified and hexose-hexose bonds are generated, respectively (Grass *et al.*, 2010).

Bacterial OTases other than PglB can be used for glycoprotein engineering. *O*-linked protein glycosylation of pilin subunits, the main components of the type IV pilus, has been found in *N. meningitidis* and *N. gonorrhoeae* (Virji *et al.*, 1993). Similar to the *N*-glycosylation system in *C. jejuni*, the genes encoding the enzymes involved in the pathway share sequence homologies with some of the enzymes from *C. jejuni* and can also be found clustered in an operon termed *pgl* (Stimson *et al.*, 1995). PglL was demonstrated to be the OTase in *N. meningitidis* that transfers a disaccharide or a trisaccharide onto Ser or Thr residues of pilin subunits (Stimson *et al.*, 1995; Aas *et al.*, 2007). The OTase in *N. gonorrhoeae* showed loss of pilin glycosylation when mutated and was named PglO (Aas *et al.*, 2007). PglL and PglO transfer oligosaccharides *en bloc* from Und-PP to the protein acceptors.

Pilin glycosylation was also described in *P. aeruginosa* (Castric, 1995). *O*-glycosylation at the C-terminal serine residue of the pilin subunits in *P. aeruginosa* was shown to occur by the acitivity of the OTase, PilO (Comer *et al.*, 2002). The glycan attached to pilin subunits consists of a trisaccharide, a single O-antigen subunit derived from the LPS biosynthetic pathway (Castric *et al.*, 2001; Castro *et al.*, 2006). Similar to PglB, the *N. meningitidis* and *P. aeruginosa* pilin *O*-glycosylation pathways were reconstituted in *E. coli* (Faridmoayer *et al.*, 2007). Expression of PglL and PilO in *E. coli* in the presence of their respective pilin

substrates and lipid-linked glycan donors resulted in the *O*-glycosylation of pilin. The glycan specificity of the *O*-OTases, PglL and PilO, is more relaxed than the *C. jejuni* PglB. It has been demonstrated that the *O*-OTases can transfer the *C. jejuni* heptasaccharide in addition to other polysaccharides. Faridmoayer and colleagues (2008) have demonstrated that PglL can transfer polysaccharides containing a hexose at the reducing end, a limitation that prevents transfer by PglB (Faridmoayer *et al.*, 2008). More recently, PglA was identified to be the *O*-OTase as demonstrated by the glycosylation of the PilA pilin protein in *Francisella tularensis* (Egge-Jacobsen *et al.*, 2011). MS analyses of the glycan structure was carried out and determined to comprise a pentasaccharide (HexNAc-Hex-Hex-HexNAc-HexNAc) attached to multiple hydroxyl groups of Ser or Thr residues.

The discovery of the bacterial OTases opens up the possibility to construct conjugate glycoproteins in *E. coli* using a recombinant glycoprotein engineering approach. All of the bacterial OTases that have been characterized so far can be used to produce recombinant glycoproteins with application in biotechnology (Table 1). The activity of PilO is limited to short glycan chains, whereas PglB and PglL can transfer polysaccharides (Faridmoayer *et al.*, 2007). Moreover, PglB activity is confined by its substrate specificity, whereas PglL has been found to transfer all tested glycans onto acceptor proteins (Feldman, 2009). Although PglL is highly non-specific, its ability to transfer any sugar that is assembled on Und-P is also a disadvantage. PglL can interfere with the synthesis of other glycan structures, such as the peptidoglycan and certain CPS types that are also

assembled on Und-P, by transferring the incorrect sugar to the acceptor protein. Therefore, the transfer of a polysaccharide of interest onto the acceptor protein is not guaranteed as PglL can be transferring any glycan produced on a lipid carrier in *E. coli*. Although PglB is more restricted to the proteins it modifies, its advantage is that the D/E-X-N-X-S/T consensus sequence can be engineered in flexible loops in proteins that are already used for glycoconjugate vaccine production, such as tetanus toxoid, cholera toxin and exotoxin A (Kowarik *et al.*, 2006b). Furthermore, PglB from *C. lari* and *D. desulfuricans* can modify asparagine residues within the shorter eukaryotic sequon N-X-S/T, making it easier to engineer the sequon into acceptor proteins. Although glycosylation at the eukaryotic sequon is made with lower efficiency, the functional characterization of PglB is a priority in order to improve its OTase activity. The number of organisms found to contain the *N*- and *O*-OTases is growing. It is predicted that new OTases with different characteristics will be found that may have improved activity for use in glycoprotein engineering.

1.5. Conclusions and thesis objectives

1.5.1. Concluding remarks

Structural comparison of distantly-related OTases draws attention to the common architecture of important catalytic domains required for function, which could not be identified by merely comparing sequence similarities. Research over several decades devoted to understanding the eukaryotic pathway can now be applied to understanding the prokaryotic process. Recognizing the factors that influence *N*-glycosylation by PglB is an important start to the elucidation of the

poorly understood mechanism. Furthermore, understanding PglB substrate specificity and the basis for recognition will improve glycoprotein engineering. PglB presents the advantage of creating acceptor proteins by simply adding the consensus sequence required for glycosylation in a surface-exposed flexible loop of a non-glycosylated protein. Moreover, developing inhibitors of the PglB enzyme would be valuable in the pursuit for antibiotics against *C. jejuni* in order to treat gastrointestinal disorders.

1.5.2. Thesis objectives

The objective of my research project is to identify additional amino acid residues that are important for *C. jejuni* PglB activity by using a bioinformatic approach and generating point mutations at specific amino acid sites. Some of the questions I would like to answer from my research include: What are the key residues in PglB that are required for function? How do the activities of sequenced *Campylobacter* PglB enzymes compare with respect to *N*-glycosylation? *In silico* analysis of a multiple sequence alignment of *Campylobacter* PglB enzymes identified a DXXK sequence located close to the highly conserved WWDXG motif. The DXXK motif is conserved in all sequenced *Camplobacter* PglBs and other epsilon- and delta-Proteobacteria. I sought to determine the function of this newly identified sequence by generating site-directed mutants of ^{475}DXXK^{478} in the *C. jejuni* PglB. Further upstream of the WWDXG motif in the PglB protein sequence, there appears to be a conserved a proline, serine or valine. I began by comparing the importance of the proline residue of *C. jejuni* PglB with the serine of *C. lari* PglB. Understanding the

mechanism of PglB transfer activity will be invaluable in creating a better OTase with enhanced *N*-glycosylation activity compared to the native *C. jejuni* protein.

Furthermore, the protein acceptor site required for *N*-glycosylation is also found within the soluble C-terminal domain of PglB (^{532}DYNQS^{536}). This raises the following questions: Is glycosylation of PglB required for its function? If so, does PglB self-glycosylation modulate its OTase and hydrolase activities? Analysis of the *C. jejuni* PglB soluble domain crystal structure revealed that the consensus sequence is located at the surface of the protein, which would allow for an autoregulatory mechanism. In this thesis, I will provide details of the research that was undertaken to functionally characterize the *Campylobacter* OTase for *N*-linked protein glycosylation.

1.6. Literature cited

Aas FE, Vik Å, Vedde J, Koomey M & Egge-Jacobsen W (2007) *Neisseria gonorrhoeae O*-linked pilin glycosylation: functional analyses define both the biosynthetic pathway and glycan structure. *Mol Microbiol* **65:** 607-624.

Abu-Qarn M & Eichler J (2007) An analysis of amino acid sequences surrounding archaeal glycoprotein sequons. *Archaea* **2:** 73-81.

Abu-Qarn M & Eichler J (2006) Protein *N*-glycosylation in Archaea: defining *Haloferax volcanii* genes involved in S-layer glycoprotein glycosylation. *Mol Microbiol* **61:** 511-525.

Abu-Qarn M, Eichler J & Sharon N (2008) Not just for Eukarya anymore: protein glycosylation in Bacteria and Archaea. *Curr Opin Struct Biol* **18:** 544-550.

Aebi M, Bernasconi R, Clerc S & Molinari M (2009) *N*-glycan structures: recognition and processing in the ER. *Trends Biochem Sci* **35:** 74-82.

Abbadi A, Mcharfi M, Aubry A, Premilat S, Boussard G & Marraud M (1991) Involvement of side functions in peptide structures: the Asx turn. Occurrence and conformational aspects. *J Am Chem Soc* **113:** 2729-2735.

Alaimo C, Catrein I, Morf L, Marolda CL, Callewaert N, Valvano MA, Feldman MF & Aebi M (2006) Two distinct but interchangeable mechanisms for flipping of lipid-linked oligosaccharides. *EMBO J* **25:** 967.

Aebi M (2006) Two distinct but interchangeable mechanisms for flipping of lipid-linked oligosaccharides. *EMBO J* **25:** 967-976.

Apweiler R, Hermjakob H & Sharon N (1999) On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim Biophys Acta* **1473:** 4-8.

Baar C, Eppinger M, Raddatz G *et al.* (2003) Complete genome sequence and analysis of Wolinella succinogenes. *Proc Natl Acad Sci USA* **100:** 11690-11695.

Barnes G, Hansen WJ, Holcomb CL & Rine J (1984) Asparagine-linked glycosylation in *Saccharomyces cerevisiae*: genetic analysis of an early step. *Mol Cell Biol* **4:** 2381-2388.

Bause E (1983) Structural requirements of *N*-glycosylation of proteins. Studies with proline peptides as conformational probes. *Biochem J* **209:** 331-336.

Belard M, Cacan R & Verbert A (1988) Characterization of an oligosaccharide-pyrophosphodolichol pyrophosphatase activity in yeast. *Biochem J* **255:** 235-242.

Berriman M, Ghedin E, Hertz-Fowler C, Blandin G, Renauld H, Bartholomeu DC, Lennard NJ, Caler E, Hamlin NE & Haas B (2005) The genome of the African trypanosome *Trypanosoma brucei. Science* **309:** 416-422.

Bhatia PK & Mukhopadhyay A (1999) Protein glycosylation: implications for *in vivo* functions and therapeutic applications. *Adv Biochem Eng Biotech* **64:** 155-202.

Bohin JP (2000) Osmoregulated periplasmic glucans in Proteobacteria. *FEMS Microbiol Lett* **186:** 11-19.

Bohin JP & Lacroix JM (2007) Osmoregulation in the periplasm. *The periplasm* (Ehrmann M, ed), pp. 325-341. ASM Press, Washington, DC.

Burda P & Aebi M (1999) The dolichol pathway of *N*-linked glycosylation. *Biochem Biophys* **1426:** 239-257.

Castric P (1995) *pilO*, a gene required for glycosylation of *Pseudomonas aeruginosa* 1244 pilin. *Microbiology* **141:** 1247-1254.

Castric P, Cassels FJ & Carlson RW (2001) Structural characterization of the *Pseudomonas aeruginosa* 1244 pilin glycan. *J Biol Chem* **276:** 26479-26485.

Castro O, Movsichoff F & Parodi AJ (2006) Preferential transfer of the complete glycan is determined by the oligosaccharyltransferase complex and not by the catalytic subunit. *Proc Natl Acad Sci USA* **103:** 14756-14760.

Chaban B, Voisin S, Kelly J, Logan SM & Jarrell KF (2006) Identification of genes involved in the biosynthesis and attachment of *Methanococcus voltae N*-linked glycans: insight into *N*-linked glycosylation pathways in Archaea. *Mol Microbiol* **61:** 259-268.

Chen MM, Glover KJ & Imperiali B (2007a) From peptide to protein: comparative analysis of the substrate specificity of *N*-Linked glycosylation in *C. jejuni. Biochemistry* **46:** 5579-5585.

Chen MM, Weerapana E, Ciepichal E, Stupak J, Reid CW, Swiezewska E & Imperiali B (2007b) Polyisoprenol specificity in the *Campylobacter jejuni N*-Linked glycosylation pathway. *Biochemistry* **46:** 14342-14348.

Chi JH, Roos J & Dean N (1996) The *OST4* gene of *Saccharomyces cerevisiae* encodes an unusually small protein required for normal levels of oligosaccharyltransferase activity. *J Biol Chem* **271:** 3132-3140.

Chia JS, Chang LY, Shun CT, Chang YY & Chen JY (2001) A 60-kilodalton immunodominant glycoprotein is essential for cell wall integrity and the maintenance of cell shape in *Streptococcus mutans*. *Infect Immun* **69:** 6987-6998.

Comer JE, Marshall MA, Blanch VJ, Deal CD & Castric P (2002) Identification of the *Pseudomonas aeruginosa* 1244 pilin glycosylation site. *Infect Immun* **70:** 2837-2845.

Dennis JW, Granovsky M & Warren CE (1999) Protein glycosylation in development and disease. *Bioessays* **21:** 412-421.

Doig P, Kinsella N, Guerry P & Trust TJ (1996) Characterization of a post-translational modification of *Campylobacter* flagellin: identification of a sero-specific glycosyl moiety. *Mol Microbiol* **19:** 379-387.

Espitia C & Mancilla R (1989) Identification, isolation and partial characterization of *Mycobacterium tuberculosis* glycoprotein antigens. *Clin Exp Immunol* **77:** 378-383.

Egge-Jacobsen W, Salomonsson EN, Aas FE *et al.* (2011) O-linked glycosylation of the PilA pilin protein of Francisella tularensis: identification of the endogenous protein-targeting oligosaccharyltransferase and characterization of the native oligosaccharide. *J Bacteriol* **193:** 5487-5497.

Faridmoayer A, Fentabil MA, Mills DC, Klassen JS & Feldman MF (2007) Functional characterization of bacterial oligosaccharyltransferases involved in *O*-linked protein glycosylation. *J Bacteriol* **189:** 8088-8098.

Faridmoayer A, Fentabil MA, Haurat MF, Yi W, Woodward R, Wang PG & Feldman MF (2008) Extreme substrate promiscuity of the *Neisseria* oligosaccharyl transferase involved in protein *O*-glycosylation. *J Biol Chem* **283:** 34596-34604.

Feldman MF (2009) Industrial exploitation by genetic engineering of bacterial glycosylation systems. *Microbial Glycobiology* (Moran A, Holst O, Brennan P & and Itzstein M, eds), pp. 903-914. Academic Press, Boston.

Feldman MF, Wacker M, Hernandez M, Hitchen PG, Marolda CL, Kowarik M, Morris HR, Dell A, Valvano MA & Aebi M (2005) Engineering *N*-linked protein glycosylation with diverse O antigen lipopolysaccharide structures in *Escherichia coli*. *Proc Natl Acad Sci USA* **102:** 3016-3021.

Freeze HH & Aebi M (2005) Altered glycan structures: the molecular basis of congenital disorders of glycosylation. *Curr Opin Struct Biol* **15:** 490-498.

Fry BN, Korolik V, ten Brinke JA, Pennings MTT, Zalm R, Teunis BJJ, Coloe PJ & van der Zeijst BAM (1998) The lipopolysaccharide biosynthesis locus of *Campylobacter jejuni* 81116. *Microbiology* **144:** 2049-2061.

Glover KJ, Weerapana E & Imperiali B (2005a) *In vitro* assembly of the undecaprenylpyrophosphate-linked heptasaccharide for prokaryotic *N*-linked glycosylation. *Proc Natl Acad Sci USA* **102:** 14255-14259.

Glover KJ, Weerapana E, Numao S & Imperiali B (2005b) Chemoenzymatic synthesis of glycopeptides with PglB, a bacterial oligosaccharyl transferase from *Campylobacter jejuni. Chem Biol* **12:** 1311-1315.

Grass S, Lichti CF, Townsend RR, Gross J & St Geme JW, 3rd (2010) The Haemophilus influenzae HMW1C protein is a glycosyltransferase that transfers hexose residues to asparagine sites in the HMW1 adhesin. *PLoS Pathog* **6:** e1000919.

Gross J, Grass S, Davis AE, Gilmore-Erdmann P, Townsend RR & St Geme JW, 3rd (2008) The Haemophilus influenzae HMW1 adhesin is a glycoprotein with an unusual N-linked carbohydrate modification. *J Biol Chem* **283:** 26010-26015.

Grunewald S, Matthijs G & Jaeken J (2002) Congenital disorders of glycosylation: a review. *Pediatr Res* **52:** 618-624.

Hebert DN, Garman SC & Molinari M (2005) The glycan code of the endoplasmic reticulum: asparagine-linked carbohydrates as protein maturation and quality-control tags. *Trends Cell Biol* **15:** 364-370.

Helenius A (1994) How *N*-linked oligosaccharides affect glycoprotein folding in the endoplasmic reticulum. *Mol Biol Cell* **5:** 253-265.

Helenius A & Aebi M (2001) Intracellular functions of *N*-linked glycans. *Science* **291:** 2364-2369.

Helenius J, Ng DTW, Marolda CL, Walter P, Valvano MA & Aebi M (2002) Translocation of lipid-linked oligosaccharides across the ER membrane requires Rft1 protein. *Nature* **415:** 447-450.

Hendrixson DR & DiRita VJ (2004) Identification of *Campylobacter jejuni* genes involved in commensal colonization of the chick gastrointestinal tract. *Mol Microbiol* **52:** 471-484.

Huffaker TC & Robbins PW (1983) Yeast mutants deficient in protein glycosylation. *Proc Natl Acad Sci USA* **80:** 7466-7470.

Ielmini MV & Feldman MF (2011) Desulfovibrio desulfuricans PglB homolog possesses oligosaccharyltransferase activity with relaxed glycan specificity and distinct protein acceptor sequence requirements. *Glycobiology* **21:** 734-742.

Igura M, Maita N, Kamishikiryo J, Yamada M, Obita T, Maenaka K & Kohda D (2008) Structure-guided identification of a new catalytic motif of oligosaccharyltransferase. *EMBO J* **27:** 234-243.

Imperiali B & Hendrickson TL (1995) Asparagine-linked glycosylation: specificity and function of oligosaccharyl transferase. *Bioorg Med Chem* **3:** 1565-1578.

Imperiali B & Rickert K (1995) Conformational implications of asparagine-linked glycosylation. *Proc Natl Acad Sci USA* **92:** 97.

Imperiali B, Shannon KL & Rickert KW (1992) Role of peptide conformation in asparagine-linked glycosylation. *J Am Chem Soc* **114:** 7942-7944.

Ivens AC, Peacock CS, Worthey EA, Murphy L, Aggarwal G, Berriman M, Sisk E, Rajandream MA, Adlem E & Aert R (2005) The genome of the kinetoplastid parasite, *Leishmania major. Science* **309:** 436-442.

Izquierdo L, Schulz BL, Rodrigues JA, Güther MLS, Procter JB, Barton GJ, Aebi M & Ferguson MAJ (2009) Distinct donor and acceptor specificities of *Trypanosoma brucei* oligosaccharyltransferases. *EMBO J* **28:** 2650-2661.

Jaffee MB & Imperiali B (2011) Exploiting topological constraints to reveal buried sequence motifs in the membrane-bound N-linked oligosaccharyl transferases. *Biochemistry* **50:** 7557-7567.

Jennings HJ & Lugowski C (1981) Immunochemistry of groups A, B, and C meningococcal polysaccharide-tetanus toxoid conjugates. *J Immunol* **127:** 1011-1018.

Jennings HJ, Roy R & Gamian A (1986) Induction of meningococcal group B polysaccharide-specific IgG antibodies in mice by using an *N*-propionylated B polysaccharide-tetanus toxoid conjugate vaccine. *J Immunol* **137:** 1708-1713.

Jervis AJ, Langdon R, Hitchen P, Lawson AJ, Wood A, Fothergill JL, Morris HR, Dell A, Wren B & Linton D (2010) Characterization of N-linked protein glycosylation in Helicobacter pullorum. *J Bacteriol* **192:** 5228-5236.

Josenhans C, Vossebein L, Friedrich S & Suerbaum S (2002) The neuA/flmD gene cluster of *Helicobacter pylori* is involved in flagellar biosynthesis and flagellin glycosylation. *FEMS Microbiol Lett* **210:** 165-172.

Karaoglu D, Kelleher DJ & Gilmore R (1995) Functional characterization of Ost3p. Loss of the 34-kD subunit of the *Saccharomyces cerevisiae* oligosaccharyltransferase results in biased underglycosylation of acceptor substrates. *J Cell Biol* **130:** 567-577.

Karlyshev AV, Everest P, Linton D, Cawthraw S, Newell DG & Wren BW (2004) The *Campylobacter jejuni* general glycosylation system is important for attachment to human epithelial cells and in the colonization of chicks. *Microbiology* **150:** 1957-1964.

Kasturi L, Chen H & Shakin-Eshleman SH (1997) Regulation of *N*-linked core glycosylation: use of a site-directed mutagenesis approach to identify Asn-Xaa-Ser/Thr sequons that are poor oligosaccharide acceptors. *Biochem J* **323:** 415-419.

Kelleher DJ & Gilmore R (2006) An evolving view of the eukaryotic oligosaccharyltransferase. *Glycobiology* **16**: 47-62.

Kindt TJ, Goldsby RA & Osborne BA (2007) *Kuby immunology.* WH Freeman, New York.

Kmiecik D, Herman V, Stroop CJM, Michalski JC, Mir AM, Labiau O, Verbert A & Cacan R (1995) Catabolism of glycan moieties of lipid intermediates leads to a single Man5GlcNAc oligosaccharide isomer: a study with permeabilized CHO cells.
Glycobiology **5:** 483-494.

Knauer R & Lehle L (1999) The oligosaccharyltransferase complex from yeast. *Biochim Biophys Acta* **1426:** 259-273.

Kohda D, Yamada M, Igura M, Kamishikiryo J & Maenaka K (2007) New oligosaccharyltransferase assay method. *Glycobiology* **17:** 1175-1182.

Kornfeld R & Kornfeld S (1985) Assembly of asparagine-linked oligosaccharides. *Annu Rev Biochem* **54:** 631-664.

Kowarik M, Numao S, Feldman MF, Schulz BL, Callewaert N, Kiermaier E, Catrein I & Aebi M (2006a) *N*-linked glycosylation of folded proteins by the bacterial oligosaccharyltransferase. *Science* **314:** 1148-1150.

Kowarik M, Young NM, Numao S, Schulz BL, Hug I, Callewaert N, Mills DC, Watson DC, Hernandez M, Kelly JF, Wacker M & Aebi M (2006b) Definition of the bacterial *N*-glycosylation site consensus sequence. *EMBO J* **25:** 1957-1966.

Kukuruzinska MA & Lennon K (1998) Protein *N*-glycosylation: molecular genetics and functional significance. *Crit Rev Oral Biol Med* **9:** 415-448.

Larsen JC, Szymanski C & Guerry P (2004) *N*-linked protein glycosylation is required for full competence in *Campylobacter jejuni* 81-176. *J Bacteriol* **186:** 6508-6514.

Lechner J & Wieland F (1989) Structure and biosynthesis of prokaryotic glycoproteins. *Annu Rev Biochem* **58:** 173-194.

Lechner J, Wieland F & Sumper M (1985) Transient methylation of dolichyl oligosaccharides is an obligatory step in halobacterial sulfated glycoprotein biosynthesis. *J Biol Chem* **260:** 8984-8989.

Lee SG, Pancholi V & Fischetti VA (2002) Characterization of a unique glycosylated anchor endopeptidase that cleaves the LPXTG sequence motif of cell surface proteins of Gram-positive bacteria. *J Biol Chem* **277:** 46912-46922.

Li L, Woodward R, Ding Y, Liu XW, Yi W, Bhatt VS, Chen M, Zhang LW & Wang PG (2010) Overexpression and topology of bacterial oligosaccharyltransferase PglB Biochem Biophys Res Commun **394:** 1069-1074.

Linton D, Allan E, Karlyshev AV, Cronshaw AD & Wren BW (2002) Identification of N-acetylgalactosamine-containing glycoproteins PEB 3 and CgpA in *Campylobacter jejuni. Mol Microbiol* **43:** 497-508.

Linton D, Dorrell N, Hitchen PG, Amber S, Karlyshev AV, Morris HR, Dell A, Valvano MA, Aebi M & Wren BW (2005) Functional analysis of the *Campylobacter jejuni* N-linked protein glycosylation pathway. *Mol Microbiol* **55:** 1695-1703.

Liu X, McNally DJ, Nothaft H, Szymanski CM, Brisson JR & Li J (2006) Mass spectrometry-based glycomics strategy for exploring *N*-linked glycosylation in eukaryotes and bacteria. *Anal Chem* **78:** 6081-6087.

Lizak C, Gerber S, Numao S, Aebi M & Locher KP (2011) X-ray structure of a bacterial oligosaccharyltransferase. *Nature* **474:** 350-355.

Logan SM (2006) Flagellar glycosylation - a new component of the motility repertoire? *Microbiology* **152:** 1249-1262.

Maita N, Nyirenda J, Igura M, Kamishikiryo J & Kohda D (2010) Comparative structural biology of Eubacterial and Archaeal oligosaccharyltransferases. *J Biol Chem* **285:** 4941-4950.

Maki M & Renkonen R (2004) Biosynthesis of 6-deoxyhexose glycans in bacteria. *Glycobiology* **14:** 1-15.

Marshall RD (1972) Glycoproteins. *Annu Rev Biochem* **41:** 673-702.

Mescher MF & Strominger JL (1976) Structural (shape-maintaining) role of the cell surface glycoprotein of *Halobacterium salinarium*. *Proc Natl Acad Sci USA* **73:** 2687-2691.

Messner P (2004) Prokaryotic glycoproteins: unexplored but important. *J Bacteriol* **186:** 2517-2519.

Mizanur RM, Zea CJ & Pohl NL (2004) Unusually broad substrate tolerance of a heat-stable archaeal sugar nucleotidyltransferase for the synthesis of sugar nucleotides. *Proc Natl Acad Sci USA* **9:** 13397-13402.

Mussar KJ, Murray GJ, Martin BM & Viswanatha T (1989) Peptide: *N*-glycosidase F: studies on the glycoprotein aminoglycan amidase from *Flavobacterium meningosepticum*. *J Biochem Biophys Meth* **20:** 53-68.

Nakagawa S, Takaki Y, Shimamura S, Reysenbach AL, Takai K & Horikoshi K (2007) Deep-sea vent ε-proteobacterial genomes provide insights into emergence of pathogens. *Proc Natl Acad Sci USA* **104:** 12146-12150.

Namboori SC & Graham DE (2008) Acetamido sugar biosynthesis in the euryarchaea. *J Bacteriol* **190:** 2987-2996.

Nasab FP, Schulz BL, Gamarro F, Parodi AJ & Aebi M (2008) All in one: *Leishmania major* STT3 proteins substitute for the whole oligosaccharyltransferase complex in *Saccharomyces cerevisiae*. *Mol Biol Cell* **19:** 3758-3768.

Nita-Lazar M, Wacker M, Schegg B, Amber S & Aebi M (2005) The NXS/T consensus sequence is required but not sufficient for bacterial *N*-linked protein glycosylation. *Glycobiology* **15:** 361-367.

Nothaft H, Liu X, McNally DJ & Szymanski CM (2010) *N*-linked protein glycosylation in a bacterial system. *Functional glycomics: methods and protocols* (Li J, ed), pp. 227-243. Humana Press, Clifton, NJ.

Nothaft H, Amber S, Aebi M & Szymanski CM (2008) *N*-linked protein glycosylation in Campylobacter. *Campylobacter* (Nachamkin I, Szymanski CM & Blaser MJ, eds), pp. 447-469. ASM Press, Washington, DC.

Nothaft H, Liu X, McNally DJ, Li J & Szymanski CM (2009) Study of free oligosaccharides derived from the bacterial *N*-glycosylation pathway. *Proc Natl Acad Sci USA* **106:** 15019-15024.

Nothaft H & Szymanski CM (2010) Protein glycosylation in bacteria: sweeter than ever Nat Rev Microbiol **8:** 765-778.

Palyada K, Threadgill D & Stintzi A (2004) Iron acquisition and regulation in *Campylobacter jejuni*. *J Bacteriol* **186:** 4714-4729.

Pathak R, Parker CS & Imperiali B (1995) The essential yeast *NLT1* gene encodes the 64 kDa glycoprotein subunit of the oligosaccharyl transferase. *FEBS Lett* **362:** 229-234.

Power PM, Roddam LF, Dieckelmann M, Srikhanta YN, Cheng Tan Y, Berrington AW & Jennings MP (2000) Genetic characterization of pilin glycosylation in *Neisseria meningitidis*. *Microbiology* **146:** 967-979.

Reiss G, Te Heesen S, Gilmore R, Zufferey R & Aebi M (1997) A specific screen for oligosaccharyltransferase mutations identifies the 9 kDa OST5 protein required for optimal activity *in vivo* and *in vitro*. *EMBO J* **16:** 1164-1172.

Roos J, Luz JM, Centoducati S, Sternglanz R & Lennarz WJ (1997) *ENP1*, an essential gene encoding a nuclear protein that is highly conserved from yeast to humans. *Gene* **185:** 137-146.

Rush JS, Gao N, Lehrman MA & Waechter CJ (2008) Recycling of dolichyl monophosphate to the cytoplasmic leaflet of the endoplasmic reticulum after the cleavage of dolichyl pyrophosphate on the lumenal monolayer. *J Biol Chem* **283:** 4087-4093.

Schoenhofen IC, McNally DJ, Vinogradov E, Whitfield D, Young NM, Dick S, Wakarchuk WW, Brisson JR & Logan SM (2006) Functional characterization of dehydratase/aminotransferase pairs from *Helicobacter* and *Campylobacter*. *J Biol Chem* **281:** 723-732.

Schwarz F, Lizak C, Fan YY, Fleurkens S, Kowarik M & Aebi M (2011) Relaxed acceptor site specificity of bacterial oligosaccharyltransferase in vivo. *Glycobiology* **21:** 45-54.

Sharma CB, Lehle L & and Tanner W (1981) *N*-glycosylation of yeast proteins. Characterization of the solubilized oligosaccharyltransferase. *Eur J Biochem* **116:** 101-108.

Silberstein S & Gilmore R (1996) Biochemistry, molecular biology, and genetics of the oligosaccharyltransferase. *FASEB J* **10:** 849-858.

Silberstein S, Collins PG, Kelleher DJ & Gilmore R (1995) The essential *OST2* gene encodes the 16-kD subunit of the yeast oligosaccharyltransferase, a highly conserved protein expressed in diverse eukaryotic organisms. *J Cell Biol* **131:** 371-383.

Spirig U, Glavas M, Bodmer D, Reiss G, Burda P, Lippuner V, Te Heesen S & Aebi M (1997) The STT3 protein is a component of the yeast oligosaccharyltransferase complex. *Mol Gen Genet* **256:** 628-637.

Spiro RG (2002) Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* **12:** 43-56.

Steen PV, Rudd PM, Dwek RA & Opdenakker G (1998) Concepts and principles of *O*-linked glycosylation. *Crit Rev Biochem Mol Biol* **33:** 151-208.

Stimson E, Virji M, Makepeace K, Dell A, Morris HR, Payne G, Saunders JR, Jennings MP, Barker S & Panico M (1995) Meningococcal pilin: a glycoprotein substituted with digalactosyl 2, 4-diacetamido-2, 4, 6-trideoxyhexose. *Mol Microbiol* **17:** 1201-1214.

Szymanski CM & Wren BW (2005) Protein glycosylation in bacterial mucosal pathogens. *Nat Rev Microbiol* **3:** 225-237.

Szymanski CM, Burr DH & Guerry P (2002) *Campylobacter* protein glycosylation affects host cell interactions. *Infect Immun* **70:** 2242-2244.

Szymanski CM, Yao R, Ewing CP, Trust TJ & Guerry P (1999) Evidence for a system of general protein glycosylation in *Campylobacter jejuni. Mol Microbiol* **32:** 1022-1030.

Te Heesen S, Knauer R, Lehle L & Aebi M (1993) Yeast Wbp1p and Swp1p form a protein complex essential for oligosaccharyl transferase activity. *EMBO J* **12:** 279-284.

Te Heesen S, Rauhut R, Aebersold R, Abelson J, Aebi M & Clark MW (1991) An essential 45 kDa yeast transmembrane protein reacts with anti-nuclear pore antibodies: purification of the protein, immunolocalization and cloning of the gene. *Eur J Cell Biol* **56:** 8-18.

Valvano MA (2008) Undecaprenyl phosphate recycling comes out of age. *Mol Microbiol* **67:** 232-235.

Varki A (1993) Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* **3:** 97-130.

Virji M, Saunders JR, Sims G, Makepeace K, Maskell D & Ferguson DJP (1993) Pilus-facilitated adherence of *Neisseria meningitidis* to human epithelial and endothelial cells: modulation of adherence phenotype occurs concurrently with changes in primary amino acid sequence and the glycosylation status of pilin. *Mol Microbiol* **10:** 1013-1028.

Voisin S, Houliston RS, Kelly J, Brisson JR, Watson D, Bardy SL, Jarrell KF & Logan SM (2005) Identification and characterization of the unique *N*-linked glycan common to the flagellins and S-layer glycoprotein of *Methanococcus voltae. J Biol Chem* **280:** 16586-16593.

Wacker M, Feldman MF, Callewaert N, Kowarik M, Clarke BR, Pohl NL, Hernandez M, Vines ED, Valvano MA & Whitfield C (2006) Substrate specificity of bacterial oligosaccharyltransferase suggests a common transfer mechanism for the bacterial and eukaryotic systems. *Proc Natl Acad Sci USA* **103:** 7088-7093.

Wacker M, Linton D, Hitchen PG, Nita-Lazar M, Haslam SM, North SJ, Panico M, Morris HR, Dell A & Wren BW (2002) *N*-linked glycosylation in *Campylobacter jejuni* and its functional transfer into *E. coli. Science* **298:** 1790-1793.

Weerapana E & Imperiali B (2006) Asparagine-linked protein glycosylation: from eukaryotic to prokaryotic systems. *Glycobiology* **16:** 91-101.

Willer T, Valero MC, Tanner W, Cruces J & Strahl S (2003) O-mannosyl glycans: from yeast to novel associations with human disease. *Curr Opin Struct Biol* **13:** 621-630.

Yan Q & Lennarz WJ (2002) Studies on the function of oligosaccharyl transferase subunits. Stt3p is directly involved in the glycosylation process. *J Biol Chem* **277:** 47692-47700.

Yan Q & Lennarz WJ (1999) Oligosaccharyltransferase: a complex multisubunit enzyme of the endoplasmic reticulum. *Biochem Bioph Res Co* **266:** 684-689.

Young NM, Brisson JR, Kelly J, Watson DC, Tessier L, Lanthier PH, Jarrell HC, Cadotte N, St. Michael F & Aberg E (2002) Structure of the *N*-linked glycan present on multiple glycoproteins in the Gram-negative bacterium, *Campylobacter jejuni. J Biol Chem* **277:** 42530-42539.

Yurist-Doutsch S, Chaban B, VanDyke DJ, Jarrell KF & Eichler J (2008) Sweet to the extreme: protein glycosylation in Archaea. *Mol Microbiol* **68:** 1079-1084.

Zeitler R, Hochmuth E, Deutzmann R & Sumper M (1998) Exchange of Ser-4 for Val, Leu or Asn in the sequon Asn-Ala-Ser does not prevent *N*-glycosylation of the cell surface glycoprotein from *Halobacterium halobium. Glycobiology* **8:** 1157-1164.

Zou W & Jennings HJ (2009) Preparation of glycoconjugate vaccines. *Carbohydrate-based vaccines and immunotherapies* (Guo Z & Boons GJ, eds), pp. 55-79. Wiley, Hoboken, NJ.

Zufferey R, Knauer R, Burda P, Stagljar I, Te Heesen S, Lehle L & Aebi M (1995) STT3, a highly conserved protein required for yeast oligosaccharyl transferase activity *in vivo*. *EMBO J* **14:** 4949-4960.

CHAPTER 2

Identification of important amino acid residues required for the function of the *Campylobacter* oligosaccharyltransferase for *N*-linked protein glycosylation

2.1. Introduction

Asparagine-(*N*)-linked protein glycosylation is one of the most prevalent protein modifications in eukaryotes. It involves the assembly of nucleotide-activated sugars on the dolichol-pyrophosphate lipid carrier at the cytoplasmic side of the endoplasmic reticulum (ER) (Burda & Aebi, 1999). The partially assembled glycan is flipped to the lumenal side of the ER membrane by an ATP-independent flippase where additional modification of the glycan occurs before it is transferred from the lipid carrier to asparagine residues within an N-X-S/T (X≠P) sequon of protein acceptors by the oliogosaccharyltransferase (OTase) (Marshall, 1972; Helenius *et al.*, 2002).

In yeast, the OTase is a multimeric enzyme that consists of at least eight different transmembrane protein subunits, each of which is required for enzymatic activity and where the STT3 subunit is the catalytic center (Yan & Lennarz, 1999). In lower eukaryotes, such as *Giardia* and kinetoplastids, the OTase is composed of a single membrane protein that consists of STT3 alone (Kelleher & Gilmore, 2006). The genomes of *Trypanosoma brucei* and *Leishmania major* reveal the presence of three and four *STT3* paralogous genes, respectively (Berriman *et al.*, 2005; Ivens *et al.*, 2005). Homologous processes of *N*-linked protein glycosylation have been found outside of the Eukaryotic domain, in both Archaea and Bacteria (Mescher *et al.*, 1976; Szymanski *et al.*, 1999). Similarly, the OTase from the thermophilic archaeon, *Pyrococcus furiosus*, is also composed of the STT3 protein alone and is capable of transferring glycans onto peptide acceptors (Kohda *et al.*, 2007).

The bacterial process of *N*-glycosylation shares many features with its eukaryotic and archaeal counterparts, but significant differences exist. In *C. jejuni*, *N*-glycosylation is a post-translational process occurring in the periplasm. In this pathway, nucleotide-activated sugars are synthesized in the cytoplasm and assembled onto an undecaprenyl-pyrophosphate-linked carrier forming a heptasaccharide at the cytoplasmic side of the inner membrane, where they are then translocated to the periplasm by an ATP-dependent flippase (Alaimo *et al.*, 2006; Linton *et al.*, 2005) and subsequently transferred from the lipid carrier to asparagine residues within the extended D/E-X_1-N-X_2-S/T consensus sequence, where X_1 and X_2 are any amino acid except proline (Kowarik *et al.*, 2006b). The central enzyme responsible for the transfer reaction is the OTase, PglB, encoded in the protein glycosylation (*pgl*) locus (Szymanski *et al.*, 1999). The *C. jejuni pgl* gene locus contains all of the genes necessary for the *N*-glycosylation process, which has been functionally reconstituted in *E. coli* to produce recombinant glycoproteins (Wacker *et al.*, 2002). In addition to *N*-glycosylation, it has also been demonstrated that PglB releases free oligosaccharides (fOS) into the periplasm in response to changes in the osmolarity of the environment and bacterial growth phase (Nothaft *et al.*, 2009).

The *N*-glycosylation pathway has also been identified in an increasing number of bacteria. For yet undetermined reasons, the pathway is exclusive to the epsilon and delta classes of proteobacteria. The *pgl* gene orthologues are conserved in all *Campylobacter* species sequenced to date, the delta-proteobacterium *Desulfovibrio desulfuricans* (Ielmini & Feldman, 2011), and in

the epsilon-proteobacteria *Wolinella succinogenes* (Baar *et al.*, 2003), *Sulfurovum* sp., *Nitratiruptor* sp. (Nothaft *et al.*, 2008), and certain *Helicobacter* species (Jervis *et al.*, 2010; Nothaft & Szymanski, 2010). It has not yet been shown that the *N*-glycosylation pathway is functional in all of these organisms, but they each contain the key enzyme, PglB. Therefore, it is predicted that these bacteria also *N*-glycosylate their proteins.

The PglB of *C. jejuni* (*Cj*PglB) is the best studied bacterial OTase to date. It is an 82 kDa single-subunit OTase which shares sequence similarity and structural organization with the STT3 subunits of eukaryotic OTase complexes (Szymanski *et al.*, 2005). It comprises 11-13 predicted hydrophobic N-terminal transmembrane domains spanning the periplasmic membrane and a soluble C-terminal domain oriented toward the periplasmic space (Li *et al.*, 2010). All OTase orthologues in all three domains of life have the signature WWDXG motif, which has been shown to be necessary for *Cj*PglB activity (Wacker *et al.*, 2002). Recently, the crystal structure of the full length PglB of *Campylobacter lari* (*Cl*PglB) has been solved, which revealed that the side chains of the serine or threonine at the +2 position of the acceptor consensus sequence interact with those of the WWD residues to form stable hydrogen bonds resulting in acceptor sequon binding (Lizak *et al.*, 2011). A crystal structure of the soluble C-terminal domain only exists for the *Cj*PglB and through structural comparisons with the archaeal AglB revealed a conserved Ile 571 within an MXXI motif (Maita *et al.*, 2010). It is suggested that the isoleucine in homologous proteins can provide

contact to the +2 Thr to aid in protein binding and/or recognition (Lizak *et al.*, 2011).

Here, we attempt to identify other amino acids that are important for the function and/or structure of *Cj*PglB. We began by analyzing *Campylobacter* PglB protein sequences through multiple sequence alignments. We have identified a DXXK sequence located close to the highly conserved WWDXG motif. The DXXK motif is conserved in all epsilon-proteobacteria and is present in at least one PglB orthologue for species that contain two PglB enzymes. We examined whether this highly conserved DXXK sequence is necessary for PglB activity. Further downstream of the DXXK motif, we found the *Campylobacter* PglB protein sequences to contain either a proline, serine or valine residue at a single position in the alignment. Since proline is known to disrupt protein structures more than serine or valine residues due to the presence of its bulky side chain, we wanted to identify the structural importance of these amino acids in this region of PglB. We began by comparing the PglB enzymes of *C. jejuni* and *C. lari*. The *N*-glycosylation system in *C. lari* is encoded by a similar *pgl* operon, which lacks a glucosyltransferase normally present in the *C. jejuni pgl* operon leading to the production of a linear hexasaccharide that is transferred onto acceptor proteins (Schwarz *et al.*, 2011). Thus, we chose to compare the importance of Pro 487 in *Cj*PglB with the corresponding residue in *Cl*PglB, Ser 493. These data identify new amino acids that are important in the structure and catalysis of PglB.

2.2. Materials and Methods

2.2.1. *Bacterial strains, plasmids, and growth conditions*

The bacterial strains and plasmids used in this study are listed in Table 2.1. *Escherichia coli* strains were grown in Luria-Bertani (LB) broth or on LB agar plates (1% Bacto™ Tryptone, 0.5% Bacto™ Yeast Extract, 1% NaCl and 1.5% agar was added for plates) at 37°C. *Campylobacter jejuni* NCTC 11168 and 81-176 strains were grown on Mueller Hinton (MH, Difco™) agar plates under microaerobic conditions (10% CO_2, 5% O_2, 85% N_2) at 37°C for 18 hrs. Ampicillin (100 ug/mL), chloramphenicol (25 ug/mL), kanamycin (25 ug/mL), trimethoprim (25 ug/mL) and tetracycline (25 ug/mL) were added to the medium as needed for selection.

2.2.2. *Construction of plasmids*

Amino acid substitutions in the *C. jejuni* and *C. lari pglB* genes were generated using oligonucleotides listed in Table 2.2. For the *C. jejuni pglB*, the single point mutants D475V and K478V were introduced using the QuikChange Site-Directed Mutagenesis Kit (Stratagene) with plasmid pMAF10 as template. All remaining single and double *pglB* point mutants were introduced following a modified three-piece ligation method described by Kato *et al.* (2008). Two PCR fragments of the $pglB_{Cj}$ gene were amplified using *Pfx* polymerase (Invitrogen) with pMAF10 as template. The N-terminal fragment of $pglB_{Cj}$ was amplified using oligonucleotide pglB*BamH*I-F with each of the oligonucleotides carrying the amino acid substitutions within the DXXK motif: D475A, K478A, AXXA, VXXV, KXXK, and KXXD. The C-terminal fragment of $pglB_{Cj}$ was amplified using

Table 2.1. Bacterial strains and plasmids used in this study

Strain or plasmid	Characteristics	Source
E. coli		
DH5α	F⁻ *endA1 hsdR17 supE44 thi-1 recA1* Δ (*argF-lacZYA*)*U169* (80d *lacZ* Δ *M15*) *gyrA96* λ^-	Invitrogen
CLM24	W3110, Δ*waaL*	Feldman *et al.*, 2005
C600 (RK212.2)	*leu thr thi lacy supE44 tonA;* pRK212.2, Amp^R, Tet^R	Figurski & Helinski, 1979
C. jejuni		
81-176	Clinical isolate	Korlath *et al.*, 1985
11168 NCTC	Clinical isolate used for genome sequencing	Parkhill *et al.*, 2000
11168-*pglB*::Kan	*pglB* mutant, Km^R	Nothaft *et al.*, 2009
11168 *pglB*(pCj-*pglB*)	*pglB* mutant complemented with pCE111-28 carrying *pglB*, Kan^R, Cm^R	Nothaft *et al.*, 2009
11168-*pglB*(pCj-*pglB*$_{mut}$)	*pglB* mutant complemented with pCE111-28 carrying *pglB*$_{mut}$, Kan^R, Cm^R	Nothaft *et al.*, 2009
Plasmids		
pACYC(*pgl*)	Encodes the *C. jejuni pgl* locus, Cm^R	Wacker *et al.*, 2002
pACYC(*pgl*-Δ*pglB*)	Encodes the *C. jejuni pgl* locus with a deletion in PglB, Cm^R	Provided by Bernadette Beadle
pWA2	Soluble periplasmic *C. jejuni acr*A_{6His} under control of Tet promoter, in pBR322, Amp^R	Feldman *et al.*, 2005
pMLBAD	Cloning vector, arabinose-inducible, Tmp^R	Lefebre & Valvano, 2002
pMAF10	HA-tagged *C. jejuni pglB* cloned in pMLBAD, Tmp^R	Feldman *et al.*, 2005
pWA1	HA-tagged *C. jejuni pglB*$_{mut}$ cloned in pMLBAD, Tmp^R	Feldman *et al.*, 2005
pET24(*Cl*pglB)	His-tagged *C. lari pglB* cloned in pET24, Kan^R	Provided by Abofu Alemka
pYB1D	*C. jejuni pglB*(D475A)$_{HA}$ cloned into pMLBAD, Tmp^R	This study
pYB2D	*C. jejuni pglB*(D475V)$_{HA}$ cloned into pMLBAD, Tmp^R	This study
pYB1K	*C. jejuni pglB*(K478A)$_{HA}$ cloned into pMLBAD, Tmp^R	This study
pYB2K	*C. jejuni pglB*(K478V)$_{HA}$ cloned into pMLBAD, Tmp^R	This study
pYB1A	*C. jejuni pglB*(D475A, K478A)$_{HA}$ cloned into pMLBAD, Tmp^R	This study
pYB1V	*C. jejuni pglB*(D475V, K478V)$_{HA}$ cloned into pMLBAD, Tmp^R	This study
pYB3K	*C. jejuni pglB*(D475K)$_{HA}$ cloned into pMLBAD, Tmp^R	This study
pYB4K	*C. jejuni pglB*(D475K, K478D)$_{HA}$ cloned into pMLBAD, Tmp^R	This study
pYBCJP	*C. jejuni pglB*(P487S)$_{HA}$ cloned into pMLBAD, Tmp^R	This study
pYBCLwt	*C. lari pglB*$_{HA}$ cloned into pMLBAD, Tmp^R	This study
pYBCLS	*C. lari pglB*(S493P)$_{HA}$ cloned into pMLBAD, Tmp^R	This study
pCE111-28	*C. jejuni* expression vector, plasmid pRY111 with σ^{28} promoter of *flaA*, Cm^R	Larsen *et al.*, 2004
pCj-*pglB*D	*C. jejuni pglB*(D475V) cloned in pCE111-28, Cm^R	This study
pCj-*pglB*K	*C. jejuni pglB*(K478V) cloned in pCE111-28, Cm^R	This study
pCj-*pglB*P1	*C. jejuni pglB*(P487S) cloned in pCE111-28, Cm^R	This study
pCl-*pglB*wt	*C. lari pglB* cloned in pCE111-28, Cm^R	This study
pCl-*pglB*S1	*C. lari pglB*(S493P) cloned in pCE111-28, Cm^R	This study

Table 2.2. Oligonucleotides used in this study

Oligonucleotide	Sequence 5'-3'	Purpose
D475V-F	CGATGTGAAAACTTTAGTAGTTGGTGGAAAGCATTTAG GTAAGG	Mutagenesis
D475V-R	CCTTACCTAAATGCTTTCCACCAACTACTAAAGTTTTCA CATCG	Mutagenesis
K478V-F	CTTTAGTAGATGGTGGAGTACATTTAGGTAAGGATAATT TTTTCC	Mutagenesis
K478V-R	GGAAAAAATTATCCTTACCTAAATGTACTCCACCATCTA CTAAAG	Mutagenesis
D475A-R	CCTTACCTAAATGTTTTCCACCCGCTACTAAAGTTTTCA CATCG	Mutagenesis
K478A-R	CCTTACCTAAATGCGCTCCACCATCTACTAAAGTTTTCA CATCG	Mutagenesis
AXXA-R	CCTTACCTAAATGCGCTCCACCCGCTACTAAAGTTTTCA CATCG	Mutagenesis
VXXV-R	CCTTACCTAAATGTACTCCACCAACTACTAAAGTTTTCA CATCG	Mutagenesis
KXXK-R	CCTTACCTAAATGTTTTCCACCTTTTACTAAAGTTTTCAC ATCG	Mutagenesis
KXXD-R	CCTTACCTAAATGATCTCCACCTTTTACTAAAGTTTTCA CATCG	Mutagenesis
pglB-C1-F	5Phos-ATAATTTTTTCCCTTCTTTTGCTTTAAGCAAAGATG	Mutagenesis
pglB-C2-F	5Phos-AAAGATGAACAAGCTGCAGCTAATATGG	Mutagenesis
Cj-P487S-R	GCTTAAAGCAAAAGATGAGAAAAAATTATCCTTACC	Mutagenesis
pglBHA*Kpn*I-R	ATATACTCGAGGTACCATGGTTAAGCGTAATCTGGAAC ATCG, *Kpn*I site underlined	Mutagenesis
pglB*Bam*HI-F	ATTAGCGGATCCTACCTGACGCTTTTTATCGC, *Bam*HI site underlined	Mutagenesis PCR, cloning
pglB*Xho*I-R	TATACTCGAGTTAAGCGTAATCTGGAACATCGTATGG, *Xho*I site underlined	PCR, cloning
pMAF10*Nde*I-1	CAGTATTTATCCATATGTACAACTATAAAG	Sequencing
pETClpglB-F	AATTCTGCAGTAATTTTGTTTAACTTTAAGAAGG	Mutagenesis
Cl-S493P-R	GCTTAAGACAAAAGAAGGGAAAAAATTATCTTTTCC	Mutagenesis
pglB-C3-F	5Phos-AAAGAACAAATTCCAGCAGCCAATATGG	Mutagenesis
ClpglBHA-R	AATTTAAGCTTTGTTAGCAGCCGGATCTCAGTGG	Mutagenesis Sequencing
ClpglB*EcoR*I-F	ATGAATTCACCATGAAACTACAACAAAATTTCACG	PCR, cloning
ClpglB*Pst*I-R	AGCTGCAGCTAAGCGTAATCTGGAACATCGTATGGGTA TCTTTTTAGCCTATAAATTTTTGC, *Pst*I site underlined	PCR, cloning
ClpglB*BamH*I-F	AAGGATCCAATAATTTTGTTTAACTTTAAGAAGG, *Bam*HI site underlined	PCR, cloning
ClpglB*Xho*I-R	TGCTCGAGCTAAGCGTAATCTGGAACATCGTATGGGTA TCTTTTTAGCCTATAAATTTTTGC, *Xho*I site underlined	PCR, cloning

*underlined amino acids indicate mutated sites within *pglB* unless otherwise indicated

oligonucleotides pglB-C1-F and pglBHA*Kpn*I-R. The N-terminal and C-terminal *pglB* fragments were digested with *EcoR*I and *Kpn*I, respectively, to generate blunt-ended PCR products that were cloned in the same sites of the pMLBAD empty vector.

The *Cj*PglB(P487S) and *Cl*PglB(S493P) single point mutants were also constructed using the three-piece ligation method. The oligonucleotides pglB*BamH*I-F and Cj-P487S-R were used to amplify the N-terminal fragment of $pglB_{Cj}$, while oligonucleotides pglB-C2-F and pglBHA*Kpn*I-R amplified the C-terminal fragment using pMAF10 as template and digesting the PCR products with *EcoR*I or *Kpn*I to clone in pMLBAD. The N-terminal fragment of $pglB_{Cl}$ was amplified using oligonucleotides pET-ClpglB-F and Cl-S493P-R, while the C-terminal fragment of $pglB_{Cl}$ was amplified with oligonucleotides pglB-C3-F and ClpglB-HA-R using pET24(*Cl*pglB) as template. The blunt-ended PCR products were then digested with either *Nde*I or *Xho*I and cloned in the same sites of the pET24b empty vector. The $pglB_{Cl}$ was then cloned in pMLBAD using oligonucleotides ClpglB*EcoR*I-F and ClpglB*Pst*I-R. Oligonucleotides pglBHA*Kpn*I-R and ClpglBHA-R encode an HA-tag to assist in following *C. jejuni* and *C. lari* PglB protein expression when analyzed by Western blot.

PCR products and empty plasmid vectors were digested with FastDigest® restriction enzymes (Fermentas) for 1 hr at 37°C and subsequently purified using the QIAquick® Gel Extraction Kit (Qiagen) following the manufacturer's instructions. A blunt-end ligation was carried out using the T4 ligase (Fermentas) with a vector:insert ratio of 1:5 (10-20 ng of vector DNA) and incubated

overnight at 16°C (Sambrook & Russell, 2001). After transformation into *E. coli* DH5α, the resulting plasmids (Table 2.1) were isolated using the GeneJET™ Plasmid Miniprep Kit (Fermentas) and confirmed by restriction analysis. DNA sequencing was carried out using the BigDye® Terminator v3.1 Cycle Sequencing Kit following the manufacturer's protocol (Applied Biosystems). Oligonucleotides pMAF10*Nde*I-1 and ClpglBHA-R were used as the sequencing primers specific for the *pglB* genes of *C. jejuni* and *C. lari*, respectively. Cycle sequencing on the ABI PRISM® 377 DNA Sequencer was performed by the Molecular Biology Service Unit (University of Alberta) to sequence the DNA. The mutant *pglB* sequences were confirmed using the Geospiza FinchTV software.

For genetic manipulations of *C. jejuni* 11168, the *pglB* point mutants were amplified by PCR with oligonucleotides pglB*Bam*HI-F and pglB*Xho*I-HA-R using Vent polymerase (NEB) and the respective pMLBAD plasmids carrying the various *pglB* mutations as template. The PCR products were digested with *Bam*HI and *Xho*I and subcloned into the *C. jejuni-E. coli* shuttle vector pCE111-28 treated with the same enzymes. After transformation into *E. coli* DH5α, the resulting plasmids (Table 2.1) were confirmed by restriction analysis and re-transformed into *E. coli* pRK212.2. The shuttle plasmids were mobilized from *E. coli* pRK212.2 to *C. jejuni* 81-176 and then complemented into a *C. jejuni* 11168 *pglB* mutant background by natural transformation as previously described (Labigne-Roussel *et al.*, 1987; Nothaft *et al.*, 2010).

2.2.3. *Expression of PglB point mutants in* E. coli *and purification of AcrA*

Plasmids containing the following constructs were co-expressed in *E. coli* CLM24: pACYC(*pgl-ΔpglB*) encoding the *C. jejuni* protein glycosylation *(pgl)* operon with a deleted PglB, pMLBAD carrying the various *pglB* alleles, and pWA2 encoding a His6-tagged *C. jejuni* glycoprotein AcrA which served as the *N*-glycosylation acceptor. Starting from a 5 mL overnight culture, *E. coli* CLM24 strains were grown in LB broth at 37°C to an OD_{600} of 0.5-0.6 and induced with 0.2% L-(+)-arabinose (wt/vol). Cells were harvested by centrifugation (3,696 x *g*, 15 min, 4°C) after a 4 hr induction. Whole cell lysates were prepared by the resuspension of cells in phosphate-buffered saline (PBS), pH 7.2, followed by sonication (3 × 1 min) with a Branson sonicator equipped with a microtip and centrifugation (16,260 x *g*, 20 min, 4°C). PglB proteins were visualized by Western blotting using HA-tag specific antibodies (Santa Cruz). For AcrA purification, whole cell lysates were adjusted to the same protein concentration, subjected to nickel affinity chromatography (NTA agarose, Qiagen), washed with 10 column volumes of 10 mM imidazole in 1xPBS and washed again with 10 column volumes of 50 mM imidazole in 1xPBS before eluting the samples with 500 uL of 500 mM imidazole in 1xPBS. Protein concentrations were obtained using the Quick Start™ Bradford Protein Assay Kit (Bio-Rad) and measured from a standard curve using known concentrations of bovine serum albumin (Bio-Rad). The glycosylation status of purified AcrA was analysed by Western blotting using AcrA-specific (Wacker *et al.*, 2002) and *C. jejuni-N*-glycan specific (hR6) antibodies (Schwarz *et al.*, 2011).

2.2.4. *Preparation of* C. jejuni *whole cell lysates*

Campylobacter whole cell extracts were prepared as previously described (Nothaft *et al.*, 2010). Briefly, *C. jejuni* 11168 NCTC strains (wildtype, *pglB* mutant, and *pglB* mutant complemented with wildtype *pglB*, inactive *pglB*, or the varioius $pglB_{DXXK}$ mutants) were thawed on MH plates from frozen glycerol stocks (-80°C), restreaked onto full plates of MH agar with appropriate antibiotics for selection, and incubated at 37°C under microaerobic conditions. Cells from two plates were harvested by centrifugation (4500 x *g*, 10 min, 4°C), resuspended with 1 mL of 1xPBS (pH 7.2), and disrupted by sonication (3 x 30 sec). Unbroken cells were removed by centrifugation (16,100 x *g*, 30 min, 4°C) and the sample supernatants were kept as whole cell lysates. Protein concentrations were determined using the NanoVue™ Plus Spectrophotometer (A_{280}) and samples were all adjusted to the same concentration for analysis. The glycosylation status of AcrA was analyzed by Western blotting using AcrA-specific antibodies.

2.2.5. *Analysis of fOS production*

Whole cell lysates from the *C. jejuni* wildtype and *pglB* mutants were adjusted to a final concentration of 5 mg/mL with 1xPBS (pH 7.2) and then freeze-dried overnight. The relative free glycan amounts from each *C. jejuni* sample were investigated as previously described (Liu *et al.*, 2006; Nothaft *et al.*, 2009), but without Pronase E digestion (in collaboration with Dr. Liu and Dr. Li, National Research Council, Ottawa). Briefly, lyophilized samples were dissolved in 1xPBS (pH 7.2) and permethylated by the addition of methyl iodide. The permethylated oligosaccharides were extracted in chloroform, washed with water, dried and

dissolved in 75% aqueous methanol. The production of fOS in each of the samples was analyzed by using an electrospray tandem mass spectrometer, the 4000 Q-Trap® mass spectrometer (Applied Biosystems/MDS Sciex), coupled to a capillary electrophoresis interface. The mass spectra were acquired with dwell times of 2.0 ms per step of 0.5 *m/z* unit in full mass scan mode. The relative amounts of free glycans were compared based on the peak area of the extracted electropherogram for each corresponding ion. Sample preparation and fOS analysis were repeated in triplicate and the average relative quantities were used.

2.2.6. *Immunodetection by Western blotting*

Western blotting was performed as described previously (Nothaft *et al.*, 2010). HA-specific, AcrA-specific or *C. jejuni*-*N*-glycan-specific (hR6) polyclonal antisera served as primary antibodies at 1:1000, 1:10,000 and 1:10,000 dilutions, respectively, for 45 min in 1% skim milk in 1xPBS with 0.05% (w/v) Tween 20 (PBST). Blots were washed three times with PBST and incubated in alkaline phosphatase conjugated anti-rabbit IgG (Santa Cruz) for 45 min, which served as the secondary antibody at 1:2,000 dilution. After washing three times with PBST, membranes were then incubated in alkaline phosphatase colour development buffer supplemented with NBT/BCIP solution (Roche).

2.3. Results

2.3.1. *Site-directed mutagenesis and expression of PglB enzymes*

Multiple sequence alignment of the *Campylobacter* PglB OTases identified a DXXK motif that is located 12 amino acids away from the highly conserved WWDXG motif (Wacker 2002). *In silico* analysis also showed that the DXXK

motif is conserved in all ε-proteobacteria (Fig. 2.1). To analyze the importance of these amino acids for PglB function, we substituted the D and K residues to generate the following single and double point mutants: AXXK, DXXA, AXXA, VXXK, DXXV, VXXV, KXXK and KXXD. Amino acid substitutions were confirmed by sequencing. Mutated *pglB* alleles were expressed in *E. coli* and PglB enzymes were visualized by Western blotting with HA-tag specific antibodies showing that the introduction of these point mutations had no effect on PglB expression and that the proteins were produced at the same level (Fig. 2.2).

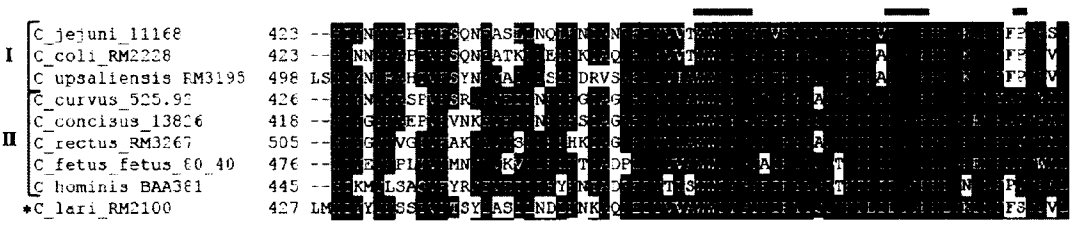

Figure 2.1. Multiple sequence alignment of *Campylobacter* PglB protein sequences. The ^{475}DXXK^{478} sequence in *C. jejuni* 11168 is located close to the highly conserved ^{457}WWDYG^{461} motif and is conserved in all epsilonproteobacteria. Downstream of the DXXK motif (position 487 in *C. jejuni* 11168) all *Campylobacter* PglB sequences possess a proline or valine, which appears to correlate with the groups I and II glycan structure divisions. The *C. lari* PglB sequence is marked with an asterisk indicating the presence of a serine at the aligned site.

2.3.2. *Investigation of the DXXK motif in PglB expressed in* E. coli

To analyze the PglB activity of the point mutants in the DXXK motif, we co-expressed the *pglB* alleles in *E. coli* CLM24 with plasmids carrying the *C. jejuni*

pgl operon with an inactive PglB and a soluble form of the His6-tagged *C. jejuni* glycoprotein AcrA, which served as the *N*-glycosylation acceptor. The glycosylation status of purified AcrA was analyzed by Western blotting using AcrA-specific and *C. jejuni-N*-glycan specific antibodies (Fig. 2.2). The wildtype PglB encoded in pMAF10 or the inactive PglB carried on pWA1 were also introduced in *E. coli* cells harboring the *C. jejuni pgl* locus and $AcrA_{6His}$ to serve as positive and negative controls, respectively. Expression of the wildtype PglB, but not the $PglB_{mut}$, resulted in fully glycosylated AcrA as previously shown (Feldman *et al.*, 2005). Loss of AcrA glycosylation was observed with the AXXA, VXXV, KXXK and KXXD mutations, but not most of the single alanine mutations with the exception of the D475V single point mutation, which showed reduced glycosylation compared to the K478V mutation.

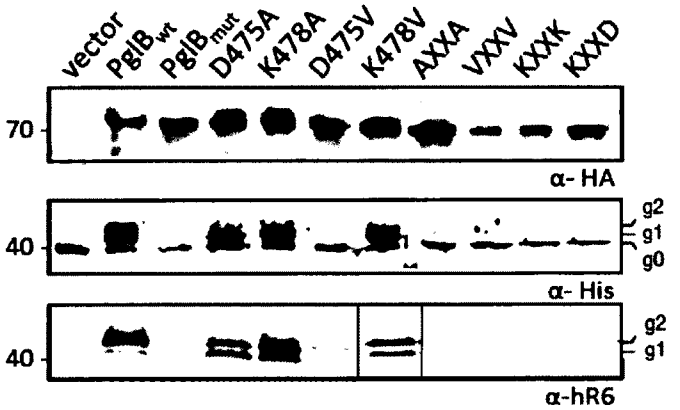

Figure 2.2. Amino acid exchanges of D and K residues were made to examine the importance of this sequence in *Cj*PglB in *E. coli.* PglB expression and activity were visualized by Western blotting with HA, His and *N*-glycan specific (hR6) antibodies as indicated. Reduction of AcrA glycosylation was observed with *pglB* alleles that carried the D457V, AXXA, KXXK, and KXXD point mutations.

Molecular masses are indicated in kDa on the left of each blot. Glycosylation state of AcrA is indicated on the right of the α-His and α-hR6 blots.

2.3.3. *Activity of PglB DXXK point mutants in* C. jejuni *and analysis of fOS production*

In vivo PglB activity was also investigated in the native host by complementation of a chromosomal *C. jejuni pglB* knockout mutant with the D475V and K478V *pglB* point mutants. The importance of the DXXK motif in *Cj*PglB was examined by analyzing the glycosylation status of AcrA by Western blotting using AcrA-specific antibodies (Fig. 2.3). The results obtained for the D475V mutant is comparable to that in *E. coli*, where glycosylation of AcrA was dramatically reduced. The K478V mutant in *C. jejuni* shows a decrease in glycosylation, while the K478V mutant expressed in *E. coli* showed no affect in PglB activity.

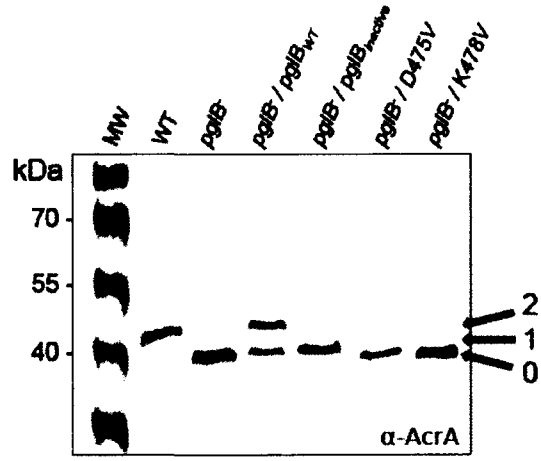

Figure 2.3. Complementation of *C. jejuni pglB* knock-out with the DXXK *pglB* point mutants. PglB activity was visualized by Western blotting with *C. jejuni* AcrA specific antibodies. The D475V and K478V point mutants showed a

reduction in glycosylation compared to the wildtype complement. Molecular masses and AcrA glycosylation state are indicated.

We also examined whether the DXXK motif is necessary for release of free oligosaccharides (fOS), a novel hydrolase function of PglB recently discovered (Nothaft *et al.*, 2009). Semi-quantitative mass spectrometry analyses of fOS levels in the *C. jejuni* PglB D to V and K to V single point mutants showed a corresponding decrease of this pathway product (Fig. 2.4).

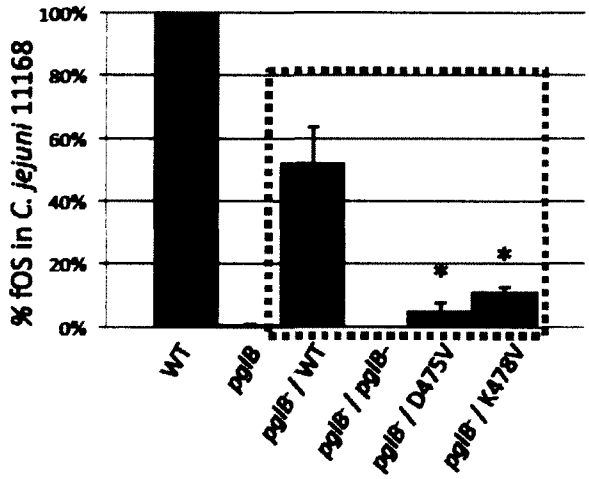

Figure 2.4. The amount of fOS release in *C. jejuni* was determined using semi-quantitative mass spectrometry. The complemented *C. jejuni pglB* mutant shows only partial complementation of fOS production (and *N*-glycosylation) (Nothaft *et al.*, 2009). However, complementation of the *pglB* mutant with the point mutations demonstrates that perturbation of PglB function affects fOS release. Results obtained are from biological triplicates; asterisks indicate $p<0.05$ relative to the complemented *C. jejuni pglB* mutant (indicated by the boxed area for comparison).

2.3.4. *Analysis of the C. jejuni and C. lari PglB proline / serine point mutants*

*Cj*PglB was also compared to *Cl*PglB through the exchange of a single amino acid in the soluble C-terminal domain close to the DXXK sequence. The proline residue in position 487 in *Cj*PglB was swapped with the corresponding serine residue in position 493 in *Cl*PglB. Amino acid substitutions were confirmed by sequencing and mutated *pglB* alleles were co-expressed in *E. coli* with $AcrA_{6His}$ and the *C. jejuni pgl* or *C. lari pgl* operon with a deleted *pglB* gene. Western blot analysis of AcrA glycosylation using anti-AcrA antibodies shows no considerable difference in the *Cj*PglB activity in *E. coli*, but shows an increase in glycosylation efficiency for the amino acid substitution in the *C. lari* PglB mutant as compared to the wildtype PglB sequence (Fig. 2.5A). In contrast, complementation of a chromosomal *C. jejuni pglB* knock-out mutant with the *Cj*PglB(P487S) point mutant shows improved glycosylation efficiency, whereas the *Cl*PglB wildtype and S493P point mutant show similar glycosylation levels to the *Cj*PglB wildtype complement (Fig. 2.5B).

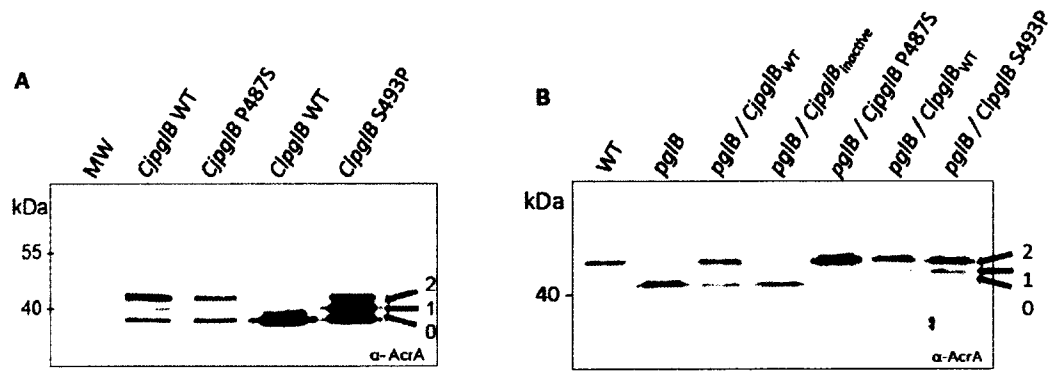

Figure 2.5. Amino acid exchanges of proline and serine residues were made to examine the importance of these sites in PglB. **A)** Mutated *pglB* alleles were co-expressed in *E. coli* with the *C. jejuni pgl* operon and $AcrA_{6His}$. PglB activity was

visualized by Western blotting with AcrA specific antibodies. Amino acid substitution of the *C. lari* PglB improves glycosylation efficiency, but not for the *C. jejuni* PglB. **B)** Complementation of a chromosomal *C. jejuni pglB* knock-out mutant with the *C. jejuni* PglB point mutant shows improved glycosylation efficiency, whereas the *C. lari* PglB wildtype and point mutant expression in *C. jejuni* show similar glycosylation levels comparable to the wildtype complement. Molecular masses and glycosylation state of AcrA are indicated.

2.4. Discussion

N-linked protein glycosylation is present in all domains of life where several prominent similarities and differences in the pathways are found. The defining event is the catalytic transfer of a glycan moiety onto a protein acceptor catalyzed by the OTase. Comparison of the processes provides further understanding of the common and distinct properties of the OTases. The archaeal *aglB* and bacterial *pglB* genes share sequence similarity with the *STT3* gene that encodes the largest subunit of the eukaryotic OTase complex, where a conserved WWDXG motif can be found within the soluble C-terminal domain (Wacker *et al.*, 2002; Yan & Lennarz, 2002).

A structure-guided sequence alignment that included the yeast *STT3* and the archaeal *Pyrococcus furiosus aglB* sequences revealed a ^{571}DXXK^{574} sequence that was demonstrated through mutagenesis studies to be essential for OTase activity in yeast (Igura *et al.*, 2008). The 2.8 Å resolution crystal structures of the C-terminal globular domains of *Cj*PglB and *P. furiosus* AglB were compared and revealed that the counterpart of the DXXK sequence in AglB is ^{568}MXXI^{571} in

PglB, and that OTase activity *in vitro* was reduced resulting from an I571A substitution (Maita *et al.*, 2010). In addition, a highly conserved aspartic acid exists in the first loop of the transmembrane region in the OTases from all domains of life (residue 54 in *Cj*PglB). Alanine mutations of residue D54 in *Cj*PglB resulted in nearly complete loss of PglB activity *in vitro* (Maita *et al.*, 2010). Recently, the crystal structure of the full-length *Cl*PglB has been solved, which provides a molecular explanation for the requirement of a Ser or Thr residue at the +2 position of the consensus sequence for *N*-linked glycosylation (Lizak *et al.*, 2011). The peptide-bound structure of PglB reveals that the β-hydroxyl group of the +2 Thr forms hydrogen bonds with each of the side chains of the WWD residues. The structure also shows that I571 provides contact to the +2 Thr; but since the Ile within the MXXI motif is not conserved in all STT3 homologues, other residues can aid in protein binding.

In this study, alignment of all *pglB* gene sequences from the epsilon-proteobacteria revealed a second conserved ^{475}DXXK^{478} sequence. In the *Cj*PglB, this sequence motif is located 12 amino acids away from the highly conserved WWDXG motif. Site-directed mutagenesis of the D and K amino acids was carried out to show the importance of these residues in OTase function. The glycosylation status of the *C. jejuni* glycoprotein AcrA was analyzed *in vivo*. Single alanine mutations of D475 and K478 in PglB showed no considerable difference in glycosylation activity compared to wildtype PglB. In addition, single valine substitutions were made to the D and K residues in PglB in order to examine the affects of hydrophobic residues at these sites, which resulted in

reduced glycosylation efficiency of D475V, but not K478V in the heterologous *E. coli* host. Recently, the DXXK motif in *Cj*PglB has also been analyzed by constructing only single point mutations in the D and K sites (Jaffee & Imperiali, 2011). The authors concluded that the 'DXXK motif' does not play an essential role in OTase activity, except that the D475 residue is likely involved in PglB function since a more significant effect was observed when this site was mutated compared to K478. However, we found that double alanine or double valine mutations at both the D and K sites resulted in the complete loss of AcrA glycosylation, an effect more drastic than the single D475 substitution. These data suggest that both D and K residues are important for the glycosylation activity of PglB, although the D475V mutation appears to destabilize the structure more than the K478V mutation. Also, consistent with our findings that PglB is responsible for fOS production in *C. jejuni* (Nothaft *et al.*, 2009), the D475V and K478V point mutants displayed a reduction in fOS release when complemented in a chromosomal *C. jejuni* PglB knock-out mutant. In addition, the OTase activity of these PglB point mutants in the native *C. jejuni* host were reduced, which compare to the levels observed for fOS release.

Analysis of the available crystal structure of the soluble portion of *Cj*PglB (in collaboration with Dr. Ken Ng, University of Calgary) revealed that the ^{475}DXXK^{478} motif is located directly above the highly conserved WWDYG motif (Fig. 2.6A). The D and K residues do not interact with each other, but appear to stabilize the structure through electrostatic interactions with R465 and D483, respectively, forming salt bridges (Fig. 2.6B). Loss of the two wild type salt

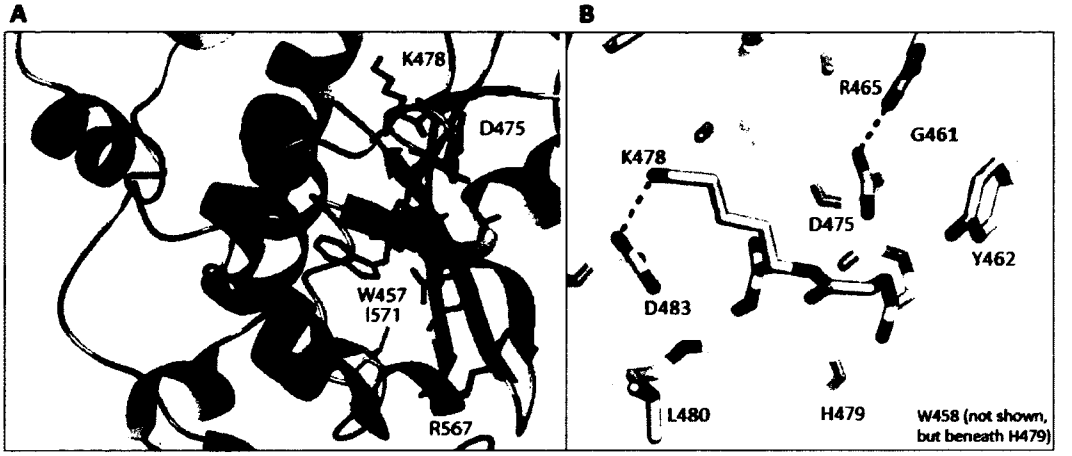

Figure 2.6. A) Close-up view of the conserved WWDYG motif in *C. jejuni* PglB. The ^{475}DXXK^{478} motif (green) is located immediately above the ^{457}WWDYG^{461} motif (orange). **B)** Highlighted stick residues are shown to interact to form salt bridges. Two salt bridges (D475:R465 and K478:D483) may hold the 475-483 loop in a conformation needed for activity. Modified from Maita *et al.* (2010).

bridges may alter the conformation of the DXXK loop away from the active conformation. Among the single point mutations, D475V is the most disruptive. Perhaps it is because V475 now forms hydrophobic interactions with G461 and Y462 to introduce an inactive loop conformation, since these residues are in very close proximity to each other. To further support this hypothesis, KXXK and KXXD point mutants were generated and were shown to further destabilize the structure by potentially disrupting both salt bridges formed by D and K residues. These results suggest that the DXXK motif plays an important structural role for PglB activity.

There are an increasing number of sequenced *Campylobacter* genomes that have been reported to encode the *N*-glycosylation pathway (Nothaft & Szymanski,

2010). More recently, the *N*-glycans and fOS from all sequenced *Campylobacter* species were analyzed and divided into two major groups based on their glycan structures (Nothaft *et al.*, manuscript in preparation). *N*-glycans from group I species are described to have a similar structure to the heptasaccharide produced by *C. jejuni* 11168, which consists of GalNAc-GalNAc-[Glc]GalNAc-GalNAc-GalNAc-diNAcBac (Young *et al.*, 2002). In contrast, group II species include the non-thermophilic *Campylobacter* species that produce a heterogeneous mixture of *N*-glycan and fOS structures. The *N*-linked protein glycosylation system in *C. lari* has been characterized in detail recently (Schwarz *et al.*, 2010; Lizak *et al.*, 2011). *Cl*PglB shares 56% sequence identity to *Cj*PglB. The glycan structure produced by *C. lari* is the hexasaccharide $[HexNAc]_5$-diNAcBac, which is similar to the *C. jejuni* heptasaccharide (group I), but lacking the glucose branch since the *C. lari* *pgl* operon does not contain the *pglI* gene encoding the glucosyltransferase (Kelly *et al.*, 2006; Schwarz *et al.*, 2010; Nothaft & Szymanski, 2010). Interestingly, the fOS structure that is released by the *C. lari* PglB contains an additional phosphate residue on the reducing end sugar, diNAcBac (Nothaft *et al.*, manuscript in preparation).

Analysis of the multiple sequence alignment of *Campylobacter* PglB protein sequences resulted in the identification of amino acid differences at a single position located several residues downstream of the DXXK motif, which appears to correlate with the group I and II glycan structure divisions (Fig. 2.1). A proline is found in the PglB OTases from group I, which aligns with valine residues found in the PglB OTases from group II. Although *C. lari* is classified in

group I, *Cl*PglB has a serine at the aligned site. Therefore, we decided to compare the *Cj*PglB and *Cl*PglB enzymes to examine whether exchanging the proline and serine residues influences PglB activity. We demonstrated that the *Cl*PglB S493P point mutant glycosylates $AcrA_{Cj}$ at a higher efficiency than the *Cl*PglB wildtype enzyme when analyzed in *E. coli*. However, no difference was observed in the glycosylation efficiency of *Cj*PglB P487S point mutant compared to the wildtype enzyme in *E. coli*. The difference in PglB activity between wildtype *Cj*PglB and *Cl*PglB suggests that the structures of the proteins are slightly different and that the proline residue in *Cj*PglB does not influence its OTase activity. Interestingly, when we analyzed these PglB variants in *C. jejuni*, there was no significant difference in the glycosylation efficiency when compared to the *Cj*PglB wildtype enzyme complement. It appears that the *N*-glycosylation system in *C. jejuni* already works efficiently as observed by the complete modification of AcrA in the wildtype strain of *C. jejuni*, which may be due to the presence of another molecular component in *C. jejuni* that is interacting with PglB, but missing in *E. coli*. This is supported by previous research that has shown partial modification of AcrA when the *C. jejuni pgl* machinery is reconstituted in *E. coli* (Wacker *et al.*, 2002; Feldman *et al.*, 2005; Kowarik *et al.*, 2006a; Ielmini *et al.*, 2010).

The recently published 3.4 Å resolution crystal structure of the full-length *Cl*PglB provides insight into the structural functions of the identified sites (Lizak *et al.*, 2011). Analysis of the crystal structure (with the help of Dr. Ng, University of Calgary) shows that S493 of *Cl*PglB is located at a kink (Fig. 2.7). Substituting the serine residue with proline may further destabilize the structure of the helix by

removing a single hydrogen bond between the side chain hydroxyl group and the main chain carbonyl group of A462. The presence of an interruption in the alpha helix at S493 likely increases the flexibility of this part of the PglB structure in comparison to a regular alpha-helix.

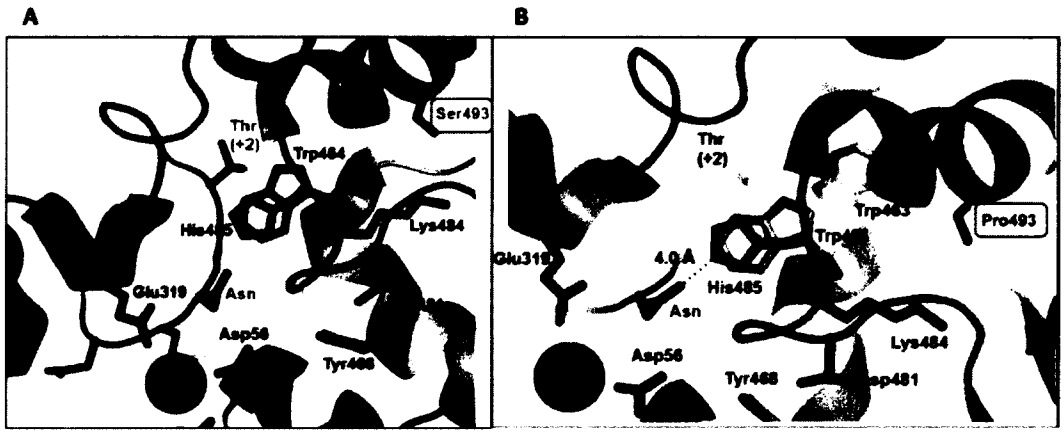

Figure 2.7. Close-up view of the *C. lari* PglB catalytic site. **A)** Ser493 of the *C. lari* PglB is located at a kink. Substituting the serine for a proline residue would alter this kink structure by introducing an even more distorted structure in the helix. **B)** Introducing a kink at this location may increase the flexibility of this part of the PglB structure in comparison to a regular alpha-helix. This interaction may become more optimal for either acceptor binding or for promoting a productive position of the acceptor side chain relative to the lipid-linked donor sugar. The catalytic pocket features the acidic side chains (Asp56 and Glu319) that coordinate the bound divalent cation (Mg^{2+}). Modified from Lizak *et al.* (2011).

Flexibility in this region may allow for more optimal binding interactions with the acceptor or may promote a more favourable position of the acceptor side chain relative to the lipid-linked donor sugar during the glycosyl transfer reaction. A

single point mutation may improve PglB activity by making slight changes in the binding of donor and acceptor substrate molecules. However, since the *E. coli* and *C. jejuni* experiments showed contrasting results, it is currently difficult to determine the influence of these amino acid substitutions without further studies.

The *N*-linked protein glycosylation pathways of *C. jejuni* and *C. lari*, encoded by the *pgl* gene clusters, have been functionally reconstituted in *E. coli* to produce recombinant glycoproteins (Wacker *et al.*, 2002; Schwarz *et al.*, 2010). This has opened up the possibility to exploit the bacterial glycosylation systems for biotechnological applications. The *Cj*PglB and *Cl*PglB both modify asparagine residues within the consensus sequence D/E-X_1-N-X_2-S/T (where X_1 and X_2 are any amino acid except proline), but it has been recently shown that the -2 position to the asparagine is dispensable for *Cl*PglB activity in *E. coli*, albeit with low transfer efficiency (Schwarz *et al.*, 2010). Our findings indicate that a single point mutation within the *Cl*PglB can improve glycosylation efficiency. It would be interesting to know whether an acceptor protein carrying the short eukaryotic sequon can also be modified with improved glycosylation efficiency using the *Cl*PglB variant created in this study. High-throughput assays capable of measuring relative PglB activities will be useful in comparing the various PglB point mutants that have been generated as well as further examining the uncharacterized OTases identified in the proteobacteria. These studies may lead to more efficient production of recombinant glycoproteins in *E. coli* for the development of novel glycoconjugate vaccines.

2.5. Acknowledgements

The authors would like to thank Abofu Alemka for providing the pET24(*Cl*pglB) plasmid, Bernadette Beadle for providing the pACYC(*pgl-ΔpglB*) plasmid, Margaret Johnson for the preliminary analysis of the *Cj*PglB soluble domain crystal structure, Markus Aebi (ETH, Zurich, Switzerland) for the hR6 and AcrA antibodies, and Mario Feldman and John Klassen for helpful suggestions. This work was supported by the Alberta Glycomics Centre. CMS is an Alberta Innovates Scholar.

2.6. Literature cited

Alaimo C, Catrein I, Morf L, Marolda CL, Callewaert N, Valvano MA, Feldman MF & Aebi M (2006) Two distinct but interchangeable mechanisms for flipping of lipid-linked oligosaccharides. *EMBO J* **25:** 967.

Baar C, Eppinger M, Raddatz G *et al.* (2003) Complete genome sequence and analysis of Wolinella succinogenes. *Proc Natl Acad Sci USA* **100:** 11690-11695.

Berriman M, Ghedin E, Hertz-Fowler C, Blandin G, Renauld H, Bartholomeu DC, Lennard NJ, Caler E, Hamlin NE & Haas B (2005) The genome of the African trypanosome *Trypanosoma brucei. Science* **309:** 416-422.

Burda P & Aebi M (1999) The dolichol pathway of *N*-linked glycosylation. *BBA-General Subjects* ***1426*:** 239-257.

Feldman MF, Wacker M, Hernandez M, Hitchen PG, Marolda CL, Kowarik M, Morris HR, Dell A, Valvano MA & Aebi M (2005) Engineering *N*-linked protein glycosylation with diverse O antigen lipopolysaccharide structures in *Escherichia coli. Proc Natl Acad Sci USA* **102:** 3016.

Figurski DH & Helinski DR (1979) Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. *Proc Natl Acad Sci USA* **76:** 1648-1652.

Helenius J, Ng DTW, Marolda CL, Walter P, Valvano MA & Aebi M (2002) Translocation of lipid-linked oligosaccharides across the ER membrane requires Rft1 protein. *Nature* **415:** 447-450.

Ielmini MV & Feldman MF (2011) Desulfovibrio desulfuricans PglB homolog possesses oligosaccharyltransferase activity with relaxed glycan specificity and distinct protein acceptor sequence requirements. *Glycobiology* **21:** 734-742.

Igura M, Maita N, Kamishikiryo J, Yamada M, Obita T, Maenaka K & Kohda D (2008) Structure-guided identification of a new catalytic motif of oligosaccharyltransferase. *EMBO J* **27:** 234-243.

Ivens AC, Peacock CS, Worthey EA, Murphy L, Aggarwal G, Berriman M, Sisk E, Rajandream MA, Adlem E & Aert R (2005) The genome of the kinetoplastid parasite, *Leishmania major. Science* **309:** 436-442.

Jaffee MB & Imperiali B (2011) Exploiting topological constraints to reveal buried sequence motifs in the membrane-bound N-linked oligosaccharyl transferases. *Biochemistry* **50:** 7557-7567.

Jervis AJ, Langdon R, Hitchen P, Lawson AJ, Wood A, Fothergill JL, Morris HR, Dell A, Wren B & Linton D (2010) Characterization of N-linked protein glycosylation in Helicobacter pullorum. *J Bacteriol* **192:** 5228-5236.

Kato Y, Arakawa N, Masuishi Y, Kawasaki H & Hirano H (2009) Mutagenesis of longer inserts by the ligation of two PCR fragments amplified with a mutation primer. *J Biosci Bioeng* **107:** 95-97.

Kelleher DJ & Gilmore R (2006) An evolving view of the eukaryotic oligosaccharyl-transferase. *Glycobiology* **16:** 47R.

Kelly J, Jarrell H, Millar L, Tessier L, Fiori LM, Lau PC, Allan B & Szymanski CM (2006) Biosynthesis of the *N*-linked glycan in *Campylobacter jejuni* and addition onto protein through block transfer. *J Bacteriol* **188:** 2427-2434.

Kohda D, Yamada M, Igura M, Kamishikiryo J & Maenaka K (2007) New oligosaccharyltransferase assay method. *Glycobiology* **17:** 1175.

Korlath JA, Osterholm MT, Judy LA, Forfang JC & Robinson RA (1985) A point-source outbreak of campylobacteriosis associated with consumption of raw milk. *J Infect Dis* **152:** 592-596.

Kowarik M, Numao S, Feldman MF, Schulz BL, Callewaert N, Kiermaier E, Catrein I & Aebi M (2006a) *N*-linked glycosylation of folded proteins by the bacterial oligosaccharyltransferase. *Science* **314:** 1148-1150.

Kowarik M, Young NM, Numao S, Schulz BL, Hug I, Callewaert N, Mills DC, Watson DC, Hernandez M, Kelly JF, Wacker M & Aebi M (2006b) Definition of the bacterial *N*-glycosylation site consensus sequence. *EMBO J* **25:** 1957-1966.

Labigne-Roussel A, Harel J & Tompkins L (1987) Gene transfer from *Escherichia coli* to *Campylobacter* species: development of shuttle vectors for genetic analysis of *Campylobacter jejuni. J Bacteriol* **169:** 5320-5323.

Larsen JC, Szymanski C & Guerry P (2004) *N*-linked protein glycosylation is required for full competence in *Campylobacter jejuni* 81-176. *J Bacteriol* **186:** 6508-6514.

Lefebre MD & Valvano MA (2002) Construction and evaluation of plasmid vectors optimized for constitutive and regulated gene expression in Burkholderia cepacia complex isolates. *Appl Environ Microbiol* **68:** 5956-5964.

Li L, Woodward R, Ding Y, Liu XW, Yi W, Bhatt VS, Chen M, Zhang LW & Wang PG (2010) Overexpression and topology of bacterial oligosaccharyltransferase PglB. *Biochem Biophys Res Commun* **394:** 1069-1074.

Linton D, Dorrell N, Hitchen PG, Amber S, Karlyshev AV, Morris HR, Dell A, Valvano MA, Aebi M & Wren BW (2005) Functional analysis of the *Campylobacter jejuni* N-linked protein glycosylation pathway. *Mol Microbiol* **55:** 1695-1703.

Liu X, McNally DJ, Nothaft H, Szymanski CM, Brisson JR & Li J (2006) Mass spectrometry-based glycomics strategy for exploring *N*-linked glycosylation in eukaryotes and bacteria. *Anal Chem* **78:** 6081-6087.

Lizak C, Gerber S, Numao S, Aebi M & Locher KP (2011) X-ray structure of a bacterial oligosaccharyltransferase. *Nature* **474:** 350-355.

Maita N, Nyirenda J, Igura M, Kamishikiryo J & Kohda D (2010) Comparative structural biology of Eubacterial and Archaeal oligosaccharyltransferases. *J Biol Chem* **285:** 4941.

Mescher MF & Strominger JL (1976) Structural (shape-maintaining) role of the cell surface glycoprotein of *Halobacterium salinarium*. *Proc Natl Acad Sci USA* **73:** 2687-2691.

Nothaft H, Amber S, Aebi M & Szymanski CM (2008) *N*-linked protein glycosylation in Campylobacter. *Campylobacter*, Vol. 3 (Nachamkin I, Szymanski CM & Blaser MJ, eds), pp. 447-469. ASM Press, Washington, DC.

Nothaft H, Liu X, McNally DJ, Li J & Szymanski CM (2009) Study of free oligosaccharides derived from the bacterial *N*-glycosylation pathway. *Proc Natl Acad Sci USA* **106:** 15019.

Nothaft H, Liu X, McNally DJ & Szymanski CM (2010) *N*-linked protein glycosylation in a bacterial system. Functional glycomics: methods and protocols, Vol. 600 (Li J, ed), pp. 227-243. Humana press, Clifton, NJ, USA.

Nothaft H & Szymanski CM (2010) Protein glycosylation in bacteria: sweeter than ever. *Nat Rev Microbiol* **8:** 765-778.

Parkhill J, Wren BW, Mungall K, Ketley JM, Churcher C, Basham D, Chillingworth T, Davies RM, Feltwell T & Holroyd S (2000) The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature* **403:** 665-668.

Reid CW, Stupak J & Szymanski CM (2010) Characterization of lipid-linked oligosaccharides by mass spectrometry. Methods in molecular biology **600:** 187.

Sambrook J & Russell DW (2001) *Molecular cloning: a laboratory manual.* CSHL press.

Schwarz F, Lizak C, Fan YY, Fleurkens S, Kowarik M & Aebi M (2011) Relaxed acceptor site specificity of bacterial oligosaccharyltransferase *in vivo*. *Glycobiology* **21:** 45-54.

Szymanski CM, Yao R, Ewing CP, Trust TJ & Guerry P (1999) Evidence for a system of general protein glycosylation in *Campylobacter jejuni*. *Mol Microbiol* **32:** 1022-1030.

Szymanski CM, Michael FS, Jarrell HC, Li J, Gilbert M, Larocque S, Vinogradov E & Brisson JR (2003) Detection of conserved N-linked glycans and phase-variable lipooligosaccharides and capsules from campylobacter cells by mass spectrometry and high resolution magic angle spinning NMR spectroscopy. *J Biol Chem* **278:** 24509-24520.

Szymanski CM & Wren BW (2005) Protein glycosylation in bacterial mucosal pathogens. *Nat Rev Microbiol* **3:** 225-237.

Wacker M, Linton D, Hitchen PG, Nita-Lazar M, Haslam SM, North SJ, Panico M, Morris HR, Dell A & Wren BW (2002) *N*-linked glycosylation in *Campylobacter jejuni* and its functional transfer into *E. coli*. *Science* **298:** 1790.

Yan Q & Lennarz WJ (1999) Oligosaccharyltransferase: a complex multisubunit enzyme of the endoplasmic reticulum. *Biochem Biophys Res Commun* **266:** 684-689.

Young NM, Brisson JR, Kelly J, Watson DC, Tessier L, Lanthier PH, Jarrell HC, Cadotte N, St. Michael F & Aberg E (2002) Structure of the *N*-linked glycan present on multiple glycoproteins in the Gram-negative bacterium, *Campylobacter jejuni*. *J Biol Chem* **277:** 42530-42539.

CHAPTER 3

Autoregulation of the bacterial oligosaccharyltransferase

3.1. Introduction

Protein glycosylation is the most common covalent modification of newly synthesized proteins in eukaryotes. Glycans play important roles for the structure and function of cellular proteins in addition to the development and growth of an organism (Varki *et al.*, 1993; Hebert *et al.*, 2005). It has been predicted that more than half of all proteins in nature are glycosylated, of which the majority are asparagine-(*N*)-linked (Apweiler *et al.*, 1999). *N*-linked protein glycosylation has been most extensively studied in the yeast *Saccharomyces cerevisiae*, where the glycan is assembled on the lipid carrier dolichol-pyrophosphate at the cytoplasmic side of the endoplasmic reticulum (ER), translocated to the lumen for additional glycan modification, and subsequently attached to the acceptor protein at asparagine residues found within the N-X-S/T (X ≠ P) consensus sequence by the oligosaccharyltransferase (OTase) (Burda & Aebi, 1999). The OTase in yeast is a multimeric protein complex comprising at least eight subunits, where the Stt3 protein is the most conserved among eukaryotes and represents the central catalytic site (Yan *et al.*, 1999). Single-subunit proteins that consist of the Stt3 alone have been found to be sufficient for OTase activity in other lower eukaryotes and archaea (Kelleher & Gilmore, 2006).

Evidence for the bacterial *N*-linked protein glycosylation (*pgl*) system was first described in the epsilonproteobacterium *Campylobacter jejuni* (Szymanski *et al.*, 1999). All of the genes necessary for the *C. jejuni* protein glycosylation machinery are encoded in a *pgl* gene locus that can be functionally transferred into *E. coli* to produce recombinant glycoproteins (Szymanski *et al.*, 1999;

Wacker *et al.*, 2002). The single subunit Stt3 homologue, PglB, was shown to be the bacterial OTase essential for *N*-glycosylation and solely responsible for modifying asparagine residues with a heptasaccharide consisting of five *N*-acetylgalactosamine residues, a glucose branch, and *N,N'*-diacetyl-bacillosamine (diNAcBac) at the reducing end (Wacker *et al.*, 2002; Young *et al.*, 2002). *In vivo* and *in vitro* assays have been dedicated to elucidate the individual roles of the *pgl* genes in *C. jejuni* (Linton *et al.*, 2005; Glover *et al.*, 2005). Homologous to the eukaryotic pathway, the bacterial process begins in the cytoplasm where the glycan is assembled on the lipid anchor undecaprenyl-pyrophosphate, flipped to the periplasm and attached to the acceptor asparagine of nascent polypeptides (Nothaft & Szymanski, 2010). In addition to its OTase activity, it has been demonstrated that PglB produces heptasaccharides as free oligosaccharides (fOS) in *C. jejuni* dependent on the bacterial growth phase and the osmotic environment (Nothaft *et al.*, 2009).

Furthermore, PglB can modify acceptor proteins with glycans of various lengths other than its native heptasaccharide, including O-antigens from *E. coli* and *Pseudomonas aeruginosa* (Feldman *et al.*, 2005). However, it has been demonstrated that PglB can only transfer glycans that have an acetyl group at the C-2 position of the reducing end sugar (Wacker *et al.*, 2006). *C. jejuni* PglB modifies asparagine residues found within the extended D/E-X_1-N-X_2-S/T (X_1, $X_2 \neq$ P) consensus sequence located in the flexible loops of proteins (Kowarik *et al.*, 2006b). However, other PglB homologous enzymes identified in the epsilonproteobacterium *Campylobacter lari* and the deltaproteobacterium

Desulfovibrio desulfuricans have been shown to transfer glycans at the eukaryotic-like N-X-S/T sequon (Ielmini & Feldman, 2011; Schwarz *et al.*, 2011).

Interestingly, the consensus sequence required for the *N*-glycosylation of acceptor proteins is also found within the soluble C-terminal domain of *C. jejuni* PglB (^{534}DYNQS^{536}). A glycopeptide fragment of *C. jejuni* PglB was previously identified from LC-MS/MS using zwitterionic hydrophilic interaction chromatography confirming that this site is indeed modified in the native protein (Scott *et al.*, 2011). In addition, *C. lari* PglB was shown to autoglycosylate at N535 and N556 (Lizak *et al.*, 2011). This raises the question of whether PglB self-glycosylation modulates its OTase and hydrolase activities.

Analysis of the available *C. jejuni* PglB soluble domain crystal structure (Maita *et al.*, 2010) revealed that the consensus sequence is located at the surface of the protein (Fig. 3.1). If glycosylation of PglB is required for function as an autoregulatory mechanism, then its location allows for the control of oligomerization or allosteric effects (Johnson, M., personal communication). In this study, we created and examined the PglB consensus sequence mutant in both the heterologous *E. coli* host and the native *C. jejuni* host and found that the PglB mutant does not affect protein glycosylation, but influences fOS production. This suggests that PglB self-glycosylation may be modulating the two enzyme activities of this OTase.

Figure 3.1. The crystal structure of the *C. jejuni* PglB soluble domain. Analysis of the PglB soluble domain crystal structure revealed that the consensus sequence (^{532}DYNQS^{536}) is located at the surface exposed flexible loop of the protein (acceptor asparagine labeled N534).

3.2. Materials and Methods

3.2.1. *Bacterial strains, plasmids, and growth conditions*

The bacterial strains and plasmids used in this study are listed in Table 3.1. *Escherichia coli* strains were grown in Luria-Bertani (LB) broth or agar (1% Bacto™ Tryptone, 0.5% Bacto™ Yeast Extract, 1% NaCl and 1.5% agar was added for plates) at 37°C. For large scale growth and protein purification, *E. coli* strains were grown in 2×YT medium (1.6% Bacto™ Tryptone, 1% Bacto™ Yeast Extract, 0.5% NaCl) at 37°C. *Campylobacter jejuni* strains were grown on Mueller Hinton (MH, Difco™) agar plates under microaerobic conditions (10% CO_2, 5% O_2, 85%

N_2) at 37°C for 18 hrs. Ampicillin (100 ug/mL), chloramphenicol (25 ug/mL), kanamycin (25 ug/mL), trimethoprim (25 ug/mL) and tetracycline (25 ug/mL) were added to the media as needed for selection.

3.2.2. *Construction of plasmids*

The asparagine residue found within the ^{532}DYNQS^{536} consensus sequence was substituted for its amide-containing complement (N534Q). The mutation was generated with the forward oligonucleotide (N523Q-F) 5'-GATGAAAGATTATCAACAAAGCAATGTG-3' and the reverse oligonucleotide (N534Q-R) 5'-CACATTGCTTTGTTGATAATCTTTCATC-3' using the QuikChange® II XL Site-Directed Mutagenesis Kit (Stratagene) following the manufacturer's protocol with plasmid pMAF10 carrying the *C. jejuni pglB* gene as template. Transformation of XL10-Gold® ultracompetent cells with the mutagenized plasmid resulted in plasmid pNQ1, which was isolated using the GeneJET™ Plasmid Miniprep Kit (Fermentas). The presence of the desired point mutation was confirmed by DNA sequencing (Molecular Biology Service Unit, University of Alberta) as carried out using the BigDye® Terminator v3.1 Cycle Sequencing Kit following the manufacturer's protocol (Applied Biosystems) with N534Q-R as the sequencing primer.

To add a hexahistidine tag to the PglB(N534Q) mutant, the *pglB* gene carrying the N534Q mutation was amplified by PCR with oligonucleotides PglB*Nde*I-F (5'-AATATATAACATATGTTGAAAAAAGAGTATTTAAAAA-ACCC-3') and PglB*Xho*I-R (5'-TATGGCTCGAGTTTAAGTTTAAAAACCTT-AGC-3') using the Platinum® *Pfx* DNA Polymerase Kit (Invitrogen) with pNQ1

Table 3.1. Bacterial strains and plasmids used in this study

Strain or plasmid	Characteristics	Source
E. coli		
DH5α	F⁻ *endA1 hsdR17 supE44 thi-1 recA1* Δ *(argF-lacZYA)U169* (80d *lacZ* Δ *M15) gyrA96 λ⁻*	Invitrogen
CLM24	W3110, Δ*waaL*	Feldman *et al.*, 2005
C43 (DE3)	F⁻ *ompT gal hsdS$_B$* (r_B^-, m_B^+) *dcm lon* λDE3 with two uncharacterized mutations	Li *et al.*, 2010
BL21 (DE3	F⁻ *ompT gal dcm lon hsdS$_B$*(r_B^- m_B^-) λ(DE3 [*lacI lacUV5*-T7 gene 1 *indl sam7 nin5*])	Wacker *et al.*, 2002
SCM7	SØ874, Δ*wec*	Alaimo *et al.*, 2006
C600 (RK212.2)	*leu thr thi lacy supE44 tonA;* pRK212.2, Amp^R, Tet^R	Figurski & Helinski, 1979
C. jejuni		
81-176	Clinical isolate	Korlath *et al.*, 1985
11168 NCTC	Clinical isolate used for genome sequencing	Parkhill *et al.*, 2000
11168-*pglB*::Kan	*pglB* mutant, Km^R	Nothaft *et al.*, 2009
11168-*pglB*(pCj-*pgl*B)	*pglB* mutant complemented with pCE111-28 carrying *pglB*, Kan^R, Cm^R	Nothaft *et al.*, 2009
11168-*pglB*(pCj-*pgl*B$_{mut}$)	*pglB* mutant complemented with pCE111-28 carrying *pglB*$_{mut}$, Kan^R, Cm^R	Nothaft *et al.*, 2009
Plasmids		
pACYC(*pgl*)	Encodes the *C. jejuni pgl* locus, Cm^R	Wacker *et al.*, 2002
pACYC(*pgl$_{mut}$*)	Encodes the *C. jejuni pgl* locus with a mutation at WWDYG in PglB, Cm^R	Wacker *et al.*, 2002
pACYC(*pgl*-Δ*pglB*)	Encodes the *C. jejuni pgl* locus with a deletion in PglB, Cm^R	Provided by Bernadette Beadle
pET24b-*acrA*	Soluble periplasmic *C. jejuni acrA*$_{His6}$ under control of T7 promoter, in pET24b, Amp^R	Wacker *et al.*, 2002
pWA2	Soluble periplasmic *C. jejuni acr*A$_{His6}$ under control of Tet promoter, in pBR322, Amp^R	Feldman *et al.*, 2005
pMLBAD	Cloning vector, arabinose-inducible, Tmp^R	Lefebre & Valvano, 2002
pMAF10	HA-tagged *C. jejuni pglB* cloned in pMLBAD, Tmp^R	Feldman *et al.*, 2005
pWA1	HA-tagged *C. jejuni pglB*$_{mut}$ cloned in pMLBAD, Tmp^R	Feldman *et al.*, 2005
pNQ1	HA-tagged *C. jejuni pglB*(N534Q) cloned in pMLBAD, Tmp^R	This study
pET24-PglB$_{WT}$	His-tagged *C. jejuni pglB* cloned in pET24, Kan^R	Glover *et al.*, 2005
pET24-PglB$_{N534Q}$	His-tagged *C. jejuni pglB*(N534Q) cloned in pET24, Kan^R	This study
pCE111-28	*C. jejuni* expression vector, plasmid pRY111 with σ^{28} promoter of *flaA*, Cm^R	Larsen *et al.*, 2004
pCj-*pglB*NQ	*C. jejuni pglB*(N534Q) cloned in pCE111-28, Cm^R	This study

as template and following the manufacturer's protocol. The oligonucleotide PglB*Xho*I-R encodes the His6-tag for purification by immobilized nickel ion chromatography and also assists with following PglB expression during Western blotting. To clone the PCR product in a pET24 vector, the pET24b-*acrA* plasmid

was first digested with *Nde*I and *Xho*I restriction enzymes (Fermentas). The DNA fragments were then separated and visualized by agarose gel electrophoresis. The pET24b vector was recovered from a 1% agarose gel and purified using the QIAquick® Gel Extraction Kit (Qiagen) according to the manufacturer's directions. The PCR product was also digested with *Nde*I-*Xho*I, purified, and cloned in the same sites of the purified pET24b vector using the T4 ligase (Fermentas) with a vector:insert ratio of 1:5 (10-20 ng of vector DNA) incubated overnight at 16°C (Sambrook & Russell, 2001). Chemically competent *E. coli* DH5α cells (Invitrogen) were transformed with the overnight ligation reaction as per the manufacturer's instructions. The resulting plasmid pET24-$PglB_{N534Q}$ was confirmed by restriction analysis.

For genetic manipulations of *C. jejuni* 11168, the *pglB* gene carrying the N534Q mutation was amplified by PCR with oligonucleotides pglB*Bam*HI-F (5'-ATTAGCGGATCCTACCTGACGCTTTTTATCGC-3') and pglB*Xho*I-HA-R (5'-TATACTCGAGTTAAGCGTAATCTGGAACATCGTATGG-3') where *Bam*HI and *Xho*I restriction sites were inserted at the flanking regions of the *pglB* gene using Vent polymerase (NEB) and pNQ1 as template. The PCR products and the chloramphenicol-resistant *C. jejuni-E. coli* shuttle vector pCE111-28 were double digested with *Bam*HI and *Xho*I restriction enzymes (Fermentas) and purified using the QIAquick® Gel Extraction Kit (Qiagen). The purified DNA fragments were ligated together using the T4 ligase (Fermentas). Transformation of the ligation reaction into *E. coli* DH5α resulted in plasmid pCj-*pglB*NQ and was confirmed by restriction analysis.

3.2.3. *Expression of the PglB point mutant in* E. coli *and purification of AcrA*

E. coli CLM24 cells were transformed in a stepwise manner with pACYC($pglB_{mut}$) encoding the *C. jejuni pgl* locus with an inactive *pglB* gene, pWA2 encoding a soluble form of the *C. jejuni* glycoprotein $AcrA_{6His}$, and pNQ1 carrying the *pglB* N534Q point mutant. For negative controls, the *E. coli* strain co-expressing pACYC($pglB_{mut}$) and pWA2 were individually expressed with the pMLBAD empty vector and pWA1 encoding an inactive *pglB* gene with a mutation in the WWDYG motif (W458A, D459A). As a positive control, pMAF10 encoding the wildtype *pglB* was expressed in *E. coli*. All of the *pglB* genes cloned in the pMLBAD vector encode an HA-tag to follow PglB expression. All strains were grown in LB broth from 5 mL overnight cultures at 37°C with shaking to an OD_{600} of 0.5-0.6 and induced with 0.2% L-arabinose (wt/vol). After 4 hrs, cells were harvested by centrifugation (3,696 x *g*, 15 min, 4°C) and the pellets were resuspended in 5% of the original culture volume in phosphate-buffered saline (PBS), pH 7.2. Samples were kept on ice and whole cell lysates were prepared by sonication (3 × 1 min) with a Branson sonicator and soluble proteins were obtained by centrifugation (16,260 x *g*, 20 min, 4°C). Protein concentrations in the supernatant were determined using the NanoVue™ Plus Spectrophotometer (GE Healthcare) at an absorbance reading of 280 nm (A_{280}). Protein samples were all set to the same concentration. PglB protein expression was confirmed by Western blot analysis using HA-tag specific antibodies.

For AcrA purification, whole cell lysates were set to the same protein concentration, subjected to nickel affinity chromatography (NTA agarose, Qiagen), washed with 10 column volumes of 20 mM imidazole in 1xPBS and eluted with 500 mM imidazole in 1xPBS. The samples were dialyzed against 1xPBS overnight at 4°C to remove imidazole and other small contaminants using the Slide-A-Lyzer Dialysis Cassettes, 3.5 K MWCO (Thermo Scientific). Protein concentrations of the AcrA samples were then determined spectrophotometrically (A_{280}). The glycosylation status of purified AcrA was analysed by Western blotting using AcrA-specific and *C. jejuni-N*-glycan specific antibodies.

3.2.4. *Complementation of $PglB_{N534Q}$ in* C. jejuni pglB *mutant*

For conjugation, the pCj-*pglB*NQ shuttle plasmid was isolated from *E. coli* DH5α and re-transformed into *E. coli* C600 harbouring the conjugative plasmid pRK212.2. The shuttle plasmid was then conjugated from *E. coli* RK212.2 to *C. jejuni* 81-176 by following the protocol of Labigne-Roussell and colleagues (1987). The *Campylobacter*-methylated plasmid was isolated using the GeneJET™ Plasmid Miniprep Kit (Fermentas) and then complemented into a *C. jejuni* 11168 *pglB* mutant background by natural transformation. The transconjugants were selected on MH agar containing the appropriate antibiotics, incubated for up to 72 hrs under microaerobic conditions, restreaked and stored as frozen glycerol stocks (-80°C) until needed. Whole cell lysates of *C. jejuni* 11168 were prepared as described in section 2.2.4. The glycosylation status of AcrA was analyzed by Western blotting using AcrA-specific antibodies.

3.2.5. Semi-quantitative mass spectrometry (sqMS)-based fOS analyses

Free glycans from the *C. jejuni* 11168 *pglB* mutants were prepared and analyzed as described in section 2.2.5. Samples were prepared to a final concentration of 5 mg/mL and analyzed using a 4000 Q-Trap® mass spectrometer (Applied Biosystems/MDS Sciex) with a capillary electrophoresis interface as previously described (Liu *et al.*, 2006; Nothaft *et al.*, 2009). fOS analysis were repeated in triplicate and the average relative quantities were determined.

3.2.6. *PglB overexpression and purification*

The plasmids pET24-$PglB_{WT}$ and pET24-$PglB_{N534Q}$ were expressed in *E. coli* C43(DE3). Starting from 10 mL overnight cultures, the *E. coli* strains were each grown in 1L of 2xYT broth at 37°C with shaking. When the cells reached an OD_{600} of 0.95-1.05, the cultures were induced by the addition of IPTG to a final concentration of 0.1 mM. The temperature was reduced to 30°C and the cell cultures were incubated for an additional 5 hrs. The following steps were performed at 4°C. Cells were harvested by centrifugation (5,000 x *g*, 15 min) and resuspended in 5% of the original culture volume in cold lysis buffer (50 mM Tris-HCl (pH 8.0), 250 mM NaCl, cOmplete™ EDTA-free protease inhibitor cocktail (Roche), DNAseI (25 μg/mL)). Cells were disrupted using a Cell Disrupter (TS Series Benchtop, Constant Systems Ltd.) followed by centrifugation (16,260 x *g*, 20 min) to remove unbroken cells. The membrane fraction was isolated from the supernatant by ultracentrifugation (200,000 x *g*, 70 min). The membrane pellets were subsequently resuspended in resuspension buffer (50 mM Tris-HCl (pH 8.0), 250 mM NaCl, 10 mM imidazole, 5% (v/v)

glycerol, 1% (w/v) DDM) and incubated overnight at 4°C with slow shaking. The insoluble fraction was then removed by a second ultracentrifugation step (200,000 x *g*, 70 min) and the supernatants were subjected to nickel affinity chromatography (NTA agarose, Qiagen) to purify $PglB_{WT}$ and $PglB_{N534Q}$.

The Ni-NTA column was pre-equilibrated with 10 column volumes of equilibration buffer (50 mM Tris-HCl (pH 8.0), 250 mM NaCl, 10 mM imidazole, 5% (v/v) glycerol, 0.5% (w/v) DDM). The supernatant was loaded onto the column, the flow through was collected, and the resin was washed with another 10 volumes of equilibration buffer. The resin was washed again with 10 volumes of washing buffer (50 mM Tris-HCl (pH 8.0), 250 mM NaCl, 50 mM imidazole, 5% (v/v) glycerol, 0.5% (w/v) DDM) and the PglB proteins were eluted with 5 x 500 µL of elution buffer (50 mM Tris-HCl (pH 8.0), 250 mM NaCl, 250 mM imidazole, 5% (v/v) glycerol, 0.5% (w/v) DDM), yielding pure fractions of $PglB_{His6}$. The elution fractions were desalted against buffer containing 50 mM Tris-HCl (pH 8.0), 5% (v/v) glycerol, and 0.5% (w/v) DDM using a PD-10 Desalting column (GE Healthcare). Protein concentrations were measured using the *DC*™ Protein Assay Kit (Bio-Rad) following the manufacturer's instructions. The PglB proteins were stored at 4°C until needed.

3.2.7. *Purification of $AcrA_{His6}$*

The plasmid pET24b-*acrA* carrying the $AcrA_{His6}$ protein was expressed in *E. coli* BL21 (DE3) and purified as described (Wacker *et al.*, 2002). Briefly, the *E. coli* strain was grown at 37°C with shaking in 1L of 2×YT broth supplemented with ampicillin. When the OD_{600} reached 0.5-0.6, the culture was induced with 0.1 mM

IPTG and was further incubated overnight at 37°C. Cells were harvested (5,000 x *g*, 15 min, 4°C), resuspended in 5% of the original culture volume of 1xPBS (pH 7.2), and lysed using the Cell Disrupter (TS Series Benchtop, Constant Systems Ltd.). The soluble fraction was isolated by centrifugation (16,260 x *g*, 20 min) and the supernatant was subjected to Ni-NTA purification as described 2.2.3. The samples were dialyzed against 1xPBS overnight at 4°C and the protein concentration of $AcrA_{His6}$ was determined spectrophotometrically (A_{280}).

3.2.8. *Preparation of LLOs*

LLOs were obtained from *E. coli* SCM7 harboring the plasmid pACYC(pgl_{mut}) using a modified protocol as described by (Reid *et al.*, 2009). Briefly, cells were grown in 1L of 2×YT and incubated overnight at 37°C with shaking. Cells were harvested by centrifugation (5,000 x *g*, 15 min, 4°C), washed with 0.9% NaCl solution, and recentrifuged (3,696 x *g*, 15 min, 4°C). Lipids were sequentially extracted from the cell pellet with 2:1 chloroform/methanol and 10:20:3 chloroform/methanol/water (2 x 20 mL). The cell debris was removed by centrifugation (3,696 x *g*, 15 min, 4°C) and the supernatants were pooled together before evaporating to dryness with low heat and stirring. Dried lipids were dissolved in 1 mL of 9:1 ethanol:methanol and further diluted 10 fold in Buffer A (10 mM Hepes (pH 7.4), 150 mM NaCl, 1 mM EDTA) and incubated with 100 µL of Soybean-agglutinin-agarose (EY-Labs) for 1 hr to enrich for LLOs. The mixture was centrifuged (2200 x *g*, 5 min, 4°C) and the resin was subsequently washed with 3 x Buffer A. Bound LLOs were then eluted with 2 mL of 0.2 M D-galactose in 100 mM Hepes (pH 7.4).

3.2.9. In vitro *PglB activity assay*

In vitro PglB activity was assayed as previously described (Faridmoayer *et al.*, 2008) with appropriate modifications. First, 30 µL of LLOs were dried under vacuum and resuspended in 10 µL of DMSO. Subsequently, 50 µL of reaction buffer (50 mM Tris-HCl (pH 8.0), 280 mM sucrose), 2 µL of 1 M $MnCl_2$, 500 ng of purified $AcrA_{His6}$, and 1 µg of PglB were added to the tube containing LLOs. The volume was adjusted to 100 µL with water and the mixture was incubated overnight at 30°C with shaking. To analyze protein glycosylation, the reaction was immediately TCA precipitated, dissolved in 20 µL of 2X SDS sample buffer, and applied to 10% SDS-PAGE. $AcrA_{His6}$ proteins were detected by Western blotting using AcrA-specific and *C. jejuni-N*-glycan specific antibodies.

3.2.10. *Western blot analysis*

Protein expression and glycosylation were analyzed by SDS-PAGE on 10% gels followed by Western transfer analysis as described in section 2.2.6. Primary antibodies were used at the following dilutions: 1:2,000 for HA-specific (Santa Cruz), 1:2,000 for His-specific (monoclonal, Santa Cruz), 1:6,000 for AcrA-specific (kindly provided by Dr. Mario Feldman), and 1:10,000 *C. jejuni-N*-glycan-specific antisera. Alkaline phosphatase conjugated anti-rabbit and anti-mouse IgG (Santa Cruz) served as the secondary antibodies at 1:2,000 dilution.

3.3. Results

3.3.1. *Site-directed mutagenesis and expression of $PglB_{N534Q}$*

To analyze the importance of self-glycosylation of PglB, the amino acid substitution of N534 to Q534 within the ^{532}DYNQS^{536} consensus sequence was

made. The mutation was generated by site-directed mutagenesis using an HA-tagged version of PglB expressed under the control of an arabinose inducible promoter. Sequencing results confirmed the amino acid substitution at the desired site. The mutated-*pglB* allele was expressed in *E. coli* and the PglB point mutant ($PglB_{N534Q}$) protein was visualized by Western blotting with anti-HA antibodies showing that the introduction of this point mutation had no affect on PglB protein expression.

3.3.2. *$PglB_{N534Q}$ activity in E. coli and C. jejuni*

To see if the point mutation had an affect on PglB function, the *in vivo* PglB activity was investigated in the heterologous *E. coli* system and by complementation of a *C. jejuni pglB* mutant. Three plasmids containing: a His6-tagged *C. jejuni* glycoprotein AcrA ($AcrA_{His6}$), the *C. jejuni* protein glycosylation *(pgl)* operon containing a deletion in *pglB*, and the $PglB_{N534Q}$ mutant were co-expressed in *E. coli* CLM24 (an O-antigen ligase mutant where the produced undecaprenyl-pyrophosphate-glycan is transferred to an acceptor protein rather than to lipid A). Periplasmic *E. coli* $AcrA_{His6}$ protein was purified by Ni-NTA affinity chromatography and examined by Western blotting with monoclonal anti-His antibodies. With respect to the wildtype PglB, which serves as a positive control and results in the expression of double-glycosylated AcrA, there was no change in glycosylation with the PglB N534Q point mutant. This was also confirmed by Western blot analysis with anti-*C. jejuni-N*-glycan specific antisera (Fig. 3.2A). The $PglB_{N534Q}$ mutant was also complemented into a *C. jejuni pglB* mutant background. Again, no considerable difference in glycosylation efficiency

was observed for the N534Q mutant as visualized by Western blotting with anti-*C. jejuni*-AcrA specific antisera (Fig. 3.2B).

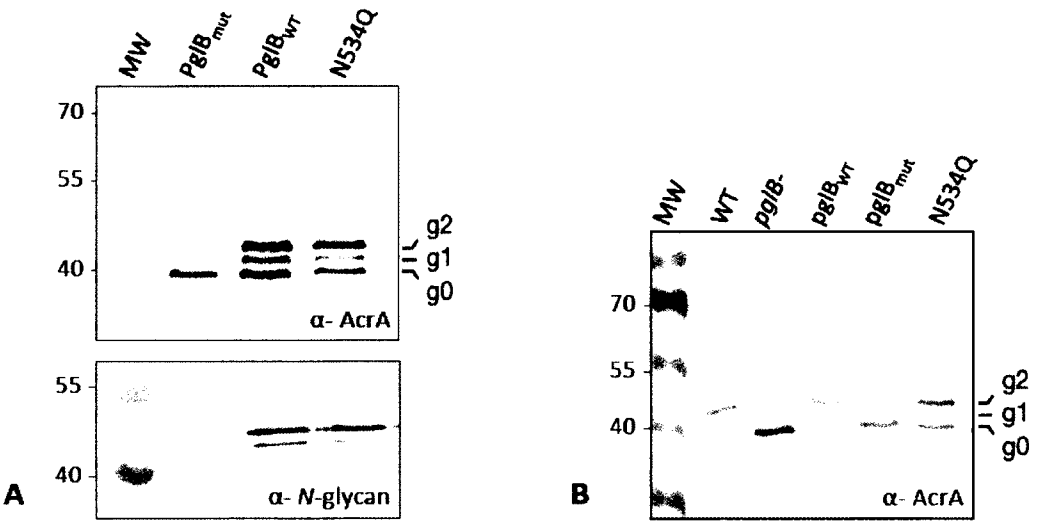

Figure 3.2. Western Blot analysis of the *pglB* N534Q mutant in *E. coli* and *C. jejuni*. **A)** $PglB_{mut}$, $PglB_{wt}$ and the N534Q mutant were each co-expressed in *E. coli* with the *C. jejuni pgl* operon and a His6-tagged *C. jejuni* glycoprotein AcrA as an *N*-glycosylation acceptor. The periplasmic *E. coli* AcrA-His proteins were purified by Ni-NTA affinity chromatography and examined by Western blotting with anti-AcrA and anti-*N*-glycan specific antisera as indicated. **B)** Complementation of a chromosomal *C. jejuni pglB* knock-out mutant with the $pglB_{wt}$, $pglB_{mut}$, and $pglB_{N534Q}$ mutant alleles. PglB activity was visualized by Western blotting with *C. jejuni* AcrA specific antibodies. Molecular weight masses are indicated in kDa on the left of each blot. The glycosylation state of AcrA is indicated on the right.

3.3.3. *Analysis of fOS production in C. jejuni*

We also examined whether modification of the consensus sequence in PglB influences the release of free oligosaccharides (fOS) in *C. jejuni* (in collaboration with Dr. Liu and Dr. Li, National Research Council, Ottawa). Although complementation of the *C. jejuni pglB* mutant with the $PglB_{N534Q}$ mutant did not show an affect on protein *N*-glycosylation, the same complement with $PglB_{N534Q}$ showed a 2-fold reduction in fOS production (Fig. 3.3). However, this $PglB_{N534Q}$ mutant shows variable levels of fOS production in *C. jejuni*. When compared to the wildtype fOS levels (set to 100%), the $PglB_{N534Q}$ mutant produced 18%, 40%, and 16% of fOS when comparing biological triplicates for analysis. The level of variation suggests that self-glycosylation may be modulating the hydrolase enzyme activity.

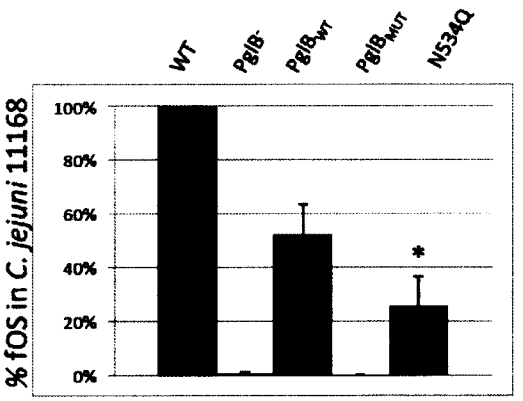

Figure 3.3. The role of the ^{532}DY<u>N</u>QS^{536} consensus sequence in *C. jejuni* PglB fOS production. The amount of fOS released in *C. jejuni* was determined using semi-quantitative mass spectrometry. The complemented *C. jejuni pglB* N534Q mutant shows a decrease in fOS production compared to the wildtype complement. Results obtained are from biological triplicates; asterisks indicate $p<0.05$.

3.3.4. In vivo *PglB self-glycosylation*

To understand the conditions under which PglB glycosylates itself, five *E. coli* C43 strains were constructed expressing: 1) the *pglB* gene alone; 2) the *pglB* gene with the *C. jejuni pgl* operon containing a *pglB* deletion (*ΔpglB*); 3) the *pglB* gene with the *C. jejuni pgl*(*ΔpglB*) operon and $AcrA_{His6}$; 4) the *pglB* gene with the *C. jejuni pgl*(*ΔpglB*) operon and an $AcrA_{His6}$ glycosylation mutant (N123L, N273L); and 5) the $pglB_{N534Q}$ mutant with the *C. jejuni pgl*(*ΔpglB*) operon and $AcrA_{His6}$. Solubilized membrane fractions containing the $PglB_{His6}$ proteins were prepared by ultracentrifugation and purified by Ni-NTA affinity chromatography. Elution fractions containing purified PglB were subjected to SDS-PAGE and stained (Fig. 3.4). Glycosylation of PglB and AcrA proteins were analyzed by Western blotting with anti-His and anti-*N*-glycan antibodies (Fig. 3.5). All of the PglB proteins were expressed at the expected mass as confirmed with His-specific antisera. Glycosylation of wildtype PglB was observed in *E. coli* strains co-expressed with the *C. jejuni pgl*(*ΔpglB*) operon and $AcrA_{His6}$ (wildtype and mutant forms). Interestingly, glycosylation of PglB was observed in the absence of the protein acceptor $AcrA_{His6}$. Moreover, the $PglB_{N534Q}$ mutant was not able to self-glycosylate, but was still able to modify $AcrA_{His6}$. Proteins samples were also probed with AcrA-specific antibodies to confirm that the *N*-glycan modified protein bands are not PglB degradation products, but indeed the $AcrA_{His6}$ protein (Fig. 3.5).

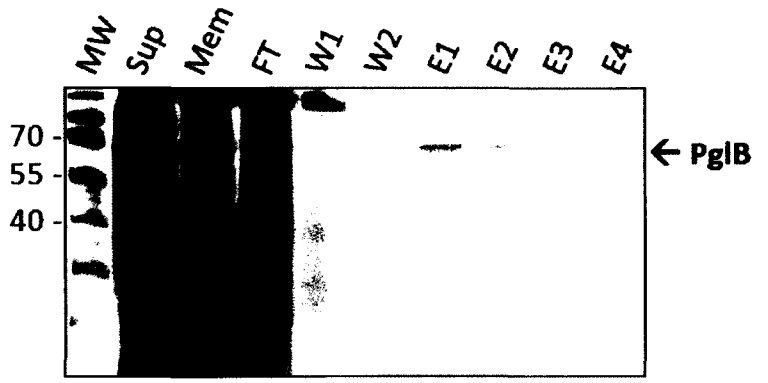

Figure 3.4. Purification of His6-tagged *C. jejuni* PglB by Ni-NTA affinity chromatography. PglB membrane fractions were purified from *E. coli* C43 strain expressing the *C. jejuni* wildtype PglB. Purified proteins were analyzed by SDS-PAGE and visualized by Coomassie Blue staining. The supernatant before ultracentrifugation of lysed *E. coli* cells was collected (Sup); membrane fractions of *E. coli* after ultracentrifugation (Mem); flow-through of samples during purification (FT); His-tagged PglB coupled to the Ni-NTA agarose washed with buffer containing 20mM imidazole (W1) and 50mM imidazole (W2); PglB-His was eluted with 1mL of elution buffer containing 250mM imidazole (E1-E4). PglB is indicated with an arrow. The molecular masses are shown in kDa on the left of the gel.

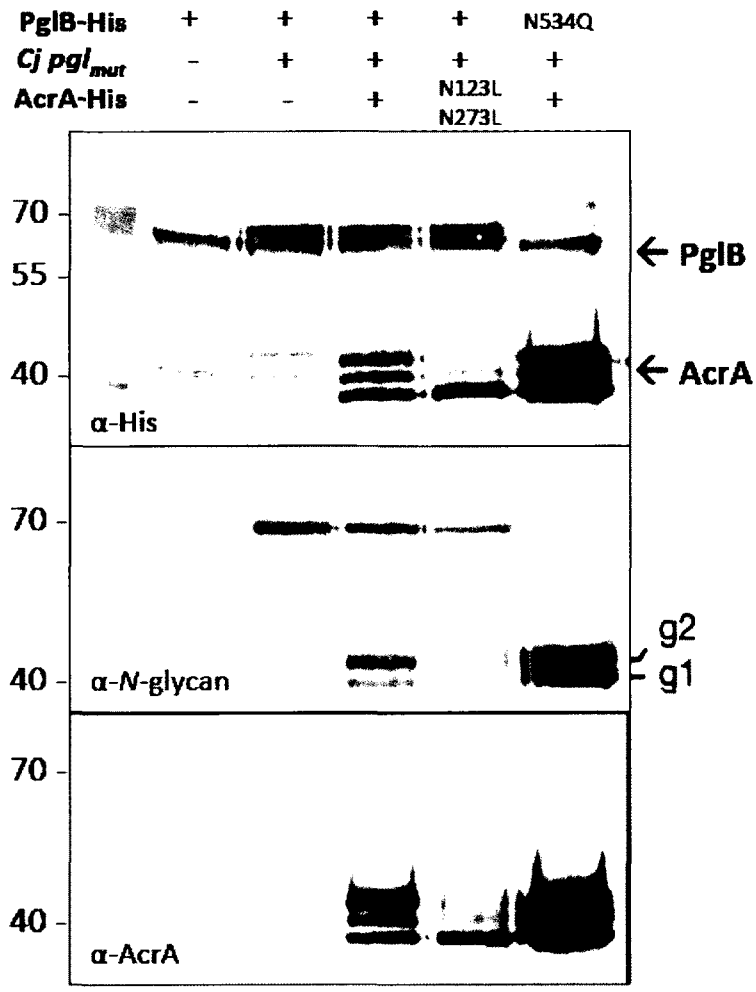

Figure 3.5. Glycosylation of *C. jejuni* PglB. PglB proteins were purified from *E. coli* C43 strain expressing PglB alone, PglB with the *C. jejuni pgl* operon, PglB with the *C. jejuni pgl* operon and AcrA acceptor protein, PglB with the *C. jejuni pgl* operon and an AcrA glycosylation mutant (N123L, N273L), as well as the PglB(N534Q) mutant with the *C. jejuni pgl* operon and AcrA acceptor protein. Purified PglB proteins were analyzed by Western blot using anti-His, anti-*N*-glycan and anti-AcrA specific antibodies. Protein bands which result from PglB degradation overlap with AcrA. PglB and AcrA are indicated with an arrow. The molecular masses are shown in kDa on the left of each blot.

3.3.5. In vitro *protein glycosylation*

Activity of the purified PglB proteins was assayed *in vitro* following a modified protocol (Faridmoayer *et al.*, 2008). The activity of the wildtype, unglycosylated form of PglB ($PglB_{WT}$), and the consensus sequence mutant $PglB_{N534Q}$ were examined. These two PglB proteins were analyzed in order to compare the hydrolase activities of the OTase, which may be influenced by the glycosylation status of the enzyme. Each of the PglB proteins was co-incubated with LLOs and AcrA overnight at 30°C. Western blot analysis of TCA precipitated proteins from the reaction mixture showed the glycosylation of AcrA using anti-*N*-glycan antibodies, which confirms that the PglB proteins are active (Fig. 3.6). Detection with anti-His antibodies, only showed the expression of AcrA, but did not show PglB, indicating that PglB is degraded over time.

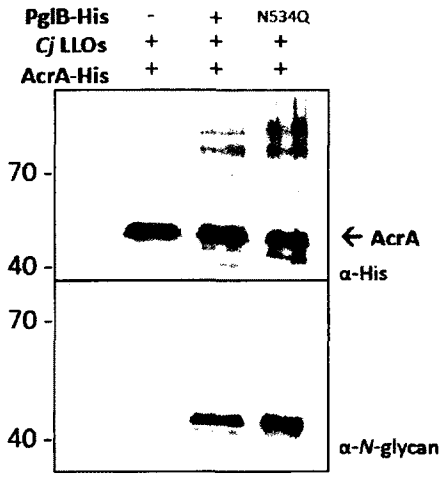

Figure 3.6. PglB activity assay. The activity of $PglB_{WT}$ and $PglB_{N534Q}$ were tested in an *in vitro* assay where purified proterins were co-incubated with LLOs and $AcrA_{His6}$. Western analysis shows the detection of AcrA in each of the reaction mixtures using anti-His antibodies and glycosylation of AcrA is observed only in

the presence of PglB as detected by anti-*N*-glycan antibodies. AcrA is indicated with an arrow. The molecular masses are shown in kDa on the left of each blot.

3.4. Discussion

Protein *N*-glycosylation has been demonstrated to play several important roles for the biology of *C. jejuni*. Disruption of the *N*-glycosylation pathway in *C. jejuni* does not affect its viability, but rather has multiple pleiotropic effects for the organism including the decrease of mouse and chicken colonization *in vivo* and the reduction of adherence and invasion of host epithelial cells *in vitro* (Szymanski *et al.*, 2002; Karlyshev *et al.*, 2004). The *N*-glycan has also been shown to be recognized by the host immune system during infection, where absence of the carbohydrate moiety results in a dramatic reduction in protein immunoreactivity with both human and animal sera (Szymanski *et al.*, 1999; Szymanski *et al.*, 2005). Although many different proteins are *N*-glycosylated in *C. jejuni*, not all protein functions are influenced by the glycan modification (Kakuda & DiRita, 2006; Davis *et al.*, 2009; Scott *et al.*, 2009). However, the *N*-glycan has been demonstrated to play a role in protein complex formation. The type IV secretion system in *C. jejuni* strain 81-176 contains the glycoprotein component VirB10 that, when mutated at its glycosylation site, resulted in reduced natural competence levels (Larsen *et al.*, 2004).

PglB is the key enzyme of the *N*-glycosylation process responsible for the transfer of oligosaccharides from the lipid carrier to asparagine residues within protein acceptors. The identification of a glycosylation site within *C. jejuni* PglB

led us to examine whether self-glycosylation of PglB affects its function. In this study, we constructed a point mutant in the consensus sequence found within the soluble domain of *C. jejuni* PglB (^{532}DY<u>N</u>QS^{536}). The $PglB_{N534Q}$ mutant was expressed and analyzed in both the heterologous *E. coli* host and the native *C. jejuni* host. Our experiments show that the $PglB_{N534Q}$ mutant does not influence protein glycosylation, but shows an affect in fOS production. This suggests that self-glycosylation may be modulating the hydrolase activity of PglB.

We also investigated the conditions under which PglB auto-glycosylates. PglB modifies itself *in vivo* when co-expressed with the *C. jejuni pgl*(*ΔpglB*) operon in the presence of either a wildtype or glycosylation mutant AcrA. It has been suggested that PglB changes into an active conformation when tightly bound to a peptide or protein at their glycosylation sequons (Lizak *et al.*, 2011). Our result that PglB self-glycosylates in the absence of AcrA, which served as an acceptor protein, suggests that PglB is binding to another PglB protein that carries the acceptor sequon in order to change into an active conformation and self-glycosylate. Specifically, PglB may not be self-glycosylating *per se*, but rather modifying other PglB proteins adjacently located along the membrane since it has been shown that hydrophobic integral membrane proteins are significant targets of *N*-glycan modification (Nothaft *et al.*, 2008; Scott *et al.*, 2009).

The crystal structure of the *C. jejuni* PglB C-terminal soluble domain showed that the consensus sequence at N534 is located in a surface exposed flexible loop (Maita *et al.*, 2010). It has been demonstrated that glycosylation sites are located in the flexible loops of folded proteins, which would permit access to

PglB (Kowarik *et al.*, 2006a). This is also supported by the crystal structure of the *C. jejuni* *N*-glycoprotein PEB3 where the key residues within the consensus sequence are well exposed on the surface (Rangarajan *et al.*, 2007). In addition, the three-dimensional structure of a truncated glycoprotein from *C. jejuni* ($AcrA^{61\text{-}210\Delta\Delta}$) containing the *N*-glycan of *C. jejuni* was determined by NMR spectroscopy demonstrating that the glycosylation site is also located in a flexible region modified with a well-defined rod-like glycan structure (Slynko *et al.*, 2009). Therefore, the location of the consensus sequence within PglB may allow for the control of allosteric effects.

N-linked glycosylation of OTase subunits in eukaryotes has been found to be important for enzyme function. In *S. cerevisiae*, the OTase is assembled into a multimeric complex composed of at least eight different subunits, each of which is required for enzymatic activity (Yan & Lennarz, 1999). Three of the essential proteins, Ost1, Wbp1, and Stt3, are *N*-linked glycoproteins. Mutagenesis studies were carried out at the *N*-glycosylation sites of each protein, but only the Stt3 glycosylation mutant exhibited defects in growth and OTase activity (Li *et al.*, 2005). These studies demonstrated that *N*-glycosylation of Stt3 is essential for its function in the OTase complex and it was suggested that the *N*-glycan might serve to support protein folding and structure (Li *et al.*, 2005).

PglB shares significant levels of sequence similarity with the Stt3 protein, which is the central catalytic subunit of the eukaryotic OTase complex (Wacker *et al.*, 2002; Yan & Lennarz, 2002). Unlike the eukaryotic Stt3 subunit, the OTase activity of PglB is not affected by removing the acceptor asparagine found within

the glycosylation site. However, fOS production decreases when PglB is not modified with the *N*-glycan, suggesting that PglB possesses an autoregulatory mechanism to modulate the hydrolase activity. Covalent modifications have been shown to regulate enzyme activity. In eukaryotes, β-N-acetylglucosamine (*O*-GlcNAc) modification of proteins at the hydroxyl groups of Ser and Thr residues has been described as a fundamental regulatory mechanism for cell division, metabolism, transcription and cell signaling (Mishra *et al.*, 2011). *O*-GlcNAc modification occurs within the nucleus and cytoplasm of the cell and involves the cycling of the *O*-GlcNAc transferase and *O*-GlcNAcase to attach and hydrolyze the *O*-GlcNAc from proteins, respectively (Hart & Akimoto, 2009). Specific roles for *O*-GlcNAc have been defined to include: signaling nuclear transport, regulating the assembly of protein complexes, and modulating proteolytic processing of proteins (Zachara & Hart, 2002). *O*-GlcNAc cycling shares some similarities with the regulation of protein phosphorylation.

In both eukaryotic and bacterial cells, protein phosphorylation plays an important role in the regulation of physiological processes (Deutscher & Saier Jr, 1988). In *E. coli*, the isocitrate dehydrogenase (IDH) is an enzyme that is involved in the Krebs cycle to control the regular flux of carbon at the critical junction between the Krebs cycle and the glyoxylate shunt (Zheng & Jia, 2010). IDH is regulated by the isocitrate dehydrogenase kinase/phosphatase, where IDH is inactivated by the addition of a phosphate and reactivated by removal of the phosphate. Another example of enzyme activity regulation is found in the bacterial RecA protein. RecA is a recombinase functioning in recombinational

DNA repair and is regulated at three levels: *recA* gene expression, RecA protein autoregulation, and modulation by other proteins (Cox, 2007). Autoregulation of the RecA protein function occurs at the C-terminal 17 amino acid residues, where removal of this peptide enhances almost all RecA functions (Lusetti *et al.*, 2003). Therefore, covalent modifications such as phosphorylation, acetylation, and even *N*-glycosylation regulate enzyme activity by modifying the properties of the enzyme, and thus its function.

Furthermore, *C. jejuni* produces free glycans in the periplasm that are present in tenfold excess to the *N*-linked form (Liu *et al.*, 2006). The release of fOS in *C. jejuni* has been reported to occur in response to the bacterial growth phase and to changes in the osmolarity of the environment, where fOS production decreases with increasing concentrations of salts and sucrose (Nothaft *et al.*, 2009). The periplasmic space of many proteobacteria have been found to contain large amounts of osmoregulated periplasmic glucans (OPGs), which are made of oligosaccharides that comprise a number of D-glucose sugars linked by β-glycosyidic bonds (Bohin, 2000). The concentration of OPGs in the periplasm increases during low osmotic shifts in order to increase the periplasmic volume of the cell (Bohin, 2000). Therefore, it has been suggested that the underlying regulatory mechanism of *C. jejuni* fOS is similar to OPGs for providing an appropriate periplasmic density of the cell (Nothaft *et al.*, 2009).

In eukaryotes, the release of fOS has been described to be generated in the lumen of the endoplasmic reticulum (ER) and in the cytosol (Suzuki & Funakoshi, 2006). Recent studies have demonstrated that fOS are derived from the processing

of *N*-glycans on misfolded glycoproteins during the ER-associated protein degradation pathway in yeast (Hirayama *et al.*, 2010). The mechanism of fOS release in the ER is unclear; however, removal of the glycans on proteins in the cytosol is achieved by peptide-*N*-glycanase (PNGase), which cleaves the amide bond in the side chain of modified asparagine residues (Suzuki & Funakoshi, 2006). It has also been described that in mammalian cells fOS are generated from lipid-linked oligosaccharides by the activity of a pyrophosphatase (Chantret & Moore, 2008). In addition, it has been suggested that fOS are generated by the OTase in the absence of a sufficient protein acceptor, and that the glycans are transferred from their lipid carriers to water instead (Spiro & Spiro, 1991; Villers *et al.*, 1994). Therefore, the role of fOS in higher eukaryotes remains unclear.

Alternatively, fOS production in *C. jejuni* is regulated by the osmotic changes in the environment. However, we observe reduction in fOS by the PglB consensus sequence mutant in standard laboratory conditions without the supplement of osmolytes as compared to its wildtype counterpart. Further investigation will be needed to unravel the biological significance of PglB self-glycosylation in *C. jejuni* in order to determine whether modification of PglB allows the enzyme to detect changes in osmolarity or whether the mutation on PglB affects its structure and function.

3.5. Acknowledgements

The authors would like to thank Abofu Alemka for providing the pCLpglB plasmid, Bernadette Beadle for providing the pACYC(*pgl*-Δ*pglB*) plasmid, Margaret Johnson for the preliminary analysis of the *Cj*PglB soluble domain crystal structure, M. Veronica Ielmini for guidance on PglB protein purification, and Markus Aebi (ETH, Zurich, Switzerland) for the hR6 antibodies. This work was supported by the Alberta Glycomics Centre. CMS is an Alberta Innovates Scholar.

3.6. Literature cited

Alaimo C, Catrein I, Morf L, Marolda CL, Callewaert N, Valvano MA, Feldman MF & Aebi M (2006) Two distinct but interchangeable mechanisms for flipping of lipid-linked oligosaccharides. *EMBO J* **25:** 967-976.

Apweiler R, Hermjakob H & Sharon N (1999) On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim Biophys Acta* **1473:** 4-8.

Bohin JP (2000) Osmoregulated periplasmic glucans in Proteobacteria. *FEMS Microbiol Lett* **186:** 11-19.

Burda P & Aebi M (1999) The dolichol pathway of *N*-linked glycosylation. *BBA-General Subjects* **1426:** 239-257.

Chantret I & Moore SE (2008) Free oligosaccharide regulation during mammalian protein N-glycosylation. *Glycobiology* **18:** 210-224.

Cox MM (2007) Regulation of bacterial RecA protein function. *Crit Rev Biochem Mol Biol* **42:** 41-63.

Davis LM, Kakuda T & DiRita VJ (2009) A *Campylobacter jejuni znuA* orthologue is essential for growth in low-zinc environments and chick colonization. *J Bacteriol* **191:** 1631-1640.

Deutscher J & Saier Jr MH (1988) Protein phosphorylation in bacteria - regulation of gene expression, transport functions, and metabolic processes Angewandte. *Chemie International Edition in English* **27:** 1040-1049.

Faridmoayer A, Fentabil MA, Haurat MF, Yi W, Woodward R, Wang PG & Feldman MF (2008) Extreme substrate promiscuity of the *Neisseria* oligosaccharyl transferase involved in protein *O*-glycosylation. *J Biol Chem* **283:** 34596-34604.

Feldman MF, Wacker M, Hernandez M, Hitchen PG, Marolda CL, Kowarik M, Morris HR, Dell A, Valvano MA & Aebi M (2005) Engineering *N*-linked protein glycosylation with diverse O antigen lipopolysaccharide structures in *Escherichia coli*. *Proc Natl Acad Sci USA* **102:** 3016.

Figurski DH & Helinski DR (1979) Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. *Proc Natl Acad Sci USA* **76:** 1648-1652.

Glover KJ, Weerapana E & Imperiali B (2005) *In vitro* assembly of the undecaprenylpyrophosphate-linked heptasaccharide for prokaryotic *N*-linked glycosylation. *Proc Natl Acad Sci USA* **102:** 14255.

Hebert DN, Garman SC & Molinari M (2005) The glycan code of the endoplasmic reticulum: asparagine-linked carbohydrates as protein maturation and quality-control tags. *Trends Cell Biol* **15:** 364-370.

Hirayama H, Seino J, Kitajima T, Jigami Y & Suzuki T (2010) Free oligosaccharides to monitor glycoprotein endoplasmic reticulum-associated degradation in *Saccharomyces cerevisiae. J Biol Chem* **285:** 12390-12404.

Ielmini MV & Feldman MF (2011) *Desulfovibrio desulfuricans* PglB homolog possesses oligosaccharyltransferase activity with relaxed glycan specificity and distinct protein acceptor sequence requirements. *Glycobiology* **21:** 734-742.

Kakuda T & DiRita VJ (2006) Cj1496c encodes a *Campylobacter jejuni* glycoprotein that influences invasion of human epithelial cells and colonization of the chick gastrointestinal tract. *Infect Immun* **74:** 4715-4723.

Karlyshev AV, Everest P, Linton D, Cawthraw S, Newell DG & Wren BW (2004) The *Campylobacter jejuni* general glycosylation system is important for attachment to human epithelial cells and in the colonization of chicks. *Microbiology* **150:** 1957-1964.

Kelleher DJ & Gilmore R (2006) An evolving view of the eukaryotic oligosaccharyltransferase. *Glycobiology* **16:** 47R.

Korlath JA, Osterholm MT, Judy LA, Forfang JC & Robinson RA (1985) A point-source outbreak of campylobacteriosis associated with consumption of raw milk. *J Infect Dis* **152:** 592-596.

Kowarik M, Numao S, Feldman MF, Schulz BL, Callewaert N, Kiermaier E, Catrein I & Aebi M (2006a) *N*-linked glycosylation of folded proteins by the bacterial oligosaccharyltransferase. *Science* **314:** 1148-1150.

Kowarik M, Young NM, Numao S, Schulz BL, Hug I, Callewaert N, Mills DC, Watson DC, Hernandez M, Kelly JF, Wacker M & Aebi M (2006b) Definition of the bacterial *N*-glycosylation site consensus sequence. *EMBO J* **25:** 1957-1966.

Labigne-Roussel A, Harel J & Tompkins L (1987) Gene transfer from *Escherichia coli* to *Campylobacter* species: development of shuttle vectors for genetic analysis of *Campylobacter jejuni. J Bacteriol* **169:** 5320-5323.

Larsen JC, Szymanski C & Guerry P (2004) *N*-linked protein glycosylation is required for full competence in *Campylobacter jejuni* 81-176. *J Bacteriol* **186:** 6508-6514.

Lefebre MD & Valvano MA (2002) Construction and evaluation of plasmid vectors optimized for constitutive and regulated gene expression in *Burkholderia cepacia* complex isolates. *Appl Environ Microbiol* **68:** 5956-5964.

Li G, Yan Q, Nita-Lazar A, Haltiwanger RS & Lennarz WJ (2005) Studies on the *N*-glycosylation of the subunits of oligosaccharyl transferase in *Saccharomyces cerevisiae. J Biol Chem* ***280*:** 1864.

Li L, Woodward R, Ding Y, Liu XW, Yi W, Bhatt VS, Chen M, Zhang LW & Wang PG (2010) Overexpression and topology of bacterial oligosaccharyltransferase PglB. *Biochem Biophys Res Commun* **394:** 1069-1074.

Linton D, Dorrell N, Hitchen PG, Amber S, Karlyshev AV, Morris HR, Dell A, Valvano MA, Aebi M & Wren BW (2005) Functional analysis of the *Campylobacter jejuni N*-linked protein glycosylation pathway. *Mol Microbiol* **55:** 1695-1703.

Liu X, McNally DJ, Nothaft H, Szymanski CM, Brisson JR & Li J (2006) Mass spectrometry-based glycomics strategy for exploring *N*-linked glycosylation in eukaryotes and bacteria. *Anal Chem* **78:** 6081-6087.

Lizak C, Gerber S, Numao S, Aebi M & Locher KP (2011) X-ray structure of a bacterial oligosaccharyltransferase. *Nature* **474:** 350-355.

Lusetti SL, Wood EA, Fleming CD, Modica MJ, Korth J, Abbott L, Dwyer DW, Roca AI, Inman RB & Cox MM (2003) C-terminal deletions of the *Escherichia coli* RecA protein. Characterization of *in vivo* and *in vitro* effects. *J Biol Chem* **278:** 16372-16380.

Maita N, Nyirenda J, Igura M, Kamishikiryo J & Kohda D (2010) Comparative structural biology of Eubacterial and Archaeal oligosaccharyltransferases. *J Biol Chem* **285:** 4941.

Nothaft H, Amber S, Aebi M & Szymanski CM (2008) *N*-linked protein glycosylation in Campylobacter. *Campylobacter*, Vol. 3 (Nachamkin I, Szymanski CM & Blaser MJ, eds), pp. 447-469. ASM Press, Washington, DC.

Nothaft H, Liu X, McNally DJ, Li J & Szymanski CM (2009) Study of free oligosaccharides derived from the bacterial *N*-glycosylation pathway. *Proc Natl Acad Sci USA* **106:** 15019.

Nothaft H & Szymanski CM (2010) Protein glycosylation in bacteria: sweeter than ever. *Nat Rev Microbiol* **8:** 765-778.

Parkhill J, Wren BW, Mungall K, Ketley JM, Churcher C, Basham D, Chillingworth T, Davies RM, Feltwell T & Holroyd S (2000) The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature* **403:** 665-668.

Rangarajan ES, Bhatia S, Watson DC, Munger C, Cygler M, Matte A & Young NM (2007) Structural context for protein *N*-glycosylation in bacteria: The structure of PEB3, an adhesin from *Campylobacter jejuni.* Protein Science: A Publication of the Protein Society **16:** 990.

Reid CW, Stupak J, Szymanski CM & Li J (2009) Analysis of bacterial lipid-linked oligosaccharide intermediates using porous graphitic carbon liquid chromatography-electrospray ionization mass spectrometry: heterogeneity in the polyisoprenyl carrier revealed. *Anal Chem* **81:** 8472-8478.

Sambrook J & Russell DW (2001) *Molecular cloning: a laboratory manual.* CSHL press.

Schwarz F, Lizak C, Fan YY, Fleurkens S, Kowarik M & Aebi M (2011) Relaxed acceptor site specificity of bacterial oligosaccharyltransferase *in vivo.* *Glycobiology* **21:** 45-54.

Scott NE, Bogema DR, Connolly AM, Falconer L, Djordjevic SP & Cordwell SJ (2009) Mass spectrometric characterization of the surface-associated 42 kDa lipoprotein JlpA as a glycosylated antigen in strains of *Campylobacter jejuni. J Proteome Res* **8:** 4654-4664.

Scott NE, Parker BL, Connolly AM *et al.* (2011) Simultaneous glycan-peptide characterization using hydrophilic interaction chromatography and parallel fragmentation by CID, higher energy collisional dissociation, and electron transfer dissociation MS applied to the *N*-linked glycoproteome of *Campylobacter jejuni. Mol Cell Proteomics* **10:** M000031-MCP201.

Slynko V, Schubert M, Numao S, Kowarik M, Aebi M & Allain FH (2009) NMR structure determination of a segmentally labeled glycoprotein using *in vitro* glycosylation. *J Am Chem Soc* **131:** 1274-1281.

Spiro MJ & Spiro RG (1991) Potential regulation of N-glycosylation precursor through oligosaccharide-lipid hydrolase action and glucosyltransferase-glucosidase shuttle. *J Biol Chem* **266:** 5311-5317.

Suzuki T & Funakoshi Y (2006) Free N-linked oligosaccharide chains: formation and degradation. *Glycoconj J* **23:** 291-302.

Szymanski CM, Yao R, Ewing CP, Trust TJ & Guerry P (1999) Evidence for a system of general protein glycosylation in *Campylobacter jejuni. Mol Microbiol* **32:** 1022-1030.

Szymanski CM, Burr DH & Guerry P (2002) *Campylobacter* protein glycosylation affects host cell interactions. *Infect Immun* **70:** 2242-2244.

Szymanski CM, Goon S, Allan B & Guerry P (2005) Protein glycosylation in *Campylobacter*. Campylobacter: molecular and cellular biology (Ketley, J.M. and Konkel, M.E., ed), pp. 259-274. Horizon Bioscience, Norfolk, U.K.

Varki A (1993) Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* **3:** 97.

Villers C, Cacan R, Mir AM, Labiau O & Verbert A (1994) Release of oligomannoside-type glycans as a marker of the degradation of newly synthesized glycoproteins. *Biochem J* **298 (Pt 1):** 135-142.

Wacker M, Linton D, Hitchen PG, Nita-Lazar M, Haslam SM, North SJ, Panico M, Morris HR, Dell A & Wren BW (2002) *N*-linked glycosylation in *Campylobacter jejuni* and its functional transfer into *E. coli. Science* **298:** 1790.

Wacker M, Feldman MF, Callewaert N, Kowarik M, Clarke BR, Pohl NL, Hernandez M, Vines ED, Valvano MA & Whitfield C (2006) Substrate specificity of bacterial oligosaccharyltransferase suggests a common transfer mechanism for the bacterial and eukaryotic systems. *Proc Natl Acad Sci USA* **103:** 7088.

Yan Q & Lennarz WJ (1999) Oligosaccharyltransferase: a complex multisubunit enzyme of the endoplasmic reticulum. *Biochem Biophys Res Commun* **266:** 684-689.

Yan Q & Lennarz WJ (2002) Studies on the function of oligosaccharyl transferase subunits. Stt3p is directly involved in the glycosylation process. *J Biol Chem* **277:** 47692.

Young NM, Brisson JR, Kelly J, Watson DC, Tessier L, Lanthier PH, Jarrell HC, Cadotte N, St. Michael F & Aberg E (2002) Structure of the *N*-linked glycan present on multiple glycoproteins in the Gram-negative bacterium, *Campylobacter jejuni. J Biol Chem* **277:** 42530-42539.

Zheng J & Jia Z (2010) Structure of the bifunctional isocitrate dehydrogenase kinase/phosphatase. *Nature* **465:** 961-965.

Chapter 4

General Discussion and Conclusions

4.1. Research purpose

My research project entailed the functional characterization of the *Campylobacter* PglB enzyme, an OTase responsible for the transfer of glycans onto asparagine residues of acceptor proteins during the process of *N*-linked protein glycosylation. The objectives of this research were to identify new amino acid residues that are important for *C. jejuni* PglB (*Cj*PglB) activity, compare sequenced *Campylobacter* PglB enzymes, and explore the possible autoregulatory mechanism of PglB. Identification of important residues that are required for the function and/or structure of PglB would further enhance our current understanding of the underlying mechanism and principles of *N*-glycosylation. Moreover, my project focused on characterizing *Campylobacter* PglB enzymes in order to improve OTase activity for glycoprotein engineering.

4.2. Summary and future directions

At the onset of this thesis project, the only residues shown to be required for *Cj*PglB activity were found within the highly conserved ^{457}WWDYG^{461} motif (W458, D459) located within the soluble C-terminal domain (Wacker *et al.*, 2002). Another conserved motif, DXXK, was identified within the C-terminal domain of a functionally equivalent archaeal OTase of *P. furiosus* and was also shown to be required for activity in eukaryotes (Igura *et al.*, 2008). The DXXK sequence is also present within *Cj*PglB (residues 519-522). Coincident with my studies, a 2.8 Å resolution crystal structure of the C-terminal soluble domain of *Cj*PglB was published, which revealed another catalytic motif in PglB where

^{568}MXXI^{571} residues occupy the same positions as the DXXK residues in the archaeal OTase (Maita *et al.*, 2010).

The first objective of my Master's project was to identify additional amino acids required for *Cj*PglB function by using a bioinformatic approach and generating point mutations at specific amino acid sites (see Chapter 2). Through multiple sequence alignments, we identified a second ^{475}DXXK^{478} motif located 12 amino acids away from the signature WWDYG motif that is conserved in all sequenced *Camplobacter* PglB sequences (residues 475-478 in *Cj*PglB) and other epsilon- and delta-Proteobacteria. I examined whether this highly conserved DXXK sequence is necessary for *Cj*PglB activity by generating amino acid substitutions at the D and K sites. *In vivo* PglB activity was investigated in both the heterologous *E. coli* system and the native *C. jejuni* host using AcrA as the model acceptor protein. I showed that protein glycosylation is considerably reduced for the D475V point mutant in *E. coli* and for the D475V and K478V point mutants in *C. jejuni*. Furthermore, the requirement of the DXXK motif for release of fOS was examined by semi-quantitative mass spectrometry (in collaboration with Dr. Jianjun Li, NRC, Ottawa). We found that the *Cj*PglB D475V and K478V single point mutants showed a corresponding decrease of this pathway product. In collaboration with Dr. Ken Ng (University of Calgary), the available crystal structure of the soluble portion of *Cj*PglB was analyzed to reveal that the DXXK motif is located directly above the highly conserved WWDYG motif. Modeling of the available crystal structure for *Cj*PglB revealed that the D and K residues do not interact with each other, but may form salt bridges to hold

PglB in its active conformation. Loss of the two wild-type salt-bridges may alter the conformation of the DXXK loop away from the active conformation. We propose that the DXXK motif plays an important structural role for PglB OTase and hydrolase activities.

Future work is still required to develop a quantitative assay that can be used to measure the OTase activity of the PglB mutants that have been constructed. Currently, the OTase efficiency of PglB was analyzed by detecting the glycosylation state of AcrA by Western blotting. A reliable, faster, and more sensitive detection method needs to be developed for glycoprotein-quantification. I have been working on an ELISA-based method to quantify PglB activity by analyzing the glycosylation state of AcrA (see Appendix 1). I was using nickel-coated 96-well plates that allow binding of His-tagged AcrA. This one-step purification method would allow me to quantify the ratio of glycosylated AcrA over the total number of AcrA bound to the wells of the plate using *N*-glycan- and AcrA-specific antisera. However, I was obtaining values with high standard deviations due to the high level of background observed in each sample. After having little success with the optimization process of this ELISA-based assay, we decided to quantify the intensity of the protein bands observed on the Western blot membranes by using densitometry.

In addition, another study is being carried out in the Szymanski lab to develop a quantitative assay to measure the OTase activity of PglB. Dr. Abofu Alemka, a post-doctoral fellow in our lab, is working in collaboration with Dr. Chris Cairo (Department of Chemistry, University of Alberta) to develop a

general protease-protection assay to monitor the OTase activity of PglB. In this approach, a short peptide is synthesized with a fluorescent dye/quencher pair signal that contains a unique protease site adjacent to the glycosylation site (D/E-X_1-N-X_2-S/T). The peptide substrate is subjected to an *in vitro* glycosylation reaction and then treated with a protease, followed by the measurement of excited light that is absorbed at a different wavelength depending on whether the peptide is intact or cleaved. This assay will be useful to compare the various *Cj*PglB mutant constructs to the wild-type *Cj*PglB enzyme generated in this study, and also for comparing activities of the PglB enzymes from sequenced *Campylobacter* species and other proteobacteria.

The second objective of my project was to compare the OTase activity of sequenced *Campylobacter* PglB enzymes (see Chapter 2). We analyzed the multiple sequence alignment of *Campylobacter* PglB enzymes once more and we identified that further downstream of the DXXK sequence there existed a conserved proline (P487 in *Cj*PglB) or serine (S493 in *Cl*PglB). The aligned amino acids of the two PglB enzymes were swapped and the *pglB* mutants were investigated in *E. coli*. Unlike the *Cj*PglB mutant that had no considerable affect on glycosylation, I showed that the *Cl*PglB point mutant glycosylates AcrA at a higher efficiency than the wildtype *Cl*PglB. However, there was no difference in glycosylation when each of the mutant variants of *Cj*PglB and *Cl*PglB were complemented into a *C. jejuni pglB* knock-out mutant as compared to their wildtype counterparts. Based on previous results showing full modification of AcrA in wildtype *C. jejuni*, I think that there might be other components of the

cell, which may be interacting with PglB that are not present in *E. coli*. This indicates to us that the *N*-glycosylation system is perfectly optimized in the native host, where expression levels of glycosyltransferases, nucleotide sugar donors, and polypeptide acceptors are synchronized. Further work is required to determine the influence of the amino acid exchanges in the *Cj*PglB and *Cl*PglB as well as the difference in activity in *E. coli* and *C. jejuni*. However, to explain the improved efficiency of the *Cl*PglB point mutant, we examined the recently published 3.4 Å resolution crystal structure of the full-length *Cl*PglB (Lizak *et al.*, 2011) in collaboration with Dr. Ng (University of Calgary). We propose that substituting the serine residue with a proline might increase the flexibility of this part of the PglB structure, which is located at a kink in the alpha-helix near the proposed protein-binding site of the enzyme. In this study, I showed that it is possible that a single amino acid substitution in PglB can improve OTase activity in *E. coli*.

Furthermore, investigating the biotechnological application of the improved *Cl*PglB variant can advance this study by possibly demonstrating glycosylation of an acceptor protein that carries the short eukaryotic sequon (**N**-X-S/T). The *Cl*PglB was previously shown to glycosylate asparagine residues found within the eukaryotic sequon in *E. coli*, albeit with very low efficiency (Schwarz *et al.*, 2010). During the later part of my studies, I have constructed a *C. jejuni* AcrA mutant (D121A, D271A) to contain the short eukaryotic sequons in order to examine whether the *Cl*PglB point mutant can also modify the **N**-X-S/T at an improved efficiency, a site that cannot be efficiently modified by *Cj*PglB (Schwarz *et al.*, 2011). A future experiment that can be carried out involves co-

expressing the AcrA (D121A, D271A) double mutant with the *C. jejuni pgl* operon containing a deletion in the *C. jejuni pglB* and the *Cj*PglB point mutant. The advantage of an OTase that can modify proteins with the short eukaryotic sequon is that the residues can be engineered into the flexible loops of the desired proteins more easily. The discovery of new OTases with enhanced activity or the improvement of previously characterized OTases by mutagenesis at specific residues of the protein is crucial for expanding current strategies for glycoengineering.

The third objective of my research project was to explore the possible autoregulatory mechanism of PglB (see Chapter 3). I constructed a point mutant in the consensus sequence of *Cj*PglB (N534Q) and expressed and analyzed the PglB mutant in both the heterologous *E. coli* host and the native *C. jejuni* host. I showed that there is no change in protein glycosylation with the *pglB* consensus sequence mutant (N534Q) in either *E. coli* or *C. jejuni*. However, in collaboration with Dr. Li (NRC, Ottawa) we showed that the N534Q mutant influences fOS production, suggesting that self-glycosylation of PglB may be modulating the two enzyme activities of PglB. I also tested the conditions under which PglB modifies itself and showed that PglB can auto-glycosylate *in vivo* in the presence of the *pgl* operon alone.

I have now optimized the purification of full-length *Cj*PglB (under the advisement of Dr. M Veronica Ielmini, University of Alberta) and showed that the purified protein is active *in vitro*. The purified active PglB enzyme can then be used to carry out the *in vitro* analysis of PglB fOS production, where the

hydrolase activities of the glycosylated and mutant PglB proteins can be measured and compared to nonglycosylated PglB. Based on the reduction in fOS release by the N534Q mutant in a *C. jejuni pglB* knockout, I hypothesize that the glycosylated form of PglB will release more fOS than the unmodified PglB and that the N534Q mutant will release less than the wildtype PglB. Future work is required to deduce the importance of PglB modification for its hydrolase activity. The semi-quantitative mass spectrometry assay that we have used in the past with Dr. Li (NRC, Ottawa) can also be used here to quantify fOS. In addition, Dr. Harald Nothaft, in our lab, is developing a complementary method using high performance anionic exchange chromatography with pulsed amperometric detection (HPAEC-PAD), a technique that can also be exploited for fOS quantitation. Additional work that can be carried out to continue this project involves the development of a real-time mass spectrometry method in collaboration with Dr. Klassen to obtain kinetic data of PglB hydrolysis of LLOs. Moreover, mass spectrometry methods can be developed in order to evaluate the relative rates of protein glycosylation by PglB, providing a quantitative assay necessary for the measurement of the OTase activity. And, the mass spectrometric analysis can be extended to obtain an estimate of PglB wildtype and mutant binding affinities for its LLO substrate. These studies will shed new light on the mechanisms of protein *N*-glycosylation and fOS release.

Currently, the glycosylation mechanism of PglB is proposed to involve a three-state catalytic cycle: i) binding of the polypeptide at the consensus sequon to form the catalytic site and activate the amide nitrogen of the acceptor asparagine,

ii) binding of the LLO that results in a nucleophilic attack of the activated amide nitrogen and formation of the glycosidic bond, and iii) disengagement of the enzyme to allow dissociation of the newly formed glycopeptides (Lizak *et al.*, 2011). However, the suggested mechanism of PglB glycosylation activity does not explain fOS production. The conserved WWDYG motif was shown to be necessary for *in vivo* and *in vitro* PglB fOS release (Nothaft *et al.*, 2009); therefore, further studies are required to elucidate the production of fOS in *C. jejuni* and its possible regulation by PglB self-glycosylation.

4.3. Conclusions and perspectives

The key objective of this thesis was to characterize the bacterial OTase in order to extend our current knowledge of the basic processes of *N*-linked protein glycosylation. Analysis of the OTase and hydrolase activities, as well as the identification of *N*-glycan structures in other organisms will shed light on the evolution of the *N*-glycosylation system across Bacteria. Also, the study of fOS production adds further complexity to the analyses of bacterial glycomes (Nothaft *et al.*, 2009), and further characterization of these free glycans is still required. Some of the research in our lab involves the analysis of *N*-linked glycans and fOS from various sequenced *Campylobacter* species. These studies have demonstrated the production of different glycan structures that vary at the species level, which might have evolved with respect to the ecological niche that each species inhabits (Nothaft *et al.*, manuscript in preparation). Although there is no direct correlation between *N*-glycosylation and pathogenicity in other *Campylobacter* species, this protein modification presents an advantage for *C. jejuni* to colonize its natural

niche (the gastrointestinal tract of chickens). Disruption of the *N*-glycosylation pathway in *C. jejuni* does not affect its viability, but disturbs many biological functions in this organism (Nothaft *et al.*, 2008) and the conservation of this process in a growing number of *Campylobacter* species supports a key role in the physiology of these bacteria (Nothaft and Szymanski, 2010).

Although bacterial *N*-glycosylation is best characterized in *C. jejuni*, there has been a surge of research in the last few years that has led to the discovery of glycosylation systems in other organisms. The discovery of the bacterial OTases has made the production of recombinant glycoproteins possible for application in biotechnology. Current debate focuses on the efficacy of the OTases for use in glycoprotein engineering. There are still a considerable number of hurdles to overcome in order to engineer glycoconjugate vaccines of any desired carbohydrate structure. However, further exploration of novel PglB enzymes and exploitation into the versatility of other OTases, such as the *Neisseria* PglL, which can be used to transfer many diverse glycan polysaccharides, including those that contain a hexose at the reducing end, remain to be optimized. Exploring the diversity of bacterial OTases will help to identify active enzymes with different specificities, which may be exploited for glycoprotein engineering.

4.4. Literature cited

Igura M, Maita N, Kamishikiryo J, Yamada M, Obita T, Maenaka K & Kohda D (2008) Structure-guided identification of a new catalytic motif of oligosaccharyltransferase. *EMBO J* **27:** 234-243.

Lizak C, Gerber S, Numao S, Aebi M & Locher KP (2011) X-ray structure of a bacterial oligosaccharyltransferase. *Nature* **474:** 350-355.

Maita N, Nyirenda J, Igura M, Kamishikiryo J & Kohda D (2010) Comparative structural biology of Eubacterial and Archaeal oligosaccharyltransferases. *J Biol Chem* **285:** 4941.

Nothaft H, Amber S, Aebi M & Szymanski CM (2008) *N*-linked protein glycosylation in Campylobacter. *Campylobacter*, Vol. 3 (Nachamkin I, Szymanski CM & Blaser MJ, eds), pp. 447-469. ASM Press, Washington, DC.

Nothaft H, Liu X, McNally DJ, Li J & Szymanski CM (2009) Study of free oligosaccharides derived from the bacterial *N*-glycosylation pathway. *Proc Natl Acad Sci USA* ***106*:** 15019.

Nothaft H & Szymanski CM (2010) Protein glycosylation in bacteria: sweeter than ever. *Nat Rev Microbiol* **8:** 765-778.

Schwarz F, Lizak C, Fan YY, Fleurkens S, Kowarik M & Aebi M (2011) Relaxed acceptor site specificity of bacterial oligosaccharyltransferase *in vivo*. *Glycobiology* **21:** 45-54.

Wacker M, Linton D, Hitchen PG, Nita-Lazar M, Haslam SM, North SJ, Panico M, Morris HR, Dell A & Wren BW (2002) *N*-linked glycosylation in *Campylobacter jejuni* and its functional transfer into *E. coli*. *Science* **298:** 1790.

Appendix

A.1. Development of an ELISA-based quantitation assay for protein glycosylation

A.1.1. Introduction

Employing a bioinformatic comparison of PglB protein sequences and subsequently generating point mutations at specific residues of the protein have allowed for the functional characterization of *Campylobacter* PglB enzymes in this thesis. The OTase activity of PglB is generally examined by immunodetection on Western blots or mobility shift analysis on SDS-PAGE. However, a quantitative assay that can be used to measure PglB *N*-glycosylation activity is needed. Such an approach would make it possible to screen glycoprotein production by PglB in a high throughput manner and to compare the activities of OTases from other species. The objective of this study was to develop an ELISA-based method for analyzing the glycosylation status of the hexahistidine-tagged AcrA protein in order to detect the different levels of glycoprotein biosynthesis when comparing PglB enzymes. The *in vivo* approach established for recombinant glycoprotein production was used, whereby the *C. jejuni pgl* operon was functionally transferred into *E. coli* in the presence of the $AcrA_{His6}$ acceptor protein and PglB (Wacker *et al.*, 2002; Feldman *et al.*, 2005).

A.1.2. Methods

The bacterial strains and plasmids used in this study are listed in Table A-1.1. The *E. coli* constructs used in this study are also listed in Tables 2.1 and 3.1 as previously described in Chapters 2 and 3 of this thesis. Expression of *C. jejuni*

$AcrA_{His6}$ (plasmid pWA2) with the *pgl* operon containing a deletion in *pglB* (pACYC(*pgl-ΔpglB*) and PglB wildtype and mutant alleles in *E. coli* CLM24 were performed as described in section 2.2.3. The purification of $AcrA_{His6}$ also follows the same protocol as described in section 2.2.3. Periplasmic extracts were performed as described (Feldman *et al.*, 2005).

Table A-1.1. Bacterial strains and plasmids used in this study

Strain or plasmid	Characteristic	Source
E. coli CLM24	W3110, Δ*waaL*	Feldman *et al.*, 2005
E. coli BL21	F⁻ *ompT gal dcm lon hsdS$_B$(r$_B$⁻ m$_B$⁻) λ*(DE3 [*lacI lacUV5*-T7 gene 1 *indI sam7 nin5*])	Wacker *et al.*, 2002
pACYC(*pgl-ΔpglB*)	Encodes the *C. jejuni pgl* locus with a deletion in PglB, Cm^R	Provided by Bernadette Beadle
pWA2	Soluble periplasmic *C. jejuni acr*A_{6His} under control of Tet promoter, in pBR322, Amp^R	Feldman *et al.*, 2005
pMLBAD	Cloning vector, arabinose-inducible, Tmp^R	Lefebre & Valvano, 2002
pMAF10	HA-tagged *C. jejuni pglB* cloned in pMLBAD, Tmp^R	Feldman *et al.*, 2005
pWA1	HA-tagged *C. jejuni pglB*$_{mut}$ cloned in pMLBAD, Tmp^R	Feldman *et al.*, 2005
pYB1D	*C. jejuni pglB*(D475A)$_{HA}$ cloned into pMLBAD, Tmp^R	This study (see Ch. 2)
pYB2D	*C. jejuni pglB*(D475V)$_{HA}$ cloned into pMLBAD, Tmp^R	This study (see Ch. 2)
pYB1K	*C. jejuni pglB*(K478A)$_{HA}$ cloned into pMLBAD, Tmp^R	This study (see Ch. 2)
pYB2K	*C. jejuni pglB*(K478V)$_{HA}$ cloned into pMLBAD, Tmp^R	This study (see Ch. 2)
pYB1A	*C. jejuni pglB*(D475A, K478A)$_{HA}$ cloned into pMLBAD, Tmp^R	This study (see Ch. 2)
pYB3K	*C. jejuni pglB*(D475K)$_{HA}$ cloned into pMLBAD, Tmp^R	This study (see Ch. 2)
pYB4K	*C. jejuni pglB*(D475K, K478D)$_{HA}$ cloned into pMLBAD, Tmp^R	This study (see Ch. 2)
pNQ1	*C. jejuni pglB*(N534Q)$_{HA}$ cloned in pMLBAD, Tmp^R	This study (see Ch. 3)

ELISA

The Pierce® nickel coated 96-well plates (Thermo Scientific) were used to analyze the glycosylation status of AcrA. The bacterial lysates can be added directly to the plates since the nickel coating allows selective binding of the His-tagged AcrA proteins. The plates were supplied pre-blocked with bovine serum albumin.

$AcrA_{His6}$ proteins were bound to the wells and prepared for immunodetection following the manufacturer's protocol. Briefly, the wells were washed three times with phosphate-buffer-saline containing 0.05% Tween (PBST). Whole cell lysates of *E. coli* CLM24 diluted in 1xPBS (pH 7.2) were normalized to the same concentration and 100 µL of each sample per well was added. The binding reaction between the $AcrA_{His6}$ proteins and the nickel plate was performed by shaking overnight at 4°C. After the wells were washed three times with PBST, the samples were incubated with two different antibodies in a volume of 100 µL/well: rabbit-α-AcrA antiserum (kindly provided by Dr. Mario Feldman) and rabbit-α-*C. jejuni*-*N*-glycan antiserum (prepared in the Szymanski lab). The samples were incubated for 1 hr at room temperature with shaking, and subsequently washed three times with PBST to remove unbound antisera. The bound $AcrA_{His6}$ samples were then incubated with 100 µL of alkaline phosphatase conjugated anti-rabbit IgG (Santa Cruz) for 1 hr. The antibody dilutions used here are described in section 2.2.6. The wells were finally washed three times with PBST followed by the colour development of bound proteins by the addition of 100 µL of 1-Step™ PNPP (*p*-nitrophenyl phosphate disodium salt) (Thermo Scientific). PNPP is a substrate used for detecting alkaline phosphatase, producing a water-soluble yellow product from the reaction that absorbs light at 405 nm (Synergy HT multi-mode microplate reader, BioTek®). The colour development reaction was measured every 5 minutes for 60 min.

SDS-PAGE and Western blotting

Confirmation of protein expression was analyzed by SDS-PAGE on 10% gels and stained by using Coomassie Blue (R-250). Immunodetection by Western blotting was performed as previously described in section 2.2.6.

A.1.3. Results and Discussion

In this study, I have been working on developing an ELISA-based method to quantify PglB activity by analyzing the glycosylation status of AcrA. Nickel-coated 96-well plates were used to allow binding of His-tagged AcrA proteins. This one-step purification method would allow me to quantify the ratio of glycosylated AcrA in comparison to the total amount of AcrA bound to the wells of the plate using AcrA and *N*-glycan specific antisera. Two sets of samples were analyzed for each *E. coli* strain co-expressing AcrA, the *pgl* operon, with the various PglB wildtype and mutant constructs. One set of protein samples was incubated with anti-AcrA serum to detect the total AcrA protein bound to the well, while the other set of protein samples was incubated with anti-*N*-glycan serum in order to detect the fraction of AcrA protein modified with the *N*-glycan. Both sets of samples were analyzed in technical triplicate. *E. coli* CLM24 cells expressing $AcrA_{His6}$ and the *C. jejuni pgl* operon with the inactive PglB (WWDYG mutant; W458A, D459A) served as negative controls to eliminate the background that may result from any non-specific binding of *E. coli* proteins or interaction with antibodies.

The AcrA proteins expressed in the various *E. coli* constructs were purified from whole cell lysates by coupling the His-tagged proteins to the nickel-

coated ELISA plates. Modification of AcrA was detected with *N*-glycan specific antisera in comparison to the total amount of AcrA protein bound to the wells using anti-AcrA antibodies. The resulting *N*-glycan/AcrA ratios are expressed in percentages after subtracting the ratio value of the $PglB_{WWDYG}$ mutant (Figure A-1.1). The OTase activities of the DXXK mutants in PglB were examined by analyzing AcrA glycosylation by Western blotting (see chapter 2). When these samples were analyzed by this ELISA-based approach, the glycosylation state of AcrA followed a similar trend comparable to the data observed by Western analysis; however, there were high standard deviations due to the high level of background observed in each sample.

In order to improve the assay, the ELISA experiment was repeated using AcrA purified from whole-cells using nickel affinity chromatography (NTA agarose, Qiagen). The proteins were dialyzed against 1xPBS to remove the imidazole that was used to elute the His-tagged proteins. The purified AcrA proteins were coupled to the nickel-coated 96-well plates and incubated with the antibodies as described above. When the levels of bound proteins were measured, high amounts of background was still observed in the analysis. To further optimize this assay, *E. coli* CLM24 expressing $AcrA_{His6}$, the *C. jejuni pgl* operon with a deleted *pglB*, and the pMLBAD empty vector, used in this study for cloning *pglB*, served as the negative control in addition to the inactive $PglB_{WWDYG}$ mutant. In lieu of using bacterial lysates, periplasmic extracts were prepared from the *E. coli* CLM24 strains. For optimization, *E. coli* constructs containing AcrA and the *pgl* operon expressing wildtype PglB, inactive $PglB_{WWDYG}$, $PglB_{N534Q}$

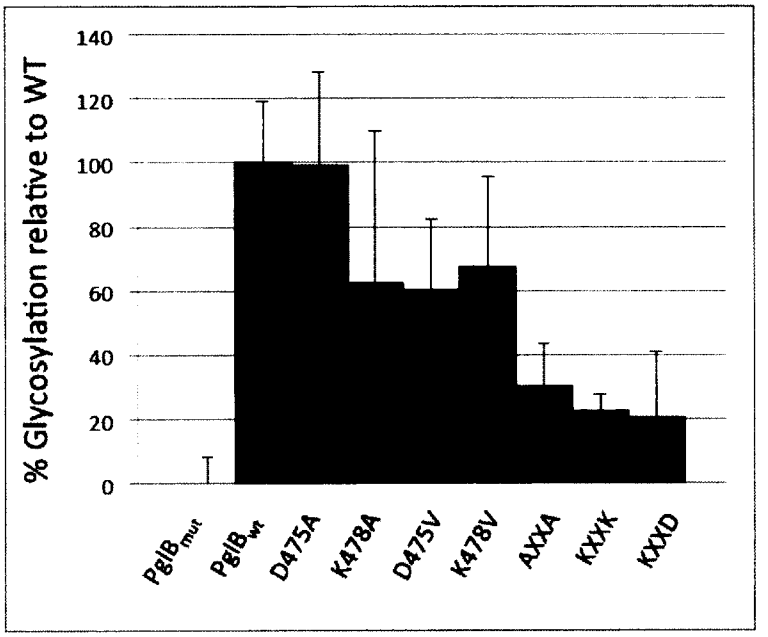

Figure A-1.1. Whole-cell lysates prepared from *E. coli* cells were applied to the nickel-coated ELISA-plate to determine the relative levels of AcrA glycosylation. *E. coli* CLM24 expressing the *C. jejuni pgl* operon with a deleted *pglB*, the *C. jejuni* $AcrA_{His6}$, and PglB proteins were analyzed. The inactive PglB (mut) was used as negative control and the wildtype PglB (wt) served as positive control. The various DXXK mutants in PglB were tested for OTase activity: D475A, K478A, D475V, K478V, AXXA, KXXK, and KXXD. The ratios between the signal detected with anti-*N*-glycan antisera and the signal detected with anti-AcrA antisera were expressed as percentages of the wildtype PglB (wt) glycosylation efficiency (100%) after the ratio-value of the inactive PglB (mut) were subtracted. Technical samples were analyzed in triplicate and averages compared. Error bars indicate standard deviation.

mutant (see chapter 3), and the empty vector were used. The periplasmic samples were analyzed by SDS-PAGE followed by immunodetection with AcrA specific antiserum confirming the overexpression of AcrA in the samples (Figure A-1.2). However, when the periplasmic samples were applied to the ELISA plates and detected with anti-*N*-glycan and anti-AcrA antibodies, the level of background was as high as observed in the whole cell lysates. Unfortunately, the empty vector control even showed reactivity to the *N*-glycan specific antisera, which is unexpected because there is no PglB enzyme expressed in this strain. However, it appears that the high background level is due to the interaction between the *N*-glycan antisera and other non-specific *E. coli* proteins.

After having little success with the optimization process of this ELISA-based method, we decided to quantify the intensity of the protein bands observed on the Western blot membranes by using a densitometer. In the meanwhile, further optimization of this ELISA-based quantitation assay is needed to measure the OTase activity of *Campylobacter* PglB enzymes and their mutant variants, including other on-going experiments in our lab. The advantage of establishing this method for quantifying protein glycosylation is to detect glycoprotein production in a faster and more quantitative approach, while eliminating additional protein purification steps. Currently, bacterial *N*-glycosylation systems have been studied by various analytical techniques, such as immunoblotting, lectin blotting, mass spectrometry, and nuclear magnetic resonance, which have allowed the detection of glycoproteins and the elucidation of *N*-glycan structures on proteins (Szymanski *et al.*, 1999; Wacker *et al.*, 2002; Young *et al.*, 2002;

Kelly *et al.*, 2006). Nonetheless, detection and quantification of glycoprotein biosynthesis using a 96-well nickel-coated ELISA plate renders this assay suitable for large-scale screening purposes and would allow more accurate and reproducible analysis compared to Western blotting.

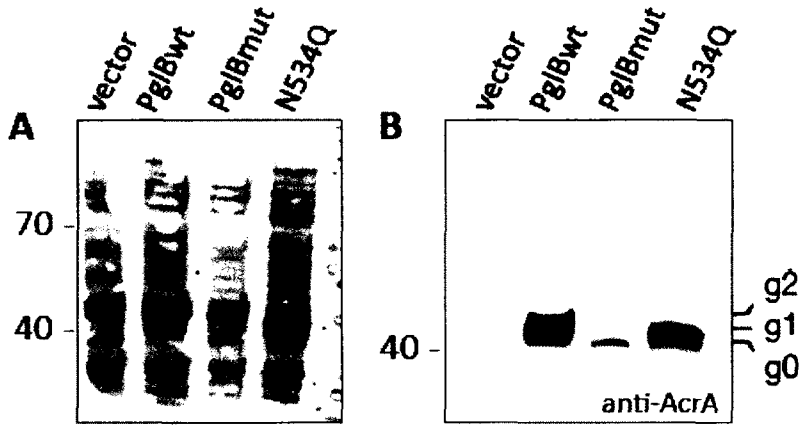

Figure A-1.2. Electrophoretic mobility shift assay and immunodetection of *E. coli* periplasmic extracts. **A)** Coomassie stain of periplasmic extracts from *E. coli* CLM24 strains analyzed by SDS-PAGE. *E. coli* strains expressing the *C. jejuni* $AcrA_{His6}$, the *C. jejuni pgl* operon containing a deletion in *pglB*, and the various PglB constructs: wildtype PglB (wt), inactive PglB (mut), PglB consensus sequence mutant (N534Q), and the negative control without PglB (vector). **B)** Immunoblot of periplasmic extracts of *E. coli* CLM24 strains. Periplasmic samples from *E. coli* expressing $AcrA_{His6}$ were run on SDS-PAGE, transferred onto a membrane, and probed with AcrA specific antibodies. Molecular masses and AcrA glycosylation state are indicated.

A.1.4. Literature cited

Feldman MF, Wacker M, Hernandez M, Hitchen PG, Marolda CL, Kowarik M, Morris HR, Dell A, Valvano MA & Aebi M (2005) Engineering *N*-linked protein glycosylation with diverse O antigen lipopolysaccharide structures in *Escherichia coli*. *Proc Natl Acad Sci USA* **102:** 3016-3021.

Kelly J, Jarrell H, Millar L, Tessier L, Fiori LM, Lau PC, Allan B & Szymanski CM (2006) Biosynthesis of the *N*-linked glycan in *Campylobacter jejuni* and addition onto protein through block transfer. *J Bacteriol* **188:** 2427-2434.

Lefebre MD & Valvano MA (2002) Construction and evaluation of plasmid vectors optimized for constitutive and regulated gene expression in *Burkholderia cepacia* complex isolates. *Appl Environ Microbiol* **68:** 5956-5964.

Szymanski CM, Yao R, Ewing CP, Trust TJ & Guerry P (1999) Evidence for a system of general protein glycosylation in *Campylobacter jejuni*. *Mol Microbiol* **32:** 1022-1030.

Wacker M, Linton D, Hitchen PG, Nita-Lazar M, Haslam SM, North SJ, Panico M, Morris HR, Dell A & Wren BW (2002) *N*-linked glycosylation in *Campylobacter jejuni* and its functional transfer into *E. coli*. *Science* **298:** 1790-1793.

Young NM, Brisson JR, Kelly J, Watson DC, Tessier L, Lanthier PH, Jarrell HC, Cadotte N, St. Michael F & Aberg E (2002) Structure of the *N*-linked glycan present on multiple glycoproteins in the Gram-negative bacterium, *Campylobacter jejuni*. *J Biol Chem* **277:** 42530-42539.

A.2. Antibody generation against the soluble domain of *C. jejuni* PglB

A.2.1. Introduction

Characterization of the bacterial *N*-linked protein glycosylation system requires a number of applicable methods that are used to clone the genes encoding the proteins of interest, separate and detect proteins by SDS-PAGE and Western blotting, and analyze the sugar composition of the oligosaccharide by mass spectrometry and NMR. Western blotting was the standard approach used in this project to detect the expression and glycosylation status of proteins. A hexa-histidine fusion protein was expressed with our protein of interest allowing its detection by using commercially available anti-His antibodies. The *C. jejuni* AcrA protein is the model acceptor protein used in this thesis because it is the best characterized *N*-glycoprotein that possesses two glycosylation sites allowing the detection of the protein and qualitatively measuring the glycosylation efficiency of the bacterial OTase.

To analyze the efficiency of PglB activity, I was comparing the glycosylation status of AcrA by PglB wildtype and mutant variants. For this thesis, the PglB proteins that were analyzed are tagged with HA to follow expression of the enzymes and to ensure the same level of protein is expressed in each mutant. However, immunodetection of the PglB proteins using HA specific polyclonal antibodies was difficult due to the high level of background observed on the membrane blot, which resulted from the unspecific binding of the antibody. In contrast, when immunodetection was performed with commercially available anti-HA monoclonal antibodies, no bands were observed on the membrane.

Although monoclonal antibodies are highly specific, the unsuccessful Western analysis may have been due to the conformational requirement of the epitope that is recognized on the HA tag. If the epitope is conformation-dependent, then the antibody will not detect denatured proteins. Therefore, the generation of PglB-specific antibodies is a valuable reagent for the continuation of this thesis and other on-going experiments in our lab.

A.2.2. Methods

Construction of plasmid

The C-terminal soluble domain of *C. jejuni* PglB (residues 428-713) was amplified by PCR with oligonucleotides PglB-SD-*Nde*I-F (5'-CAGTATTTATCCATATGTACAACT-ATAAAG-3') and PglB-SD-*Xho*I-R (5'-ATTACTCGAGAATTTTAAGTTTAAAAAC-CTTAGC-3') using the Vent polymerase (New England BioLabs®) and pMAF10 as template. The cloning vector used was the pET24b-*acrA*, which was double digested with *Nde*I-*Xho*I to separate the *acrA* insert by agarose gel electrophoresis and replace with the PglB(SD) insert. For this, the PCR product was also digested with *Nde*I and *Xho*I and cloned into the empty pET24b vector that encodes a His tag. *E. coli* DH5α cells (Invitrogen) were transformed with the ligation reaction and the resulting plasmid, pET24-PglB(SD), was confirmed by restriction analysis.

Expression and purification of the PglB soluble domain

The pET24-PglB(SD) plasmid was transformed into *E. coli* BL21 and grown on LB agar plates containing 25 μg/mL of kanamycin (final concentration). For protein expression, *E. coli* was grown in 1L LB at 37°C until the OD_{600} reached

0.5-0.6, the cells were induced with 0.1 mM IPTG and incubated for an additional 3 hrs at 37°C. Cells were then harvested by centrifugation (3,696 x *g*, 15 min, 4°C) and the pellets were resuspended in 5% of the original culture volume of PBS, pH 7.2. Whole cell lysates were prepared by sonication (3 x 1 min) using a Branson sonicator and the samples were centrifuged once more (16,260 x *g*, 20 min, 4°C) to isolate the soluble proteins. The wet pellet was dissolved with the Inclusion Body Solubilization Reagent (Thermo Scientific) to retrieve expressed PglB(SD) proteins from inclusion bodies and to maximize the yield of solubilized protein, following the manufacturer's protocol. Unbroken cells and cell debris were removed by centrifugation (16,260 x *g*, 20 min, 4°C). The supernatant was collected, subjected to nickel affinity chromatography (NTA agarose, Qiagen) to purify PglB(SD)$_{His6}$, and dialyzed against 1xPBS, pH 7.2, as previously described (section 3.2.3). The protein concentration was determined to be approximately 300 μg/mL.

Table A-2.1. Bacterial strains and plasmids used in this study

Strain or plasmid	Characteristic	Source
E. coli DH5α	F⁻ *endA1 hsdR17 supE44 thi-1 recA1* Δ (*argF-lacZYA*)*U169* (80d *lacZ* Δ *M15*) *gyrA96* λ^-	Invitrogen
E. coli CLM24	W3110, Δ*waaL*	Feldman *et al.*, 2005
E. coli BL21	F⁻ *ompT gal dcm lon hsdS$_B$*(r_B^- m_B^-) λ(DE3 [*lacI lacUV5*-T7 gene 1 *indl sam7 nin5*])	Wacker *et al.*, 2002
pMAF10	HA-tagged *C. jejuni pglB* cloned in pMLBAD, Tmp^R	Feldman *et al.*, 2005
pET24b-*acrA*	Soluble periplasmic *C. jejuni acrA*$_{His6}$ under control of T7 promoter, in pET24b, Amp^R	Wacker *et al.*, 2002

Rabbit immunization

Dr. Harald Nothaft prepared purified antigen for rabbit immunization in our lab. New Zealand White Rabbits were immunized with the antigen using a 10 week immunization protocol (approved Animal Care Committee protocol No. 717). After the concentration of the PglB(SD) antigen was measured, a stable mixture comprised of the antigen and the adjuvant was prepared prior to injection (a total of 5 mL with 500 μg of protein). Injections into rabbits were carried out by the personnel of the Animal Facility (University of Alberta). After an initial subcutaneous injection (3 sites, 1.5 mL each) of antigen using Freund's complete adjuvant (in a 1:1 ratio with the antigen), booster doses of the antigen mixed with Freund's incomplete adjuvant (in a 1:1 ratio with the antigen) were given (3 sites, 1.5 ml each) in weeks 4 and 8. A test bleed was taken at week 6. After 10 weeks, rabbits were sacrificed and serum was prepared by cooling the obtained blood for 60 min on ice followed by centrifugation (10,000 x *g*, 20 min, 4°C). Blood sera in the supernatants were analyzed for PglB(SD)-specific antibodies by Western blotting against whole cell lysates from strain *C. jejuni* 11168.

Immunodetection by Western blotting

Proteins were analyzed by SDS-PAGE on 12% gels and transferred onto PVDF membrane as previously described in section 2.2.6. The membrane was probed with PglB(SD)-specific antisera in a 1:2000 dilution against various concentrations of antigen.

Expression of full-length $PglB_{His6}$

In chapter 3 of this thesis, PglB activity was assayed *in vitro* as described in section 3.2.9. For expression of $AcrA_{His6}$ and $PglB_{His6}$, His-specific and PglB(SD)-specific antibodies were used for Western analysis. The difference in antibody sensitivity was tested in this study.

A.2.3. Results and Discussion

The His-tagged PglB(SD) protein was expressed in *E. coli* BL21 and purified by Ni-NTA affinity chromatography (Figure A-2.1). Purified PglB(SD) protein was mixed with adjuvant, followed by multiple injections given to rabbits in order to develop strong immune response against the antigen. The serum, which contains all the antibodies produced by the animal, was collected and separated from the blood cells by centrifugation. The serum was then tested (by Dr. Nothaft in our lab) to evaluate the sensitivity of the polyclonal antibodies. Immunodetection with a suitable dilution of the antibody (1:2000) was carried out using different amounts of antigen (Figure A-2.2). The purified PglB(SD) protein that was used for rabbit immunization was used as a positive control and whole cell lysates of wildtype *C. jejuni* 11168 were tested for the detection of full-length PglB, while the *C. jejuni* 11168 *pglB* mutant was used as a negative control. As expected, the PglB(SD) was detected with the polyclonal antisera and there was no reactivity with the antibody against the *C. jejuni* 11168 *pglB* mutant since the PglB protein is not expressed in this strain. However, there was also no detection observed for whole cell lysates of the wildtype *C. jejuni* strain, which was

surprising because PglB is produced in this strain as previously described by demonstrating the *N*-glycosylation of proteins (Nothaft *et al.*, 2009).

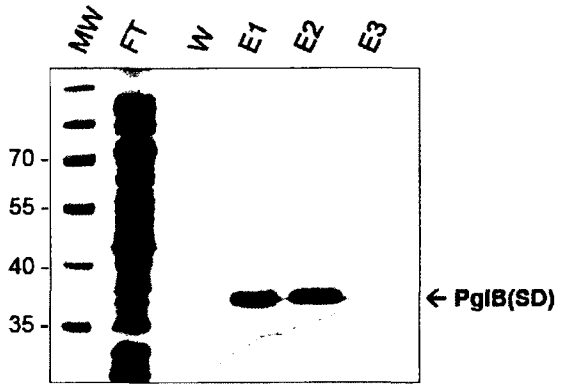

Figure A-2.1. Purification of His6-tagged *C. jejuni* PglB soluble domain (SD) by Ni-NTA affinity chromatography. The PglB(SD) protein was expressed in *E. coli* BL21, purified from inclusion bodies, analyzed by SDS-PAGE and visualized by Coomassie Blue staining. Solubilized proteins were subjected to column purification and the flow through was collected (FT); the column resin was washed with 20 mM imidazole in 1xPBS (W); and eluted three times with 1 mL of 500 mM imidazole in 1xPBS (E1-E3). Molecular weight (MW) markers and PglB(SD) are indicated.

Reactivity of the PglB(SD)-specific antisera was also tested against *E. coli* BL21 cells overexpressing the PglB(SD) by Western analysis, but there was a very high level of background observed on the membrane (data not shown). The advantages of producing polyclonal antibodies for Western analysis are that they are relatively easy to generate and recognize both conformational and sequential epitopes of the antigen. However, the disadvantage is that the serum will contain a

variety of antibodies with various affinities to other antigens produced in *E. coli*; thus, resulting in non-specific binding to *E. coli* whole cells. To increase the specificity of the antibodies for PglB(SD), the serum was cross-adsorbed against *E. coli* whole cell lysates in order to remove *E. coli* antibodies that cross-react with *E. coli* proteins. I prepared *E. coli* BL21 lysates, from which proteins were transferred onto membranes and probed with the antisera. Unbound antibodies were presumed to contain PglB(SD)-specific antibodies and were collected after overnight incubation before continuing with secondary anti-rabbit antibody incubation and development of the Western blot. This cross-adsorption method of eliminating cross-reactivity of non-specific *E. coli* antibodies was repeated 3 times by Western blotting, and the unbound PglB(SD)-specific antisera were pooled together for further immunodetection studies.

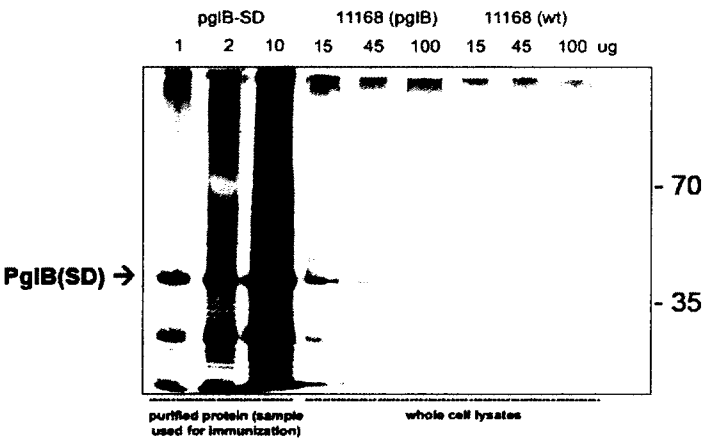

Figure A-2.2. Western Blot analysis of the polyclonal rabbit-anti-PglB(SD) antibodies for use against *C. jejuni* whole-cell extracts. The sensitivity of the antibody was tested against the indicated amounts of antigen: purified PglB(SD) as positive control (1μg, 2μg, 10μg); lysates of *C. jejuni* 11168 *pglB* mutant as

negative control (15μg, 45μg, 100μg), and lysates of *C. jejuni* 11168 wildtype strain to test for antibody reactivity (15μg, 45μg, 100μg). Proteins were probed with antisera at a dilution of 1:2000. Molecular weight markers and PglB(SD) are indicated. Expected masses of PglB(SD) and full-length PglB are 37 kDa and 70 kDa, respectively.

In Chapter 3 of this thesis, the *in vitro* protein glycosylation activity of full-length PglB was examined (section 3.3.5). The expression of His-tagged PglB and His-tagged AcrA was analyzed by immunoblotting with monoclonal anti-His antibodies, which only detected AcrA protein expression, but not that of PglB indicating degradation of the protein over time. For this study, I repeated the *in vitro* analysis of PglB activity by incubating $PglB_{His6}$ with $AcrA_{His6}$ and LLOs at various conditions: 1hr at 30°C, overnight at 30°C, and overnight at room temperature. The different incubation conditions were tested to examine PglB degradation. Immunodetection of protein expressions with anti-His and anti-PglB(SD) antibodies was carried out (Figure A-2.3). The full-length $PglB_{His6}$ and the $AcrA_{His6}$ proteins were incubated alone and were included as controls. As expected, the monoclonal anti-His antibody allowed the detection of both PglB and AcrA proteins. Surprisingly, the cross-adsorbed polyclonal antibodies that were prepared for the detection of PglB also showed expression of both PglB and AcrA proteins when analyzed by Western blotting. These data suggest that the polyclonal antiserum is specific to the His-tag, but not to PglB(SD). Interestingly, PglB expression was observed by Western analysis using the polyclonal

antiserum when incubated for 1hr in the *in vitro* assay, but could not be detected with the monoclonal anti-His antibody under the same condition. For both of the blots analyzed, the amount of protein used in each sample was kept constant. Therefore, the intensity of the bands on the membranes correlates with the specificity of the antibodies for protein detection.

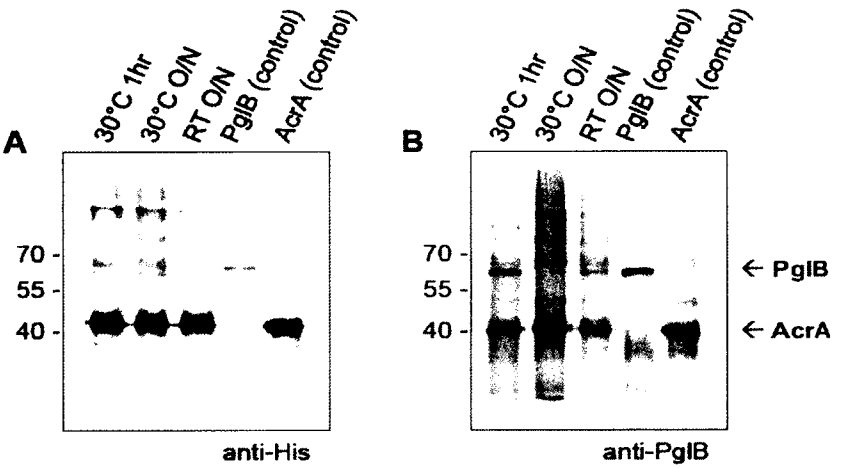

Figure A-2.3. Immunodetection of full-length $PglB_{His6}$ with (A) His-specific and (B) PglB(SD)-specific antisera. His-tagged *C. jejuni* PglB and His-tagged AcrA proteins were expressed in *E. coli* and purified by nickel affinity chromatography. Purified proteins were incubated with crude *C. jejuni* LLOs *in vitro* for 1hr at 30°C and overnight at 30°C or at room temperature. TCA-precipitated proteins were analyzed by Western blotting using anti-His and anti-PglB(SD) at dilutions 1:2000 and 1:500, respectively. Molecular weight markers and PglB and AcrA proteins are indicated.

Furthermore, when the specificity of the PglB antiserum was analyzed against *E. coli* strains over-expressing His-tagged PglB(SD) and HA-tagged full-length PglB, immunodetection with the antiserum was observed only with $PglB(SD)_{His6}$ (Nothaft, H., personal communication). It appears that the serum prepared for detection of PglB(SD) was generated against the His6-tag, which possibly contains antibodies that recognize specific epitopes of PglB(SD) at the C-terminal end of the protein near the His6-tag. The polyclonal antiserum also appears to be more specific for the PglB(SD) than the full-length PglB since Western blotting showed greater reactivity against 1 μg of $PglB(SD)_{His6}$ (Figure A.2.2) compared to 20 μg of $PglB(full)_{His6}$ (Figure A.2.3). These polyclonal antibodies might recognize conformational epitopes that vary between the two PglB proteins, which may be due to the difference in protein fold and structure. Although proteins are degraded in the rabbit after injection and presented to the immune system as fragments, antibodies can be generated to recognize various epitopes that are both linear and conformational. In spite of this, conformational epitopes are usually no longer recognized by their antibody after denaturation by SDS-PAGE.

For future directions on the generation of PglB-specific antibodies, the full-length PglB protein should be used for rabbit immunization because it might stimulate the desired immune compared to the response observed against the PglB(SD). I have optimized the over-expression and purification of His6-tagged PglB(full) (see Chapter 3). Moreover, Bernadette Beadle, in our lab, is continuing to optimize the construct to be used for the generation of PglB-specific antibodies.

She has already cloned the full-length *pglB* gene into a vector containing a His6-tag and has inserted a protease cleavage site immediately before the tag. The concept is to express PglB(full)$_{His6}$ in *E. coli*, purify the protein by nickel affinity chromatography, and cleave the His6-tag using a protease, such as Factor Xa. The availability of PglB-specific antibodies will be valuable for any future studies in characterizing the *C. jejuni* PglB enzyme.

A.2.4. Literature cited

Feldman MF, Wacker M, Hernandez M, Hitchen PG, Marolda CL, Kowarik M, Morris HR, Dell A, Valvano MA & Aebi M (2005) Engineering *N*-linked protein glycosylation with diverse O antigen lipopolysaccharide structures in *Escherichia coli*. *Proc Natl Acad Sci USA* **102:** 3016.

Nothaft H, Liu X, McNally DJ, Li J & Szymanski CM (2009) Study of free oligosaccharides derived from the bacterial *N*-glycosylation pathway. *Proc Natl Acad Sci USA* **106:** 15019.

Wacker M, Linton D, Hitchen PG, Nita-Lazar M, Haslam SM, North SJ, Panico M, Morris HR, Dell A & Wren BW (2002) *N*-linked glycosylation in *Campylobacter jejuni* and its functional transfer into *E. coli*. *Science* **298:** 1790.